



**HAL**  
open science

# Sex determination and differentiation in the brown algae

Josselin Guéno

► **To cite this version:**

Josselin Guéno. Sex determination and differentiation in the brown algae. Botanique. Sorbonne Université, 2019. English. NNT : 2019SORUS629 . tel-03560845

**HAL Id: tel-03560845**

**<https://theses.hal.science/tel-03560845>**

Submitted on 7 Feb 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.

# Sorbonne Université

Ecole Doctorale 515 Complexité du Vivant

*UMR 8227 CNRS – Sorbonne Université*

*Laboratoire de Biologie Intégrative des Modèles Marins Equipe*

*Génétique des Algues*

## **Détermination et différenciation sexuelle chez les algues brunes**

*Sex determination and differentiation in the brown algae*

Par Josselin Guéno

Thèse de doctorat de Génétique et biologie du développement

Dirigée par Susana M. Coelho et J. Mark Cock

Devant un jury composé de :

Dr. Yann Guiguen	Inra LPGP rennes	Rapporteur
Dr. Gareth Pearson	Universidade do Algarve, Faro	Rapporteur
Dr Leila Tirichine Delacour	Université de Nantes	Examinatrice
Dr Frédérique Le Roux	Sorbonne Université - CNRS	Présidente du jury
Dr Susana M. Coelho	Sorbonne Université – CNRS	Directrice de thèse
Dr J. Mark Cock	Sorbonne Université - CNRS	Co-directeur de thèse



*À ma famille,*

*À Aurélia,*

*"In a hole in the ground there lived a hobbit."*

John Ronald Reuel Tolkien 1937

Voilà maintenant quatre ans que j'ai débarqué à Roscoff, à la recherche d'un cadre de vie utopique en bord de mer. Il est vrai que lorsque je me suis renseigné sur la géographie, les photos étaient attrayantes, même édéniques. Malheureusement, en arrivant au mois de Janvier, la douche est froide et l'eau salée ! On appréhende, on se demande si on a bien fait de choisir cette ville et si les photos n'étaient pas qu'une façade pour leurrer. Mais derrière cette brume de mer qui nappe la côte déchiquetée, si on s'accroche et on persiste à vouloir explorer et dompter cette côte sauvage, si on arrive à voir au-delà des apparences, alors on découvre quelque chose que l'on n'avait pas envisagé ; on s'y plait. La vie à Roscoff, c'est l'allégorie d'une thèse, ce qui en fait un endroit parfait pour en faire une. J'ai apprécié chacun des moments passés ici, et je vous en remercie tous.

Tout d'abord je voudrais remercier mes directeurs de thèse, Susana Coelho et Mark Cock, pour m'avoir accueilli au sein de leur équipe, pour m'avoir fait confiance pendant quatre années, et pour avoir su garder un équilibre entre autonomie et encadrement, ce qui m'a permis d'acquérir des bases solides pour la suite.

Je tiens également à remercier les autres chercheurs qui m'ont aiguillé et conseillé pendant les comités de thèses : Dr Leila Tirichine Delacour, Dr Frédérique Le Roux et Dr Julia Morales. Et comme le hasard fait bien les choses, je remercie deux d'entre elles : Dr Leila Tirichine Delacour, Dr Frédérique Le Roux pour leur présence dans mon jury de thèse. Je remercie également Dr Yann Guiguen et Dr Gareth Pearson pour avoir accepté d'être les deux rapporteurs de cette thèse.

Je tiens à remercier toutes les personnes qui ont pu m'aider pour la réalisation de ce travail. Je pense à Guillaume Cossard, Aga Lipinska et Olivier Godfroy, pour le nombre de fois (que je ne compte plus) où je suis venu les voir pour mes problèmes de bio-informatique. Mais également, à Zofia Nehr, Laurent Peres et Delphine Scornet pour leur aide lors de mes différentes expérimentations.

Je remercie les membres de l'équipe « Génétique des algues » avec qui j'ai eu plaisir à travailler et échanger. Je remercie également les chercheurs, chercheuses, post-doc, techniciens et doctorants du laboratoire de biologie intégrative des modèles marins pour leur accueil et leur bonne humeur.

Trois années dans le même environnement forgent forcément des liens. Je remercie toutes les personnes avec qui j'ai tissé de très bons souvenirs : Laure Mignerot et Romain Fusté pour ces merveilleux week-ends, mon ventre vous remercie, mon foie moins. Je remercie Simon Bourdareau qui m'a vu grandir (pas physiquement je le suis déjà bien assez) et qui m'a accompagné pendant 3 ans. Je remercie également Hervé Rabillé, Adèle James, Céline Conan, Florent Souchoux, Yacine Badis, Florian Ponteaux et Elodie Rolland (Don't stop me now) pour les moments passés. Je remercie tout particulièrement Haiqin Yao pour sa présence lors des 3 années de thèse, et qui a permis de rendre la rédaction de thèse plus agréable (surtout avec les hots pots).

Je tiens à remercier mes parents et mon frère, qui m'ont soutenu et encouragé malgré la distance (oui Roscoff-Paris c'est loin). Je voudrais remercier également Nathalie, Gaël et Gwen pour leur soutien et le refuge apporté les week-ends. Je remercie également Souna (elle se reconnaîtra !).

Pour finir, le meilleur pour la fin ... enfin la meilleure, je remercie plus que tout Aurélia, pour m'avoir soutenu, supporté et encouragé, plus que ce que j'aurais pu imaginer. Elle a énormément contribué, à sa façon, à l'épanouissement de cette thèse. Je t'aime plus que trois fois mille.

## Table des matières

<b>Chapter 1</b> .....	<b>11</b>
<b>General introduction</b> .....	<b>11</b>
<b>1. Evolution of eukaryotic sex</b> .....	<b>12</b>
1.1. The origin of sex .....	12
1.2. The maintenance of sex .....	14
1.3. Advantages of meiotic sex .....	15
<b>2. Genetic sex determination: sex chromosomes</b> .....	<b>17</b>
2.1 Different types of sex chromosome .....	17
2.2 Sex chromosome evolution .....	19
2.3 Sex chromosome structure: The sex-specific and the pseudoautosomal regions of the sex chromosome.....	22
<b>3. Genetic mechanisms that determine sex in eukaryotes: sex determination cascades</b> .....	<b>24</b>
3.1 Sex determination cascades .....	24
3.2 Master sex determination genes .....	25
3.3 "Slave" genes .....	28
3.4 Female determination .....	29
3.5 "Masters change, slaves remain" .....	31
<b>4. Epigenetic mechanisms that determine sex in eukaryotes</b> .....	<b>32</b>
4.1 Environmental sexual determination (ESD) .....	32
4.2 Continuum between GSD and ESD.....	33
4.3 Chromatin dynamics and regulation of sex determination cascades .....	35
<b>5. Sex differentiation: sex biased gene expression</b> .....	<b>36</b>
<b>6. The role of chromatin in the regulation of gene expression:</b> .....	<b>38</b>
6.1 Chromatin, the carrier of genetic information. ....	40

6.2 DNA methylation .....	41
6.3 Canonical histones and histone variants.....	42
6.4 Histone modifications .....	43
<b>7. The brown alga <i>Ectocarpus</i>: a new model to study the evolution of sex determination .45</b>	
7.1 Context .....	45
7.2 The life cycle of <i>Ectocarpus</i> .....	47
7.3 Sex determination in the brown algae.....	48
7.4 Sex differentiation in brown algae.....	49
<b>Objectives.....</b>	<b>50</b>
<b>Chapter 2.....</b>	<b>53</b>
<b><i>Mechanisms of sex determination: functional analysis of a candidate sex-determination factor in Ectocarpus</i> .....</b>	<b>53</b>
<b>1. Introduction.....</b>	<b>54</b>
<b>2. Material and methods .....</b>	<b>59</b>
2.1 Plasmid construction.....	59
Plasmid for DAP-seq and Production of anti-HMG antibodies .....	59
Plasmids for Yeast Two-hybrid bait constructs .....	61
2.2 Yeast Two-Hybrid cDNA library .....	61
2.3 Yeast Two-Hybrid Assay .....	62
Strain genotypes.....	62
Selection of bait and prey combinations in yeast.....	62
Auto-activation bait .....	62
Mating and screening for prey-bait interactions .....	63
Extraction and cloning of prey plasmids.....	63
2.4 HMG and Yeast Two-Hybrid .....	63
2.5 HMG protein production .....	65
2.6 DAP-seq .....	65
2.7 Production of anti-HMG antibodies.....	66
<b>3. Results .....</b>	<b>68</b>

<b>3.1 Identification of HMG-interacting proteins</b> .....	<b>68</b>
<b>3.2 Construction of a Y2H library</b> .....	<b>68</b>
Interaction analysis .....	70
Identification of the cause of the low mating efficiency .....	72
<b>3.3 Genome-wide search for HMG binding sites <i>in vivo</i></b> .....	<b>74</b>
<b>3.4 Genome-wide search for HMG binding sites <i>in vitro</i></b> .....	<b>76</b>
<b>4. Discussion and conclusion</b> .....	<b>78</b>
<b>4.1 Results of the search for proteins that interact with the HMG protein Ec-13_001750</b> .....	<b>78</b>
<b>4.2 Results of the search for DNA-binding sites of the HMG protein Ec-13_001750.</b> ....	<b>82</b>
<b>Chapter 3</b> .....	<b>89</b>
<b><i>Epigenetic modifications associated with sex determination and differentiation in the brown alga Ectocarpus</i></b> .....	<b>89</b>
<b>1. Introduction:</b> .....	<b>90</b>
<b>2. Chromatin landscape associated with sexual differentiation in a UV sex determination system</b> .....	<b>92</b>
<b>2.1 Summary</b> .....	<b>92</b>
<b>2.2 Introduction</b> .....	<b>93</b>
<b>2.3 Results</b> .....	<b>95</b>
The chromatin landscape of <i>Ectocarpus</i> sp. ....	95
Identification of histone PTMs associated with gene activation and gene repression.....	97
Dynamic landscape of chromatin states on sex-biased genes .....	98
Sex-specific chromatin reprogramming of sex-biased genes.....	98
The chromatin landscape of the <i>Ectocarpus</i> sex-chromosomes.....	100
Chromatin states and gene expression levels on the sex chromosome.....	100
<b>2.4 Discussion</b> .....	<b>101</b>
Broad conservation of the histone code in the brown algae.....	101
Widespread sex-specific chromatin reprogramming in <i>Ectocarpus</i> gametophytes .....	102
Unique chromatin organisation features in the U and V sex chromosomes.....	102
<b>2.5 Methods</b> .....	<b>103</b>
Biological Material .....	103
Genome-wide detection of histone PTMs.....	105



Detection of chromatin states.....	106
Coverage of each chromatin state .....	107
<b>2.6 Supplemental Tables Legends .....</b>	<b>117</b>
<b>2.7 Acknowledgements.....</b>	<b>118</b>
<b>2.8 References .....</b>	<b>118</b>
<b>3. Conclusion.....</b>	<b>123</b>
<b>Chapter 4.....</b>	<b>125</b>
<b><i>A partially sex-reversed giant kelp sheds light into the mechanisms of sexual differentiation in the brown alga .....</i></b>	<b>125</b>
<b>1. Introduction .....</b>	<b>126</b>
<b><i>2. A partially sex-reversed giant kelp sheds light into the mechanisms of sexual differentiation in the brown algae .....</i></b>	<b>127</b>
<b>2.1 Summary .....</b>	<b>127</b>
<b>2.2 Introduction .....</b>	<b>128</b>
<b>2.3 Results .....</b>	<b>130</b>
<b>Identification of a male <i>M. pyrifera</i> line exhibiting female-like morphology.....</b>	<b>130</b>
<b>Phenotypic characterisation of the feminized <i>M. pyrifera</i> line .....</b>	<b>130</b>
<b>Mpyr-13-4 gametes do not produce the male-attracting pheromone lamoxiren .....</b>	<b>131</b>
<b>Gene expression in Mpyr-13-4 compared with wild-type male and female gametophytes .....</b>	<b>131</b>
<b>Analysis of differential gene expression in wild-type male, wild type female and Mpyr-13-4 gametophytes.....</b>	<b>132</b>
<b>Sex-biased gene expression in wild-type and variant Mpyr-13-4 lines .....</b>	<b>132</b>
<b>Evolutionary analysis of genes involved in sexual differentiation .....</b>	<b>133</b>
<b>Predicted functions of sex-biased genes.....</b>	<b>134</b>
<b>Expression of male SDR genes in the Mpyr-13-4 line .....</b>	<b>134</b>
<b>2.4 Discussion .....</b>	<b>134</b>
<b>Mpyr-13-4 is partially sex reversed .....</b>	<b>134</b>

The transcriptional landscape underlying sexual differentiation in the giant kelp .....	136
Evolution of sex-biased genes in the giant kelp .....	136
<b>2.5 Materials and Methods.....</b>	<b>138</b>
Biological material .....	138
Pheromone measurements.....	138
Generation of transcriptomic sequence data .....	139
Identification of sex-biased genes .....	140
Evolutionary analysis .....	140
Flow cytometry.....	140
<b>2.6 Supplemental Tables.....</b>	<b>141</b>
<b>2.7 Acknowledgements.....</b>	<b>141</b>
<b>2.8 References .....</b>	<b>141</b>
<b>Chapter 5.....</b>	<b>153</b>
<b>General discussion .....</b>	<b>153</b>
<b>1. Characterisation of the candidate male sex-determination gene in Ectocarpus, the HMG domain gene Ec-13_001750. ....</b>	<b>154</b>
<b>2. Chromatin modifications in male and female gametophytes .....</b>	<b>156</b>
<b>3. Molecular changes associated with partial sex-reversal shed light into the mechanisms of sex differentiation in the giant kelp .....</b>	<b>162</b>
<b>References.....</b>	<b>165</b>
<b>Appendix UV Chromosomes and Haploid Sexual Systems.....</b>	<b>189</b>





# Chapter 1

## General introduction

## 1. Evolution of eukaryotic sex

What is sex? Sex can be defined simply as a cycle involving alternation between meiosis and syngamy (the fusion of haploid gametes to form a diploid zygote). The classic view of sex is that it allows genetic mixing and recombination of genetic information, which in turn allows offspring to receive diverse genetic heritages and to possess new genetic combinations. Thus, mutations create variation and sex creates new combinations of these variations. However, this recombination-related sex advantage is only a long-term advantage. Recombination requires a population with sufficient genotype diversity to be advantageous and requires long-time periods to produce adaptive genotypes (Bürger, 1999). In the short term, other (asexual) strategies are more adaptive and can destabilize and replace sexual reproduction. Thus, two fundamental questions remain about sex: why did it appear and why is it maintained?

### 1.1. The origin of sex

Current analysis indicates that the last eukaryotic common ancestor (LECA) was capable of full meiotic sex (Speijer, 2016). As its name suggests, meiotic sex involves meiosis. Meiosis is an ancient, highly conserved process in eukaryotic life cycles, and a shared component of sexual reproduction for all eukaryotes (Speijer et al., 2015). Eukaryotic sex involves alternation between a haploid and a diploid phase, with meiosis mediating the transition from the diploid to the haploid phase, and gamete fusion (syngamy) reconstituting a diploid genome. Eukaryotic sex is different from bacterial sex (conjugation), which does not involve meiosis although there is an exchange of genetic material between two individuals. Bacterial sex involves unidirectional transfer (donor/receptor) of genetic material. Also, in bacteria, sex only creates variability through homologous recombination for a small fraction of the genome (Smith et al., 1991) whereas in eukaryotes, sex produces variability across the entire genome.

The original eukaryotic life cycle (Figure 1) was probably clonal, interrupted by episodic sex triggered by external or internal stressors. The question is therefore: why did the LECA develop the ability to achieve sexual reproduction and in response to what kind of stresses?

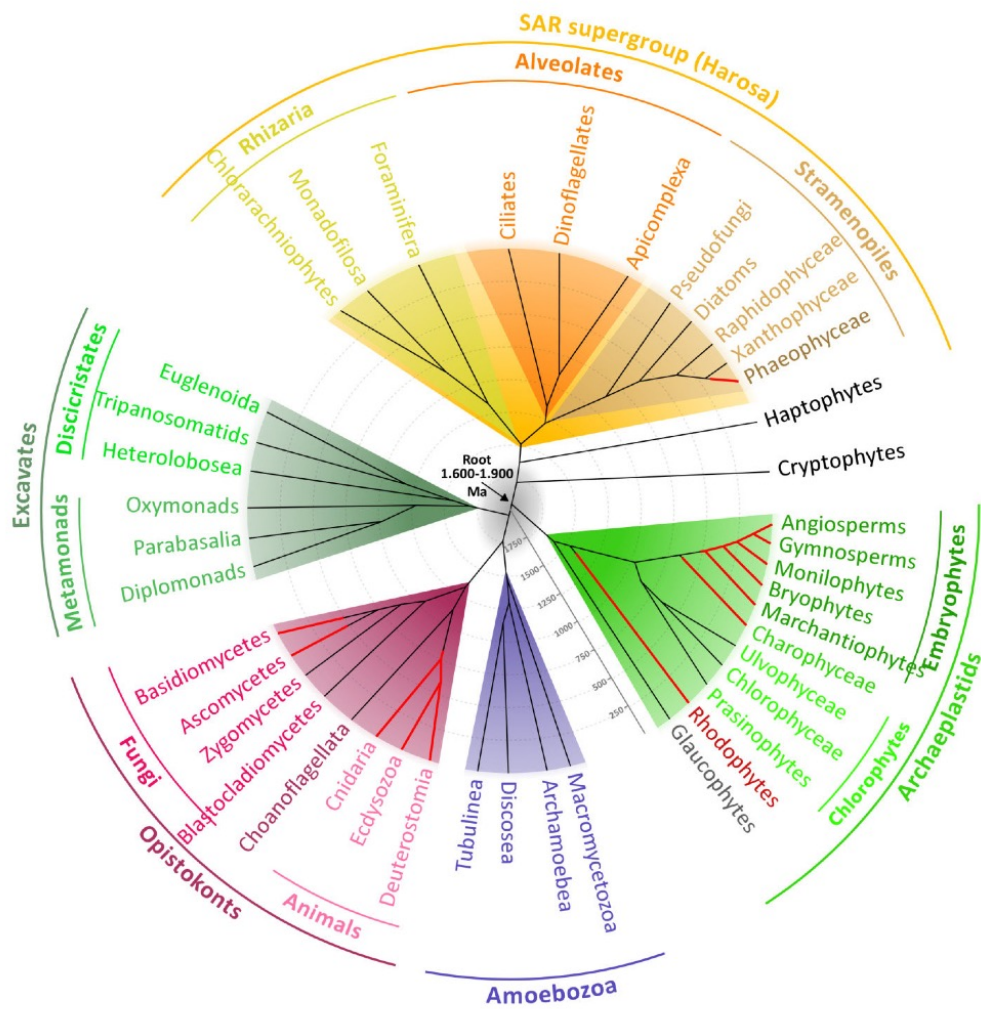


Figure 1. **Time-scaled phylogenetic tree of the eukaryotes.** Absolute time in million years (Ma) is given based on (Parfrey et al., 2011). Lineages where complex multicellularity has arisen are indicated by red branches. Simplified lineage phylogenies are based on: for Excavates, (Simpson et al., 2006); for Opisthokonts, (Cavalier-Smith et al., 2015, 2014); for Archaeplastids, (Umen, 2014); for Haptophytes and Cryptophytes, (Burki et al., 2016); for SAR supergroup, (Burki et al., 2007).

Evolution of sex in eukaryotes was most probably favoured by a high mutation rate due to the acquisition of the endosymbiont, as this (proto)mitochondrion generated internal reactive oxygen species (Speijer, 2016). It is known that oxidative stress can cause physical damage to DNA (Slupphaug et al., 2003). Recent research has uncovered multiple links between stress conditions, such as the formation of reactive oxygen species (ROS), and meiotic sex. Consistent with this idea, (Nedelcu, 2005) showed that a sexual inducer (SI) and several SI-induced extracellular matrix protein genes in *Volvox carteri* are induced under various stress conditions. This multicellular green alga cannot initiate its sexual cycle in the presence of antioxidants. Furthermore, an iron chelator (iron being a cofactor in the formation of ROS) inhibits sexual induction in this species (Nedelcu and Michod, 2003). Meiotic sex would therefore make it possible to solve problems related to DNA damage through homologous recombination by creating recombinant genotypes that lack, or have reduced numbers of, mutated loci. Diploidisation through syngamy during eukaryotic sex allows homologous chromosomes to be brought together in a single cell, allowing repair of damaged DNA using undamaged copies as a template. This mechanism is particularly effective when there is double strand damage (Bernstein et al., 2011). However, this explanation alone is not sufficient as there are populations of asexual diploid eukaryotic asexual populations where homologous chromosomes are present allowing homologous recombination repair of DNA lesions (Otto and Lenormand, 2002).

## 1.2. The maintenance of sex

At first glance, sexual reproduction appears to be of little benefit, particularly when the cost of meiosis (or the cost of male production) is taken into account. This cost translates into a reproduction rate that is half that observed in an asexual population. Moreover, a gene present in a female has only a one-in-two chance of being transmitted to her offspring. If the purpose of sex is to generate diversity, what is the immediate benefit for a female to share half of her genetic heritage with a male to produce diversity whose benefit to her offspring is uncertain?

Meiotic sex has a disadvantage called the « double cost of sex » (Maynard Smith 1978). It is initially characterized by the cost of male production. Sexually reproducing females need a male in order to reproduce and half of her offspring will be male. Thus, only half of the offspring will be able to produce new offspring. In contrast, an asexual female reproduces independently and all of her offspring can produce further offspring. As a result, an asexual female will reproduce faster and generate more offspring than a sexual female. Moreover, an asexual female



transmits twice as many copies of her genes as a sexual female, for the same energy expended. Indeed, the latter must share the genetic heritage of her offspring with the male (Otto and Lenormand, 2002).

Furthermore, the double cost of sex is not the only disadvantage of sexual reproduction, it is also necessary to find a sexual partner, which can be a difficult task, exposing the individual to additional risks of predation, contracting diseases, and requiring considerable energy. For example, if an organism is sessile or almost sessile, finding a mate can be challenging. Even if an organism can actively search for a mate, finding a partner may be difficult if the energetic costs of mate searching are high. Some organisms have evolved hermaphroditism and can self-fertilize, a strategy that can limit at least part of the cost of mate searching (Eppley and Jesson, 2008). On the other hand, sex enhances the potential for evolution because recombination expands the range of gene combinations exposed to selection. The ability to parcel together beneficial genes or to separate them from harmful mutations is among the key long-term benefits of sex. In the short term, however, recombination is costly because it breaks up successful gene combinations (Lewis, 1987). It is thus surprising that sexual species are so common, and have not rapidly become asexual, or (in the case of hermaphrodites) completely self-fertilising.

### 1.3. Advantages of meiotic sex

(Lehtonen et al., 2012) defined the cost of sex as being equivalent to the magnitude of the minimum compensatory benefits that enable sexual individuals to avoid being outcompeted by asexual individuals. For example, if sex has an exactly twofold cost, then it would also have to have benefits of at least a twofold magnitude to persist in competition with asexuals. So, there must be advantages that are at least equivalent to the costs of sex to allow sexual reproduction to be maintained in eukaryotes.

The first point concerns a disadvantage of asexual reproduction. This is a phenomenon known as Muller's ratchet, which is a metaphorical representation of a physical mechanism, called a ratchet, that prevents a system from returning to its original state. The idea was developed by Muller in 1964 to model the evolution of an exclusively asexually reproducing haploid population and is based on the fact that unfavourable mutations regularly occur in all living species. Muller's ratchet is a process by which the genomes of an asexual population accumulate deleterious mutations in an irreversible manner. The accumulation of unfavourable mutations within this asexual lineage would lead to its disappearance, since each new generation would

have at least as many mutations as the parents. However, in organisms with sexual reproduction, meiotic recombination would allow the elimination of some of these deleterious mutations each generation (Kondrashov, 1988). This has been termed restorative meiosis. Recombination mixes genomes and therefore creates individuals with reduced mutation load, avoiding Muller's ratchet.

Sexuality is generally seen as a generator of genetic variability. Two phenomena associated with sexual reproduction explain how this variability is generated: recombination and segregation. These processes allow the descendants to be genetically different from their parents and from each other. The question is therefore whether descendants generated by sexual reproduction that are genetically different from their parents have improved fitness compared with clones. The perceived advantage depends on the temporal point of view. It may appear disadvantageous in the short term to break up parental allelic combinations that have confer a survival advantage. On the other hand, in the long term, generating genetic variability makes it possible to better respond to natural selection. But the selection pressure must be high to counterbalance the numerical advantage of asexual reproduction (twice as many progeny per parent). An important selective pressure in eukaryotes is due to host/parasite interactions. These interactions are expected to lead to co-evolution originating in an arms race according to the Red Queen Theory. A number of studies support the arms race model. For example, the Red Queen hypothesis was tested using *Caenorhabditis elegans* as a host and the bacterium *Serratia marcescens* as parasite. *Caenorhabditis elegans* populations, propagating either sexually or by self-fertilization or using a combination of these two strategies, were exposed to the bacterium. Self-fertilizing populations were rapidly driven extinct while outcrossing sexual populations were not. Thus, in this case, coevolving pathogens selected for biparental sex (Morran et al., 2011).

On an evolutionary scale, almost all eukaryotic species with an asexual mode of reproduction are recent and this observation has been interpreted to indicate that species may become clonal, and could presumably invade asexually multiplying genotypes as a result of their improved rate of reproduction (reduced cost of sex) but that this strategy would not be viable in the long term and would lead to the extinction of these species due to a lack of adaptive capacity (Gouyon, 1999). Such asexual lineages are sometimes referred to as "evolutionary dead ends".

Meiotic sex is a complex two-step process initiated by syngamy, which involves the fusion of two haploid cells (gametes) to form a diploid zygote, and ending with reduction to haploidy through meiosis. Meiosis is a reducing mechanism which produces four haploid cells from a

single diploid cell. Interestingly, in sexual lineages, there are often mechanisms preventing the fusion of some combinations of gametes. Depending on the lineage, the range of permissible partners is controlled by systems referred to as ‘mating types’ or ‘sexual systems’. These systems are generally defined according to the sizes of the gametes (Parker 2011). The term mating type is employed for isogamous systems, where fusion occurs between gametes of the same size, but only if they belong to complementary mating types. Sexes produce gametes of different sizes: one small, male (sperm) and one large, female (ova) gamete. A system is termed anisogamous if the size of the gametes differs but both gametes are mobile and oogamous if gamete size is different but only the small gametes are mobile (Wiese et al., 1979). Sexual systems therefore include a factor that determines the production of different size gametes. We will focus more specifically here on anisogamous and oogamous systems.

There is a wide variety of sex-determination mechanisms, including both epigenetic and genetic systems. The difference between genetic sex determination (GSD) and epigenetic sex determination (ESD) is that genetic sex is determined by a genetic element, independent of the environmental conditions. (Valenzuela et al. 2003).

In GSD systems, sex is determined by a sex chromosome. These chromosomes originate from intralocus sexual conflicts. Intralocus sexual conflict occurs when a trait encoded by the same genetic locus in the two sexes has different optima in males and females, such as the gene controlling gamete size in anisogamous species (Mills et al., 2012)

## 2. Genetic sex determination: sex chromosomes

### 2.1 Different types of sex chromosome

In species with separate sexes, sex determination often has a genetic basis, and in a wide diversity of taxa a pair of cytologically distinguishable ‘sex chromosomes’ are found such that the chromosomal complements in males and females differ. Sex chromosomes have evolved independently several times in eukaryotes. In organisms that mate as diploids, the sex chromosome system takes two forms. In XY systems, males are heterogametic, i.e. they have both types of sex chromosome (X and Y) while females are homogametic (XX). In ZW systems, common in birds and reptiles, the male is the homogametic sex (ZZ) and the female is heterogametic (ZW) (Figure 2), we speak then of diplo-genotypic sex determination (D-GSD). These two systems can coexist within the same genus (Ezaz et al., 2006).

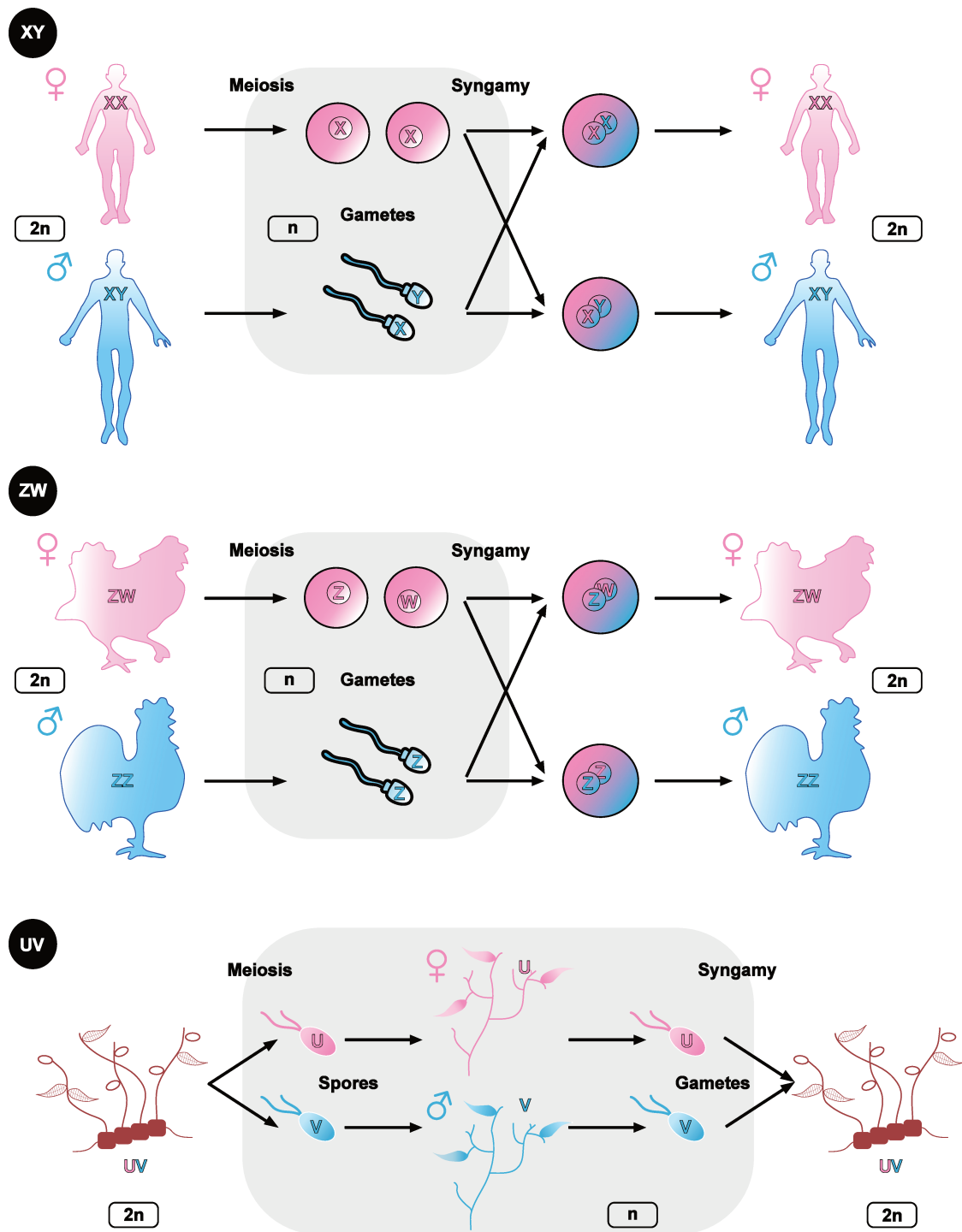


Figure 2. **Diversity of types of sexual systems (XY, ZW and UV).** In organisms with diploid life cycles, sex is determined in the diploid phase of the life cycle, after fertilization. In XY systems, the sex of the embryo depends on the chromosome carried by the sperm cell, X or Y. In ZW it is the female egg that determines the sex of the individual. In organisms such as some algae and mosses, that alternate between gametophyte and sporophyte generations (haploid-diploid life cycles), sex is expressed during the haploid (gametophyte) phase of the life cycle. The sexual system in this case is called UV systems. In contrast to XY and ZW systems, sex in UV systems depends on whether the spores receive a U or a V chromosome after meiosis (not at the fertilization stage). Based on (Bachtrog et al., 2011)

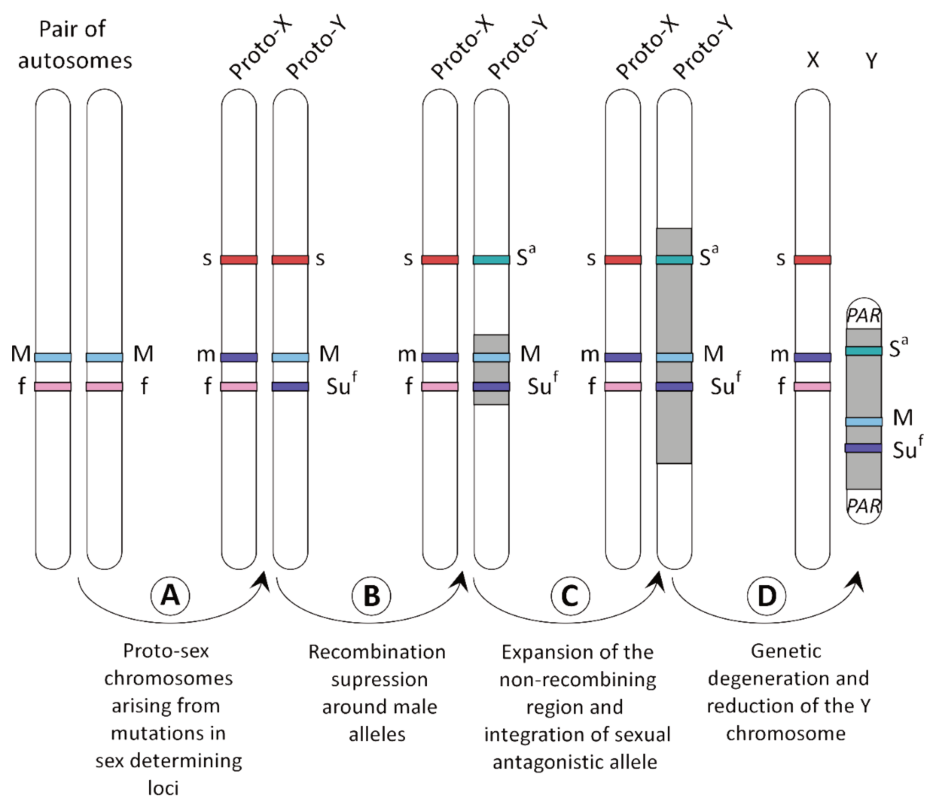
In species with an independent haploid phase in their life cycle, such as mosses and many algae, sex can be determined in the haploid gametophyte by a UV sex chromosome system: females have a U chromosome, whereas males have a V (the heterogametic UV diploid sporophyte is asexual) (Bachtrog et al., 2011; Coelho et al., 2018). This type of system is referred to as haplo-genotypic sex determination (H-GSD) (Figure 2)

## 2.2 Sex chromosome evolution

A model for sex chromosome evolution has been developed, which posits that sex chromosomes originate from an ordinary pair of autosomes harbouring a sex-determining locus between which recombination becomes suppressed (Charlesworth, 1991). The appearance of a sterility gene for females on the proto-Y together with a promoter gene for males allows the transition from a hermaphroditic population to a separate sexes population. The presence of these genes selects for reduced recombination, through chromosomal rearrangements. The loss of recombination triggers a host of evolutionary processes, including Muller's Ratchet, background selection and genetic hitchhiking that lead to the loss of gene activity and pseudogenization (Bachtrog, 2013; Bergero and Charlesworth, 2009). (Figure3)

SDRs have very variable sizes depending on the organism, ranging from the size of a locus to an entire chromosome. Extension of reduced recombination from the sex-determining locus to adjacent parts of the sex chromosome is probably the result of sexual antagonistic selection in which alleles at one locus have different relative fitness in males and females (Bachtrog et al., 2011). Reduced recombination between a locus under sexual antagonistic selection and the SDR, links genes that determine one of the sexes with alleles that increase the fitness of that sex. Therefore, selection can promote the expansion of the non-recombinant SDR to include many more genes than those essential to sexual determination. Once recombination is lost, the effectiveness of natural selection to maintain beneficial mutations and to purge deleterious mutations is reduced. Indeed, in non-recombinant chromosomes, genetic linkage between beneficial and deleterious mutations probably leads to the accumulation of deleterious mutations (background selection) and reduced purifying selection (Hill-Robertson effects). In the short term, the accumulation of mutations and transposable elements can lead to the expansion of the non-recombinant sex-determining region (Bachtrog, 2013).

In addition, if a sexually antagonistic gene occurs near the non-recombining sex-determining region (SDR), one way to resolve conflict is to fix the advantageous gene in the good sex and remove it from the disadvantaged sex. Such fixation of a sexual antagonistic (SA) gene in one



#### Legend

- f  Female promoting recessive allele
- M  Male promoting dominant allele
- m  Recessive male sterility allele
- Su<sup>f</sup>  Female suppressing dominant allele
- s  Gene allele
- S<sup>a</sup>  Allele beneficial for males but harmful for females
- Non-recombining region
- PAR Pseudo-autosomal region

**Figure 3: Sex chromosome evolution in a XY system.** In a hermaphrodite population, a pair of homologous chromosomes carries the “M” and “f” alleles. (A) “M” mutates into a recessive male-sterility allele (m) which causes the emergence of females and the dichotomy of proto-x and proto-Y chromosomes. On the proto-Y, “f” mutates into a dominant “Su<sup>f</sup>” allele, causing female-sterility and the appearance of males. (B) Between the proto-Y chromosome and its homologue the proto- X chromosome, suppression of recombination around male alleles (M and Su<sup>f</sup>) is favoured creating a non-recombining sex determining region. The “s” gene undergoes mutation on the proto-Y chromosome to create a sexual antagonistic allele (S<sup>a</sup>) that benefits the male but harms the female. (C) On the male proto-Y chromosome expansion of the non-recombining region to include S<sup>a</sup> is favoured. (D) The lack of recombination on the Y chromosome induces accumulation of deleterious mutations, genetic degeneration and genes loss resulting in a smaller male Y chromosome. The non-recombining region is not spread throughout the Y chromosome, the pseudoautosomal region (PAR) can still recombine with the corresponding region on the X. (Adapted from Charlesworth et al., 2005)

sex can be achieved by the expansion of the non-recombining region (the SDR) to include this gene (Rice, 1996). Theoretical models predict that SA alleles should accumulate on the pseudoautosomal region (PAR) close to the non-recombining sex-determining region (Charlesworth et al., 2014; Otto et al., 2011) and drive the evolution and expansion of the SDR. An example for an H-GSD system are the genes that affect gamete size in *Volvox carteri*, such as *Mat3* which is autosomal in *Chlamydomonas* (Umen and Goodenough, 2001). Linkage of *Mat3* to *Mid* contributed to the expansion of the U and V chromosome in *Volvox*, trapping several other genes within the non-recombining region (Ferris et al., 2010). For D-GSD systems, an example would be the human *Zfy* gene, a *Sry*-linked testis-specific transcription factor, previously thought to be the sex-determiner (Decarpentrie et al., 2012).

In the long term, the accumulation of mutations can lead to a decrease in the size of the heterogenetic sex chromosome Y due to gene loss following genetic degeneration and also to transposition of Y- genes to autosomes to escape degeneration (Hughes et al., 2015). The X chromosome, on the other hand, does not suffer the same evolutionary processes as the Y because it continues to recombine in females. Interestingly, recent studies indicate expansion of the non-recombining region (NRR) in the anther-smut fungus *Microbotryum lychnidis-dioicae* in the absence of sexually antagonistic selection. This fungus has mating-type chromosomes with large regions devoid of recombination but have no ‘female’ or ‘male’ functions (isogamous gametes), and sexual antagonism should not, therefore, have driven the development of the NRR. In fact, there has been (i) no enrichment in genes that are upregulated in the haploid phase (when cells express alternative mating types) and (ii) no differentially expressed genes have been detected between mating types (Bazzicalupo et al., 2019). This means that other processes must be evoked to explain the evolution of the non-recombining regions in these organisms where sexual antagonism appears to be inexistent.

Most theories about the evolution of sex chromosomes are based on studies of organisms in which sex is determined at the diploid stage (e.g., animals and land plants) (Bergero and Charlesworth, 2009; Charlesworth, 2016; Wright et al., 2016). However, suppression of recombination in sex-determining regions can also occur in organisms in which sex is determined at the haploid stage, such as algae or bryophytes (Coelho et al., 2018). The V and U chromosomes in H-GSD systems are sex limited, being present only in male and female individuals, respectively. As with XY or ZW systems, U and V sex chromosomes are expected to evolve through loss of recombination (Immler and Otto, 2015). Differences in selection between the sexes should lead to closer linkage of female beneficial alleles on the U

chromosome and male beneficial alleles on the V chromosome and this may lead to expansion of the non-recombining region. Moreover, theoretical work predicts that U and V chromosomes should have similar characteristics to those of Y and W chromosomes, namely similar size and similar rates and degrees of degeneration (Bull, 1978; Immler and Otto, 2015). As in XY systems, suppression of recombination during the diploid phase leads to the accumulation of deleterious mutations. But this accumulation should affect both chromosomes symmetrically and to a lesser degree than for W or Y chromosomes because UV chromosomes function in a haploid context where deleterious mutations cannot be masked, allowing purifying selection to act on deleterious recessive mutations. U and V chromosomes also differ from X, Y, Z and W chromosomes because they are subjected to a less extensive decrease in effective population size. For genes that are only expressed during the diploid phase, detrimental mutations can accumulate, but only insofar as they are compensated by the other sex chromosome (Bull, 1978). Thus, theories suggest that the relatively long time spent by the U and V chromosomes in the haploid phase also plays a role in the slower degeneration of SDRs through exposure of the chromosomes to purifying selection (Immler and Otto, 2015).

### 2.3 Sex chromosome structure: The sex-specific and the pseudoautosomal regions of the sex chromosome

As discussed above, sex chromosomes include a non-recombining region, called the sex-determining region (SDR), which is bordered by one or two regions called the pseudoautosomal region(s) (PAR), which undergo recombination (Otto et al., 2011). The PAR is a unique genomic region, exhibiting some features of autosomes, but also influenced by partial sex linkage. In contrast to the SDR, the PARs of sex chromosomes maintain similarity between alleles of the same gene because they undergo homologous pairing and recombination. Therefore, PAR genes are expected to evolve in a similar manner to autosomal genes (Ellis et al., 1990). However, because of the proximity to the SDR, the PARs are expected to display specific evolutionary dynamics (Otto et al., 2011).

PAR size varies depending on the species (Charlesworth, 2017). Recently evolved sex chromosomes tend to have larger PARs than older sex chromosomes, but correlation between age and PAR size is not always strict (Charlesworth et al., 2005; Otto et al., 2011; Whittle et al., 2015). Thus, in species that have recently acquired a sex chromosome, a large variation in the size of the PAR can be observed, ranging from a rather limited size, as in *Gasterosteus aculeatus* where the PAR represents 15.8% of the sex chromosome (Ross and Peichel, 2008),



to a larger size in papaya with an PAR that represents 83% of the sex chromosome (Yu et al., 2009). In species with older sex chromosomes, such as mice, the PAR represents only 1% of the total size of the Y chromosome and in humans, 4.6% (Flaquer et al., 2008; Perry et al., 2001). The availability of extensive data for the bird group (ZW system) has shown that there is a large variation in the size of the PAR within this group (Zhou et al., 2014). For example, in the paleognath group, including emu and ostrich, 65% of the Z chromosome recombines, while only 1% is recombinant in the white-throated tinamou.

Theoretical predictions of NRR expansion suggest that the PAR should evolve towards zero recombination (Otto et al., 2011). However, some PARs, such as those in eutherian mammals, have been maintained over substantial periods of evolutionary time. Several hypotheses can explain the maintenance of the PAR. The first, as stated above, is its role in ensuring the correct segregation of chromosomes by allowing pairing during meiosis (Rouyer et al., 1986). A second mechanism to explain the persistence of PAR is ongoing translocation of autosomal regions containing genes that are sexually antagonistic (Lenormand, 2003). Such translocations are particularly favourable for genomic regions evolving under sexual selection (Charlesworth and Charlesworth, 1980). A final mechanism to explain maintenance of the PAR is the resolution of sexual antagonism by the differential expression of PAR genes between the sexes (sex-biased gene expression), as in such cases it is not necessary for the genes to be subsumed into the SDR to limit their expression to one sex, it is sufficient to express them only in the sex they are advantageous for. Note that, in diploid systems, the existence of XY males and ZW females (which are heterozygous for the sex chromosomes) allows the accumulation of recessive alleles on the X and Z chromosomes that are beneficial for the heterogametic sex. (Rice, 1984) noted that, despite the fact that an X chromosome spends only one third of its time in the male germline, a perfectly recessive allele of an X-linked gene that is favourable to the hemizygous sex (hereafter males) is much more likely to spread than an autosomal counterpart. This is because selection would act strongly on the hemizygotously expressed favourable effects, whereas the deleterious effects in females would initially be masked, owing to heterozygosity in females. In addition, dominant female-beneficial alleles can accumulate on the X chromosome because this chromosome is found preferentially in females (Charlesworth et al., 1987; Rice, 1984).

Study of the PAR of the sex chromosomes of the brown alga *Ectocarpus* has provided a better understanding of how these regions evolve in UV systems. The *Ectocarpus* PARs recombine at a similar rate to autosomal regions of the genome and yet exhibit many features typical of

non-recombining regions (Luthringer et al., 2015). Interestingly, the *Ectocarpus* PAR also appears to play a role during the non-sexual, sporophyte generation. Indeed, recent studies revealed an accumulation of physically-linked clusters of genes with increased expression in the sporophyte (i.e., silenced in the gametophyte) in the PAR (Luthringer et al., 2015).

### 3. Genetic mechanisms that determine sex in eukaryotes: sex determination cascades

#### 3.1 Sex determination cascades

The existence of males and females is due to two processes: sex determination and sexual differentiation. A distinction between these two processes is often made in the biological literature. Sex determination is meant to designate the developmental step at which an individual's fate is irremediably directed towards either the male and female condition, whereas sexual differentiation describes the subsequent developmental steps, during which the male and female phenotypes are progressively built up (Beukeboom and Perrin, 2014).

Sex determination can be mediated by a single gene, located within the SDR (such as *SRY* in therians), while the resulting development of gonads (testis or ovary in animals) is due to primary sex differentiation (Capel, 2017). But there are good reasons to reject this fundamental distinction if we want to understand the evolution of the mechanisms that govern gender determination. These two phenomena are interconnected in a sex determination cascade which is a network of interacting genes that are hierarchically ordered: at the bottom of these cascades lie major effectors, which orchestrate male or female development (to switch on the correct differentiation programme), often described as ‘slave’ genes (Herpin and Schartl, 2015). For example, across the vertebrates, the same genes are often involved in cascades leading to gonad development. Genes from the DM-domain DMRT1 family, for example, play a role in sexual differentiation pathway in most species, but have moved up the hierarchy in some species to occupy the top of the cascade, i.e. have become the sex-determining gene in these species (Picard et al., 2015). For example, loss of *DMRT1* in the male mouse embryo induces transdifferentiation of Sertoli cells into granulosa cells despite the presence of an active *SRY* gene (the sex-determining factor in mammals) (Matson et al., 2011). We can therefore see that differentiation and determination are closely related because even if an *SRY* gene is present, abnormal expression of a differentiation gene can lead to a change of sex. In consequence, we

define sex determination as the whole process that leads to the development of differential reproductive organs (Uller and Helanterä, 2011).

### 3.2 Master sex determination genes

The first genetic clue of the involvement of a gene in sex determination came when Teophilus Shickel Painter, in 1923, discovered that all males were XY and all females were XX. This led to the conclusion that genetic sex was determined when a sperm carrying an X or a Y fuses with an egg carrying an X.

In 1959 two groups examined the circumstances in humans causing Turner's syndrome and Klinefelter's syndrome. In these individuals, the presence of a single X chromosome (XO = Turner syndrome) results in the expression of female characters. In addition, in individuals with Klinefelter's syndrome (XXY), male characters were observed. Thus, sex determination is independent of the number of X chromosomes but depends only on the presence of a Y chromosome. In conclusion, for mammals, the presence of the Y chromosome is sufficient for male determinism and this chromosome carries the genetic information required for male sex determination.

In 1990, a sex-determining gene was identified in the SDR of the Y chromosome in mice: *SRY* (Berta et al., 1990; Sinclair et al., 1990). Conclusive evidence appeared shortly afterwards using transgenic mice. XY mice without *SRY* developed ovaries and followed the female sexual development pathway (Gubbay et al., 1990; Lovell-Badge and Robertson, 1990) and XX mice with a transgenic autosomal copy of *SRY* developed testicles and male secondary sexual characteristics, although they were sterile in the absence of a Y-chromosome, which is required for spermatogenesis (Koopman et al., 1991). Thus, the relationship between the Y chromosome and sex determination was clearly established.

*SRY* allows cells to acquire male cell fate. As stated above, before the action of this gene, the cells are indistinguishable and can have become either male or female: i.e. they are in a state of bipotentiality (DiNapoli and Capel, 2008). *SRY* is expressed in the support cells in the gonads of XY individuals. Support cells are bi-potential cells that can develop into either Sertoli cells in males or follicular cells in females. Primary sex determination, at least in mammals, appears to depend on the cell fate decision that occurs in this precursor population when cells are directed to assume one of the two fates. Spatial and temporal regulation of *SRY* is therefore essential. Gonads represent a unique environment because ectopic expression of *SRY* outside this tissue does not lead to differentiation of Sertoli cells (Kidokoro et al., 2005).

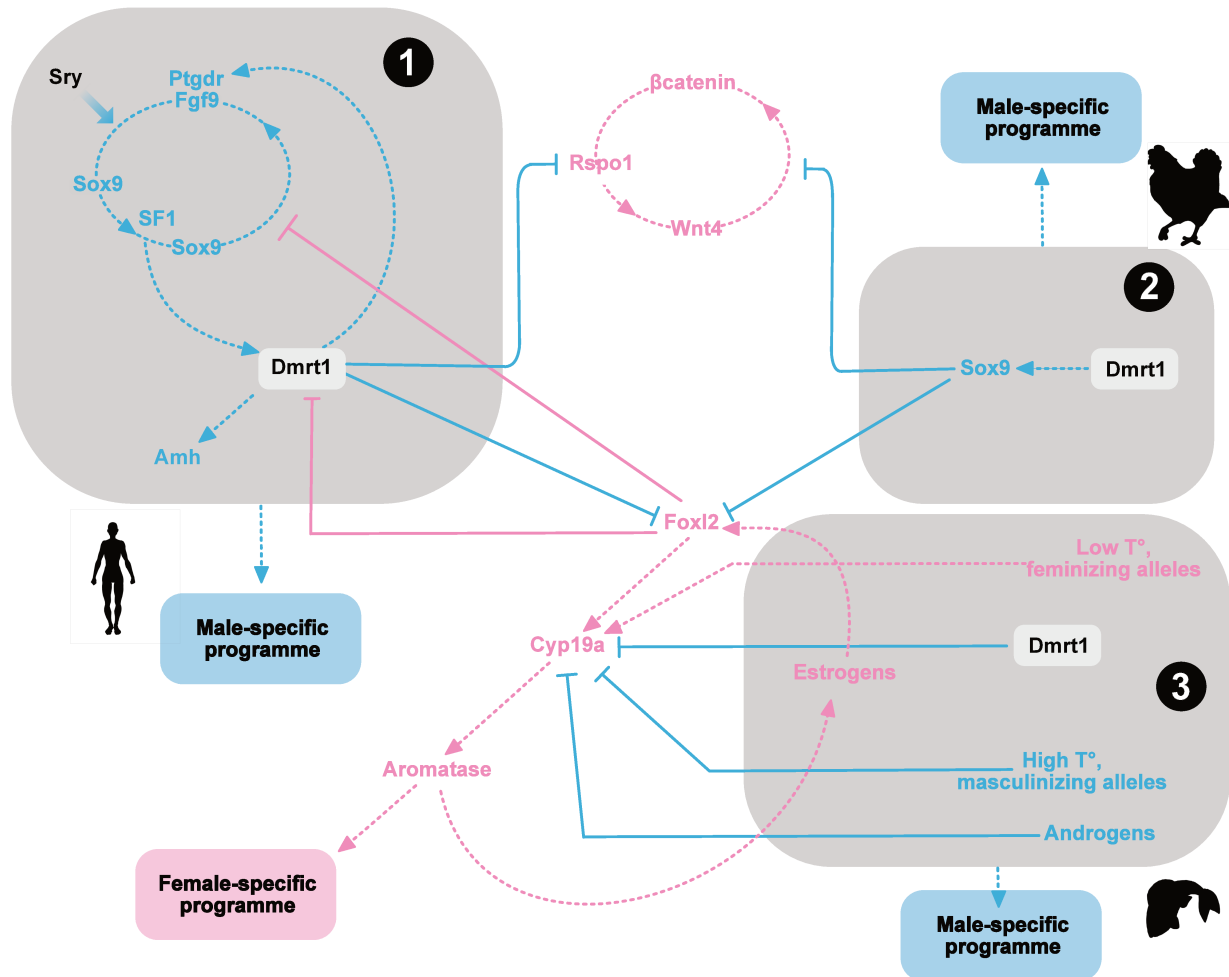


Figure 4. **Sex determination cascade.** (1) The eutherian sex determination pathway. Males (blue) carry the Y chromosome with the dominant *Sry* gene, which activates autosomal *Sox9* in co-operation with *Sf1*. *Sox9* turns on *Fgf9* which in turn upregulates *Dmrt1*. Then *Dmrt1* activate *Amh* and lead to testis development. *Dmrt1* is not only important for maintaining expression of the male pathway but also for suppression of the two female networks (pink). These two female networks involve the *Foxl2* and the *Wnt/β-catenin* signaling pathways. In the absence of *Sry* in XX female (pink), *Sox9* and *Fgf9* are not expressed at levels sufficient to inhibit *Foxl2*, which therefore, in turn, inactivates the *Sox9/Fgf9/Dmrt1* loop, and activates the *Rspo1/Wnt4/β-catenin* pathway, leading to aromatase expression and ovary development. Solid lines define negative regulations. Dashed lines designate positive regulations. (2) Avian sex-determination pathway. Males (blue) have two Z chromosomes and thus a double dose of *Dmrt1*. *Dmrt1* suppresses, possibly through *Sox9*, both the *Rspo1/Wnt4/β-catenin* and the *Foxl2/Cyp19a* pathways, resulting in androgen production and testis development. Females (pink), have a single Z chromosome and thus a single copy of *Dmrt1*. Suppression of *Sox9* results in activation of the *Rspo1/Wnt4/β-catenin* and the *Foxl2/Cyp19a* pathways, oestrogen production and ovary development. (3) Sex determination in fish. Up-regulation of *Cyp19a* by *Foxl2* depends on differential methylation of its promoter by a mix of environmental and genetics effects. High temperature and/or masculinizing alleles lead to high methylation levels that inhibit aromatase expression; the resulting androgens and *Dmrt1* expression drive male development. Conversely, low temperature and/or feminizing alleles, result in low methylation levels, so that aromatase expression converts enough androgens into oestrogens to inhibit *Dmrt1* expression and drive ovary development.

Maintenance of gonadal identity in differentiated gonads is a result of the cross-inhibition activities of *Dmrt1* and *Foxl2*. A critical equilibrium between these conflicting pathways underlies the bipotentiality of the gonadal somatic cells. Tipping the balance in one direction or the other will regulate gonadal fate as a consequence of activation of the male or female pathways. In addition, although male determination factors vary between species, the genes involved in the male and female differentiation pathways directly downstream of this the cascade (i.e. the "slaves") appear to be preserved.

The SRY protein is the founding member of the Sox (Sry-related high mobility group (HMG) box) transcription factor family. The characteristic that defines this family is the presence of an HMG DNA-binding domain. SOX proteins recognize and bind similar DNA motives (A/TAACAA/T) (Mertin et al., 1999). SRY is not the only HMG protein involved in the sex-determination cascade. Sox9, for example, also play a crucial role in sex determination in therians (Gonen and Lovell-Badge, 2019; Kent et al., 1996).

HMG protein involvement in sex determination is not exclusive to mammals. Proteins of this family are also involved, for example, in the determination of mating type in yeast or in the fungus *Phycomyces blakesleeanus* (Idnurm et al., 2008).

Within the animal kingdom, additional classes of gene can also mediate sex determination, notably *TGFβ* in fish such as *Odontesthes hatcheri* (Hattori et al., 2012) and *Oreochromis niloticus* (Li et al., 2015). *TGFβ*, called *amhy* in these species, is a member of the Ahm (anti-müllerian hormone) family that is present only in XY individuals. *Gsdf* (Gonadal soma derived growth factor on the Y chromosome) appears to have supplanted *DMY* as the major sex determination gene in *Oryzias luzoensis* (Myosho et al., 2012). *DMY* is Y-linked sex-determination gene involved in male sex determination in teleost fish such as *Oryzias latipes* or *Oryzias curvinotus* and is a duplicate of the autosomal gene *DMRT1* (Matsuda et al., 2002; Nanda et al., 2002). In trout, the major gene is thought to be *sdY* (sexual dimorphic on the Y chromosome), which has no homology with other known sex determination genes (Yano et al., 2012).

Members of the large DM domain family of genes are involved in sex determination in many animal species. This family was first identified in flies and nematodes, where *Doublesex* (*Dsx*) (Burtis and Baker, 1989; Hildreth, 1965) and *Male abnormal 3* (*Mab3*) (Raymond et al., 1998; Shen and Hodgkin, 1988), respectively, play key roles in sex differentiation. *Dsx* and *Mab3* related transcription factor 1 (*Dmrt1*) was later identified as a key player in the sex-determination pathway of vertebrates. *DMRT1*, which is a member of the DMRT family (for DSX/MAB3-related transcription factor), plays a major sex-determination role in birds. Paralogues of *DMRT1* have also taken on a master sex determination roles in some fish and amphibian lineages. These paralogues mainly act as master regulators of sex determination in ZW system. (Figure 4). For example, in birds sex determination is controlled by dosage of the *Dmrt1* gene on the Z: males have two copies of *DMRT1*, whereas females have only one. Overexpression of *DMRT1* in the gonads of genotypic (ZW) females causes male development,

while suppression of *DMRT1* expression in genotypic (ZZ) males induces female development (Lambeth et al., 2014; Smith et al., 2009).

Thus, *DMRT1* appears to be both a master regulator of sex determination in some species with ZW systems and downstream of the master regulator, in the sexual determination cascade, in other systems, such as XY (Picard et al., 2015). This evidence further underlines the overlap between sex determination and differentiation processes.

### 3.3 "Slave" genes

We have previously seen that young embryos are bi-potential, i.e. have the ability to become either male or female regardless of karyotype: sex is not determined immediately at conception. One crucial issue, given this bipotentiality is to make sure that an individual's commitment is unambiguous; sexual differentiation must ensure proper development towards one or the other sex, and not towards intermediate phenotypes. Sexual identity must also be maintained throughout life. This is one of the crucial roles of the 'slave' genes in the sexual determination cascade. These slaves are the 'central gear' that convey the sexual identity decision from the main effectors in an unambiguous manner (Herpin and Schartl, 2015).

For example, in therians, *Dmrt* genes play a central switch role, by turning on the male programme and turning off the female program (Herpin and Schartl, 2011). *Dmrt1* accomplishes both functions in vertebrates: it inactivates female development by repressing the *Foxl2* and *Rspo1/Wnt4/β-catenin* pathways and activates male development both by triggering (together with *Sox9* and *Fgf9*) downstream male-specific effectors and by maintaining male identity over time, through a feed-back loop based on *Sox9/Sox8* (Figure 4). Loss of *Dmrt1* in embryonic male mice induces transdifferentiation of Sertoli cells into granulosa cells (Matson et al. 2011). The same antagonistic functions are fulfilled in insects by alternative splicing of *Dsx* (*Doublesex*) (Hoshijima et al., 1991). Finally, *Dmrt1* also promotes the maintenance of male sex by maintaining expression of an essential gene of the cascade, *Sox9* (Matson et al., 2011) (Figure 4). In mammals, *Sox9* is expressed in the bi-potential gonads of both genetic sexes and then later, at the time of sex determination, it is overexpressed in the XY gonads and under-expressed in the XX gonads so that it remains expressed in the adult testicles (Jakob and Lovell-Badge, 2011). *Sox9* is the direct target of the *Sry* sex determination factor. *Sry* joins forces with *Sf1* (Steroidogenic factor 1) to bind to the TESCO regulatory element (Testis-specific enhancer of *Sox9*), resulting in positive regulation of *Sox9* (Sekido and Lovell-Badge, 2008). Moreover, *Sox9* self-activates by associating with *Sf1* and binding to the TESCO

element (Sekido and Lovell-Badge, 2009) (Figure 4). The structure of this pathway indicates that Sry is the short-term trigger that initiates sexual determination, while Sox9 allows male characters to be maintained via a self-activating loop. In insects, Tra (Transformer) performs the same function as Sox9 but in the female. Tra is involved in specific splicing of the *dsx* transcript to induce female differentiation. Moreover, Tra ensures its self-activation (Steinmann-Zwicky, 1994). These intermediate factors thus constitute a memory device; the initial trigger is usually only expressed transiently, during a limited window of embryonic development, where it tips the balance towards the male or female fate. The downstream network then establishes an autoregulatory loop that retains memory of the initial decision and ensures lifelong continuity.

### 3.4 Female determination

Sex is a bistable equilibrium, and sex determination a battle between male- and female-specific programmes, occurring through a complex network of genes that repress each other (Kim and Capel, 2006). So far, we have highlighted the importance of male sex determination in eukaryotes, implying that male sex determination is active while female sex determination is passive or occurs by default. However, female sex determination does not always occur by default in all species (Nef and Vassalli, 2009). For example in *Drosophila*, the initial trigger of sex is dependent on the ratio of the number of X chromosomes versus the haploid autosome complement (X:A ratio). In the female, an X:A ratio of one will enable transcription of the *Sex lethal* gene (*Sxl*), a splicing regulator. The SXL protein will then promote female-specific splicing of *Transformer* (*Tra*), a direct downstream target, and lead to the production of functional TRA proteins. Similarly, a complex of the TRA and TRA-2 proteins will then favour female-specific splicing of *Doublesex* (*Dsx*, a Dmrt1 homolog) gene transcripts. This results in the production of the female-type DSX protein DSXF, which initiates up-regulation of the downstream gene-regulatory network for female development. In males, an X:A ratio of 0.5 will prevent production of the SXL protein and result in the production of the male-specific splice form of the *Tra* gene, which is the default form. This splice variant translates to produce a non-functional protein due to a premature stop codon. In the absence of TRA, the male-specific splice form of the *Dsx* gene will be produced by default. The male-type DSX protein DSX M then orchestrates the downstream gene-regulatory network that determines male development (Cho et al., 2007; Cline and Meyer, 1996) (Figure 5).

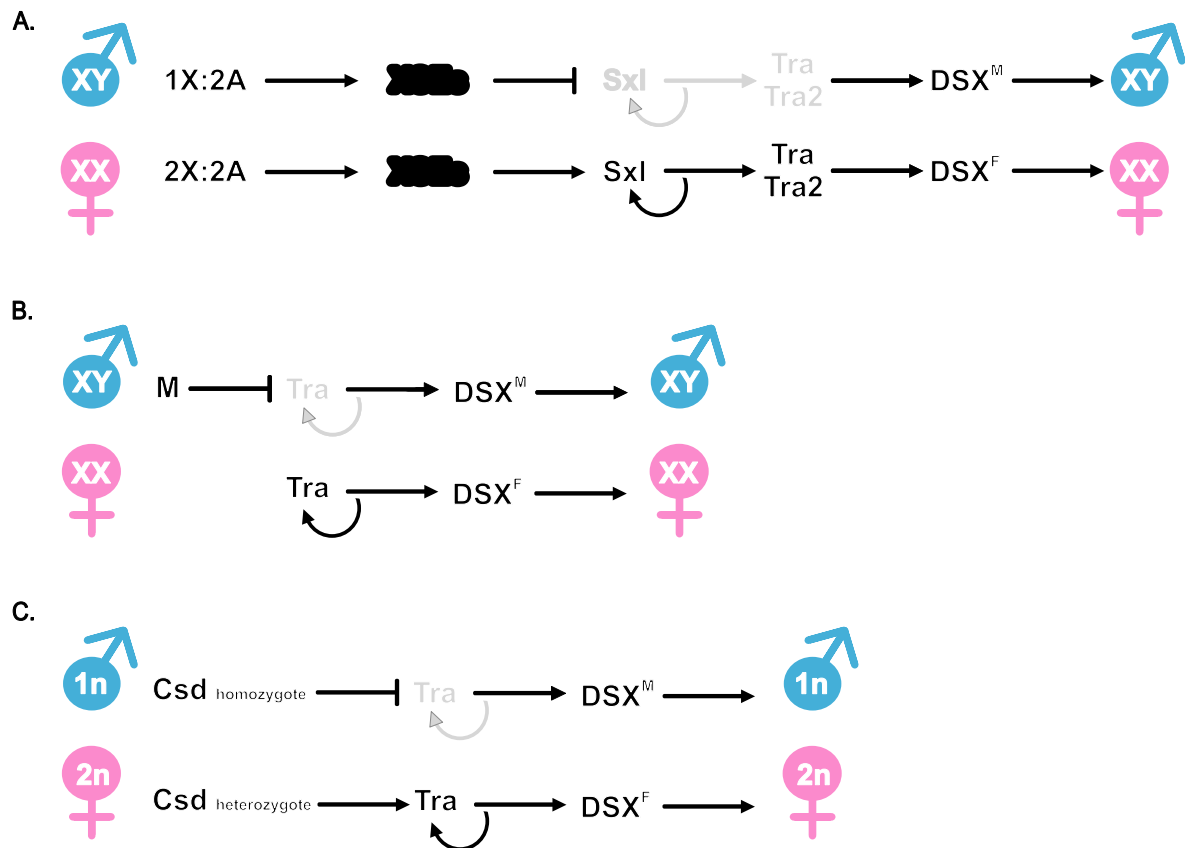


Figure 5. **Sex determination pathways in insects.** (A) The *Drosophila* sex-determination pathway. Males have a single X chromosome which carries several genes (XSEs) but the gene dosage is too low for activation of *Sxl*. This results in male-specific splicing of *Dsx* and male development. Embryos with two X chromosomes have double doses of XSEs which activate *Sxl*. *Sxl* is autoregulated and induces in turn female-specific splicing of *Tra* and of *Dsx*. *Dsx* transcriptionally regulates downstream sex differentiation genes. Activation steps are indicated by arrows and inhibition steps by bars; active genes are black and repressed genes in grey; the higher dose effects of XSEs in females are depicted in bold, the lower dose effects of XSEs in male are depicted by empty letters. (B) The dipteran sex determination pathway. Males carry a Y chromosome with a dominant male determiner M which blocks the *Tra* autoregulation loop, resulting in male-specific splicing of TRA and DSX, and male development. In embryos without a Y chromosome, female-specific splicing of TRA and of DSX causes female sexual differentiation. Activation steps are indicated by arrows and inhibition steps by bars; active genes are in black; repressed genes in grey. (C) The hymenopteran sex determination pathway. Male embryos are haploid and carry a single allele of the *Csd* gene that results in male-specific TRA and DSX and male development. Diploid female embryos carry two different *Csd* alleles. This activates the autoregulation of *Tra* which induces the female-specific splice form of DSX and female sexual differentiation. Activation steps are indicated by arrows and inhibition steps by bars.; active genes are in black; repressed genes in grey. Adapted from (Beukeboom and Perrin, 2014).



In mammals, recently, it has also been shown that female sex determination is not completely passive. (Zhao et al., 2017) have identified a previously unknown role for the *COUP-TFII* gene (or *NR2F2* in humans) in the development of Wolff's ducts, which is an embryological element of the mesonephros that gives origin to male genital glands and ducts. In the female, it is the Müller ducts that develop to give birth to female organs such as the fallopian tubes, uterus, cervix, and the upper two third of the vagina. Female mouse embryos (which have two X chromosomes) that had been genetically modified not to have *COUP-TFII* had both Müller and Wolff ducts derivatives at birth. Differentiation of the Müller duct seems normal in these mice, but the Wolff duct is abnormally preserved. This maintenance is not due to hormonal action of androgen. *COUP-TFII* could specifically participate in the regression of the Wolff's duct in females, highlighting active repression of the male pathway in the female.

Moreover, the *Rspo1/Wnt4/β-catenin* pathway induces female development. The *Wnt4* canonical signalling pathway is crucial for ovarian differentiation and for the regulation and formation of Müller ducts (Parr and McMahon, 1998) or steroidogenesis (Lapointe and Boerboom, 2011). Inactivation of *Wnt4* or *Rspo1* leads to masculinisation of the female gonad (Tomizuka et al., 2008), while overexpression of either gene leads to feminization of the male gonad (Jordan et al., 2003). *Wnt-4* and *Rspo1* therefore appear to be key factors in ovarian differentiation. In non-mammalian vertebrates, studies show an involvement of the canonical *Wnt* pathway during sexual differentiation in species with various modes of sexual determinism such as chickens (*ZW/ZZ GSD*) (Smith et al., 2008), lizards (*ESD*) (Tripathi and Raman, 2010) and fish such as the medaka, where overexpression of *Rspo1* in XY gonads induces sexual reversion (Zhou et al., 2016). These observations suggest a conserved role for the *Wnt* pathway in sex determination in vertebrates (Figure 4).

