



Article

# Size and Content of the Sex-Determining Region of the Y Chromosome in Dioecious *Mercurialis annua*, a Plant with Homomorphic Sex Chromosomes

Paris Veltsos <sup>1,2,\*</sup> , Guillaume Cossard <sup>1</sup>, Emmanuel Beaudoin <sup>3</sup>, Genséric Beydon <sup>4</sup>,  
Dessislava Savova Bianchi <sup>1</sup> , Camille Roux <sup>1,5</sup>, Santiago C. González-Martínez <sup>1,6</sup>  
and John R. Pannell <sup>1</sup>

<sup>1</sup> Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; guillaume.cossard@unil.ch (G.C.); dessislava.savovabianchi@unil.ch (D.S.B.); camille.roux.1983@gmail.com (C.R.); santiago.gonzalez-martinez@pierroton.inra.fr (S.C.G.-M.); John.Pannell@unil.ch (J.R.P.)

<sup>2</sup> Department of Biology, Jordan Hall, 1001 East Third Street, Indiana University, Bloomington, IN 47405, USA

<sup>3</sup> Faculty of Biology and Medicine, University of Lausanne, Bâtiment Génopode, 1014 Lausanne, Switzerland; emmanuel.beaudoin@unil.ch

<sup>4</sup> National Centre for Genomic Resources (CNRGV), 24 Chemin de Borde Rouge—Auzeville—CS52627, 31326 Castanet Tolosan Cedex, France; genseric.beydon@toulouse.inra.fr

<sup>5</sup> CNRS, University of Lille, UMR 8198—Evo-Eco-Paleo, F-59000 Lille, France

<sup>6</sup> BIOGECO, INRA, University of Bordeaux, 33610 Cestas, France

\* Correspondence: parisveltsos@gmail.com

Received: 10 April 2018; Accepted: 23 May 2018; Published: 29 May 2018



**Abstract:** Dioecious plants vary in whether their sex chromosomes are heteromorphic or homomorphic, but even homomorphic sex chromosomes may show divergence between homologues in the non-recombining, sex-determining region (SDR). Very little is known about the SDR of these species, which might represent particularly early stages of sex-chromosome evolution. Here, we assess the size and content of the SDR of the diploid dioecious herb *Mercurialis annua*, a species with homomorphic sex chromosomes and mild Y-chromosome degeneration. We used RNA sequencing (RNAseq) to identify new Y-linked markers for *M. annua*. Twelve of 24 transcripts showing male-specific expression in a previous experiment could be amplified by polymerase chain reaction (PCR) only from males, and are thus likely to be Y-linked. Analysis of genome-capture data from multiple populations of *M. annua* pointed to an additional six male-limited (and thus Y-linked) sequences. We used these markers to identify and sequence 17 sex-linked bacterial artificial chromosomes (BACs), which form 11 groups of non-overlapping sequences, covering a total sequence length of about 1.5 Mb. Content analysis of this region suggests that it is enriched for repeats, has low gene density, and contains few candidate sex-determining genes. The BACs map to a subset of the sex-linked region of the genetic map, which we estimate to be at least 14.5 Mb. This is substantially larger than estimates for other dioecious plants with homomorphic sex chromosomes, both in absolute terms and relative to their genome sizes. Our data provide a rare, high-resolution view of the homomorphic Y chromosome of a dioecious plant.

**Keywords:** bacterial artificial chromosomes (BAC); RNA sequencing (RNAseq); genetic map; transposable element; gene density

## 1. Introduction

Most flowering plants are hermaphroditic or monoecious, but dioecy has evolved frequently and is found in about half of all plant families [1,2]. Although sex in many animals is determined by environmental triggers [3], in almost all dioecious plants studied so far it appears to be determined genetically (though see [4,5]), usually at a single genetic locus within a non-recombining sex-determining region (SDR) on a sex chromosome [6,7]. Many plant sex chromosomes are cytologically heteromorphic (currently known for 19 species in four families), but cytological differences between males and females are not evident in others (20 species in 13 families; reviewed in [7,8]). Indeed, closely related dioecious species may often differ in terms of their degree of heteromorphism [6,7]. We might expect the magnitude of the cytological difference between homologous sex chromosomes to increase with their age, as the result of the progressive genetic degeneration of the Y or W chromosome (in species with XY or ZW systems, respectively). However, although there is some evidence for this expectation [9], there are many exceptions. For instance, in *Coccinia grandis* (Cucurbitaceae), a species that evolved dioecy about three million years ago [10], the X and Y chromosomes are highly divergent, with a 10% elongation of the Y compared to the X. In contrast, the Y chromosome is smaller than the X in the palm genus *Phoenix*, in which the sex chromosomes may have diverged >50 million years ago [11–13].

