Evidence for a genetic sex determination in Cnidaria, the Mediterranean red coral (Corallium rubrum)

M. Pratlong¹,², A. Haguenauer¹, S. Chenesseau¹, K. Brener³, G. Mitta³, E. Toulza³, M. Bonabaud⁴, S. Rialle⁴, D. Aurelle¹,† and P. Pontarotti²,†

¹Aix Marseille Univ, Avignon Université, CNRS, IRD, IMBE, Marseille, France
²Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
³Perpignan Via Domitia Univ, IHPE UMR 5244, CNRS, IFREMER, Montpellier Université,
Perpignan, France
⁴UMS BioCampus-MGX Montpellier GenomiX, Institut de Génomique Fonctionnelle,
141 rue de la Cardonnille, 34094 Montpellier Cedex 05, France

Cite this article: Pratlong M et al. 2017 Evidence for a genetic sex determination in Cnidaria, the Mediterranean red coral (Corallium rubrum). R. Soc. open sci. 4: 160880. http://dx.doi.org/10.1098/rsos.160880

Received: 3 November 2016
Accepted: 1 February 2017

Subject Category:
Genetics

Subject Areas:
evolution/genomics/ecology

Keywords:
genetic sex determination, Corallium rubrum, RAD-sequencing, Cnidaria

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3699073.

Sexual reproduction is widespread among eukaryotes, and the sex-determining processes vary greatly among species. While genetic sex determination (GSD) has been intensively described in bilaterian species, no example has yet been recorded among non-bilaterians. However, the quasi-ubiquitous repartition of GSD among multicellular species suggests that similar evolutionary forces can promote this system, and that these forces could occur also in non-bilaterians. Studying sex determination across the range of Metazoan diversity is indeed important to understand better the evolution of this mechanism and its lability. We tested the existence of sex-linked genes in the gonochoric red coral (Corallium rubrum, Cnidaria) using restriction site-associated DNA sequencing. We analysed 27 461 single nucleotide polymorphisms (SNPs) in 354 individuals from 12 populations including 53 that were morphologically sexed. We found a strong association between the allele frequencies of 472 SNPs and the sex of individuals, suggesting an XX/XY sex-determination system. This result was confirmed by the identification of 435 male-specific loci. An independent test confirmed that the amplification of these loci enabled us to identify males with absolute certainty. This is the first demonstration of a GSD system among non-bilaterian species and a new example of its convergence in multicellular eukaryotes.
1. Introduction

Sexual reproduction is ubiquitous among eukaryotes [1], and there is a wealth of literature on the evolutionary advantages of sex [2]. Although sex is widely shared, the corresponding mechanisms and the sex-determination systems vary greatly among species [3,4]. In gonochoric species, the sexual identity of individuals is defined by sex-determination systems, going from purely genetic sex determination (GSD), to purely environmental sex determination (ESD) where the same genotype can produce both male and female phenotypes depending on environmental conditions [3]. GSD is observed, for example, in mammals, where sex chromosomes are present. The echiurian *Bonellia viridis* gives an example of ESD, where larvae recruiting on a female will develop into males, but otherwise become female [5]. While GSD has been intensively studied in bilaterian species, one example has been recorded so far among non-bilaterians such as Porifera, Cnidaria and Ctenophora [3,4]. However, the quasi-ubiquitous repartition of GSD among eukaryotes suggests that similar evolutionary forces repeatedly led to the evolution of GSD, and these forces could occur also in non-bilaterians. This lack of evidence of GSD in non-bilaterians is probably the consequence of a reduced number of model organisms in these groups, and sometimes of the difficulty to identify separate sexes. Nevertheless, studying sex determination along the range of Metazoan diversity is important to understand better the evolution of this mechanism and its lability. For example, understanding the ancestral state of sex-determination systems in Metazoans requires studying them in the main branches of the phylogenetic tree.