### 3.5 "Masters change, slaves remain"

The downstream components of the master sex determination gene appear to be more conserved from an evolutionary point of view than the master regulator and seem to converge towards the regulation of some central common effectors. An example, described earlier, that illustrates this paradigm is the master sex-determining gene in mammals, *SRY*. No corresponding homologue has been detected outside therian mammals (marsupials and placentals). Conversely, genes that act downstream of *SRY*, encoding either transcription factors (*SOX9*, *DMRT1*) or components of signalling pathways (*TGF-β/Amh*, *Wnt4/β-catenin*, *Hedgehog*) and genes involved in *SRY*

regulation (*SFI*, *WT1*) have known or presumed roles in gonadogenesis or gonadal differentiation in several vertebrate species, and even in some non-vertebrate deuterostomes and protostomes. These results suggest that a central paradigm of sex determination is that "masters change, slaves remain" (Graham et al., 2003; Haag and Doty, 2005; Herpin and Scharl, 2008) (Figure 4).

Another example is *dsx*, which is spliced sex-specifically and has distinct DsxM and DsxF regulatory activities in dipterans, moths, beetles, and Hymenoptera (Cho et al., 2007; Oliveira et al., 2009; Salvemini et al., 2011; Suzuki et al., 2005). In all these insects, the *tra* transcript is also spliced sex-specifically and controls the sex-specific splicing of *dsx*, suggesting that sex determination based on alternative splicing of *dsx* is common to all insects (Gempe and Beye, 2011; Verhulst et al., 2010). Insects also illustrate divergence of sex determination and a tendency for "slave" pathways to be conserved. For example, in *Drosophila*, it is the balance between the number of X chromosomes and the number of autosomes that triggers differentiation, in dipterans it is a male determinant on the Y chromosome that induces the male pathway and in hymenoptera, the male is haploid and has only one *Csd* allele, which induces the male pathway. In all three cases, the *Tra* gene is inhibited in males, inducing male-specific splicing of the *Dsx* gene (Beukeboom and Perrin, 2014) (Figure 5).

## 4. Epigenetic mechanisms that determine sex in eukaryotes

### 4.1 Environmental sexual determination (ESD)

In environmental sexual determination (ESD) systems the activation of sex-determination cascades is influenced by the action of external factors such as temperature. ESD is phylogenetically dispersed, being found in diverse taxa such as plants, nematodes, amphipods, molluscs, fish and amniotic vertebrates (Janzen and Phillips, 2006). This type of sex determination can involve responses to either biotic or abiotic factors. Abiotic factors can include photoperiod, pH, oxygen level, food availability or temperature.

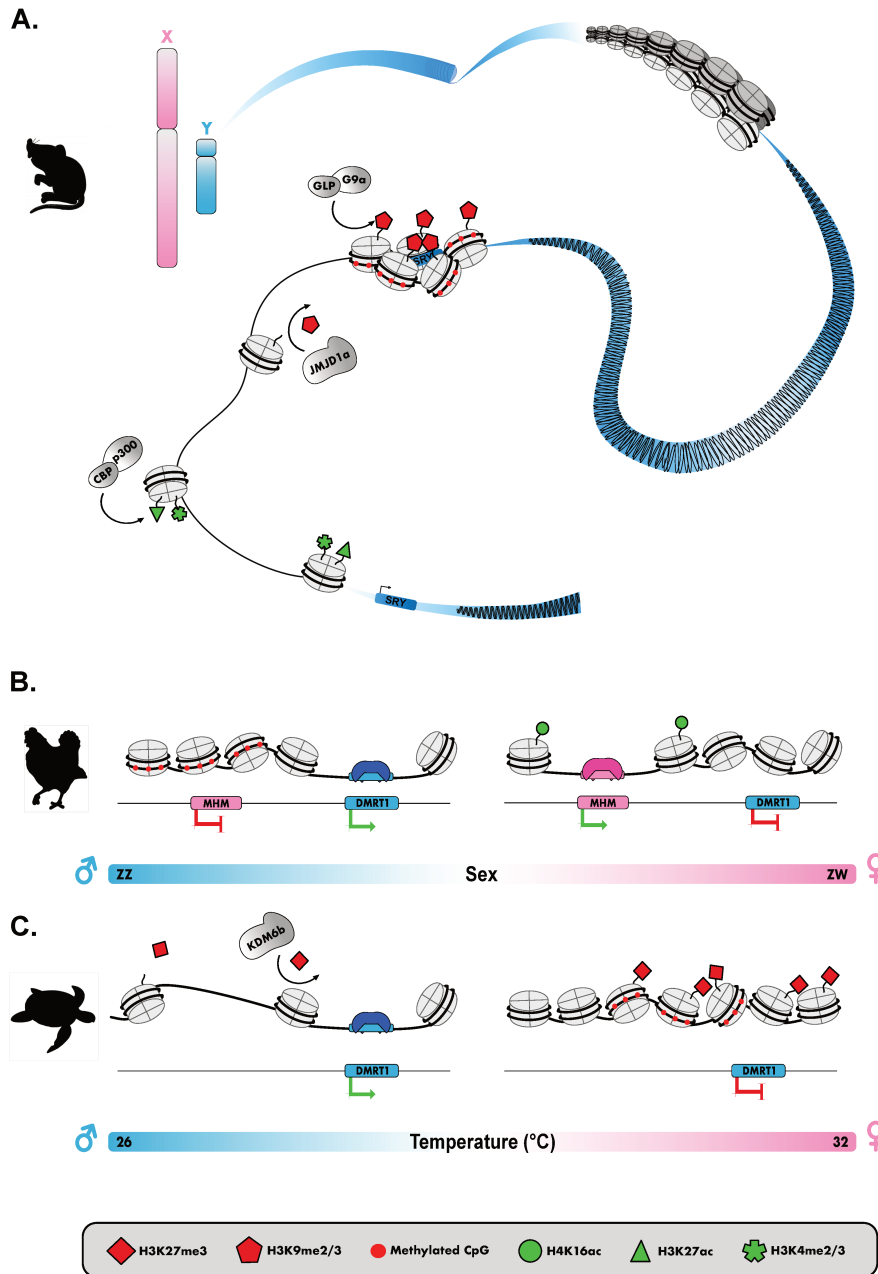
For example, the temperature at which embryos are incubated will determine their sex in most turtles and crocodiles and in some fish (Bull and Vogt, 1979; Ospina-Álvarez and Piferrer, 2008; Woodward and Murray, 1993). As far as biotic factors are concerned, social cues can determine which sex will develop. In the worm *Bonellia viridis* sex determination depends on where the larva settles. If the larva settles on the seafloor, it develops into a macroscopic female. On the contrary if a larva settles on a female, it migrates inside the female and develops as a

microscopic male (Berec et al., 2005). In some fishes the presence or absence of the other sex can trigger the development of a male or a female (Godwin et al., 2003). In *Crepidula fornicata*, individuals create a mound and the development of either males or females depends on their position on this mound (Coe, 1936). Furthermore, social cues can also have an important role in sex determination in plants, mosses or in reef fishes (Banks et al., 1993; Godwin, 2009; Tanurdzic and Banks, 2004).

The theoretical advantage of ESD is that it provides a population with the ability to adapt to unstable and changing environments by modulating the sex ratio of the population (Bull, 1985). If favourable growth conditions improve female fitness over male fitness, then embryos should develop into females in such an environment, and into males otherwise. On the other hand, this capacity of adaptation of the population has a cost in that the sex ratio is very dependent on the environment, leading to strong sex ratio bias if the conditions are favourable to one of the sexes over a long period of time.

#### 4.2 Continuum between GSD and ESD

Interestingly, frequent evolutionary transitions between GSD and ESD have been observed in phylogenetic trees, for example in turtles and lizards, based on the use of cytological techniques to identify sex chromosomes (Janzen and Phillips, 2006). GSD and ESD have long been seen as resulting from distinct and mutually exclusive processes, occurring with different timings, and underlain by different mechanisms but recent studies indicate a continuum between GSD and ESD with many systems being mixed, making this dichotomous view less acceptable. For example, zebra fish sex determination is mediated by a combination of genetic and environmental factors. The different genomic regions involved in sex determination identified by Bradley et al. (2011) account for only 16% of the phenotypic variance of sex. Much of the remainder is due to environmental or random effects. Environmental effects may also co-occur with well-defined male or female heterogamety. In medaka fish, maleness is normally determined by the extra copy of *Dmrt1* on the Y chromosome (*DmY*) (Masuyama et al., 2012). However, the autosomal copy of *Dmrt1* is upregulated at high temperature (32°C), leading some XX individuals to develop as males (Hattori et al., 2007; Sato et al., 2005).



**Figure 6. Epigenetic regulation of male sex-determining factors** (A) Epigenetic regulation of *Sry* in mice. Prior to sex determination *Sry* is repressed and enriched for the GLP/G9a-mediated, repression-associated histone modification H3K9me2 and for methylated CpGs sites. *Sry* activation requires Jmjd1a-mediated removal of H3K9me2, CBP/p300-mediated deposition of H3K27ac, and accumulation of the activation-associated H3K4me2 at its promoter. CpGs become demethylated. (B) Epigenetic regulation of *DMRT1* in chickens. In males (ZZ, left), a region neighboring the male-determining gene *DMRT1* is hypermethylated (MHM) and repressed. In females (ZW, right), MHM is hypomethylated and enriched for the activation-associated histone modification H4K16ac. The open chromatin conformation of MHM in females enables transcription of a lncRNA that inhibits expression of *DMRT1*. (C) Epigenetic regulation of *Dmrt1* in red-eared slider turtles. During embryogenesis, turtles develop as males if eggs are incubated at lower temperatures (blue, left) and as females at higher temperatures (pink, right). At male-producing temperatures, *Kdm6b* is upregulated. KDM6b-mediated removal of the repression-associated mark H3K27me3 and demethylation of CpGs at the *Dmrt1* promoter is required for *Dmrt1* activation and subsequent testis development. At female-producing temperatures, enrichment of H3K27me3 and CpG methylation at the *Dmrt1* promoter inhibits its expression and leads to ovary development.

### 4.3 Chromatin dynamics and regulation of sex determination cascades

Epigenetic sex determination therefore involves external factors that will influence the development of male or female individuals. But epigenetics, and more particularly chromatin dynamics, is also involved in the regulation of sex determination genes in GSD systems. Indeed, if we take the case of *Sry*, this gene must be expressed in specific tissues and at a specific time. This pattern of expression is partly mediated by chromatin dynamics, which is also involved in the regulation of many developmental mechanisms.

In at least nine species of Akodon South American grass mice, females can be either XX or XY (Sánchez et al., 2010). In *Akodon azarae*, although the *Sry* gene is present with no apparent mutations within the coding region, delayed or deficient expression levels due to epigenetic modifications are responsible for sex reversal in XY females. In this species, Y chromosomes that escape epigenetic silencing give rise to XY males. A high proportion of XY females carry a rearranged X chromosome capable of triggering a normal female ovarian phenotype in XY animals (Veyrunes et al., 2010). The gene responsible for the X's feminizing influence has not yet been identified.

There is evidence for involvement of epigenetic mechanisms in both the initiation and in the stabilization and maintenance of sex determination in humans and mice. For example, an unmethylated CTCF-binding site mapped upstream of the human *Sry* gene in white blood cells was associated with enrichment of the histone H3 lysine 9 trimethylation (H3K9me3) mark, indicating recruitment of Polycomb repressive complex 2 (PRC2) to silence the locus (Singh et al., 2011). Consistent with the idea that activation of *Sry* requires depletion of H3K9me3, XY mice deficient for the H3K9-demethylating enzyme JMJD1A showed an increase in H3K9 dimethylation and a decrease in the activation-associated mark H3K4 trimethylation across the locus, leading to a high frequency of sex reversal (Kuroki et al., 2013). Thus, before male determination, *Sry* is repressed and enriched for the repressive modification of the histone H3K9me2 deposited by GLP/G9a. Activation of *Sry* requires the elimination of H3K9me2 by *JmJd1a* but also the demethylation of the promoter region and accumulation of the activation-associated marks H3K4me2 and H3K27ac as a result of the activity of CBP/p300 (Carré et al., 2018) (Figure 6).

Further evidence for epigenetic regulation of *Sry* came from studies of *chromobox protein homologue 2* (*Cbx2*). Loss of function of *Cbx2* in mice led to hypoplastic gonads and male-to-

female sex reversal, which could be rescued by forced expression of *Sry* (Katoh-Fukui et al., 2012, 1998). This role for CBX2 as an activator of *Sry* was unexpected, given the role of this protein in the repressive complex PRC1 (Lanzuolo and Orlando, 2012), although the effect on *Sry* may be mediated by repression of a repressor.

Consistent with the role of *Cbx2* in mice, a human patient with male-to-female sex reversal was found to carry a mutation in *CBX2* (Biaison-Lauber et al., 2009). Chromatin immunoprecipitation of CBX2 in Sertoli-like cells identified genomic targets associated with both male and female pathways (Eid et al., 2015). Several lines of evidence suggest that the female pathway must be silenced to establish the male pathway and vice versa (Maatouk et al., 2017). In cases where Wnt signalling is not silenced, the male pathway is not stabilized, despite the activation of *Sry* and *Sox9* (Jameson et al., 2012; Maatouk et al., 2008). CBX2 may be involved in blocking expression of genes associated with female fate, which would otherwise disrupt commitment to male fate.

SRY is not the only master regulator that is controlled by chromatin dynamics. For example, in ZZ males chickens, a region neighbouring the male-determining gene *DMRT1* is hypermethylated (MHM) and repressed. In females (ZW), MHM is hypomethylated and enriched for the activation-associated histone mark H4K16ac. The open chromatin conformation of MHM in females enables transcription of a lncRNA that inhibits expression of *DMRT1* (Bisoni et al., 2005; Itoh et al., 2011; Teranishi et al., 2001). (Figure 6)

Finally, during embryogenesis, turtles develop as males if eggs are incubated at lower temperatures and as females at higher temperatures. At male-producing temperatures, *Kdm6b* is upregulated. KDM6b-mediated removal of the repression-associated mark H3K27me3 and demethylation of CpGs at the promoter of *Dmrt1* is required for its activation and subsequent testis development. At female-producing temperatures, enrichment of H3K27me3 and CpG methylation at the promoter of *Dmrt1* inhibits its expression and leads to ovary development (Ge et al., 2017) (Figure 6).

## 5. Sex differentiation: sex biased gene expression

Differences between males and females are largely due to differences in the expression of autosomal genes (Grath and Parsch, 2016). In fact, only 10% of sex reversal cases are linked to a mutation in *SRY* (DiNapoli and Capel, 2008), indicating that other critical genes are involved in the development of the male sex. These genes are mostly downstream of the sex

determination cascade and are part of a genetic pathway responsible for male sex differentiation. In many animal and plant species, male morphology, physiology and behaviour differ significantly when compared to females. All these differences are mediated by the action of genes that are differentially expressed between the two sexes. This differential expression leads to the development of sexual dimorphism.

At the genetic level, these majority of these genes are present in both males and females. Indeed, differences between the sexes may be completely absent, as is the case for environmental sex determination, or differences may occur but be mainly limited to a single sex-specific chromosome (with exceptions such as the duck-billed platypus, for example, which has ten sex chromosomes). Examples include the Y chromosomes of mammals, the W chromosomes of birds and the U and V chromosome of some alga.

To understand how sexual differentiation is implemented, transcriptional profiling methods have been used to compare gene expression in females and males on a genome-wide scale. These studies have shown that the multitudes of different morphologies resulting from limited genetic divergence are correlated with differential gene expression of many autosomal genes. Most morphological differences between females and males are caused by the differential expression of genes that are present in both sexes, a phenomenon known as sex-biased gene expression (Ellegren and Parsch, 2007). Sex-biased genes can be classified as either male-biased or female-biased, depending on which sex has the highest expression level. Studies of different tissues in mammals, birds and fish have shown that the expression of sex-biased genes tends to be highest in the gonads and lower in other tissues, such as the heart, spleen, and brain (Catalán et al., 2012; Goldman and Arbeitman, 2007; Harrison et al., 2015; Mank et al., 2008; Parisi et al., 2003; Pointer et al., 2013; Wong et al., 2014). This is not surprising, as the gonads show much greater sexual dimorphism than these other tissues, although some exceptions are to be noted, such as the *D. melanogaster* fly Malpighian tubule which is analogous to the mammalian kidney and thought to perform the same functions in both sexes, shows a relatively high level of sex-biased expression (Huylmans and Parsch, 2014). In addition, the sex-biased gene expression is highly variable over the life of an individual. In general, the amount and the magnitude of sex-biased gene expression tends to increase during development, with low levels in embryonic stages and high levels in sexually mature adults (Mank et al., 2010; Perry et al., 2014). However, this is not necessarily the case. For example, in chicken gonads, 35% of genes show sex-biased expression only in adults, 5% show sex-biased expression only during embryonic stages, and only 1% show consistent sex-biased expression across all stages (Mank

et al., 2010), whilst in *D. melanogaster* gonads, 10% of genes show sex-biased expression only in adults, 11% show sex-biased expression only during larval stages, and 30% show consistent sex-biased expression across all stages.

Elucidating the molecular mechanisms that lead to sex-biased gene expression is essential to understanding gene regulation, epigenetics, and developmental biology. Although the early stages of sex determination are well understood in a diverse range of species (Bachtrog et al., 2014), the molecular mechanisms responsible for regulating the sex-biased expression of hundreds, or even thousands, of genes in developing and adult organisms are largely unknown. One possible mechanism regulating the differential regulation of these genes in the two sexes is physical alteration of the DNA or chromatin without modification of the DNA sequence. This is commonly referred to as epigenetic modification or chromatin dynamics and can occur through mechanisms such as DNA methylation and histone modification. DNA methylation, affects gene regulation in all kingdoms of life (Bird, 2002). For example, DNA methylation plays an important role in differentiating female morphs (workers and queens) in the honeybee (*Apis mellifera*) (Elango et al., 2009; Kucharski et al., 2008). The study of chromatin dynamics is therefore crucial to understand the underlying mechanisms of sex differentiation between males and females.

## 6. The role of chromatin in the regulation of gene expression:

Epigenetics is the study of changes in gene expression, that do not involve modification of the DNA sequence (Riggs and Porter, 1996). Epigenetic processes, including DNA methylation, histone modification and various RNA-mediated processes, are thought to influence gene expression mainly at the transcriptional level; however, other steps (for example, translation) may also be regulated epigenetically (Goldberg et al., 2007).

It is important to note that, unlike genetic mutations which directly and permanently affect the expression of genes, epigenetic changes are reversible. Specifically, the mechanisms that govern epigenetic modifications are reversible, as methylation is associated with demethylation and acetylation with deacetylation, for example.

Although all the cells in an organism contain essentially the same DNA, cell types and functions differ because of qualitative and quantitative differences in gene expression.



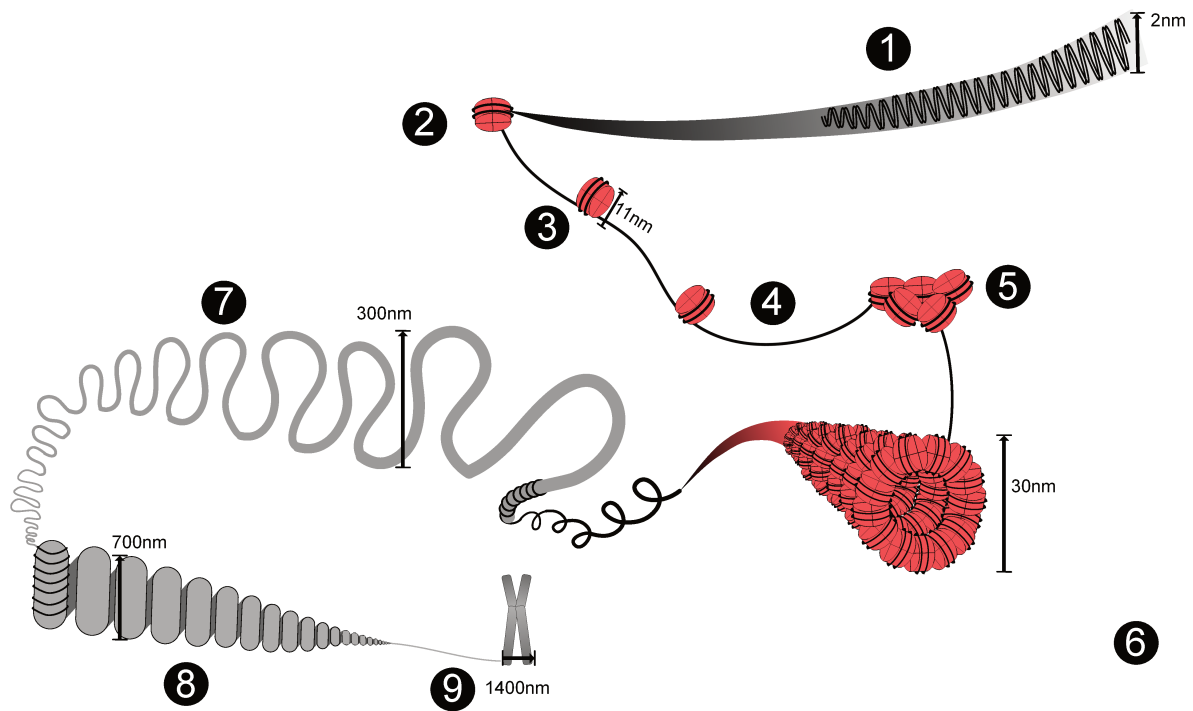


Figure 7. **Chromatin compaction.** Chromosomal DNA is packaged inside microscopic nuclei with the help of histones. These are positively-charged proteins that strongly adhere to negatively-charged DNA and form complexes called nucleosomes. Each nucleosome is composed of DNA wound 1.65 times around eight histone proteins. Nucleosomes associate to form a 30-nanometer chromatin fiber, which forms loops averaging 300 nanometers in length. The 300 nm fibers are compressed and folded to produce a 250 nm-wide fiber, which is tightly coiled into the chromatid of a chromosome. (1) At the simplest level, chromatin is a double stranded helical structure of DNA. (2) DNA is complexed with histones to form nucleosomes. (3) Each nucleosome consists of eight histone proteins around which the DNA wraps 1.65 times. (4) When the chromatin is decondensed, we speak about euchromatin. (5) On the other hand, when the chromatin is compacted, we speak about heterochromatin. (6) The nucleosomes associate to produce a 30-nm fiber (7) that forms loops averaging 300 nm in length. (8) The 300-nm fibers are compressed and folded to produce a 250 nm wide fiber. (9) Tight coiling of the 250-nm fiber produces the chromatid of a chromosome. Adapted from (Pierce and Benjamin 2013).

Since each tissue has very different genetic needs, very precise mechanisms of gene regulation must be put in place. The existence of distinct epigenetic profiles explains why the same genotype can generate different phenotypes (Berger et al., 2009).

Physical access to DNA, enabled by epigenetic modifications, is a dynamic property of chromatin that allows the establishment and maintenance of cellular identity. Chromatin is a complex of DNA and proteins found in eukaryotic cells (Figure 7). Chromatin is organized into two main forms throughout the genome, depending on whether it is compacted (heterochromatin) or open (euchromatin). This accessibility landscape changes dynamically in response to external stimuli or developmental factors. Chromatin organization is mediated by an interaction network involving enhancers, promoters, insulators and chromatin-binding factors, and the entire network regulates gene expression.

### 6.1 Chromatin, the carrier of genetic information.

In eukaryotic cells, DNA is condensed in the nucleus in the form of a nucleoprotein structure, called chromatin (Figure 7). Beyond the DNA condensation function, chromatin plays an essential dynamic role in many cellular processes such as DNA transcription, replication, recombination and repair (Koprinarova et al., 2016; Li et al., 2007).

The basic unit of chromatin is the nucleosome. The latter is composed of a histone octamer around which 147 pairs of DNA bases are wrapped (Figure 7). Each octamer consists of two H2A-H2B dimers and one H3-H4 tetramer (Kornberg and Thomas, 1974). The linker histone H1 participates in the formation of higher order structures of compacted chromatin (Fyodorov et al., 2018). Compaction is highest during mitosis when the chromosome condenses. In non-mitotic cells, chromatin is distributed in the nuclei in condensed regions of heterochromatin and in open regions of euchromatin (Woodcock and Ghosh, 2010).

The two states of compaction, heterochromatin and euchromatin, have distinct roles. For example, transcription is largely restricted to euchromatin. There are two types of heterochromatin: constitutive and optional. Constitutive heterochromatin has a permanently condensed architecture, few genes and is enriched with repeated sequences. Optional heterochromatin is also transcriptionally inactive but has the potential to convert to euchromatin, allowing transcription in certain spatial/temporal contexts (e.g. developmental stages or specific stages of the cell cycle) or regulated by hereditary (monoallelic expression of genes subject to parental genomic imprinting) (Trojer and Reinberg, 2007). DNA compaction can affect the actions of different proteins that require direct access to DNA such as

transcriptional factors, RNA polymerase, and enzymes involved in DNA recombination and repair. During embryogenesis, for example, the amount of facultative heterochromatin increases as unwanted sets of genes are progressively shut down until at maturity, a cell expresses only the genes appropriate for that tissue (Woodcock and Ghosh, 2010). The opposite also occurs, for example, with differentiated cells that are reprogrammed to become stem cells. These events are generally accompanied by profound changes in histone variants, histone modifications and the presence of chromatin architectural proteins (CAPs) (Woodcock and Ghosh, 2010).

## 6.2 DNA methylation

DNA methylation is perhaps the best characterized chemical modification of chromatin. This modification was found in 1948 when Hotchkiss discovered that cytosine residues could be methylated on the fifth carbon of the pyrimidic cycle if they were followed of a guanine in the DNA sequence (Hotchkiss, 1948). In mammals, nearly all DNA methylation occurs on cytosine residues of CpG dinucleotides. Regions of the genome that have a high density of CpGs are referred to as CpG islands and DNA methylation of these islands correlates with transcriptional repression (Goll and Bestor, 2005).

The cytosine residues of these CpG dinucleotides are methylated by C5-Cytosine DNA methyltransferases (DNMT). In mammals, DNMT3a and DNMT3b are responsible of *de novo* deposition of methyl groups, whereas DNMT1 is required for their maintenance during chromosome replication (Chen and Li, 2006). Binding of transcription factors to promoters can be reduced or prevented following methylation of the DNA in these regions (O'Malley et al., 2016; Yin et al., 2017). This direct effect on transcription factor binding contributes to the repressive effect of methylated C5-Cytosine on gene expression. Finally, methylated CpG dimers are bound by methyl-CpG-binding protein, which recruits repressor complexes associated with histone deacetylase activities (Deaton and Bird, 2011). Note that DNA can also be methylated on adenine residues by N6-A methyltransferases (N6-A MTases) (Iyer et al., 2016).

DNA methylation plays a role in many cellular processes from fungi to mammals, including silencing of repetitive and centromeric sequences; X chromosome inactivation in female mammals; and mammalian imprinting, all of which can be stably maintained (Yang and Kuroda, 2007). DNA methylation provides a stable, heritable, and critical component of epigenetic regulation, in addition to histone post-translational modifications.

### 6.3 Canonical histones and histone variants

We have seen previously that there are four canonical histones, H2A/H2B and H3/H4. These histones form the nucleosome core particle. All four core histones have a three dimensional structure with disordered N- and C-terminal tails and a central histone-fold domain (HFD). This domain, which is conserved in the archaea and the eukaryotes, consists of three alpha propellers separated by two short loops (Arents and Moudrianakis, 1995).-The HFD facilitates histone heterodimerisation during nucleosome assembly.

In addition to these 4 canonical histones, there are several histone variants, which can replace conventional histones within nucleosomes. Histone variants differ from canonical histones mainly in that their genes are expressed outside the S phase and the proteins are incorporated into chromatin independently of DNA replication (Li et al., 2007). These histone variants differ in their tails (MacroH2A; Doyen et al., 2006), in the structure of their histone-fold domain (H2ABdb; Doyen et al., 2006b) or simply at a number of residues (H3.3; Henikoff and Ahmad, 2005). Some histone variants have specialized functions in processes such as DNA repair (H2AX, H2A.Z or macroH2A), spermatogenesis (TSH2B) and kinetochore assembly (CENP-A) (De Rop et al., 2012; Montellier et al., 2013; Yuan et al., 2010).

Histone variants exhibit differences in timing of expression (i.e. are DNA replication-independent) and their mRNAs have different characteristics from those of canonical histones. The incorporation of histone variants into chromatin induces nucleosome restructuring and helps to create functionally distinct chromatin domains. Several histone variants have been implicated in the regulation of cellular processes such as DNA repair and transcriptional activity. Replacement of canonical histones by histone variants changes the position of nucleosomes and therefore influences gene expression (Portela and Esteller, 2010; Santenard and Torres-Padilla, 2009). More particularly, histone variants are involved in the inactivation of the X chromosome in female mammals (Chadwick and Willard, 2002). For example, the histone variant MacroH2A is enriched in the compacted chromatin of the inactive X chromosome and therefore appears to play a role in transcriptional repression during meiotic sex chromosome inactivation (Chadwick and Willard, 2001). The histone-fold domain and long C-terminal tail of MacroH2A are involved in transcriptional repression of genes through interaction with other proteins (Kozlowski et al., 2018) and in sterically blocking the binding of transcription factors and coactivators (Abbott et al., 2005; Kamakaka and Biggins, 2005; Talbert and Henikoff, 2010).

Unlike other H2A histones, the H2A-Bdb variant does not have a C-terminal tail. This property is thought to lead to destabilization of nucleosomes, facilitating transcription. The H2A-Bdb variant is located in active X chromosomes and in autosomes (Li et al., 2007).

Histone variants appear to retain most of the post-translational modification sites of canonical histones, allowing recognition of nucleosomes by chromatin regulatory proteins (Li et al., 2007; McKittrick et al., 2004), which is another mechanism involved in chromatin dynamics.

## 6.4 Histone modifications

We have previously seen that nucleosome formation involves the histone fold domain at globular C-terminal end of histone proteins. Their N-terminal domains also play important roles. These N-terminal domains emerge from the surface of nucleosomes and are the preferred target of many enzymes that mediate post-translational histone modifications. These modifications include methylation, acetylation and ubiquitination but also the more recently discovered propionylation, butyrylation, crotonylation, malonylation, succinylation, formylation, citrullination, phosphorylation and hydroxylation (Huang et al., 2014). Most of these modifications occur on lysines, serines and threonines, which are very abundant in histones. Post-translational histone modifications are also a key process in chromatin dynamics since they directly influence the accessibility of chromatin (Bannister and Kouzarides, 2011; Sharma et al., 2010). Thus, depending on the residue deposited and the type of modification present, histone modifications can lead to the activation or repression of genes.

Changes in histones modify chromatin by two main mechanisms. The first corresponds to modifications that directly influence the overall structure of the chromatin. For example, acetylation and phosphorylation of histones effectively reduces the positive charge of histones, which can disrupt electrostatic interactions between histones and DNA. This probably leads to a less compact chromatin structure, thus facilitating access to DNA for proteins such transcription factors. The second mechanism involves modifying the regulation (positive or negative) of the binding of effector molecules. Effector molecules are capable of recognizing certain histone modifications. For example, HP1 (heterochromatin protein 1) recognizes the mark associated with repressive heterochromatin, H3K9me3. This interaction is important for maintaining the structure of heterochromatin (Bannister et al., 2001).

Histone acetylation is a histone post-translational modification that controls the compaction of chromatin. (Allfrey et al., 1964) were the first to highlight this phenomenon. It has since been shown that lysine acetylation is highly dynamic and is regulated by the opposing action of two

families of enzymes: histone acetyltransferases (HATs) and histone deacetylases (HDACs) (Xhemalce et al., 2011). Acetylation the N-terminal tails of histones on (positively charged) lysine residues weakens interaction with DNA, which is negatively charged. Chromatin then adopts a relaxed structure that is more permissive to transcription (Bannister and Kouzarides, 2011; Yang and Seto, 2007). In addition to its direct action on the accessibility of chromatin, histone acetylation also promotes transcription by creating binding sites for proteins involved in gene activation. This is the case, for example, for proteins of the bromodomain family, which are known to function as transcriptional regulators (Denis et al., 2006). Bromodomain proteins combine with HATs to form chromatin remodelling complexes at acetylated lysines (Bannister and Kouzarides, 2011; Yang and Seto, 2007). Unlike HATs, HDACs reverse the acetylation of lysine residues, restoring their positive charge and stabilizing the chromatin architecture. Thus, HDACs are mainly associated with repressive complexes. Unlike HATs, these enzymes are not very specific and can deacetylate multiple sites within histones (Arrowsmith et al., 2012; Bannister and Kouzarides, 2011).

Histone methylation occurs on lysine or arginine residues and may involve mono-, di- or tri-methylation and mono- or di-methylation, respectively (Ng et al., 2009). Like acetylation, histone methylation is a dynamic process that involves two classes of antagonistic enzymes: histone methyltransferases (HMTs) which, through their SET domain, catalyse the transfer of a methyl group, and demethylase proteins of either the Jumonji protein family (Tsukada et al., 2006) or lysine-specific histone demethylase (LSD) (Shi et al., 2004). Unlike acetylation, histone methylation does not alter the charge of the histone protein and therefore has no direct effect on chromatin structure (Arrowsmith et al., 2012; Bannister and Kouzarides, 2011). Therefore, changes in chromatin dynamics induced by histone methylation depend on the binding of proteins and specific complexes to these post translational modifications (PTMs). Thus, depending on the type of fixed complex, histone methylation leads to activation or repression of genes. For example, tri-methylation of lysine 4 on histone H3 (H3K4me3) is associated with transcriptionally active promoters (Liang et al., 2004) while trimethylation of H3K9 (H3K9me3) and H3K27 (H3K27me3) occurs at the promoters of repressed genes (Arrowsmith et al., 2012; Bannister and Kouzarides, 2011; Kouzarides, 2007; Sharma et al., 2010).

The various post-translational histone modifications described above act together in a complex dynamic that has been termed the histone code. Histones can undergo post-translational modifications on various sites simultaneously. Collectively, these modifications constitute the

code and define the structure and function of a chromatin region (Ernst and Kellis, 2010; Strahl and Allis, 2000). Modification can also interact sequentially, with some PTMs inducing or inhibiting other PTMs. This communication between PTMs can take place in cis (i.e. on the same histone), in trans (i.e. on two different histone proteins), or even between two different nucleosomes (Lee et al., 2010). There are multiple mechanisms at the origin of this communication. For example, competitive antagonism may occur if modifications such as methylation or acetylation target the same residue. In addition, some changes are dependent on the prior occurrence of another mark (e.g. H2B ubiquitinylation results in H3K4 methylation) (Kim et al., 2009).

To summarise, post-translational histone modifications are information patterns interpreted by large multi-protein complexes that contribute to chromatin remodelling and transcription regulation.

## 7. The brown alga *Ectocarpus*: a new model to study the evolution of sex determination

### 7.1 Context

Research on sex determination has focused mainly on animals, plants and fungi. Moreover, accumulated data on genetic sex determination mainly concerns XY and ZW systems. Thus, information on sex determination in other eukaryotes is very limited.

In order to fill this gap, studies of phylogenetically distant species from large eukaryotic groups need to be carried out. Brown algae belong to the stramenopile group, a phylogenetically group almost as far from the green lineage (Archaeplastida) as from animals (opisthokonts) (over 1 billion years old). In addition, brown algae are one of only a small number of eukaryotic lineages to have acquired complex multicellularity.

Analysis of brown algal life cycles, mechanisms of sex determination and differentiation and the evolution of brown algal sex chromosomes can be particularly interesting in relation to other large eukaryotic groups, due to their evolutionary distance. The mechanisms identified may be either conserved or completely novel compared to other lineages, in both cases providing important insights into the history of eukaryotic evolution.

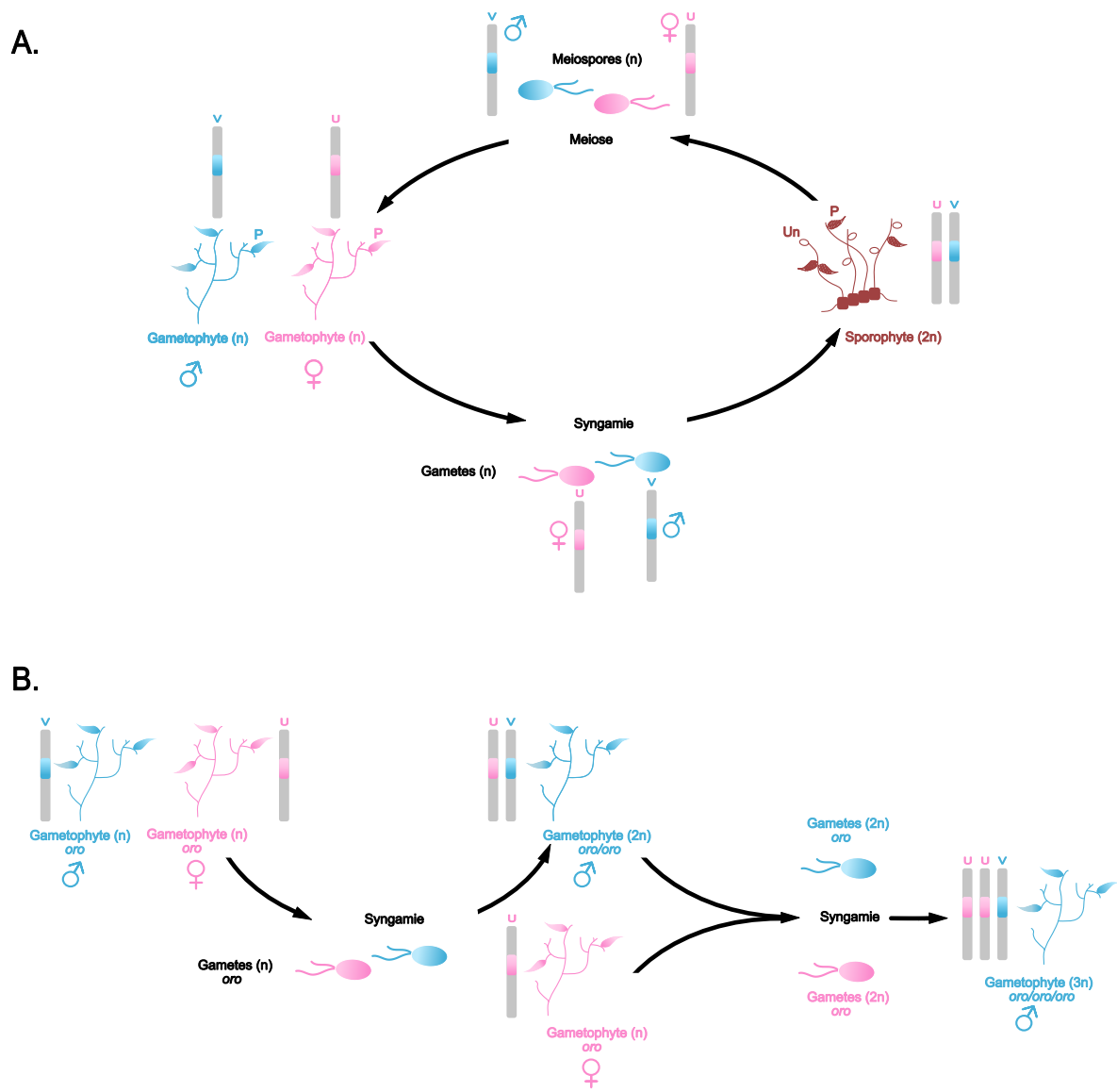


Figure 8. **The life cycle of the brown alga *Ectocarpus*.** (A) The haploid-diploid life cycle of the brown alga *Ectocarpus* sp. Un: Unilocular sporangia, P (brown): Plurilocular sporangia on the sporophyte and P (blue and pink): plurilocular gametangia in the gametophyte. (B) Life cycle of the *Ectocarpus* sp. mutant *oro*. A mutation at the *ORO* locus strongly affects the gametophyte/sporophyte transition. A cross between male and female gametophytes that are both mutant for the *oro* gene gives a diploid gametophyte (homozygous for the *oro* mutation) instead of a diploid sporophyte. This diploid gametophyte, although having both U and V chromosomes, is phenotypically male. A cross between this male diploid *oro/oro* mutant gametophyte and an *oro* mutant female gives rise to a triploid gametophyte that is again male. This shows that the V chromosome induces maleness, regardless of the presence of U chromosomes (Ahmed et al.).



Brown algae are multicellular, photosynthetic organisms found almost exclusively in the marine environment. Most of the species diversity is found in cold water regions. Brown algae are found mainly in intertidal areas, a particularly stressful environment (with significant abiotic variation), resulting in many interesting adaptations and surprising diversity, including in terms of sex determination mechanisms and life cycles.

*Ectocarpus* is a genetic and genomic model for the brown algae, which has received particular attention since the publication of a complete male genome sequence (Cock et al., 2010). *Ectocarpus* is a small brown filamentous alga that can reach 30 cm in nature, but can become fertile in the laboratory at 1-3 cm in length (Charrier et al., 2008). This alga has UV sex chromosomes (Ahmed et al., 2014) and has been used to study UV sexual systems along with several other species such as *Ulva partita* (Yamazaki et al., 2017), *Chlamydomonas reinhardtii* (De Hoff et al., 2013), *Gonium pectoral* (Hamaji et al., 2016), *Volvox* (Ferris et al., 2010) and *Marchantia polymorpha* (Bowman et al., 2017).

## 7.2 The life cycle of *Ectocarpus*

*Ectocarpus* has a haploid-diploid life cycle that involves alternation between two independent heteromorphic multicellular generations, the gametophyte and the sporophyte (Figure 8A). The diploid phase is characterized by the development of an asexual sporophyte from a zygote. The sporophyte is composed of basal (prostrate) filaments that are attached to the substrate and upright filaments that develop from the basal filaments. Sporophytes bear two types of spore-containing structure, plurilocular and unilocular sporangia, which develop on the upright filaments. Plurilocular sporangia produce spores through mitosis (mito-spores) that develop into new haploid sporophytes. Unilocular sporangia, on the other hand, produce spores via meiosis (meio-spores). The meio-spores develop into male or female haploid gametophytes depending on the inherited sex chromosome (V or U, respectively) (Ahmed et al., 2014).

Gametophytes produce only one type of structure at maturity: plurilocular gametangia. It is in these structures that flagellated gametes are produced. The gametes are released into the environment and fuse with a gamete of the opposite sex (fertilisation) to produce diploid zygotes. In addition, gametes that do not find partners are able to develop into haploid sporophytes via a parthenogenetic mechanism (partheno-sporophytes) and these partheno-sporophytes are morphologically and functionally identical to diploid sporophytes (Peters et al., 2008).

Studies have been conducted on the different mechanisms that govern life cycle progression in *Ectocarpus*. Two mutants that affect the transition from the gametophyte to the sporophyte have been identified. These two mutants have been named *ouroboros* (*oro*) and *samsara* (*sam*). Both of these mutant exhibit conversion of the sporophyte into a functional gametophyte (Coelho et al., 2011) (Figure 8B). Parthenotes derived from either *oro* or *sam* gametes develop as parthenogametophytes instead of partheno-sporophytes. Both *ORO* and *SAM* are predicted to encode transcription activator-like effectors homeodomain (TALE HD) transcription factors. The TALE HD transcription factor gene family is already known in other species, such as *Chlamydomonas* for example, for its involvement in life cycle progression (Lee et al., 2008).

Genetic analysis of the *oro* mutant has provided essential information about sex-determining factors in the *Ectocarpus*. When two *oro* mutant strains (male and female) are crossed, a diploid gametophyte develops instead of a diploid sporophyte. The diploid gametophyte carries a pair of UV sex chromosomes and produces diploid UV gametes that are capable of fusing with female gametes. This suggests that the diploid gametophyte is male despite the presence of the U chromosome. This experiment can be further extended, since fertilization between a diploid *oro/oro* UV gamete and a female *oro* U gamete, results in a triploid UUV male gametophyte (Ahmed et al., 2014) (Figure 8B). Thus, we find a pattern identical to the Klinefelter syndrome in humans (XXY), i.e. regardless of the number of X or U chromosomes, the presence of a Y or V male chromosome is sufficient and necessary for the determination of the male sex. If we draw a parallel with the research carried out in humans, we can deduce that there is a dominant factor on the V chromosome that allows male determination.

### 7.3 Sex determination in the brown algae

In a recent study, (Lipinska et al., 2017) conducted a systematic search for orthologues of the SDR genes of *Ectocarpus* sp. in the genomes of nine brown algal species which are at different evolutionary distances from the reference species *Ectocarpus* sp. The only gene that was consistently male-limited in all the species studied was the HMG domain protein-encoding gene *Ec-13\_001750*.

HMG transcriptional factors are known to be the male-determining factor in many species, such as Sry in humans for example. The transcript of *Ec-13\_001750* was more than 10-fold more abundant in mature gametophytes (i.e. during production of male gametes) than at other stages of development (Ahmed et al., 2014). This gene therefore appears to be a strong candidate for

the male sex-determining gene in *Ectocarpus*, and possibly in several other brown algal species. However, sex-reversed mutants or reverse genetic tools will be needed to confirm that this gene is the *bona fide* master sex determination gene in *Ectocarpus*.

#### 7.4 Sex differentiation in brown algae

The candidate sex-determining, SDR-located HMG gene in *Ectocarpus*, *Ec-13\_001750*, is predicted to direct implementation of the male differentiation pathway. This pathway consists of a cascade of genes that lead to the expression of sexual dimorphism between male and female individuals. This dimorphism consists of morphological, physiological and behavioural differences. Most of this phenotypic difference is mediated by differential gene expression between the two sexes (Ellegren and Parsch, 2007) and this differential gene expression can affect a significant proportion of the genome, up to 50% in *Drosophila* for example (Assis et al., 2012). Given that only a small portion of the genes involved in sexual differentiation are found in the SDR (Ellegren and Parsch, 2007), it is important not only to characterize brown algal SDRs but also to compare gene expression between the two sexes to fully understand the genetic basis of sexual dimorphism in this group.

Fewer than 12% of the genes in the *Ectocarpus* genome (i.e., 1,947 genes) exhibit sex-biased gene expression (Ahmed et al., 2014). This low percentage is correlated with the low level of sexual dimorphism observed in *Ectocarpus*. *Ectocarpus* is quasi-isogamous (male and female gametes are very similar in size) and it is difficult to distinguish a male gametophyte from a female gametophyte based on morphology. One of the possible reasons for this low dimorphism may be a low level of sexual selection due to the way gametes are dispersed in the environment. Theoretically, low levels of dimorphism could be the result of dioecy (or dioicy in the case of brown algae) only having evolved recently in a particular lineage so that sexual selection has not occurred long enough to establish marked sexual dimorphisms (Barrett and Hough, 2013). However, dioicy is thought to have evolved early in the brown algal lineage and it is therefore unlikely that the latter hypothesis explains the apparent low level of sexual dimorphism in this lineage.

## Objectives

The goals of my PhD were the following (Figure 9):

- 1) To gain knowledge about the pathways controlling sex determination in *Ectocarpus* by characterizing the candidate master male sex determining gene *Ec-13\_001750* (chapter 2).
- 2) To understand the molecular mechanisms regulating sexual differentiation, by examining the role of chromatin modifications in regulating gene expression in males versus females using *Ectocarpus* as a model system (chapter 3) and by studying variant brown algal strains that are partially sex-reversed (chapter 4).

I have also been involved in a review paper about UV chromosome and haploid sexual systems that can be found in the Annex.

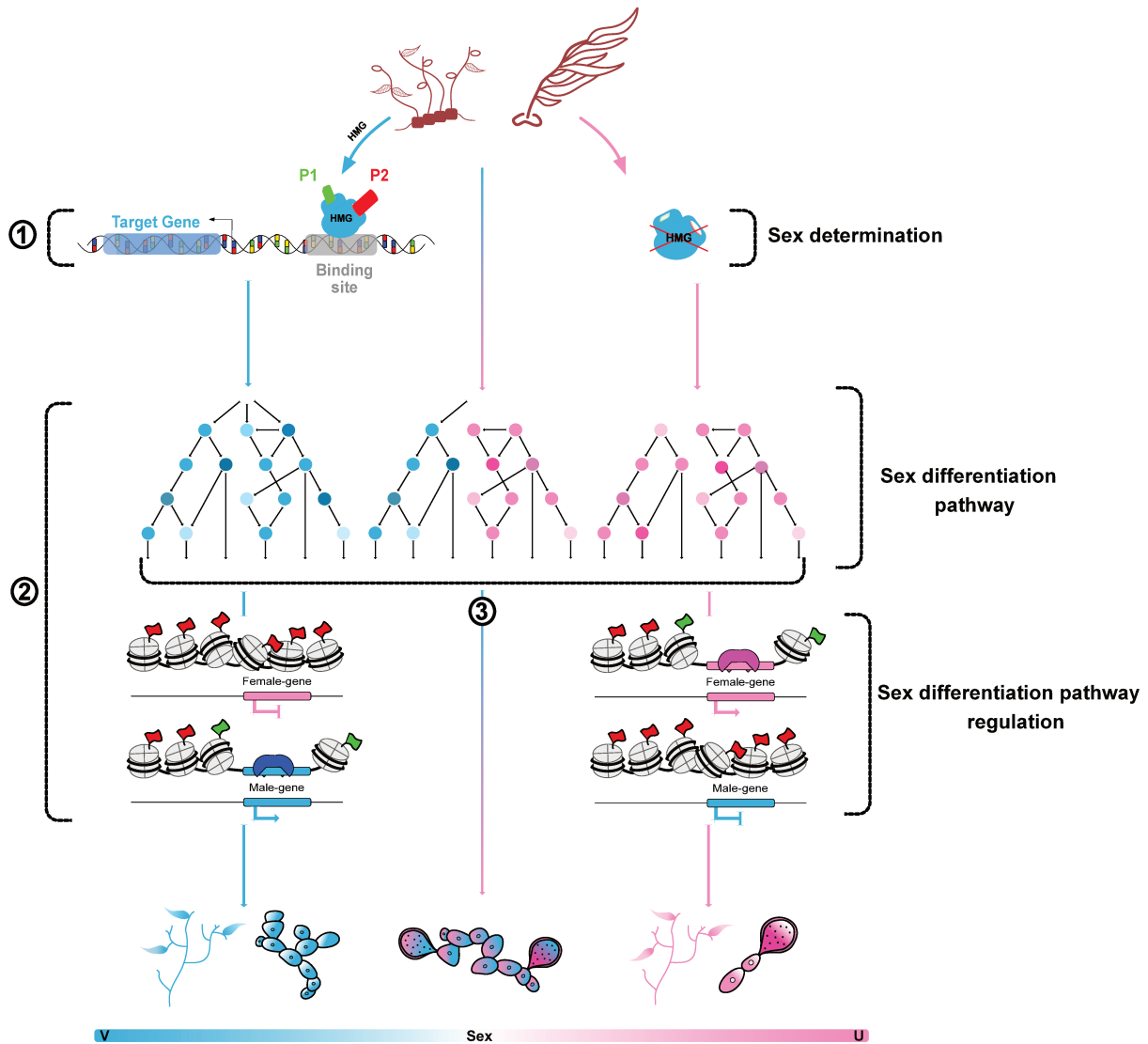


Figure 9. **Overview of the thesis objectives**

(1) Characterization of the V-chromosome linked protein HMG-sex in the brown alga *Ectocarpus* sp., a candidate sex determining factor. The objective of this study was to identify the DNA binding sites of the HMG protein, as well as partner proteins that interact with it (P1/P2).

(2) Identification of epigenetic modifications associated with differential expression of genes involved in sex differentiation in the brown alga *Ectocarpus* sp.. This part of my PhD focused on post-translational histone modifications and aimed to highlight differences in chromatin landscapes between males and females. We will focus here on the path of sexual differentiation, induced by the presence of the putative master regulator in males and by its absence in females.

(3) Study of a variant line of the brown alga *Macrocystis pyrifera* that is partially sex-reversed. This study focused on the sexual differentiation pathways in male, female and in the feminized male strain, including a comparison of sex-biased gene expression. The aim was to understand the genetic mechanisms that govern sexual differentiation.



# Chapter 2

## Mechanisms of sex determination: functional analysis of a candidate sex-determination factor in *Ectocarpus*

This chapter presents experiments that were carried out to determine the DNA binding specificity and to identify possible protein partners of the HMG domain protein Ec-13\_001750, a V-specific, putative transcription factor that is likely to be the master male sex determining factor in *Ectocarpus*. The DAP-seq method was used with the aim of identifying *in vitro* binding sites of the HMG protein and we implemented an experiment using the yeast two-hybrid (Y2H) system to search for interacting proteins.

Reverse genetics tools (CRISPR-cas9) are being developed for *Ectocarpus* and this methodology could be used in the future to validate the function of the HMG as a master factor in sex determination. However, currently the methodology is not yet fully functional. Therefore, we decided to use the DAP-seq, ChIP-seq and Y2H methods to investigate the putative role of Ec-13\_001750 in sex determination.

## 1. Introduction

*Ec-13\_001750* is a HMG transcription factor gene carried by the male V chromosome, located within the sex-determining region (SDR) of that chromosome. It represents an attractive candidate for the male determining factor in the brown alga *Ectocarpus* sp. Several lines of evidence point towards a role for this gene as the master sex-determining factor. First, the HMG domain gene has been shown to be conserved across nine different male SDRs in different brown algae (Lipinska et al., 2017) and recent results in our group extend this to 12 species of brown algae, spanning 160 MY of brown algal evolution. Second, the expression pattern of this gene is consistent with a role in sex determination in several brown algae species, as it exhibits a significant increase in transcript abundance during gametogenesis (Ahmed et al., 2014; Lipinska et al., 2017). Finally, HMG domain transcriptional factors have been shown to be involved in gender determination across a range of species, including humans and the yeast *Saccharomyces cerevisiae* (Kurjan, 1985; Sinclair et al., 1990)14/11/2019 14:36:00.

Transcription factors (TFs) are proteins that bind specifically to DNA sequences and enhance or repress target gene expression (Spitz and Furlong, 2012), being the main players in transcriptional regulation in eukaryotes. TFs respond to a broad range of stimuli to coordinate many different processes including cell cycle progression, cell differentiation and development (Caramori et al., 2019). Recognition of target sites by TFs can involve binding to specific nucleotide sequences but binding can also be influenced by 3D structure and flexibility of both the TFs and their binding sites, by the presence of cofactors, by chromatin accessibility and by nucleosome occupancy (for example for pioneer TFs that bind to nucleosomal DNA) and by DNA methylation (Slattery et al., 2014). The regulatory genomic sequences targeted by TFs are found mainly in non-coding intergenic or intronic DNA, with few exceptions (Stergachis et al., 2013). Although pioneering studies in *S. cerevisiae* provided a foundation for the understanding of TF biology, the noncoding regulatory landscape in this organism is easier to analyse than for metazoan eukaryotes. The regulatory genomic sequences targeted by TFs are found mainly in non-coding intergenic or intronic DNA, with few exceptions (Stergachis et al., 2013). Although pioneering studies in *S. cerevisiae* provided a foundation for the understanding of TF biology, the noncoding regulatory landscape in this organism is easier to analyse than for metazoan eukaryotes. Indeed, for *S. cerevisiae*, most of the regulatory DNA sequences of a given gene are located a few hundred base pairs from its transcription start site (TSS) (Lin et al., 2010).



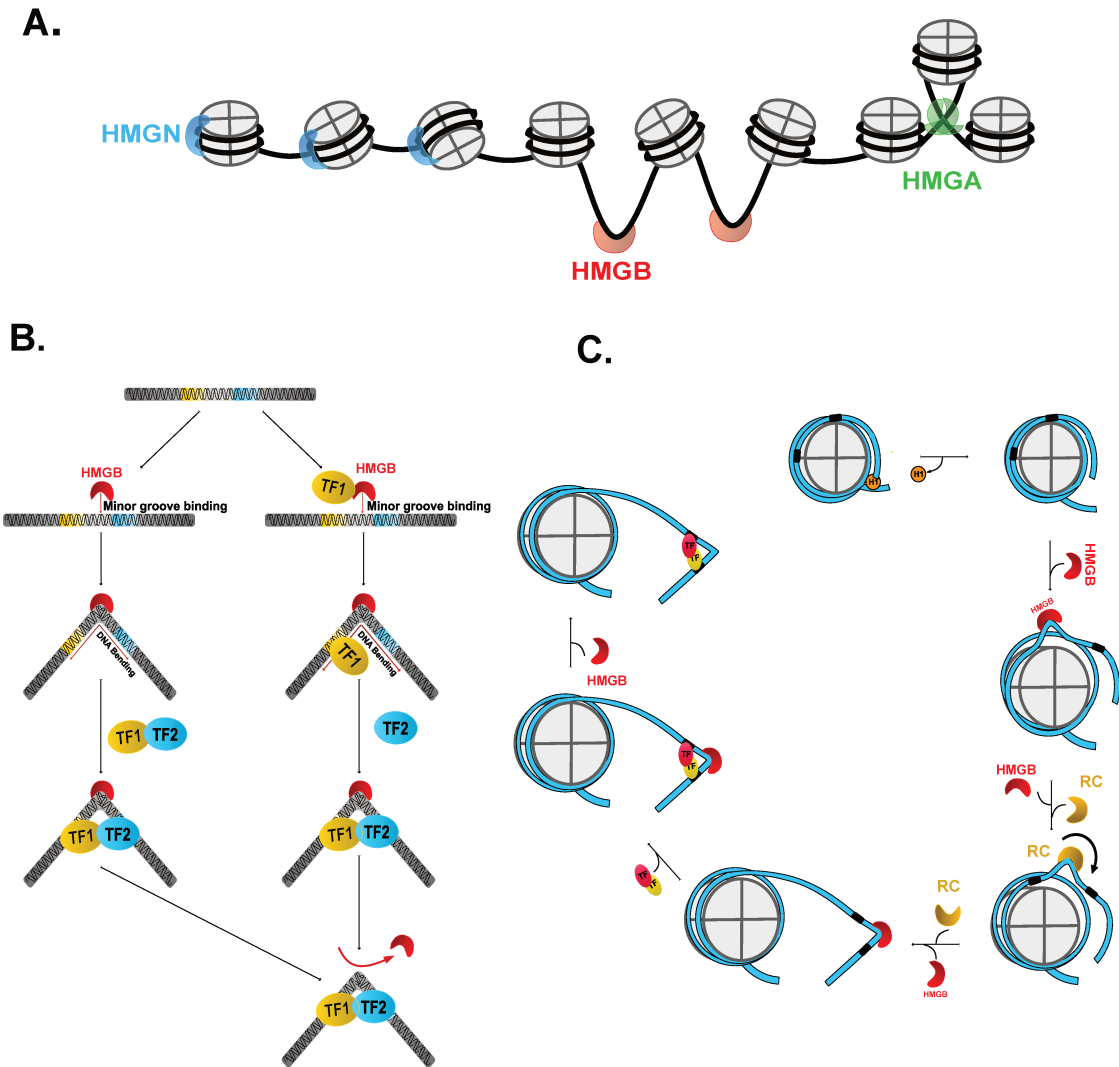


Figure 1. Mechanisms of **HMG binding to DNA** (A) Mode of action of the three HMG families: HMGN in blue, HMGB in red and HMGA in green. (B) A model depicting the putative role of HMGB proteins as architectural factors in transcription. (Right) HMGB interacts with a transcription factor (TF1) and directs it to its specific-binding DNA site which is pre-bent by HMGB. A ternary complex TF1–HMGB–DNA is formed. Following the binding of the HMG protein, the DNA bends. Another transcription factor (TF2) is attracted to the complex TF1–DNA–HMGB, followed by the release of HMGB from the complex. (Left) Alternatively, the specific DNA sequence is bent by HMGB resulting in enhanced binding of TFs without direct interactions with HMGB. (C) A model for HMGBs functioning as factors promoting nucleosome mobility and accessibility to specific DNA sites. Histone H1, like HMGB, binds at the DNA entry points into the nucleosome (histone octamer and DNA are in grey and blue, respectively). Binding of H1 and HMGB is mutually exclusive, and HMGB can displace histone H1 from the nucleosome (step 1), most likely by interaction of the negatively charged C-tail of HMGB with positively charged flanking sequences of H1. HMGB binding at the edge of the nucleosome results in DNA bending and formation of a tight “DNA loop” to which HMGB is bound (step 2). The DNA loop represents an anchoring site for the remodeling complex (RC) (step 3), the binding of which is enhanced by HMGB. HMGB dissociates from the loop upon binding of RC, followed by the directional propagation of the looped DNA (bound to RC) around the histone octamer (step 3), rendering the sequence-specific site (black box) accessible for transcription factor (TF) binding. Release of the remodeling complex from the nucleosome and positioning of the sequence-specific site outside the histone octamer (step 4) is accompanied by pre-bending of the DNA sequence by HMGB (step 4), followed by TF binding (step 5) and HMGB dissociation (step 6). The model presented is based on (Štros et al., 2009).

In metazoans, on the other hand, regulatory sequences are often located tens of kilobases (kb) or even megabases (Mb) from the TSS of the target gene. These distal elements may be upstream or downstream of the target gene (El-Kasti et al., 2012; Li et al., 2013). Thus, although the basic properties of TFs are preserved (recognition of a binding site and induction/repression of the transcription of a target gene), the mechanisms of action of TFs are diverse, ranging from simple mechanisms such as in yeast to more complex ones, such as in metazoans. In general, the binding of a TF to DNA is not sufficient to trigger gene expression. TFs generally need to recruit other proteins or complexes such as chromatin remodelling factors and histone modifying enzymes (methyltransferases and acetyltransferases), which modify the local chromatin environment, resulting in nucleosome compaction and post-translational histone modifications in the neighbourhood of transcription start sites (Bai and Morozov, 2010; Barrett and Wood, 2008; Kadonaga, 1998; Voss and Hager, 2014). TFs can also be recruited, or be recruited by other TFs, to stabilize accessibility to the DNA and to improve transcription activity.

The HMG superfamily is composed of three subgroups: HMGA, HMGB and HMGN (Figure 1A). The HMG box is the functional motif of the HMGB proteins (HMG-box protein), the AT hook is the functional motif of the HMGA family, and the nucleosomal binding domain (NBD) is the functional motif of the HMGN family (Postnikov and Bustin, 2016). In recent years, HMG box domains have been found in various DNA-binding proteins including TFs and subunits of chromatin-remodelling complexes. The HMG-box can bind non-B-type DNA structures (bent, kinked and unwound) with high affinity, and also distort DNA by bending/looping and unwinding (Štros et al., 2007) (Figure 1B). HMG-box proteins are usually classified into two major groups distinguished by their abundance, function and DNA specificity. The first group consists of proteins containing two HMG-boxes. In general, two or more HMG-boxes are mostly found in abundant HMG-box proteins with little or no DNA sequence specificity. The second group of HMG-box proteins is highly diverse and consists of much less abundant proteins having mostly a single HMG-box. The single HMG-box proteins recognize specific DNA sequences and generally have a role as a transcription factor (HMG TFs) (Štros et al., 2009).

In contrast to most TFs, which bind the major groove of DNA, HMG TFs bind the minor groove inducing a 60°-70° bend in the double helix. DNA binding modulates the 3D architecture of the DNA, leading to speculation that TF proteins containing a HMG-box act remotely by

creating a bending point (Figure 1B). DNA bending could bring distant sites into contact with each other to facilitate the interaction of transcription factors (Bianchi and Beltrame, 1998). Proteins with HMG boxes are relatively abundant, the HMG-box is quite conserved between species and has a low specificity of DNA-binding target sequence (Weir et al., 1993). However, regions adjacent to a HMG-box show a high degree of sequence divergence in amino acid sequence, even between closely related species (Sessa and Bianchi, 2007). Thus, in humans, there are about forty proteins with the HMG-box of the Sox family. These HMG-box proteins are involved in developmental processes, such as the genes of the Sox family or even the *Sry* gene involved in male sex determination in mammals (Schepers et al., 2002).

In mammals, either ovaries or testes are formed from the bipotential genital ridges during embryogenesis. Activity of the Y-linked HMG domain transcription factor *Sry* induces testis differentiation in males (Smith et al., 2009). In addition to directly bending DNA and altering the 3D conformation of chromatin, *Sry* can act as an epigenetic regulator by interacting with chromatin-modifying complexes. In fetal gonads, *Sry* interacts with the Krüppel-associated box only (KRAB-O) protein and its obligatory co-repressor Krab-associating protein 1 (KAP1) (Oh and Lau, 2006; Peng et al., 2009). KAP1 then recruits heterochromatin protein 1 (Hp1), HDACs and the SETDB1 methyltransferase, which function as gene silencers by creating a repressive chromatin environment. Therefore, *Sry* may have a dual function in early sex determination by activating male-determining genes by directly binding to regulatory elements (such as binding to the testis-specific enhancer TESCO upstream of *Sox9*; Sekido and Lovell-Badge, 2008), and by repressing the female pathway through recruitment of KRAB-O/KAP1 chromatin-mediated repression machinery.

In *Ectocarpus*, 14 genes are predicted to encode HMG-box proteins, including the gene *Ec-13\_001750*, which encodes the male sex-determination candidate. Unlike *Sry*, *Ec-13\_001750* has two HMG domains predicted (Figure 2). There are two sub-families of HMG-box protein: one comprising proteins with a single sequence-specific HMG-box, the other encompassing relatively non-sequence-specific DNA-binding proteins with several HMG boxes (Laudet et al., 1993). Although the two types of HMG-box protein do not recognize DNA in the same way, they perform essentially the same action: they increase the flexibility and stability of the DNA. On the other hand, both effects are obtained at much lower protein concentrations for HMG-box proteins with two HMG boxes (McCauley et al., 2007).

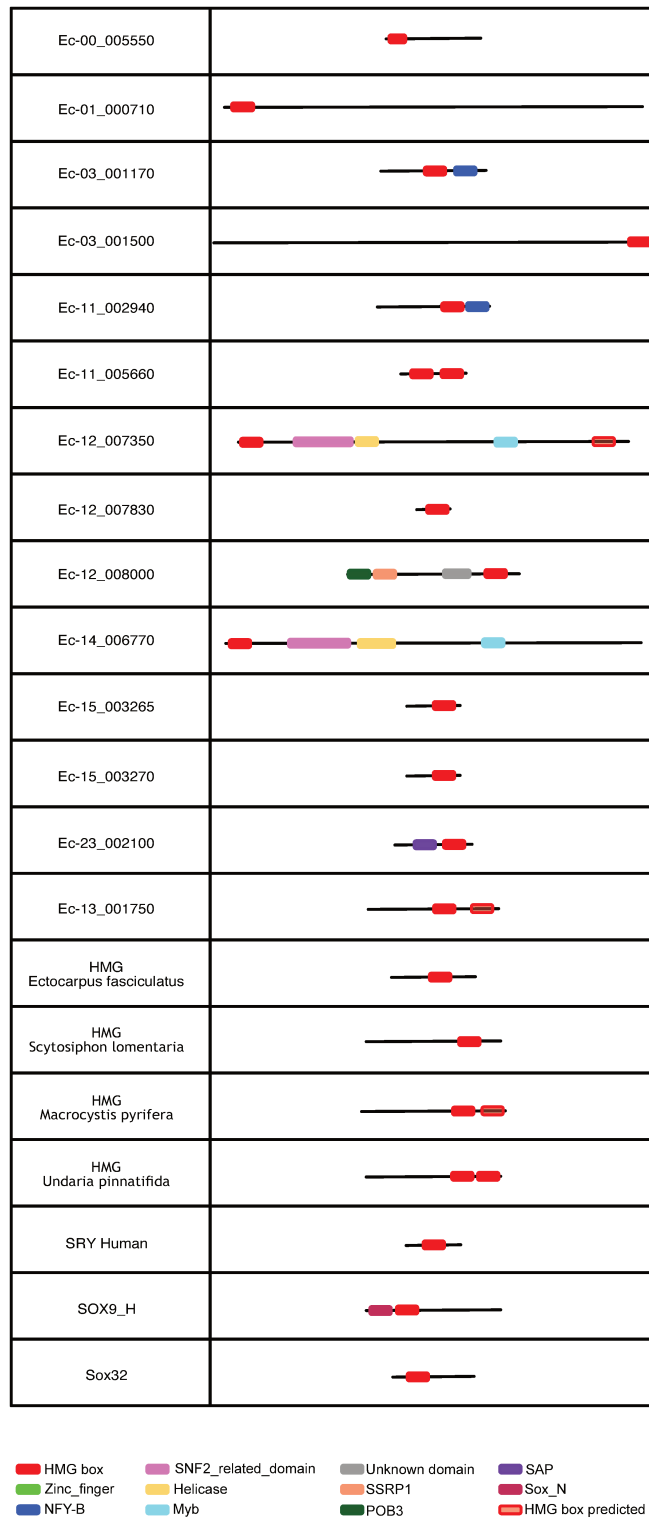


Figure 2. **Domain structure of *Ectocarpus* sp. HMG domain proteins.** Schematic representation on scale of the different domains present on the autosomal HMGs of *Ectocarpus* (Ec-XX\_XXXXXX), HMG-sex of *Ectocarpus* (Ec-13\_001750), the V-chromosome-linked HMGs from four other brown algae, the master regulator of male determination in humans (*Sry*), the master regulator of male determination in zebra fish (Sox32) and the human protein Sox9, involved in the male differentiation pathway.