Sex chromosomes are expected to diverge in length through the accumulation of repetitive sequences in the non-recombining chromosome (a phenomenon that appears to be particularly common in plants with heteromorphic sex chromosomes) or through the loss of genes [7]. Both processes can be attributed to the reduced efficacy of purifying selection in the SDR following suppressed recombination [14], which may have evolved because of advantages in linking the sex-determining locus with alleles that differentially benefit male or female fitness [15], or because of the accumulation of repetitive elements themselves [16]. Whereas the importance of sexually antagonistic selection in the evolution of suppressed recombination is plausible [15], there is still little empirical evidence for it. We also do not understand why X and Y (or Z and W) chromosomes of some species diverge rapidly, quickly becoming heteromorphic, while others remain homomorphic. We also remain largely ignorant about how the size of the SDR relates to its content, not only for species with putatively large SDR, such as *Silene latifolia* [17], but also for species with homomorphic sex chromosomes. Importantly, while species with homomorphic sex chromosomes may have an SDR that is small (even restricted to a single gene), their SDR might also be large.

Characterization of the physical size of the SDR is challenging because of its often highly repetitive content. It is possible to compare the DNA content of males and females and attribute the difference to the sex chromosomes. As expected, such comparisons have revealed larger differences in species with heteromorphic sex chromosomes (see Supplementary Table S1, and [7]). However, to determine the actual size of the non-recombining SDR of the chromosome of the heterogametic sex (Y or W), it is necessary to identify markers on the sex chromosome, and to use them to build genetic maps that estimate the region that does not recombine in one sex, e.g., [18,19] (Supplementary Table S1). Alternatively, sex-linked markers may be used to identify, and then to sequence, bacterial artificial chromosomes (BACs) that contain long sections of the SDR [20], with subsequent assembly and potential chromosome-walking [21,22]. We take this approach in the current paper to assess the size and content of the SDR of the Y chromosome of the diploid dioecious plant *Mercurialis annua*.

*Mercurialis annua* is a polyploid complex that shows striking variation in its sexual system, ranging from diploid populations that have fully separate sexes and a homomorphic XY sex-determination system, to androdioecy (where males co-occur with hermaphrodites) and monoecy [23–26]. The unusual variation in the sexual system of the *M. annua* complex lends itself to testing a number of hypotheses about the selection of combined versus separate sexes [27–29] and the evolution of sex determination and sex chromosomes [24,30]. Previous work, based on de novo sequencing, genetic mapping of SNPs from open reading frames (ORFs) segregating in crossing families, and genome capture from males and females sampled across the species range (representing about 7% of *M. annua*

genome; [31]), has allowed the assembly of the diploid *M. annua* genome into eight linkage groups (corresponding to the  $2n = 16$  chromosomes of the diploid karyotype) and the identification of 568 sex-linked transcripts on the largest linkage group (i.e., chromosome 1), representing about 33% of the genes on the chromosome [30]. Genome and transcriptome analyses suggest that the SDR of *M. annua* is mildly degenerate, with a single gene interrupted by a premature stop codon and the accumulation of transposable elements and other repetitive DNA [30].

Here, we aimed (1) to estimate the physical length of the diploid *M. annua* SDR in relation to the size of the Y chromosome and to the rest of the genome, and (2) to characterize its content and genomic structure, including the identification of additional sex-linked genes that were either not previously mapped (due to the absence of suitable variation), or that were not among the ORFs used for previous mapping [30]. To estimate the size of the SDR, we identified male-specific polymerase chain reaction (PCR) products on ORFs, and used them to identify and sequence Y-linked BACs from two males. We inferred the minimum size of the SDR in terms of the size of the sex-linked region in the genetic map of *M. annua* [30] associated with the BACs. Our analysis suggests that the non-recombining SDR of the Y chromosome of *M. annua* is among the largest known for a homomorphic plant sex chromosome.