Cnidarians display various sexual systems, from hermaphroditism (simultaneous or sequential) to gonochorism, and sexual reproduction can take place at different stages, polyp or medusae [6]. The corresponding system determinations are poorly known: a few ESD examples have been reported but no examples of GSD have so far been confirmed among non-bilaterians [6]. A cytogenetic analysis has shown a clear evidence of potential sex chromosomes in a scleractinian [7] but the role of these chromosomes in sex determination remains to be studied as this species is hermaphroditic. Gonochorism is highly predominant in octocorals (89% of the species [8]), even if cases of rapid transition between gonochorism and hermaphroditism have been demonstrated in the genus *Alcyonium* [9]. The relative stability of gonochorism in octocorals makes them interesting models for the study of sex-determination systems. This could, for example, correspond to an evolutionary trap, which would stabilize a sex-determination system [4], or, conversely, to a variety of sex-determination systems, but with a selective pressure for gonochorism.

The red coral (*Corallium rubrum*) is a long-lived gonochoric octocoral, with an age at first reproduction of around 7–10 years [10,11]. This harvested species is the object of ecological and population genetic studies [12,13] for conservation and management purposes. Although hermaphroditic colonies have been mentioned in the first description of the reproduction of this species [14], it is considered to be gonochoric. Red coral individuals from the two sexes are morphologically identical at the macroscopic scale, and the sex can be identified microscopically after dissection only during the period of gametogenesis (from May to September) [15]. Elucidating the sex-determination factors in this species would be useful for a better understanding of its biology and of its potential response to environmental change. It would also widen our knowledge of sex-determination systems in cnidarians. The identification of GSD may be difficult in groups where cytogenetic analyses are problematic to implement, as is often the case in non-model species, or when species lack visually heteromorphic sex chromosomes [16,17]. Thus, our goal here was to test the existence of sex-linked genes in the precious red coral, with a population genomic approach. We used restriction site-associated DNA sequencing (RAD-Seq) applied to sexed individuals, as proposed in [6]. Our results point to an XX/XY sex-determination system in this species, and we developed a polymerase chain reaction (PCR)-based protocol for sexing.

2. Material and methods

2.1. Sampling and DNA extraction

*C. rubrum* colonies were collected by scuba diving at two depths of two sites in three geographical regions of the Mediterranean Sea (Marseille, Banyuls, Corsica) between February and August 2013 (electronic supplementary material, figure S1 and table S1). Thirty individuals per site and depth were collected (360 individuals), preserved in 95% ethanol and stored at −20 °C. To validate our results, 40 additional individuals were collected in one site near Marseille in June 2016 and conserved in both 95%
ethanol and formaldehyde. Total genomic DNA was extracted according to the protocol of [18], followed by a purification using Qiagen DNeasy blood and tissue spin columns.

2.2. Morphological sex identification

Samples fixed in 95% ethanol from the Corsica populations were decalcified in 10% EDTA adjusted to a pH of 7.4 with NaOH for 48 h and dehydrated in graded alcohols, cleared in xylene substitute (Neo-clear VWR) and embedded in paraffin wax. Sections (7 µm) were cut using a rotary microtome. Sections were stained using trichrome of Masson variant of Goldner protocols and examined using a Leica DMLB. The sex of individuals from the Marseille population collected in June 2016 and preserved in formaldehyde was identified under a dissecting microscope after dissection.

2.3. Restriction site-associated DNA sequencing

Twelve RAD libraries were prepared according to the protocol described in [19], with small modifications (see the electronic supplementary material, methods). Libraries were sequenced on an Illumina HiSeq2000 using 100 bp single-end reads, at the Biology Institute of Lille (IBL, UMR 8199 CNRS) and at the MGX sequencing platform in Montpellier (France). The Stacks pipeline [20,21] was used for the loci de novo assembly and genotyping. We applied several filters to the resulting dataset in order to filter for poor-quality single nucleotide polymorphisms (SNPs) and artefacts due paralogous sequences (electronic supplementary material, table S2 and methods).