This chapter will focus on two objectives: 1) identification of proteins that interact with the *Ectocarpus* HMG protein Ec-13\_001750 and may therefore act as cofactors to regulate gene expression during sex determination and differentiation (section 3.1 and 3.2, a search for HMG interacting proteins was carried out using the Yeast Two-Hybrid (Y2H) system and a cDNA library representing the gametophyte transcriptome) and 2) identification of the specific DNA binding sequence of Ec-13\_001750 and the target sites of this protein in the genome (section 3.3 and 3.4). Two different approaches were attempted to detect the genomic binding sites of this protein. The first was an *in vitro* approach called DNA Affinity Purification sequencing (DAP-seq, Bartlett et al., 2017; O'Malley et al., 2016). DAP-seq is not the first method that has been used to effectively identify TF binding sites in genomic DNA. PB-seq (Guertin et al., 2012), DIP-chip (Liu et al., 2005) and DAP-ChIP (Rajeev et al., 2014) have already been used successfully in other species. The second approach used was an *in vivo* methodology, chromatin immunoprecipitation with nucleotide resolution through exonuclease treatment, unique barcode and single ligation (ChIP-nexus, He et al., 2015).

## 2. Material and methods

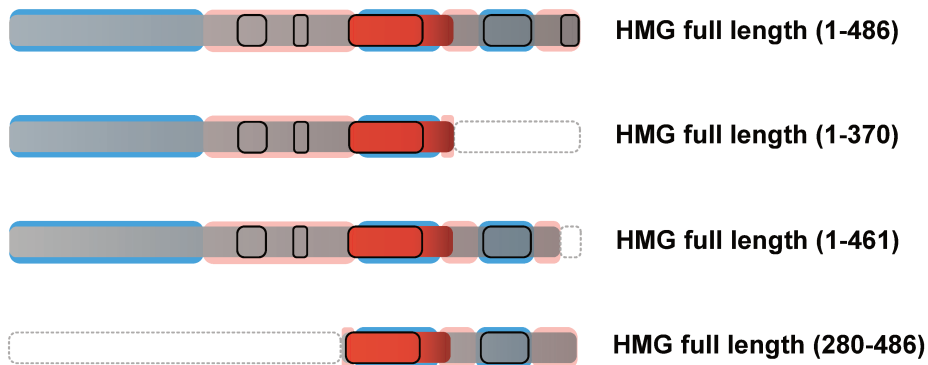
### 2.1 Plasmid construction

All constructs were generated using the single-step directional cloning In-Fusion Cloning system (Takara Bio/Clontech). First, the DNA fragment to be inserted was amplified by PCR using the high fidelity CloneAmp HiFi PCR premix and oligonucleotides containing, at the 5' end, a minimum of 15 bases pairs which were homologous to the plasmid region near the insertion site (multiple cloning site, MCS). Amplification products were purified on an agarose gel and extracted using the NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel). Then, *in vitro* homology-based recombination was used to fuse the target linearized vector and the DNA insert. Finally, recombination products were transformed into *Escherichia coli* Stellar Cells (Clontech). In-frame ligation of each construct was confirmed by a Sanger sequencing.

#### Plasmid for DAP-seq and Production of anti-HMG antibodies

Different fragments of the HMG domain protein coding gene (corresponding to amino acids 0-486, 192-486, 293-486, 302-486, 192-467 and 1-467) were inserted into *EcoRI/BamHI*-digested pGex 4T-1 containing the glutathione S-transferase (GST) gene. The construct is under the control of the strong and constitutive hybrid pTAC promoter and a *lacI/lacO* induction system. (Figure 3B)

A.



B.

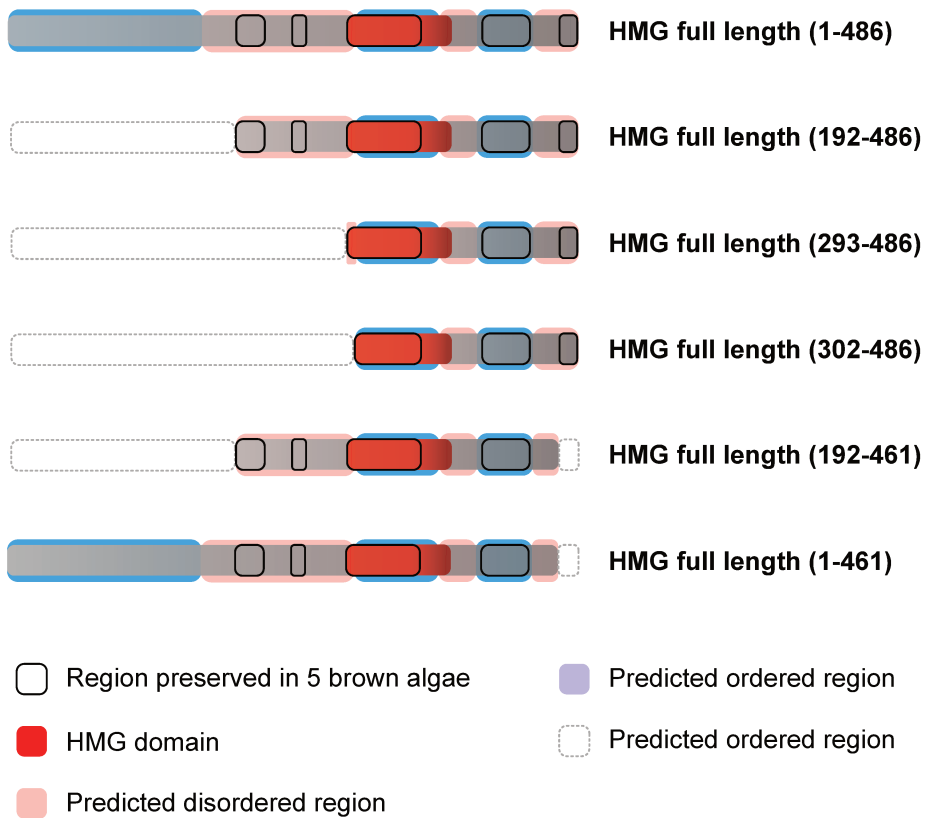


Figure 3. **Schematic representation of HMG constructs used for in this study** (A) Schematic representation of the different constructions made to solve the problem of self-activation of the HMG protein in the yeast two-hybrid assay. The constructions were cloned in plasmid pGBKT7. (B) Schematic representation of the different constructions generated to produce the HMG protein. The constructions were cloned in plasmid pGex 4T-1. See methods for details.

## Plasmids for Yeast Two-hybrid bait constructs

The full-length HMG-domain protein coding gene (0-486) coding sequence and various sub-fragments of this sequences (1-370, 280-486 and 1-462) were inserted into an *EcoRI/BamHI* digested pGBKT7 plasmid (Takara Bio/Clontech) (Figure 3A). This plasmid encodes the DNA-binding domain of the *Saccharomyces cerevisiae* GAL4 transcription factor following by a 34 aa linker containing a Myc tag which allows production of a GAL4-DBD – Myc – Protein product (where DBD is the DNA binding domain).

## 2.2 Yeast Two-Hybrid cDNA library

Total RNA was extracted from 200 mg of 3-week-old gametophyte of the strain Ec32 using the RNeasy Plant Mini kit (Qiagen). cDNA library and plasmid integration were carried out using 500 ng of total RNA with the high throughput Make Your Own “Mate & Plate” Library System (Takara Bio/Clontech). cDNA first-strand synthesis was performed using the SMART MMLV Reverse Transcriptase and the CDS III oligonucleotide containing a poly-dT stretch. SMART III oligonucleotide was added to the reverse transcription following manufacturer recommendations. Next, single-strand cDNAs were amplified by long-distance PCR using the high fidelity Advantage2 Polymerase and primers which bind the CDS III and SMART III regions added during the cDNA first-strand synthesis. Amplified double-strand cDNAs were purified on CHROMA SPIN columns and the size range of the cDNA was checked on an agarose gel.

Purified cDNAs containing the flanking SMART III and CDS III sequences and the linearized pGADT7-Rec plasmid were co-transformed into *S. cerevisiae* strain Y187 using an optimized lithium acetate-mediated protocol available in the Yeastmaker Yeast Transformation System (Takara Bio/Clontech). This co-transformation protocol uses the homologous recombination machinery of the yeast cell to generate recombinant clones between the target plasmid and the cDNA library *via* SMARTIII and CDS III sequences. Transformants were grown on selective SD/-Leu agar plates and incubated at 30°C for 4 days. It is necessary to measure the recombination rate i.e. if the transformants contained at least one plasmid construct encoding recombinants between the transcriptional activation domain (TAD) of GAL4 and a random coding region of the cDNA almost full length. Then, transformants were harvested using YPDA + 25 % glycerol and frozen at -80°C.

## 2.3 Yeast Two-Hybrid Assay

### Strain genotypes

The GAL4 system uses the functional reconstitution of the GAL4 protein (interaction between the DBD and TAD domains of the GAL4 protein) to activate reporter genes which complement auxotrophic phenotypes or produce reporter proteins. The Y2H Gold and Y187 strains are unable to grow in minimal medium that lacks specific amino acids or nucleic acid precursors such as leucine (*leu2*), tryptophane (*trp1*), histidine (*his3*) and adenine (*ade2*). Moreover, both strains carry deletions of the *Gal4* and *Gal80* (negative regulator of Gal4) genes. In addition, the Y2H Gold strain carries three reporter constructs GAL2 – Ade2, LYS2:GAL1 – His3 and MEL1 – Aur1-C. The three reporter constructs contain an upstream activating sequence that is recognised by the GAL4-DBD. The coding regions of these reporter genes correspond to *ADE2* and *HIS3* which complement the *ade2* and *his3* genotypes and the *AURI-C* gene, which confers resistance to Aureobasidin A. As the Y2H Gold and Y187 strains are of opposite mating type (MATa and MAT $\alpha$ , respectively), genetic crosses between a Y2H Gold clone and a Y187 clone are possible (Figure 3)

### Selection of bait and prey combinations in yeast

Bait constructs (pGBKT7 plasmids) were individually transformed into the *S. cerevisiae* Y2H Gold strain (Matchmaker Gold Yeast Two-Hybrid System from Takara Bio/Clontech) and grown on selective SD/-Trp agar plates with autotrophy being conferred by the *TRP3* gene (on the pGBKT7 plasmids). Prey constructs (pGADT7-Rec plasmids) were transformed into the Y187 strain and grown on selective SD/-Leu plates with autotrophy being conferred by the *LEU2* gene (on the pGADT7-Rec plasmids). Again, transformations were performed using the Yeastmaker Yeast Transformation System.

### Auto-activation bait

It is necessary to verify that the HMG bait protein does not autonomously activate the reporter genes of the Y2H yeast in the absence of a GAL4 activation domain. In addition, it is also necessary to verify that the HMG protein does not interact directly with the GAL4 activation dominance. The auto-activation tests thus prevent false positives. First, Y2H cells transformed with the plasmid pGBKT7-HMG were spread on a selective medium SD/-Trp/-Ade/-His/+Aureobasidin A agar plate and incubated for 3 days. In parallel, Y2H transformed with the plasmid pGBKT7-HMG was mated with Y187 transformed with the empty plasmid pGADT7-Rec. Finally, yeasts were spread on a selective medium SD/-Trp/-Leu/-Ade/-



His/+Aureobasidin A agar plate and incubated for 3 days. In the absence of auto-activation, diploid yeast are not expected to grow.

#### Mating and screening for prey-bait interactions

An overnight culture of the bait construct was incubated in SD/-Trp liquid medium until the optical density at 600 nm reached 0.8. Then, cells were pelleted and resuspended in 4 ml SD/-Trp. One millilitre of the prey library (from a -80°C stock) was combined with the bait in 45 ml 2x YPDA. This mating culture was incubated for 24 hours at 30°C on an orbital shaker. Yeast were spread on selective SD/-Trp/-Leu/+Aureobasidin A agar plates at 30°C and incubated for 3 days. This medium reduces the number of diploid colonies by keeping only those with were resistant to Aureobasidin A. Screening for interaction between the prey and bait proteins was carried out by transferring the diploid yeast colonies to a high stringency selective SD/-Trp/-Leu/-Ade/-His/+Aureobasidin A.

#### Extraction and cloning of prey plasmids

Prey plasmids were extracted from positive diploid yeast colonies with the Easy Yeast Plasmid Isolation kit (Takara Bio/Clontech). These plasmids were transformed into *E. coli* Stellar Cells strain and selected on LB agar plates with ampicillin. Plasmids were then extracted from *E. coli* clones and sequenced. Sequences were blasted against the *Ectocarpus* genome and transcript database. GAL4-TAD – protein fusion constructs were sequenced to verify that they were in-frame and did not contain any mutations.

### 2.4 HMG and Yeast Two-Hybrid

To evaluate the influence of *Ectocarpus* HMG on the mating-type of MATa yeasts (Y2H), 11 different crosses were made (Table 1). The cloning and transformation techniques were the same as those described above. An *Ectocarpus* autosomal HMG was chosen for its similarity to Ec-13\_001750, in terms of size and protein domains. Each mating was spread on a selective SD/-Leu/-Trp medium, to select only diploid yeasts. The yeast colonies that develop were then counted.

Mating Y187 x Y2H	Number of colonies
pGAD $\emptyset$ x pGBKHMGsex	12,10 <sup>6</sup>
pGBK HMGsex x pGAD $\emptyset$	13,5.10 <sup>6</sup>
pGAD $\emptyset$ x pGBK $\emptyset$	>19.10 <sup>6</sup>
pGAD $\emptyset$ x pGBKHMGauto	>19.10 <sup>6</sup>
pGBK HMGauto x pGAD $\emptyset$	15,5.10 <sup>6</sup>
pGADlibrary Ga x pGBK HMGsex	25000
pGADlibrary Ga x pGBK $\emptyset$	31000
pGADlibrary Ga x pGBK HMGauto	6000
pGADlibrary Sp x pGBK HMGsex	287500
pGADlibrary Sp x pGBK $\emptyset$	468750
pGADlibrary SP x pGBK HMGauto	393750

Table 1. Table summarizing the crosses carried out to evaluate the impact of the *Ectocarpus* HMG-sex protein on mating efficiency in yeast. HMGsex = Ec-13\_001750; HMGauto = HMG autosomal from *Ectocarpus*;  $\emptyset$  = empty plasmid; Ga = gametophyte; Sp = sporophyte.

## 2.5 HMG protein production

pGex4T-1 plasmids encoding the different HMG protein fusions were transformed into *E. coli* Rosetta-gami™(DE3) strain for expression. Bacteria were grown in a lysogeny broth medium (LB) to a DO 600 of 0.5. Then, 500mM NaCl and 2mM Betaine (final concentration) were added to the culture medium and bacteria growth continued at 47°C for 1 hour. The bacteria were then incubated at 20°C for 35 minutes. Protein expression was induced at 20°C for 20 hours with 0.2 mM IPTG. Bacterial cultures were pelleted for 20 minutes at 4200 rpm and 4°C, and resuspended in 10ml of Tris-EDTA pH9 (TE) for 800ml of culture. Anti-protease complete (Roche) was added to 1x final concentration. Mechanically lysis through a French press was applied to the bacterial culture and 100 µl of DNase was added to the lysed cells. The lysed cells were centrifuged at 14000 rpm for 90 min at 4°C. The supernatant was filtered through a 0.45 µm filter.

## 2.6 DAP-seq

DNA Affinity Purification (DAP) experiments were carried out following the protocol from (O'Malley et al., 2016) and (Bartlett et al., 2017) with some modifications proposed by D. Scornet in the Algal Genetics group. Genomic DNA from *Ectocarpus* strain Ec32 was fragmented to a target size of 200 bp using a Covaris M220 ultrasonicator. Fragmented DNA was purified and concentrated using AMPure XP magnetic beads (Agencourt). Ten micrograms of blunt-ended, fragmented DNA was obtained with the NEBNext End Repair Module (New England Biolabs) and purified with the NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel). Next, desoxyribo-adenine was added to the blunt ends with the NEBNext dA-Tailing Module (New England Biolabs) and the fragments were purified again with the NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel). Double-stranded adaptors were ligated to the dA-tailed, fragmented DNA using the NEBNext Quick Ligation Module (New England Biolabs), followed by purification with AMPure XP magnetic beads (Agencourt) and elution in 52 µl of 10 mM Tris HCl pH 8.

A GST-(HMG 1-467) construct was produced under the conditions described in the previous section 1.5., To obtain sufficiently concentrated HMG protein for the DAP-seq it was necessary to use a 10K centricon, until a volume of 2ml of lysate was obtained. Recombinant proteins were immobilized on paramagnetic particles (MagneGST Glutathione Particles - Promega) and washed three times with 250µl of MagneGST Binding/Wash Buffer. Ligated DNA was mixed with MagneGST magnetic beads and incubated for one hour at room temperature on a rotating

agitator. The beads were then washed four times with MagneGST Binding/Wash Buffer and boiled for 10 minutes at 98°C in a thermal cycler. Free DAP-DNA was amplified with the Illumina TruSeq Universal and Illumina TruSeq Index primers to allow multiplexing. Sequencing was carried out on an Illumina HiSeq4000 and single-end sequencing primers over 150 cycles.

Reads were trimmed using Cutadapt v1.8.3 and mapped onto the *Ectocarpus* genome v3.0 with Bowtie v1 using default parameters in paired-end mode. Signal depth was calculated with the Deeptools plotCoverage tool (Ramírez et al., 2016). Peak calling was carried out with the MACS2 (Zhang et al., 2008) callpeak module and a minimal FDR of 0.01 (-q 0.01). Duplicates were removed using samtools markdup in the package samtools (v1.9) (Li et al., 2009). Fasta sequences of bound regions were retrieved using the BedTools getfasta module. Motifs were found with the MEME program set with any number of repetitions (-mod anr), minimal motif size of 6 bp (-minw 6), maximal motif size of 8 bp (-maxw 8) for 5 motifs (-nmotifs 5). *De novo* motif discovery was also performed using the KMAC module of GEM peak caller software using either all peaks or peaks that were not localized in transposable elements (Guo et al., 2017, 2012).

## 2.7 Production of anti-HMG antibodies

Recombinant GST – HMG (residues 1-467) was produced according to the method used in the section 2.1.1. Then the protein was purified on a 5ml GSTtrap FF GE column (General Electric) using an Akta purifier. The lysate was injected and then washed with a 1X Phosphate-Buffered Saline (PBS). Finally, GST-HMG protein was eluted with an elution buffer (reduced glutathione 20mM). For elution to occur correctly, the column must be incubated for 10 mins in the elution buffer.

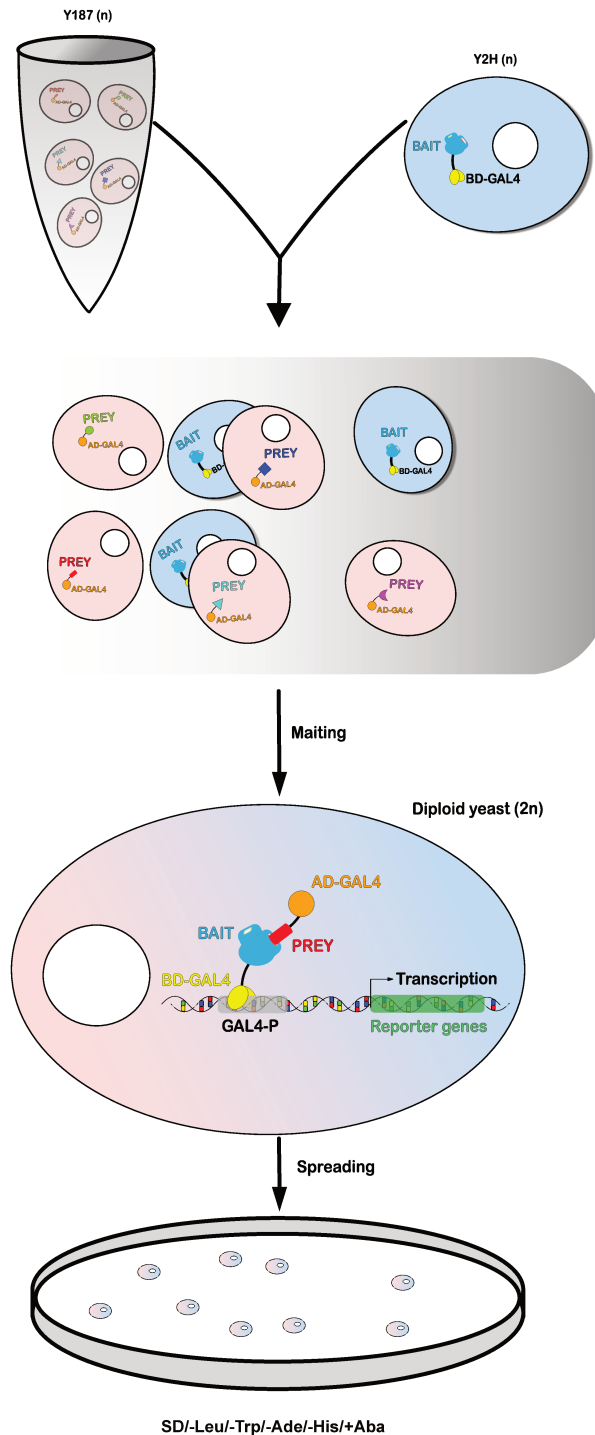


Figure 4. **Yeast two-hybrid library assay.** The Y2H strain expresses the bait protein of interest, fused with the DNA binding site of the GAL4 protein (BD-GAL4). Y187 strain yeasts, which constitute the yeast library, express prey proteins fused with the GAL4 protein activation site (AD-GAL4). Mating takes place between yeast Y2H and the yeast Y187 library, which are of different mating types, respectively a and alpha. After mating, the diploid yeasts obtained possess express both the bait and prey proteins. If there is an interaction between the bait protein and a prey protein, the GAL4 protein becomes functional. The protein-bait/prey-protein heterodimer binds to the GAL4 protein binding site (GAL4-P) upstream of reporter genes and activates expression. The reporter genes are responsible for autotrophy for the amino acid histidine and adenine, as well as for resistance to the antibiotic aureobasidine. Thus, in the presence of an interaction between bait and prey proteins, diploid yeasts are able to grow on a strict SD/-Leu/-Trp/-Ade/-His/+Aba medium.

Two rabbits were immunized per recombinant protein following the 87-day program (Eurogentec). Pre-immune bleeds (hereafter called PPI) were performed before the first injection and served as negative controls. Then, 100 µg of purified protein were used per rabbit and per injection. Four injections were performed at 0, 14, 28 and 56 days. The final bleed was carried out to sample serum containing antibodies (hereafter called SAB) on the 87th day.

### 3. Results

#### 3.1 Identification of HMG-interacting proteins

The ability of HMG to interact with other proteins was assessed using a GAL4-based Y2H assay. The Y2H assay is based on restoration of functionality to a split TF protein (in our case the yeast TF GAL4) whose DNA binding domain (BD) and the transcriptional activation domain (AD) are each separately fused to different proteins (the prey and bait proteins) to test for interaction between these two proteins. If interaction between the two proteins of interest occurs, the function of the split transcription factor is restored triggering the expression of reporter genes (Figure 3). Plasmids containing the BD-fused and AD-fused constructions can be introduced into yeast strains of opposite sex (MATa and MATα). Thus, thanks to mating, it is possible to analyse the interaction between a prey protein and a bait protein.

#### 3.2 Construction of a Y2H library

Our study included a bait protein (HMG) but no identified prey protein. In order to screen the largest possible panel of prey proteins, a library of prey proteins from *Ectocarpus* sp. gametophytes had to be prepared. Using the recombination machinery of *S. cerevisiae*, it is possible to rapidly construct full-length transcript libraries with any mRNA source. We therefore constructed a Y2H GAL4-TAD recombinant library using mRNA extracted from gametophytes of male strain Ec32. The library contained 6,400,000 independent clones with a mean insert size about 2 kbp. Considering that an efficient library should have more than 1,000,000 independent clones, our library can be considered of good quality. In addition, the total number of yeast cells present was  $3.3 \times 10^7$  cfu/ml, which was higher than the recommended  $1 \times 10^7$  cfu/ml. Finally, screening of 40 random clones showed that the library contained about 99% recombinants. Taken together, these results show that the library had all the characteristics necessary for an efficient Y2H assay.

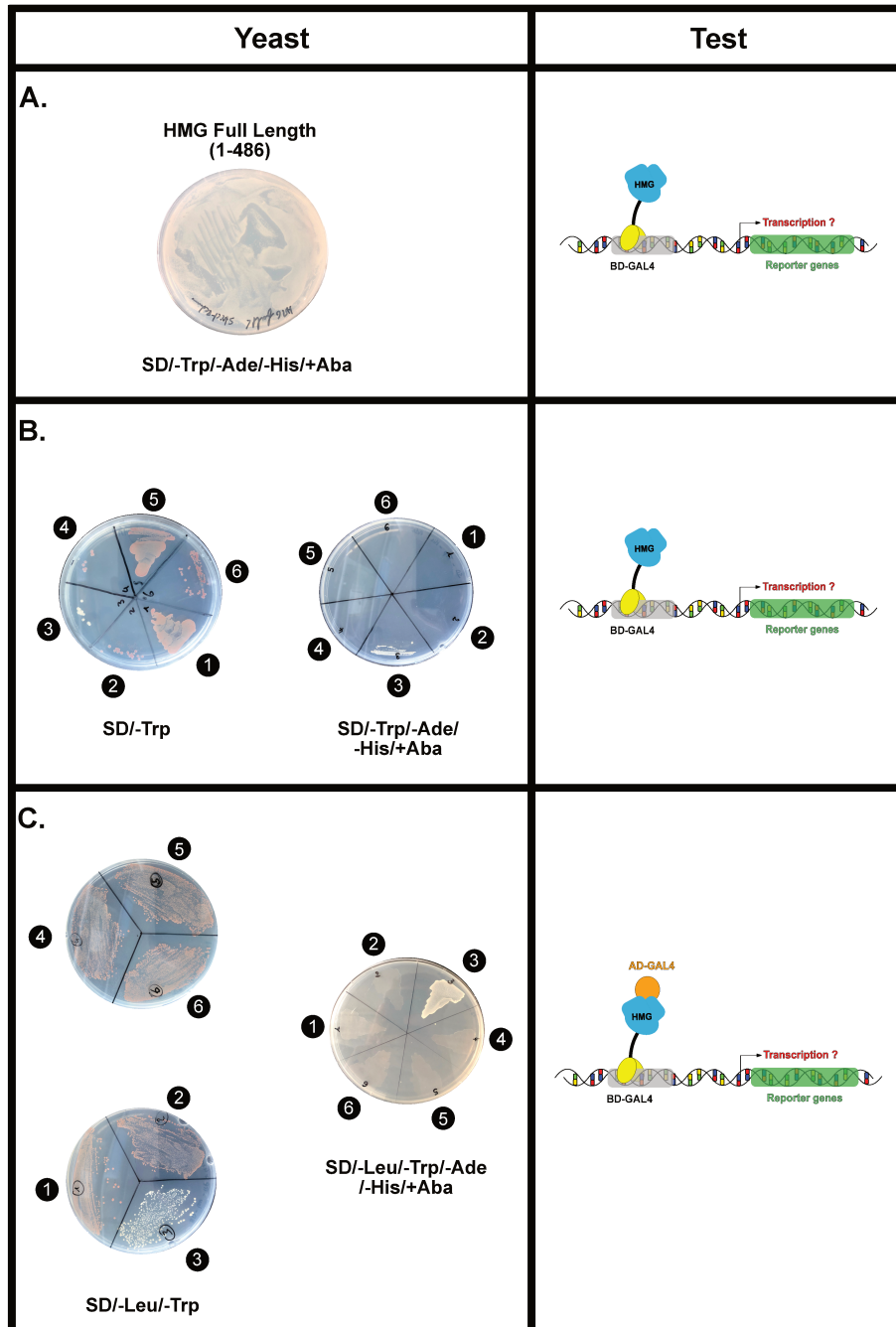


Figure 5. **Self-activation tests.** (A) Y2H yeasts with the construction BD-GAL4-HMG-sex 1-486 were spread on strict medium SD/-Trp/-Ade/-His/+Aba to determine whether HMG 1-486 was able to activate reporter genes in the absence of the AD-GAL4 domain. The yeasts grew, thus indicating self-activation of reporter genes by the HMG protein construct. (B) Yeasts transformed with 6 different constructions: (1) pGBK empty, (2) plasmid negative control p53, (3) pGBK AD/BD GAL4 positive control, (4) pGBK HMG 1-370, (5) pGBK HMG 280-486, (6) HMG 1-462. The haploid yeasts were first spread on SD -Trp medium as a growth control, then on SD/-Trp/-Ade/-His/+Aba medium to determine whether the construct autonomously activated the expression of reporter genes without the activation domain of GAL4 (AD- GAL4). (C) The yeasts were then mated with a yeast transformed with an empty pGAD plasmid (i.e. only having the AD-GAL4 domain). They were then spread on the SD -Leu -Trp medium as a growth control, then on the SD -Leu -Trp -Ade - His +Aba medium to determine whether the HMG protein interacted with the AD-GAL4 domain alone. In both tests (B) and (C), only yeasts with the positive control construction pGBK AD/BD GAL4 reconstituted GAL4 and grew. There was therefore no self-activation of reporter genes by the HMG protein in either case.

The full-length coding region of *Ec-13\_001750* (0-486 amino acid) was fused to the GAL4-BD domain, in order to obtain a bait protein. For a bait protein to be used in a Y2H analysis, it is necessary to verify that it does not autonomously activate reporter genes in the absence of an interaction with a prey protein (Figure 4). The Y2H strain with the fusion protein (full-length HMG and GAL4-DBD domain) was therefore spread on a selective medium (requiring the activation of reporter genes to develop) and on non-selective medium as a control. The full-length HMG construction activated the expression of reporter genes in the absence of prey protein (Figure 4A), indicating that it is auto-activating. To solve this problem, it was necessary to identify the region of the HMG protein that activated the reporter genes. We therefore established three different constructs of the HMG protein (1-370, 280-486 and 1-462), focusing on the regions conserved between the different homologous HMGs present on the V chromosomes of four other brown algae (*Ectocarpus fasciculatus*, *Scytosiphon lomentaria*, *Macrocystis pyrifera* and *Undaria pinnatifida*). We also used the predicted intrinsic disorder to determine conserved regions (Figure 5A). These three constructs did not auto-activate the reporter genes, suggesting that both the N-terminal and C-terminal parts are required for the autoactivation (Figure 4B-C).

### Interaction analysis

Based on the above results, high-throughput interaction analysis was carried out by mating the strain Y2H expressing the GAL4-BD–HMG (1-462) fusion protein with strain Y187 containing the GAL4-AD library. Mating on strict selection medium (SD/-Leu/-Trp/+Aba) resulted in the development of only three yeast colonies (Figure 6). The three colonies were transferred on a more stringent medium (SD/-Leu/-Trp/-Ade/-His +Aba). Only two colonies also developed on this stringent medium, indicating they were true positives (Figure 6). However, the extraction of the prey plasmid was technically impossible, preventing the identification of the potential HMG interactant.

The low number of colonies that developed on the strict medium was surprising. Mating controls (two replicate experiments) were carried out in parallel to evaluate the number of diploid yeast. These control experiments resulted in 14,300 and 16,700 colonies in the replicate 1 and replicate 2, respectively. These numbers correspond to a very low level of mating, compared to the 1,250,000 colonies obtained during a double hybrid experiment between the ORO transcription factor (Bourdareau, 2018) and a yeast library made from sporophyte mRNA.





Therefore, it appears that the low mating efficiency may explain the low number of yeasts that developed on the strict medium (SD/-Leu/-Trp/+Aba).

Two hypotheses can be put forward to explain this low level of mating. Either yeast mating was disturbed by the HMG protein or the yeast library has lost its efficiency over time.

#### Identification of the cause of the low mating efficiency

In order to understand the reason for the low mating efficiency, we performed an experiment involving several crosses (Table 1). For clarity, Ec-13\_001750 is referred to as "HMG-sex" and the other HMG gene as "HMG-autosomal". The HMG-sex protein was expressed in both strain Y2H (MATa) and strain Y187 (MAT $\alpha$ ) to verify its influence on mating when expressed in the MATa strain. In parallel, an autosomal HMG from *Ectocarpus* was used as a control, in order to determine if any influence on mating was due specifically to the HMG-sex protein. HMG-autosomal was chosen to have characteristics similar to HMG-sex, in terms of size and HMG domains. Finally, various crosses were also carried out with the gametophyte library and the sporophyte control library in order to identify potential problems related to the gametophyte library.

The results of these crosses showed that Y2H pGBK HMG-sex / Y187 pGAD  $\Phi$  (where "/" denotes a cross) and Y2H pGAD  $\Phi$  / Y187 pGBK HMG-sex had the same number of colonies. Moreover, the number of colonies was not different from the control mating Y2H pGBK  $\Phi$  / Y187 pGAD  $\Phi$  nor from the matings Y2H pGBK HMG-autosomal / Y187 pGAD  $\Phi$  and Y2H pGAD  $\Phi$  / Y187 pGBK HMG-autosomal (Table X). This result indicated that HMG-sex did not influence mating efficiency in MATa yeast strains (Table 1). On the other hand, the different crosses carried out using two different yeast libraries (gametophyte and sporophyte) showed a low mating efficiency for the gametophyte library compared to the sporophyte library, whether for the pGBK  $\Phi$ , the pGBK HMG-sex or the pGBK HMG-autosomal (Table 1). Taken together, these results indicated that the Y2H library constructed from the gametophyte mRNA had a low mating rate and was therefore responsible for the low level of diploids obtained after mating (Figure 7A). In addition, a titration was carried out on the gametophyte library to determine the evolution of the number of colonies. At a dilution of 1,000, the library produced only 88 colonies compared with 400 at the time of its creation and 440 for the sporophyte library (Figure 7B). All these results are consistent with a problem with the gametophyte Y2H library. This problem was most probably responsible for the low efficiency of mating and the failure of this approach to identify interactors of the HMG-sex protein.

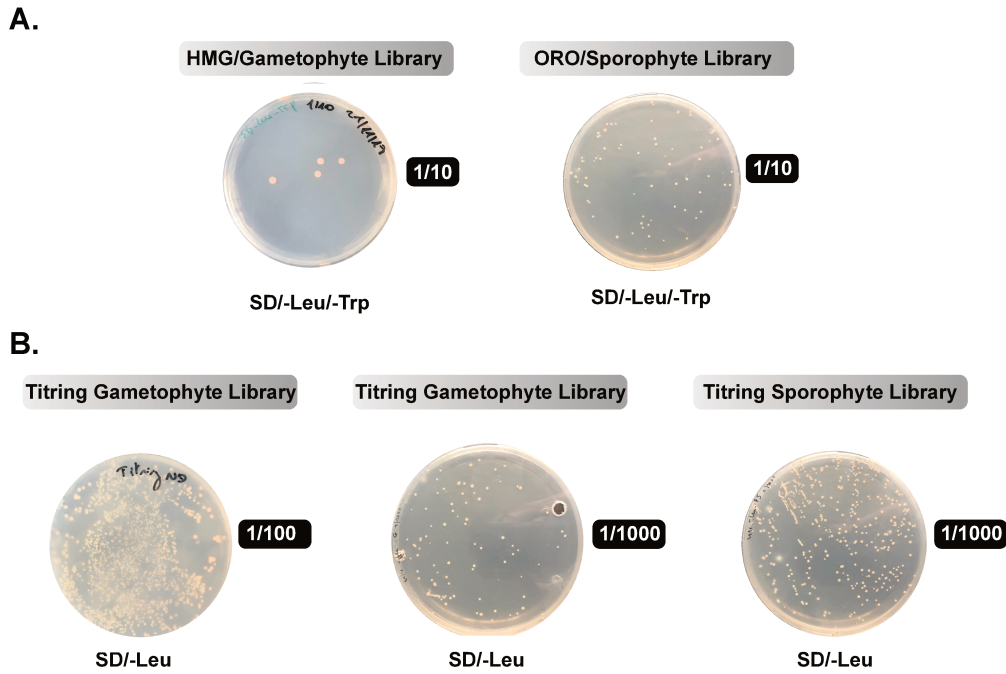


Figure 7. **Y2H library quality control.** (A) Low mating efficiency between Y2H pGBKT7 HMG-sex and the gametophyte Y2H library. (left) Mating efficiency test (HMG/Gametophyte Library) (right) Control mating efficiency test (ORO/Sporophyte Library). The number of diploid yeasts was much lower for the HMG/Gametophyte Library cross than for the ORO/Sporophyte Library control cross, suggesting a mating problem. (B) (from left to right) Titre of the gametophyte Y2H library when it was constructed, titre of the gametophyte YH2 library one year later. Titre of the control sporophyte YH2 library. The gametophyte YH2 library produced fewer colonies than the sporophyte yeast library, both after construction and one year later.

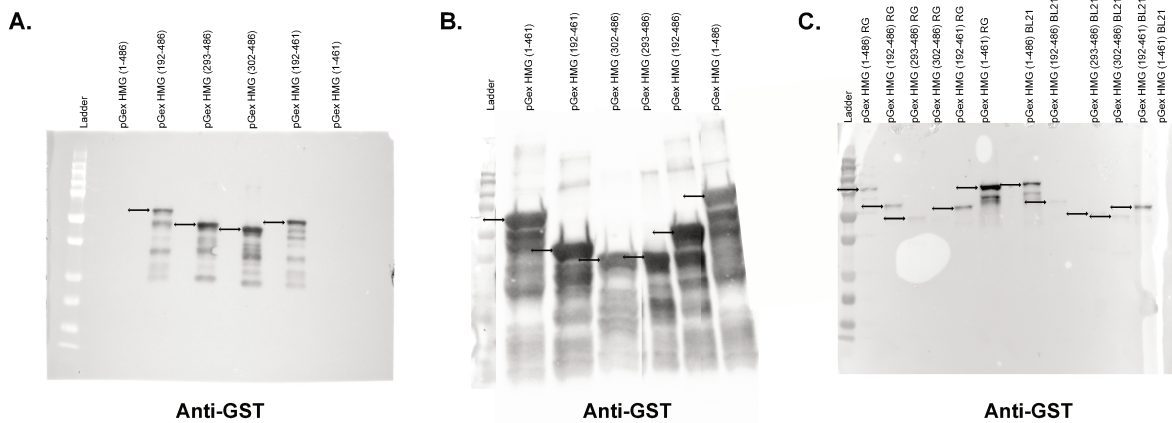


Figure 8. **Expression of HMG-sex protein constructs in *E. coli*:** (A) Western blot of the HMG-sex full-length protein (1-486) and 5 constructions with a GST tag expressed in bacteria strain BL21. (B) Western blot of the HMG-sex full-length protein (1-486) and 5 constructions with a GST tag expressed in the bacterial strain Rosetta gami. (C) Western blot of the HMG-sex full-length protein (1-486) and 5 constructions with a GST tag expressed in the bacteria strain Rosetta gami (RG) and in strain BL21. The production of the proteins was carried out under thermo-saline stress. All western blots were probed with an anti-GST antibody. Size expected are indicated by an arro

### 3.3 Genome-wide search for HMG binding sites *in vivo*

This section describes an attempt to produce an anti-HMG-sex antibody for use in ChIP-nexus experiments aimed at identifying HMG-sex binding sites in the genome of *Ectocarpus in vivo*.

The HMG-sex coding region was cloned in an overexpression vector pGex4T1 downstream of a glutathione S-transferase (GST) tag. When cloning the full length HMG-sex sequence (1-486 amino acids), most transformed bacteria contained only an empty circular plasmid without the region encoding the HMG protein. A screen of 160 bacteria identified one clone transformed by a plasmid with the HMG 1-486 insert. However, after sequencing, it was found that a mutation had occurred during cloning resulting in a stop codon at position 345 in the HMG domain. This mutation was corrected by *in vitro* mutagenesis and a series of sub-constructs were then obtained by deletion of sequences from the pGex-HMG(1-486) construct. Overexpression of the full length protein (1-486 aa) in bacteria was not efficient (Figure 8A). However, among the different constructions designed to promote protein overexpression, the construction pGex HMG(1-461) didn't allow expression of a recombinant protein in BL21. The constructions pGex HMG(1-192), pGex HMG(293-486), pGex HMG(302-486) and pGex HMG(192-461) were expressed but exhibited a high level of protein degradation. pGex HMG(1-461) and pGex HMG (1-486) were expressed at a high level in Rosetta-gami, but also with a high level of degradation (Figure 8B). However, the level of degradation was greatly reduced when bacteria were subjected to thermo-saline stress (Figure 8C) and the construction pGex HMG(1-461) has the highest level of expression. In conclusion, the pGex-HMG(1-461) construction with thermo-saline stress in Rosetta-gami exhibited the best expression level and these conditions were subsequently used for the overexpression of the HMG protein.

It was concluded from the above experiments that the application of thermo-saline stress was necessary to stabilize the protein. In addition, since these experiments involved heterologous expression of a eukaryotic protein in a prokaryotic system, the property of Rosetta-gami to alleviate codon bias seemed to be necessary for the proper expression of the HMG protein.

Two independent batches of antibodies were raised against the HMG transcription factor in rabbits using the protein expressed from the construct pGex-HMG-sex(1-461).

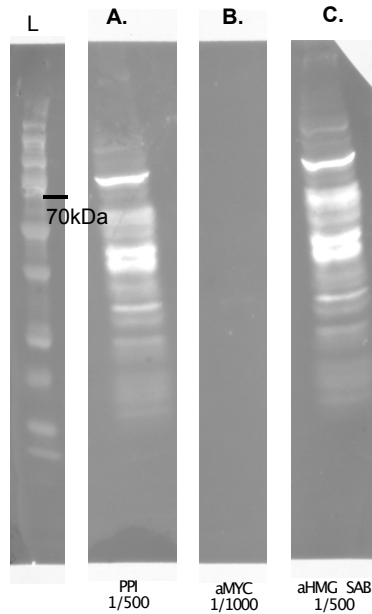


Figure 9. **Western blot of HMG-sex protein expressed *in vitro* with the TNT Quick Coupled Transcription/Translation System with rabbit polyclonal antibodies raised against HMG-sex.** (A) Western blot with the pre-immune serum (PPI). (B) Western blot with an anti-Myc antibody as a control of the presence of the HMG protein. (C) Western blot with the anti-HMG-sex antibody. Size expected: 70kDa

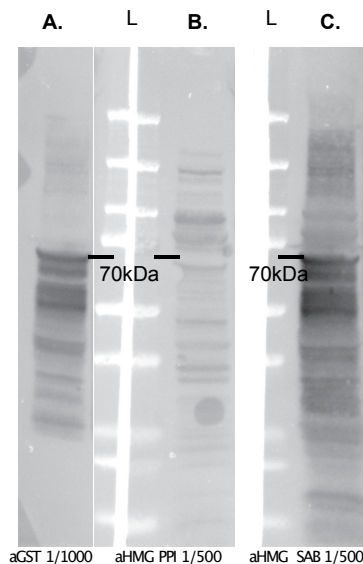


Figure 10. **Western blot of HMG-sex protein expressed in strain Rosetta gami with rabbit polyclonal antibodies raised against HMG-sex.** The western blot was performed on bacterial lysate without purification of the HMG-sex protein (1-461). (A) Western blot with an anti-GST as a control of the presence of the HMG protein in the lysate. (B) Western blot with the pre-immune serum (PPI.) (C) Western blot with the anti-HMG-sex antibody. Size expected: 70kDa

For each batch, pre-immune sera were sampled before the first immunisation. To verify the functionality of the antibodies, HMG protein was produced with a MYC tag using recombinant clones in pGADT7-AD and the TNT Quick Coupled Transcription/Translation System (Promega). Expression of the Myc-tagged protein was verified using an anti-Myc antibody. The two sets of anti-HMG antibodies did not recognize the HMG protein (Figure 9). However, since the anti-Myc antibody did not recognize the Myc-HMG protein either, it is possible that the protein was not expressed in this experiment. Western blotting of a bacterial lysate containing the fusion protein GST-HMG was performed using the anti-HMG antibodies and with an anti-GST antibody as control. The HMG protein was detected by anti-GST antibodies and by the serum containing antibodies (SAB). In addition, background noise is very important due to the degradation of HMG protein and bacterial proteins. It is therefore difficult to distinguish whether HMG protein is recognized by pre-immune serum (PPI) (Figure 10). To conclude, in order to perform the ChIP-seq experiment, it will be necessary to purify the HMG protein to verify if the PPI does not recognize the HMG<sub>sex</sub> protein. In addition, a western blot on a male gametophyte protein extract will have to be performed to verify whether the HMG protein is recognized by the antibody in *in vivo* concentrations, and to measure the background noise captured by the anti-HMG antibody. Finally, given the presence of 13 HMG-box domain proteins in *Ectocarpus*, it will be necessary to verify the cross-reaction of the antibody with these proteins.

### 3.4 Genome-wide search for HMG binding sites *in vitro*.

As an alternative to the ChIP-nexus experiment, we employed a recently published method called DAP-seq (Bartlett et al., 2017; O'Malley et al., 2016). The DAP-seq approach investigates interactions between recombinant TFs and genomic DNA *in vitro*. This method does not require antibodies nor transgenic lines that overexpress the TF. As genomic DNA is used, it can detect binding capacities that are dependent on the genomic context.

The protein fusion GST-HMG-sex(1-461) was expressed in *E. coli* and then bound via GST to magnetic beads MagneGST Glutathione Particles. In parallel, *Ectocarpus* genomic DNA was extracted and sonicated to obtain DNA fragments of between 200 and 400 bp. The sonicated DNA and the recombinant protein bound to the beads were co-incubated to allow the HMG protein to capture DNA fragments by binding to its target sequences. The recovered DNA was then sequenced and mapped to the reference genome. MACS2 (Zhang et al., 2008) was used to analyse enrichment in reads corresponding to HMG protein binding sites.

We identified 87 and 84 peaks respectively in the first and second replicates. The background level of peaks in the GST control was 29 and 18 common with both HMG replicates. Twenty peaks were common to both HMG replicates but absent from the GST control.

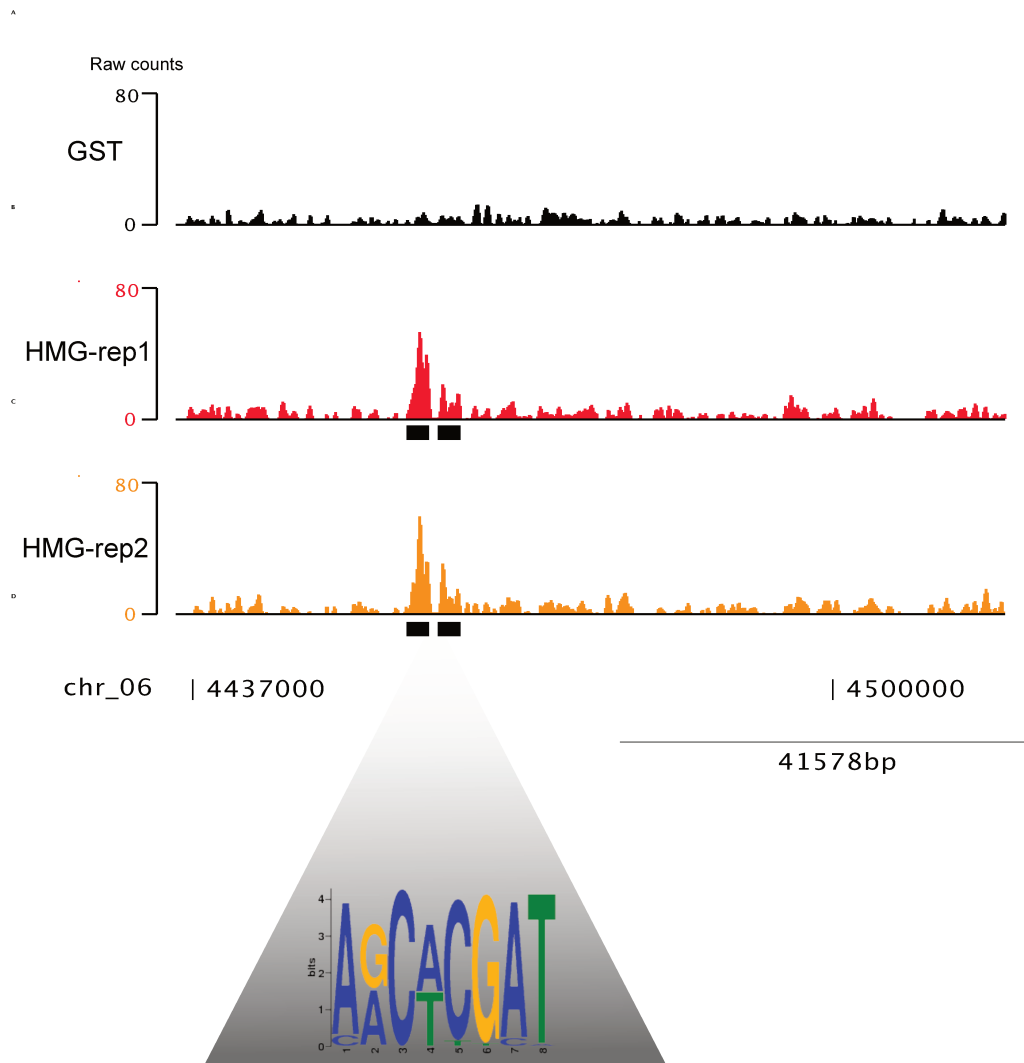


Figure 11: Example of a genomic region of the chromosome 6 showing distributions of mapped DAP-seq reads for the protein HMG-sex (Ec-13\_001750) and the GST control and the consensus DNA binding sequence of the HMGsex protein. Black blocs represent peaks detected by MACS2.

Although the number of peaks obtained was unexpectedly low, Manual curation of the read mapping for both HMG-sex replicates, the GST controls and the input control using IGV confirmed that the data was homogeneous at the level of coverage and that (Figure 11) all the peaks corresponded to bona fide peaks compared with the GST input.

The number of peaks identified was insufficient to carry out a statistically robust search for conserved binding motifs but a search of motifs was nevertheless carried out to have an overview of the sequences common to the discovered peaks. The aim was to assess the nature of the enrichment sites to determine whether or not they had characteristics that would indicate that they were artefactual. If this is not the case, it can be interesting to obtain a candidate consensus DNA binding sequence even if the consensus is not supported statistically as this may provide information for further experiments (Figure 11). Both MEME and KMAC were used to analyse the 20 HMG-specific peaks to detect conserved motifs. Both analyses detected the following motif: AA/GCA/TCGAT. This motif does not correspond to a repeated sequence, which would have been indicative of artefactual peaks. The AA/GCA/TCGAT motif may therefore be considered as candidate binding site that will have to be confirmed by further experimentation.

## 4. Discussion and conclusion

### 4.1 Results of the search for proteins that interact with the HMG protein Ec-13\_001750

HMG transcription factors, including SRY and the related SOX proteins, are part of the HMGB (HMG-box protein) subgroup of the HMG superfamily (Figure 2 ; Postnikov and Bustin, 2016). These proteins, which are known to be involved in many developmental processes (She and Yang, 2015), often function in association with partner proteins. For SOX-partner complexes, partner proteins not only influence the specific recognition of the binding sites of on target genes, but also determine transcription activities and significantly improve the activation/repression potential (Kamachi and Kondoh, 2013; Wilson and Koopman, 2002). For example, SRY associates with SF1 and WT1 to activate target gene expression, while interaction with KRAB-O allows negative regulation of target gene expression (Matsuzawa-Watanabe et al., 2003; Oh et al., 2005; Oh and Lau, 2006; Peng et al., 2009; Polanco et al., 2009; Sekido and Lovell-Badge, 2008). Ec-13\_001750 (HMG-sex) is predicted to be a member



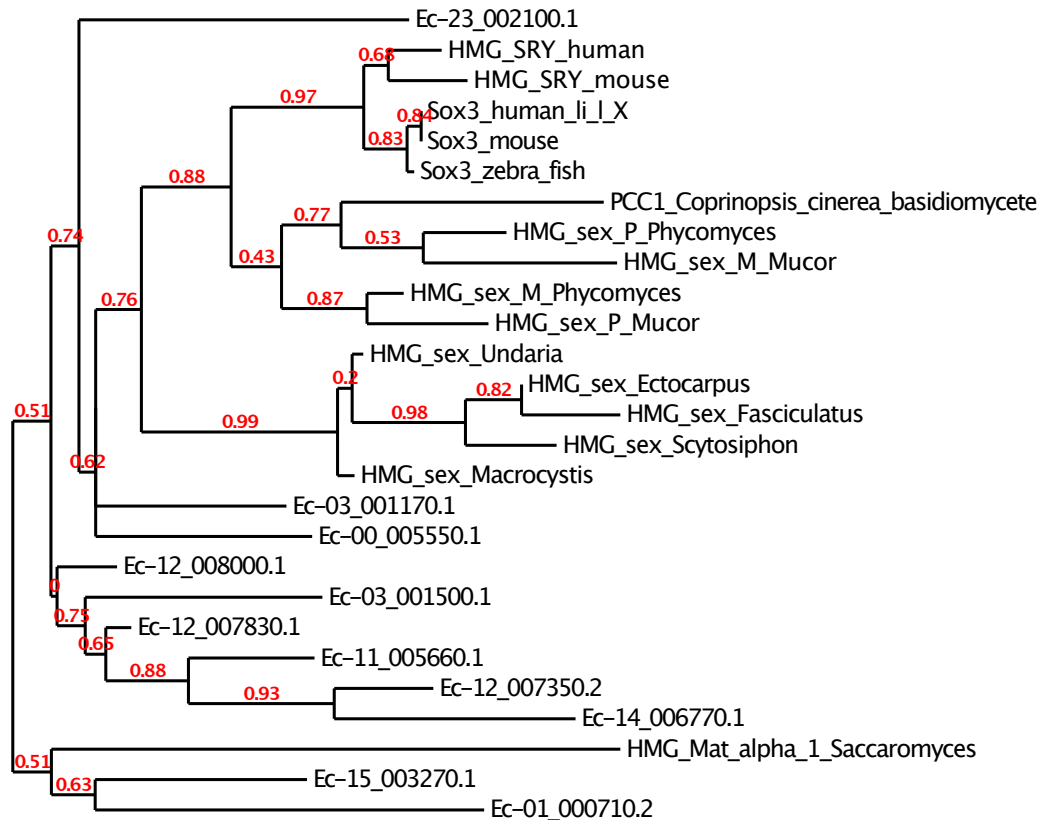


Figure 12. **Phylogenetic analysis of diverse HMG-box domains.** Unrooted maximum likelihood tree based on an alignment of HMG-box domains from diverse species. The nomenclature starting with Ec represents HMG-boxes from proteins encoded by *Ectocarpus* autosomal genes. SRY and SOX3 are derived from the same ancestral gene. *Mucor*, *Mucor mucedo*; *Phycomyces*, *Phycomyces blakesleeanus*; *Ectocarpus*, *Ectocarpus siliculosus*; *Undaria*, *Undaria pinnatifida*; *Fasciculatus*, *Ectocarpus fasciculatus*; *Scytosiphon*, *Scytosiphon lomentaria*; *Macrocyctis*, *Macrocyctis pyrifer*. Protein sequences were aligned with Muscle and the phylogenetic tree was generated using PhyML and 100 bootstrap replicates. Bootstrap support values are indicated at each node.

of the HMGB family since it has an HMG-box. In addition, it is more closely related to HMGs known to be involved in male sex determination, than to the other autosomal HMG-box proteins encoded by the *Ectocarpus* genome (Figure 12). Our objective in this chapter was to understand how this protein behaved in comparison to other HMGs involved in sex determination.

It was not possible to identify proteins that interact with the HMG-sex protein because problems were encountered when a gametophyte Y2H library was screened. The mating between yeasts carrying the Y2H library and the HMG bait construction resulted in too few diploid yeasts to provide potential interactants. One hypothesis we considered to explain the low number of

diploids obtained was that the HMG-sex protein influenced the yeast mating. Mating is controlled in *Saccharomyces cerevisiae* by two loci (Figure 13): MATa and MAT $\alpha$ . MAT $\alpha$  codes for two proteins MAT $\alpha$ 1 and MAT $\alpha$ 2. MAT $\alpha$ 1 is a transcription factor of the HMG family that activates the expression of  $\alpha$ -specific genes while MAT $\alpha$ 2 is a homeodomain protein that acts as a repressor of  $\alpha$ -specific genes (Haber, 2012). The *Ectocarpus* HMG protein was expressed in MATa strain Y2H. It was therefore possible that the presence of both HMG-sex and MAT $\alpha$ 1 in this strain could have suppressed mating, although these two proteins are relatively distant from a phylogenetic point of view compared to other transcription factors of the HMG family that are involved in sex determination (Figure 12). However, HMGs are transcription factors with low DNA binding specificity (Thomas and Travers, 2001). It was therefore possible that HMG-sex could have modified the mating type and thus impacted mating with the Y2H library. However, the results of various test crosses indicated that HMG-sex had no impact on mating (section 3.1.3).

An alternative hypothesis, was that the experiment failed to detect interactors due to problems with the quality of the gametophyte Y2H library. Several parameters that were measured immediately following the construction of the gametophyte Y2H library indicated that the library was of good quality. These included the number of colonies per milliliter and the recombination rate. However, when the library was titered at a later date, it contained five times fewer yeast per ml than when it was created, indicating a problem of stability during storage. For comparison, a Y2H library constructed from sporophyte mRNA maintained a constant number of yeast cells over time.

One solution to this problem for the future would be to construct a new gametophyte Y2H library. Alternatively, other methods could be tried, such as co-immunoprecipitation, crosslinking protein interaction analysis or a pull-down assay with whole cell extract.

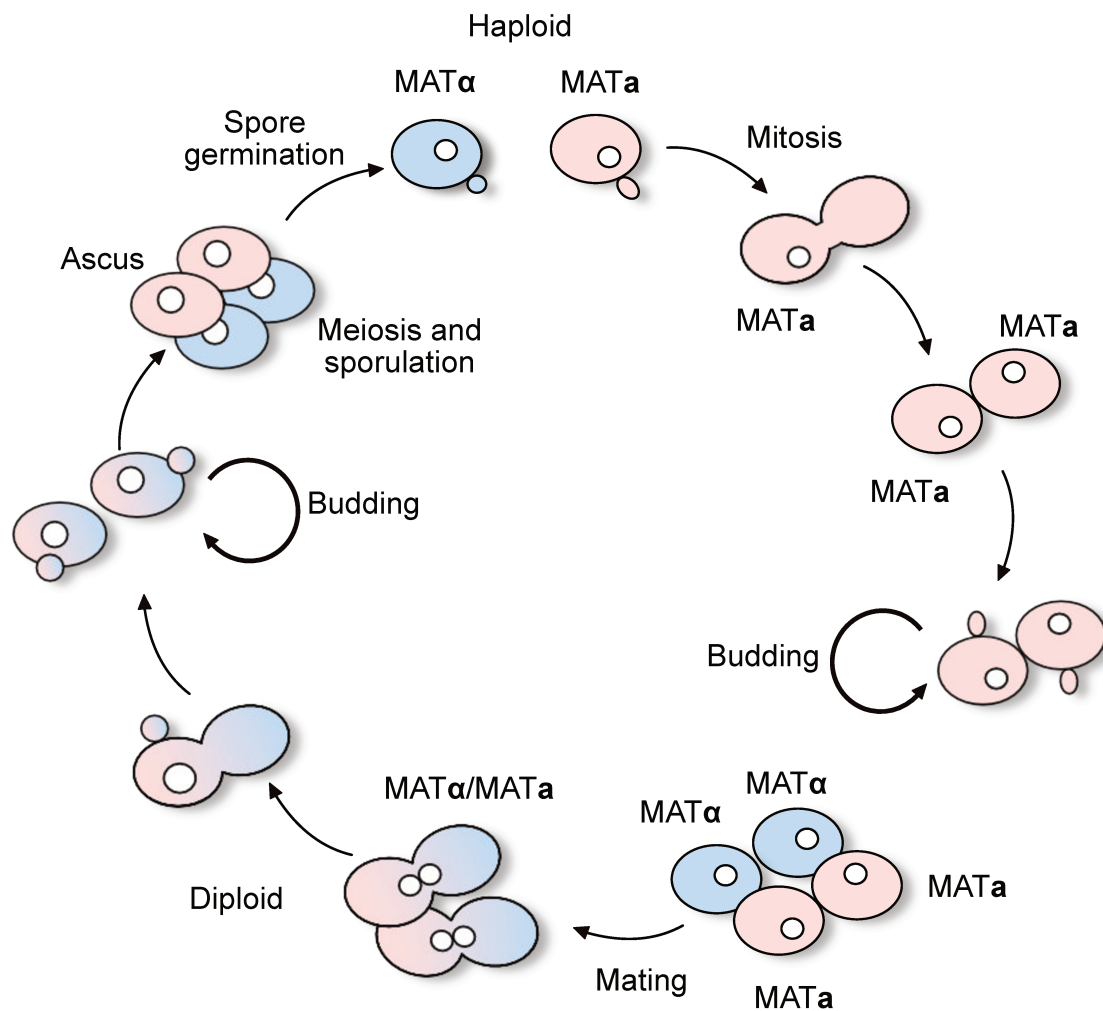


Figure 13. **Yeast life cycle.** Homothallic life cycle of *Saccharomyce cerevisiae*. Homothallic MATa (light pink) and MAT $\alpha$  (light blue) yeasts can reproduce vegetatively by mitosis following budding. In addition, these cells can conjugate to form a zygote that gives rise to MATa/MAT $\alpha$  diploids (mix light pink and light blue). Heterothallic cells grow vegetatively by budding until they exhaust their nutrients and enter stationary phase. Under nitrogen starvation, diploids undergo meiosis and sporulation to produce four haploid spores (two MATa and two MAT  $\alpha$ ) in an ascus. The spores germinate and grow vegetatively and can repeat the homothallic cycle. Adapted from (Haber, 2012).

If one of these methods detects potential interactors, it would then be necessary to validate the interaction using an alternative method, which could be a specific Y2H assay between the two proteins. If an interactor could be identified, the different constructions of the HMG-sex protein domain would be useful to identify the protein-protein interaction site(s).

Work carried out with HMG-box proteins from other species has provided some indication of what sort of interacting proteins these experiments might be expected to find. These could include proteins involved in enhanceosome formation or in chromatin remodelling (although a direct parallel with Sry interactants, for example, cannot be made given the absence of genes encoding KRAB domain proteins in the *Ectocarpus* genome). The function of HMG-box proteins as architectural factors is determined not only by their DNA-binding properties (establishing proper DNA conformation), but also by their ability to interact with a plethora of proteins to promote formation of complex nucleoprotein structures. For example, one proposed mechanism for HMG-box protein architectural function in transcription assumes that HMG-box proteins can facilitate binding of the sequence-specific proteins through direct interactions with transcription factors and pre-bending of the DNA target sequences (Štros, 2010). (Figure 1).

#### 4.2 Results of the search for DNA-binding sites of the HMG protein Ec-13\_001750.

HMG-box proteins are protein-chromatin interactors that bind to the minor groove of B-form DNA, significantly widening the groove and introducing a bend of 90° or more into the backbone. HMG-box protein proteins also bind with high affinity to already distorted DNA structures such as four-way junctions, bulges, kinks and modified DNA containing cisplatin adducts (Pil and Lippard, 1992). The two main mechanisms that enable HMG-box proteins to function as chromatin “architectural” proteins are induction of DNA bends and recognition of distorted DNA structures. For example, it has been suggested that DNA bending induced by HMG-box proteins produces an allosteric transition structure that promotes recognition and binding by other proteins during the formation of functional multiprotein DNA complexes. On the other hand, HMG-box protein recognition and binding to already distorted DNA conformations is thought to be analogous to enzymes recognizing molecular structures that resemble transition states between reagents and products and, as a consequence, these interactions would influence the rate of formation of multiprotein DNA complexes (Agresti and Bianchi, 2003). HMG-box protein bind to DNA in a different manner to other transcription factors and it has therefore been suggested that they can act as pioneering factors when binding

to chromatin and nucleosomes (Figure 1) (Slattery et al., 2014). Indeed, HMG-box protein can modulate structural changes in chromatin by association with ATP-dependent chromatin remodelers resulting in nucleosome sliding. This sliding is accompanied by chromatin decompaction and DNA accessibility for other transcriptional factors.

The ChIP-seq experiments, necessary for the detection of DNA binding sites for HMG-sex protein, could be performed if additional verifications are carried out on anti-HMG antibodies. The SAB antibodies, produced by injection of HMG-sex protein that had been expressed in *E. coli*, seem to recognize HMG-sex protein in a western blot. However, a weak signal is visible in western blot for the pre-immune serum (PPI) for the lysis of the bacteria. Thus, it will be necessary to verify the recognition of the HMG protein by the antibody on a purified protein, without observing any PPI signal. In addition, a western blot will have to be made on a protein extract from male gametophytes of *Ectocarpus*, in order to verify the recognition of the HMG protein by the antibody in protein concentrations *in vivo* and to evaluate the level of non-specificity of the antibody. Finally, given the presence of 13 other HMG-box domain proteins in *Ectocarpus*, it will be necessary to verify that SAB antibodies do not cross-react with these proteins.

Another approach to overcome problems of protein specificity would be to select a single peptide specific to the HMG protein. This solution would also circumvent any problems associated with overexpressing the HMG protein in bacteria and may provide a means to limit cross-reaction with other HMG proteins present in *Ectocarpus*. There are 14 HMG domain proteins in *Ectocarpus* in total and, as the current antibodies are polyclonal, they could potentially recognize other HMG domains. Production of monoclonal antibodies would be another means to increase recognition specificity.

DAP-seq has been proposed as a simple and efficient alternative to more technically difficult methods such as ChIP-seq or SELEX-seq (Bartlett et al., 2017; O'Malley et al., 2016). When applied to several *A. thaliana* TFs, DAP-seq identified up to 80% of the peaks found using ChIP-seq. Since DAP-seq involves binding TFs to naked DNA *in vitro*, the test TF can potentially bind to all sequences corresponding to its DNA binding motif, including regions that may not be accessible *in vivo*. It has been suggested that DNA recognition involves a conjunction between sequence and shape characteristics (Slattery et al., 2014). DNA shape, i.e. DNA structural characteristics such as groove width, roll, helix torsion, helical torsion (Zhou et al., 2013), explain the biologically relevant results obtained in *in vitro* experiments using

methods such as genomic context protein binding microarrays (gcPBM) or DAP-seq (Gordân et al., 2013). Local DNA shape for any base pair in the genome is dependent on the nucleic acid composition within a window of generally less than 10 bp.

For the DAP-seq approach, although the number of peaks detected was relatively low, we could detect a candidate DNA binding sequence for the HMG-sex. Indeed, only 20 reproducible, HMG-specific peaks were detected, significantly lower than what is normally detected in a DAP-seq experiment (Bartlett et al., 2017). Pattern discovery was nevertheless used to determine if these 20 peaks corresponded to artefacts due to the presence of a repeated sequence. This was not the case, since both MEME and KMAC detected the motif AA/GCA/TCGAT. This motif can therefore be considered as a candidate binding motif for the HMG-sex protein. Interestingly, the SRY and SOX proteins bind DNA at a similar binding motif (A/TAACAA/ T).

Several hypotheses can explain the low number of peaks detected during the DAP-seq experiment. We first compared the number of good quality reads before mapping with the number of mapped reads. For all of our samples, more than half of the reads did not map to the genome, with a maximum of 43% unmapped reads for HMG-sex replicate 1. The large proportion of unmapped reads could be due to two possible factors: DNA contamination or the sequence quality of the reference genome. The *Ectocarpus* genome assembly is of good quality (Cormier et al., 2017), suggesting that contamination with bacteria DNA during the extraction step may be the most likely explanation. This hypothesis was verified by an unmapped read blast, revealing that more than one in three reads was of bacterial origin. Another factor that could have prevented the detection of peaks corresponding to HMG-sex binding sites would be poor read coverage. Coverage was between 8.6x and 11.4x, depending on the sample, and about 20% of the genome had at least 15 overlapping reads (i.e. a coverage of 15x) for all samples, which corresponds to a good coverage. Read coverage should therefore have been sufficient and therefore was not the reason for the absence of detectable peaks (Figure 14). In conclusion, it would therefore be possible that the low number of peaks detected could be a consequence of the low read mapping rate of reads, preventing peaks from emerging from the background noise. Finally, the HMG protein used for the DAP-seq experiment was truncated, corresponding to the C-terminal region (462-486) and it is possible that this impacted the conformation of the protein, preventing efficient binding to DNA target sequences.