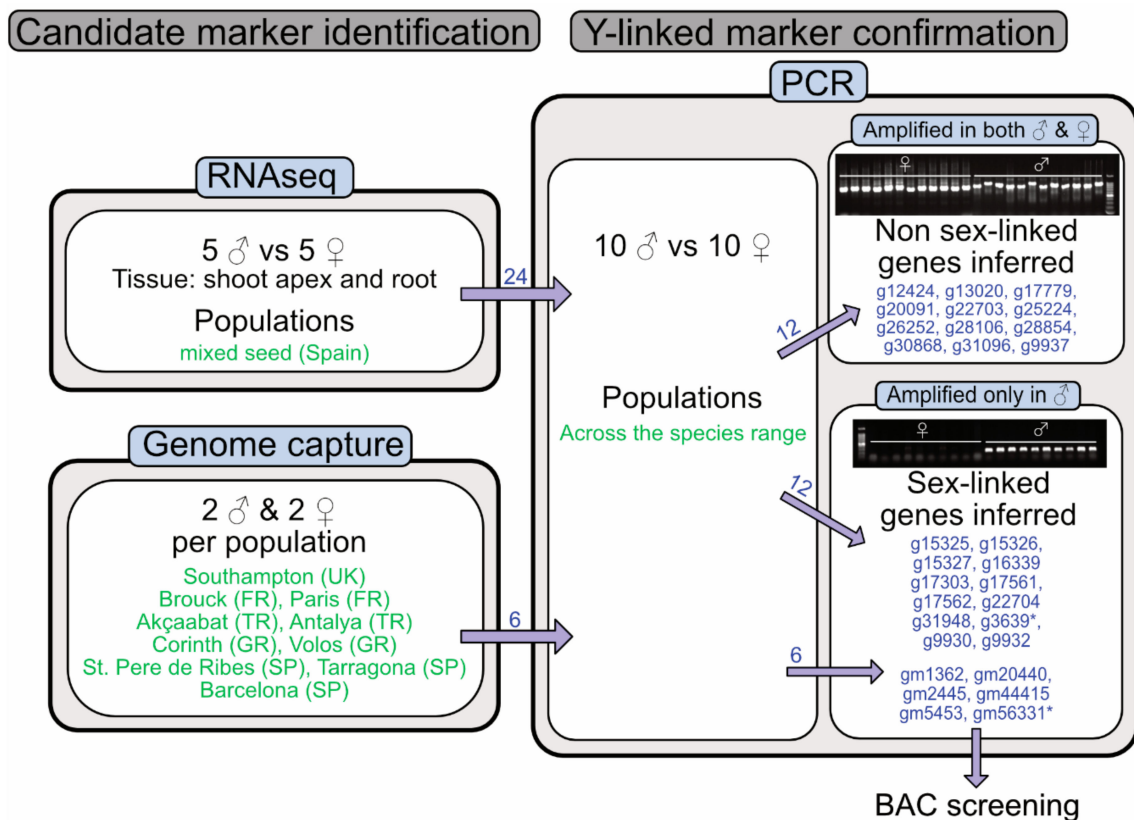
## 2. Materials and Methods

### 2.1. Overview

Our approach is summarized in Figure 1. We used male-specific gene expression [32] and male-specific genome sequences from a large sample of males and females of *M. annua*, sampled from across the species' range [31], to identify potential Y-linked markers. We verified the Y-linked status of these markers by male-specific PCR amplification and used Sanger sequencing to verify that the amplified sequences were similar to the expected transcripts. We then used PCR reactions to probe a BAC library constructed from two diploid *M. annua* males, collected near Lausanne in Switzerland. Finally, we used long-read (Pacific Biosciences, Menlo Park, CA, USA) sequencing to assemble contigs of the identified Y-linked BAC sequences for subsequent analysis of content and genomic structure.