2.4. Identification of sex-linked loci

The presence of sex-linked loci, i.e. loci common to the two sex chromosomes but presenting differences in allele frequencies between sex (figure 1a), was explored by performing a principal component analysis (PCA) using the package adegenet in R [22,23]. This analysis was performed on the total dataset (12 populations). The dataset was centred and missing data were replaced by the mean allele frequency for each locus.

2.5. Identification of sex-specific loci

We filtered the Stacks catalogue in order to search among morphologically sexed individuals for loci present in all individuals from one sex (one read or more by individual) and absent in all individuals from the other sex (no read detected) (i.e. sex-specific loci; figure 1a).

2.6. Real-time polymerase chain reaction

Primers were designed for six putative male-specific sequences using IDT online tool (http://eu.idtdna.com/Primerquest/Home/Index) (primer sequences in electronic supplementary material, table S3). To avoid a specific hybridization, we selected loci presenting no blast hit results other than themselves against the Stacks catalogue. The presence or absence of target genomic sequences was assessed by real-time PCR on 1 µl of the same DNA extracts of the 58 morphologically sexed individuals from Corsica used for RAD-Seq, and for the 39 supplementary morphologically sexed individuals from Marseille (see the electronic supplementary material, methods). The difference of amplification between males and females was tested with a Wilcoxon–Mann–Whitney test in R [23].

3. Results and discussion

3.1. Morphological identification of sex

The sexing of the 58 individuals from four populations in Corsica was undertaken on the basis of gonad analysis. Among them, 25 males and 28 females were formally identified (hereafter called morphological males and females). Five individuals were sexually undetermined because of the quality of the tissues, or due to a sexually immature stage of development. Among the 40 individuals from the sampling of June 2016 in Marseille, we identified 15 males and 24 females (one individual was sexually undetermined).
Figure 1. (Caption opposite.)
Figure 1. (Opposite, top) (a) Example of a genetic sex-determination system. Sex-specific markers (blue) are present on the Y chromosomes in the XX/XY system and in the W chromosome in the ZW/ZZ system. Sex-linked markers (green) are common to the two sex chromosomes but should present differences in allele frequencies between sex because of the recombination arrest (adapted from [15]). (b) Mean allele frequencies expected in the case of sex-linked markers fixed on the X/Z and on the Y/W chromosomes. For those markers, individuals of the heterogametic sex (males in the XX/XY system, females in the ZW/ZZ system) should be heterozygous and individuals of the homogametic sex (females in the XX/XY system, males in the ZW/ZZ system) should be homozygous. (c) Mean allele frequencies of loci with the highest contribution to the PCA fifth axis observed in PCA-identified males and females. (d,e) Distribution of loci absent in all morphologically sexed individuals of one sex in morphologically sexed individuals of the opposite sex. (f) Results of the real-time PCR amplification (threshold cycle Ct) of one male-specific locus (Locus_139082) on morphologically sexed males and females from the Corsica populations. The Ct indicates the number of PCR cycles necessary to reach a threshold value; a low Ct value indicates a high amplification rate from the corresponding sample. Results for the six loci tested were similar (electronic supplementary material, figure S3). (g) Result of the real-time PCR amplification (threshold cycle Ct) of one male-specific locus (Locus_139082) on morphologically sexed males and females from the Marseille population. Results for the six loci tested were similar (electronic supplementary material, figure S4).

3.2. RAD-tag sequencing and quality filtering

RAD-tag sequencing generated an average of 187 ± 21 million reads per library before any quality filtering. The quality filtering step enabled us to remove an average of 2.02% of reads without a correct restriction enzyme cut site, an average of 1.69% of reads with ambiguous barcodes and an average of 1.61% of reads with low quality score. An average of 183 ± 22 million reads per library (95% of total reads) were retained (electronic supplementary material, table S4) with an average of 5.7 million reads per individual. After assembly (see details in electronic supplementary material, methods), 138 810 SNPs were successfully genotyped in at least 75% of individuals from all populations (electronic supplementary material, table S2). Individuals were sequenced with a mean coverage of 37 reads per individual per locus. Morphologically sexed females presented a mean coverage of 37.9 reads per individual per locus, and males 45.6 reads per individual per locus. A total of 27 461 SNPs remained after several steps of filtration (electronic supplementary material, table S2 and methods).