It is important to take precautions with regard to the motive discovered by the Dap-seq. Indeed, in addition to the small number of DNA binding sites found, a study of the protein domains present in the HMG-sex protein of *Ectocarpus* revealed the presence of a second HMG domain downstream of the first, which appears to be preserved in other brown algae (Figure 2). HMG-box proteins are usually classified into two major groups, which exhibit differences in their abundance, function and DNA specificity. The first group consists of proteins containing two HMG boxes. These proteins generally have little or no DNA sequence binding specificity,

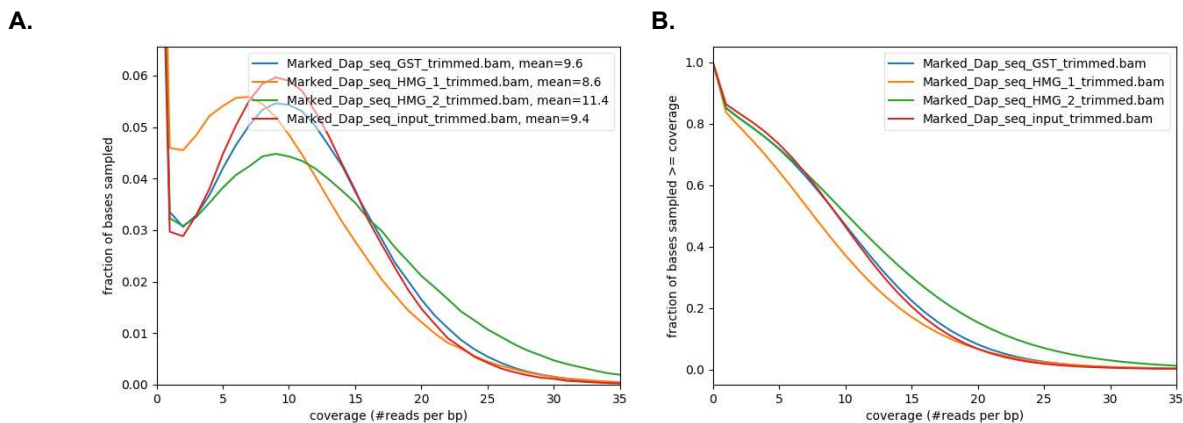


Figure 14. **Plot coverage of the sequencing depth of a Dap-seq samples** : (A) simply representation of the frequencies of the found read coverages. The mean coverage is indicated in the box. This graph indicated for example that 6% of the input sample are covered 10 times. (B) The second diagram shows us which fraction of the genome at which sequencing depth. For example, 43 % of the input sample have at least a coverage of 10

although some non-sequence specific proteins, such as plant HMG-box proteins, contain a single HMG box(Štros, 2010). The second group of HMG-box proteins is highly diverse and consists of much less abundant proteins, most of which have a single HMG-box (examples include the SRY and SOX proteins). HMG-box proteins with a single HMG domain recognize specific DNA sequences. The difference between sequence-specific and non-sequence-specific minor DNA groove binding by HMG boxes depends on changes in only a few amino acid residues. For example, sequence-specific HMG-box proteins like SRY contain a conserved asparagine at position 10 whereas non-sequence-specific HMG-box proteins have a conserved serine (Travers, 2000). In the HMG-sex protein, residue 10 is a serine, which would suggest non-specific DNA binding.

The DAP-seq experiment should be repeated in order to confirm the candidate DNA binding sequence. Repetition of the experiment will also make it possible to verify whether the number of peaks found was due to an experimental problem or whether it is replicable.

A consensus DNA binding sequence does not appear to be the only way to interact with DNA. Indeed, some HMG-box proteins can interact with nucleosomes (Figure 1C) (Bianchi and Agresti, 2005). The chromatin context is not preserved in DAP-seq experiments, therefore a complementary ChIP-seq experiment would be necessary to identify binding sites in an *in vivo* context. If a consensus binding sequence is detected in future DAP-seq experiments, it will be necessary to validate the motif by an alternative method such as a protein binding micro-array (PBM) (Godoy et al., 2011). In addition, it would be interesting to analyse the locations of binding sites in the genome. The detection of binding sites in the vicinity of SBGs would provide further support for their functionality and would allow the identification of candidates for genes directly regulated by HMG-sex, acting as a TF.

CRISPR-cas9 mutagenesis is currently being developed for *Ectocarpus* in the group. This method would represent a very effective means to investigate the functional role of HMG-sex. Alternatively, another reverse genetic method, such as screening of a tilling library, could also be used to investigate HMG-sex function. The expected phenotype would be phenotypically female with a male genetic background. This phenotype was observed in mice carrying a mutation in the *Sry* gene (McElreavey et al., 1992).

It has been shown that 12% of *Ectocarpus* genes exhibit sex-biased expression during the gametophyte generation (Lipinska et al., 2015). The majority of these genes are autosomal and therefore present in both male and female genomes. A major objective for the future will be to link the putative master regulator of male sex determination in *Ectocarpus* with the regulation of genes involved in the downstream differentiation pathway. This regulation may involve chromatin modification. In mammals, SRY interacts with the Krüppel-associated box only (KRAB-O) protein, which is involved in chromatin remodelling. Oh et al. (2005) suggested that the interaction between SRY and KRAB-O and its obligatory co-repressor Krab-associating protein 1 (KAP1) recruits heterochromatin protein 1 (Hp1), HDACs and the SETDB1 methyltransferase, which function as gene silencers by creating a repressive chromatin environment at female-specific genes. On the other hand, SRY activates the male determination pathway by directly inducing transcription of the *Sox9* gene. Interestingly, the SOX9 protein is itself involved in chromatin remodelling. SOX9 helps remove epigenetic signatures of transcriptional repression and establish active-promoter and active-enhancer marks (Liu et al., 2018). Moreover, HMG proteins, in general, are known to be involved in chromatin dynamics. Indeed, they are classed as architectural nucleosome-binding proteins, along with variants of the H1 protein family (Bianchi and Agresti, 2005; Bustin, 1999; Kasinsky et al., 2001). HMG



proteins bind to nucleosomes and induce local and global changes in the chromatin structure that affect various DNA-dependent activities, including transcription, recombination and repair. These proteins can also play a role in epigenetic regulation, since their interaction with chromatin affects the levels of histone modification (Gerlitz et al., 2009). Thus, our hypothesis is that, in *Ectocarpus*, males and females have sex-specific chromatin landscapes, and this difference is associated with the specific expression of sex-biased genes.

To test this hypothesis, a study of the dynamics of chromatin between male and female in the brown alga *Ectocarpus* was carried out using a ChIP-seq method and antibodies that recognised post-translational histone modifications. The ChIP-seq was accompanied by an RNA-seq analysis of males and females to link the expression of sexually-biased genes to the chromatin landscape.



## Chapter 3

# Epigenetic modifications associated with sex determination and differentiation in the brown alga *Ectocarpus*

## 1. Introduction:

Sex determination is a key developmental process that requires both the activation of the sex determination pathway for one gender (e.g. the female) and the repression of sex determination pathways for the other gender (e.g. the male). Sex determination is followed by a process of sexual differentiation involving genes differentially expressed between the male and the female, referred to as sex-biased genes (SBGs). Although multiple mechanisms are involved in the process of differentiation, chromatin modification appears to be an important component.

Epigenetics is the study of changes to gene function that do not involve modifications of the DNA sequence. Between 1940 and 1956, Conrad Waddington proposed the concept of an epigenetic landscape to describe the process of cell fate commitment during development. He envisioned that a cell progresses towards a differentiated state through a series of fate decisions that are stabilized by changes to its epigenome. These changes maintain cell-type-specific gene expression patterns that channel the cell along certain pathways while eliminating other possible pathways the cell might take (Garcia-Moreno et al., 2018). This mechanism of determinism represents a pluripotent system because it exhibits bipotentiality. For comparison, in pluripotent systems, such as embryonic stem cells (ESCs), promoters of developmental genes are often marked by both H3K4me3 and H3K27me3 histone modifications, two modifications that tend to be associated with gene activation and gene repression, respectively (Azuara et al., 2006; Bernstein et al., 2006). This “bivalency” plays a key role in pluripotency by maintaining these genes in a poised state for rapid activation or repression during differentiation. As pluripotent cells differentiate, chromatin becomes more specific by epigenetic deposition of DNA methylation and histone modifications that establish heritable cell-type-specific gene expression patterns (Atlasi and Stunnenberg, 2017). Based on this observation, one hypothesis is that the chromatin landscape in XX and XY bipotential progenitor cells is similar enabling equal access to promoters of sex-determining genes and regulatory elements of both sexes but that, after sex has been determined, the chromatin landscape becomes more sex-specific to fix either male or female fate and repress the alternative pathway (Garcia-Moreno et al., 2018). Upregulation of either the male or female sex determination pathway is accompanied by mutually antagonistic mechanisms that repress the alternative cell fate (Chassot et al., 2008; Kim et al., 2006; Matson et al., 2011; Ottolenghi et al., 2007; Uhlenhaut et al., 2009). At the chromatin level, this regulation is reflected in a reorganization of histone marks around sex-determining genes. Genes associated with the female pathway lose "repressive" marks when

the female pathway is activated and accumulate “active” marks, whilst genes associated with the male pathway lose "repressive" marks and accumulate “active” marks when the male pathway is activated.

This idea of bivalence has mainly been stated for mammals. But since the bivalence mechanism of histone PTMs seems to be conserved in all animals (Lesch et al., 2016), it would also be interesting to determine whether the chromatin landscape differs between males and females in distantly related species with different genetic sex determination system such as brown algae with their UV sex determination systems.

Evidence of bipotentiality in brown algae has been obtained for a male gametophyte mutant of the brown alga *Macrocystis* that was feminised in the absence of a female U chromosome (Müller et al., unpublished).

The establishment of sex, following the transition from the sporophyte to the gametophyte generation, requires genetic re-programming, and given the bipotential nature of gametophytes, we can hypothesize the existence of a chromatin landscape that becomes sex-specific to determine male or female fate.

In the previous chapter, we discuss how HMG sex determination genes modify chromatin states in animals. HMG transcription factors establish active or inactive chromatin regions and specifically regulate a limited number of genes (Bianchi and Agresti, 2005). This chapter will provide a link between the master sex-determining factor and the regulation of downstream effector genes (sex-biased genes) through an analysis of histone post-translational modifications in male and female strains of *Ectocarpus* sp. The analysis focused particularly on chromatin modifications at sex-biased genes and on the sex chromosomes.

The work described in this chapter represents the major results of the thesis. I carried out most of the work described. My contribution consisted of both experimental and bioinformatic analyse.

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

Josselin Gueno<sup>1</sup>, Simon Bordareau<sup>1</sup>, Guillaume Cossard<sup>1</sup>, Olivier Godfroy<sup>1</sup>, Agnieszka Lipinska<sup>1</sup>, Leila Tirichine<sup>2</sup>, J. Mark Cock<sup>1\*</sup>, Susana M. Coelho<sup>1\*</sup>

<sup>1</sup>Sorbonne Université, UPMC Univ Paris 06, CNRS, Algal Genetics Group, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, F-29688, Roscoff, France. <sup>2</sup>Université de Nantes, CNRS, UFIP, UMR 6286, F-44000 Nantes, France

\*Correspondence: [coelho@sb-roscoff.fr](mailto:coelho@sb-roscoff.fr); [cock@sb-roscoff.fr](mailto:cock@sb-roscoff.fr)

### 2.1 Summary

In many eukaryotes, such as dioicous mosses and many algae, sex is determined by UV sex chromosomes and expressed during the haploid, gametophyte stage. Male and female developmental programs are initiated by the presence of a U or V-specific regions, but phenotypic differentiation is largely driven by autosomal sex-biased gene expression. How the sex-biased transcriptional program in UV sexual systems is governed remains unknown. Here, we set out to understand the mechanisms involved in the regulation of sex-biased gene expression in *Ectocarpus*, a brown alga with UV sex chromosomes. We examine the chromatin landscape across different genomic regions, by contrasting five key histone modifications (H3K4me3, H3K27Ac, H3K9Ac, H3K36me3, H4K20me3) in near-isogenic male and female lines, and define 13 chromatin states across the *Ectocarpus* genome. We show that for individual genes, chromatin states differ dramatically between the two sexes, even if the relative proportions of the different chromatin states across the genome remain relatively stable. Males and females exhibit different histone posttranslational modifications at sex-biased genes, implying sex-specific mechanisms of chromatin-based gene regulation. Surprisingly, despite the fact that U and V are largely identical, recombining chromosomes with a very small sex-specific region, our results reveal not only a dramatically distinct chromatin landscape in the sex-chromosomes compared with autosomes but also distinct epigenetic profiles between the male and female sex-specific regions.

## 2.2 Introduction

Epigenetic chromatin modifications play an important role in determining patterns of gene expression in eukaryotes (Allis and Jenuwein, 2016). Groups of correlated histone modifications (so called chromatin states) are associated with functionally distinct regions of the genome such as heterochromatic regions and regions of either active transcription or repression (Kouzarides, 2007). Genome-wide maps of histone post-translational modifications provide the opportunity to decipher the complex combinatorial regulatory code of histone modifications and represents an effective means to understand the functional significance of chromatin remodelling in several developmental processes. Accordingly, chromatin profiling has increased our knowledge about epigenetic events associated with differentiation, cellular reprogramming and early development across several lineages, (e.g (Atlasi and Stunnenberg, 2017; Hota and Bruneau, 2016; Wu et al., 2018)). However, the relationships between sexual differentiation, gene expression and histone post translational modifications (PTMs) have not been characterised in detail, and very few studies have carried out genome-wide comparative analysis of chromatin states in males versus females (but see (Brown and Bachtrog, 2014), reviewed in (Ratnu et al., 2017)). Indeed, while genome-wide expression profiling studies across a broad range of taxa have often demonstrated extensive sex-biased gene expression underlying phenotypic sexual differentiation (Grath and Parsch, 2016), the role of chromatin structure on the regulation of sex-specific differences is still ill defined.

Autosomal sex biased gene expression explains a great proportion of expression differences between the sexes (reviewed in (Grath and Parsch, 2016)), but many species with genetic sex determination have largely differentiated sex chromosomes, and in organisms with XY or ZW sex determination systems, sex chromosomes often exhibit unusual, sex-specific, patterns of gene expression. For instance, in males of the fruit fly *Drosophila* the Y is transcriptionally repressed and the X is hyper-transcribed (Baker et al., 1994). Chromosome-wide transcriptional modifications are controlled by changes in chromatin states (Straub and Becker 2007, Lemos et al 2010, Steinemann and Steinemann 2005; Straub and Becker 2007) and chromatin states also play important roles in heterochromatin formation and dosage compensation. The Y chromosome has been suggested to be a heterochromatin sink in *Drosophila* (Brown and Bachtrog, 2014). In *Drosophila* the euchromatin/heterochromatin balance differs between the sexes because males have more transposable elements on the Y (Brown and Bachtrog, 2014), Yasuhara and Wakimoto 2008, Zhou et al. 2012). In schistosomes (which have ZW sex

chromosome systems), chromatin remodelling within the Z-specific region is associated with dosage compensation (Picard et al., 2019).

So far, no information is available concerning the dynamics of chromatin modifications in organisms with UV sexual systems, such as mosses and algae (Coelho et al., 2018). In these systems, sex is expressed during the haploid phase of the life cycle and depends on inheritance of U and V sex chromosomes in the spores, which determine whether the adult, multicellular individual is female or male, respectively (Bachtrog et al., 2014; Coelho et al., 2019). UV systems are fundamentally different from XY and ZW systems (Coelho et al., 2019; Umen and Coelho, 2019). For example, no dosage compensation has been described in UV systems, and Meiotic Sex Chromosome Inactivation (MSCI) is not expected to occur because meiosis takes place in the diploid sporophyte, where sex is not expressed. Moreover, whereas in diploid systems the Y and W accumulate repeats and show evidence for degenerative evolution, both U and V of haploid sexual systems appear to evolve symmetrically, having for instance similarly low gene density compared with autosomes and being overall relatively preserved despite their age (Ahmed et al., 2014). Heterochromatin is recruited to regions with high repeat densities, resulting in a correlation between repetitive DNA sequence and the tendency of a genomic region to adopt a heterochromatic appearance (Dorer and Henikoff, 1994; Lippman et al., 2004; Sentmanat and Elgin, 2012). Because the U and V sex chromosomes have similar repeat contents and non-recombining regions of comparable size (Ahmed et al., 2014), heterochromatin enrichment, relative to autosomes, is likely to be similar on both the U and V specific regions, and males and females are not expected to differ substantially in terms of chromatin environment.

Similarly to diploid systems, gene expression patterns of UV sex chromosomes differ from those of autosomes (reviewed in (Coelho et al., 2019)). In the latter, sex-linked genes are upregulated during the haploid, gametophyte phase of the life cycle (Ahmed et al., 2014; Lipinska et al., 2017). Remarkably, genes located on the pseudo-autosomal regions (PARs) of the U and V sex chromosomes exhibit unique gene expression patterns, being enriched in both life cycle-related genes (sporophyte-biased genes) and female-biased genes (Lipinska et al., 2015). Moreover, PAR genes display unusual structural features compared with autosomal genes in terms of their GC content, TE content and intron sizes (Luthringer et al., 2015b).

Here, we investigate the sex-specific chromatin landscape of *Ectocarpus*, a model brown alga with a UV sexual system and low levels of sexual dimorphism. We show that although the distribution of genome-wide chromatin marks is similar between males and females, a large



proportion of the genes is associated with different chromatin states in males and females. In particular, patterns of histone PTMs on genes that exhibit sex-biased expression differ substantially in males and females. Finally, our results reveal a dramatically distinct configuration of histone modifications associated with the U and V sex chromosomes compared with autosomes.

## 2.3 Results

### The chromatin landscape of *Ectocarpus* sp.

Near-isogenic male and female gametophyte lines (Table S1) were used to generate sex-specific ChIP-seq profiles for five different histone post-translational modifications (PTMs)(Figure S1): H3K4me3, H3K9ac, H3K27ac, H3K36me3 and H4K20me3. H3K4me3 is a near-universal chromatin modification that has been found at the transcription start sites (TSS) of active genes in a broad range of eukaryotes whose levels reflect the amount of transcription (Howe et al., 2017). H3K9ac is an active chromatin mark often associated with ongoing transcription (Heintzman et al., 2007). H3K27ac, an important mark that can distinguish between active and poised enhancer elements (Creyghton et al., 2010), H3K36me3, is a gene body mark associated with active gene transcription (Shilatifard, 2006). H4K20me3 is a repressive, promotor-localised constitutive heterochromatin mark (Schotta et al., 2004).

Given the large phylogenetic distance and the independent evolution of multicellularity in the brown algae (Cock et al., 2010), there is the possibility that these marks do not have the same functions in brown algae. However, a previous study in *Ectocarpus* has assayed six histone PTMs marks (Bourdareau, 2018) and has provided evidence for similar roles for at least four of the marks (H3K36me3, H3K27ac, H3K9ac, H3K4me3). Surprisingly, this study did not find evidence for tri-methylation of H3K27 in the *Ectocarpus* chromatin, therefore we chose to assay here tri-methylation of H3K20 as a putative repressive chromatin mark. Together, the five histone modifications used in our study are expected to provide a broad view of the *Ectocarpus* sp. genome chromatin landscape.

Figure 1 provides an overview of the chromatin states identified in *Ectocarpus* sp. and their distribution across four different classes of gene. Thirteen chromatin states (i.e., different combinatorial patterns of histone PTMs) were defined based analysis of the genome-wide distribution patterns of the five histone PTMs using MACS2 (Zhang et al., 2008) and SICER (Xu et al., 2014) (Figure 1A). States S9-S13 consisted of combinations of histone marks that

are usually associated with active transcription (H3K36me3, H3K27ac, H3K9ac, H3K4me3) (Bourdareau, 2018). States S2-S8 all included H4K20me3, which is generally associated with gene repression (Schotta et al., 2004), in most cases in addition to one or more of the above gene activation associated marks. State 1 corresponded to a ‘background’ state, i.e., domains that were not enriched for any of the histone PTMs assayed. An example of chromatin profiles across the *Ectocarpus* genome is given in Figure 1B.

### Chromatin states are correlated with patterns of gene expression

To correlate the observed chromatin states with gene expression patterns, RNA-seq data was generated using the same biological samples as were used for the ChIP-seq analysis (see methods). Four classes of genes were defined: transcribed genes (TPM $\geq$ 5th percentile), silent genes (TPM<5th percentile), housekeeping genes (defined as having values of less than 0.25 for the tissue specificity index tau (see methods) and tissue-specific genes (represented by basal-system specific genes, see methods).

The most common chromatin state for the actively transcribed genes (29.8% and 30.2% in males and females respectively) was S13, which corresponds to the co-localisation of all four of the histone PTMs that are generally associated with gene activation (H3K36me3, H3K27ac, H3K9ac, H3K4me3; Figure 1B, Table S2). For the ‘silent’ class of genes, S1 (no detectable histone PTM peak) was the most common state (45.4% and 43% in males and females, respectively; Figure 1B, Table S2). Finally, there was no clearly preferred state associated with the tissue-specific gene class, the most common state being S12 (H3K4me3, H3K9ac, H3K27ac; 20.9% and 18.3% in males and females, respectively; Figure 1B, Table S2).

When the relative proportions of the chromatin states were compared between males and females for each of the four gene classes (transcribed, silent, housekeeping and tissue-specific), similar patterns were observed in the two sexes, with some minor differences (Figure 1C). For example, less than 1% of the transcribed genes corresponded to state 8 (i.e., combination of H3K36me3 and H4K20me3) in females, compared to 2.4% in males (Table S2). Remarkably, while the proportion of genes associated with each chromatin state was similar in males and females, genome wide, the same genes were not always associated with the same chromatin state in the two sexes. Indeed, about 40% of the transcribed genes were associated with different states in males and females (Table S3). Taken together, these results indicate that, for individual genes, chromatin states can differ markedly between the two sexes, even if the relative proportions of the different chromatin states across the genome remain relatively stable.

## Identification of histone PTMs associated with gene activation and gene repression

To further investigate the relationship between the observed chromatin states and gene expression, transcript abundances in both males and females were plotted for the sets of genes corresponding to each chromatin state. A clear trend towards increasingly higher levels of transcript abundance was correlated with the gradual acquisition of the histone PTMs H3K9ac, H3K27ac, H3K4me3 and H3K36me3 (in the following order: H3K9ac followed by H3K9ac/H3K27ac, then by H3K9ac/H3K27ac/H3K4me3 and finally by H3K9ac/H3K27ac/H3K4me3/H3K36me3; Figure 2A; Table S3, S4). Again, these observations support the proposed association of these histone PTMs with gene activation (Bourdareau, 2018). In contrast, in pairwise comparisons, sets of genes corresponding to chromatin states that included H4K20me3 consistently exhibited lower transcript abundance than sets of genes with equivalent chromatin states without H4K20me3 (e.g. transcript abundance was significantly lower for S5 than for S13; Wilcox test, p-value= 4.463E-18 Figure 2A; Table S3, S4). These results are consistent with H4K20me3 playing a role in the repression of gene expression in *Ectocarpus*. Note that the background state S1 corresponds to domains that are not associated to any of the histone modifications assayed, and *Ectocarpus* genes associated to state S1 exhibited very low transcript abundance (Figure 2A, Table S3, S4).

Analysis of the RNA-seq data also indicated some minor differences between the sexes. For example, genes in chromatin state S1, S11 and S12 had significantly higher expression levels in females compared with males (pairwise Wilcoxon, p-value=2.4E-7; p-value=0.02 and p-value=0.001, respectively). Conversely, genes in chromatin state S2 and S3 had lower expression levels in females than in males (pairwise Wilcoxon, p-value=6.3E-8, p-value=3.4E-8; Figure 2A, Table S4, Table S5).

To further investigate the link between chromatin states and transcript abundances in males versus females, we classified states S1 and S2 as ‘repressive’ chromatin states (absence of marks and presence of H4K20me3), while states S9-S13 represent “active” chromatin states (presence of at least one canonical activation mark H3K9ac, H3K27ac, H3K4me3 and/or H3K36me3). Note that we did not include in this analysis genes belonging to states S3-S7 because they exhibit a combination of repressive (H4K20me3) and active marks and are expressed at intermediate levels (Figure 2A). As expected, genes that are in an ‘active’ chromatin environment in both sexes are expressed at higher levels than those that are associated to ‘repressive’ states (Figure 2B; pair-wise Wilcox test, p-value<2E-16).

Interestingly, levels of gene expression in males and females are also significantly different at genes that are in an ‘active’ chromatin environment in one sex, and a repression-associated state in the other (Figure 2B; pair-wise Wilcoxon test,  $p\text{-value} < 2E-16$ ). Thus, sex-specific differences in the chromatin state of genes are associated with their sex-specific expression patterns.

### Dynamic landscape of chromatin states on sex-biased genes

To investigate the role of chromatin modifications during sexual differentiation in *Ectocarpus*, we examined the chromatin states associated with genes that showed sex-biased expression patterns. A comparison of gene expression patterns in male versus female near-isogenic lines (Table S5), based on RNA-seq data generated using the same biological samples as were used for the ChIP-seq analysis, identified a total of 316 genes that exhibited a sex-biased pattern of expression ( $p_{\text{adj}} < 0.05$ , fold change  $> 2$ , TPM  $> 1$ ).

The presence of activation-associated chromatin marks H3K36me3, H3K9me3, H3K27ac and H3K4me3 were associated with significantly higher expression levels of sex-biased genes in males (Wilcoxon test,  $p\text{-value} = 0.0123$ ), but not in females (Wilcoxon test,  $p\text{-value} = 0.1882$ ; Figure S3A, B).

To analyse the chromatin modifications associated with sex-biased gene expression, transitions between chromatin states in males and females were evaluated on a gene-by-gene basis. This analysis showed that 57.3% of male-biased genes and 53.7% of female-biased genes had different chromatin states in males and females (Figures S2A; Table S6). These proportions were significantly higher than those observed for unbiased genes (40%; Fisher exact test,  $p\text{-value} = 6.24E-4$  and  $1.83E-4$  for males and females respectively). Sex-biased genes have low expression breadth (Lipinska et al., 2015), and in *Drosophila*, genes with low breadth of expression have unique chromatin patterns (Filion et al., 2010), so we also compared chromatin states in male and females for genes with low expression breadth (tissue-specific genes). Again, the proportions of sex-biased genes that changes states between males and females was significantly higher than those observed for tissue-specific genes (45%; Figures S2A, Table S6; Fisher exact test  $p\text{-value} = 1.32E-3$ ). Taken together these observations underscore the highly dynamic landscape of histone PTMs on sex-biased genes in males and females.

### Sex-specific chromatin reprogramming of sex-biased genes

For the set of male-biased genes there was a striking difference between the proportions of the chromatin states in males compared to females: in males, chromatin states that included the repression-associated mark H4K20me3 were rare, whereas states that included activation-

associated marks (H3K9ac, H3K27ac, H3K4me3 and/or H3K36me3, but not H4K20me3) were common: the proportion represented by chromatin states S2 to S7 decreased from 42% in females to 18.9% in males whilst the proportion represented by states S10 to S13 increased from 33.3% in females to 66.3% in males (Figure 3A, 3B; Table S2). Overall, almost half (44%) of the male-biased genes exhibited a transition from a state that lacked the repressive mark H4K20me3 (S9-S13) in males to a state that included the repressive mark H4K20me3 or a background state with none of the associated marks in females (S1-S8) (Figure 3C; Table S6, S7). These chromatin state modifications were consistent with a functional link between the presence and absence of activation-associated and repression-associated histone PTMs and differences in the abundances of the transcripts of sex-biased genes between sexes.

Surprisingly however, female-biased genes exhibit much weaker modifications of their chromatin states when males and females were compared. There was some tendency for chromatin states that included the repression-associated mark H4K20me3 to be less frequent in females and for states that included activation-associated marks (H3K9ac, H3K27ac, H3K4me3 and/or H3K36me3) to be common but the differences were much less marked than for the male-biased genes. (Figure 3A, 3B; Table S2). Only 14% of the female biased genes that are associated with active marks (S9-S13) in females presented a transition to a state associated to the repressive mark H4K20me3 in males (Figure 3C; Table S6, S7).

In *Drosophila*, the chromatin landscape of sex-biased genes is predominantly dictated by their low expression breadth and tissue specificity, and not particularly by their sex-biased expression patterns (Brown and Bachtrog, 2014). Comparison of chromatin states changes in males and females for tissue-specific genes revealed a very similar pattern to female-biased genes (but not to male-biased genes) with weak modifications in their chromatin states in males versus females (Table S6). Taken together, these results suggest that *Ectocarpus* sex-biased genes exhibit sex-specific chromatin reprogramming and that male-biased genes are specifically regulated in males versus females, whereas female-biased genes behave indistinctly to any other tissue-specific gene.

In conclusion, differences in expression levels of sex-biased genes between males and females correlate with changes in chromatin states and the latter are consistent with the proposed associations between specific histone PTMs and gene activation or repression. Surprisingly, however, sex-specific modifications to chromatin states were much more marked for male-biased than for female-biased genes.

## The chromatin landscape of the *Ectocarpus* sex-chromosomes

In organisms with diploid sexual systems (XY or ZW), the sex chromosomes have been shown to exhibit different patterns of histone PTMs compared with autosomes (Brown and Bachtrög, 2014; Picard et al., 2019). In *Drosophila* males, for instance, active chromatin is enriched on the X, relative to females, due to dosage compensation of the hemizygous X, whereas the inactivated mammalian X chromosome of females is characterized by DNA methylation and widespread repressive chromatin marks (Brockdorff and Turner, 2015; Lucchesi et al., 2005). On the chicken Z chromosome, a local and female-specific hyperacetylation of the fourth histone (H4K16Ac) has been described (Bisoni et al., 2005) and epigenetic analysis underlined the specialized chromatin landscape of the Z-specific region of *S. mansooni*, which is more permissive than that of the autosomal regions in both male and female *S. mansooni* (Picard et al., 2019).

A similar marked difference between sex chromosomes and autosomes was observed in *Ectocarpus* (Figure 4, Table S8, Figure S4, S5). Relative proportions of the chromatin states showed some variance between autosomes but genes on the sex chromosomes exhibited very different patterns to the autosomes (Figure 4A). The U and V sex chromosomes presented a significant dearth of genes in the activation-associated states S12 and S13 compared to autosomes (permutation tests U versus autosomes,  $p\text{-value}_{S12}=0.047$  and  $p\text{-value}_{S13}=0.039$ ; permutation tests V versus autosomes,  $p\text{-value}_{S12}=0.046$  and  $p\text{-value}_{S13}=0.037$ , respectively). Furthermore, the V-chromosome was significantly enriched in repression-associated states compared with autosomes, specifically state S2 ( $p\text{-value} = 0.025$ ), S4 ( $p\text{-value} = 0.021$ ), S5 (only in the PAR,  $p\text{-value} = 0.0078$ ) and S8 (on the SDR,  $p\text{-value} = 0.028$ ; Figure 4A, Table S10).

The vast majority of the genes within the female SDR (i.e., U-specific genes) were in state S1, i.e., exhibited none of the assayed marks, whereas V-specific genes were mostly in state S7 (deposit of all marks), state 8 (H3K36me3 and H4K20me3) with some genes in state 13 (all marks except H4K20me3) (Figure 4B-D; Table S9). However, the low number of SDR genes did not allow us to rule out that this result is due to chance (100 000 permutations tests on Pearson's  $X^2$  statistic).

## Chromatin states and gene expression levels on the sex chromosome

Gene expression levels and deposition of activation-associated chromatin marks were highly correlated for autosomal genes (see above, Figure 2A). Surprisingly, the effect of three

chromatin states (S7, S12, S13) on transcript abundance was significantly reduced when associated with female PAR genes (Table S11). Similarly, in the PAR of males, S7 and S13 were associated with a significantly lower increase in transcript abundance compared to their effect on autosomal gene expression (Table S11). Moreover, within the non-recombining SDRs, we found no correlation between levels of expression of SDR genes and deposit of chromatin marks (likelihood ratio tests, p-value = 0.460 and p-value = 0.304 for female and male SDR, respectively; Figure 4E).

Taken together, our observations suggest that the effect of chromatin states on transcript abundance are highly dependent on the genomic location of genes, and the same chromatin states do not correspond to the same level of transcriptional change in genes located in autosomes and sex chromosomes.

## 2.4 Discussion

Here, we have used five histone modifications marks to identify 13 biologically relevant active and repressed chromatin states in males and females of a brown alga with haploid UV sex determination, yielding the first characterisation of the sex-specific chromatin landscape in the third most developmentally complex eukaryotic lineage.

### Broad conservation of the histone code in the brown algae

Our analysis of the genome-wide distribution of several selected PTMs indicated that H3K4me3, H3K9ac, H3K36me3 and H3K27ac were associated with actively transcribed genes, whereas the canonical repressive mark H4K20me3 was associated with decreased gene expression. States associated with H3K36me3 were characteristic of broadly expressed genes, and this mark was less present in genes with narrow expression patterns. A similar H3K36me3 pattern has been described for *Drosophila* (Brown and Bachtrog, 2014; Filion et al., 2010), underscoring the conservation of this signature across distantly related lineages. Overall, therefore, our results are in line with the observed roles of chromatin remodelling and histone PTMs in plant and animal lineages (Baroux et al., 2011; Margueron and Reinberg, 2010; She and Baroux, 2015). These observations indicating that histone modification functions have a deep evolutionary origin and were probably present in the eukaryotic ancestor, supporting the idea that histone PTMs act as a conserved histone code mediating establishment and maintenance of epigenetic states.

### Widespread sex-specific chromatin reprogramming in *Ectocarpus* gametophytes

Although genome-wide the proportion of each of the chromatin states in males and females did not differ substantially, almost 40% of the *Ectocarpus* genes were in a different chromatin state when in a male or female context. This observation indicates that despite the low level of sexual dimorphism and relatively modest sex-biased gene expression (Lipinska et al., 2015; Luthringer et al., 2015a), *Ectocarpus* exhibits conspicuous sex-specific chromatin reprogramming.

The sex-specific differences in the chromatin state of genes manifested themselves as sex-specific expression patterns, with sex-biased genes, and specifically male-biased genes in *Ectocarpus* having radically distinct chromatin profiles compared with narrowly expressed genes. This observation suggests a specific regulation of sex-biased genes in males versus females, in contrast to *Drosophila* where the tissue specificity of sex-biased genes dominates chromatin landscape (Brown and Bachtrog, 2014).

Remarkably, chromatin remodelling of sex-biased genes in males was substantially more dynamic than in females, and this pattern is consistent with a similarly dynamic gene expression level modifications reported for sex-biased genes in males compared with females (Lipinska et al., 2015). The female developmental program may therefore be considered more like a ‘default’ program, involving less transcriptional modifications than the male developmental program which appears to be actively deployed. These results support previous evidence using genetic approaches suggesting that the female program is triggered by default in *Ectocarpus* in the absence of a male master sex determining factor(s) on the V-specific region (Ahmed et al., 2014).

### Unique chromatin organisation features in the U and V sex chromosomes

Sex-specific differences in chromatin states were exceptionally prominent on the U and V sex chromosomes, where about half (48%) of the genes had sex-specific chromatin states (Table S3). In *D. melanogaster*, a species with XY sexual system, epigenetic sex differences in regulation of gene expression have also been observed, and most sex-specific differences in the chromatin landscape are through to be a consequence of the sex chromosomes, due to dosage of the X and the presence of the Y in males (Brown and Bachtrog, 2014; Lemos et al., 2010; Yasuhara and Wakimoto, 2008; Zhou et al., 2012).

In organisms with UV sexual systems, the U and V sex-specific regions are both non-recombining and exhibit relatively similar features (Ahmed et al., 2014; Avia et al., 2018; Mignerot and Coelho, 2016). Consequently, epigenetic differences in male versus females on



the sex chromosomes are expected to be less obvious. Surprisingly, however, our data revealed a marked distinct pattern of histone PTMs in the male (V) versus female (U) SDRs. For example, H3K36me3, a mark involved in dosage compensation that is usually enriched in the X chromosome (Bell et al., 2008), is enriched on the V-specific, but totally absent from U-specific region. Moreover, our analysis of genome-wide chromatin patterns revealed conspicuous differences in the distribution of chromatin states in the PAR of the *Ectocarpus* sex chromosomes. Our observations suggest therefore that the effect of chromatin states on transcript abundance depend on genomic location of genes, and the same chromatin states do not correspond to the same level of transcriptional change in genes located in autosomes and sex chromosomes. Hence, gene expression in the U and V sex chromosome may be regulated by different epigenetic processes than autosomal gene expression, presumably involving histone PTMs that have not been assayed in this study. Further investigations employing more histone PTMs marks would be needed to further apprehend the extraordinary features of these chromosomes.

## 2.5 Methods

### Biological Material

Male (Ec457) and female (Ec460) *Ectocarpus* sp. lines (Table S1) were generated by crossing brother and sister gametophytes for five and six generations, respectively (Ahmed et al., 2014). Male and female strains had, therefore, essentially identical genetic backgrounds apart from the sex locus. The near-isogenic male and female gametophytes were cultured for 13 days as previously described (Coelho et al., 2012) at 13°C in autoclaved natural sea water supplemented with 300  $\mu$ l/L Provasoli solution (PES), with a light:dark cycle of 12:12 h (20  $\mu$ mol photons.m<sup>-2</sup>.s<sup>-1</sup>) using daylight-type fluorescent tubes. All manipulations were performed under a laminar flow hood in sterile conditions.

### Comparisons of male and female transcriptomes using RNA-seq

RNA-seq analysis was carried out to compare the relative abundances of gene transcripts between male and female samples. RNA for transcriptome analysis was extracted from the same duplicate male and female cultures as were used for the ChIP-seq analysis. For each sex, total RNA was extracted from a mix of 90 gametophytes each, using the Qiagen Mini kit (<http://www.qiagen.com>). RNA quality and quantity were assessed using an Agilent 2100

bioanalyzer, associated with Qubit2.0 Fluorometer using Qubit RNA BR assay kit (Invitrogen, Life Technologies, Carlsbad, CA, USA), as described previously (Lipinska et al., 2015, 2017). For each replicate sample, total RNA was quantified and cDNA was synthesized using an oligo-dT primer. The cDNA was fragmented, cloned, and sequenced by Fasteris (CH-1228 Plan-les-Ouates, Switzerland) using an Illumina HiSeq 4000 set to generate 150-bp single-end reads. Table S1 details the accession numbers for the sequenced data.

Data quality was assessed using fastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>; assessed May 2019). Reads were trimmed and filtered using cutadapt (Martin, 2011) with a quality threshold of 33 (quality-cutoff) and a minimal size of 30-bp.

Filtered reads were mapped to the *Ectocarpus* sp. genome (Cock et al., 2010; Cormier et al., 2017a) using TopHat2 with the Bowtie2 aligner (Kim et al., 2013). More than 85% of the sequencing reads from each library could be mapped to the genome (Table S1).

The mapped sequencing data were then processed with featureCounts (Liao et al., 2014) to obtain counts for sequencing reads mapped to genes. Gene expression levels were represented as transcripts per million (TPMs). Genes with expression values below the fifth percentile of all TPM values calculated per sample were considered not to be expressed and were removed from the analysis. This resulted in a total of 18,462 genes that were considered to be expressed.

Differential expression analysis was performed with the DESeq2 package (Bioconductor) (Love et al., 2014). Genes were considered to be male-biased or female-biased if they exhibited at least a twofold difference (FC) in expression between sexes with a false discovery rate (FDR) < 0.05. Full lists of sex-biased genes can be found in Table S5.

To calculate breadth of expression we employed the tissue-specificity index tau (Yanai et al., 2005) using published expression data from nine tissues or stages of the life cycle (female and male immature and mature gametophytes, mixed male and female gametophytes, parthenosporophytes, upright parthenosporophyte filaments, basal parthenosporophyte filaments, diploid sporophytes) from *Ectocarpus* sp. (Cormier et al., 2017a; Lipinska et al., 2015, 2017, 2019; Luthringer et al., 2015b). Tissue-specific genes were taken from basal system-specific genes detected in (Godfroy et al., 2017).

## Genome-wide detection of histone PTMs

Male versus female *Ectocarpus* sp. gametophyte ChIP-seq experiments were carried for H3K4me3, H3K9ac, H3K27ac, H3K36me3, H4K20me3, and three controls (an input control corresponding to sonicated DNA, histone H3 and immunoglobulin G monoclonal rabbit (IgG)) as in (Bourdareau, 2018). RNA-seq data (see above) was generated from the same samples, to ensure that the histone PTM and gene expression data were fully compatible. For ChIP-seq, 2.8 g (corresponding to 2800 individual gametophytes) of *Ectocarpus* tissue was fixed for five minutes in seawater containing 1% formaldehyde and the formaldehyde eliminated by rapid filtering followed by incubation in PBS containing 400 mM glycine. Nuclei were isolated by grinding in liquid nitrogen and in a Tenbroeck Potter in nuclei isolation buffer (0.1% triton X-100, 125 mM sorbitol, 20 mM potassium citrate, 30 mM MgCl<sub>2</sub>, 5 mM EDTA, 5 mM β-mercaptoethanol, 55 mM HEPES at pH 7.5 with complete ULTRA protease inhibitors), filtering through Miracloth and then washing the precipitated nuclei in nuclei isolation buffer with and then without triton X-100. Chromatin was fragmented by sonicating the purified nuclei in nuclei lysis buffer (10 mM EDTA, 1% SDS, 50 mM Tris-HCl at pH 8 with cComplete ULTRA protease inhibitors) in a Covaris M220 Focused-ultrasonicator (duty 25%, peak power 75, cycles/burst 200, duration 900 sec at 6°C). The chromatin was incubated with an anti-histone PTM antibody (anti-H4K20me3, anti-H3K4me3, and anti-H3K9ac, Cell Signal Technology; anti-H3K27ac, Millipore; anti-H3K36me3, Abcam) overnight at 4°C and the immunoprecipitation carried out using Dynabeads protein A and Dynabeads protein G. Following immunoprecipitation and washing, a reverse cross-linking step was carried out by incubating for at least six hours at 65°C in 5 M NaCl and the samples were then digested with Proteinase K and RNase A. Purified DNA was analysed on an Illumina HiSeq 4000 platform with a single-end sequencing primer over 50 cycles. At least 20 million reads were generated for each immunoprecipitation. The ChIP-seq dataset has been deposited in the NCBI Gene Expression Omnibus database under the accession numbers described in Table S1.

Quality control of the sequence data was carried out using FASTQC (Andrews, 2016). Poor quality sequence was removed and the high quality sequence was trimmed with Cutadapt (Hansen et al., 2016; Martin, 2011). Reads smaller than 15 bp were eliminated. Illumina reads were mapped onto the *Ectocarpus* genome (Cormier et al., 2017b) using Bowtie (Langmead et al., 2009). Duplicates were removed by samtools markdup in the package samtools (v 1.9) (Li et al., 2009).

The analysis was carried out for two biological replicates for each PTM in both the male and female samples. Pearson correlation analysis of replicates was performed with multiBamSummary and then by plotCorrelation (v3.1.2 deepTools) (Ramirez et al., 2014). Replicate samples were strongly correlated (Pearson correlations higher than 0.92, Figure S7). To identify peaks and regions of chromatin mark enrichment, each data set, after combining data for biological replicates, was analyzed separately for male and female gametophyte. Peaks corresponding to regions enriched in H3K4me3, H3K9ac and H3K27ac were identified using the MACS2 (version 2.1.1) callpeak module (minimum FDR of 0.01) and refined with the MACS2 bdgpeakcall and bdgbroadcall modules (Zhang et al., 2008). The input used for this peak calling was input data from sonicated DNA.

H3K36me3 and H4K20me3 were analyzed using SICER (v1.1) (minimum FDR of 0.01) ((Xu et al., 2014; Zang et al., 2009) and a window size of 200 bp and a gap size of 400 pb. SICER is better adapted to identify disperse ChIP-enriched regions (i.e., broad domains or islands).

In order to determine the coverage of each of the histone marks, the signal was normalized using the Signal Extraction Scaling (SES) method (Diaz et al., 2012) with input from DNA sonication for H3K4me3, H3K9ac and H3K27ac and H3 input for H3K36me3 and H4K20me3.

For both the identified MACS2 peaks and SICER regions, the number of male reads was compared to the number of female reads after normalization using the total number of reads in common peaks (FRiP = Fraction of reads in peaks). This method takes into account any bias between male and female gametophytes in the percentage of reads in peaks/regions. The log<sub>2</sub> fold change (female/male) was then calculated using the R package DiffBind (Stark and Brown, 2011). Heatmaps, average tag graphs and coverage tracks were plotted using EaSeq software (Lerdrup et al., 2016).

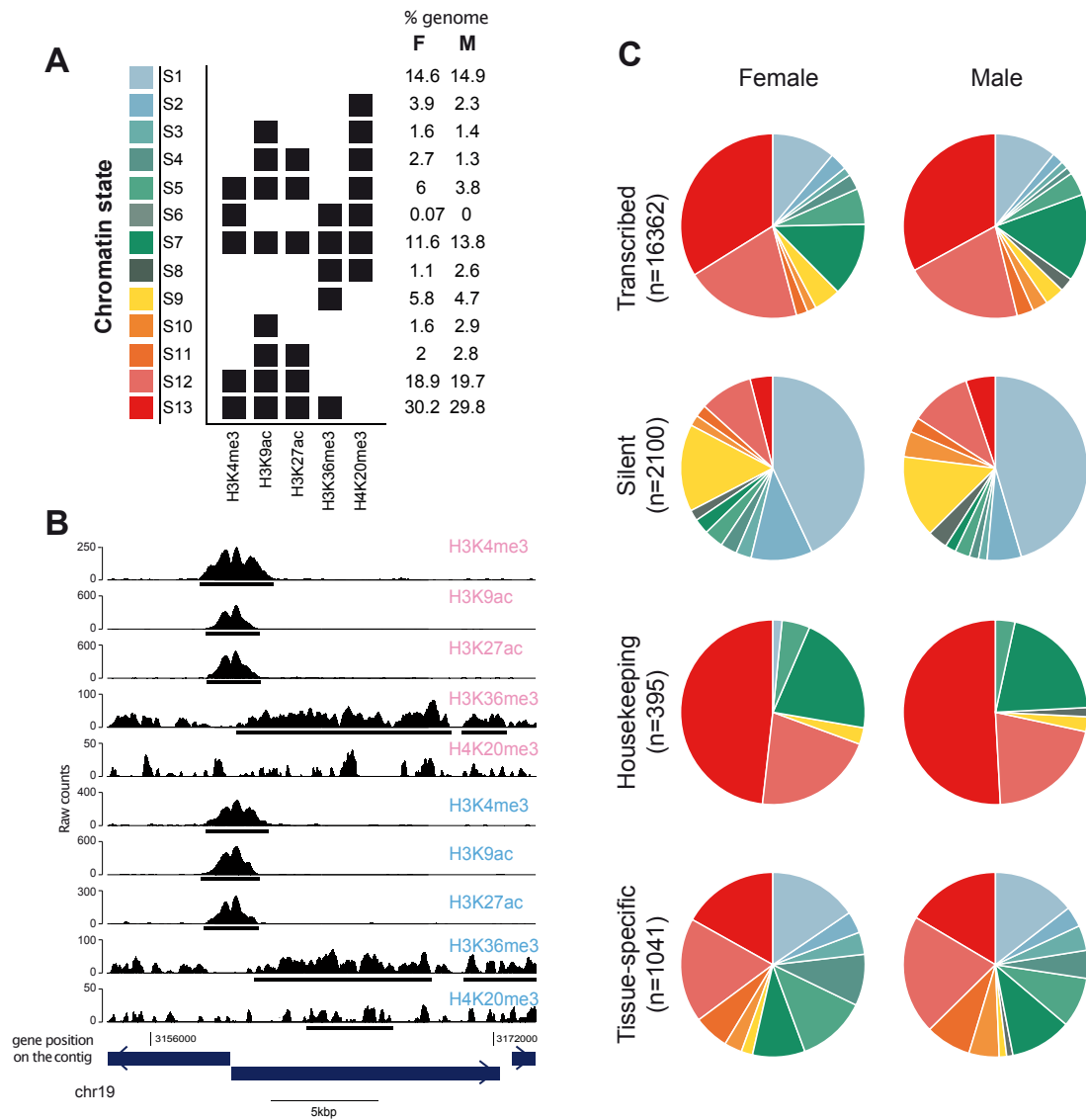
### Detection of chromatin states

To identify the patterns of histone PTM marks (or combination of PTMs) associated with each gene (i.e., chromatin states), we used bedtools intersect (aka intersectBed) in the Bedtools software (v2.26)(Quinlan and Hall, 2010). A total of 13 combinations of histone PTM marks (S1 to S13) were detected. Note that only chromatin states that are present in more than 1% of the genes were taken in consideration.

### Coverage of each chromatin state

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

Circos graphs were generated using Circos software (Krzywinski et al., 2009). To test for significant differences in the conservation of chromatin states between sex-biased and unbiased genes, we used mixed generalised linear models with a binomial distribution, modelling conserved vs non-conserved states as a function of bias. We then performed a likelihood ratio test with a null model to assess the significance of Bias. Statistical analysis was performed in R.

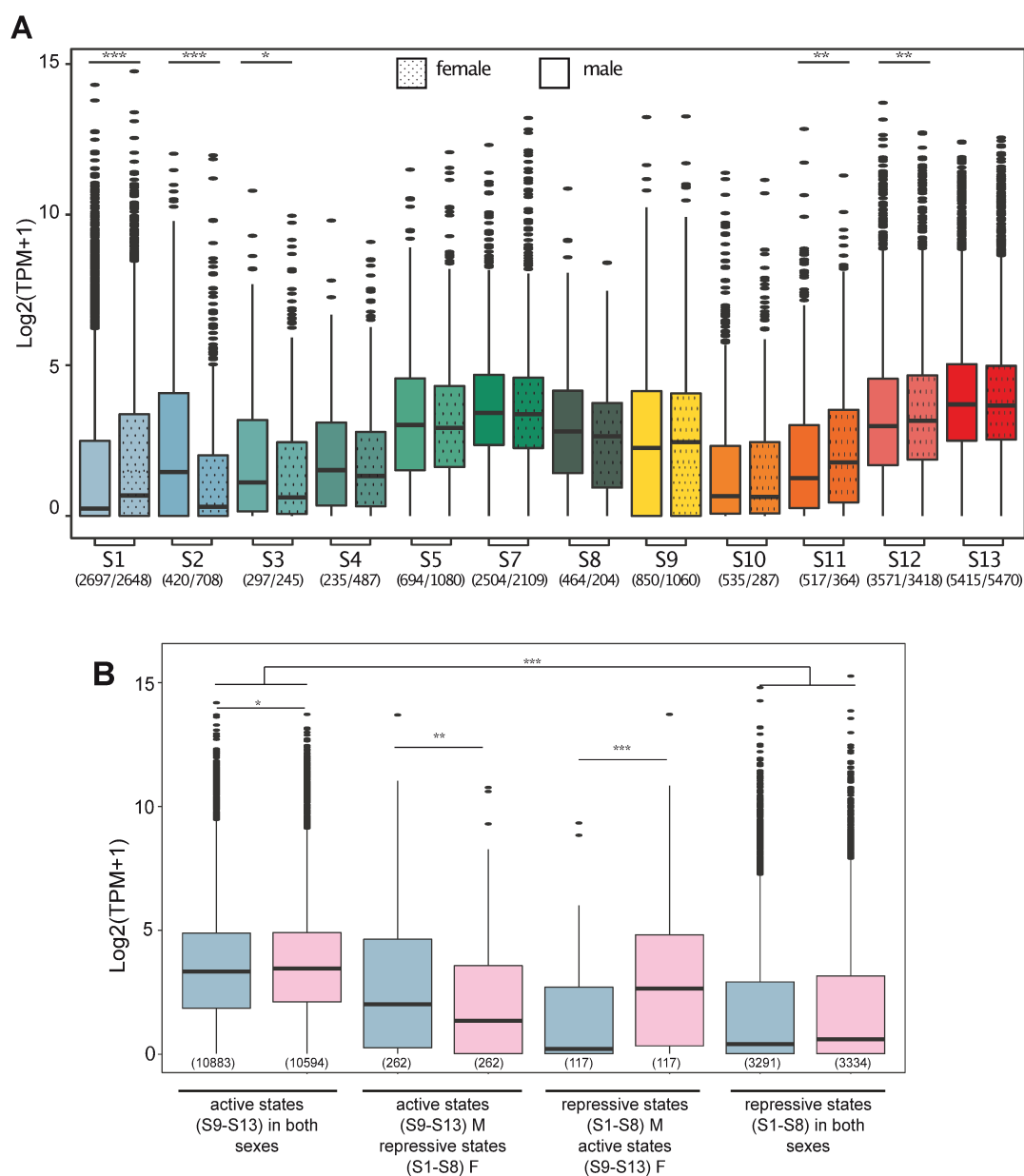


**Figure 1. The chromatin landscape of male and female *Ectocarpus* sp.**

A) Summary of the 13 chromatin states detected in *Ectocarpus* sp. Percentages of the total gene set associated with each chromatin state in males (M) and females (F) are shown to the right.

B) Representative region of the chromosome 19 showing profiles of mapped ChIP-seq reads for the five histone PTMs in males and females. Black bars represent the peaks detected by SICER and MACS2. Blue bars represent genes.

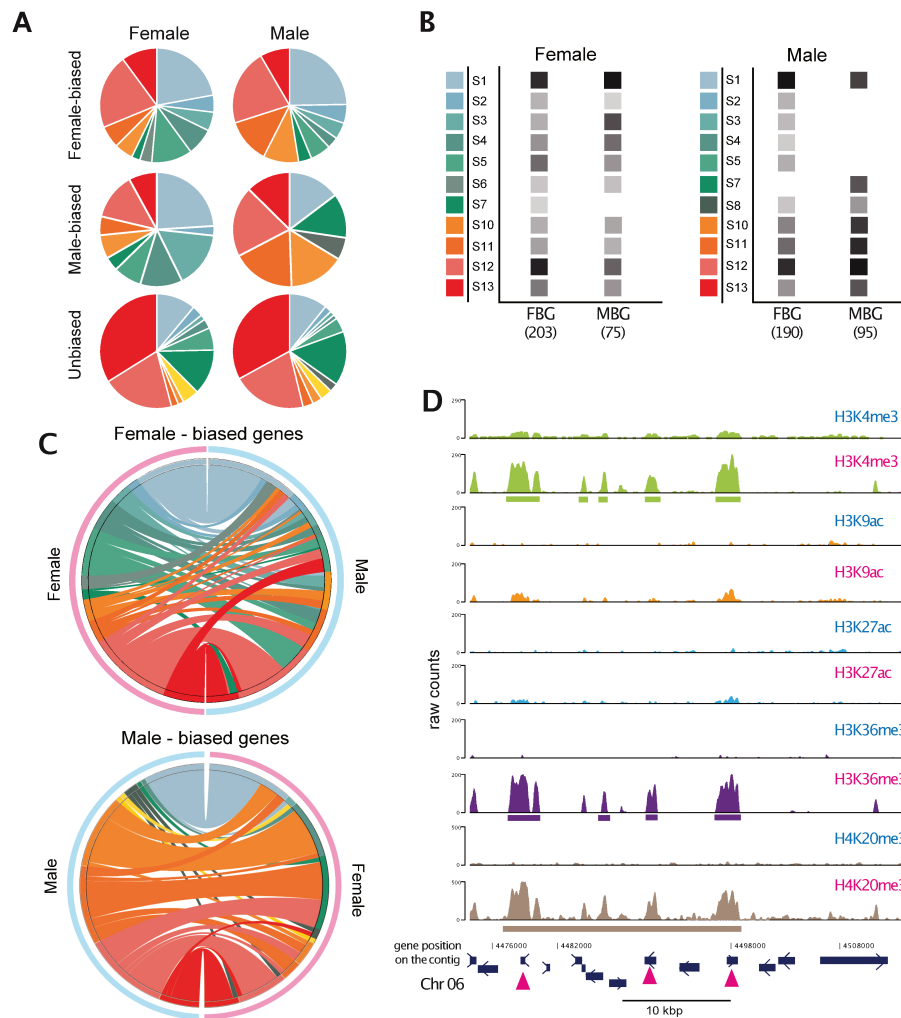
C) Proportions of each chromatin state associated with transcribed (TPM $\geq$ 5th percentile), silent (TPM<5th percentile), housekeeping and tissue (basal)-specific genes.



**Figure 2. Gene expression and chromatin states.**

A) Transcript abundance, measured as  $\log_2(\text{TPM}+1)$ , for genes associated with different chromatin states in males and females. The colour code is the same as that used in Figure 1A. Total number of genes associated with each state are indicated in brackets (males/females). Asterisks above plots represent significant differences in gene expression (pair-wise Wilcoxon test, \* $p$ -value $<0.05$ , \*\* $p$ -value $<0.01$ , \*\*\* $p$ -value $<0.001$ ).

B) Abundance of transcripts (measured as  $\log_2(\text{TPM}+1)$ ) of genes exhibiting either 'activation associated' (S9 to S13) or 'repression associated' (S1 to S8) chromatin states in females (pink) and males (blue). Numbers in brackets indicate the number of genes in each class. Asterisks indicate significant differences (\*,  $p>0.05$ , \*\* $p$ -value $<0.01$ ).



**Figure 3. Alterations of histone PTM patterns at sex-biased genes in *Ectocarpus* sp. males and females.**

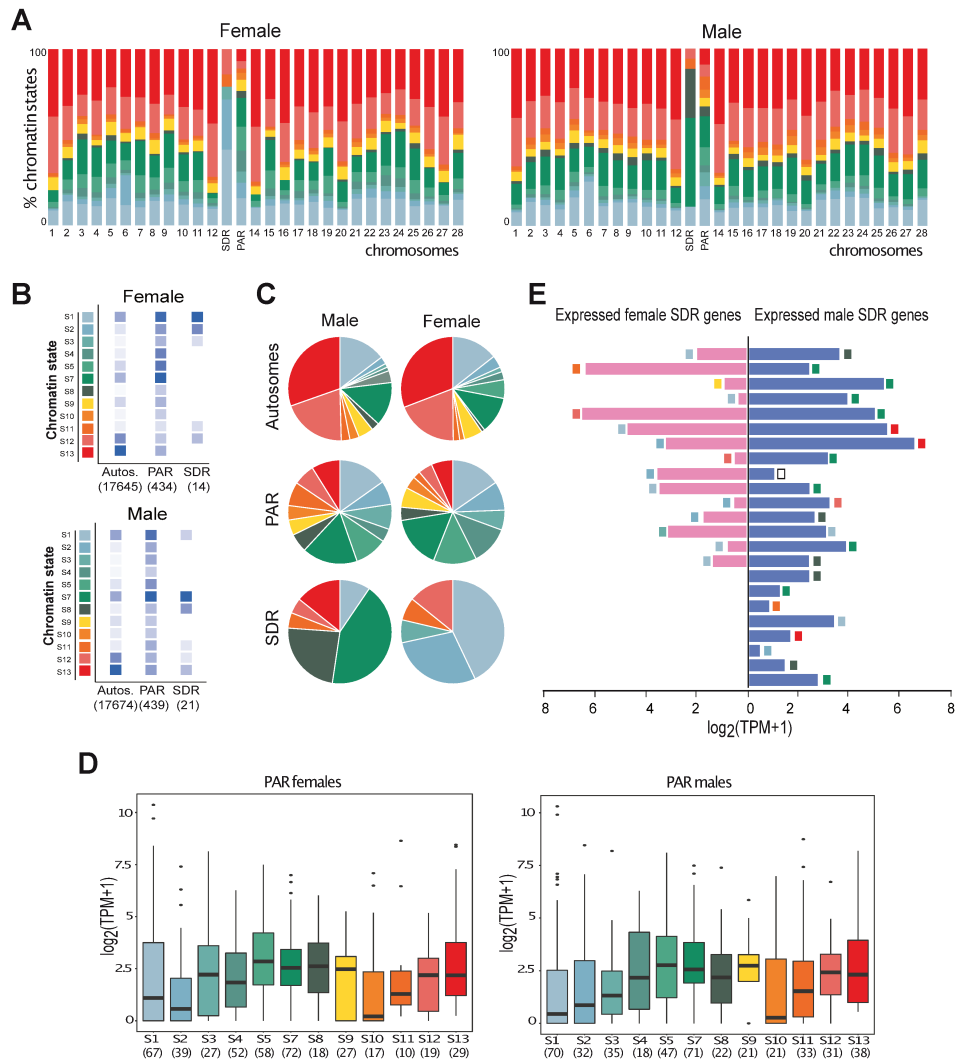
A) Distribution of chromatin states associated with female-biased, male-biased and unbiased genes in females (left) and males (right).

B) Relative proportions of chromatin states at sex-biased genes between in males and females. States corresponding to less than 1% of genes are not represented. The intensity of the grey gradient squares is proportional to the number of genes corresponding to each state. Coloured squares represent the different chromatin states. The total numbers of genes analysed for each condition is given in brackets. See Table S6 for the number of genes represented by each grey square. FBG: female-biased genes; MBG: male-biased genes.

C) Circos plots comparing chromatin states associated with female-biased (above) and male-biased (below) genes in chromatin from females (pink) and males (blue). The colour code for chromatin states on genes is the same as that used in Figure 1A. Each link corresponds to the transition from a state in the sex on the left of the circos plot to a state in the sex on the right of the circos plot. Chromatin states corresponding to less than 1% of the genes are not represented.

D) Representative chromatin profiles for female-biased genes (indicated by pink arrowheads) from chromosome 6 (EA-seq, (Lerdrup et al., 2016)). The histone PTMs indicated in blue correspond to those of the male, and in pink to those of the female *Ectocarpus*. The horizontal bars under each track correspond to peaks called by either MACS2 (H3K4me3, H3K9ac and H3K27ac) and SICER (H3K36me3 and H4K20me3).





**Figure 4. Chromatin landscape on the U and V sex chromosomes compared with autosomes.**

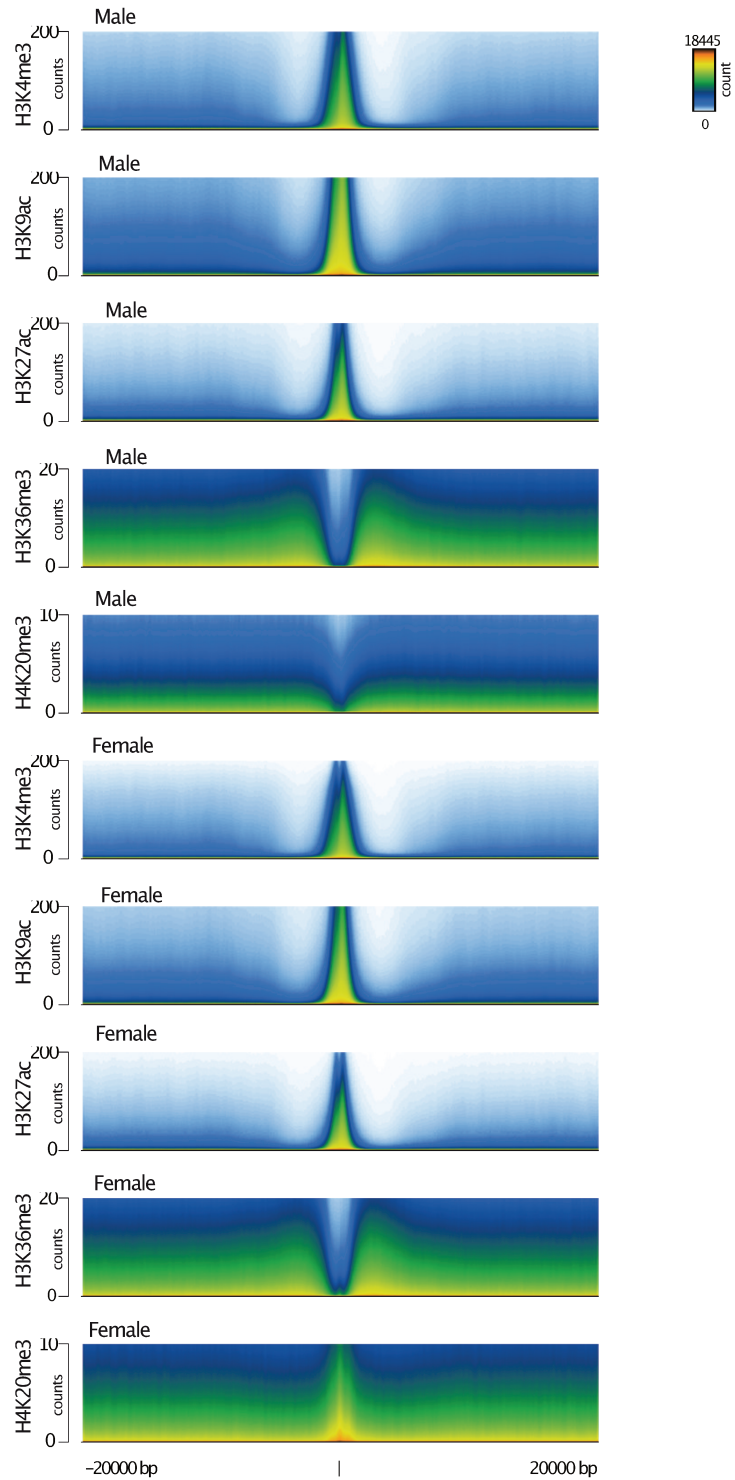
A) Percentage of genes associated with each of the 13 chromatin states for each autosome and for the SDR and PAR regions of the sex chromosome in female (left panel) and male (right panel) individuals.

B) Proportions of chromatin states in different chromosomes (autosomes) or chromosomal regions (SDR and PAR) in *Ectocarpus* female and male strains. States corresponding to less than 1% of genes are not represented. The intensity of the blue colour is proportional to the number of genes in each state (see also Table S8). The total number of genes in each genomic region are represented in brackets. Coloured squares represent the combinatorial patterns of histone PTMs. Autos., autosomes.

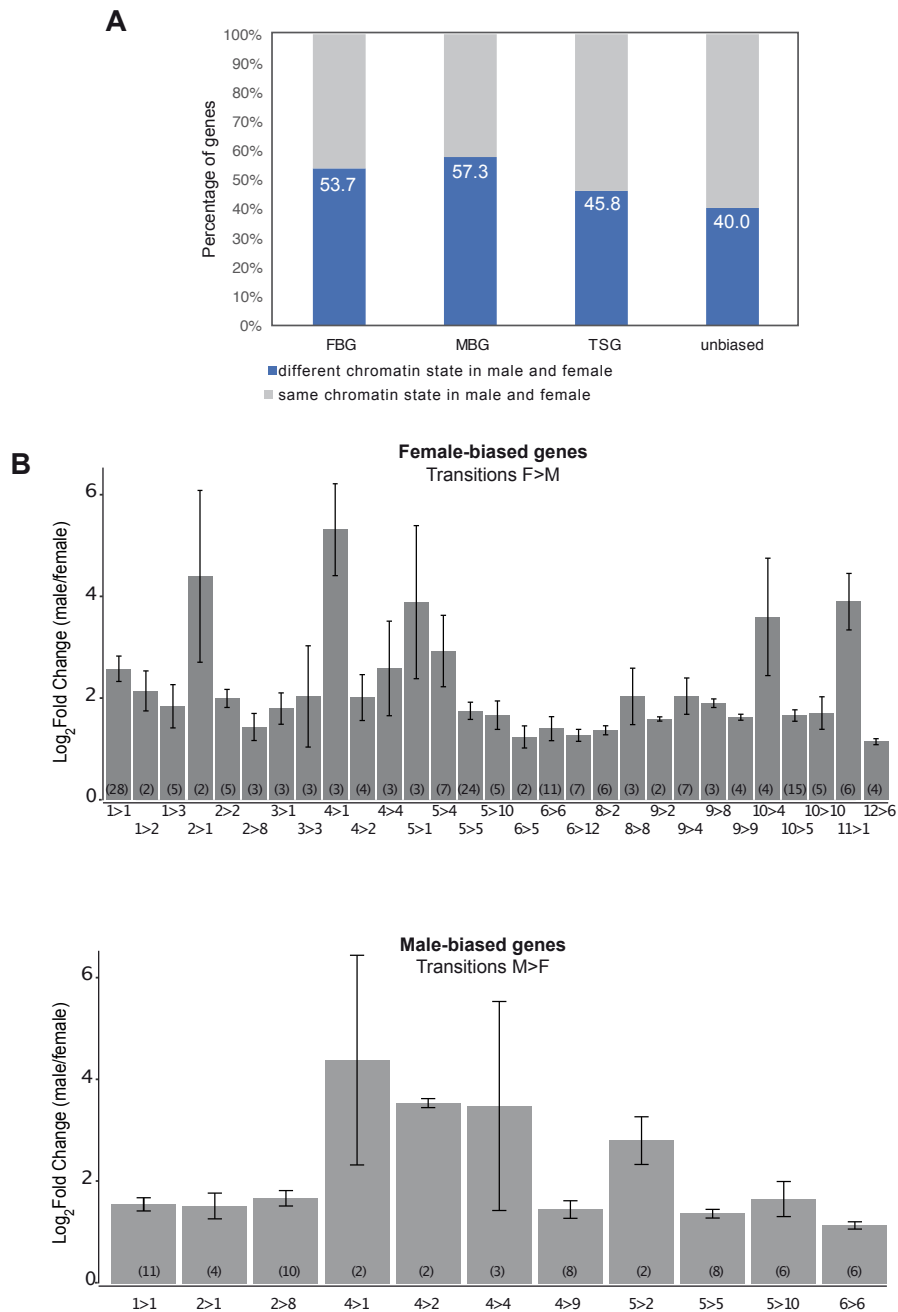
C) Proportions of chromatin states associated with autosomal, PAR and SDR genes in males and females. The color code used is the same as in Figure 1A.