3.3. Identification of sex-linked loci

While the first four principal components of the PCA highlighted neutral population genetic differences [24], the fifth axis (1.58% of explained variance) separated individuals from all populations in two clear groups that matched the male/female repartition of morphologically sexed individuals (figure 2). Only three individuals that were morphologically female appeared among males in the PCA. From this PCA, we identified potential females and males among individuals for which sex determination was not performed on the basis of the visible separation in two groups. Considering this separation, we postulate that the dataset comprised 169 females and 185 males, corresponding to a balanced sex ratio \( p = 0.40 \), electronic supplementary material, table S5). The sex ratio was balanced also inside each population, except in ELV12, where it was significantly biased towards male individuals (70% of males; \( p = 0.03 \), electronic supplementary material, table S5). There were 472 SNPs with a contribution to the fifth axis of the PCA higher than 1% (electronic supplementary material, figure S2); these loci displayed the same allele fixed in almost all morphological females, and were at the heterozygous state in almost all morphological males. This observation was confirmed with PCA-identified males and females (figure 1b,c). The genotypes of these 472 SNPs enabled the sexing of 95% of individuals for the whole dataset. Among these SNPs, when considering only morphologically sexed individuals, 379 SNPs were fixed in all females, 59 were at the heterozygous state in all males (the remaining SNPs were at the heterozygous state in almost all males). Fifty-five SNPs were in these two categories and were therefore diagnostic of sex for the morphologically sexed individuals. Identification of such SNPs that were homozygous in females and heterozygous in males suggested an XX/XY sex-determination system in the red coral, with a non-recombining XY-like region. Of the SNPs leading the fifth PCA axis, 347 were fixed in all PCA-identified females. However, none of these SNPs was at the heterozygous state in 100% of PCA-identified males. SNPs that were strictly homozygous in females and heterozygous in males were found inside each population and geographical region, but none was common to the three geographical regions. Furthermore, even if we did not observe markers that were heterozygous in 100% of males, these markers remain at the heterozygous state in the majority of males: 303 of these 472 sex-linked markers were at the heterozygous state in more than 70% of all males. The absence of
markers diagnostic of sex in the overall dataset may indicate that these loci were submitted to a low but non-null recombination rate, as is the case when sexual chromosomes have recently diverged, or near the boundary of pseudoautosomal regions, where recombination is more frequent than in fully sex-linked regions [25,26]. It could also be the result of polymorphism within the restriction sites on the Y chromosome (i.e. allele dropout) making it impossible to observe the Y allele with RAD-Seq in some populations [27]. A low rate of mutation, recombination or allele dropout could suffice to make a marker go from diagnostic (heterozygous in 100% of males) to sex-linked but not diagnostic (heterozygous in less than 100% of males). For these different reasons, it may be difficult to identify strictly diagnostic loci common to all populations for such species with a strong genetic structure. Nevertheless, even if diagnostic markers of sex were not found at the scale of the overall dataset, the multilocus analysis enabled us to identify clearly males and females. Sex-linked polymorphisms identified by RAD-Seq have been found in the pistachio (*Pistacia vera*) and the salmon louse (*Lepeophtheirus salmonis*), where authors identified markers being heterozygous in females and homozygous in males, and suggested a ZW/ZZ system [28,29]. These polymorphisms have also been found in the Atlantic halibut (*Hippoglossus hippoglossus*), the Nile tilapia (*Oreochromis niloticus*) and the date palm (*Phoenix dactylifera*), where an XX/XY system was suggested [30–32]. Here, we highlighted a strong signal of sex-linked markers in our population genomic dataset. Considering the pattern of allele frequencies of these loci, they may be detected as being under balanced selection, which can easily lead to a misinterpretation of the signal when these loci are not expected (especially in species whose sex-determining system is unknown).