D) Transcript abundance, measured as  $\log_2(\text{TPM}+1)$ , for genes associated with different chromatin states in the PAR of males and females. The colour code is the same as that used in Figure 1A. Total number of genes associated with each state are indicated in brackets.

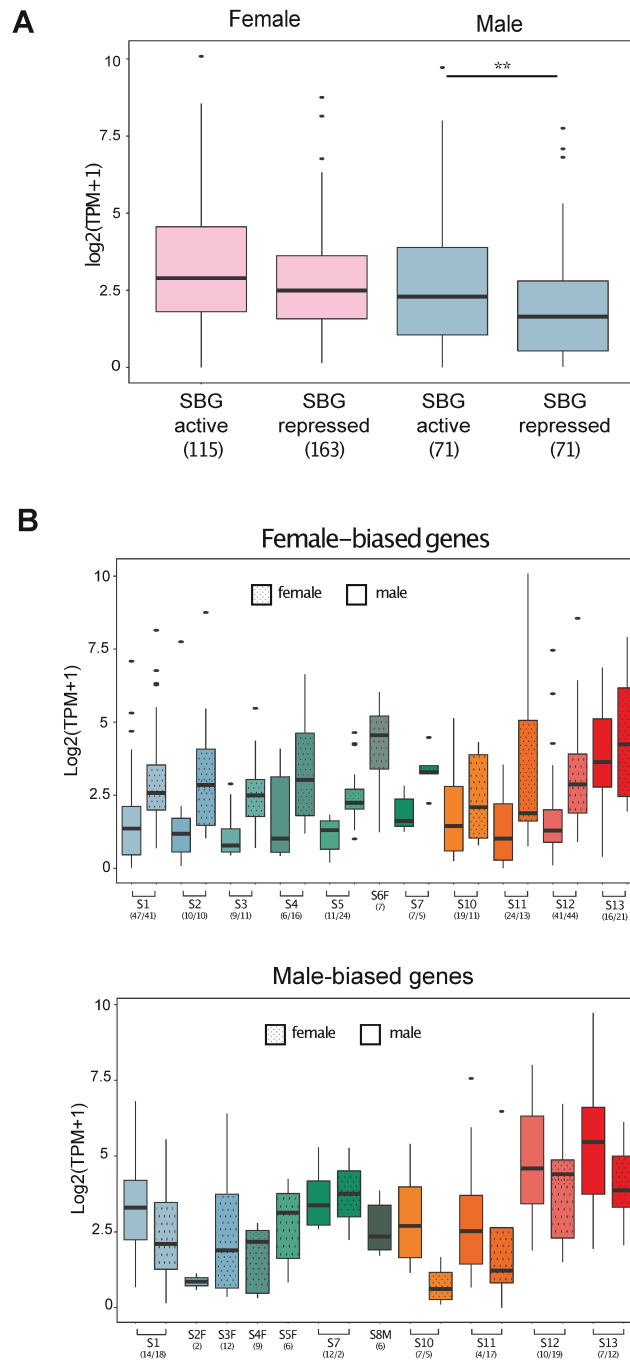
E) Expression levels measured as  $\log_2(\text{TPM}+1)$  of *Ectocarpus* genes located in the female (pink) and male (blue) SDRs. Coloured squares represent chromatin states corresponding to the colour code indicated in Figure 1A. Empty squares represents chromatin states represented in less than 1% of genes genome-wide (see also Table S4).



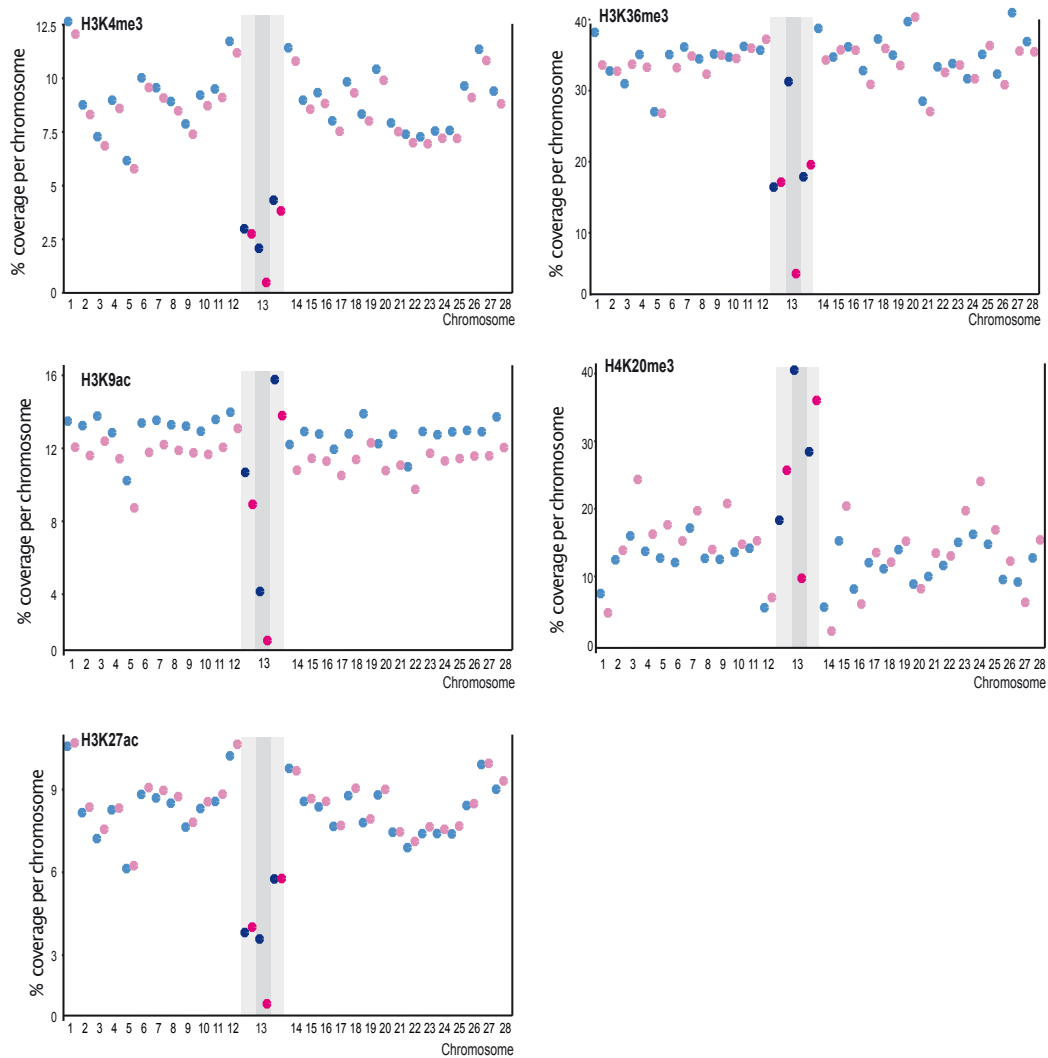
**Figure S1. Male and female enrichment profiles for the five histone PTMs centred on the transcription start sites (TSS, black bar).** The scale at the left of each volcano plot corresponds to the average tag count.



**Figure S2. Transitions between chromatin states in males and females. A)** Proportion of genes that exhibit changes in chromatin states when in a male or female context. MBG, male-biased genes; FBG, female-biased genes; TSG, tissue-specific genes. **B)** Median fold changes in abundances of sex-biased gene transcripts associated with transitions between chromatin states in males compared with females. The chromatin state transitions are schematized as “x > y” under each barplot, which corresponds to the transition from a state “x” to a state “y”. For the female-biased genes (upper panel), the state transition is from female to male. For the male-biased genes plot (down panel), the state transition is from male to female. The number of genes associated with each transition are indicated in brackets.



**Figure S3. Correlation between sex biased genes (SBG) states and sex biased genes expression in female and male.** A) Transcript abundance and chromatin states in males and females. ‘Active’ states correspond to S9-S13 and ‘repressive’ states to S1-S8. The left area (pink) corresponds to a female context and the right one (blue) to a male context. Values in brackets indicate number of analysed genes. Asterisks above the plots indicate significant differences (pair-wise Wilcoxon test, \*\*p-value<0.01). B) Transcript abundance, measured as log<sub>2</sub>(TPM+1), for genes associated with each of the different chromatin states in sex biased genes, in males and females. The colour code is the same as that used in Figure 1A. Total number of genes associated with each state are indicated in brackets.



**Figure S4: Percentage of genes marked with specific histone PTMs for the SDRs, PAR and autosomes in *Ectocarpus* male and females.** Scatter plot showing the chromosome coverage (percent of each chromosome) for the five histone PTM assayed (H3K4me3, H3K9ac, H3K27ac, H3K36me3, H4K20me3). Light blue and light pink represent coverage in male and female, respectively. Dark blue and red dots correspond to coverage in V and U sex chromosomes, respectively. Light shading indicates the two PARs and dark shading the non-recombining, sex specific region (SDR).

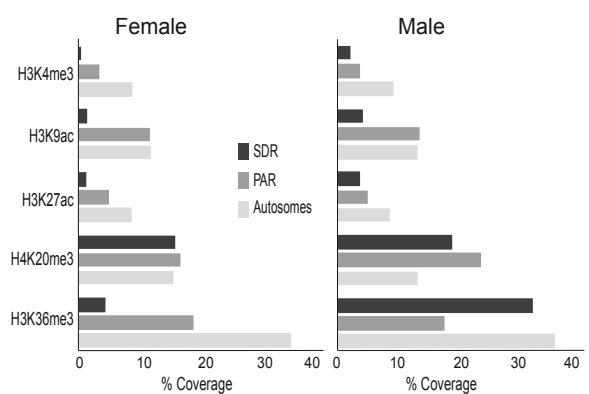


Figure S5. Percent coverage (normalised per number of genes) of each histone PTM across different genomic regions (PAR, SDR and autosomes) for female (left) and male (right) *Ectocarpus*.

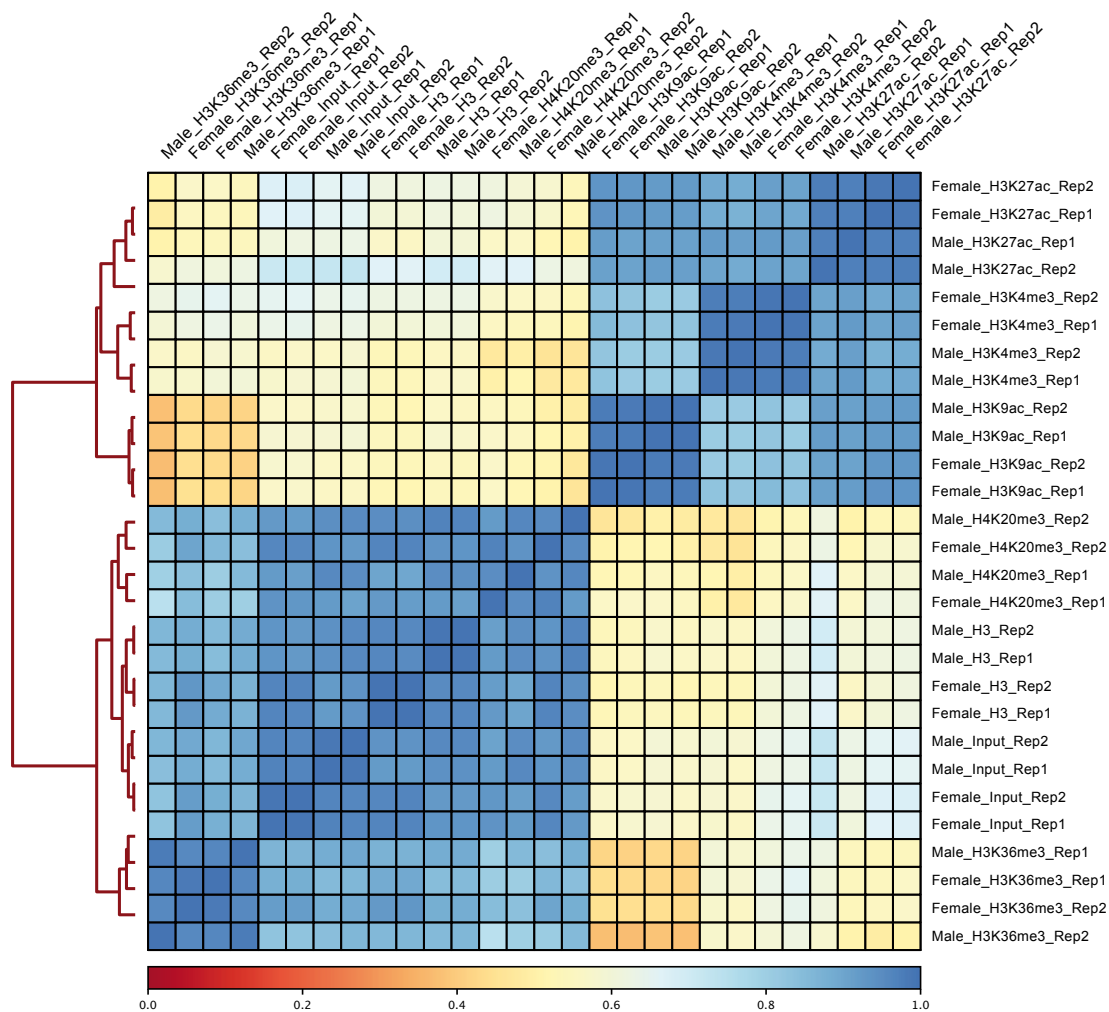


Figure S6. Pearson correlation scores for comparisons of the genomic distributions of ChIP-seq signal peaks for the five histone PTMs.

## 2.6 Supplemental Tables Legends

**Table S1. Strains used, database accessions numbers and RNA sequencing statistics**

**Table S2. Distribution of chromatin states across different types of genes in *Ectocarpus* male and female gametophytes.** The table presents the percentage of genes in each of the defined chromatin states. In blue the states associated with repression of transcription, in red the states associated with active transcription.; grey: state that corresponds to silent domains that are not enriched for any of the histone PTMs assayed. Transcribed genes correspond to genes with TPM>5th percentile. Silent genes correspond to genes with TPM<5th percentil. 'Global' indicates all genes in the genome. 'Basal' indicate genes specifically expressed in the basal systems (Lipinska et al., 2017).

**Table S3. Transcript abundances (measured as TPM) and associated chromatin states for all *Ectocarpus* genes in males and females.**

**Table S4. Statistical tests for the association between expression level and chromatin states.**

**Table S5. Sex-biased genes levels of expression (measured as TPM) in male and female samples.**

**Table S6. Transitions between chromatin states observed for male-biased, female-biased and tissue specific genes in comparisons of males and females.**

**Table S7. Chromatin states for sex biased genes for individual chromosomes in males and females.** Values indicate the number of sex-biased genes in a given chromosome that are associated with a specific combination of chromatin states.

**Table S8. Coverage of the five histone PTMs across male and female genomes.** The sex chromosome (chromosome 13) is divided in PAR1 (pseudo-autosomal region 1), SDR (sex-determining region) and PAR2 (pseudo-autosomal region 2).

**Table S9. Histone PTMs on SDR sex-specific genes (see also Figure 4D).**

**Table S10. Permutation tests performed to study chromatin states in different genomic regions.**

**Table S11. Statistical tests employed to investigate the correlation between transcript abundance and chromatin states in different genomic regions.**

## 2.7 Acknowledgements

We thank Swann Floc'hlay for advice on the Chip-seq analysis, Thomas Broquet for discussions on the statistical analysis and Maxim Bruto for help with the Circus visualisation. This work was supported by the CNRS, Sorbonne Université and an ERC grant to S.M.C. (638240).

## 2.8 References

- Ahmed S, Cock JM, Pessia E, Luthringer R, Cormier A, Robuchon M, Sterck L, Peters AF, Dittami SM, Corre E, Valero M, Aury J-M, Roze D, Van de Peer Y, Bothwell J, Marais GAB, Coelho SM. 2014. A haploid system of sex determination in the brown alga *Ectocarpus* sp. *Curr Biol* **24**:1945–1957. doi:10.1016/j.cub.2014.07.042
- Allis CD, Jenuwein T. 2016. The molecular hallmarks of epigenetic control. *Nature Reviews Genetics* **17**:487.
- Andrews S. 2016. FastQC A Quality Control tool for High Throughput Sequence Data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Atlasi Y, Stunnenberg HG. 2017. The interplay of epigenetic marks during stem cell differentiation and development. *Nat Rev Genet* **18**:643–658. doi:10.1038/nrg.2017.57
- Avia K, Lipinska AP, Mignerot L, Montecinos AE, Jamy M, Ahmed S, Valero M, Peters AF, Cock JM, Roze D, Coelho SM. 2018. Genetic Diversity in the UV Sex Chromosomes of the Brown Alga *Ectocarpus*. *Genes (Basel)* **9**. doi:10.3390/genes9060286
- Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L, Hahn MW, Kitano J, Mayrose I, Ming R, Perrin N, Ross L, Valenzuela N, Vamosi JC. 2014. Sex determination: why so many ways of doing it? *PLoS Biol* **12**:e1001899. doi:10.1371/journal.pbio.1001899
- Baker BS, Gorman M, Marin I. 1994. Dosage compensation in *Drosophila*. *Annu Rev Genet* **28**:491–521. doi:10.1146/annurev.ge.28.120194.002423
- Baroux C, Raissig MT, Grossniklaus U. 2011. Epigenetic regulation and reprogramming during gamete formation in plants. *Current Opinion in Genetics & Development* **21**:124–133. doi:10.1016/j.gde.2011.01.017
- Bell O, Conrad T, Kind J, Wirbelauer C, Akhtar A, Schubeler D. 2008. Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in *Drosophila melanogaster*. *Mol Cell Biol* **28**:3401–3409. doi:10.1128/MCB.00006-08
- Bisoni L, Batlle-Morera L, Bird AP, Suzuki M, McQueen HA. 2005. Female-specific hyperacetylation of histone H4 in the chicken Z chromosome. *Chromosome Res* **13**:205–214. doi:10.1007/s10577-005-1505-4
- Bourdareau S. 2018. Contrôle génétique et épigénétique des transitions du cycle de vie chez l'algue brune *Ectocarpus* sp.
- Brockdorff N, Turner BM. 2015. Dosage compensation in mammals. *Cold Spring Harb Perspect Biol* **7**:a019406. doi:10.1101/cshperspect.a019406
- Brown EJ, Bachtrog D. 2014. The chromatin landscape of *Drosophila*: comparisons between species, sexes, and chromosomes. *Genome Res* **24**:1125–1137. doi:10.1101/gr.172155.114



- Cock JM, Sterck L, Rouzé P, Scornet D, Allen AE, Amoutzias G, Anthouard V, Artiguenave F, Aury J-M, Badger JH, Beszteri B, Billiau K, Bonnet E, Bothwell JH, Bowler C, Boyen C, Brownlee C, Carrano CJ, Charrier B, Cho GY, Coelho SM, Collén J, Corre E, Da Silva C, Delage L, Delaroque N, Dittami SM, Doulebeau S, Elias M, Farnham G, Gachon CMM, Gschloessl B, Heesch S, Jabbari K, Jubin C, Kawai H, Kimura K, Kloareg B, Küpper FC, Lang D, Le Bail A, Leblanc C, Lerouge P, Lohr M, Lopez PJ, Martens C, Maumus F, Michel G, Miranda-Saavedra D, Morales J, Moreau H, Motomura T, Nagasato C, Napoli CA, Nelson DR, Nyvall-Collén P, Peters AF, Pommier C, Potin P, Poulain J, Quesneville H, Read B, Rensing SA, Ritter A, Rousvoal S, Samanta M, Samson G, Schroeder DC, Séguens B, Strittmatter M, Tonon T, Tregear JW, Valentin K, von Dassow P, Yamagishi T, Van de Peer Y, Wincker P. 2010. The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature* **465**:617–621. doi:10.1038/nature09016
- Coelho SM, Gueno J, Lipinska AP, Cock JM, Umen JG. 2018. UV Chromosomes and Haploid Sexual Systems. *Trends Plant Sci* **23**:794–807. doi:10.1016/j.tplants.2018.06.005
- Coelho SM, Mignerot L, Cock JM. 2019. Origin and evolution of sex-determination systems in the brown algae. *New Phytologist* **10.1111/nph.15694**. doi:10.1111/nph.15694
- Coelho SM, Scornet D, Rousvoal S, Peters NT, Dartevelle L, Peters AF, Cock JM. 2012. How to cultivate *Ectocarpus*. *Cold Spring Harb Protoc* **2012**:258–261. doi:10.1101/pdb.prot067934
- Cormier A, Avia K, Sterck L, Derrien T, Wucher V, Andres G, Monsoor M, Godfroy O, Lipinska A, Perrineau M-M, Van De Peer Y, Hitte C, Corre E, Coelho SM, Cock JM. 2017a. Re-annotation, improved large-scale assembly and establishment of a catalogue of noncoding loci for the genome of the model brown alga *Ectocarpus*. *New Phytol* **214**:219–232. doi:10.1111/nph.14321
- Cormier A, Avia K, Sterck L, Derrien T, Wucher V, Andres G, Monsoor M, Godfroy O, Lipinska A, Perrineau M-M, Van De Peer Y, Hitte C, Corre E, Coelho SM, Cock JM. 2017b. Re-annotation, improved large-scale assembly and establishment of a catalogue of noncoding loci for the genome of the model brown alga *Ectocarpus*. *New Phytol* **214**:219–232. doi:10.1111/nph.14321
- Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences* **107**:21931–21936. doi:10.1073/pnas.1016071107
- Diaz A, Park K, Lim DA, Song JS. 2012. Normalization, bias correction, and peak calling for ChIP-seq. *Stat Appl Genet Mol Biol* **11**:Article 9. doi:10.1515/1544-6115.1750
- Dorer DR, Henikoff S. 1994. Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* **77**:993–1002. doi:10.1016/0092-8674(94)90439-1
- Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, van Steensel B. 2010. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **143**:212–224. doi:10.1016/j.cell.2010.09.009
- Godfroy O, Uji T, Nagasato C, Lipinska AP, Scornet D, Peters AF, Avia K, Colin S, Laure M, Motomura T, Cock JM, Coelho SM. 2017. DISTAG/TBCCd1 Is Required for Basal Cell Fate Determination in *Ectocarpus*. *Plant Cell*. doi:10.1105/tpc.17.00440

- Grath S, Parsch J. 2016. Sex-Biased Gene Expression. *Annu Rev Genet* **50**:29–44. doi:10.1146/annurev-genet-120215-035429
- Hansen P, Hecht J, Ibn-Salem J, Menkuec BS, Roskosch S, Truss M, Robinson PN. 2016. Q-nexus: a comprehensive and efficient analysis pipeline designed for ChIP-nexus. *BMC Genomics* **17**:873. doi:10.1186/s12864-016-3164-6
- Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**:311–318. doi:10.1038/ng1966
- Hota SK, Bruneau BG. 2016. ATP-dependent chromatin remodeling during mammalian development. *Development* **143**:2882–2897. doi:10.1242/dev.128892
- Howe FS, Fischl H, Murray SC, Mellor J. 2017. Is H3K4me3 instructive for transcription activation? *Bioessays* **39**:1–12. doi:10.1002/bies.201600095
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**:R36. doi:10.1186/gb-2013-14-4-r36
- Kouzarides T. 2007. Chromatin modifications and their function. *Cell* **128**:693–705. doi:10.1016/j.cell.2007.02.005
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res* **19**:1639–1645. doi:10.1101/gr.092759.109
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**:R25. doi:10.1186/gb-2009-10-3-r25
- Lemos B, Branco AT, Hartl DL. 2010. Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict. *Proc Natl Acad Sci U S A* **107**:15826–15831. doi:10.1073/pnas.1010383107
- Lerdrup M, Johansen JV, Agrawal-Singh S, Hansen K. 2016. An interactive environment for agile analysis and visualization of ChIP-sequencing data. *Nat Struct Mol Biol* **23**:349–357. doi:10.1038/nsmb.3180
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**:2078–2079. doi:10.1093/bioinformatics/btp352
- Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**:923–930. doi:10.1093/bioinformatics/btt656
- Lipinska AP, Serrano-Serrano ML, Cormier A, Peters AF, Kogame K, Cock JM, Coelho SM. 2019. Rapid turnover of life-cycle-related genes in the brown alga. *Genome Biol* **20**:35. doi:10.1186/s13059-019-1630-6
- Lipinska AP, Toda NRT, Heesch S, Peters AF, Cock JM, Coelho SM. 2017. Multiple gene movements into and out of haploid sex chromosomes. *Genome Biol* **18**:104. doi:10.1186/s13059-017-1201-7
- Lipinska, Cormier A, Luthringer R, Peters AF, Corre E, Gachon CMM, Cock JM, Coelho SM. 2015. Sexual dimorphism and the evolution of sex-biased gene expression in the brown alga *ectocarpus*. *Mol Biol Evol* **32**:1581–1597. doi:10.1093/molbev/msv049

- Lippman Z, Gendrel A-V, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, Carrington JC, Doerge RW, Colot V, Martienssen R. 2004. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**:471–476. doi:10.1038/nature02651
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:550. doi:10.1186/s13059-014-0550-8
- Lucchesi JC, Kelly WG, Panning B. 2005. Chromatin remodeling in dosage compensation. *Annu Rev Genet* **39**:615–651. doi:10.1146/annurev.genet.39.073003.094210
- Luthringer R, Cormier A, Peters AF, Cock JM, Coelho SM. 2015a. Sexual dimorphism in the brown algae. *Perspectives in Phycology* **1**:11–25.
- Luthringer R, Lipinska AP, Roze D, Cormier A, Macaisne N, Peters AF, Cock JM, Coelho SM. 2015b. The Pseudoautosomal Regions of the U/V Sex Chromosomes of the Brown Alga *Ectocarpus* Exhibit Unusual Features. *Mol Biol Evol* **32**:2973–2985. doi:10.1093/molbev/msv173
- Margueron R, Reinberg D. 2010. Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet* **11**:285–296. doi:10.1038/nrg2752
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet* 10–12. doi:https://doi.org/10.14806/ej.17.1.200
- Mignerot L, Coelho SM. 2016. The origin and evolution of the sexes: Novel insights from a distant eukaryotic lineage. *C R Biol* **339**:252–257. doi:10.1016/j.crv.2016.04.012
- Picard MAL, Vicoso B, Roquis D, Bulla I, Augusto RC, Arancibia N, Grunau C, Boissier J, Cosseau C. 2019. Dosage Compensation throughout the *Schistosoma mansoni* Lifecycle: Specific Chromatin Landscape of the Z Chromosome. *Genome Biol Evol* **11**:1909–1922. doi:10.1093/gbe/evz133
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–842. doi:10.1093/bioinformatics/btq033
- Ramirez F, Dunder F, Diehl S, Gruning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**:W187–191. doi:10.1093/nar/gku365
- Ratnu VS, Emami MR, Bredy TW. 2017. Genetic and epigenetic factors underlying sex differences in the regulation of gene expression in the brain. *J Neurosci Res* **95**:301–310. doi:10.1002/jnr.23886
- Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**:1251–1262. doi:10.1101/gad.300704
- Sentmanat MF, Elgin SCR. 2012. Ectopic assembly of heterochromatin in *Drosophila melanogaster* triggered by transposable elements. *Proc Natl Acad Sci USA* **109**:14104–14109. doi:10.1073/pnas.1207036109
- She W, Baroux C. 2015. Chromatin dynamics in pollen mother cells underpin a common scenario at the somatic-to-reproductive fate transition of both the male and female lineages in *Arabidopsis*. *Front Plant Sci* **6**:294. doi:10.3389/fpls.2015.00294
- Shilatifard A. 2006. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* **75**:243–269. doi:10.1146/annurev.biochem.75.103004.142422
- Stark r, Brown g. 2011. DiffBind: differential binding analysis of ChIP-Seq peak data.
- Umen J, Coelho S. 2019. Algal Sex Determination and the Evolution of Anisogamy. *Annu Rev Microbiol*. doi:10.1146/annurev-micro-020518-120011

- Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li T, Shi S, Zhang N, Duan F, Ming J, Zhang X, Niu W, Song W, Jin H, Guo Y, Dai S, Hu L, Fang L, Wang Q, Li Y, Li W, Na J, Xie W, Sun Y. 2018. Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* **557**:256–260. doi:10.1038/s41586-018-0080-8
- Xu S, Grullon S, Ge K, Peng W. 2014. Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. *Methods Mol Biol* **1150**:97–111. doi:10.1007/978-1-4939-0512-6\_5
- Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safran M, Domany E, Lancet D, Shmueli O. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**:650–659. doi:10.1093/bioinformatics/bti042
- Yasuhara JC, Wakimoto BT. 2008. Molecular landscape of modified histones in *Drosophila* heterochromatic genes and euchromatin-heterochromatin transition zones. *PLoS Genet* **4**:e16. doi:10.1371/journal.pgen.0040016
- Zang C, Schones DE, Zeng C, Cui K, Zhao K, Peng W. 2009. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**:1952–1958. doi:10.1093/bioinformatics/btp340
- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**:R137. doi:10.1186/gb-2008-9-9-r137
- Zhou J, Sackton TB, Martinsen L, Lemos B, Eickbush TH, Hartl DL. 2012. Y chromosome mediates ribosomal DNA silencing and modulates the chromatin state in *Drosophila*. *Proc Natl Acad Sci U S A* **109**:9941–9946. doi:10.1073/pnas.1207367109

### 3. Conclusion

The expression of sex-biased genes must be regulated in both male and female organisms to allow a single pathway of sexual differentiation to be determined. Thus, in males, the male sex differentiation pathway must be activated and the female differentiation pathway must be repressed, and vice versa in females. This regulation involves epigenetic mechanisms. A disruption of this regulation could lead to sex-reversion, as has been observed in mammals. For example, mutations that affect proteins *Cbx2* and *Jmjd1a*, which are involved in the regulation of the chromatin dynamics, result in female phenotypes in genetically male individuals (Kato-Fukui et al., 1998; Kuroki et al., 2013). Recently, a male mutant of the brown alga *Macrocystis* was shown to exhibit female characteristics. A study of sex-biased gene expression in this mutant, compared to the patterns in wild type male and female individuals, would provide a better understanding of the mechanisms that govern sexual differentiation in brown algae.



## Chapter 4

A partially sex-reversed giant kelp  
sheds light into the mechanisms of  
sexual differentiation in the brown  
alga

## 1. Introduction

Across vertebrates, antagonism between core male- and female-promoting gene networks is now recognized as crucial to the establishment and maintenance of gonadal fate (Capel, 2017). The discovery of revertant mutants for sex made it possible to highlight the balance between the male and female differentiation pathways. The differentiation pathway of one sex must be active while the one governing the other must be suppressed, in order to maintain gender integrity. For example, in mice, a loss of function of the FOXL2 gene, involved in female differentiation, leads to a transdifferentiation of mature ovaries into testes (Uhlenhaut et al., 2009). Inversely, testicular cells will become ovarian cells if the masculinizing transcription factor DMRT1 is lost (Matson et al., 2011), suggesting that gonadal bipotentiality presents a mechanism through which female-male and male-female gonadal sex reversal may be regulated.

Although the literature is very extensive on the mechanisms that control and maintain sexual differentiation in animals and plants (Bachtrog et al., 2014; Pannell, 2017), very little knowledge is currently available on the mechanisms that control sexual differentiation in brown algae. A feminised male variant line has recently been discovered in the giant kelp *Macrocystis pyrifera*, a brown alga that has a UV sex determination system. Thus, this study is helping us to understand the role of sex biased genes expression in the initiation of the male and female developmental programs. The feminised male strain was first characterized phenotypically, both in terms of its morphology, its ability to produce fertile gametes, or its ability to produce female pheromones, and then genetically in terms of overall gene expression and sex biased genes. The availability of this giant kelp feminised line, together with near-isogenic male and female lines, has provided a means to access transcriptional events underlying the initiation of the male and female sexual differentiation programmes in this ecologically important organism and to disentangle the role of sex-linked genes and autosomal gene expression in the initiation of the male and female developmental programs. My participated in this study included bioinformatics analysis and phenotypic characterisation of the male, female and variant lines.



## 2. A partially sex-reversed giant kelp sheds light into the mechanisms of sexual differentiation in the brown algae

Dieter Muller<sup>1</sup>, Enora Gaschet<sup>1</sup>, Olivier Godfroy<sup>1</sup>, Josselin Gueno<sup>1</sup>, Agnieszka P. Lipinska<sup>1</sup>, Guillaume Cossard<sup>1</sup>, Maritta Kunert<sup>2</sup>, Akira F. Peters<sup>3</sup>, Renato Westermeier<sup>4</sup>, Wilhelm Boland<sup>2</sup>, J. Mark Cock<sup>1</sup>, Susana M. Coelho<sup>1\*</sup>

<sup>1</sup>Sorbonne Université, UPMC Univ Paris 06, CNRS, Algal Genetics Group, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, F-29688, Roscoff, France. <sup>2</sup>Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology, Jena, Germany. <sup>3</sup>Bezhin Rosko, 29250 Santec, France. <sup>4</sup>Universidad Austral de Chile, Instituto de Acuicultura, Valdivia, Chile.

\*correspondence: [coelho@sb-roscoff.fr](mailto:coelho@sb-roscoff.fr)

### 2.1 Summary

The giant kelp *Macrocystis pyrifera* possesses a UV sex determination system that is analogous to the one found in the model brown alga *Ectocarpus* and mosses. In UV sexual systems, sex is determined during the haploid phase of the life cycle and males have a V chromosome whereas females have a U chromosome. Previous work in the model *Ectocarpus* revealed that the brown algal V chromosome has a dominant role in male sex determination and the female developmental program is triggered in the absence of the male master sex determination gene(s). Here, we describe the identification of a genetically male giant kelp strain presenting phenotypic features typical of a female, despite lacking the U-specific region. The conversion to the female developmental program is however incomplete, because gametes of this feminised male are unable to produce the sperm-attracting pheromone lamoxiren, and consequently are not fully functional. We identify the transcriptomic pathways underlying the male and female specific developmental programs and show that the phenotypic feminisation of the variant strain is associated with both feminisation and de-masculinisation of gene expression patterns. In particular, a candidate sex-determining gene on the V-specific region was significantly downregulated in the feminised strain. Our results reveal the molecular changes associated with sexual differentiation in a UV system with strong gamete dimorphism, disentangling the role of sex-linked genes and autosomal gene expression in the initiation of the male and female

developmental programs. Overall, the data presented here imply that the U-specific region in the giant kelp is not required to initiate the female developmental program, however, this region appears to be critical to produce a fully functional female, arguing against female being a fully ‘default’ sex in this species.

## 2.2 Introduction

Females and males often differ dramatically in appearance and behaviour. These differences, which are referred to as sexual dimorphism, are principally the result of natural and/or sexual selection for traits that influence the fitness of each sex. Genetically, however, females and males are nearly identical differing by only a few genes located on sex-specific chromosomes (such as Y chromosomes in mammals, W chromosomes in birds, or U and V chromosomes in mosses and many algae (Coelho et al., 2018; Umen and Coelho, 2019)). Consequently, sexually dimorphic traits are a result of differential expression of genes that are present in both sexes (Grath and Parsch, 2016).

Autosomal sex biased gene expression patterns in diploid XY and ZW systems have been intensively investigated in recent years (reviewed in (Grath and Parsch, 2016)). Sex-biased genes, especially genes with male-biased expression in XY systems, often show elevated rates of both protein sequence and gene expression divergence between species, which could have a number of causes, including sexual selection, sexual antagonism, and relaxed selective constraint (Grath and Parsch, 2016; Whittle and Extavour, 2019). In contrast, knowledge about sex-biased gene expression, sexual dimorphism and the link between sex-chromosomes and control of sex-biased gene expression in UV haploid sexual systems remains relatively scarce. In the brown alga *Ectocarpus*, which has a UV sexual system, less than 10% of the transcriptome is sex-biased, and both male and female sex biased genes evolve faster than unbiased genes, (Lipinska et al., 2015b). Contrasting studies of diverse sexual species can be informative regarding the influence of the sex-determination system and the reproductive mode on expression differences between the sexes.

In UV systems, sex is determined by a male or a female sex-determining region (SDR). Sex is determined at meiosis: if a spore inherits the U-chromosome (containing a U-specific region), it will develop into a female gametophyte that at maturity will produce female structures (oogonia) and female gametes (eggs). If the spore inherits a V-chromosome, it will develop into a male gametophyte, producing male reproductive structures (antheridea), where male gamete cells are produced by mitosis at maturity. Sexual dimorphism can be minor, as in *Ectocarpus*

where male and females differ very little, but sexual dimorphism may be relatively extensive in algae such as kelps (Luthringer et al., 2015).

The SDRs of UV systems contain sex-determining gene(s) that initiate the male/female program, by regulating sex-specific patterns of expression of downstream effector autosomal genes. In *Ectocarpus*, genetic analysis has shown that the male V-specific region is dominant over the female U-specific region, and that female developmental program is expressed in the absence of the VSR (and presence of the U-specific region)(Ahmed et al., 2014; Müller, 1975). These results suggest that a master male sex-determining gene(s) is located on the V-specific region (Ahmed et al., 2014; Lipinska et al., 2017) and that female sex is initiated by default, in absence of the VSR (Ahmed et al., 2014). The female developmental program would therefore mainly rely on autosomal gene expression, whose patterns would be triggered in the absence of the male master sex-determining gene(s). If this is the case, the U specific region may not be strictly necessary for the female program to be initiated, and silencing the male master sex-determining gene in a male individual would be sufficient to initiate the female program. Following this reasoning, a mutant or a strain whose male master SD gene(s) is impaired would be expected to develop into a phenotypic female.

We describe here the identification of a variant strain of the giant kelp *Macrocystis pyrifera* that exhibits a range of morphological features typical of a female, despite being genetically male. The availability of this line, together with near-isogenic male and female lines, provided a means to access the molecular events underlying the initiation of the male versus female programme in this ecologically important organism and to disentangle the role of sex-linked genes and autosomal gene expression in the initiation of sex-specific developmental programs. *M. pyrifera* and *Ectocarpus* sex chromosomes are derived from a common ancestral sex chromosome and share a candidate sex-determining gene (Lipinska et al., 2017). We show that a considerable proportion of the transcriptome of the giant kelp is sex-biased, and that the feminisation of the variant line is correlated with extensive feminisation and demasculinization of gene expression patterns. Two genes within the VSR (Lipinska et al., 2017) exhibit a significantly reduction in transcript abundance in the feminised line, suggesting their likely involvement in sexual differentiation. Contrary to the idea that the female program is initiated entirely by default, the variant strain is not a fully functional female suggesting that the U specific region in this organism is required to fully express the female developmental program. Our results provide an illustration of how male and female-specific developmental programs may, at least partially, be uncoupled from sex chromosome identity.

## 2.3 Results

### Identification of a male *M. pyrifera* line exhibiting female-like morphology

In *M. pyrifera*, sex is determined genetically during the (haploid) gametophyte generation. The presence of shared orthologues in the SDRs of the Ectocarpales and the giant kelp suggests that the U and V chromosomes in these organisms are derived from the same ancestral autosome (Lipinska et al., 2017). Meiosis occurs during the sporophyte generation, producing haploid spores that develop into either male or female gametophytes. Although sexual dimorphism in brown algae is less conspicuous than in animals, male and female gametophytes in the giant kelp exhibit sexual dimorphism in terms of size of vegetative cells and reproductive structures (Lipinska et al., 2015a; Westermeier et al., 2007). Male gametophytes produce small male gametes (sperm) in specialised sexual structures called antheridia, whereas female gametophytes produce large female gametes (eggs) in oogonia (Figure 1).

During a screen of colchicine treatment of *M. pyrifera* gametophytes, we identified a genetically male line (Mpyr-13-4) that exhibited morphological features resembling those of female gametophytes (Figure 1). To determine whether the observed phenotypes were due to polyploidy or aneuploidy induced by the colchicine treatment, we measured the ploidy level of Mpyr-13-4 using flow cytometry. This analysis showed no evidence for chromosome doubling (Figure 1A).

### Phenotypic characterisation of the feminized *M. pyrifera* line

A detailed morphometric study of gametophytes of the Mpyr-13-4 strain showed the cells of this strain were intermediate in size between those of cells of male and female wild-type strains, i.e., the cells were significantly larger than those of the wild-type male strain (Wilcoxon test  $W = 31$ ,  $p\text{-value} = 6.27e-10$ ) (Figure 1B, S1).

Following induction of gametogenesis, Mpyr-13-4 gametophytes formed reproductive structures (Figure 1C-F) that strongly resembled wild-type female oogonia (Figure 1G-H) and not male antheridia (Figure 1I-J). At maturity, Mpyr-13-4 gametophytes produced egg-like cells, lacking flagella, with diameters ranging from 10-40  $\mu\text{m}$ . Wild-type eggs are typically around 40  $\mu\text{m}$  in diameter (Figure 1K-L).

In absence of fertilisation by gametes of the opposite sex, wild-type female gametes of *M. pyrifera* initiate parthenogenetic development within 48h. Wild type male gametes do not

undergo parthenogenesis. Mpyr-13-4 gametes, regardless of their size, initiated parthenogenetic development within 48 hours (Figure 1M).

Note that the female-like phenotypes of Mpyr-13-4 were stably maintained through (asexual) vegetative reproduction, and were not affected by culture conditions (10°C, 15°C, high/low white light, red light).

Taken together, these analyses indicated that Mpyr-13-4 gametophytes exhibited several phenotypic features more typical of females despite being genetically male, including gametophyte cell size, gamete cell size and parthenogenetic capacity.

### Mpyr-13-4 gametes do not produce the male-attracting pheromone lamoxiren

When Mpyr-13-4 gametes were confronted with active sperm from an unrelated, wild-type *M. pyrifera* male line, no interaction was observed, and no zygotes were produced.

Wild-type female gametes attract male gametes by producing a pheromone (lamoxiren; (Maier et al., 2001)). We therefore investigated whether Mpyr-13-4 was capable of producing laminoxiren. MS/MS analysis of *M. pyrifera* detected lamoxiren in control wild-type female gametophytes, but failed to detect lamoxiren or other C11 hydrocarbons in feminised Mpyr-13-4 fertile gametophytes (Figure 2A-C).

### Gene expression in Mpyr-13-4 compared with wild-type male and female gametophytes

To determine whether the morphological feminisation of the variant male strain was associated with changes in terms of transcriptional landscape compared with wild type male and female gametophytes, transcript abundances were measured by RNA-seq analysis of wild-type male, wild type female and Mpyr-13-4 gametophytes (Table S1 and methods for details).

Transcript abundances, measured as transcript per million (TPM) were strongly correlated between biological replicates of each sample (Figure S2B). This analysis identified 17,922, 18,436 and 17,744 expressed genes (defined as TPM>5<sup>th</sup> percentile) in wild-type males, wild type females and the Mpyr-13-4 strain, respectively, (Figure S2B) of the total of 22,242 genes that have been annotated in the *M. pyrifera* genome (Table S2). Hence, 81%, 82% and 80% of the total number of annotated genes were detected as being expressed in wild-type male, wild-type female and Mpyr-13-3 gametophytes, respectively. The majority (app. 94%) of the transcriptome was expressed in all three samples but a slightly greater number of genes were

uniquely expressed in wild-type females than in wild-type males and Mpyr-13-4 (3%, 1.8% and 0.69% of their transcriptomes, respectively; Figure S2B). The variant Mpyr13-4 strain shared more expressed genes with wild type females than with wild type males (610 and 327 genes respectively; Figure S2B).

### Analysis of differential gene expression in wild-type male, wild type female and Mpyr-13-4 gametophytes

DEseq2 analyses identified 5,442 genes that were differentially expressed between males and females, indicating that a considerable proportion (24.5%) of the transcriptome of *M. pyrifera* exhibits sex-biased expression (Table S2). Approximately the same numbers of genes were found to be male-biased (2785) and female-biased (2657) (Table S2, Figure 3A). Consistent with the female-like phenotype of Mpyr13-4, more differentially expressed genes were identified when this strain was compared with wild-type males (20.8%) than when it was compared with WT females (5.8%) (Table S2, Figure 3A).

### Sex-biased gene expression in wild-type and variant Mpyr-13-4 lines

Hierarchical clustering of expression levels was used to visualize gene transcription patterns for wild-type male, female and variant Mpyr-13-4 samples. Wild-type female and Mpyr-13-4 samples clustered together both when sex-biased gene expression was analysed and when the expression of non-sex-biased genes (unbiased genes) was analysed, indicating that the transcriptome of the Mpyr-13-4 strain was more similar to that of a wild type female than to that of a male (Figure 3B). Principal component analysis (PCA) further supported this conclusion (Figure S2C).

In order to examine how sex-biased gene expression is affected by the phenotypic feminisation of the variant Mpyr-13-4 strain, we next focused on the sets of genes that had been defined as sex-biased in the wild type strains. The median level of expression of the male-biased gene set in Mpyr-13-4 gametophytes (median TPM=0.66) was 79.4% lower than that observed in wild-type males (median TPM=3.211; Wilcoxon test,  $p < 2.2E-16$ , Figure 4A), suggesting these genes are transcriptionally "de-masculinized" in Mpyr-13-4. Female-biased genes were expressed at a higher level in Mpyr-13-4 (i.e. 31.6% more) than in wild-type males (Wilcoxon test,  $p < 2.2E-16$ , Figure 4A), suggesting that Mpyr-13-4 is transcriptionally "feminized". However, some male-biased genes were upregulated in Mpyr-13-4 compared to the wild type female indicating that there was not a complete shift to the female transcriptional program. A subset of the set of

female-biased genes that was completely silenced in wild type males was expressed in Mpyr-13-4. Conversely, male biased genes that were silenced in wild type females were expressed in Mpyr-13-4 (Figure 4B).

We also noticed that whilst wild-type females strongly activated female biased genes and downregulated male biased genes, the transcript abundance of male and female biased genes in males was less obvious (Figure S3, Table S3). Mpyr-13-4 showed, again, a pattern very similar to that observed for wild-type females, with the male-biased gene set being significantly down regulated and the female-biased gene set being only slightly upregulated, on average (Figure S3).

To examine in more detail the relationship between degree of sex-biased gene expression (fold change in transcript abundance between the wild type female and male) and transcript abundance, the sex-biased genes were grouped according to fold-change (FC) differences between the male and female samples and median transcript abundances plotted for wild type males, wild-type females and Mpyr-13-4. The global pattern of demasculinization and feminization of Mpyr-13-4 transcription appeared to be strongly correlated with the degree of sex-bias (Figure 4C). Down-regulation of the male-biased gene set in Mpyr-13-4 was more pronounced for the genes that showed the highest levels of male-bias, suggesting that the male-biased genes that exhibit the greatest fold changes make the greatest contribution to male-specific traits in the variant strain. Similarly, stronger upregulation was observed for the genes that exhibited the highest levels of female-bias (Figure 4C).

Taking together, these results indicate the transcriptome of Mpyr-13-4 is both demasculinised and feminised, with the strongest effect on genes that exhibit strong sex-biased expression patterns.

### Evolutionary analysis of genes involved in sexual differentiation

To test for differences in rates of evolutionary divergence between different categories of sex-biased and unbiased genes, we calculated levels of nonsynonymous (dN) and synonymous (dS) substitution using pairwise comparisons with orthologs from another kelp species, *Saccharina japonica*. The results of this analysis indicated that genes exhibiting male-biased expression patterns in *M. pyrifera* gametophytes evolve significantly faster (i.e., had higher dN/dS values) than female-biased or unbiased genes (Mann–Whitney U test,  $P < 0.01$ ) (Figure 5, Table S4). Female-biased genes were found to have evolved more slowly than unbiased genes.

### Predicted functions of sex-biased genes

An analysis of gene ontology (GO) terms associated with the sex-biased genes and with the genes differentially expressed in Mpyr-13-4 was carried out using BLAST2GO (Conesa and Götz, 2008). The aim of this analysis was to search for enrichment in particular functional groups and to relate gene function to phenotypic feminisation. Significant enrichment of specific GO categories related to microtubule- and flagella-related processes was detected for genes that were down-regulated in Mpyr-13-4 compared with wild-type males (Table S4). These genes may be involved in the production of flagellated gametes, which are not produced by Mpyr-13-4. Note that these two GO categories were also enriched in the set of male-biased genes expressed in male gametes identified by (Lipinska et al., 2013) and in fertile male gametophytes (Lipinska et al., 2015b, 2019). The female-biased genes that were upregulated in Mpyr-13-4 compared to wild type males were enriched in GO terms related to metabolism and photosynthesis (Table S4).

### Expression of male SDR genes in the Mpyr-13-4 line

The feminised phenotype of the Mpyr-13-4 strain would be consistent with modification of a sex-determining gene or genes carried by the V sex chromosome. We therefore analysed the expression patterns of *M. pyrifera* orthologues of *Ectocarpus* male SDR genes (Lipinska et al., 2017) using RNA-seq data. These 10 genes have been previously shown to be male-linked in *M. pyrifera* (Lipinska et al., 2017). For the majority (8/10) of these genes, similar transcript abundances were detected in wild-type males and the Mpyr-13-4 strain (Figure 6). However, two of the genes, gHMG.13001750 and gSDR.13001840, were markedly downregulated in the Mpyr-13-4 strain compared with the wild type male. The orthologs of these two genes in *Ectocarpus* sp. are Ec-13-001750 and Ec-13-001840, respectively (Figure 5). Both genes are part of the group of genes that have been conservatively sex-linked across all brown algae species investigated so far (Lipinska et al., 2017).

## 2.4 Discussion

### Mpyr-13-4 is partially sex reversed

Gametophytes of the giant kelp strain Mpyr-13-4 exhibited female phenotypic characteristics despite being genetically male. We used near-isogenic wild type male and female lines to examine whether the degree of sex-biased gene expression is associated with phenotypic sexual



dimorphism, and to understand the role of demasculinization and feminization in gene expression in encoding the variant morphology. Our results reveal a simultaneous demasculinization and feminisation of the transcriptome of the variant line and underline the association between sex-specific phenotypes and sex-specific transcriptomic patterns.

One particularly interesting phenomena highlighted by our study was the capacity of the feminised line to undergo parthenogenesis, a trait that is typical of females of oogamous brown algal species (Luthringer et al., 2015). In the brown alga *Ectocarpus*, parthenogenesis is a complex genetic trait under the control of the SDR, together with at least two additional autosomal quantitative trait loci (Mignerot et al., 2019), and it has been suggested that either the male SDR produces a repressor of parthenogenesis, or, alternatively, the female SDR produces an activator of parthenogenesis (in either case the activator or repressor could be directly encoded by the SDR or produced indirectly as part of the male or female sex-differentiation programs). Taking into consideration the parthenogenetic development of the egg-like cells of the Mpyr-13-4 strain (which lacks the female SDR), observed in this study, the most likely mechanisms involved in parthenogenesis are either 1) a repressor of parthenogenesis being produced either directly or indirectly by the male SDR, as part of the male-differentiation program or 2) parthenogenesis being induced as part of the female sex-differentiation program independently of the female SDR.

Despite the phenotypic resemblance to a female, Mpyr-13-4 is unable to attract male gametes, indicating that this line is incompletely sex-reversed. MS/MS analyses revealed no evidence for pheromone production in the eggs-like cells of the variant line, so it appears that the factor(s) required to initiate the pheromone pathway is absent in this line. Examination of the predicted functions of a list of 592 genes that were expressed in wild type females but not in wild type males nor Mpyr-13-4 revealed a enrichment in receptor signalling functions, and interestingly, in receptor kinase signalling. This pathway is involved in hormone signalling cascades in animals (Rawlings et al., 2004). It is therefore likely that among these genes there are components of the pheromone biosynthetic cascade or its regulators.

The phenotype of the Mpyr-13-4 suggests that the female SDR may be necessary to fully express the female program of development, in particular pheromone production. Pheromone production may therefore be controlled by a female SDR gene. Alternatively, sexually antagonistic interactions between genes located in the SDR and the sexual development pathway controlled by the male master gene may negatively impact reproductive fitness (including pheromone pathways) when gamete type does not match the SDR genotype.

Evidence for the latter mechanism has been suggested for the green alga *Volvox* (Geng et al., 2014).

### The transcriptional landscape underlying sexual differentiation in the giant kelp

Sex-biased gene expression has been characterised for the brown alga *Ectocarpus*, a near-isogamous species with a low level of sexual dimorphism (Lipinska et al., 2015b; Luthringer et al., 2015). Here we show that substantially more genes are sex-biased in the giant kelp (24% of the transcriptome, compared with less than 10% in *Ectocarpus* (Lipinska et al., 2015b). The higher proportion of sex-biased genes is consistent with the higher level of phenotypic sexual dimorphism in *M. pyrifera*, where male and female gametophytes have clearly distinct morphologies (Muller et al., 1979).

Sex-specific patterns of gene expression appear to be more dynamic in *M. pyrifera* females than in males. This pattern is, surprisingly, the opposite to *Ectocarpus*, where males actively increase the expression level of male biased genes (and strongly silence female biased genes) whereas females are less dynamic.

### Evolution of sex-biased genes in the giant kelp

The rapid rates of evolution sex-biased genes observed in animals are thought to be due to a combination of natural selection, sexual selection, and relaxed purifying selection (Grath and Parsch, 2016; Parsch and Ellegren, 2013). Although sexual dimorphism is generally more marked in animals compared to plants, male and female flowering plants with separate sexes also experience conflicts over trait optima and are subject to natural and sexual selection leading to a range of phenotypic sexual differences (Barrett, 2013). Studies of patterns of differential male and female gene expression in plants indicate that sex-biased gene expression plays a role in the evolution of sexual dimorphism (Harkess et al., 2015, Jiang et al., 2012; Zemp, Tavares, & Widmer, 2015).

Our study revealed that the genes involved in sex-specific development in *M. pyrifera* show a broadly similar pattern of sequence evolution to that seen in animals, with male-biased genes showing elevated rates of evolution compared to female-biased and unbiased genes, consistent with the notion that sex-specific selection may be stronger in males than females in this species. This contrasts with previous studies in another brown alga species, *Ectocarpus* sp., where both male- and female-biased genes exhibited faster evolutionary rates than unbiased genes (Lipinska et al., 2015b). *Ectocarpus* has a low level of sexual dimorphism, so this may result

in more sexual selection for both male and female biased genes compared with unbiased genes. Kelps have more conspicuous sexual dimorphism with males producing higher amount of sperm (Muller et al., 1979) compared with eggs, providing scope for sperm competition and higher levels of sexual selection in males.

In fish, where both male- and female-biased genes exhibit accelerated evolution, it has been proposed that this symmetry may be driven by the fact that these organisms have external fertilization (Yang et al., 2016). This idea is consistent with the symmetric evolutionary rates of male- and female-biased genes in *Ectocarpus*; where both male and female gametes are released in the seawater. Compared with internal fertilization in mammals, *Drosophila*, and birds, fish and algae such as *Ectocarpus* need to produce large quantities of female gametes, thus the selection pressure on females may be as strong as for males, driving rapid evolution of both female- and male-biased genes. In *Macrocystis* as in other kelp species, female gametes remain attached to the parental gametophyte and male gametes swim towards the eggs, attracted by the pheromone. Therefore males have to produce large quantities of rapidly swimming male gametes, and there may be more scope for male competition than for female competition.

Animal male-biased genes tend to have a smaller expression breadth than other genes, which could lead to a relaxation of selection pressure and increase their evolutionary rate, and a similar process may be occurring in *M. pyrifera*. Surprisingly, female-biased genes evolve slower compared with unbiased genes. No data on expression breadth are however currently available for *M. pyrifera* genes, so more information about gene expression in several tissues and life cycle stages would be necessary to further understand the evolutionary rates of male and female biased genes in this kelp species.

## 2.5 Materials and Methods

### Biological material

A *Macrocystis pyrifera* sporophyte was collected at Curahoe, Chile. Meiospores were isolated and clonal cultures of male and female gametophytes were established and propagated as described by Westermeier et al. (2006). Male and female gametophyte clones (Westermeier et al. 2009; 2010) were selected for the present study (Figure S1). Axenic sub-clones were initiated by antibiotic treatment as described by Müller et al. (2008) and maintained on 1% agar in seawater with transfers every 3-months. For clone isolation, culture medium was prepared with a commercial salt mixture (hw-Professional, Wiegandt, Krefeld Germany) in demineralized water and adjusted to 3% salinity with an optical refractometer. Nutrients were added with 20 ml/L PES-enrichment (Coelho et al., 2012; Starr and Zeikus, 2004). Colchicine treatment was performed using a disk of filter paper of 6 mm diameter loaded with 1 mg of colchicine (Fluka, Honeywell Research Chemicals), which was placed in the centre of an agar plate filled with gametophyte material, with good contact between agar and paper in order to allow diffusion of the colchicine into the agar. The agar plates were sealed with parafilm and subjected to culture conditions ( $12 \pm 2^\circ\text{C}$  and  $2\text{-}3 \mu\text{E m}^{-2} \text{sec}^{-1}$  from daylight type fluorescent lamps for 14:10 (light:dark) cycles for 12 to 16 weeks until selected regenerates were isolated.

### Pheromone measurements

Female eggs of *M. pyrifera* produce a pheromone (lamoxiren) that attracts male gametes. To induce gametogenesis in wild-type female and Mpyr-13-4 lines, light intensity was increased to  $30 \mu\text{E m}^{-2} \text{sec}^{-1}$  and culture medium was refreshed every 4 days. Oogonia and eggs were produced between 9-16 days after culture in these conditions.

For each of the tested strains, 25 mL of fertile gametophyte cultures containing  $6.5 \times 10^4$  eggs (wild type female) or  $3.1 \times 10^5$  eggs (Mpyr-13-4) were introduced in 50 mL Greiner Cellstar tissue culture tubes placed horizontally to increase surface area.

All algal pheromones known are hydrophobic, cycloaliphatic unsaturated hydro-carbons, comprising eight to eleven carbon atoms [1]. The sperm-releasing pheromone in the Laminariales carries an additional epoxy moiety, but is still a volatile and hydrophobic compound (Maier et al., 2001). To efficiently trap these pheromones, which are released in only minute amounts from fertile gametes, Solid Phase Micro-extraction (SPME) was used (Maier et al., 1996). SPME fibers and the holder were obtained from Supelco (Bellefonte, PA,

USA). For volatile extraction, a poly-dimethyl-siloxane (100  $\mu\text{m}$  PDMS) fiber was used (red fiber). Prior to use, the SPME fibers were conditioned according to the manufacturer's instructions. After 14 h exposure time of the fiber in the culture medium, the fiber was thermally desorbed in the gas chromatographer (GC/MS) injection port followed by separation of the volatiles under programmed conditions using an ISQ LT or Trace 1310 (Thermo Fisher Scientific GmbH, Dreieich, Germany) device equipped with a ZB5 column (30 m, 0.25 mm I.D., 0.25  $\mu\text{m}$  film thickness) linked to a guard column (10 m, Phenomenex, Aschaffenburg, Germany). Helium (1.5  $\text{ml}\cdot\text{min}^{-1}$ ) served as the carrier gas. Separation of compounds was achieved under programmed conditions. 50°C (2 min isotherm), followed by heating at 10°C  $\text{min}^{-1}$  to 200°C and at 50 °C  $\text{min}^{-1}$  to 280°C. The GC injector (splitless, splitless time 2 min), transfer line and ion source were set at 230, 280 and 250°C, respectively. Mass spectra were recorded in electron impact (EI) mode at 70 eV, 35–350 m/z.

### Generation of transcriptomic sequence data

The algal strains used, sequencing statistics, and accession numbers are listed in Table S1.

RNA-seq analysis was carried out to compare the relative abundances of gene transcripts in the different samples. For each sample, total RNA was extracted from 2 independent bulks of app. 1000 male individuals and 2 bulks of 1000 female individuals (two biological replicates for each sex) using the Qiagen Mini kit (<http://www.qiagen.com>) as previously described (Arun et al., 2019; Lipinska et al., 2015b). RNA from wild type male and female and variant Mpyr-13-4 line pools was extracted using the protocol described by (Apt et al., 1995). RNA quality and quantity was assessed using an Agilent 2100 bioanalyzer, associated with an RNA 6000 Nano kit.

For each replicate, the RNA was quantified and cDNA was synthesized using an oligo-dT primer. The cDNA was fragmented, cloned, and sequenced by Fasteris (CH-1228 Plan-les-Ouates, Switzerland) using an Illumina Hi-seq 2000 set to generate 150-bp single-end reads. Data quality was assessed using the FastQC (Wingett et al. 2018). Reads were trimmed and filtered using Trimmomatic (Bolger et al., 2014) with average quality > 28, a quality threshold of 24 (base calling) and a minimal size of 60 bp.

Filtered reads were mapped to the *M. pyrifera* genome (Lipinska et al., 2019) using TopHat2 (Kim et al., 2013) with the Bowtie2 aligner (Langmead and Salzberg, 2012).

More than 70% of the sequencing reads for each library could be mapped to the genome (Table S1). The mapped sequencing data were then processed with FeatureCounts (Liao et al., 2014) to obtain counts for sequencing reads mapped to exons and counts by gene.

Expression values were represented as transcript per million (TPM) and a filter of TPM>5th percentile was applied to remove noise and genes with very low expression levels. This resulted in a total of 19208 genes with expression values above the threshold.

Differential expression analysis was performed with the DESeq2 package (Love et al. 2014) (Bioconductor) using an adjusted P-value cut-off of 0.05 and a minimal fold-change of 2. Full lists of sex-biased genes can be found in Table S3.

### Identification of sex-biased genes

The filtering steps described above yielded a set of expressed genes in the transcriptome that were then classified based on their sex-expression patterns. Genes were considered to be male-biased or female-biased if they exhibited at least a twofold difference in expression between generations with a false discovery rate (FDR) of < 0.05. Sex-biased genes were defined as sex-specific when the TPM was below the fifth percentile for one of the sexes.

### Evolutionary analysis

To estimate evolutionary rates (non-synonymous to synonymous substitutions, dN/dS) we used single copy ortholog genes between *Macrocystis pyrifera* and *Saccharina japonica* described in (Lipinska et al., 2019). Protein sequences were aligned with Tcoffee (M-Coffee mode) (Notredame et al., 2000) and translated back to nucleotide sequence using Pal2Nal (Suyama et al., 2006). Gapless alignments that exceeded 100 bp were analysed with CodeML (F3x4 model of codon frequencies, runmode = -2) implemented in the Phylogenetic Analysis by Maximum Likelihood (PAML4) suit (Yang, 2007)). Genes with saturated synonymous substitution values (dS > 1) were excluded from the analysis.

### Flow cytometry

Flow cytometry was performed as described before for brown algal tissues (Bothwell et al., 2010). Gametophyte tissue was finely cut with a razor blade and nuclei were isolated by suspension in nuclei buffer (30 mM MgCl<sub>2</sub>, 120 mM trisodium citrate, 120 mM sorbitol, 55 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), pH 8, 5 mM EDTA supplemented with 0.1% (v/v) Triton X-100 and 5 mM sodium bisulfite; pH 8.0), and their

DNA content was measured immediately by flow cytometry. Between 600 and 13, 200 nuclei were analyzed in each sample. Wild type gametophytes were considered to be haploid and were used as an internal reference for the determination of ploidy. The nucleic acid-specific stain SYBR Green I (<http://www.invitrogen.com>) was used at a final dilution of 1:10,000. Samples were analyzed using a FACSort flow cytometer (<http://www.bsbiosciences.com>).

## 2.6 Supplemental Tables

**Table S1.** List of strains used and genome and transcriptome assembly statistics.

**Table S2.** Number of biased genes from DEseq2 analysis and categories of sex-biased genes with different levels of FC between the three samples. Only genes with TPM>5th percentile in at least one of the samples were considered for the analysis.

**Table S3.** Differential expression levels (DEseq2, FC>2, padj<0.05) between wild-type male, wild-type female and feminized Mpyr-13-4 samples.

**Table S4.** Gene expression (measured as TPM) in wild type males, wild type females and Mpyr-13-4 variant strain.

## 2.7 Acknowledgements

We thank Dominique Marie for help with the flow cytometry analysis. This work was supported by the CNRS, Sorbonne Université and an ERC grant to S.M.C. (638240).

## 2.8 References

- Ahmed S, Cock JM, Pessia E, Luthringer R, Cormier A, Robuchon M, Sterck L, Peters AF, Dittami SM, Corre E, Valero M, Aury J-M, Roze D, Van de Peer Y, Bothwell J, Marais GAB, Coelho SM. 2014. A haploid system of sex determination in the brown alga *Ectocarpus* sp. *Curr Biol CB* **24**:1945–1957. doi:10.1016/j.cub.2014.07.042
- Almeida LV, Coqueiro-Dos-Santos A, Rodriguez-Luiz GF, McCulloch R, Bartholomeu DC, Reis-Cunha JL. 2018. Chromosomal copy number variation analysis by next generation sequencing confirms ploidy stability in *Trypanosoma brucei* subspecies. *Microb Genomics* **4**. doi:10.1099/mgen.0.000223
- Apt KE, Clendennen SK, Powers DA, Grossman AR. 1995. The gene family encoding the fucoxanthin chlorophyll proteins from the brown alga *Macrocystis pyrifera*. *Mol Gen Genet MGG* **246**:455–464.
- Arun A, Coelho SM, Peters AF, Bourdareau S, Peres L, Scornet D, Strittmatter M, Lipinska AP, Yao H, Godfroy O, Montecinos GJ, Avia K, Macaisne N, Troadec C, Bendahmane A, Cock JM. 2019. Convergent recruitment of TALE homeodomain life cycle regulators

- to direct sporophyte development in land plants and brown algae. *eLife* **8**. doi:10.7554/eLife.43101
- Barrett SCH. 2013. The evolution of plant reproductive systems: how often are transitions irreversible? *Proc Biol Sci* **280**:20130913. doi:10.1098/rspb.2013.0913
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma Oxf Engl* **30**:2114–2120. doi:10.1093/bioinformatics/btu170
- Bothwell JH, Marie D, Peters AF, Cock JM, Coelho SM. 2010. Role of endoreduplication and apomeiosis during parthenogenetic reproduction in the model brown alga *Ectocarpus*. *New Phytol* **188**:111–121. doi:10.1111/j.1469-8137.2010.03357.x
- Coelho SM, Gueno J, Lipinska AP, Cock JM, Umen JG. 2018. UV Chromosomes and Haploid Sexual Systems. *Trends Plant Sci* **23**:794–807. doi:10.1016/j.tplants.2018.06.005
- Coelho SM, Scornet D, Rousvoal S, Peters NT, Darteville L, Peters AF, Cock JM. 2012. How to cultivate *Ectocarpus*. *Cold Spring Harb Protoc* **2012**:258–261. doi:10.1101/pdb.prot067934
- Conesa A, Götz S. 2008. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics* **2008**:619832. doi:10.1155/2008/619832
- Geng S, De Hoff P, Umen JG. 2014. Evolution of sexes from an ancestral mating-type specification pathway. *PLoS Biol* **12**:e1001904. doi:10.1371/journal.pbio.1001904
- Gershoni M, Pietrokovski S. 2014. Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. *Nat Commun* **5**:4438. doi:10.1038/ncomms5438
- Grath S, Parsch J. 2016. Sex-Biased Gene Expression. *Annu Rev Genet* **50**:29–44. doi:10.1146/annurev-genet-120215-035429
- Harrison PW, Wright AE, Zimmer F, Dean R, Montgomery SH, Pointer MA, Mank JE. 2015. Sexual selection drives evolution and rapid turnover of male gene expression. *Proc Natl Acad Sci U S A* **112**:4393–4398. doi:10.1073/pnas.1501339112
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**:R36. doi:10.1186/gb-2013-14-4-r36
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**:357–359. doi:10.1038/nmeth.1923
- Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma Oxf Engl* **30**:923–930. doi:10.1093/bioinformatics/btt656
- Lipinska, Ahmed S, Peters AF, Faugeron S, Cock JM, Coelho SM. 2015a. Development of PCR-Based Markers to Determine the Sex of Kelps. *PLoS ONE* **10**:e0140535. doi:10.1371/journal.pone.0140535
- Lipinska AP, D’hondt S, Van Damme EJ, De Clerck O. 2013. Uncovering the genetic basis for early isogamete differentiation: a case study of *Ectocarpus siliculosus*. *BMC Genomics* **14**:909–909. doi:10.1186/1471-2164-14-909
- Lipinska AP, Serrano-Serrano ML, Cormier A, Peters AF, Kogame K, Cock JM, Coelho SM. 2019. Rapid turnover of life-cycle-related genes in the brown algae. *Genome Biol* **20**:35. doi:10.1186/s13059-019-1630-6
- Lipinska AP, Toda NRT, Heesch S, Peters AF, Cock JM, Coelho SM. 2017. Multiple gene movements into and out of haploid sex chromosomes. *Genome Biol* **18**:104. doi:10.1186/s13059-017-1201-7



- Lipinska, Cormier A, Luthringer R, Peters AF, Corre E, Gachon CMM, Cock JM, Coelho SM. 2015b. Sexual dimorphism and the evolution of sex-biased gene expression in the brown alga *Ectocarpus*. *Mol Biol Evol* **32**:1581–1597. doi:10.1093/molbev/msv049
- Luthringer R, Cormier A, Peters AF, Cock JM, Coelho SM. 2015. Sexual dimorphism in the brown algae. *Perspectives in Phycology* **1**:11–25.
- Maier I, Hertweck C, Boland W. 2001. Stereochemical specificity of lamoxirene, the sperm-releasing pheromone in kelp (Laminariales, Phaeophyceae). *Biol Bull* **201**:121–125. doi:10.2307/1543327
- Maier I, Pohnert G, Pantke-Boecker S, Boland W. 1996. Solid-phase microextraction and determination of the absolute configuration of the *Laminaria digitata* (Laminariales, Phaeophyceae) spermatozoid-releasing pheromone. *Naturwissenschaften* **83**:378–379.
- Mignerot L, Avia K, Luthringer R, Lipinska AP, Peters AF, Cock JM, Coelho SM. 2019. A key role for sex chromosomes in the regulation of parthenogenesis in the brown alga *Ectocarpus*. *PLoS Genet* **15**:e1008211. doi:10.1371/journal.pgen.1008211
- Müller DG. 1975. Sex expression in aneuploid gametophytes of the brown alga *Ectocarpus siliculosus* (Dillw.) Lyngb. *Arch Protistenk* **117**:297–302.
- Muller DG, Gassmann G, Luning K. 1979. Isolation of a spermatozoid-releasing and -attracting substance from female gametophytes of *Laminaria digitata*. *Nature* **279**:430–431. doi:10.1038/279430a0
- Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**:205–217. doi:10.1006/jmbi.2000.4042
- Parsch J, Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nat Rev Genet* **14**:83–87. doi:10.1038/nrg3376
- Pröschel M, Zhang Z, Parsch J. 2006. Widespread Adaptive Evolution of *Drosophila* Genes With Sex-Biased Expression. *Genetics* **174**:893. doi:10.1534/genetics.106.058008
- Rawlings JS, Rosler KM, Harrison DA. 2004. The JAK/STAT signaling pathway. *J Cell Sci* **117**:1281. doi:10.1242/jcs.00963
- Starr R, Zeikus J. 2004. UTEX—The Culture Collection of Algae at the University of Texas at Austin. *J Phycol* **29**:1–106. doi:10.1111/j.0022-3646.1993.00001.x
- Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* **34**:W609–612. doi:10.1093/nar/gkl315
- Umen J, Coelho S. 2019. Algal Sex Determination and the Evolution of Anisogamy. *Annu Rev Microbiol*. doi:10.1146/annurev-micro-020518-120011
- Westermeier R, Patiño D, G. Müller D. 2007. Sexual compatibility and hybrid formation between the giant kelp species *Macrocystis pyrifera* and *M. integrifolia* (Laminariales, Phaeophyceae) in Chile. doi:10.1007/s10811-006-9126-7
- Whittle CA, Extavour CG. 2019. Selection shapes turnover and magnitude of sex-biased expression in *Drosophila* gonads. *BMC Evol Biol* **19**:60. doi:10.1186/s12862-019-1377-4
- Yang L, Zhang Z, He S. 2016. Both Male-Biased and Female-Biased Genes Evolve Faster in Fish Genomes. *Genome Biol Evol* **8**:3433–3445. doi:10.1093/gbe/evw239
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**:1586–1591. doi:10.1093/molbev/msm088

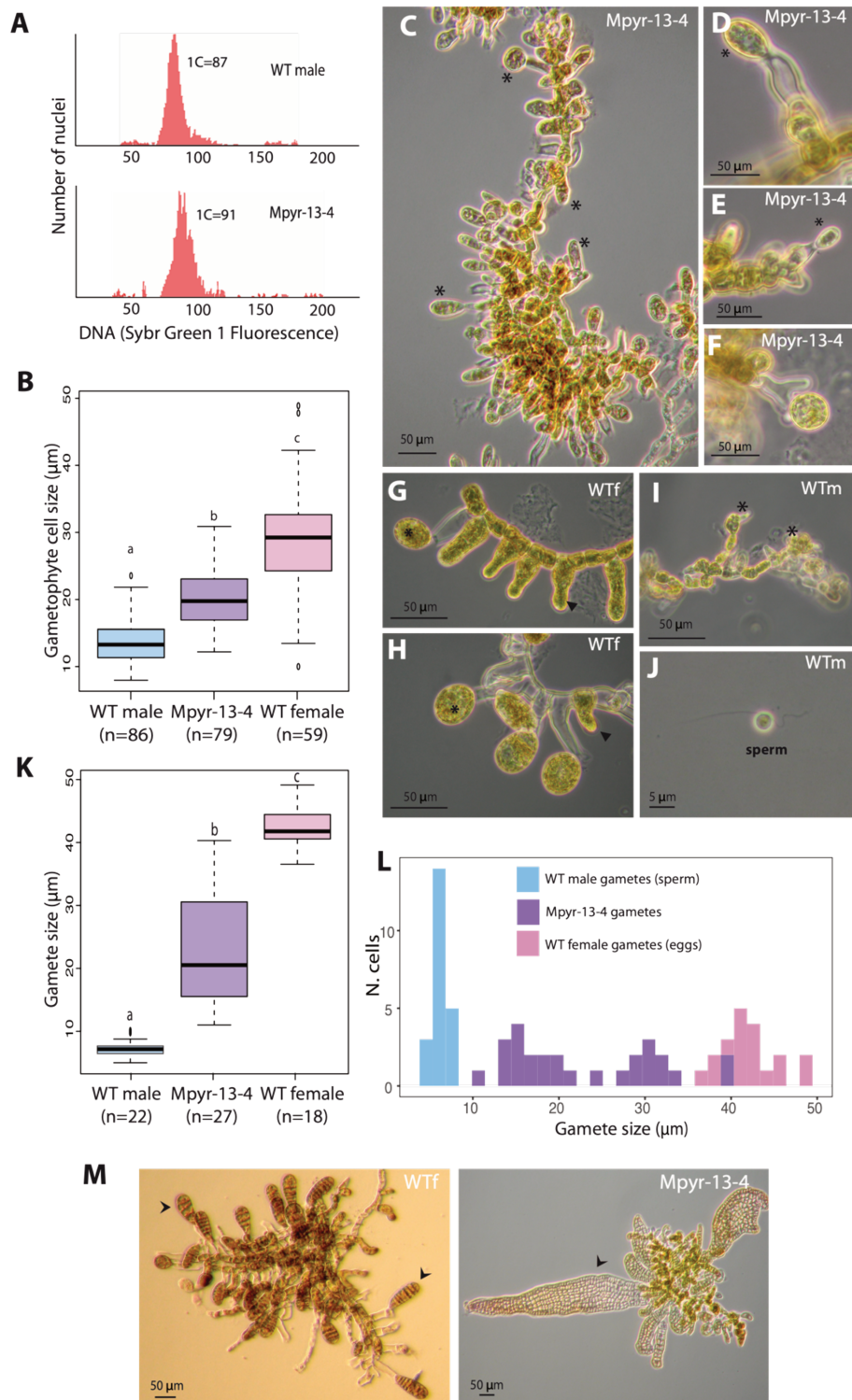


Figure 1. Phenotypic characterisation of wild-type male and female gametophytes and the variant Mpyr-13-4 line. (A) Flow cytometry analysis of a wild type male line and the Mpyr-13-4 line; (B) Sizes of gametophyte cells in a wild-type male, a wild-type female and a Mpyr-13-4 individual. (C-F) Fertile Mpyr-13-4 gametophytes showing several egg-like structures of different sizes (asterisks). (G-H) Fertile wild-type female gametophyte with extruded eggs (asterisks) and oogonia (arrow head); (I) Fertile wild-type male gametophyte with antheridia (asterisks) and sperm (J). (K) Gamete sizes in wild-type males, wild type

females and Mpyr-13-4. Different letters above the plots indicate significant differences (Wilcoxon rank sum test,  $p$ -value < 0.001). (L) Distribution of the gamete sizes for wild-type male, wild-type female and Mpyr-13-4 variant lines. (M) Parthenogenetic sporophytes in wild type females (left panel; arrow heads) and in Mpyr-13-4 (right panel, arrow heads).

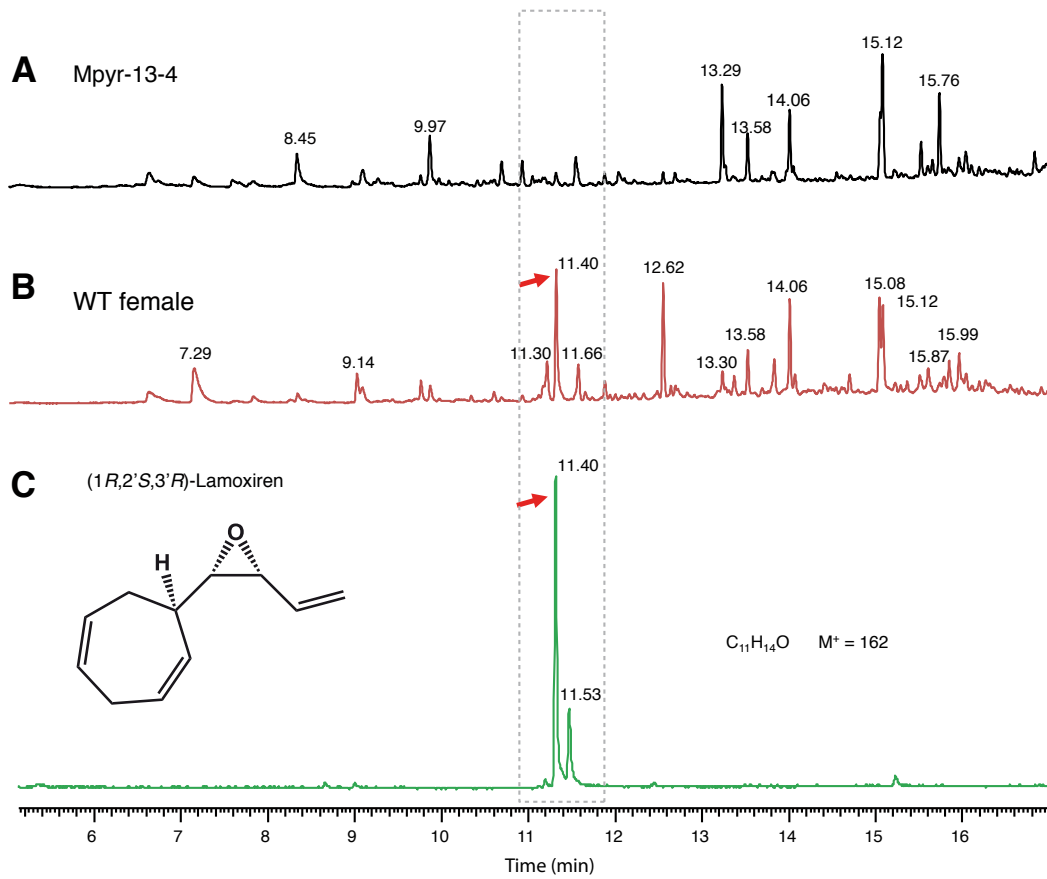
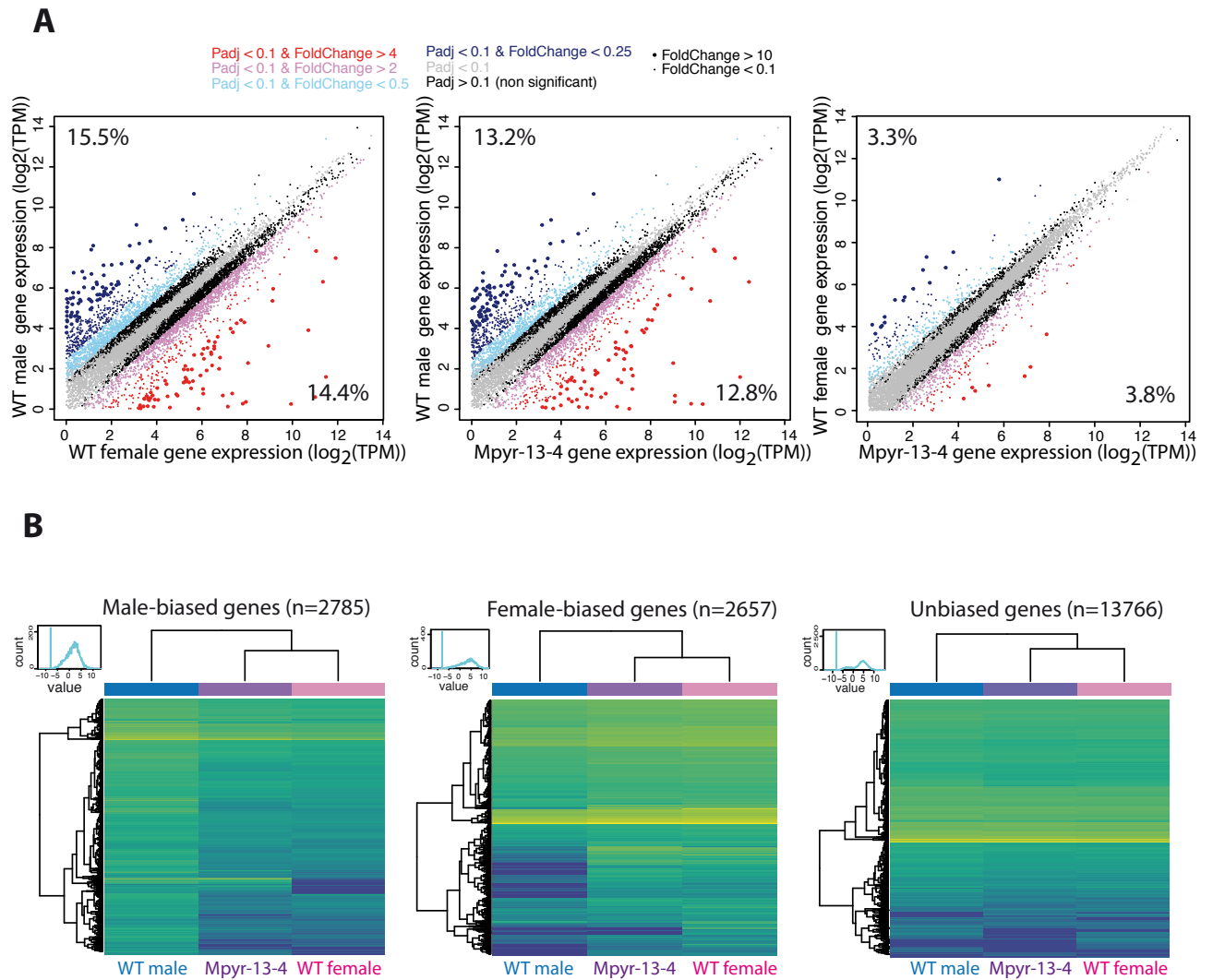
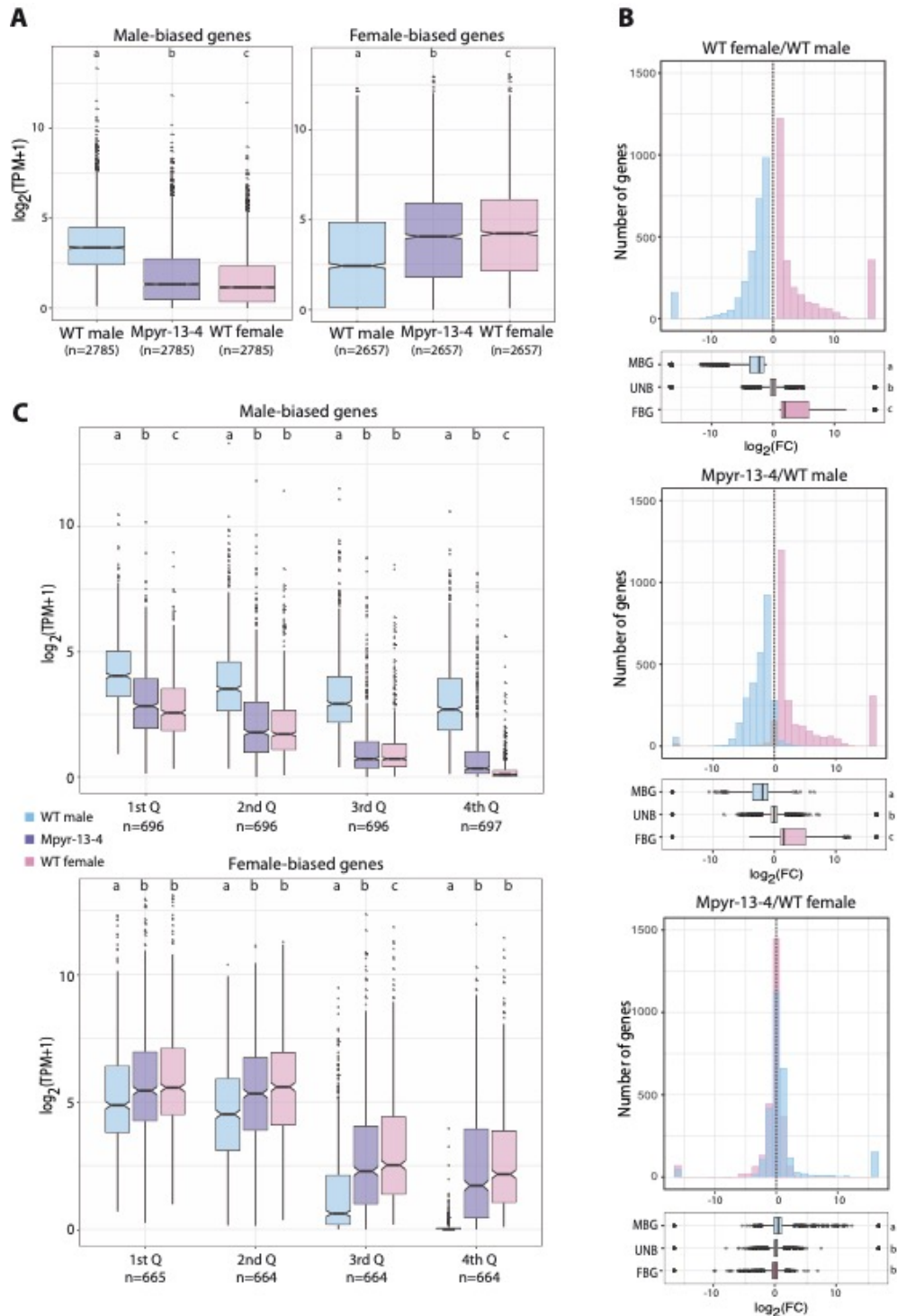


Figure 2. GS/MS assay for lamoxiren in the Mpyr-13-4 line (A), wild type females (B). The GS/MSprofile for lamoxiren is shown in (C).

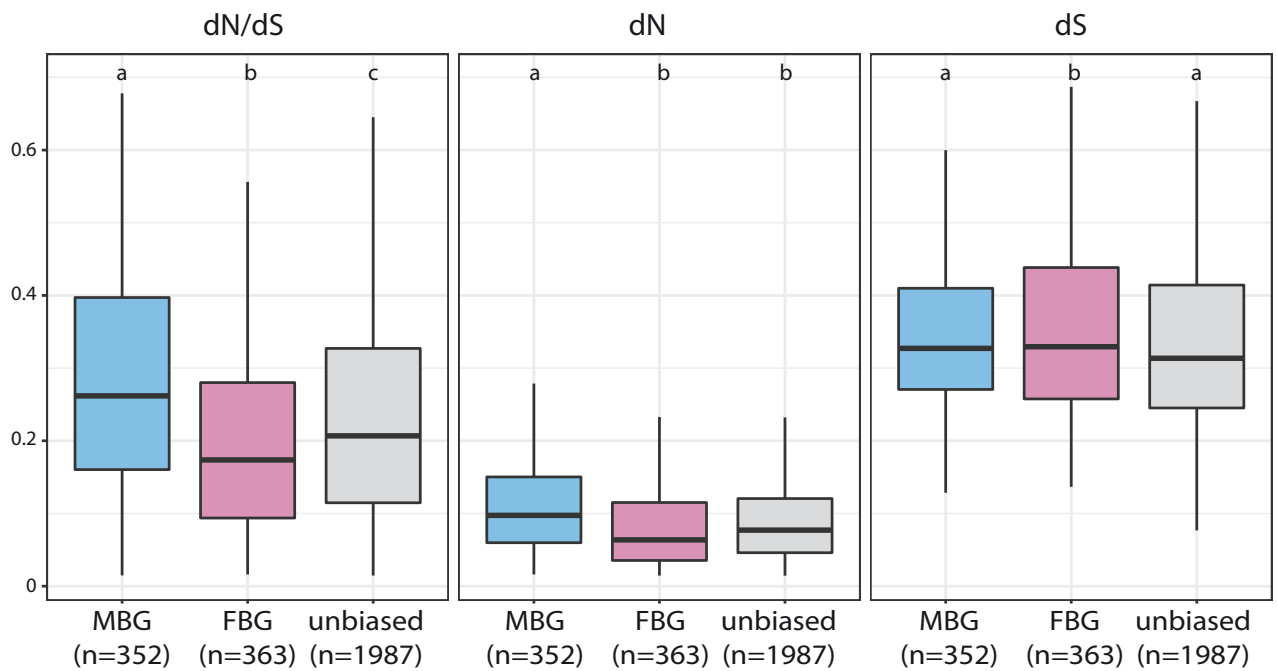


**Figure 3.** Sex-biased gene expression (A) Comparisons of gene transcript abundances (TPMs) in wild-type male, wild-type female and the Mpyr-13-4 line. Genes whose expression was 0 in one of the samples were removed from the plot. The percentages of biased genes in each sample are indicated in the upper left hand and lower right hand corners. (B) Heat maps and hierarchical clustering of gene expression for wild-type male, wild-type female and Mpyr-13-4 variant line. Shown is the relative expression for autosomal male-biased, female-biased and unbiased genes. Hierarchical gene clustering is based on Euclidean distance for average  $\log_2$  expression of each gene for the three samples.

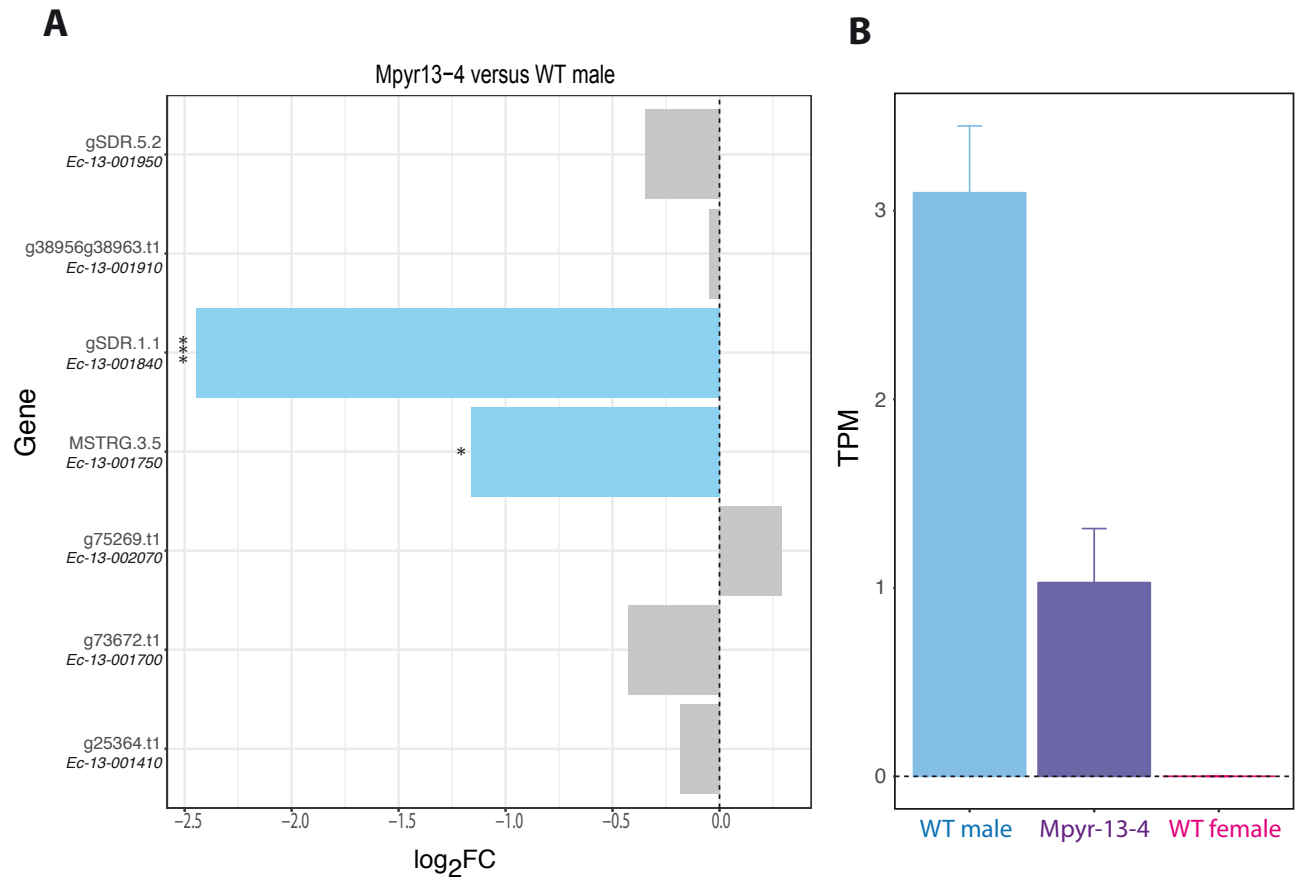


**Figure 4.** (A) Abundances of transcripts (TPM) of male- and female-biased genes in wild-type males, wild type females and Mpyr-13-4. (B) Transcript abundance fold changes in pairwise comparisons between wild-type males, wild type females and in Mpyr-13-4. **Upper panel:** wild-type female versus wild type male. Positive values in the x axis correspond to higher expression in females, negative values to higher expression in males. **Middle panel:** gene expression in Mp-13-4 versus wild-type male gametophytes. Positive values on the x axis correspond to higher expression in the variant Mpyr-13-4, negative values to higher expression in wild-type males. **Lower panel:** Positive values on the x axis correspond to higher expression in Mpyr-13-4, negative values to higher expression in wild type females. Colours indicate male-biased (blue), female-biased (pink) and unbiased (white) genes. For clarity, unbiased genes are omitted from the histogram. Different letters indicate significant differences (Wilcoxon rank sum test,  $p$ -value < 0.01). (C) Male-biased (upper panel) and female-biased (lower

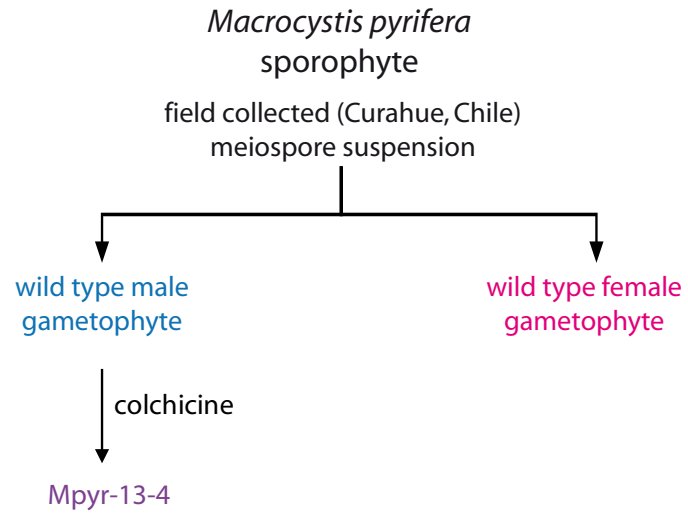
panel) genes ranked by level of bias ( $\log_2FC$ ). Numbers of genes are indicated above the plots. Note that only genes with  $TPM > 0$  are included in the plots. Significant differences detected following Wilcoxon tests ( $p$ -values  $< 0.01$ ) are indicated as different letters.



**Figure 5.** Rates of evolution of male, female and unbiased genes. Pairwise dN, dS and dN/dS ratios were calculated by comparing orthologous sequences from *M. pyrifera* and *S. japonica*. A) Ratio of nonsynonymous to synonymous substitutions (dN/dS), (B) synonyms substitutions C) nonsynonymous substitutions. Whiskers extend to the most extreme data point, excluding outliers that exceeded 1.5x the interquartile range. Significant differences detected following Wilcoxon tests ( $p$ -values  $< 0.01$ ) are indicated as different letters.



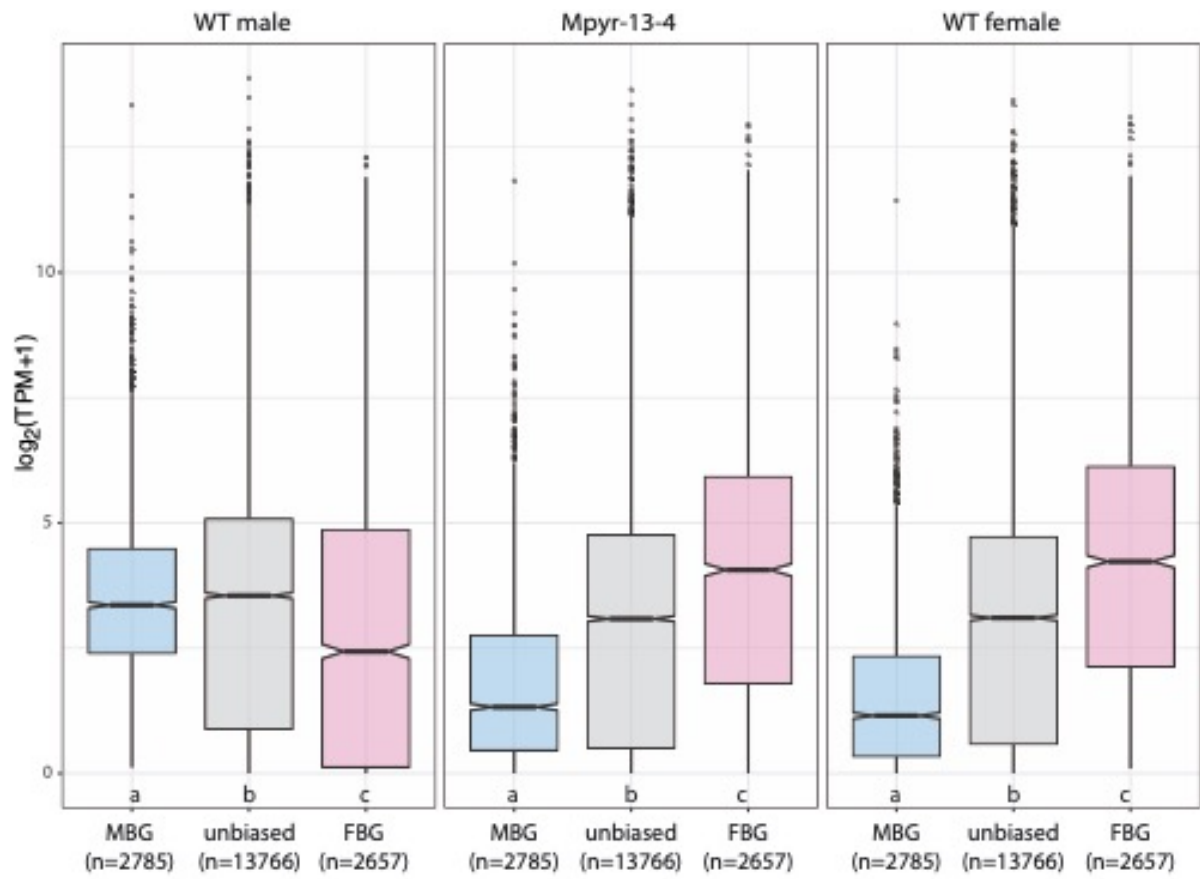
**Figure 6.** Differential expression (DESeq2) of *M. pyrifera* orthologues of *Ectocarpus* male SDR genes in Mpyr-13-4 versus wild-type males. Negative values represent downregulation (log<sub>2</sub>FC) in the variant Mpyr-13-4 strain relative to the wild-type male. The significance level of Wilcoxon rank sum test is indicated by asterisks (\*<0.01; \*\*\* < 0.0001).



**Figure S1.** Pedigree of the strains used in this study.







**Figure S3** - Levels of expression ( $\log_2(\text{TPM}+1)$ ) of sex-biased genes in wild-type males, feminised Mpyr-13-4 and wild type females. MBG: male-biased genes; FBG: female-biased genes.

# Chapter 5

## General discussion

## 1. Characterisation of the candidate male sex-determination gene in *Ectocarpus*, the HMG domain gene *Ec-13\_001750*.

Analysis of the expression of sex-linked genes during the *Ectocarpus* life-cycle indicated that the male-specific gene *Ec-13\_001750* was a strong candidate for the male sex-determining gene because it was highly expressed during the mature male gametophyte generation. The gene *Ec-13\_001750* codes for a HMG box protein, a class of protein that is known to be involved in male sex determination in most mammals (Kashimada and Koopman, 2010) and in mating type determination in fungi (Idnurm et al., 2008). Moreover, *Ec-13\_001750* was the only non-gametologue SDR gene that was conserved in the male-specific region of six diverse brown algal species (Lipinska et al., 2017). In addition, this is one of the two male-specific genes that shows a dramatic reduction of expression in a feminised strain of giant kelp, and this is a further argument to make this gene a strong candidate for male-sex determination (Müller et al., unpublished).

Ideally, elimination of this gene or inhibition of its expression by directed mutagenesis methods such as CRISPR-cas9, would be the best way to study its role in male sex-determination in brown algae. Genetic transformation is being developed for *Ectocarpus*, and RNAi silencing using dsRNA has been used (Macaisne et al., 2017) but is not very efficient. Alternative reverse genetic approaches could also be used, for example via the establishment of a TILLING machine. A TILLING protocol already exists for *Ectocarpus* (Godfroy, Perrinau et al. unpublished). If the function of *Ec-13\_00175* can be confirmed by one of these techniques, it will raise important questions about the evolution of sex-determination genes in eukaryotes. Indeed, HMG factors would then have been implicated in male sex determination in three phylogenetically distant eukaryotic supergroups (animals, fungi and brown algae), which have very different sexual systems (XY, mating type and UV).

Other methods may also provide information about the function of *Ec-13\_00175*. In particular, it would be interesting to identify proteins that interact with *Ec-13\_001750* in the cell and compare these results with those that have been obtained for other male sex determination regulatory masters in other lineages. It would also be interesting to determine whether *Ec-13\_00175* binds to DNA and to characterise the binding sites. Identification of DNA binding sites upstream of sex-biased genes involved in male and female differentiation would reinforce the hypothesis that this gene plays a role in sex determination.

Ec-13\_001750 is a member of the HMG box protein (HMGB) subfamily of the HMG superfamily. There are two subgroups within the HMGB transcription factor subfamily: those with a single HMG box and those with more than one HMG box. HMGB transcription factors with two or more HMG boxes have little or no DNA sequence specificity, unlike HMGB transcription factors with a single HMG box (Štros, 2010). The SRY protein has a single HMG box and binds to a specific consensus sequence (A/TAACAA/T) (Harley et al., 1994). Ec-13\_001750 has two HMG boxes. It is therefore possible that the binding sites of this protein do not correspond to a conserved consensus sequence motif but rather that binding depends on DNA conformation. However, there are HMGBs in plants that have only one HMG box but do not recognize specific sequences (Štros, 2010). There may also be exceptions to this general rule regarding binding site specificity among HMGBs with two HMG boxes. It was therefore important to carry out the DAP-seq experiment to attempt to confirm the specific DNA binding motifs for Ec-13\_001750. The DAP-seq experiments carried out in this thesis identify a candidate binding site motif for Ec-13\_001750 but these experiments should be repeated to ensure that the low number of binding sites identified was not the result of problems with the implementation of the methodology.

Moreover, a ChIP-seq approach would be extremely useful to determine whether Ec-13\_001750 binds to specific DNA sites *in vivo*. HMGB proteins are known to interact with chromatin, which they actively remodel (Agresti and Bianchi, 2003), and are predicted to promote nucleosome mobility and accessibility of proteins to specific DNA sites (Figure 1C). Although DAP-seq retains many of the tissue/cell line-specific secondary modifications and features present in genomic DNA such as cytosine methylation, the effect of additional genomic elements (such as chromatin accessibility and histone modification) are not reflected in DAP-seq datasets. DAP-seq therefore provides information about TF binding specificities but in a chromatin-free context. One powerful way to overcome this limitation of the DAP-seq method is to analyse the results of DAP-seq assays together with tissue-specific chromatin accessibility information derived from methods such as DNase-seq, ATAC-seq, and MNase-seq (Buenrostro et al., 2015; Rizzo and Sinha, 2014; Song and Crawford, 2010). For example, integration of DAP-seq and DNase hypersensitivity data from multiple *Arabidopsis* tissue types showed that DAP-seq captures *in vivo* binding sites that correspond to multiple tissue-specific binding events (O'Malley et al., 2016). This type of analysis offers a cost-effective means to assess the TF binding landscape across many tissues and cell-types without having to perform thousands of individual ChIP-seq experiments. Since chromatin context information is not available

through DAP-seq, a chromatin dimension can be added to the DAP-seq results by approaching what can be obtained in ChIP-seq. Thus, if we look for the binding sites of a transcription factor, the binding sites detected by the DAP-seq but located in a condensed chromatin region predicted by ATAC-seq can be eliminated. Conversely, in our case, since some HMGB proteins can bind to nucleosomes, it would be possible to detect the binding regions at the level of this nucleosomes.

A Y2H analysis was carried out with aim of identifying Ec-13\_001750-interacting proteins. Due to a possible defect of the Y2H library that was constructed using the *Ectocarpus* gametophyte mRNA, it was not possible to identify protein partners. To continue this line of research, a new Y2H library would need to be constructed and the screen repeated.

Alternative techniques could also be used to identify Ec-13\_001750 interactors. These include co-immunoprecipitation, crosslinking protein interaction analysis or pull-down assays using whole cell extract. If interacting partners are identified by these methods, pull down assays or a directed Y2H analyses should be performed to independently confirm the interactions. Similar searches for proteins that interact with SRY identified proteins involved in chromatin remodelling. SRY was found to associate with proteins such as SF1 and WT1 to activate target gene expression and with KRAB-O to negatively regulate target gene expression (Matsuzawa-Watanabe et al., 2003; Oh et al., 2005; Oh and Lau, 2006; Peng et al., 2009; Polanco et al., 2009; Sekido and Lovell-Badge, 2008). Since HMG proteins are generally involved in chromatin remodelling, it is possible that similar (chromatin regulatory) proteins may interact with Ec-13\_001750.

## 2. Chromatin modifications in male and female gametophytes

This thesis has provided considerable advances for the study of epigenetic regulation of sex determination in the brown algae *Ectocarpus*. Previous studies have described the chromatin landscape in diatoms, which are unicellular Straminopiles (Lin et al., 2012; Veluchamy et al., 2015); but our study together with (Bordereau, 2018) provides for the novel informations about epigenetic regulation in a complex multicellular Straminopile. Also, very few studies have been carried out to compare chromatin landscapes at the whole genome scale in males and females for any eukaryotic lineage (Brown and Bachtrog, 2014). The work presented here provided the first combined transcriptomic and epigenetic analysis of sexual differentiation in the brown algal lineage. We analysed six different histone post-translational modifications (H3K4me3,

H3K9ac, H3K27ac, H3K36me3, H4K20me3 and H3K27me3). Analysis of these chromatin marks in males and females indicated that H3K4me3, H3K9ac and H3K27ac were associated with actively transcribed genes and were localized at the TSSs of genes. The activation-associated mark H3K36me3 was also associated with active transcription but was localized on gene bodies. Presence of the repression-associated mark H4K20me3 was associated with decreased gene expression. Overall, our results are therefore consistent with the observed roles of chromatin remodelling and histone PTMs in plants and animals (Baroux et al., 2011; Margueron and Reinberg, 2010; She and Baroux, 2015) indicating that these functions have deep evolutionary origins and were present in the common eukaryotic ancestor of animals, plants and brown algae. Nevertheless, interestingly the histone PTM H3K27me3 has not been detected in *Ectocarpus* using mass spectrometry (Bordereau 2019). This mark plays an important role in transcriptional silencing in animals and plants and is deposited by Polycomb Repressive Complex 2 (PRC2) (Boyer et al., 2006; Zheng and Chen, 2011). PRC2 is not present in *Ectocarpus*, supporting the fact that PTM is not essential for transcriptional repression in *Ectocarpus*. Thus, despite the presence of histone post-translational modifications that are functionally preserved between brown algae, animals and plants, brown algae seem to diverge on some mechanisms related to chromatin dynamics. This divergence also concerns another transcriptional repression mechanism linked to chromatin dynamics: DNA methylation on the 5<sup>th</sup> position of cytosine (5mC), which is also absent in *Ectocarpus*, but present in the majority of plants and animals (with some exceptions like *C. elegans* which does not have 5mC but has an N6-methylation adenine (6mA) (Greer et al., 2015)). *Ectocarpus* must therefore have other epigenetic mechanisms responsible for transcriptional repression, which are currently unknown. It has been highlighted thanks to the study of landscape chromatin carried out in *Ectocarpus*, that chromatin states that include no activation-associated marks correspond to the states associated with the strongest transcriptional repression (i.e. chromatin states S1, no histone PTMs detected, and S2, only H4K20me3). Two hypotheses can be put forward. Thus, on the one hand, the absence of activation-associated marks show that do not necessarily need to repress a gene actively to switch it off. On the other hand, state S1 includes all genes that are not marked with any of the five histone PTMs analysed in the study but these genes may have additional marks associated with gene repression that have not yet been analysed in ChIP-seq. The histone PTM analysis, in *Ectocarpus* was, has made it possible to identify combinations of marks present at the level of genes. These different combinations were used to define chromatin states. The proportion of the genome that could be assigned to each of these chromatin states

was almost identical for male and female samples. Interestingly however, nearly 40% of individual genes were assigned to different chromatin states in males compared with females. This observation suggests that it may be necessary for the proportions of each state to be maintained at equivalent levels at the genome level but that transitions between states can occur (at high frequency) for individual genetic loci. Tendencies to maintain the relative proportions of chromatin states, genome-wide, have also been reported for other organisms such as *Drosophila* (Francisco and Lemos, 2014).

In addition, it is interesting to note that, despite the low level of sexual dimorphism and the relatively small number of sex-biased genes in *Ectocarpus* compared with animals (Lipinska et al., 2015; Luthringer et al., 2015), the former exhibited modifications of the chromatin states of 40% of its genes when males were compared with females. Thus, differences in chromatin states between the sexes not only occur at sex-biased genes, but also at other genes in the genome.

The patterns of histone PTMs at sex-biased genes differed in males and females. In females, many male-biased genes (MBGs) were decorated with both activation-associated marks and the repression-associated mark H4K20me<sub>3</sub>, while H4K20me<sub>3</sub> was absent from the same genes in males. MBGs appeared to be in a bivalent state in females, since they were decorated with both activation-associated and repression-associated marks. This state of bivalence has been observed in embryonic stem cells (ESCs) where promoters of developmental genes are often marked by both the activation-associated mark H3K4me<sub>3</sub> and the repression-associated mark H3K27me<sub>3</sub> (Azucara et al., 2006; Bernstein et al., 2006). This bivalency plays a key role in pluripotency by maintaining genes in a poised state for rapid activation or repression during differentiation. We could therefore hypothesise that H4K20me<sub>3</sub> is removed in *Ectocarpus* males to allow expression of MBGs. The female developmental program may be considered a "default" program since the transition of PTM histone at the level of male biased genes between male and female has a sex-specific pattern, unlike female biased genes where the transition of PTM histone between male and female is very close to a regulation observed for tissue specific genes. In addition, the male program appears to be actively deployed in males, unlike the female program in females. These results support previous evidence obtained using a genetic approach, which suggested that the female program is triggered by default in *Ectocarpus* (Ahmed et al., 2014). To investigate if the female is the default program and that the MBGs are in a bivalent state, a ChIP-seq analysis should be performed on the sporophyte. This analysis would provide a global picture of changes in epigenetic modification of genes during the transition from the



asexual to the sexual generation. Interestingly, the PAR of the U and V sex chromosomes have more differences in their chromatin states than when autosomes are compared between males and females. In other species such as *Silene latifolia*, which has an XY sex chromosome system (Filatov, 2005). The *S. latifolia* Y chromosome was partially depleted of PTMs associated with transcriptionally active chromatin, and bore these marks only in the pseudo-autosomal region. However, the X chromosome of the male and one of the X chromosomes of the female were enriched in activation-associated histone PTMs (Bačovský et al., 2019). Similarly, activation-associated histone PTMs were strongly enriched on the *Drosophila* male X, whereas the inactivated X chromosome of female mammals is characterized by DNA methylation and repression-associated histone PTMs (Brockdorff and Turner, 2015; Lucchesi et al., 2005). However, all these differences were associated with either dosage compensation-related modifications of the homogamous chromosome or with enrichment of repeated elements on the Y chromosome, mechanisms that are also found in *D. melanogaster* or the pea aphid (Brown and Bachtrog, 2014; Lemos et al., 2010; Richard et al., 2017; Yasuhara and Wakimoto, 2008). Dosage compensation is not necessary in *Ectocarpus* since sex is expressed during the haploid phase, so the epigenetic differences between the U and the V probably have another function. Moreover, in contrast to the situation observed for autosomes, there was no correlation between expression level and chromatin state for sex chromosome genes. We therefore lack information to explain the regulation of genes on these chromosomes. Other epigenetic marks could be exploited, such as the H4K16ac mark has been detected by mass spectrometry in *Ectocarpus* (Bourdareau et al., unpublished). Indeed, H4K16ac is present on male X chromosomes in *Drosophila* and the pea aphid. H4K16ac is located at TSSs where it increases the accessibility of chromatin (Gelbart and Kuroda, 2009; Richard et al., 2017). In addition, H3K4me2 is located at gene promoters on the Y chromosome in mice (Garcia-Moreno et al, 2018) and H3K9me2, H3K9me3 and H4K20me are enriched in the inactivated X chromosomes of female mammals (Chow and Heard, 2009). These histone PTMs have all been detected in *Ectocarpus*. These marks are therefore good candidates for future in-depth analyses on the landscape chromatin of the U and V chromosomes in *Ectocarpus*.

A diploid male *Ectocarpus* gametophyte has been constructed that carries both the U and V sex chromosomes. Analysis of this male diploid strain would provide information about the behaviour of U and V sex chromosomes when they are present in the same cell because the V chromosome is dominant over the U chromosome in this context and could lead to the discovery of mechanisms of inactivation of the U chromosome. It would be interesting to analyse patterns

of histone PTMs on the U chromosome in this strain and to compare these patterns both with those of a U chromosome in a haploid female and with the U chromosome in the diploid (UV) sporophyte generation, where sex is not expressed. This analysis would allow changes in histone PTMs to be correlated with the transition from an asexual to a sexual stage and to evaluate the effect of gender and the presence of the V chromosome on histone modifications at U chromosome genes. In addition, studying the chromatin dynamics of U and V chromosome between the sporophyte and gametophyte generation would make it possible to understand regulatory events upstream of sex determination. For example, repression of the *Sry* gene has been correlated with an enrichment in the repression-associated, GLP/G9a-mediated H3K9me2 mark. *Sry* activation requires Jmjd1a-mediated removal of H3K9me2, CBP/p300-mediated deposition of H3K27ac and accumulation of the H3K4me2 at its promoter. It would therefore be interesting to look at how the *Ectocarpus Ec-13\_001750* gene is regulated during of male sexual determination.

Finally, it would be interesting to analyse histone PTMs during sex determination and differentiation in a broad range of brown algae to evaluate the evolution of sex-related epigenetic processes within this lineage. Brown algal sexual systems are highly diverse, they include both haploid and diploid genetic sex-determination systems, species with separate sexes (both dioicy and dioecy) and co-sexuality (both monoicy and monoecy) and isogamous, anisogamous and oogamous gamete systems (Luthringer et al., 2015; Coelho et al., 2019) (Figure 1). It would be interesting to correlate sex-related epigenetic features with these various reproductive traits. For example, systems with more pronounced sexual dimorphism may exhibit more complex mechanisms of sexual differentiation than systems with low sexual dimorphism. In other species sex chromosome structure is often correlated with gamete dimorphism, with small mating type loci tending to be found in isogamous species and sex chromosomes with large non-recombining regions in anisogamous and oogamous species (Lehtonen et al., 2016). Isogamy is thought to be the ancestral state of eukaryotes (Maynard Smith 1982). It would also be of interest to correlate the number of SBGs in each brown algal species with the level of sexual dimorphism, such as primary sexual characteristics (i.e. gametangia), or secondary sexual characteristics, which may be morphological (gamete size, gametophyte size) or physiological.

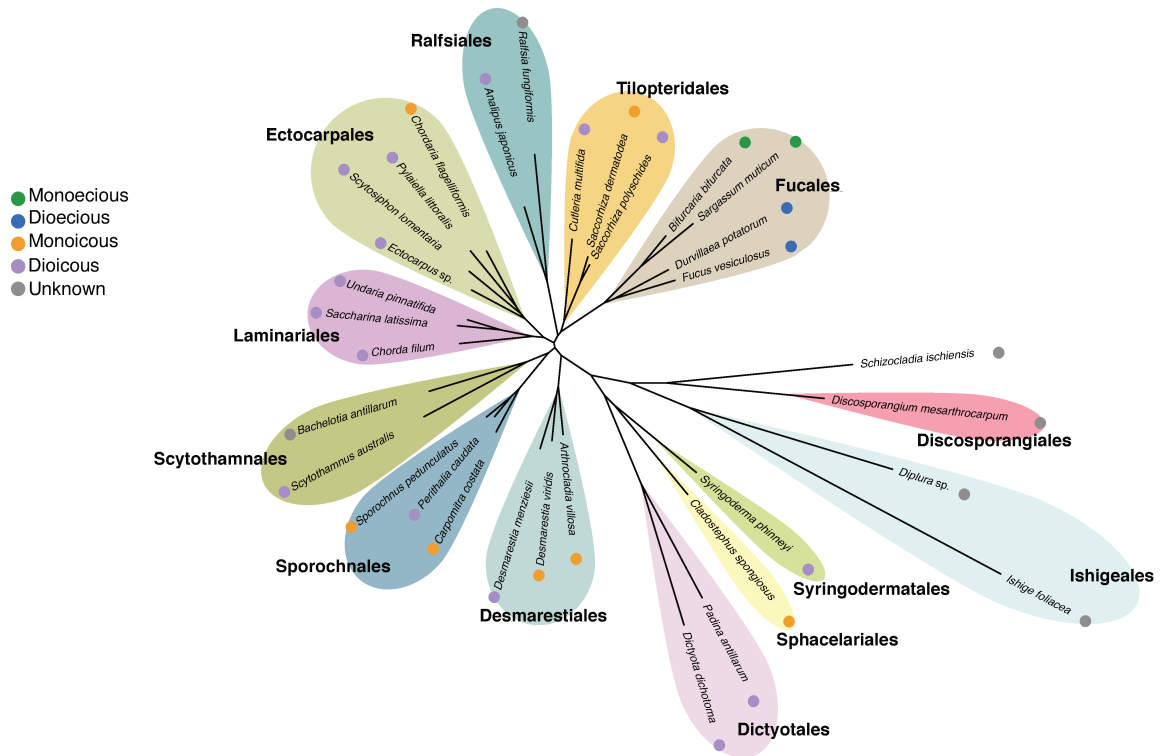


Figure 1. Diversity of sexual systems in the brown algae. (a) Evolutionary tree of the brown algae showing the major orders. Adapted from Coelho et al. (2019)

It has been shown that the number of SBGs and the level of sexual dimorphism are positively correlated in turkeys (Pointer et al., 2013). Similarly, more genes were found to be sex-biased in the oogamous giant kelp *Macrocystis* (24% of the transcriptome) than in quasi-isogamous *Ectocarpus* (less than 10% of the transcriptome), which is less dimorphic. In addition, it was shown in this thesis that *Ectocarpus* sex-biased genes have different histone PTM patterns in males and females, and that the differences are particularly marked for MBGs. Based on these observations, it would be interesting to analyse changes in histone PTMs during sexual differentiation, both for the U and V chromosomes and for the rest of the genome, in both an isogamous species with few SBGs, an anisogamous and in an oogamous species with many SBGs. We could thus have an overall view on the evolution of the epigenetic regulation mechanisms of sexual differentiation, between an ancestral character (isogamy) and a less ancestral character (oogamy), via intermediate characters (quasi-isogamy and anisogamy)

This sort of comparative analysis may provide insights into the relationship between histone PTM modifications and regulation of the expression of sex-biased genes. Indeed, in *Ectocarpus*, the regulation of MBGs follows a sex-specific pattern in males and females, unlike FBGs, which follow a pattern close to that of specific tissue genes. Is this pattern due to the quasi-isogamy of this brown alga? Indeed, given the low sexual dimorphism, male and female traits almost do not differ, and therefore do not require the implementation of a complex female differentiation pathway. Thus, would the no sex specific regulation of female SBGs be observed in an oogamous context where there is stronger dimorphism between male and female gametes?

It would therefore be interesting to correlate patterns of histone PTMs on brown algal sex-biased genes with levels of sexual dimorphism in isogamous, quasi-isogamous, anisogamous and oogamous species.

Finally, in the future, when information is available about the genomic binding sites of the HMG protein Ec-13\_001750 (if specific binding sites exist) and about the biological function of hypothetical partner proteins, it would be interesting to correlate this information with the distributions of histone PTMs across the genome during sexual differentiation.

### 3. Molecular changes associated with partial sex-reversal shed light into the mechanisms of sex differentiation in the giant kelp

Recent studies suggest that the giant kelp *Macrocystis pyrifera* possesses a UV sex determination system that shares a common evolutionary origin with that of the model brown

alga *Ectocarpus*. This study describes the identification of a genetically male giant kelp strain that has phenotypic features typical of a female. However, conversion to the female developmental program is incomplete, because gametes of this feminised male do not produce the sperm-attracting pheromone lamoxiren, and consequently are not fully functional.

A transcriptomic analysis that focused on MBGs and FBGs indicated that the phenotypic feminization of the variant strain was associated with both upregulation of FBGs and downregulation of MBGs. In addition, searches for orthologues of male and female SDR genes from *Ectocarpus* indicated that this strain lacks a U chromosome, suggesting that the U SDR is not necessary for at least partial deployment of the female developmental program. Thus, male and female individuals appear to be bipotential, in the sense that male individuals have the necessary tools for female development in the absence of U. Thus, to fully express the female program in *Macrocystis* it is required both an activation of the female differentiation pathway with a parallel repression of the male differentiation pathway.

This suggests that there are antagonistic mechanisms for regulating the expression of sexually biased genes, one allowing the activation of the sexual differentiation pathway and the other allowing the repression of the opposite sex differentiation pathway, which is characteristic of bipotential systems (e.g. mammals) (Garcia-Moreno et al., 2018). Regulation of sex-biased genes may involve chromatin modifications. It would therefore be interesting to extend this study by carrying out ChIP-seq experiments to compare post-translational histone modifications in the variant strain with patterns in wild type males and females, especially for SBGs (Figure 2). It would be interesting to determine if the expression patterns of sex-biased genes are correlated with patterns of histone PTMs in wild type males and females, and in the variant strain (Figure 2). In addition, it would be interesting to compare these results with those obtained for *Ectocarpus* (chapter 3), to determine whether histone PTMs of sex-biased genes have the same pattern in *Macrocystis*. This study could also address the question as to there is a different mechanism of chromatin regulation by PTM histones depending on sexual dimorphism.

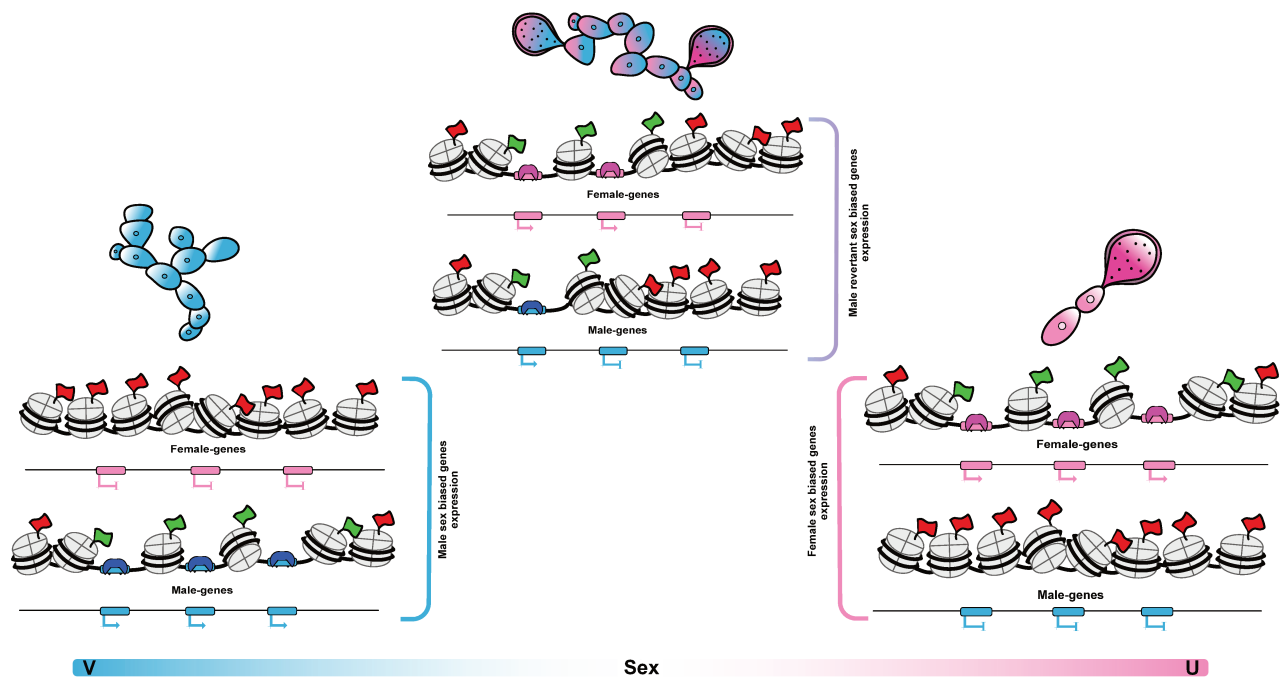


Figure 2. **Hypothetical model for epigenetic control of sex biased genes by histone PTMs in the brown alga *M. pyrifera*.**

The *Macrocystis* orthologue of the *Ectocarpus* candidate male sex-determining gene is significantly under-regulated in the feminised strain. Master regulators of sex, such as *Sry* in mammals or *Dmrt1* in birds and fish, are regulated by epigenetic mechanisms (Garcia-Moreno et al., 2018). For example, *Jmjd1a*-deficient mice exhibit male-to-female sex conversion (Kuroki et al., 2013). *Jmjd1a* encodes a histone demethylase that acts specifically on the repression-associated histone modification H3K9me2. *Jmjd1a*-dependent demethylation of H3K9me2 is required for subsequent H3K4me2 accumulation at the *Sry* promoter, leading to an open chromatin conformation and *Sry* activation. But, the decrease expression of the male candidate for male determination in *Macrocystis* may also be due to a mutation in one of the transcription factors responsible for its activation. That is why it would be important to have information about the sequence of the variant versus the sequence of the wild type to fully understand the molecular mechanisms underlying this feminisation. However, it is possible that no mutations in genes are found, and in this case the phenotype would be epigenetic. The sequencing is underway and we will know more information in the near future.