### 3.4. Identification of sex-specific loci

We identified 435 loci present in all 25 morphological males and absent in the 28 morphological females (figure 1a,d,e). To avoid any bias caused by an eventual misidentification of the sex of the three misclassified females, these individuals have not been taken into account for this analysis. To confirm the male specificity of these 435 loci, we targeted six of them in real-time PCR of morphologically sexed individuals. The six loci could be amplified in 100% of morphologically sexed males. Almost no amplification signal was observed in 100% of morphologically sexed females (figure 1f; electronic supplementary material, figure S3; $p < 0.01$ in all tests, electronic supplementary material, table S6). The presence of male-specific loci supports our previous hypothesis of a system with male heterogamety (i.e. XX/XY) [16,17,33]. Finally, in order to confirm that the male-specific markers identified from only one geographical region were not the result of random divergence between sexes, we applied the real-time PCR test to the 40 individuals from the additional June 2016 sampling in Marseille. The six...
male-specific primer pairs enabled us to amplify 100% of morphologically sexed males by real-time PCR. Almost no signal was observed in 22 of the morphologically sexed females (figure 1g; electronic supplementary material, figure S4; p < 0.01 in all tests, electronic supplementary material, table S6). Two females presented an amplification profile of the six markers similar to that of the males, as was also the case for three females from Corsica. This confirms that these loci are male-specific, and that they are conserved between the two distant populations.

3.5. Cross-validation

Finally, as a validation test, we crossed the results obtained by PCA and by sex-specific loci by searching the presence of the 435 male-specific loci in all 354 individuals (figure 3). Twenty-four per cent of PCA-identified males possessed all 435 male-specific loci and 80% of the individuals possessed 90% of these loci. Some of these loci may have been lost during library preparation and sequencing, or due to allele dropout, thereby explaining their absence in some of the males. Furthermore, the male-specific loci have been found from 25 individuals, and it is likely that some of them have been detected erroneously if the pattern of technical missing data for a locus followed by chance the male/female distribution. All of the 435 male-specific loci were absent in 40% of PCA-identified females, and 97% of PCA-identified females contained less than 1% of male-specific loci. Figure 3 illustrates the correlation between the three methods of sex identification presented here (coordinate on the fifth axis of the PCA, number of male-specific loci and morphological identification). We observed an extremely good correlation between these three methods, the groups of males and females being well defined in each case. The three misclassified females identified from the PCA also possessed a high number of male-specific loci, and seemed to be genetically male. Such incongruence between sex genotype and phenotype may result from an environmental sexual reversal during sexual differentiation [34], or from the existence of females with the XY genotype, as already described in several mammals [35]. Finally, four individuals that have not been sexed morphologically presented female genotypes but between 15 and 65 male-specific loci. Further analysis is needed to determine the sexual identity of these individuals.

The number of sex-specific markers should depend on the size of the non-recombining region and on the divergence between X and Y, from a minimal Y-specific region for homomorphic sex chromosomes to several sex-specific markers in heteromorphic sex chromosomes [33]. The 435 male-specific loci detected here suggested at least 218 male-specific PstI sites, which is far higher than the number of sex-specific loci in other similar studies based on RAD-Seq [17,33,36]. However, considering the specificity of each study (genome size, frequency of the restriction site and parameters used for the loci assembly), we could not directly compare the number of sex-linked markers or the number of male-specific markers.
Finally, the functional annotation of loci in the Stacks catalogue (see the electronic supplementary material, methods) enabled us to identify a homologue of a double-sex and mab3-related transcription factor (Dmrt). The Dmrt family of transcription factors is involved in sex determination in numerous metazoans, and the conservation of this function in cnidarians has recently been shown [37,38]. In our case, this gene was neither a sex-specific marker nor a sex-linked marker, which remains compatible with a role in sexual differentiation.

4. Conclusion

In summary, the exploration both of sex-linked polymorphisms and of sex-specific loci enabled us to identify an XX/XY genetic sex-determination system in the red coral. This is the first time such an identification has been made for non-bilaterian species, and is a new example of the evolution of GSD in multicellular organisms [3,39]. As gonochorism is predominant among octocorals [8], they constitute a promising group for further exploration of GSD among Cnidaria. In hexacorals, the sister-group of octocorals [40,41], there is apparently a higher diversity of reproduction systems than in octocorals [6]. Extending our study to other anthozoans would be useful in order to test scenarios of evolution of reproduction systems in this group [6]. Furthermore, our study has shown that signal of GSD, if it is unexpected, can be misinterpreted in non-model organisms (for example because sex-linked markers could be detected as being under balanced selection). Our research also emphasizes that population genomic datasets should be analysed and interpreted by taking into account the possibility of GSD in non-model organisms. We provide here real-time PCR primers that will enable the identification of red coral males and females, and that will facilitate the monitoring of the population dynamics of this emblematic species, which is being increasingly submitted to anthropic pressures such as harvesting and global change [42–44].

Ethics. The experiments performed during this study are in agreement with the French law on animal experiments (law no. R 214-87 to 90 of ‘Code rural’). The sampling was authorized by the French Ministry of Ecology, Sustainable Development, Transports and Housing (authorization no. 009, date: January 2013; authorization no. 943, date: 17 December 2013; authorization no. 150, date: 12 March 2013).

Data accessibility. The raw DNA sequences are available in the Short Read Archive (SRA) database under accession no. SRR5186771-SRR5187129 [45]. The filtered SNP file is available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.rs7bm [46].

Authors’ contributions. D.A., P.P., A.H., G.M. and E.T. conceived the project. A.H. and M.P. performed library preparation. M.B. performed the library sequencing. S.R. performed the bioinformatic quality control of sequenced libraries. A.H. performed the real-time PCR. S.C. morphologically determined the sex of individuals. M.P. analysed the data. M.B. performed the real-time PCR. S.C. morphologically determined the sex of individuals. M.P. analysed the data. M.B. performed the library sequencing. S.R. performed the bioinformatic quality control of sequenced libraries. A.H. performed the real-time PCR. S.C. morphologically determined the sex of individuals. M.P. analysed the data. M.B. performed the library sequencing. S.R. performed the bioinformatic quality control of sequenced libraries. A.H. performed the real-time PCR. S.C. morphologically determined the sex of individuals. M.P. analysed the data. M.B. performed the library sequencing. S.R. performed the bioinformatic quality control of sequenced libraries. A.H. performed the real-time PCR. S.C. morphologically determined the sex of individuals. M.P. analysed the data. M.B. performed the library sequencing. S.R. performed the bioinformatic quality control of sequenced libraries. A.H. performed the real-time PCR. S.C. morphologically determined the sex of individuals. M.P. analysed the data. M.B. performed the library sequencing. S.R. performed the bioinformatic quality control of sequenced libraries. Authors declare we have no competing interests.

Funding. This project has been funded by the ADACNI programme of the French National Research Agency (ANR) (project no. ANR-12-ADAP-0016; http://adacni.imbe.fr) and ECCOREV Research Federation (FR 3098). This work is a contribution to the Labex OT-Med (no. ANR-11-LABX-0061) funded by the French Government ‘Investissements d’Avenir’ programme of the French National Research Agency (ANR) through the A*MIDEX project (no. ANR-11-IDEX-0001-02).

Acknowledgements. We thank Nicolas Fernandez and Béatrice Loriod from the Marseille TGML platform for their invaluable help and advice with the preparation of the RAD libraries; Véronique Dhennin from the Biology Institute of Lille; and the team of the MGX platform for the sequencing of the RAD libraries. We thank the molecular biology service of the IMBE and Frédéric Zuberer of the Pytheas Institute for their support in sampling. We thank the Scandola Natural Reserve, especially Jean-Marie Dominici, and also thank Alexander Ereskovsky for his advice in red coral sexing.

References