## References

- Abbott, D.W., Chadwick, B.P., Thambirajah, A.A., Ausió, J., 2005. Beyond the Xi: macroH2A chromatin distribution and post-translational modification in an avian system. *J. Biol. Chem.* 280, 16437–16445. <https://doi.org/10.1074/jbc.M500170200>
- Agresti, A., Bianchi, M.E., 2003. HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.* 13, 170–178. [https://doi.org/10.1016/S0959-437X\(03\)00023-6](https://doi.org/10.1016/S0959-437X(03)00023-6)
- Ahmed, S., Cock, J.M., Pessia, E., Luthringer, R., Cormier, A., Robuchon, M., Sterck, L., Peters, A.F., Dittami, S.M., Corre, E., Valero, M., Aury, J.-M., Roze, D., Van de Peer, Y., Bothwell, J., Marais, G.A.B., Coelho, S.M., 2014. A Haploid System of Sex Determination in the Brown Alga *Ectocarpus* sp. *Curr. Biol.* 24, 1945–1957. <https://doi.org/10.1016/j.cub.2014.07.042>
- Allfrey, V.G., Faulkner, R., Mirsky, A.E., 1964. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc. Natl. Acad. Sci.* 51, 786–794. <https://doi.org/10.1073/pnas.51.5.786>
- Arents, G., Moudrianakis, E.N., 1995. The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc. Natl. Acad. Sci. U. S. A.* 92, 11170–11174.
- Arrowsmith, C.H., Bountra, C., Fish, P.V., Lee, K., Schapira, M., 2012. Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discov.* 11, 384–400. <https://doi.org/10.1038/nrd3674>
- Assis, R., Zhou, Q., Bachtrog, D., 2012. Sex-biased transcriptome evolution in *Drosophila*. *Genome Biol. Evol.* 4, 1189–1200. <https://doi.org/10.1093/gbe/evs093>
- Atlasi, Y., Stunnenberg, H.G., 2017. The interplay of epigenetic marks during stem cell differentiation and development. *Nat. Rev. Genet.* 18, 643–658. <https://doi.org/10.1038/nrg.2017.57>
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merckenschlager, M., Fisher, A.G., 2006. Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* 8, 532–538. <https://doi.org/10.1038/ncb1403>
- Bachtrog, D., 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.* 14, 113–124. <https://doi.org/10.1038/nrg3366>
- Bachtrog, D., Kirkpatrick, M., Mank, J.E., McDaniel, S.F., Pires, J.C., Rice, W., Valenzuela, N., 2011. Are all sex chromosomes created equal? *Trends Genet. TIG* 27, 350–357. <https://doi.org/10.1016/j.tig.2011.05.005>
- Bachtrog, D., Mank, J.E., Peichel, C.L., Kirkpatrick, M., Otto, S.P., Ashman, T.-L., Hahn, M.W., Kitano, J., Mayrose, I., Ming, R., Perrin, N., Ross, L., Valenzuela, N., Vamosi, J.C., The Tree of Sex Consortium, 2014. Sex Determination: Why So Many Ways of Doing It? *PLoS Biol.* 12, e1001899. <https://doi.org/10.1371/journal.pbio.1001899>
- Bačovský, V., Houben, A., Kumke, K., Hobza, R., 2019. The distribution of epigenetic histone marks differs between the X and Y chromosomes in *Silene latifolia*. *Planta* 250, 487–494. <https://doi.org/10.1007/s00425-019-03182-7>

- Bai, L., Morozov, A.V., 2010. Gene regulation by nucleosome positioning. *Trends Genet.* 26, 476–483. <https://doi.org/10.1016/j.tig.2010.08.003>
- Banks, J.A., Hickok, L., Webb, M.A., 1993. The Programming of Sexual Phenotype in the Homosporous Fern *Ceratopteris richardii*. *Int. J. Plant Sci.* 154, 522–534.
- Bannister, A.J., Kouzarides, T., 2011. Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. <https://doi.org/10.1038/cr.2011.22>
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., Kouzarides, T., 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124. <https://doi.org/10.1038/35065138>
- Barrett, R.M., Wood, M.A., 2008. Beyond transcription factors: The role of chromatin modifying enzymes in regulating transcription required for memory. *Learn. Mem.* 15, 460–467. <https://doi.org/10.1101/lm.917508>
- Barrett, S.C.H., Hough, J., 2013. Sexual dimorphism in flowering plants. *J. Exp. Bot.* 64, 67–82. <https://doi.org/10.1093/jxb/ers308>
- Bartlett, A., O'Malley, R.C., Huang, S.C., Galli, M., Nery, J.R., Gallavotti, A., Ecker, J.R., 2017. Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* 12, 1659–1672. <https://doi.org/10.1038/nprot.2017.055>
- Bazzicalupo, A.L., Carpentier, F., Otto, S.P., Giraud, T., 2019. Little Evidence of Antagonistic Selection in the Evolutionary Strata of Fungal Mating-Type Chromosomes (*Microbotryum lychnidis-dioicae*). *G3 Bethesda Md* 9, 1987–1998. <https://doi.org/10.1534/g3.119.400242>
- Berec, L., Schembri, P.J., Boukal, D.S., 2005. Sex determination in *Bonellia viridis* (Echiura: Bonelliidae): population dynamics and evolution. *Oikos* 108, 473–484. <https://doi.org/10.1111/j.0030-1299.2005.13350.x>
- Berger, S.L., Kouzarides, T., Shiekhatar, R., Shilatifard, A., 2009. An operational definition of epigenetics. *Genes Dev.* 23, 781–783. <https://doi.org/10.1101/gad.1787609>
- Bergero, R., Charlesworth, D., 2009. The evolution of restricted recombination in sex chromosomes. *Trends Ecol. Evol.* 24, 94–102. <https://doi.org/10.1016/j.tree.2008.09.010>
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L., Lander, E.S., 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326. <https://doi.org/10.1016/j.cell.2006.02.041>
- Bernstein, H., Bernstein, C., Michod, R.E., 2011. Meiosis as an Evolutionary Adaptation for DNA Repair. *DNA Repair.* <https://doi.org/10.5772/25117>
- Berta, P., Hawkins, J.B., Sinclair, A.H., Taylor, A., Griffiths, B.L., Goodfellow, P.N., Fellous, M., 1990. Genetic evidence equating SRY and the testis-determining factor. *Nature* 348, 448–450. <https://doi.org/10.1038/348448a0>
- Beukeboom, L., Perrin, N., 2014. *The Evolution of Sex Determination*. Oxford University Press, Oxford, New York.
- Bianchi, M.E., Agresti, A., 2005. HMG proteins: dynamic players in gene regulation and differentiation. *Curr. Opin. Genet. Dev.* 15, 496–506. <https://doi.org/10.1016/j.gde.2005.08.007>



- Bianchi, M.E., Beltrame, M., 1998. Flexing DNA: HMG-Box Proteins and Their Partners. *Am. J. Hum. Genet.* 63, 1573–1577. <https://doi.org/10.1086/302170>
- Biason-Lauber, A., Konrad, D., Meyer, M., DeBeaufort, C., Schoenle, E.J., 2009. Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the CBX2 gene. *Am. J. Hum. Genet.* 84, 658–663. <https://doi.org/10.1016/j.ajhg.2009.03.016>
- Bird, A., 2002. DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21. <https://doi.org/10.1101/gad.947102>
- Bisoni, L., Battle-Morera, L., Bird, A.P., Suzuki, M., McQueen, H.A., 2005. Female-specific hyperacetylation of histone H4 in the chicken Z chromosome. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 13, 205–214. <https://doi.org/10.1007/s10577-005-1505-4>
- Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S.S., Althoff, F., Araki, T., Arteaga-Vazquez, M.A., Balasubramanian, S., Barry, K., Bauer, D., Boehm, C.R., Briginshaw, L., Caballero-Perez, J., Catarino, B., Chen, F., Chiyoda, S., Chovatia, M., Davies, K.M., Delmans, M., Demura, T., Dierschke, T., Dolan, L., Dorantes-Acosta, A.E., Eklund, D.M., Florent, S.N., Flores-Sandoval, E., Fujiyama, A., Fukuzawa, H., Galik, B., Grimanelli, D., Grimwood, J., Grossniklaus, U., Hamada, T., Haseloff, J., Hetherington, A.J., Higo, A., Hirakawa, Y., Hundley, H.N., Ikeda, Y., Inoue, K., Inoue, S.-I., Ishida, S., Jia, Q., Kakita, M., Kanazawa, T., Kawai, Y., Kawashima, T., Kennedy, M., Kinose, K., Kinoshita, T., Kohara, Y., Koide, E., Komatsu, K., Kopischke, S., Kubo, M., Kyojuka, J., Lagercrantz, U., Lin, S.-S., Lindquist, E., Lipzen, A.M., Lu, C.-W., De Luna, E., Martienssen, R.A., Minamino, N., Mizutani, Masaharu, Mizutani, Miya, Mochizuki, N., Monte, I., Mosher, R., Nagasaki, H., Nakagami, H., Naramoto, S., Nishitani, K., Ohtani, M., Okamoto, T., Okumura, M., Phillips, J., Pollak, B., Reinders, A., Rövekamp, M., Sano, R., Sawa, S., Schmid, M.W., Shirakawa, M., Solano, R., Spunde, A., Suetsugu, N., Sugano, S., Sugiyama, A., Sun, R., Suzuki, Y., Takenaka, M., Takezawa, D., Tomogane, H., Tsuzuki, M., Ueda, T., Umeda, M., Ward, J.M., Watanabe, Y., Yazaki, K., Yokoyama, R., Yoshitake, Y., Yotsui, I., Zachgo, S., Schmutz, J., 2017. Insights into Land Plant Evolution Garnered from the *Marchantia polymorpha* Genome. *Cell* 171, 287-304.e15. <https://doi.org/10.1016/j.cell.2017.09.030>
- Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., Bell, G.W., Otte, A.P., Vidal, M., Gifford, D.K., Young, R.A., Jaenisch, R., 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349–353. <https://doi.org/10.1038/nature04733>
- Brockdorff, N., Turner, B.M., 2015. Dosage Compensation in Mammals. *Cold Spring Harb. Perspect. Biol.* 7, a019406. <https://doi.org/10.1101/cshperspect.a019406>
- Brown, E.J., Bachtrog, D., 2014. The chromatin landscape of *Drosophila*: comparisons between species, sexes, and chromosomes. *Genome Res.* 24, 1125–1137. <https://doi.org/10.1101/gr.172155.114>
- Buenrostro, J., Wu, B., Chang, H., Greenleaf, W., 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.* Ed. Frederick M Ausubel A1 109, 21.29.1-21.29.9. <https://doi.org/10.1002/0471142727.mb2129s109>
- Bull, J.J., 1985. Sex determining mechanisms: An evolutionary perspective. *Experientia* 41, 1285–1296. <https://doi.org/10.1007/BF01952071>

- Bull, J.J., 1978. Sex Chromosomes in Haploid Dioecy: A Unique Contrast to Muller's Theory for Diploid Dioecy. *Am. Nat.* 112, 245–250. <https://doi.org/10.1086/283267>
- Bull, J.J., Vogt, R.C., 1979. Temperature-dependent sex determination in turtles. *Science* 206, 1186–1188. <https://doi.org/10.1126/science.505003>
- Bürger, R., 1999. Evolution of Genetic Variability and the Advantage of Sex and Recombination in Changing Environments. *Genetics* 153, 1055–1069.
- Burki, F., Kaplan, M., Tikhonenkov, D.V., Zlatogursky, V., Minh, B.Q., Radaykina, L.V., Smirnov, A., Mylnikov, A.P., Keeling, P.J., 2016. Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proc. R. Soc. B Biol. Sci.* 283, 20152802. <https://doi.org/10.1098/rspb.2015.2802>
- Burki, F., Shalchian-Tabrizi, K., Minge, M., Skjaveland, A., Nikolaev, S.I., Jakobsen, K.S., Pawlowski, J., 2007. Phylogenomics reshuffles the eukaryotic supergroups. *PloS One* 2, e790. <https://doi.org/10.1371/journal.pone.0000790>
- Burtis, K.C., Baker, B.S., 1989. *Drosophila* doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* 56, 997–1010. [https://doi.org/10.1016/0092-8674\(89\)90633-8](https://doi.org/10.1016/0092-8674(89)90633-8)
- Bustin, M., 1999. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol. Cell. Biol.* 19, 5237–5246. <https://doi.org/10.1128/mcb.19.8.5237>
- Capel, B., 2017. Vertebrate sex determination: evolutionary plasticity of a fundamental switch. *Nat. Rev. Genet.* 18, 675–689. <https://doi.org/10.1038/nrg.2017.60>
- Caramori, G., Ruggeri, P., Mumby, S., Atzeni, F., Adcock, I.M., 2019. Transcription Factors, in: ELS. American Cancer Society, pp. 1–8. <https://doi.org/10.1002/9780470015902.a0005278.pub3>
- Carré, G.-A., Siggers, P., Xipolita, M., Brindle, P., Lutz, B., Wells, S., Greenfield, A., 2018. Loss of p300 and CBP disrupts histone acetylation at the mouse Sry promoter and causes XY gonadal sex reversal. *Hum. Mol. Genet.* 27, 190–198. <https://doi.org/10.1093/hmg/ddx398>
- Catalán, A., Hutter, S., Parsch, J., 2012. Population and sex differences in *Drosophila melanogaster* brain gene expression. *BMC Genomics* 13, 654. <https://doi.org/10.1186/1471-2164-13-654>
- Cavalier-Smith, T., Chao, E.E., Snell, E.A., Berney, C., Fiore-Donno, A.M., Lewis, R., 2014. Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts (animals, fungi, choanozoans) and Amoebozoa. *Mol. Phylogenet. Evol.* 81, 71–85. <https://doi.org/10.1016/j.ympev.2014.08.012>
- Cavalier-Smith, T., Fiore-Donno, A.M., Chao, E., Kudryavtsev, A., Berney, C., Snell, E.A., Lewis, R., 2015. Multigene phylogeny resolves deep branching of Amoebozoa. *Mol. Phylogenet. Evol.* 83, 293–304. <https://doi.org/10.1016/j.ympev.2014.08.011>
- Chadwick, B.P., Willard, H.F., 2002. Cell cycle-dependent localization of macroH2A in chromatin of the inactive X chromosome. *J. Cell Biol.* 157, 1113–1123. <https://doi.org/10.1083/jcb.200112074>

- Chadwick, B.P., Willard, H.F., 2001. Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum. Mol. Genet.* 10, 1101–1113. <https://doi.org/10.1093/hmg/10.10.1101>
- Charlesworth, B., 1991. The evolution of sex chromosomes. *Science* 251, 1030–1033. <https://doi.org/10.1126/science.1998119>
- Charlesworth, B., Jordan, C.Y., Charlesworth, D., 2014. The evolutionary dynamics of sexually antagonistic mutations in pseudoautosomal regions of sex chromosomes. *Evol. Int. J. Org. Evol.* 68, 1339–1350. <https://doi.org/10.1111/evo.12364>
- Charlesworth, D., 2017. Evolution of recombination rates between sex chromosomes. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160456. <https://doi.org/10.1098/rstb.2016.0456>
- Charlesworth, D., 2016. Plant Sex Chromosomes. *Annu. Rev. Plant Biol.* 67, 397–420. <https://doi.org/10.1146/annurev-arplant-043015-111911>
- Charlesworth, D., Charlesworth, B., 1980. Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genet. Res.* 35, 205–214. <https://doi.org/10.1017/S0016672300014051>
- Charlesworth, D., Charlesworth, B., Marais, G., 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95, 118–128. <https://doi.org/10.1038/sj.hdy.6800697>
- Charrier, B., Coelho, S.M., Le Bail, A., Tonon, T., Michel, G., Potin, P., Kloareg, B., Boyen, C., Peters, A.F., Cock, J.M., 2008. Development and physiology of the brown alga *Ectocarpus siliculosus*: two centuries of research. *New Phytol.* 177, 319–332. <https://doi.org/10.1111/j.1469-8137.2007.02304.x>
- Chassot, A.-A., Ranc, F., Gregoire, E.P., Roepers-Gajadien, H.L., Taketo, M.M., Camerino, G., de Rooij, D.G., Schedl, A., Chaboissier, M.-C., 2008. Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. *Hum. Mol. Genet.* 17, 1264–1277. <https://doi.org/10.1093/hmg/ddn016>
- Chen, T., Li, E., 2006. Establishment and Maintenance of DNA Methylation Patterns in Mammals, in: Doerfler, W., Böhm, P. (Eds.), *DNA Methylation: Basic Mechanisms, Current Topics in Microbiology and Immunology*. Springer, Berlin, Heidelberg, pp. 179–201. [https://doi.org/10.1007/3-540-31390-7\\_6](https://doi.org/10.1007/3-540-31390-7_6)
- Cho, S., Huang, Z.Y., Zhang, J., 2007. Sex-Specific Splicing of the Honeybee doublesex Gene Reveals 300 Million Years of Evolution at the Bottom of the Insect Sex-Determination Pathway. *Genetics* 177, 1733–1741. <https://doi.org/10.1534/genetics.107.078980>
- Cline, T.W., Meyer, B.J., 1996. Vive la différence: males vs females in flies vs worms. *Annu. Rev. Genet.* 30, 637–702. <https://doi.org/10.1146/annurev.genet.30.1.637>
- Cock, J.M., Sterck, L., Rouzé, P., Scornet, D., Allen, A.E., Amoutzias, G., Anthouard, V., Artiguenave, F., Aury, J.-M., Badger, J.H., Beszteri, B., Billiau, K., Bonnet, E., Bothwell, J.H., Bowler, C., Boyen, C., Brownlee, C., Carrano, C.J., Charrier, B., Cho, G.Y., Coelho, S.M., Collén, J., Corre, E., Da Silva, C., Delage, L., Delaroque, N., Dittami, S.M., Doulebeau, S., Elias, M., Farnham, G., Gachon, C.M.M., Gschloessl, B., Heesch, S., Jabbari, K., Jubin, C., Kawai, H., Kimura, K., Kloareg, B., Küpper, F.C., Lang, D., Le Bail, A., Leblanc, C., Lerouge, P., Lohr, M., Lopez, P.J., Martens, C., Maumus, F., Michel, G., Miranda-Saavedra, D., Morales, J., Moreau, H., Motomura, T., Nagasato, C., Napoli, C.A., Nelson, D.R., Nyvall-Collén, P., Peters, A.F., Pommier, C., Potin, P., Poulain, J., Quesneville, H., Read, B., Rensing, S.A., Ritter, A., Rousvoal, S., Samanta, M., Samson, G., Schroeder, D.C., Ségurens, B., Strittmatter, M., Tonon,

- T., Tregear, J.W., Valentin, K., von Dassow, P., Yamagishi, T., Van de Peer, Y., Wincker, P., 2010. The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature* 465, 617–621. <https://doi.org/10.1038/nature09016>
- Coe, W.R., 1936. Sexual phases in *Crepidula*. *J. Exp. Zool.* 72, 455–477. <https://doi.org/10.1002/jez.1400720306>
- Coelho, S.M., Godfroy, O., Arun, A., Le Corguillé, G., Peters, A.F., Cock, J.M., 2011. OUROBOROS is a master regulator of the gametophyte to sporophyte life cycle transition in the brown alga *Ectocarpus*. *Proc. Natl. Acad. Sci. U. S. A.* 108, 11518–11523. <https://doi.org/10.1073/pnas.1102274108>
- Coelho, S.M., Gueno, J., Lipinska, A.P., Cock, J.M., Umen, J.G., 2018. UV Chromosomes and Haploid Sexual Systems. *Trends Plant Sci.* 23, 794–807. <https://doi.org/10.1016/j.tplants.2018.06.005>
- Cormier, A., Avia, K., Sterck, L., Derrien, T., Wucher, V., Andres, G., Monsoor, M., Godfroy, O., Lipinska, A., Perrineau, M.-M., Van De Peer, Y., Hitte, C., Corre, E., Coelho, S.M., Cock, J.M., 2017. Re-annotation, improved large-scale assembly and establishment of a catalogue of noncoding loci for the genome of the model brown alga *Ectocarpus*. *New Phytol.* 214, 219–232. <https://doi.org/10.1111/nph.14321>
- De Hoff, P.L., Ferris, P., Olson, B.J.S.C., Miyagi, A., Geng, S., Umen, J.G., 2013. Species and population level molecular profiling reveals cryptic recombination and emergent asymmetry in the dimorphic mating locus of *C. reinhardtii*. *PLoS Genet.* 9, e1003724. <https://doi.org/10.1371/journal.pgen.1003724>
- De Rop, V., Padeganeh, A., Maddox, P.S., 2012. CENP-A: the key player behind centromere identity, propagation, and kinetochore assembly. *Chromosoma* 121, 527–538. <https://doi.org/10.1007/s00412-012-0386-5>
- Deaton, A.M., Bird, A., 2011. CpG islands and the regulation of transcription. *Genes Dev.* 25, 1010–1022. <https://doi.org/10.1101/gad.2037511>
- Decarpentrie, F., Vernet, N., Mahadevaiah, S.K., Longepied, G., Streichemberger, E., Aknin-Seifer, I., Ojarikre, O.A., Burgoyne, P.S., Metzler-Guillemain, C., Mitchell, M.J., 2012. Human and mouse ZFY genes produce a conserved testis-specific transcript encoding a zinc finger protein with a short acidic domain and modified transactivation potential. *Hum. Mol. Genet.* 21, 2631–2645. <https://doi.org/10.1093/hmg/dds088>
- Denis, G.V., McComb, M.E., Faller, D.V., Sinha, A., Romesser, P.B., Costello, C.E., 2006. Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines. *J. Proteome Res.* 5, 502–511. <https://doi.org/10.1021/pr050430u>
- DiNapoli, L., Capel, B., 2008. SRY and the standoff in sex determination. *Mol. Endocrinol. Baltim. Md* 22, 1–9. <https://doi.org/10.1210/me.2007-0250>
- Doyen, C.-M., An, W., Angelov, D., Bondarenko, V., Mietton, F., Studitsky, V.M., Hamiche, A., Roeder, R.G., Bouvet, P., Dimitrov, S., 2006a. Mechanism of polymerase II transcription repression by the histone variant macroH2A. *Mol. Cell. Biol.* 26, 1156–1164. <https://doi.org/10.1128/MCB.26.3.1156-1164.2006>
- Doyen, C.-M., Montel, F., Gautier, T., Menoni, H., Claudet, C., Delacour-Larose, M., Angelov, D., Hamiche, A., Bednar, J., Faivre-Moskalenko, C., Bouvet, P., Dimitrov, S., 2006b.

- Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. *EMBO J.* 25, 4234–4244. <https://doi.org/10.1038/sj.emboj.7601310>
- Eid, W., Opitz, L., Biason-Lauber, A., 2015. Genome-wide identification of CBX2 targets: insights in the human sex development network. *Mol. Endocrinol. Baltim. Md* 29, 247–257. <https://doi.org/10.1210/me.2014-1339>
- Elango, N., Hunt, B.G., Goodisman, M.A.D., Yi, S.V., 2009. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc. Natl. Acad. Sci. U. S. A.* 106, 11206–11211. <https://doi.org/10.1073/pnas.0900301106>
- El-Kasti, M.M., Wells, T., Carter, D.A., 2012. A novel long-range enhancer regulates postnatal expression of *Zeb2*: implications for Mowat-Wilson syndrome phenotypes. *Hum. Mol. Genet.* 21, 5429–5442. <https://doi.org/10.1093/hmg/dds389>
- Ellegren, H., Parsch, J., 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat. Rev. Genet.* 8, 689–698. <https://doi.org/10.1038/nrg2167>
- Ellis, N., Taylor, A., Bengtsson, B.O., Kidd, J., Rogers, J., Fellow, P.G., 1990. Population structure of the human pseudoautosomal boundary. *Nature* 344, 663–665. <https://doi.org/10.1038/344663a0>
- Eppley, S.M., Jesson, L.K., 2008. Moving to mate: the evolution of separate and combined sexes in multicellular organisms. *J. Evol. Biol.* 21, 727–736. <https://doi.org/10.1111/j.1420-9101.2008.01524.x>
- Ernst, J., Kellis, M., 2010. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* 28, 817–825. <https://doi.org/10.1038/nbt.1662>
- Ezaz, T., Stiglec, R., Veyrunes, F., Marshall Graves, J.A., 2006. Relationships between Vertebrate ZW and XY Sex Chromosome Systems. *Curr. Biol.* 16, R736–R743. <https://doi.org/10.1016/j.cub.2006.08.021>
- Ferris, P., Olson, B.J.S.C., De Hoff, P.L., Douglass, S., Casero, D., Prochnik, S., Geng, S., Rai, R., Grimwood, J., Schmutz, J., Nishii, I., Hamaji, T., Nozaki, H., Pellegrini, M., Umen, J.G., 2010. Evolution of an expanded sex-determining locus in *Volvox*. *Science* 328, 351–354. <https://doi.org/10.1126/science.1186222>
- Filatov, D.A., 2005. Evolutionary History of *Silene latifolia* Sex Chromosomes Revealed by Genetic Mapping of Four Genes. *Genetics* 170, 975–979. <https://doi.org/10.1534/genetics.104.037069>
- Flaquer, A., Rappold, G.A., Wienker, T.F., Fischer, C., 2008. The human pseudoautosomal regions: a review for genetic epidemiologists. *Eur. J. Hum. Genet.* 16, 771–779. <https://doi.org/10.1038/ejhg.2008.63>
- Francisco, F.O., Lemos, B., 2014. How Do Y-Chromosomes Modulate Genome-Wide Epigenetic States: Genome Folding, Chromatin Sinks, and Gene Expression. *J. Genomics* 2, 94–103. <https://doi.org/10.7150/jgen.8043>
- Fyodorov, D.V., Zhou, B.-R., Skoultchi, A.I., Bai, Y., 2018. Emerging roles of linker histones in regulating chromatin structure and function. *Nat. Rev. Mol. Cell Biol.* 19, 192–206. <https://doi.org/10.1038/nrm.2017.94>

- Garcia-Moreno, S.A., Plebanek, M.P., Capel, B., 2018. Epigenetic regulation of male fate commitment from an initially bipotential system. *Mol. Cell. Endocrinol.* 468, 19–30. <https://doi.org/10.1016/j.mce.2018.01.009>
- Ge, C., Ye, J., Zhang, H., Zhang, Y., Sun, W., Sang, Y., Capel, B., Qian, G., 2017. *Dmrt1* induces the male pathway in a turtle species with temperature-dependent sex determination. *Dev. Camb. Engl.* 144, 2222–2233. <https://doi.org/10.1242/dev.152033>
- Gelbart, M.E., Kuroda, M.I., 2009. *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* 136, 1399–1410. <https://doi.org/10.1242/dev.029645>
- Gempe, T., Beye, M., 2011. Function and evolution of sex determination mechanisms, genes and pathways in insects. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 33, 52–60. <https://doi.org/10.1002/bies.201000043>
- Gerlitz, G., Hock, R., Ueda, T., Bustin, M., 2009. The dynamics of HMG protein–chromatin interactions in living cells This paper is one of a selection of papers published in this Special Issue, entitled CSBMCB’s 51st Annual Meeting – Epigenetics and Chromatin Dynamics, and has undergone the Journal’s usual peer review process. *Biochem. Cell Biol.* 87, 127–137. <https://doi.org/10.1139/O08-110>
- Godoy, M., Franco-Zorrilla, J.M., Pérez-Pérez, J., Oliveros, J.C., Lorenzo, O., Solano, R., 2011. Improved protein-binding microarrays for the identification of DNA-binding specificities of transcription factors. *Plant J. Cell Mol. Biol.* 66, 700–711. <https://doi.org/10.1111/j.1365-313X.2011.04519.x>
- Godwin, J., 2009. Social determination of sex in reef fishes. *Semin. Cell Dev. Biol., Environmental Regulation of Sex Dtermination in Vertebrates* 20, 264–270. <https://doi.org/10.1016/j.semcdb.2008.12.003>
- Godwin, J., Luckenbach, J.A., Borski, R.J., 2003. Ecology meets endocrinology: environmental sex determination in fishes. *Evol. Dev.* 5, 40–49. <https://doi.org/10.1046/j.1525-142x.2003.03007.x>
- Goldberg, A.D., Allis, C.D., Bernstein, E., 2007. Epigenetics: A Landscape Takes Shape. *Cell* 128, 635–638. <https://doi.org/10.1016/j.cell.2007.02.006>
- Goldman, T.D., Arbeitman, M.N., 2007. Genomic and Functional Studies of *Drosophila* Sex Hierarchy Regulated Gene Expression in Adult Head and Nervous System Tissues. *PLOS Genet.* 3, e216. <https://doi.org/10.1371/journal.pgen.0030216>
- Goll, M.G., Bestor, T.H., 2005. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* 74, 481–514. <https://doi.org/10.1146/annurev.biochem.74.010904.153721>
- Gonen, N., Lovell-Badge, R., 2019. Chapter Eight - The regulation of *Sox9* expression in the gonad, in: Capel, B. (Ed.), *Current Topics in Developmental Biology, Sex Determination in Vertebrates*. Academic Press, pp. 223–252. <https://doi.org/10.1016/bs.ctdb.2019.01.004>
- Gordân, R., Shen, N., Dror, I., Zhou, T., Horton, J., Rohs, R., Bulyk, M.L., 2013. Genomic regions flanking E-box binding sites influence DNA binding specificity of bHLH transcription factors through DNA shape. *Cell Rep.* 3, 1093–1104. <https://doi.org/10.1016/j.celrep.2013.03.014>
- Gouyon, 1999. Sex: a pluralist approach includes species selection. (One step beyond and it’s good.). *J. Evol. Biol.* 12, 1029–1030. <https://doi.org/10.1046/j.1420-9101.1999.00130.x>

- Graham, P., Penn, J.K.M., Schedl, P., 2003. Masters change, slaves remain. *BioEssays* 25, 1–4. <https://doi.org/10.1002/bies.10207>
- Grath, S., Parsch, J., 2016. Sex-Biased Gene Expression. *Annu. Rev. Genet.* 50, 29–44. <https://doi.org/10.1146/annurev-genet-120215-035429>
- Greer, E.L., Blanco, M.A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., Hsu, C.-H., Aravind, L., He, C., Shi, Y., 2015. DNA Methylation on N6-Adenine in *C. elegans*. *Cell* 161, 868–878. <https://doi.org/10.1016/j.cell.2015.04.005>
- Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P., Lovell-Badge, R., 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* 346, 245–250. <https://doi.org/10.1038/346245a0>
- Guertin, M.J., Martins, A.L., Siepel, A., Lis, J.T., 2012. Accurate prediction of inducible transcription factor binding intensities in vivo. *PLoS Genet.* 8, e1002610. <https://doi.org/10.1371/journal.pgen.1002610>
- Guo, Y., Mahony, S., Gifford, D.K., 2012. High Resolution Genome Wide Binding Event Finding and Motif Discovery Reveals Transcription Factor Spatial Binding Constraints. *PLOS Comput. Biol.* 8, e1002638. <https://doi.org/10.1371/journal.pcbi.1002638>
- Guo, Y., Tian, K., Zeng, H., Guo, X., Gifford, D.K., 2017. A novel k-mer set memory (KSM) motif representation improves regulatory variant prediction. *bioRxiv* 130815. <https://doi.org/10.1101/130815>
- Haag, E.S., Doty, A.V., 2005. Sex determination across evolution: connecting the dots. *PLoS Biol.* 3, e21. <https://doi.org/10.1371/journal.pbio.0030021>
- Haber, J.E., 2012. Mating-Type Genes and MAT Switching in *Saccharomyces cerevisiae*. *Genetics* 191, 33–64. <https://doi.org/10.1534/genetics.111.134577>
- Hamaji, T., Mogi, Y., Ferris, P.J., Mori, T., Miyagishima, S., Kabeya, Y., Nishimura, Y., Toyoda, A., Noguchi, H., Fujiyama, A., Olson, B.J.S.C., Marriage, T.N., Nishii, I., Umen, J.G., Nozaki, H., 2016. Sequence of the *Gonium pectorale* Mating Locus Reveals a Complex and Dynamic History of Changes in Volvocine Algal Mating Haplotypes. *G3 GenesGenomesGenetics* 6, 1179–1189. <https://doi.org/10.1534/g3.115.026229>
- Harley, V.R., Lovell-Badge, R., Goodfellow, P.N., 1994. Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res.* 22, 1500–1501.
- Harrison, P.W., Wright, A.E., Zimmer, F., Dean, R., Montgomery, S.H., Pointer, M.A., Mank, J.E., 2015. Sexual selection drives evolution and rapid turnover of male gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 112, 4393–4398. <https://doi.org/10.1073/pnas.1501339112>
- Hattori, R.S., Gould, R.J., Fujioka, T., Saito, T., Kurita, J., Strüssmann, C.A., Yokota, M., Watanabe, S., 2007. Temperature-Dependent Sex Determination in Hd-rR Medaka *Oryzias latipes*: Gender Sensitivity, Thermal Threshold, Critical Period, and DMRT1 Expression Profile. *Sex. Dev.* 1, 138–146. <https://doi.org/10.1159/000100035>
- Hattori, R.S., Murai, Y., Oura, M., Masuda, S., Majhi, S.K., Sakamoto, T., Fernandino, J.I., Somoza, G.M., Yokota, M., Strüssmann, C.A., 2012. A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination. *Proc. Natl. Acad. Sci. U. S. A.* 109, 2955–2959. <https://doi.org/10.1073/pnas.1018392109>

- He, Q., Johnston, J., Zeitlinger, J., 2015. ChIP-nexus enables improved detection of in vivo transcription factor binding footprints. *Nat. Biotechnol.* 33, 395–401. <https://doi.org/10.1038/nbt.3121>
- Henikoff, S., Ahmad, K., 2005. Assembly of Variant Histones into Chromatin. *Annu. Rev. Cell Dev. Biol.* 21, 133–153. <https://doi.org/10.1146/annurev.cellbio.21.012704.133518>
- Herpin, A., Schartl, M., 2015. Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. *EMBO Rep.* 16, 1260–1274. <https://doi.org/10.15252/embr.201540667>
- Herpin, A., Schartl, M., 2011. Dmrt1 genes at the crossroads: a widespread and central class of sexual development factors in fish. *FEBS J.* 278, 1010–1019. <https://doi.org/10.1111/j.1742-4658.2011.08030.x>
- Herpin, A., Schartl, M., 2008. Regulatory putsches create new ways of determining sexual development. *EMBO Rep.* 9, 966–968. <https://doi.org/10.1038/embo.2008.182>
- Hildreth, P.E., 1965. Doublesex, a Recessive Gene That Transforms Both Males and Females of *Drosophila* into Intersexes. *Genetics* 51, 659–678.
- Hoshijima, K., Inoue, K., Higuchi, I., Sakamoto, H., Shimura, Y., 1991. Control of doublesex alternative splicing by transformer and transformer-2 in *Drosophila*. *Science* 252, 833–836. <https://doi.org/10.1126/science.1902987>
- Hotchkiss, R.D., 1948. The Quantitative Separation of Purines, Pyrimidines, and Nucleosides by Paper Chromatography. *J. Biol. Chem.* 175, 315–332.
- Huang, H., Sabari, B.R., Garcia, B.A., Allis, C.D., Zhao, Y., 2014. SnapShot: histone modifications. *Cell* 159, 458–458.e1. <https://doi.org/10.1016/j.cell.2014.09.037>
- Hughes, J.F., Skaletsky, H., Koutseva, N., Pyntikova, T., Page, D.C., 2015. Sex chromosome-to-autosome transposition events counter Y-chromosome gene loss in mammals. *Genome Biol.* 16, 104. <https://doi.org/10.1186/s13059-015-0667-4>
- Huylmans, A.K., Parsch, J., 2014. Population- and sex-biased gene expression in the excretion organs of *Drosophila melanogaster*. *G3 Bethesda Md* 4, 2307–2315. <https://doi.org/10.1534/g3.114.013417>
- Idnurm, A., Walton, F.J., Floyd, A., Heitman, J., 2008. Identification of the sex genes in an early diverged fungus. *Nature* 451, 193–196. <https://doi.org/10.1038/nature06453>
- Immler, S., Otto, S.P., 2015. The evolution of sex chromosomes in organisms with separate haploid sexes. *Evol. Int. J. Org. Evol.* 69, 694–708. <https://doi.org/10.1111/evo.12602>
- Itoh, Y., Kampf, K., Arnold, A.P., 2011. Possible differences in the two Z chromosomes in male chickens and evolution of MHM sequences in Galliformes. *Chromosoma* 120, 587–598. <https://doi.org/10.1007/s00412-011-0333-x>
- Iyer, L.M., Zhang, D., Aravind, L., 2016. Adenine methylation in eukaryotes: Apprehending the complex evolutionary history and functional potential of an epigenetic modification. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 38, 27–40. <https://doi.org/10.1002/bies.201500104>
- Jakob, S., Lovell-Badge, R., 2011. Sex determination and the control of Sox9 expression in mammals. *FEBS J.* 1002–1009. [https://doi.org/10.1111/j.1742-4658.2011.08029.x@10.1002/\(ISSN\)1742-4658\(CAT\)FreeReviewContent\(VI\)Reviews1011](https://doi.org/10.1111/j.1742-4658.2011.08029.x@10.1002/(ISSN)1742-4658(CAT)FreeReviewContent(VI)Reviews1011)



- Jameson, S.A., Natarajan, A., Cool, J., DeFalco, T., Maatouk, D.M., Mork, L., Munger, S.C., Capel, B., 2012. Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet.* 8, e1002575. <https://doi.org/10.1371/journal.pgen.1002575>
- Janzen, F.J., Phillips, P.C., 2006. Exploring the evolution of environmental sex determination, especially in reptiles. *J. Evol. Biol.* 19, 1775–1784. <https://doi.org/10.1111/j.1420-9101.2006.01138.x>
- Jordan, B.K., Shen, J.H.-C., Olaso, R., Ingraham, H.A., Vilain, E., 2003. Wnt4 overexpression disrupts normal testicular vasculature and inhibits testosterone synthesis by repressing steroidogenic factor 1/beta-catenin synergy. *Proc. Natl. Acad. Sci. U. S. A.* 100, 10866–10871. <https://doi.org/10.1073/pnas.1834480100>
- Kadonaga, J.T., 1998. Eukaryotic Transcription: An Interlaced Network of Transcription Factors and Chromatin-Modifying Machines. *Cell* 92, 307–313. [https://doi.org/10.1016/S0092-8674\(00\)80924-1](https://doi.org/10.1016/S0092-8674(00)80924-1)
- Kamachi, Y., Kondoh, H., 2013. Sox proteins: regulators of cell fate specification and differentiation. *Development* 140, 4129–4144. <https://doi.org/10.1242/dev.091793>
- Kamakaka, R.T., Biggins, S., 2005. Histone variants: deviants? *Genes Dev.* 19, 295–310. <https://doi.org/10.1101/gad.1272805>
- Kashimada, K., Koopman, P., 2010. Sry: the master switch in mammalian sex determination. *Development* 137, 3921–3930. <https://doi.org/10.1242/dev.048983>
- Kasinsky, H.E., Lewis, J.D., Dacks, J.B., Ausió, J., 2001. Origin of H1 linker histones. *FASEB J.* 15, 34–42. <https://doi.org/10.1096/fj.00-0237rev>
- Katoh-Fukui, Y., Miyabayashi, K., Komatsu, T., Owaki, A., Baba, T., Shima, Y., Kidokoro, T., Kanai, Y., Schedl, A., Wilhelm, D., Koopman, P., Okuno, Y., Morohashi, K., 2012. Cbx2, a polycomb group gene, is required for Sry gene expression in mice. *Endocrinology* 153, 913–924. <https://doi.org/10.1210/en.2011-1055>
- Katoh-Fukui, Y., Tsuchiya, R., Shiroishi, T., Nakahara, Y., Hashimoto, N., Noguchi, K., Higashinakagawa, T., 1998. Male-to-female sex reversal in M33 mutant mice. *Nature* 393, 688–692. <https://doi.org/10.1038/31482>
- Kent, J., Wheatley, S.C., Andrews, J.E., Sinclair, A.H., Koopman, P., 1996. A male-specific role for SOX9 in vertebrate sex determination. *Dev. Camb. Engl.* 122, 2813–2822.
- Kidokoro, T., Matoba, S., Hiramatsu, R., Fujisawa, M., Kanai-Azuma, M., Taya, C., Kurohmaru, M., Kawakami, H., Hayashi, Y., Kanai, Y., Yonekawa, H., 2005. Influence on spatiotemporal patterns of a male-specific Sox9 activation by ectopic Sry expression during early phases of testis differentiation in mice. *Dev. Biol.* 278, 511–525. <https://doi.org/10.1016/j.ydbio.2004.11.006>
- Kim, J., Guermah, M., McGinty, R.K., Lee, J.-S., Tang, Z., Milne, T.A., Shilatifard, A., Muir, T.W., Roeder, R.G., 2009. RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* 137, 459–471. <https://doi.org/10.1016/j.cell.2009.02.027>
- Kim, Y., Capel, B., 2006. Balancing the bipotential gonad between alternative organ fates: A new perspective on an old problem. *Dev. Dyn.* 235, 2292–2300. <https://doi.org/10.1002/dvdy.20894>

- Kim, Y., Kobayashi, A., Sekido, R., DiNapoli, L., Brennan, J., Chaboissier, M.-C., Poulat, F., Behringer, R.R., Lovell-Badge, R., Capel, B., 2006. Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. *PLoS Biol.* 4, e187. <https://doi.org/10.1371/journal.pbio.0040187>
- Kondrashov, A.S., 1988. Deleterious mutations and the evolution of sexual reproduction. *Nature* 336, 435–440. <https://doi.org/10.1038/336435a0>
- Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., Lovell-Badge, R., 1991. Male development of chromosomally female mice transgenic for Sry. *Nature* 351, 117–121. <https://doi.org/10.1038/351117a0>
- Koprinarova, M., Schneckeburger, M., Diederich, M., 2016. Role of Histone Acetylation in Cell Cycle Regulation. *Curr. Top. Med. Chem.* 16, 732–744. <https://doi.org/10.2174/1568026615666150825140822>
- Kornberg, R.D., Thomas, J.O., 1974. Chromatin structure; oligomers of the histones. *Science* 184, 865–868. <https://doi.org/10.1126/science.184.4139.865>
- Kouzarides, T., 2007. Chromatin modifications and their function. *Cell* 128, 693–705. <https://doi.org/10.1016/j.cell.2007.02.005>
- Kozłowski, M., Corujo, D., Hothorn, M., Guberovic, I., Mandemaker, I.K., Blessing, C., Sporn, J., Gutierrez-Triana, A., Smith, R., Portmann, T., Treier, M., Scheffzek, K., Huet, S., Timinszky, G., Buschbeck, M., Ladurner, A.G., 2018. MacroH2A histone variants limit chromatin plasticity through two distinct mechanisms. *EMBO Rep.* 19, e44445. <https://doi.org/10.15252/embr.201744445>
- Kucharski, R., Maleszka, J., Foret, S., Maleszka, R., 2008. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319, 1827–1830. <https://doi.org/10.1126/science.1153069>
- Kurjan, J., 1985. Alpha-factor structural gene mutations in *Saccharomyces cerevisiae*: effects on alpha-factor production and mating. *Mol. Cell. Biol.* 5, 787–796. <https://doi.org/10.1128/MCB.5.4.787>
- Kuroki, S., Matoba, S., Akiyoshi, M., Matsumura, Y., Miyachi, H., Mise, N., Abe, K., Ogura, A., Wilhelm, D., Koopman, P., Nozaki, M., Kanai, Y., Shinkai, Y., Tachibana, M., 2013. Epigenetic regulation of mouse sex determination by the histone demethylase Jmjd1a. *Science* 341, 1106–1109. <https://doi.org/10.1126/science.1239864>
- Lambeth, L.S., Raymond, C.S., Roeszler, K.N., Kuroiwa, A., Nakata, T., Zarkower, D., Smith, C.A., 2014. Over-expression of DMRT1 induces the male pathway in embryonic chicken gonads. *Dev. Biol.* 389, 160–172. <https://doi.org/10.1016/j.ydbio.2014.02.012>
- Lanzuolo, C., Orlando, V., 2012. Memories from the polycomb group proteins. *Annu. Rev. Genet.* 46, 561–589. <https://doi.org/10.1146/annurev-genet-110711-155603>
- Lapointe, E., Boerboom, D., 2011. WNT signaling and the regulation of ovarian steroidogenesis. *Front. Biosci. Sch. Ed.* 3, 276–285. <https://doi.org/10.2741/s151>
- Laudet, V., Stehelin, D., Clevers, H., 1993. Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Res.* 21, 2493–2501. <https://doi.org/10.1093/nar/21.10.2493>
- Lee, J.-H., Lin, H., Joo, S., Goodenough, U., 2008. Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family. *Cell* 133, 829–840. <https://doi.org/10.1016/j.cell.2008.04.028>

- Lee, J.-S., Smith, E., Shilatifard, A., 2010. The language of histone crosstalk. *Cell* 142, 682–685. <https://doi.org/10.1016/j.cell.2010.08.011>
- Lehtonen, J., Jennions, M.D., Kokko, H., 2012. The many costs of sex. *Trends Ecol. Evol.* 27, 172–178. <https://doi.org/10.1016/j.tree.2011.09.016>
- Lemos, B., Branco, A.T., Hartl, D.L., 2010. Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict. *Proc. Natl. Acad. Sci. U. S. A.* 107, 15826–15831. <https://doi.org/10.1073/pnas.1010383107>
- Lenormand, T., 2003. The Evolution of Sex Dimorphism in Recombination. *Genetics* 163, 811–822.
- Lesch, B.J., Silber, S.J., McCarrey, J.R., Page, D.C., 2016. Parallel evolution of male germline epigenetic poisoning and somatic development in animals. *Nat. Genet.* 48, 888–894. <https://doi.org/10.1038/ng.3591>
- Lewis, W.M., 1987. The cost of sex. *Experientia. Suppl.* 55, 33–57. [https://doi.org/10.1007/978-3-0348-6273-8\\_2](https://doi.org/10.1007/978-3-0348-6273-8_2)
- Li, B., Carey, M., Workman, J.L., 2007. The role of chromatin during transcription. *Cell* 128, 707–719. <https://doi.org/10.1016/j.cell.2007.01.015>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, L., Zhang, J.A., Dose, M., Kueh, H.Y., Mosadeghi, R., Gounari, F., Rothenberg, E.V., 2013. A far downstream enhancer for murine *Bcl11b* controls its T-cell specific expression. *Blood* 122, 902–911. <https://doi.org/10.1182/blood-2012-08-447839>
- Li, M., Sun, Y., Zhao, J., Shi, H., Zeng, S., Ye, K., Jiang, D., Zhou, L., Sun, L., Tao, W., Nagahama, Y., Kocher, T.D., Wang, D., 2015. A Tandem Duplicate of Anti-Müllerian Hormone with a Missense SNP on the Y Chromosome Is Essential for Male Sex Determination in Nile Tilapia, *Oreochromis niloticus*. *PLOS Genet.* 11, e1005678. <https://doi.org/10.1371/journal.pgen.1005678>
- Liang, G., Lin, J.C.Y., Wei, V., Yoo, C., Cheng, J.C., Nguyen, C.T., Weisenberger, D.J., Egger, G., Takai, D., Gonzales, F.A., Jones, P.A., 2004. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7357–7362. <https://doi.org/10.1073/pnas.0401866101>
- Lin, X., Tirichine, L., Bowler, C., 2012. Protocol: Chromatin immunoprecipitation (ChIP) methodology to investigate histone modifications in two model diatom species. *Plant Methods* 8, 48. <https://doi.org/10.1186/1746-4811-8-48>
- Lin, Z., Wu, W.-S., Liang, H., Woo, Y., Li, W.-H., 2010. The spatial distribution of cis regulatory elements in yeast promoters and its implications for transcriptional regulation. *BMC Genomics* 11, 581. <https://doi.org/10.1186/1471-2164-11-581>
- Lipinska, A., Cormier, A., Luthringer, R., Peters, A.F., Corre, E., Gachon, C.M.M., Cock, J.M., Coelho, S.M., 2015. Sexual Dimorphism and the Evolution of Sex-Biased Gene Expression in the Brown Alga *Ectocarpus*. *Mol. Biol. Evol.* 32, 1581–1597. <https://doi.org/10.1093/molbev/msv049>

- Lipinska, A.P., Toda, N.R.T., Heesch, S., Peters, A.F., Cock, J.M., Coelho, S.M., 2017. Multiple gene movements into and out of haploid sex chromosomes. *Genome Biol.* 18, 104. <https://doi.org/10.1186/s13059-017-1201-7>
- Liu, C.-F., Angelozzi, M., Haseeb, A., Lefebvre, V., 2018. SOX9 is dispensable for the initiation of epigenetic remodeling and the activation of marker genes at the onset of chondrogenesis. *Development* 145, dev164459. <https://doi.org/10.1242/dev.164459>
- Liu, X., Noll, D.M., Lieb, J.D., Clarke, N.D., 2005. DIP-chip: Rapid and accurate determination of DNA-binding specificity. *Genome Res.* 15, 421–427. <https://doi.org/10.1101/gr.3256505>
- Lovell-Badge, R., Robertson, E., 1990. XY female mice resulting from a heritable mutation in the primary testis-determining gene, Tdy. *Development* 109, 635–646.
- Lucchesi, J.C., Kelly, W.G., Panning, B., 2005. Chromatin remodeling in dosage compensation. *Annu. Rev. Genet.* 39, 615–651. <https://doi.org/10.1146/annurev.genet.39.073003.094210>
- Luthringer, R., Lipinska, A.P., Roze, D., Cormier, A., Macaisne, N., Peters, A.F., Cock, J.M., Coelho, S.M., 2015. The Pseudoautosomal Regions of the U/V Sex Chromosomes of the Brown Alga *Ectocarpus* Exhibit Unusual Features. *Mol. Biol. Evol.* 32, 2973–2985. <https://doi.org/10.1093/molbev/msv173>
- Maatouk, D.M., DiNapoli, L., Alvers, A., Parker, K.L., Taketo, M.M., Capel, B., 2008. Stabilization of beta-catenin in XY gonads causes male-to-female sex-reversal. *Hum. Mol. Genet.* 17, 2949–2955. <https://doi.org/10.1093/hmg/ddn193>
- Maatouk, D.M., Natarajan, A., Shibata, Y., Song, L., Crawford, G.E., Ohler, U., Capel, B., 2017. Genome-wide identification of regulatory elements in Sertoli cells. *Dev. Camb. Engl.* 144, 720–730. <https://doi.org/10.1242/dev.142554>
- Macaisne, N., Liu, F., Scornet, D., Peters, A.F., Lipinska, A., Perrineau, M.-M., Henry, A., Strittmatter, M., Coelho, S.M., Cock, J.M., 2017. The *Ectocarpus* IMMEDIATE UPRIGHT gene encodes a member of a novel family of cysteine-rich proteins with an unusual distribution across the eukaryotes. *Development* 144, 409–418. <https://doi.org/10.1242/dev.141523>
- Mank, J.E., Hultin-Rosenberg, L., Webster, M.T., Ellegren, H., 2008. The unique genomic properties of sex-biased genes: Insights from avian microarray data. *BMC Genomics* 9, 148. <https://doi.org/10.1186/1471-2164-9-148>
- Mank, J.E., Nam, K., Brunström, B., Ellegren, H., 2010. Ontogenetic complexity of sexual dimorphism and sex-specific selection. *Mol. Biol. Evol.* 27, 1570–1578. <https://doi.org/10.1093/molbev/msq042>
- Masuyama, H., Yamada, M., Kamei, Y., Fujiwara-Ishikawa, T., Todo, T., Nagahama, Y., Matsuda, M., 2012. Dmrt1 mutation causes a male-to-female sex reversal after the sex determination by Dmy in the medaka. *Chromosome Res.* 20, 163–176. <https://doi.org/10.1007/s10577-011-9264-x>
- Matson, C.K., Murphy, M.W., Sarver, A.L., Griswold, M.D., Bardwell, V.J., Zarkower, D., 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature* 476, 101–104. <https://doi.org/10.1038/nature10239>
- Matsuda, M., Nagahama, Y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., Morrey, C.E., Shibata, N., Asakawa, S., Shimizu, N., Hori, H., Hamaguchi, S., Sakaizumi, M.,

2002. DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* 417, 559–563. <https://doi.org/10.1038/nature751>
- Matsuzawa-Watanabe, Y., Inoue, J.-I., Semba, K., 2003. Transcriptional activity of testis-determining factor SRY is modulated by the Wilms' tumor 1 gene product, WT1. *Oncogene* 22, 7900–7904. <https://doi.org/10.1038/sj.onc.1206717>
- McCauley, M.J., Zimmerman, J., Maher, L.J., Williams, M.C., 2007. HMGB Binding to DNA: Single and Double Box Motifs. *J. Mol. Biol.* 374, 993–1004. <https://doi.org/10.1016/j.jmb.2007.09.073>
- McElreavey, K.D., Vilain, E., Boucekkine, C., Vidaud, M., Jaubert, F., Richaud, F., Fellous, M., 1992. XY Sex reversal associated with a nonsense mutation in SRY. *Genomics* 13, 838–840. [https://doi.org/10.1016/0888-7543\(92\)90164-N](https://doi.org/10.1016/0888-7543(92)90164-N)
- McKittrick, E., Gafken, P.R., Ahmad, K., Henikoff, S., 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci. U. S. A.* 101, 1525–1530. <https://doi.org/10.1073/pnas.0308092100>
- Mertin, S., McDowall, S.G., Harley, V.R., 1999. The DNA-binding specificity of SOX9 and other SOX proteins. *Nucleic Acids Res.* 27, 1359–1364. <https://doi.org/10.1093/nar/27.5.1359>
- Mills, S.C., Koskela, E., Mappes, T., 2012. Intralocus sexual conflict for fitness: sexually antagonistic alleles for testosterone. *Proc. R. Soc. B Biol. Sci.* 279, 1889–1895. <https://doi.org/10.1098/rspb.2011.2340>
- Montellier, E., Boussouar, F., Rousseaux, S., Zhang, K., Buchou, T., Fenaille, F., Shiota, H., Debernardi, A., Héry, P., Curtet, S., Jamshidikia, M., Barral, S., Holota, H., Bergon, A., Lopez, F., Guardiola, P., Pernet, K., Imbert, J., Petosa, C., Tan, M., Zhao, Y., Gérard, M., Khochbin, S., 2013. Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev.* 27, 1680–1692. <https://doi.org/10.1101/gad.220095.113>
- Morran, L.T., Schmidt, O.G., Gelarden, I.A., Parrish, R.C., Lively, C.M., 2011. Running with the Red Queen: host-parasite coevolution selects for biparental sex. *Science* 333, 216–218. <https://doi.org/10.1126/science.1206360>
- Myosho, T., Otake, H., Masuyama, H., Matsuda, M., Kuroki, Y., Fujiyama, A., Naruse, K., Hamaguchi, S., Sakaizumi, M., 2012. Tracing the Emergence of a Novel Sex-Determining Gene in Medaka, *Oryzias luzonensis*. *Genetics* 191, 163–170. <https://doi.org/10.1534/genetics.111.137497>
- Nanda, I., Kondo, M., Hornung, U., Asakawa, S., Winkler, C., Shimizu, A., Shan, Z., Haaf, T., Shimizu, N., Shima, A., Schmid, M., Schartl, M., 2002. A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11778–11783. <https://doi.org/10.1073/pnas.182314699>
- Nedelcu, A.M., 2005. Sex as a response to oxidative stress: stress genes co-opted for sex. *Proc. R. Soc. B Biol. Sci.* 272, 1935–1940. <https://doi.org/10.1098/rspb.2005.3151>
- Nedelcu, A.M., Michod, R.E., 2003. Sex as a response to oxidative stress: the effect of antioxidants on sexual induction in a facultatively sexual lineage. *Proc. Biol. Sci.* 270 Suppl 2, S136-139. <https://doi.org/10.1098/rsbl.2003.0062>
- Nef, S., Vassalli, J.-D., 2009. Complementary pathways in mammalian female sex determination. *J. Biol.* 8, 74. <https://doi.org/10.1186/jbiol173>

- Ng, S.S., Yue, W.W., Oppermann, U., Klose, R.J., 2009. Dynamic protein methylation in chromatin biology. *Cell. Mol. Life Sci. CMLS* 66, 407–422. <https://doi.org/10.1007/s00018-008-8303-z>
- Oh, H.J., Lau, Y.-F.C., 2006. KRAB: A partner for SRY action on chromatin. *Mol. Cell. Endocrinol.* 247, 47–52. <https://doi.org/10.1016/j.mce.2005.12.011>
- Oh, H.J., Li, Y., Lau, Y.-F.C., 2005. Sry Associates with the Heterochromatin Protein 1 Complex by Interacting with a KRAB Domain Protein. *Biol. Reprod.* 72, 407–415. <https://doi.org/10.1095/biolreprod.104.034447>
- Oliveira, D.C.S.G., Werren, J.H., Verhulst, E.C., Giebel, J.D., Kamping, A., Beukeboom, L.W., Zande, L.V.D., 2009. Identification and characterization of the doublesex gene of *Nasonia*. *Insect Mol. Biol.* 18, 315–324. <https://doi.org/10.1111/j.1365-2583.2009.00874.x>
- O'Malley, R.C., Huang, S.-S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., Ecker, J.R., 2016. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* 165, 1280–1292. <https://doi.org/10.1016/j.cell.2016.04.038>
- Ospina-Álvarez, N., Piferrer, F., 2008. Temperature-Dependent Sex Determination in Fish Revisited: Prevalence, a Single Sex Ratio Response Pattern, and Possible Effects of Climate Change. *PLOS ONE* 3, e2837. <https://doi.org/10.1371/journal.pone.0002837>
- Otto, S.P., Lenormand, T., 2002. Resolving the paradox of sex and recombination. *Nat. Rev. Genet.* 3, 252–261. <https://doi.org/10.1038/nrg761>
- Otto, S.P., Pannell, J.R., Peichel, C.L., Ashman, T.-L., Charlesworth, D., Chippindale, A.K., Delph, L.F., Guerrero, R.F., Scarpino, S.V., McAllister, B.F., 2011. About PAR: The distinct evolutionary dynamics of the pseudoautosomal region. *Trends Genet.* 27, 358–367. <https://doi.org/10.1016/j.tig.2011.05.001>
- Ottolenghi, C., Pelosi, E., Tran, J., Colombino, M., Douglass, E., Nedorezov, T., Cao, A., Forabosco, A., Schlessinger, D., 2007. Loss of *Wnt4* and *Foxl2* leads to female-to-male sex reversal extending to germ cells. *Hum. Mol. Genet.* 16, 2795–2804. <https://doi.org/10.1093/hmg/ddm235>
- Pannell, J.R., 2017. Plant Sex Determination. *Curr. Biol.* 27, R191–R197. <https://doi.org/10.1016/j.cub.2017.01.052>
- Parfrey, L.W., Lahr, D.J.G., Knoll, A.H., Katz, L.A., 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl. Acad. Sci. U. S. A.* 108, 13624–13629. <https://doi.org/10.1073/pnas.1110633108>
- Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J., Eastman, S., Oliver, B., 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* 299, 697–700. <https://doi.org/10.1126/science.1079190>
- Parr, B.A., McMahon, A.P., 1998. Sexually dimorphic development of the mammalian reproductive tract requires *Wnt-7a*. *Nature* 395, 707–710. <https://doi.org/10.1038/27221>
- Peng, H., Ivanov, A.V., Oh, H.J., Lau, Y.-F.C., Rauscher, F.J., 2009. Epigenetic gene silencing by the SRY protein is mediated by a KRAB-O protein that recruits the KAP1 co-repressor machinery. *J. Biol. Chem.* 284, 35670–35680. <https://doi.org/10.1074/jbc.M109.032086>

- Perry, J., Palmer, S., Gabriel, A., Ashworth, A., 2001. A Short Pseudoautosomal Region in Laboratory Mice. *Genome Res.* 11, 1826–1832. <https://doi.org/10.1101/gr.203001>
- Perry, J.C., Harrison, P.W., Mank, J.E., 2014. The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. *Mol. Biol. Evol.* 31, 1206–1219. <https://doi.org/10.1093/molbev/msu072>
- Peters, A.F., Scornet, D., Ratin, M., Charrier, B., Monnier, A., Merrien, Y., Corre, E., Coelho, S.M., Cock, J.M., 2008. Life-cycle-generation-specific developmental processes are modified in the immediate upright mutant of the brown alga *Ectocarpus siliculosus*. *Dev. Camb. Engl.* 135, 1503–1512. <https://doi.org/10.1242/dev.016303>
- Picard, M.A.-L., Cosseau, C., Mouahid, G., Duval, D., Grunau, C., Toulza, È., Allienne, J.-F., Boissier, J., 2015. The roles of Dmrt (Double sex/Male-abnormal-3 Related Transcription factor) genes in sex determination and differentiation mechanisms: Ubiquity and diversity across the animal kingdom. *C. R. Biol.* 338, 451–462. <https://doi.org/10.1016/j.crv.2015.04.010>
- Pil, P.M., Lippard, S.J., 1992. Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science* 256, 234–237. <https://doi.org/10.1126/science.1566071>
- Pointer, M.A., Harrison, P.W., Wright, A.E., Mank, J.E., 2013. Masculinization of Gene Expression Is Associated with Exaggeration of Male Sexual Dimorphism. *PLOS Genet.* 9, e1003697. <https://doi.org/10.1371/journal.pgen.1003697>
- Polanco, J.C., Wilhelm, D., Mizusaki, H., Jackson, A., Browne, C., Davidson, T., Harley, V., Sinclair, A., Koopman, P., 2009. Functional analysis of the SRY—KRAB interaction in mouse sex determination. *Biol. Cell* 101, 55–67. <https://doi.org/10.1042/BC20080061>
- Portela, A., Esteller, M., 2010. Epigenetic modifications and human disease. *Nat. Biotechnol.* 28, 1057–1068. <https://doi.org/10.1038/nbt.1685>
- Postnikov, Y.V., Bustin, M., 2016. Functional interplay between histone H1 and HMG proteins in chromatin. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* 1859, 462–467. <https://doi.org/10.1016/j.bbagr.2015.10.006>
- Rajeev, L., Luning, E.G., Mukhopadhyay, A., 2014. DNA-affinity-purified Chip (DAP-chip) Method to Determine Gene Targets for Bacterial Two component Regulatory Systems. *J. Vis. Exp. JoVE.* <https://doi.org/10.3791/51715>
- Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., Manke, T., 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165. <https://doi.org/10.1093/nar/gkw257>
- Raymond, C.S., Shamu, C.E., Shen, M.M., Seifert, K.J., Hirsch, B., Hodgkin, J., Zarkower, D., 1998. Evidence for evolutionary conservation of sex-determining genes. *Nature* 391, 691–695. <https://doi.org/10.1038/35618>
- Rice, W.R., 1996. Evolution of the Y Sex Chromosome in Animals Y chromosomes evolve through the degeneration of autosomes. *BioScience* 46, 331–343. <https://doi.org/10.2307/1312947>
- Rice, W.R., 1984. SEX CHROMOSOMES AND THE EVOLUTION OF SEXUAL DIMORPHISM. *Evol. Int. J. Org. Evol.* 38, 735–742. <https://doi.org/10.1111/j.1558-5646.1984.tb00346.x>

- Richard, G., Legeai, F., Prunier-Leterme, N., Bretaudeau, A., Tagu, D., Jaquiéry, J., Le Trionnaire, G., 2017. Dosage compensation and sex-specific epigenetic landscape of the X chromosome in the pea aphid. *Epigenetics Chromatin* 10, 30. <https://doi.org/10.1186/s13072-017-0137-1>
- Riggs, A.D., Porter, T.N., 1996. Overview of Epigenetic Mechanisms. *Cold Spring Harb. Monogr. Arch.* 32, 29–45. <https://doi.org/10.1101/0.29-45>
- Rizzo, J.M., Sinha, S., 2014. Analyzing the Global Chromatin Structure of Keratinocytes by MNase-Seq, in: Turksen, K. (Ed.), *Epidermal Cells: Methods and Protocols, Methods in Molecular Biology*. Springer, New York, NY, pp. 49–59. [https://doi.org/10.1007/7651\\_2014\\_77](https://doi.org/10.1007/7651_2014_77)
- Ross, J.A., Peichel, C.L., 2008. Molecular cytogenetic evidence of rearrangements on the Y chromosome of the threespine stickleback fish. *Genetics* 179, 2173–2182. <https://doi.org/10.1534/genetics.108.088559>
- Rouyer, F., Simmler, M.C., Johnsson, C., Vergnaud, G., Cooke, H.J., Weissenbach, J., 1986. A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes. *Nature* 319, 291–295. <https://doi.org/10.1038/319291a0>
- Salvemini, M., Mauro, U., Lombardo, F., Milano, A., Zazzaro, V., Arcà, B., Polito, L.C., Saccone, G., 2011. Genomic organization and splicing evolution of the doublesex gene, a *Drosophila* regulator of sexual differentiation, in the dengue and yellow fever mosquito *Aedes aegypti*. *BMC Evol. Biol.* 11, 41. <https://doi.org/10.1186/1471-2148-11-41>
- Sánchez, A., Marchal, J.A., Romero-Fernández, I., Pinna-Senn, E., Ortiz, M.I., Bella, J.L., Lisanti, J.A., 2010. No differences in the Sry gene between males and XY females in Akodon (Rodentia, Cricetidae). *Sex. Dev. Genet. Mol. Biol. Evol. Endocrinol. Embryol. Pathol. Sex Determ. Differ.* 4, 155–161. <https://doi.org/10.1159/000309780>
- Santenard, A., Torres-Padilla, M.-E., 2009. Epigenetic reprogramming in mammalian reproduction: Contribution from histone variants. *Epigenetics* 4, 80–84. <https://doi.org/10.4161/epi.4.2.7838>
- Sato, T., Endo, T., Yamahira, K., Hamaguchi, S., Sakaizumi, M., 2005. Induction of Female-to-Male Sex Reversal by High Temperature Treatment in Medaka, *Oryzias latipes*. *Zool. Sci.* 22, 985–988. <https://doi.org/10.2108/zsj.22.985>
- Schepers, G.E., Teasdale, R.D., Koopman, P., n.d. Sox genes. We hope that this will provide a useful framework for comparative and functional studies in a range of developmental model systems. *Dev. Cell* 4.
- Sekido, R., Lovell-Badge, R., 2009. Sex determination and SRY: down to a wink and a nudge? *Trends Genet.* 25, 19–29. <https://doi.org/10.1016/j.tig.2008.10.008>
- Sekido, R., Lovell-Badge, R., 2008. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* 453, 930–934. <https://doi.org/10.1038/nature06944>
- Sessa, L., Bianchi, M.E., 2007. The evolution of High Mobility Group Box (HMGB) chromatin proteins in multicellular animals. *Gene* 387, 133–140. <https://doi.org/10.1016/j.gene.2006.08.034>
- Sharma, S., Kelly, T.K., Jones, P.A., 2010. Epigenetics in cancer. *Carcinogenesis* 31, 27–36. <https://doi.org/10.1093/carcin/bgp220>



- She, Z.-Y., Yang, W.-X., 2015. SOX family transcription factors involved in diverse cellular events during development. *Eur. J. Cell Biol.* 94, 547–563. <https://doi.org/10.1016/j.ejcb.2015.08.002>
- Shen, M.M., Hodgkin, J., 1988. *mab-3*, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. *Cell* 54, 1019–1031. [https://doi.org/10.1016/0092-8674\(88\)90117-1](https://doi.org/10.1016/0092-8674(88)90117-1)
- Shi, Yujiang, Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., Shi, Yang, 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119, 941–953. <https://doi.org/10.1016/j.cell.2004.12.012>
- Simpson, A.G.B., Inagaki, Y., Roger, A.J., 2006. Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of “primitive” eukaryotes. *Mol. Biol. Evol.* 23, 615–625. <https://doi.org/10.1093/molbev/msj068>
- Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.-M., Lovell-Badge, R., Goodfellow, P.N., 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346, 240–244. <https://doi.org/10.1038/346240a0>
- Singh, N.P., Madabhushi, S.R., Srivastava, S., Senthilkumar, R., Neeraja, C., Khosla, S., Mishra, R.K., 2011. Epigenetic profile of the euchromatic region of human Y chromosome. *Nucleic Acids Res.* 39, 3594–3606. <https://doi.org/10.1093/nar/gkq1342>
- Slattery, M., Zhou, T., Yang, L., Dantas Machado, A.C., Gordân, R., Rohs, R., 2014. Absence of a simple code: how transcription factors read the genome. *Trends Biochem. Sci.* 39, 381–399. <https://doi.org/10.1016/j.tibs.2014.07.002>
- Slupphaug, G., Kavli, B., Krokan, H.E., 2003. The interacting pathways for prevention and repair of oxidative DNA damage. *Mutat. Res.* 531, 231–251. <https://doi.org/10.1016/j.mrfmmm.2003.06.002>
- Smith, C.A., Roeszler, K.N., Ohnesorg, T., Cummins, D.M., Farlie, P.G., Doran, T.J., Sinclair, A.H., 2009. The avian Z-linked gene *DMRT1* is required for male sex determination in the chicken. *Nature* 461, 267–271. <https://doi.org/10.1038/nature08298>
- Smith, C.A., Shoemaker, C.M., Roeszler, K.N., Queen, J., Crews, D., Sinclair, A.H., 2008. Cloning and expression of *R-Spondin1* in different vertebrates suggests a conserved role in ovarian development. *BMC Dev. Biol.* 8, 72. <https://doi.org/10.1186/1471-213X-8-72>
- Smith, J.M., Dowson, C.G., Spratt, B.G., 1991. Localized sex in bacteria. *Nature* 349, 29–31. <https://doi.org/10.1038/349029a0>
- Song, L., Crawford, G.E., 2010. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb. Protoc.* 2010, pdb.prot5384. <https://doi.org/10.1101/pdb.prot5384>
- Speijer, D., 2016. What can we infer about the origin of sex in early eukaryotes? *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150530. <https://doi.org/10.1098/rstb.2015.0530>
- Speijer, D., Lukeš, J., Eliáš, M., 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc. Natl. Acad. Sci.* 112, 8827–8834. <https://doi.org/10.1073/pnas.1501725112>
- Spitz, F., Furlong, E.E.M., 2012. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 13, 613–626. <https://doi.org/10.1038/nrg3207>

- Steinmann-Zwicky, M., 1994. Sex determination of the *Drosophila* germ line: tra and dsx control somatic inductive signals. *Development* 120, 707–716.
- Stergachis, A.B., Haugen, E., Shafer, A., Fu, W., Vernot, B., Reynolds, A., Raubitschek, A., Ziegler, S., LeProust, E.M., Akey, J.M., Stamatoyannopoulos, J.A., 2013. Exonic transcription factor binding directs codon choice and affects protein evolution. *Science* 342, 1367–1372. <https://doi.org/10.1126/science.1243490>
- Strahl, B.D., Allis, C.D., 2000. The language of covalent histone modifications. *Nature* 403, 41–45. <https://doi.org/10.1038/47412>
- Štros, M., 2010. HMGB proteins: Interactions with DNA and chromatin. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* 1799, 101–113. <https://doi.org/10.1016/j.bbagr.2009.09.008>
- Štros, M., Launholt, D., Grasser, K.D., 2007. The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cell. Mol. Life Sci.* 64, 2590–2606. <https://doi.org/10.1007/s00018-007-7162-3>
- Štros, M., Polanská, E., Štruncová, S., Pospíšilová, Š., 2009. HMGB1 and HMGB2 proteins up-regulate cellular expression of human topoisomerase II $\alpha$ . *Nucleic Acids Res.* 37, 2070–2086. <https://doi.org/10.1093/nar/gkp067>
- Suzuki, M.G., Funaguma, S., Kanda, T., Tamura, T., Shimada, T., 2005. Role of the male BmDSX protein in the sexual differentiation of *Bombyx mori*. *Evol. Dev.* 7, 58–68. <https://doi.org/10.1111/j.1525-142X.2005.05007.x>
- Talbert, P.B., Henikoff, S., 2010. Histone variants--ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.* 11, 264–275. <https://doi.org/10.1038/nrm2861>
- Tanurdzic, M., Banks, J.A., 2004. Sex-determining mechanisms in land plants. *Plant Cell* 16 Suppl, S61-71. <https://doi.org/10.1105/tpc.016667>
- Teranishi, M., Shimada, Y., Hori, T., Nakabayashi, O., Kikuchi, T., Macleod, T., Pym, R., Sheldon, B., Solovei, I., Macgregor, H., Mizuno, S., 2001. Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 9, 147–165. <https://doi.org/10.1023/a:1009235120741>
- Thomas, J.O., Travers, A.A., 2001. HMG1 and 2, and related ‘architectural’ DNA-binding proteins. *Trends Biochem. Sci.* 26, 167–174. [https://doi.org/10.1016/S0968-0004\(01\)01801-1](https://doi.org/10.1016/S0968-0004(01)01801-1)
- Tomizuka, K., Horikoshi, K., Kitada, R., Sugawara, Y., Iba, Y., Kojima, A., Yoshitome, A., Yamawaki, K., Amagai, M., Inoue, A., Oshima, T., Kakitani, M., 2008. R-spondin1 plays an essential role in ovarian development through positively regulating Wnt-4 signaling. *Hum. Mol. Genet.* 17, 1278–1291. <https://doi.org/10.1093/hmg/ddn036>
- Travers, A., 2000. Recognition of distorted DNA structures by HMG domains. *Curr. Opin. Struct. Biol.* 10, 102–109. [https://doi.org/10.1016/S0959-440X\(99\)00056-1](https://doi.org/10.1016/S0959-440X(99)00056-1)
- Tripathi, V., Raman, R., 2010. Identification of Wnt4 as the ovary pathway gene and temporal disparity of its expression vis-a-vis testis genes in the garden lizard, *Calotes versicolor*. *Gene* 449, 77–84. <https://doi.org/10.1016/j.gene.2009.09.001>
- Trojer, P., Reinberg, D., 2007. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Mol. Cell* 28, 1–13. <https://doi.org/10.1016/j.molcel.2007.09.011>

- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., Zhang, Y., 2006. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811–816. <https://doi.org/10.1038/nature04433>
- Uhlenhaut, N.H., Jakob, S., Anlag, K., Eisenberger, T., Sekido, R., Kress, J., Treier, A.-C., Klugmann, C., Klasen, C., Holter, N.I., Riethmacher, D., Schütz, G., Cooney, A.J., Lovell-Badge, R., Treier, M., 2009. Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell* 139, 1130–1142. <https://doi.org/10.1016/j.cell.2009.11.021>
- Uller, T., Helanterä, H., 2011. From the Origin of Sex-Determining Factors to the Evolution of Sex-Determining Systems. *Q. Rev. Biol.* 86, 163–180. <https://doi.org/10.1086/661118>
- Umen, J.G., 2014. Green algae and the origins of multicellularity in the plant kingdom. *Cold Spring Harb. Perspect. Biol.* 6, a016170. <https://doi.org/10.1101/cshperspect.a016170>
- Umen, J.G., Goodenough, U.W., 2001. Control of cell division by a retinoblastoma protein homolog in *Chlamydomonas*. *Genes Dev.* 15, 1652–1661. <https://doi.org/10.1101/gad.892101>
- Veluchamy, A., Rastogi, A., Lin, X., Lombard, B., Murik, O., Thomas, Y., Dingli, F., Rivarola, M., Ott, S., Liu, X., Sun, Y., Rabinowicz, P.D., McCarthy, J., Allen, A.E., Loew, D., Bowler, C., Tirichine, L., 2015. An integrative analysis of post-translational histone modifications in the marine diatom *Phaeodactylum tricornutum*. *Genome Biol.* 16, 102. <https://doi.org/10.1186/s13059-015-0671-8>
- Verhulst, E.C., van de Zande, L., Beukeboom, L.W., 2010. Insect sex determination: it all evolves around transformer. *Curr. Opin. Genet. Dev., Developmental mechanisms, patterning and evolution* 20, 376–383. <https://doi.org/10.1016/j.gde.2010.05.001>
- Veyrunes, F., Chevret, P., Catalan, J., Castiglia, R., Watson, J., Dobigny, G., Robinson, T.J., Britton-Davidian, J., 2010. A novel sex determination system in a close relative of the house mouse. *Proc. R. Soc. B Biol. Sci.* 277, 1049–1056. <https://doi.org/10.1098/rspb.2009.1925>
- Voss, T.C., Hager, G.L., 2014. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat. Rev. Genet.* 15, 69–81. <https://doi.org/10.1038/nrg3623>
- Weir, H.M., Kraulis, P.J., Hill, C.S., Raine, A.R., Laue, E.D., Thomas, J.O., 1993. Structure of the HMG box motif in the B-domain of HMG1. *EMBO J.* 12, 1311–1319.
- Whittle, C.A., Votintseva, A., Ridout, K., Filatov, D.A., 2015. Recent and Massive Expansion of the Mating-Type-Specific Region in the Smut Fungus *Microbotryum*. *Genetics* 199, 809–816. <https://doi.org/10.1534/genetics.114.171702>
- Wiese, L., Wiese, W., Edwards, D.A., 1979. Inducible Anisogamy and the Evolution of Oogamy from Isogamy. *Ann. Bot.* 44, 131–139. <https://doi.org/10.1093/oxfordjournals.aob.a085712>
- Wilson, M., Koopman, P., 2002. Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. *Curr. Opin. Genet. Dev.* 12, 441–446. [https://doi.org/10.1016/S0959-437X\(02\)00323-4](https://doi.org/10.1016/S0959-437X(02)00323-4)
- Wong, R.Y., McLeod, M.M., Godwin, J., 2014. Limited sex-biased neural gene expression patterns across strains in Zebrafish (*Danio rerio*). *BMC Genomics* 15, 905. <https://doi.org/10.1186/1471-2164-15-905>

- Woodcock, C.L., Ghosh, R.P., 2010. Chromatin Higher-order Structure and Dynamics. *Cold Spring Harb. Perspect. Biol.* 2, a000596–a000596. <https://doi.org/10.1101/cshperspect.a000596>
- Woodward, D.E., Murray, J.D., 1993. On the Effect of Temperature-Dependent Sex Determination on Sex Ratio and Survivorship in Crocodylians. *Proc. Biol. Sci.* 252, 149–155.
- Wright, A.E., Dean, R., Zimmer, F., Mank, J.E., 2016. How to make a sex chromosome. *Nat. Commun.* 7, 12087. <https://doi.org/10.1038/ncomms12087>
- Xhemalce, B., Dawson, M.A., Bannister, A.J., 2011. Histone Modifications, in: *Reviews in Cell Biology and Molecular Medicine*. American Cancer Society. <https://doi.org/10.1002/3527600906.mcb.201100004>
- Yamazaki, T., Ichihara, K., Suzuki, R., Oshima, K., Miyamura, S., Kuwano, K., Toyoda, A., Suzuki, Y., Sugano, S., Hattori, M., Kawano, S., 2017. Genomic structure and evolution of the mating type locus in the green seaweed *Ulva partita*. *Sci. Rep.* 7, 1–17. <https://doi.org/10.1038/s41598-017-11677-0>
- Yang, P.K., Kuroda, M.I., 2007. Noncoding RNAs and Intranuclear Positioning in Monoallelic Gene Expression. *Cell* 128, 777–786. <https://doi.org/10.1016/j.cell.2007.01.032>
- Yang, X.-J., Seto, E., 2007. HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* 26, 5310–5318. <https://doi.org/10.1038/sj.onc.1210599>
- Yano, A., Guyomard, R., Nicol, B., Jouanno, E., Quillet, E., Klopp, C., Cabau, C., Bouchez, O., Fostier, A., Guiguen, Y., 2012. An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*. *Curr. Biol. CB* 22, 1423–1428. <https://doi.org/10.1016/j.cub.2012.05.045>
- Yasuhara, J.C., Wakimoto, B.T., 2008. Molecular Landscape of Modified Histones in *Drosophila* Heterochromatic Genes and Euchromatin-Heterochromatin Transition Zones. *PLOS Genet.* 4, e16. <https://doi.org/10.1371/journal.pgen.0040016>
- Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F., Nitta, K.R., Taipale, M., Popov, A., Ginno, P.A., Domcke, S., Yan, J., Schübeler, D., Vinson, C., Taipale, J., 2017. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* 356. <https://doi.org/10.1126/science.aaj2239>
- Yu, Q., Tong, E., Skelton, R.L., Bowers, J.E., Jones, M.R., Murray, J.E., Hou, S., Guan, P., Acob, R.A., Luo, M.-C., Moore, P.H., Alam, M., Paterson, A.H., Ming, R., 2009. A physical map of the papaya genome with integrated genetic map and genome sequence. *BMC Genomics* 10, 371. <https://doi.org/10.1186/1471-2164-10-371>
- Yuan, J., Adamski, R., Chen, J., 2010. Focus on histone variant H2AX: to be or not to be. *FEBS Lett.* 584, 3717–3724. <https://doi.org/10.1016/j.febslet.2010.05.021>
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>
- Zhao, F., Franco, H.L., Rodriguez, K.F., Brown, P.R., Tsai, M.-J., Tsai, S.Y., Yao, H.H.-C., 2017. Elimination of the male reproductive tract in the female embryo is promoted by COUP-TFII in mice. *Science* 357, 717–720. <https://doi.org/10.1126/science.aai9136>

- Zheng, B., Chen, X., 2011. Dynamics of histone H3 lysine 27 trimethylation in plant development. *Curr. Opin. Plant Biol.* 14, 123–129. <https://doi.org/10.1016/j.pbi.2011.01.001>
- Zhou, L., Chakraborty, T., Zhou, Q., Mohapatra, S., Nagahama, Y., Zhang, Y., 2016. Rspo1-activated signalling molecules are sufficient to induce ovarian differentiation in XY medaka (*Oryzias latipes*). *Sci. Rep.* 6. <https://doi.org/10.1038/srep19543>
- Zhou, Q., Zhang, J., Bachtrog, D., An, N., Huang, Q., Jarvis, E.D., Gilbert, M.T.P., Zhang, G., 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* 346, 1246338. <https://doi.org/10.1126/science.1246338>
- Zhou, T., Yang, L., Lu, Y., Dror, I., Dantas Machado, A.C., Ghane, T., Di Felice, R., Rohs, R., 2013. DNASHape: a method for the high-throughput prediction of DNA structural features on a genomic scale. *Nucleic Acids Res.* 41, W56-62. <https://doi.org/10.1093/nar/gkt437>



# Appendix UV Chromosomes and Haploid Sexual Systems

## Review

## UV Chromosomes and Haploid Sexual Systems

Susana Margarida Coelho,<sup>1,\*</sup> Josselin Gueno,<sup>1</sup> Agnieszka Paulina Lipinska,<sup>1</sup> Jeremy Mark Cock,<sup>1</sup> and James G. Umen<sup>2,\*</sup>

The evolution of sex determination continues to pose major questions in biology. Sex-determination mechanisms control reproductive cell differentiation and development of sexual characteristics in all organisms, from algae to animals and plants. While the underlying processes defining sex (meiosis and recombination) are conserved, sex-determination mechanisms are highly labile. In particular, a flow of new discoveries has highlighted several fascinating features of the previously understudied haploid UV sex determination and related mating systems found in diverse photosynthetic taxa including green algae, bryophytes, and brown algae. Analyses integrating information from these systems and contrasting them with classical XY and ZW systems are providing exciting insights into both the universality and the diversity of sex-determining chromosomes across eukaryotes.

### Sex Chromosomes and Sex Determination: Variations on a Theme

Meiotic sex and recombination, and the resulting alternation between haploid and diploid life-cycle phases, is an ancestral, highly conserved process [1]. Meiotic recombination creates genetic variation by generating new combinations of gene variants (alleles). Many eukaryotes produce gametes of equal size (isogamy), and this is thought to be the ancestral state [2], while others have evolved to produce differentiated male and female gametes (anisogamy and oogamy). In isogamous species the distinct types of gamete are referred to as mating types, while in anisogamous/oogamous species the gametes are defined as either male (the smaller-sized gamete) or female (the larger-sized gamete). Separate male/female sexes have arisen independently and repeatedly during eukaryotic evolution, and are specified by a bewildering diversity of mechanisms ranging from purely genetic to epigenetic, or some combination of the two [3].

The sex of an individual may be determined during either the haploid or the diploid phase of the life cycle (Box 1). When sex is determined genetically in organisms with haploid-phase sex-determination systems, the chromosomes that contain the **sex-determining region** (SDR; see Glossary) are referred to as **U and V sex chromosomes** [4]. In anisogamous and oogamous organisms, females carry a U chromosome, whereas males carry a V chromosome. UV chromosomes are relatively common among eukaryotes, and arose independently in different eukaryotic groups several times during evolution (Figure 1). However, for many years research has focused exclusively on XY and ZW systems, leaving UV chromosomes and haploid sex determination largely neglected. The UV designation pertains to organisms with male and female sexes and not to those with isogamous mating types, but this distinction is somewhat artificial. For example, in volvocine algae the chromosome containing the *MT*–mating locus in isogamous species corresponds phylogenetically to the V chromosome carried by males in anisogamous/oogamous species, and the chromosome with the *MT*+ locus corresponds to the U chromosome [5]. We retain the conventional usages of UV and

### Highlights

Evolutionary theory predicts some common features shared by diploid XY or ZW chromosomes and UV chromosomes, but also some unique properties of UV systems.

Haploid U and V sex chromosomes are found in diverse eukaryotes and are likely to have originated from mating-type chromosomes.

The origin of sexes from mating types in volvocine algae can be traced between UV and mating-type chromosomes, and through conservation of the sex-determining gene *MID*.

New data from algae and early-diverging plants confirm evolutionary predictions and reveal unexpected properties of UV systems, such as asymmetric chromosome expansions and the accumulation of diploid-phase genes in pseudoautosomal regions.

*MID*-related or other classes of RWP-RK transcription factor genes in green algae and HMG domain-encoding genes in brown algae have emerged as candidate sex-determining genes.

<sup>1</sup>Sorbonne Université, Centre National de la Recherche Scientifique (CNRS), Integrative Biology of Marine Models (LBI2M), Station Biologique de Roscoff (SBR), 29680 Roscoff, France  
<sup>2</sup>Donald Danforth Plant Science Center, St. Louis, MO 63132, USA

\*Correspondence: [coelho@sb-roscoff.fr](mailto:coelho@sb-roscoff.fr) (S.M. Coelho) and [jumen@danforthcenter.org](mailto:jumen@danforthcenter.org) (J.G. Umen).



mating-type chromosomes here, but note that theoretical predictions and empirical data support similar evolutionary processes operating in both. This review focuses on recent advances in the characterization of haploid sex-determination systems in photosynthetic eukaryotes. It is important to note, however, that many fungi have haploid sexual systems with properties similar to those of the systems described here (e.g., [6]). Several recent papers have focused on haploid mating types and sex determination in fungi (e.g., [7–9]).

### Modeling the Evolution of UV Chromosomes

The repeated independent evolution of sex chromosomes across the eukaryotes represents a remarkable example of genomic convergence because these chromosomes share many properties. The evolution of dimorphic sex chromosomes is assumed to be driven largely by reduced recombination that arises to maintain tight linkage between sex-determining and/or sex-related genes (Box 2; reviewed in [4,10]). Compared to autosomes, sex chromosomes experience different mutation and recombination rates, effective population sizes, and levels of sexual selection, and these differences may profoundly affect their evolution. In diploid systems, the X or the Z chromosomes experience two different environments depending on whether they are in the homomorphic state and recombination can occur (XX and ZZ), or in the heteromorphic state (XY or ZW) where recombination is blocked, thus making the evolution of the X versus Y or Z versus W chromosomes inherently asymmetric. By contrast, in haploid systems the female U and the male V experience largely similar and symmetrical recombination environments because they are either unpaired in the haploid gametophyte or paired heteromorphically in UV sporophyte diploids [11]. As a result, both the U and the V SDRs are expected to exhibit the degenerative effects of arrested recombination and reduced effective population size to a similar extent. Moreover, because U and V chromosomes function in the context of an extended haploid life-cycle phase when only one sex chromosome is present in each cell, deleterious mutations are expected to be more efficiently purged than from diploid-phase sexual systems where there is greater opportunity for sheltering of deleterious alleles [11,12]. Consequently, degeneration of UV chromosomes is expected to occur more slowly than for diploid-phase sex chromosomes (XY or ZW). Note, however, that deleterious mutations in SDR genes can be masked if these genes function during the diploid sporophyte phase, and this may allow both the U and V chromosomes to degenerate to some extent [13,14]. Unlike XY or ZW chromosomes, where gene loss on the non-recombining portion of the Y or W chromosome is a prevalent long-term outcome, the fate of non-recombining genes in UV systems is more likely to be allelic differentiation between the U and V copies where neither copy can be lost, but over time, polymorphisms in either member of the **gametolog** pair will become fixed through a combination of drift, positive selection, and/or hitchhiking. As a result, UV systems are predicted to exhibit gametolog differentiation and not gene loss as a long-term outcome of suppressed recombination [13]. It has also been suggested that changes in the size of the U or V should involve predominantly sequence gains, including additions of beneficial (but not essential) genes and/or relatively neutral sequences such as repeats and transposons, rather than gene loss [11]. Although U and V chromosomes are expected to evolve similarly, verbal models predict they may exhibit some asymmetry if sexual selection is stronger in one of the sexes [11]. Recent mathematical modeling predicted a gradual decrease in the amount of recombination between U and V chromosomes and addition of **strata** via successive inversions or rearrangements in flanking sequences, exactly as in diploid sex chromosome systems (Box 2). Note that the theoretical predictions described above are valid both for UV chromosomes and chromosomes carrying **mating-type loci** (MTLs) provided there is even minor differential selection between the two mating types that favors decreased recombination [13].

### Glossary

- Dioecious:** diploid-phase sex determination with genetically distinct sporophytes corresponding to each sex or mating type.
- Dioicous:** haploid-phase sex or mating-type determination with genetically distinct gametophytes corresponding to each sex or mating type.
- Gametologs:** pairs of orthologous MTL/SDR genes that are derived from a single ancestral gene located within the non-recombining region of opposite sex chromosomes (X and Y, Z and W, or U and V) or in the *MT*<sup>−</sup> and *MT*<sup>+</sup> MTLs.
- Gonochoric:** used in metazoans with same meaning as dioecious (genetically determined sexes).
- Heterothallism:** mating incompatibility between genetically identical individuals, that is, species with genetically determined sexes or mating types. Usually used with microbial eukaryotes.
- Homothallism:** mating compatibility between genetically identical individuals, that is, epigenetic sex determination. Usually used with microbial eukaryotes.
- Mating-type locus (MTL):** locus that determines mating type in isogamous heterothallic species.
- Monoicous:** haploid-phase sex determination or mating-type determination where both gamete types are produced by the same gametophyte.
- Monoecious:** diploid-phase sex determination or mating-type determination where both male and female sporophytes are produced by the same individual, that is, epigenetic sex determination. In angiosperms, the term 'hermaphrodite' is used specifically to denote the very common case where separate male and female organs are present in the same flower ('perfect flowers'), while monoecious refers to plants where the same individual produces distinct male and female flowers.
- Pseudoautosomal regions (PARs):** recombining regions flanking the MTL/SDR on the sex or mating-type chromosomes.
- Sex chromosome:** the chromosome in an organism with male and female gametes

## The Empirical Era – A Range of Haploid Sexual Systems Revealed by Next-Generation Sequencing

Over the past decade a growing amount of information has become available about the structure (Table 1) and evolution of UV chromosomes and MTLs, and the taxonomic breadth of haploid sexual systems under study has increased considerably. The sections below highlight some recent advances in representative UV systems from the chlorophyte, bryophyte, and brown algal lineages.

### Volvocine Algae

#### Overview

Volvocine algae are a related group of chlorophytes (green algae) that collectively form a fascinating study set for the evolution of sex and sex chromosomes. Although volvocine algae are not a formal taxonomic grouping, the multicellular members form a monophyletic clade including genera that exhibit different degrees of sexual dimorphism ranging from isogamy to anisogamy and oogamy [15]. The isogamous unicellular species *Chlamydomonas reinhardtii* is an outgroup that is included in the volvocines owing to its high degree of relatedness to the multicellular members [16]. The volvocines and most other green algae have a haplontic life cycle where vegetative haploid cells or individuals can reproduce mitotically, but under appropriate conditions transition to a sexual phase that involves gametic differentiation and mating to form a diploid zygotic resting spore. Under favorable conditions the environmentally resistant spores reawaken and undergo meiosis to produce new haploid vegetative progeny. The majority of volvocine algae are **heterothallic (dioicous)**, and the lineage as a whole appears to be ancestrally heterothallic, but **homothallic (monoicy)** has arisen independently within the volvocines multiple times [17]. Transitions between monoicy and dioicy in volvocine algae are summarized in Box 3. The MTLs/SDRs of volvocine algae control not only mating type/sexual differentiation but also govern uniparental organelle inheritance [18,19]. Complete MTL/SDR sequences are currently available for five volvocine species including isogamous (*Chlamydomonas reinhardtii*, *Gonium pectorale*, *Yamagishiella unicocca*), anisogamous (*Eudorina* sp.), and oogamous (*Volvox carteri*) representatives [20–22]. A series of studies on volvocine MTLs/SDRs, starting with the well-established model *C. reinhardtii* [23], have documented results that are in agreement with predictions for UV chromosome or MTL evolution, as well as unexpected findings that were not anticipated by theoretical models. The molecular-genetic basis of mating-type determination and sex determination in volvocine algae is continuous throughout the lineage where a conserved transcription factor gene, *MID*, is found in either the *minus* mating type of the MTL or in the V (male) chromosome SDR in all dioicous species [20,22,24–26] (Box 4). A few additional mating type-linked or sex-linked genes with functions in the sexual cycle have also been described in volvocine algae including the *MT+*/female gamete-fusion protein coding gene *FUS1* and the *MT-*/male gene *MTD1*, although none are as universally conserved as *MID* [18,22].

#### Structure and Molecular Evolution of Volvocine MTLs/SDRs

In models of UV chromosome evolution, the non-recombining MTL/SDR can expand through additional rearrangements or insertions of sequences that were formerly autosomal, leading to formation of distinct 'strata' whose residence times in the non-recombining region can be estimated based on divergence between gametologs (in the case of rearrangements) or between the MTL/SDR-linked and autosomal copies of sequences (in the case of insertions) (see Figure I in Box 2) [27,28]. It might be expected that the history of volvocine MTLs/SDRs could be elucidated based on structural and molecular comparisons of shared regions and/or gene content. On the contrary, the five known MTLs/SDRs

(anisogamy or oogamy) that carries the SDR.

**Sex-determining region (SDR):** a region of a sex chromosome that contains the locus that determines sex, often an extensive, non-recombining region spanning many kilobases.

**Sexual antagonism:** pertains to genes or alleles that increase reproductive fitness when expressed in one sex or mating type but decrease reproductive fitness when expressed in the opposite sex or mating type.

**Strata:** chromosomal regions that have become part of the non-recombining MTL/SDR at different evolutionary times.

**U and V chromosomes:** female (U) and male (V) chromosomes that determine sex during the haploid phase of the life cycle.

## Box 1. Life Cycles and Sex Determination

A typical eukaryotic sexual life cycle involves alternating phases with diploid-to-haploid transitions occurring through meiosis, and haploid-to-diploid transitions through syngamy (i.e., fusion of two haploid gametes) (Figure 1). Life cycles can be defined as diplontic, haplontic, or haplodiplontic depending on whether mitotic divisions (cell proliferation or multicellular growth) occur during the diploid phase, during the haploid phase, or during both phases, respectively. For diplontic organisms (e.g., humans), sex determination occurs during the diploid phase, while in haplontic organisms (e.g., the green alga *Volvox*) sex determination is in the haploid phase. For an organism with a haplodiplontic life cycle, however, sex can be determined during either the diploid (e.g., *Silene latifolia*) or the haploid (e.g., *Marchantia polymorpha*) phase of the life cycle. Different terminologies are used to clearly distinguish between diploid- and haploid-phase sex-determination systems. For example, for diploid-phase systems, organisms are **monoecious** if the same individual produces gametes of both sexes (e.g., *Zea mays*) and **dioecious** (or **gonochoric**) if individuals produce either male or female gametes, but not both (e.g., *Silene latifolia*). For haploid-phase sexual systems, however, the equivalent terms are **monoicous** (i.e., individuals produce both gamete types, e.g., mosses) and **dioicous** (i.e., sperm and eggs are produced by genetically distinct male or female individuals, e.g., *Marchantia polymorpha*).

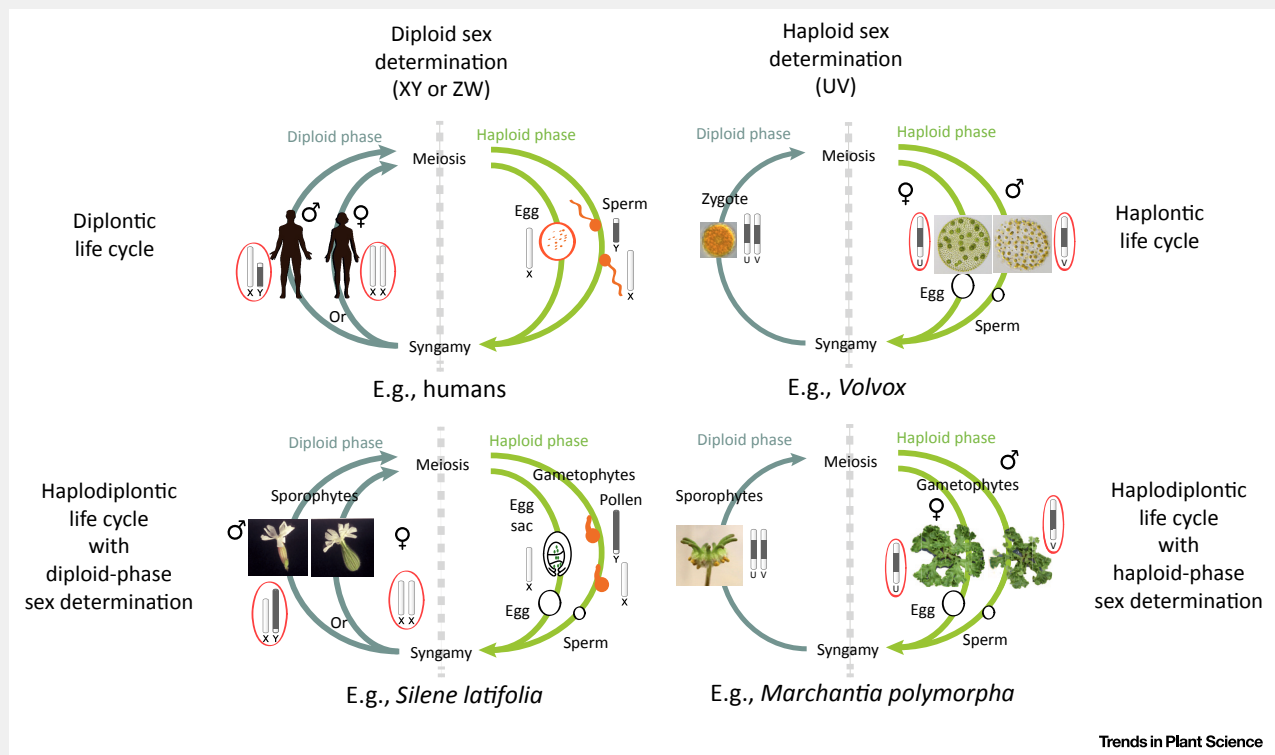
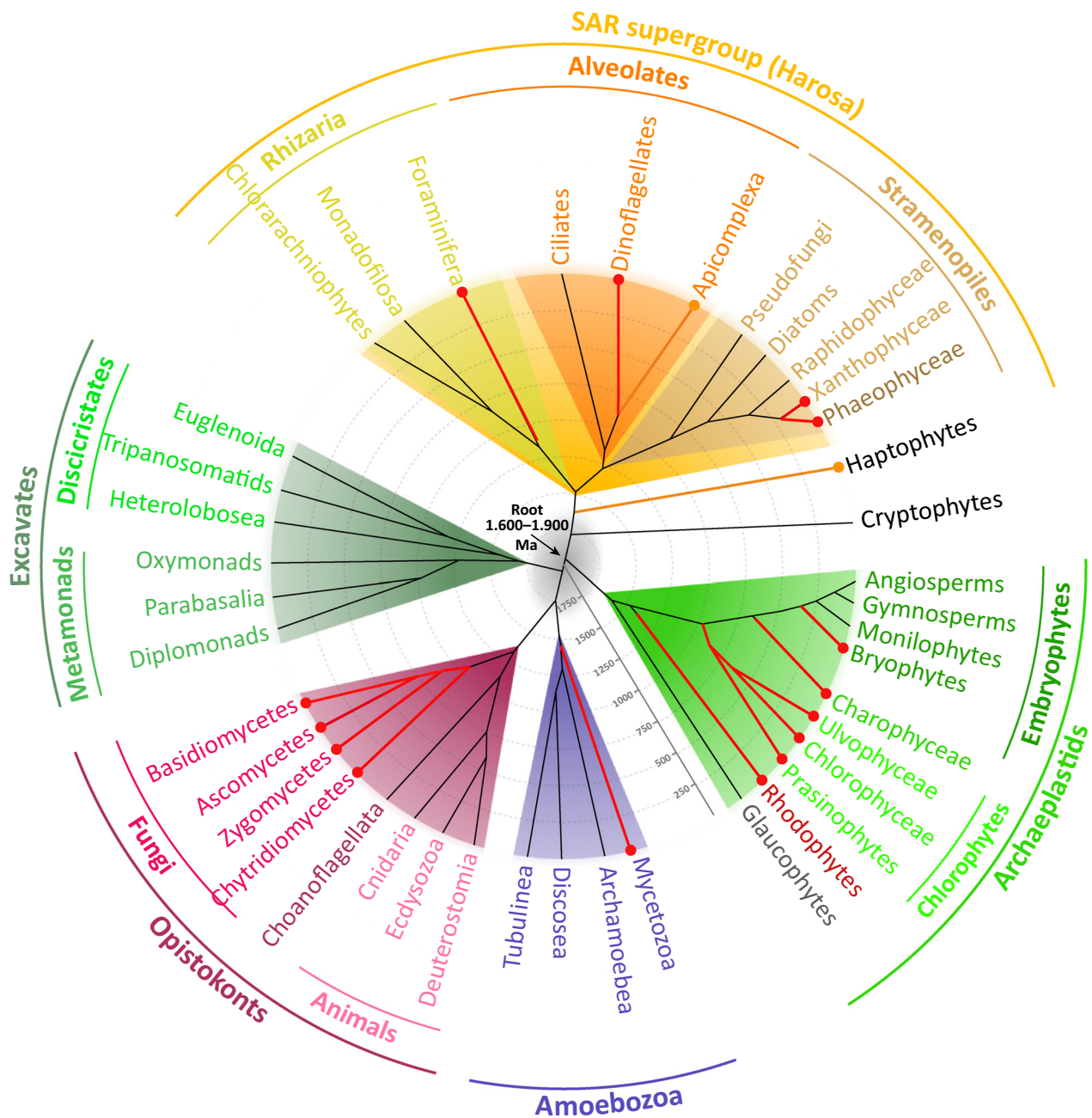


Figure 1. Relationship between Life Cycle and the Sex-Determination System in Diverse Eukaryotes.

of volvocine algae, which vary greatly in size (from 7 kb to >1 Mb), do not share strata or structural features; and few of their genes appear to be long-term permanent residents [22]. This lack of structural continuity suggests relatively frequent turnover of volvocine MTLs/SDRs. Interestingly, however, the U/V or mating-type chromosome in which the volvocine MTLs/SDRs reside has remained the same across the lineage, a conclusion inferred from the conservation of genes in and around the MTLs/SDRs from different species [20,22]. This contrasts with some animal systems, for example, where new SDRs have emerged on what were originally autosomes as a result of genes on these chromosomes evolving a primary sex-determining role during evolution [29–31]. It is not known what mechanisms might bias the volvocine MTL/SDR towards remaining on the same chromosome when it turns over.



Trends in Plant Science

Figure 1. Emergence of UV Systems across the Eukaryotic Tree of Life. Absolute time in million years (Ma) is given based on [73]. Lineages where haploid sexual systems have been found are indicated by red branches. Lineages where haploid sexual systems are suspected to have emerged are indicated in orange. Simplified lineage phylogenies are based on [74] for Excavata, [75,76] for Opisthokonta, [77] for Archaeplastida, [78] for Haptophyta and Cryptophyta, and [79] for the SAR (Stramenopila, Alveolata, and Rhizaria) supergroup.

## Box 2. Models for the Evolution of UV Systems Compared to XY/ZW Systems

In the commonly accepted model for sex chromosome evolution, sex chromosomes evolve from autosomes, initially by the acquisition of a sex-determining locus. Emergence of sexually antagonistic alleles at loci in close proximity to the sex-determining locus selects for recombination suppression between the X and Y (or Z and W) chromosomes, resulting in formation of a first stratum, which undergoes heterochromatinization. Once recombination is arrested on the Y or W chromosome, genes without sex-specific benefits often become pseudogenes. The non-recombining region can expand with the acquisition of additional sexually antagonistic alleles and further recombination suppression, leading to additional strata (spatial clusters of XY or ZW gametologs with similar degrees of divergence). Strata have been observed in mammals, birds, fish and plants (reviewed in [3]). The lack of recombination leads to the accumulation of repetitive DNA, which can lead to a short-term increase in the size of the Y or W, but which typically results in large-scale deletions, a large reduction in physical size of the sex-limited chromosome, and highly heteromorphic sex chromosomes.

Figure 1 shows possible mechanisms involved in the evolution of the non-recombining sex- or mating type-determining regions on UV or haploid mating-type chromosomes (B–D) compared to diploid sex-chromosome systems (A). Note that an XY system is illustrated, but similar processes are expected to occur in ZW systems. Sexual antagonism has been proposed as a main driving force for the expansion of the SDR in diploid systems (reviewed in [27]). In UV systems, the Otto and Immler model predicts that reduced recombination is also favored provided that different alleles have different levels of fitness in males and female backgrounds. Note that the theoretical predictions of Immler and Otto are valid both for UV chromosomes that carry SDRs and for chromosomes with MTLs [13]. In isogamous and near-anisogamous organisms, forces other than sexual antagonism may contribute to the expansion of the non-recombining region [6]. These include for instance the capture and shelter of deleterious alleles in a permanently heterozygous state, or the fixation of neutral rearrangements by drift in one gametolog [6]. The non-recombining region can expand symmetrically on the U and V (B, C) chromosomes, but expansion can also occur independently in only one of the haplotypes by transposition of loci to one of the SDRs (D) (e.g., [55]).

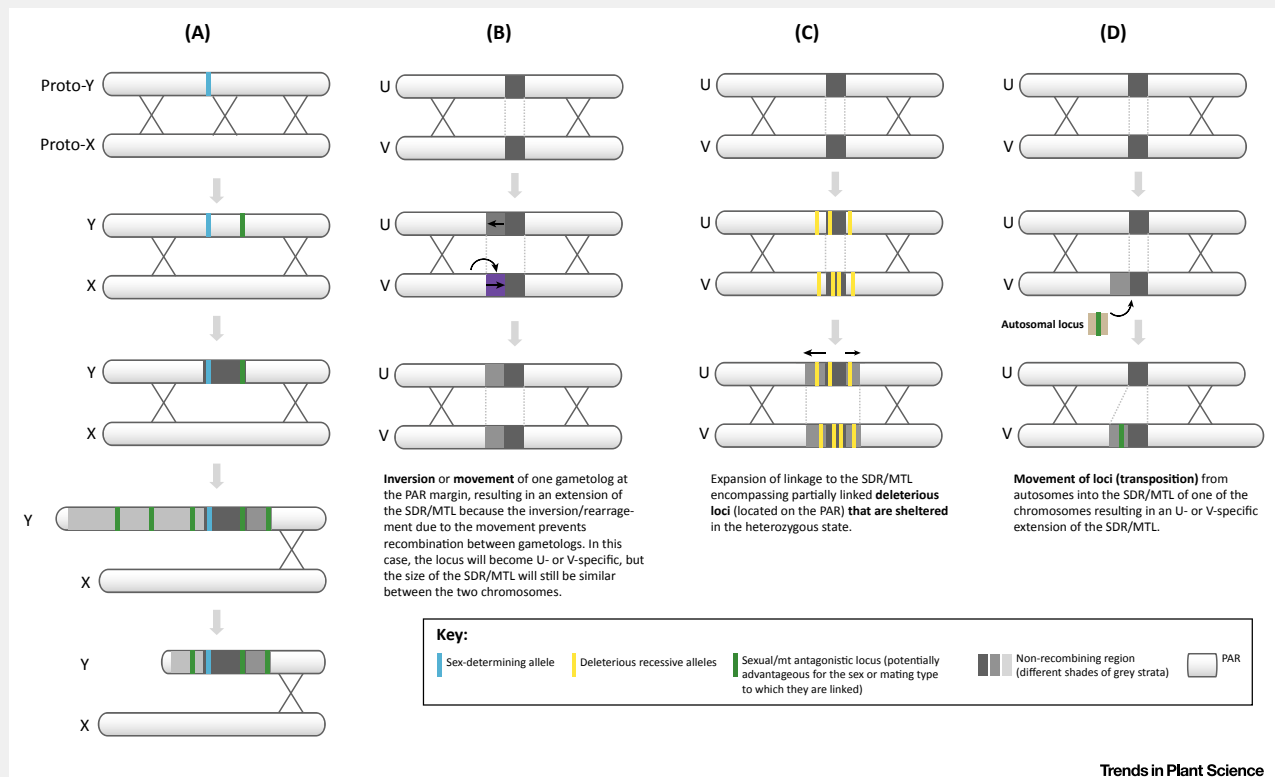


Figure 1. Schematic Model for the Evolution of UV Systems Compared to XY/ZW Systems.

## Recombination and Gametolog Divergence Patterns

Four volvocine MTLs/SDRs (*Chlamydomonas*, *Gonium*, *Yamagishiella*, and *Eudorina*) appear youthful based on very low divergence between gametologs, lack of gametolog decay, and relative paucity of intergenic repeats [21,22,32]. However, as in the case for some amphibian sex chromosomes, youthful-appearing MTLs/SDRs may not be so young [33]. A more extensive study of gametolog divergence and stratum formation in *Chlamydomonas* revealed that, although crossover recombination is undetectable across the MTL and flanking **pseudoautosomal**

Table 1. Structural Characteristics of UV SDRs across Different Organisms<sup>a,b</sup>

	Total sequence (Mb)			Number of genes			Gene density (genes/Mb)			Average gene length (bp)		
	VSR	USR	Genome	VSR <sup>c</sup>	USR <sup>c</sup>	Genome	VSR	USR	Genome	VSR	USR	Genome
<i>Ectocarpus</i> sp.	0.92	0.93	205	20 (11)	22 (11)	15779	22.82	23.7	76.9	25710	18836	6974
<i>U. partita</i>	1.0	1.5	N.d.	46 (23)	67 (23)	N.d.	46	44.6	N.d.	N.d.	N.d.	N.d.
<i>C. reinhardtii</i>	0.20	0.40	111	25 (22)	35 (22)	17732	118 (122)	109 (111)	159.6	3489	3482	3895
<i>G. pectorale</i>	0.5 <sup>d</sup>	0.37 <sup>d</sup>	149	24 (21)	24 (21)	17990	46	58	121	4726	3981	3993
<i>Y. unicocca</i>	0.17 <sup>d</sup>	0.27 <sup>d</sup>	134 <sup>e</sup>	18 (17)	18 (17)	N.d.	109	67.2	N.d.	4811	5052	N.d.
<i>Eudorina</i> sp.	0.007	0.09	184 <sup>e</sup>	3 (2)	3 (2)	N.d.	428	33.3	N.d.	1437	5112	N.d.
<i>V. carteri</i>	1.13	1.51	131	60 (50)	55 (50)	14958	54	39	114	6062	7198	5300
<i>M. polymorpha</i>	6	4.37	220	105 (19)	74 (20)	19470	17.6	17.4	88.5	N.d.	N.d.	N.d.
	Average intron length			GC (%)			Repeats (%)					
	VSR	USR	Genome	VSR	USR	Genome	VSR	USR	Genome			
<i>Ectocarpus</i> sp.	3605	3691	702	51.29	44.74	54.02	N.d.	N.d.	23.00			
<i>U. partita</i>	N.d.	N.d.	N.d.	N.d.	N.d.	N.d.	N.d.	N.d.	N.d.			
<i>C. reinhardtii</i>	358	354	420	61	60	64.1	N.d.	N.d.	N.d.			
<i>G. pectorale</i>	279	176	350	61	59.7	64.5	N.d.	N.d.	N.d.			
<i>Y. unicocca</i>	N.d.	N.d.	N.d.	60.3	60.1	61.1	N.d.	N.d.	N.d.			
<i>Eudorina</i> sp.	N.d.	N.d.	N.d.	51.4	53.9	61.0	N.d.	N.d.	N.d.			
<i>V. carteri</i>	584	618	400	53	52	56.1	70.00	72	21			
<i>M. polymorpha</i>	N.d.	N.d.	392	N.d.	N.d.	N.d.	74.7	70.8	22.2			

<sup>a</sup>Abbreviations: N.d., not determined; USR, U-specific region; VSR, V-specific region.

<sup>b</sup>References for *Ectocarpus* sp. [54,55], *U. partita* [52], *C. reinhardtii* [20,21,32]; *G. pectorale* [21]; *Y. unicocca* and *Eudorina* sp., T. Hamaji and H. Nozaki, personal communication; *V. carteri* [20,21]; *M. polymorpha* [43,44].

<sup>c</sup>The number of gametologs is indicated in brackets.

<sup>d</sup>Minimum size estimates for VSR and USR regions.

<sup>e</sup>Genome size is shown for the MT+ or female strain of *Yamagishiella* and *Eudorina*, respectively.

region (PAR) [23], there has clearly been a history of gene conversion between gametologs and between PAR genes that neighbor the MTL, evidenced by extensive allele sharing between mating haplotypes [32]. Using mutant or transgenic strains with reversed mating types, a test was performed for recombination when the two MTLs were collinear during meiosis (i.e., MT+/MT+ or MT-/MT-). Surprisingly, the homozygous MTL strains behaved differently from each other, and normal rates of recombination within the MTL were observed after meiosis with MT+/MT+ diploids, but no MTL recombination in MT-/MT- diploids (which had normal recombination outside the MTL [32]). It was inferred from this asymmetric behavior that there may be recombination suppressor sequences in the *Chlamydomonas* MT- that can act independently of sequence rearrangements. Another difference between the *Chlamydomonas* MT haplotypes is the insertion in MT+ of three autosome-derived regions and of a large tandem segmental repeat which together make the MT+ region (~396 kb) significantly larger than the MT- (~204 kb) [32]. Whether this size asymmetry is connected to the recombination asymmetry of the two MT haplotypes is unknown, but in any event it represents an example of a size difference arising in the SDR/MTL of a haploid system in the absence of any apparent differential selection between mating types or **sexual antagonism**. Interestingly, genes within each of the three autosome-derived regions in the MT+ mentioned above exhibited a wide range of neutral divergence rates from their autosomal counterparts, and could not be reliably binned into strata [32]; the relative timing of autosomal sequence additions to the *Chlamydomonas* MTL is therefore unclear.

### Box 3. Transitions between Monoicy and Dioicy

Transitions between dioicy and monoicy are common across all the eukaryotic groups that have haploid sex determination. This type of transition has occurred several times in volvocine algae [17], and, in mosses, transitions between dioicy and monoicy are very frequent and appear to have occurred a few hundred times [40,48]. In the brown algae, separate haploid sexes (dioicy) is clearly the ancestral state, with, again, several independent transitions to monoicy [53]. However, the evolutionary forces and the proximate mechanisms driving these transitions are still poorly understood. Monoicy is associated with polyploidy in mosses and liverworts [48], suggesting that diploid bisexual gametophytes originated from unreduced spores of UV diploid sporophytes of dioicous species – and thus possessed all the genes necessary for both male and female sexual functions. The male and female factors in some mosses are codominant, leading to monoicy when both the male and female haplotypes are present in the same gametophyte [65]. However, in other haploid systems the situation seems to be more complicated. Monoicous and dioicous hornworts have similar chromosome numbers [48], arguing against polyploidy as a mechanism for transition to monoicy. In addition, dominance has been observed in some UV systems, for example, the V or the U chromosome is dominant in *Ectocarpus* [54] and *Marchantia*, respectively [65]. In such systems, UV polyploids (i.e., diploid gametophytes) are not hermaphroditic, and transitions are unlikely to have been driven by changes in ploidy, suggesting the existence of alternative mechanisms. Epigenetic silencing of a master dominant male sex-determining gene in specific tissues could lead to monoicy by producing female organs if femaleness is the default state (as appears to be the case in *Ectocarpus* [54] and possibly *Volvox* [5]). Cases of monoicy have been reported in genetically male kelps [58]. Similarly, transitions from dioicy to monoicy in volvocine algae [17] may be related to epigenetic control of expression for the dominant sex-determining gene *MID* [5,66]. In fungi, transitions from heterothallism to homothallism have often evolved following gene capture. For instance, in many ascomycetes homothallic strains are not heterozygous diploids but instead contain copies of both mating types that can be alternatively expressed [67,68]. Additional studies on the evolution of reproductive traits and the correlation between life-cycle and reproductive features will be necessary to understand the molecular mechanisms, ultimate causes, and evolutionary consequences of the transitions between sex-determination modes.

### Box 4. Volvocine Algae and the Evolution of Sexes from Mating Types

Among the different UV systems that have been characterized to date, volvocine algae are unique in having members that span the whole range from isogamy to anisogamy to oogamy, and these organisms are therefore ideal models to study transitions between these states [22]. Despite their history of MTL/SDR structural turnover, the volvocine algae have retained homologous regulatory mechanisms for mating-type specification by the RWP-RK family transcription factor gene *MID* (minus dominance), which is found in either the *minus* mating type or the male SDR of all dioicous volvocine species characterized to date [20,22,24,25]. In *Chlamydomonas* the presence/absence of a *MID* gene, whose expression is induced by nitrogen deprivation, is the major determinant of minus/plus sexual differentiation [24,69]. A test of *MID* function in *Volvox* showed that, like the case in *Chlamydomonas*, the presence/absence of the *Volvox MID* ortholog is the key determinant of spermatogenic/oogenic development for germ cell precursors that are formed in response to a pheromone called sex inducer [5]. However, in the sex-reversed strains generated by ectopic *MID* expression in females, or by RNAi knockdown in males, the distinctly female or male patterns of germ-cell precursor formation (characterized by numbers, sizes, and positions of germ-cell precursors) were unaltered, indicating that sex-related developmental functions other than *MID* are encoded in the SDRs of the *Volvox* UV chromosomes. Moreover, while the germ cells that were formed in the sex-reversed strains of *Volvox* were functional, they had a variety of defects indicating that gene content in the male and female SDRs of *Volvox* has become masculinized and feminized, respectively [5]. It was hypothesized that the high degree of gametolog differentiation and expanded SDR size in *Volvox* might stem partly from sexual antagonism operating in this oogamous mating system [20]. Charlesworth [70] originally proposed a minimal two-locus model for the evolution of anisogamy where an allelically dimorphic gamete size-control gene comes into tight linkage with a sex-determining region. A prediction from this model is the presence of one or more dimorphic gamete size control genes being present in the SDRs of oogamous/anisogamous volvocines but not isogamous ones. However, more recent results have shown that anisogamy is compatible with a highly reduced SDR as is found in *Eudorina* sp. where the male haplotype spans 7 kb and contains only three genes, only one of which, *MID*, is likely to be related to sex or mating [22]. Thus, gametolog divergence as observed in *Volvox* may arise secondarily after anisogamy has already been established [20,71]. Unexpectedly, another recent study found that the *MID* gene from an isogamous volvocine genus, *Gonium*, which is more closely related to *Volvox* than is *Chlamydomonas*, could induce spermatogenesis in *Volvox* females [72]. This finding indicates that divergence of *Mid* function in volvocine algae was not the primary driver towards anisogamy/oogamy, and suggests that changes in other parts of the *Mid* regulatory network (interacting proteins and/or target genes) that are not encoded in the SDR were responsible for the evolution of sexually dimorphic gametes. The elucidation of *Mid* binding sites and target genes in selected volvocine species may shed light on how the *Mid* network expanded in multicellular volvocine algae to drive the evolution of gamete dimorphism.

In *Gonium*, *Yamagishiella*, and *Eudorina* MTLs/SDRs there are no insertions of autosome-derived sequences and very little neutral divergence between gametologs [21,22]. Whether this lack of gametolog differentiation is due to youthfulness of the MTLs/SDR, ongoing homogenization through gene conversion, or some combination of the two is not known. By contrast, the

*Volvox carteri* SDR is very different from the other four volvocine MTLs/SDR in terms of its large size (>1 Mb), low gene density, extensive gametolog differentiation, reduced codon adaptation, and completely arrested recombination that spans several speciation events [20]. These properties of the *Volvox* SDR make it conform more closely than the other volvocine MTLs/SDRs to the predicted properties of a 'mature' UV sex chromosome system.

### Ostreococcus

Prasinophyte marine picoalgae in the genus *Ostreococcus* have compact genomes (13 Mb) that are likely the result of genome reduction, but scans for meiotic genes and population genetic studies both support the retention of a sexual cycle in this group [34–37]. A recent study examined a candidate mating-type chromosome (Chr 2) from different *Ostreococcus tauri* isolates and found two divergent haplotypes designated *M*<sup>–</sup> and *M*<sup>+</sup>, with candidate MTL regions of 650 kb and 450 kb, respectively [34]. The genes within these candidate MTLs are suppressed for recombination and have high interhaplotype divergence that extends through one or more speciation events. These results indicate that the two candidate MTL haplotypes of *Ostreococcus* have persisted in the population for up to 600 Ma since the origins of the class Mamiellophyceae, which includes the genus *Micromonas* where evidence of meiotic genes and a possible MTL have also been described [38]. Interestingly, the *M*<sup>–</sup> candidate MTL haplotype has a predicted RWP-RK transcription factor gene that may function in mating-type determination like the volvocine algal gene *MID* (Box 4), although the function of this gene remains to be tested. While population genomics and protein-coding gene predictions provide compelling evidence for an active sexual cycle in *Ostreococcus*, sex has not been directly observed in this genus, and a formal association between the candidate *M*<sup>–</sup> and *M*<sup>+</sup> haplotypes with mating-type differentiation has yet to be made.

### Bryophytes

Bryophytes (liverworts, hornworts, and mosses) are an informal taxonomic grade of early-diverged land plants (embryophytes) which retain an ancestral gametophyte-dominant life cycle with reduced diploid sporophyte generation. Their sexual cycles are oogamous, and monoicous and dioicous species are found among members of all three groups [39,40] (also see Box 3). Many bryophytes are amenable to cytological evaluation, and surveys have revealed that some dioicous species possess dimorphic sex chromosomes that are easily distinguishable based on their sizes, while others appear to have homomorphic sex chromosomes [41].

The dioicous liverwort *Marchantia polymorpha* is an emerging model for early embryophyte evolution and developmental studies, including sex determination [42]. Partial characterization of its male (V) chromosome [43] has been followed more recently by full genome sequencing of male and female strains where candidate SDRs for both U and V chromosomes (referred to as X and Y in Bowman 2017 [44] and older literature) were identified [44]. The complete *M. polymorpha* genome is around 226 Mb in size with a V chromosome of around 10 Mb and a U chromosome estimated to be ~20 Mb [43]. The V chromosome has a ~4 Mb male-specific, low-complexity repeat region, while the U chromosome has a larger, less well characterized presumed repeat region that is at least partly composed of rDNA repeats [45,46]. The relatively gene-rich/high-complexity portions of the U and V chromosomes encompass 4.4 Mb and 6.0 Mb, respectively, and will be referred to as SDRs, but it should be noted that additional male- or female-specific genes could be contained in the non-assembled repeat regions of the *Marchantia* U and V.

The *Marchantia* U and V SDRs contain 75 and 99 total genes, respectively, including 20 gametologs that are expressed primarily in the vegetative phase and encode conserved green-lineage proteins (Table 1) [43,44]. The gametologs are saturated for neutral substitutions,



indicating long-term absence of recombination that likely extends back to the origins of the class Marchantiopsida. In addition, the male and female SDRs show signs of degeneration, with fivefold lower gene densities compared to autosomes and abundant transposon-derived sequences. A strikingly high proportion of the V-specific genes with detectable expression were expressed mainly during sexual reproduction (53/62), and several of them had annotated motility functions that are likely associated with spermatogenesis [43,44]. For U-specific genes, a majority with detectable expression (23/39) were also preferentially expressed during the reproductive phase [44], and among them may be one or more feminizer loci that can dominantly determine sex in diploid *Marchantia* gametophytes, which differentiate as sterile females [47]. Several MYB-family predicted transcription factors are among the sex-induced U-specific genes in *Marchantia*, but they do not appear to be conserved in related species. Notably, many of the sex-induced female genes were located on small scaffolds that are presumably embedded in repeat regions and could not be easily assembled, and there may be additional undetected U chromosome candidate feminizer genes that are not present in the current U chromosome assembly [44]. Overall, the *Marchantia* UV chromosomes conform to predictions regarding loss of recombination, gametolog differentiation, accumulation of non-coding repeat sequences, and preferential retention or acquisition of male/female-specific genes on the U/V.

Although bryophyte sex chromosome cytology has been extensively described [41], there has been little molecular characterization of UV systems outside *Marchantia* [48]. One notable example is the moss *Ceratodon purpureus* whose heteromorphic sex chromosomes are around fivefold larger than the autosomes and which have been genetically mapped [49]. There is currently no published genome sequence for *Ceratodon*, but several UV-linked gametologs and autosomal protein coding genes were identified and characterized from a population genetic study, together with their orthologs from closely related sister taxa [50]. This study uncovered significantly different levels of divergence between two subsets of the gametolog pairs, suggesting that at least two strata contributed to formation of the present-day *Ceratodon* sex chromosome. With the addition of genome sequences, including assembled U and V chromosomes, *Ceratodon* could become another very informative model for understanding the evolution of UV systems.

### *Ulva*

*Ulva partita* is a multicellular green algal species that exhibits a typical haplodiplontic life cycle with isomorphic gametophyte and sporophyte generations (Box 1). Male and female gametophytes are morphologically indistinguishable, and produce slightly anisogamous gametes with two mating types, *MT*<sup>−</sup> and *MT*<sup>+</sup> [51]. Recent sequencing of the UV chromosomes of this species revealed that its MTL spans 1–1.5 Mb of highly rearranged non-recombining sequence, with 46 and 67 genes, respectively, in the *MT*<sup>+</sup> and *MT*<sup>−</sup> haplotypes, of which about half are gametologs ([52]; Table 1). Suppression of U/V recombination appears to have preceded the diversification of the Ulvales (at ~166 Ma ago, <http://timetree.org/>). Like the case in volvocine algae, no obvious strata could be detected. This could be because they do not exist or because strata are no longer detectable owing to extensive rearrangements or divergence.

A particularly interesting feature of *Ulva* is that the gametophyte and sporophyte generations are isomorphic, and are therefore likely to require expression of similar genes in both life-cycle phases. As a result, the majority of the genome may be exposed to haploid purifying selection and its MTL should evolve under similar constraints as in haploid-dominant UV systems. Consistent with this idea, the *Ulva* MTL showed signs of weak degeneration, evidenced by relaxed codon usage for a subset of genes, decreased expression of haplotype-specific genes,

and lower gene density, although transposable element density was comparable with that of autosomes. Intriguingly, the *Ulva partita* MT– contains a gene of the RWP-RK family (*RWP1*) that exhibits an expression pattern consistent with a role in reproduction. The relationships between *RWP1*, the volvocine algal *MID* clade (Box 4), and other *RWP* genes from the green lineage were not clearly resolved in phylogenetic reconstructions, leaving open the question of whether *RWP1* is a *MID* ortholog or was convergently recruited for a putative (but still untested) role in *Ulva* sex determination.

### Brown Algae

The filamentous brown alga *Ectocarpus* has a haplodiplontic life cycle. Gametophytes and sporophytes are slightly dimorphic and there is a small but significant difference in size between male and female gametes [53]. The *Ectocarpus* UV sex chromosome SDRs exhibit accumulation of repeated DNA and low gene density (Table 1). The U and V SDRs have similar sizes and each contains a few dozen genes, about half of which are members of gametolog pairs [54]. Many of the *Ectocarpus* SDR genes have autosomal copies, and in some cases sex-specific genes appear to have moved into the SDR very recently [54,55]. Both the male and female SDRs show clear signs of degeneration, despite the predicted action of purifying selection during the haploid phase of the life cycle. *Ectocarpus* gametophytes and sporophytes are morphologically dimorphic [56], thus U- or V-specific genes that are expressed during the diploid sporophyte phase are expected to be sheltered and are therefore free to degenerate [11,57]. Consistent with this prediction, genes belonging to gametolog pairs that have a role during *Ectocarpus* gametophyte development have largely escaped degeneration and, conversely, a subset of genes that are expressed during the diploid phase show signs of greater degeneration. Interestingly, most female-specific SDR genes are weakly expressed and many are pseudogenized. This is consistent with the male sex being dominant, and suggests that female fate may be engaged in the absence of the master male sex-determining factor [54]. Interestingly, in another brown alga with UV chromosomes (*Undaria pinnatifida*), genetically male (V-bearing) individuals in some field populations may develop both male antheridial and female oogonial structures on the same individual (i.e., are monoicous [58]). Taken together, these findings suggest that the U SDR is not necessary for an individual to become a functional female. Note, however, that decreased fitness was observed in these monoicous strains, suggesting that some of the genes in the female SDR increase female fitness [58].

Comparative analysis of the U and V SDRs of several brown algal species has indicated that recombination between these two regions halted more than 100 Ma ago [54]. At least 26 genes are estimated to have been present in the ancestral SDR, and, although a set of six SDR genes have been consistently sex-linked over the 100 Ma period, there has been a remarkable level of gene traffic in and out the SDR, much the same as in volvocine algae (see above). Conserved sex-linked genes include two pairs of gametologs and a male sex-specific gene, which is strongly upregulated at fertility in *Ectocarpus* [55]. This male-specific gene is a predicted HMG-domain transcription factor. Interestingly, HMG-domain protein coding genes are also involved in mating-type and sex determination in fungi and mammals [59,60]. Given the dominance of male sexual differentiation in *Ectocarpus* UV diploid gametophytes, this gene is a strong candidate for the sex-determining gene, but confirmation of this hypothesis awaits functional validation. Currently no method is available to generate stable gene knockouts in *Ectocarpus*, although RNAi is an effective method for transient gene knockdown [61]. Naturally occurring strains, such as the monoicous *U. pinnatifida* strains described above [58], represent a potentially interesting resource for such mechanistic analyses if they use the same sex-determining gene as *Ectocarpus*.

The study of brown algal UV chromosomes has also shed light on the evolution of the PAR, a genomic region that has been largely understudied even in diploid systems. Recent studies on the *Ectocarpus* sp. PAR revealed an accumulation of physically linked clusters of genes with increased expression in the sporophyte (i.e., silenced in the gametophyte) in the *Ectocarpus* PAR. A mathematical modeling approach indicated that the PAR of UV systems is a favorable location for genes with an advantage for the sporophyte (provided that there is a difference in the strength of selection when they occur in male or females [62]). These results highlighted the potential impacts of life-cycle features on the evolution of UV sexual systems.

### Concluding Remarks and Future Perspectives

Our understanding of UV sex chromosomes has progressed rapidly in recent years. Bringing together information from the diverse systems currently under study, some general conclusions can already be drawn. UV sex chromosomes exhibit many of the unusual features identified in XY and ZW systems, such as the presence of often extensive non-recombining regions characterized by low gene densities and at least some evidence of gene degeneration. Specific theoretical predictions, such as symmetric evolution of the U and V SDRs, and a tendency for at least some gene function to be conserved within the SDR owing to haploid selection, have been partially confirmed. However, not surprisingly, the reality is more complex than theoretical predictions. The analyses of UV systems have even provided some novel insights in areas that have not been looked at in detail using XY or ZW systems. These include, for example, gene traffic in and out of the SDR, structural and evolutionary features of the PAR, and the evolution of sex chromosomes from MTLs.

Despite these impressive beginnings, a great deal still remains to be learned about UV sexual systems. There are marked differences between the patterns and rates of evolution of different UV systems, and the role of factors such as the degree of gamete and/or sporophyte/gametophyte dimorphism in influencing these differences needs to be investigated. Evolutionary strata have been detected in UV SDRs, but it is not yet clear whether they are a general feature of UV systems or why such regions do not always stably persist across related taxa. Sex chromosomes are known to play an important role during speciation (e.g., [31,63,64]), but this is another aspect of UV systems that has not yet been investigated. Importantly, more information is needed about how genes on UV chromosomes function in sex determination. A related question is the extent to which sexually antagonistic loci play a role in UV chromosome evolution (see Outstanding Questions). Sex-determining genes have been identified, or strong candidates are available, for several UV systems, but further work is needed in this area to obtain a general picture of UV sex-determining genes and the pathways they control. The finding that RWP-RK-encoding genes are linked to MTLs or SDRs in different green algal taxa poses the possibility of deep homology for sex determination in the chlorophytes, and the finding of an HMG-encoding gene as a possible brown algal sex-determining gene suggests a potentially intriguing convergence with fungi and metazoans. Another unknown aspect of sex determination for most UV systems is the degree to which both the U and the V are actively involved in determining sex, and, in instances where this is not the case, whether this might lead to asymmetry between the patterns of U and V chromosome evolution. A related question involves the mechanism by which UV systems and epigenetic sex-determination systems transition back and forth. Finally, the origins of UV chromosomes and their relationships to more ancestral mating systems remain unexplored in most taxa. On a broader evolutionary scale, the fucales within the brown algae and some bryophyte species provide opportunities to understand how ancestral UV systems might have evolved into XY or ZW chromosomal systems. More extensive genome characterization among key taxa coupled with functional analyses will be expected to provide important insights into all these exciting questions in the coming years.

### Outstanding Questions

How do sexes evolve from mating systems in organisms with haploid sex determination?

What are the mechanisms underlying the initiation of recombination suppression on nascent UV sex chromosomes?

What is the extent to which sexually antagonistic loci play a role in UV chromosome evolution?

What factors govern longevity and stability versus frequent turnover in UV systems?

Does the length of the haploid phase of the life cycle influence the rate of degeneration of the sex-linked region in UV systems through an effect on the strength of purifying selection?

Why does gene conversion between gametologs persist in some UV systems and not others?

What are the proximate mechanisms and evolutionary forces that drive transitions in sex determination between dioicy/monoicy and between haploid/diploid-phase sexual systems?

Are repeated findings of genes encoding RWP-RK or HMG transcription factors as putative sex-determination genes in UV systems the results of deep homology or convergence?

### Acknowledgments

Work in the Algal Genetics group is supported by the CNRS, Sorbonne Université, and the European Research Council (grant agreement 638240). Work in the laboratory of J.G.U. is supported by National Institutes of Health grants R01GM078376 and R01GM126557, and by National Science Foundation grants MCB1515220 and MCB1616820. The authors wish to thank Nicolas Perrin for fruitful discussions on this manuscript, Simon Bourdareau for help with Figure 1, Takashi Hamaji and Hisayoshi Nozaki for data on volvocine algal mating loci, and Deborah Charlesworth, Takayuki Kohchi, Shohei Yamaoka, and Sa Geng for providing the photographs shown in Figure 1 in Box 1.

### References

- Speijer, D. *et al.* (2015) Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8827–8834
- Lehtonen, J. *et al.* (2016) What do isogamous organisms teach us about sex and the two sexes? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371, 20150532
- Bachtrog, D. *et al.* (2014) Sex determination: why so many ways of doing it? *PLoS Biol.* 12, e1001899
- Bachtrog, D. *et al.* (2011) Are all sex chromosomes created equal? *Trends Genet.* 27, 350–357
- Geng, S. *et al.* (2014) Evolution of sexes from an ancestral mating-type specification pathway. *PLoS Biol.* 12, e1001904
- Branco, S. *et al.* (2017) Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *Proc. Natl. Acad. Sci. U. S. A.* 114, 7067–7072
- Billiard, S. *et al.* (2012) Sex, outcrossing and mating types: unsolved questions in fungi and beyond. *J. Evol. Biol.* 25, 1020–1038
- Nieuwenhuis, B.P.S. and James, T.Y. (2016) The frequency of sex in fungi. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150540
- Heitman, J. (2015) Evolution of sexual reproduction: a view from the fungal kingdom supports an evolutionary epoch with sex before sexes. *Fungal Biol. Rev.* 29, 108–117
- Wright, A.E. *et al.* (2016) How to make a sex chromosome. *Nat. Commun.* 7, 12087
- Bull, J.J. (1978) Sex chromosomes in haploid dioecy: a unique contrast to Muller's theory for diploid dioecy. *Am. Nat.* 112, 245–250
- Haldane, J.B.S. (1933) The part played by recurrent mutation in evolution. *Am. Nat.* 67, 5–19
- Immmler, S. and Otto, S.P. (2015) The evolution of sex chromosomes in organisms with separate haploid sexes. *Evolution* 69, 694–708
- Lewis, K.R. (1961) The genetics of bryophytes. *Trans. Br. Bryol. Soc.* 4, 111–130
- Umen, J.G. and Olson, B.J.S.C. (2012) Genomics of volvocine algae. *Adv. Bot. Res.* 64, 185–243
- Herron, M.D. and Michod, R.E. (2008) Evolution of complexity in the volvocine algae: transitions in individuality through Darwin's eye. *Evol. Int. J. Org. Evol.* 62, 436–451
- Hanschen, E.R. *et al.* (2018) Repeated evolution and reversibility of self-fertilization in the volvocine green algae. *Evol. Int. J. Org. Evol.* 72, 386–398
- Goodenough, U. *et al.* (2007) Sex determination in *Chlamydomonas*. *Semin. Cell Dev. Biol.* 18, 350–361
- Miyamura, S. (2010) Cytoplasmic inheritance in green algae: patterns, mechanisms and relation to sex type. *J. Plant Res.* 123, 171–184
- Ferris, P. *et al.* (2010) Evolution of an expanded sex determining locus in *Volvox*. *Science* 328, 351–354
- Hamaji, T. *et al.* (2016) Sequence of the *Gonium pectorale* mating locus reveals a complex and dynamic history of changes in volvocine algal mating haplotypes. *G3 Bethesda Md* 6, 1179–1189
- Hamaji, T. *et al.* (2018) Anisogamy evolved with a reduced sex-determining region in volvocine green algae. *Commun. Biol.* 1, 17
- Ferris, P.J. and Goodenough, U.W. (1994) The mating-type locus of *Chlamydomonas reinhardtii* contains highly rearranged DNA sequences. *Cell* 76, 1135–1145
- Ferris, P.J. and Goodenough, U.W. (1997) Mating type in *Chlamydomonas* is specified by mid, the minus-dominance gene. *Genetics* 146, 859–869
- Hamaji, T. *et al.* (2008) Identification of the minus-dominance gene ortholog in the mating-type locus of *Gonium pectorale*. *Genetics* 178, 283–294
- Nozaki, H. *et al.* (2006) Males evolved from the dominant isogamic mating type. *Curr. Biol.* 16, R1018–1020
- Charlesworth, D. (2017) Evolution of recombination rates between sex chromosomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372, 20160456
- Fraser, J.A. *et al.* (2004) Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *PLoS Biol.* 2, e384
- Myosho, T. *et al.* (2015) Turnover of sex chromosomes in celebensis group medaka fishes. *G3 Bethesda Md* 5, 2685–2691
- Miura, I. (2017) Sex determination and sex chromosomes in amphibia. *Sex. Dev.* 11, 298–306
- Kitano, J. and Peichel, C.L. (2012) Turnover of sex chromosomes and speciation in fishes. *Environ. Biol. Fishes* 94, 549–558
- De Hoff, P.L. *et al.* (2013) Species and population level molecular profiling reveals cryptic recombination and emergent asymmetry in the dimorphic mating locus of *C. reinhardtii*. *PLoS Genet.* 9, e1003724
- Stöck, M. *et al.* (2011) Ever-young sex chromosomes in European tree frogs. *PLoS Biol.* 9, e1001062
- Blanc-Mathieu, R. *et al.* (2017) Population genomics of picophytoplankton unravels novel chromosome hypervariability. *Sci. Adv.* 3, e1700239
- Derelle, E. *et al.* (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unravels many unique features. *Proc. Natl. Acad. Sci. U. S. A.* 103, 11647–11652
- Palenik, B. *et al.* (2007) The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci. U. S. A.* 104, 7705–7710
- Grimsley, N. *et al.* (2010) Cryptic sex in the smallest eukaryotic marine green alga. *Mol. Biol. Evol.* 27, 47–54
- Worden, A.Z. *et al.* (2009) Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324, 268–272
- Haig, D. (2016) Living together and living apart: the sexual lives of bryophytes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371
- McDaniel, S.F. *et al.* (2013) Recurrent evolution of dioecy in bryophytes. *Evol. Int. J. Org. Evol.* 67, 567–572
- Allen, C. (1945) The genetics of bryophytes II. *Bot. Rev.* 11, 260–287
- Bowman, J.L. (2016) A brief history of *Marchantia* from Greece to genomics. *Plant Cell Physiol.* 57, 210–229
- Yamato, K.T. *et al.* (2007) Gene organization of the liverwort Y chromosome reveals distinct sex chromosome evolution in a haploid system. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6472–6477
- Bowman, J.L. *et al.* (2017) Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* 171, 287–304.e15
- Fujisawa, M. *et al.* (2003) Evolution of ribosomal DNA unit on the X chromosome independent of autosomal units in the liverwort *Marchantia polymorpha*. *Chromosome Res.* 11, 695–703

46. Ishizaki, K. *et al.* (2002) Multicopy genes uniquely amplified in the Y chromosome-specific repeats of the liverwort *Marchantia polymorpha*. *Nucleic Acids Res.* 30, 4675–4681
47. Haupt, A. (1932) Beitrage zur Zytologie der Gattung *Marchantia* (L.). *Zeitsch. Ind. Abst. Vererbungslehre* 62, 367–428
48. Renner *et al.* (2017) The sex chromosomes of bryophytes: recent insights, open questions, and reinvestigations of *Frullania dilatata* and *Plagiochila asplenioides*. *J. Syst. Evol.* 55, 333–339
49. McDaniel, S.F. *et al.* (2007) A linkage map reveals a complex basis for segregation distortion in an interpopulation cross in the moss *Ceratodon purpureus*. *Genetics* 176, 2489–2500
50. McDaniel, S.F. *et al.* (2013) Recent gene-capture on the UV sex chromosomes of the moss *Ceratodon purpureus*. *Evol. Int. J. Org. Evol.* 67, 2811–2822
51. Kagami, Y. *et al.* (2008) Sexuality and uniparental inheritance of chloroplast DNA in the isogamous green alga *Ulva compressa* (Ulvothyceae). *J. Phycol.* 44, 691–702
52. Yamazaki, T. *et al.* (2017) Genomic structure and evolution of the mating type locus in the green seaweed *Ulva partita*. *Sci. Rep.* 7, 11679
53. Luthringer, R. *et al.* (2015) Sexual dimorphism in the brown algae. *Perspect. Phycol.* 1, 11–25
54. Ahmed, S. *et al.* (2014) A haploid system of sex determination in the brown alga *Ectocarpus* sp. *Curr. Biol.* 24, 1945–1957
55. Lipinska, A.P. *et al.* (2017) Multiple gene movements into and out of haploid sex chromosomes. *Genome Biol.* 18, 104
56. Lipinska, A.P. *et al.* (2018) Rapid turnover of life-cycle genes in the brown algae. *BioRxiv*. Published online March 30, 2018. <https://www.biorxiv.org/content/early/2018/03/29/290809>
57. Otto, S.P. *et al.* (2015) Evolution of haploid selection in predominantly diploid organisms. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15952–15957
58. Li, J. *et al.* (2014) Zoospore-derived monoecious gametophytes in *Undaria pinnatifida* (Phaeophyceae). *Chin. J. Oceanol. Limnol.* 32, 365–371
59. Marshall Graves, J.A. and Peichel, C.L. (2010) Are homologies in vertebrate sex determination due to shared ancestry or to limited options? *Genome Biol.* 11, 205
60. Idnurm, A. *et al.* (2008) Identification of the sex genes in an early diverged fungus. *Nature* 451, 193–196
61. Farnham, G. *et al.* (2013) Gene silencing in *Fucus* embryos: developmental consequences of RNAi-mediated cytoskeletal disruption. *J. Phycol.* 49, 819–829
62. Luthringer, R. *et al.* (2015) The pseudoautosomal regions of the UV sex chromosomes of the brown alga *Ectocarpus* exhibit unusual features. *Mol. Biol. Evol.* 32, 2973–2985
63. Irwin, D.E. (2018) Sex chromosomes and speciation in birds and other ZW systems. *Mol. Ecol.* Published online February 14, 2018. <http://dx.doi.org/10.1111/mec.14537>
64. Bracewell, R.R. *et al.* (2017) Rapid neo-sex chromosome evolution and incipient speciation in a major forest pest. *Nat. Commun.* 8, 1593
65. Allen, C.E. (1935) The genetics of bryophytes. *Bot. Rev.* 1, 269–291
66. Yamamoto, K. *et al.* (2017) Molecular evolutionary analysis of a gender-limited MID ortholog from the homothallic species *Volvox africanus* with male and monoecious spheroids. *PLoS One* 12, e0180313
67. Thynne, E. *et al.* (2017) Transition from heterothallism to homothallism is hypothesised to have facilitated speciation among emerging *Botryosphaeriaceae* wheat-pathogens. *Fungal Genet. Biol.* 109, 36–45
68. Yun, S.-H. *et al.* (1999) Evolution of the fungal self-fertile reproductive life style from self-sterile ancestors. *Proc. Natl. Acad. Sci.* 96, 5592
69. Lin, H. and Goodenough, U.W. (2007) Gametogenesis in the *Chlamydomonas reinhardtii* minus mating type is controlled by two genes, MID and MTD1. *Genetics* 176, 913–925
70. Charlesworth, B. (1978) The population genetics of anisogamy. *J. Theor. Biol.* 73, 347–357
71. Hiraike, R. *et al.* (2013) The evolution of male–female sexual dimorphism predates the gender-based divergence of the mating locus gene MAT3/RB. *Mol. Biol. Evol.* 30, 1038–1040
72. Geng, S. *et al.* (2018) Evolutionary divergence of the sex-determining gene MID uncoupled from the transition to anisogamy in volvocine algae. *Dev. Camb. Engl.* 145, <http://dx.doi.org/10.1242/dev.162537> publication date 04/09/2018
73. Parfrey, L.W. *et al.* (2011) Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl. Acad. Sci. U. S. A.* 108, 13624–13629
74. Simpson, A.G.B. *et al.* (2006) Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of 'primitive' eukaryotes. *Mol. Biol. Evol.* 23, 615–625
75. Cavalier-Smith, T. *et al.* (2014) Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts (animals, fungi, choanozoans) and Amoebozoa. *Mol. Phylogenet. Evol.* 81, 71–85
76. Cavalier-Smith, T. *et al.* (2015) Multigene phylogeny resolves deep branching of Amoebozoa. *Mol. Phylogenet. Evol.* 83, 293–304
77. Umen, J.G. (2014) Green algae and the origins of multicellularity in the plant kingdom. *Cold Spring Harb. Perspect. Biol.* 6, a016170
78. Burki, F. *et al.* (2016) Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centroleleida, Haptophyta and Cryptista. *Proc. Biol. Sci.* 283, 20152802
79. Burki, F. *et al.* (2007) Phylogenomics reshuffles the eukaryotic supergroups. *PLoS One* 2, e790