### 2.2. Identification and Confirmation of Y-Linked Markers

We used two different approaches to identify Y-linked markers. First, we selected 24 genes that were expressed in at least two of five males and in none of five females in at least two sample types. Sample types were either roots or shoots, sampled either before or during flowering [32]. Of these 24 genes, 12 showed male-specific PCR amplification, based on an assay of ten males and ten females sampled across the species' range (Figure 1). Additionally, we identified six more Y-linked markers on the basis of sequences that were obtained by genome capture from males, but not females, sampled from across the species range (details in [31]). These additional six sequences were also confirmed by PCR, as explained above. One of the six sequences identified by this second approach was identical to one of the 12 sequences identified on the basis of gene expression data. In total, therefore, we identified  $12 + 6 - 1 = 17$  sex-linked markers, which likely represent Y-specific transcripts, or parts thereof (Supplementary Table S2). All sequences were amplified with the same PCR program, which used a 55 °C annealing temperature, 1.5 min amplification time and 30 cycles. All sex-linked markers were Sanger-sequenced to confirm that PCR products were indeed the expected gene sequences (see Supplementary File S1).



**Figure 1.** Graphical summary of diploid *Mercurialis annua* populations used, and the techniques we employed. Sampled populations are indicated in green, whereas polymerase chain reaction (PCR) products are indicated in blue. The PCR products marked with an asterisk represent the same sequence. The numbers next to the arrows indicate successful transcript attribution. The RNA sequencing (RNAseq) data come from [32], and the genome capture data from [31].

### 2.3. Bacterial Artificial Chromosome Library Construction, Sequencing and Assembly

*Mercurialis annua* leaves from two males were collected in November 2013 from wild plants growing on the campus of the University of Lausanne (Switzerland). Following storage of leaf material at  $-80^{\circ}\text{C}$ , high molecular-weight DNA was obtained from these samples using nucleus extractions at the Centre National de Ressources Génomiques Végétales (CNRGV), Toulouse, France. The DNA was fragmented and ligated into BAC vectors (pIndigoBAC-5) before transformation. The resulting BAC library was screened for the presence of the 17 Y-specific PCR markers described above, and we selected 18 recombinant colonies that amplified targeted markers. DNA extracted from these colonies was sequenced at the Centre of Integrative Genomics, University of Lausanne, using Pacific Biosciences (PacBio) technology.

Specifically, the BAC DNA was sheared in a Covaris g-TUBE (Covaris, Woburn, MA, USA) to obtain fragments with a mean length of 6 Kb. After shearing, the DNA size distribution was checked on a fragment analyzer (Advanced Analytical Technologies, Ames, IA, USA). About 300 ng were obtained in 150  $\mu\text{L}$  at 12,000 rpm or  $13,684 \times g$  on a Heraeus Biofuge Fresco (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and was then concentrated by SpeedVac (Thermo Fisher Scientific) to 4  $\mu\text{L}$ . Barcoded adapters were added to each BAC during ligation, and BACs were pooled for sequencing. Multiplexing was performed using the SMRTbell Barcoded Adapter Prep Kit #100-465-800 (Pacific Biosciences). The sheared DNA (1.3  $\mu\text{g}$ ) was used to prepare each SMRTbell library with the PacBio SMRTbell Template Prep Kit 1 (Pacific Biosciences) according to the manufacturer's

recommendations. BAC DNA was also sheared to 30 Kb and 50 Kb, to obtain two follow-up libraries (pool1 and pool2), without multiplexing.

Each library was sequenced on one SMRT cell with P6/C4 chemistry and MagBeads on a PacBio RSII instrument (Pacific Biosciences) at movie lengths of 240 min for the multiplexed library and 360 min for the follow-up libraries. Assembly was performed using the PacBio module ‘RS\_HGAP\_Assembly.2’ in SMRTpipe version v2.3.0 (Pacific Biosciences). The multiplexed library assembly did not result in many circularized individual BACs, possibly because of the limited fragment length required for multiplexing. We therefore mixed equal quantities of DNA from the BACs, split into two independent pools (pool1, pool2), and we sequenced the corresponding long reads. We used the partial assembly of the multiplexed BAC library to identify reads that only mapped to BACs of known origin (from demultiplexing), and we assembled those longer reads independently. After successfully circularizing the longer read contigs, we removed the vector sequence and linearized the BAC contigs. We used Geneious v9 [33] to assemble these BAC contigs into scaffolds. We used the scaffolds, which contained an overlapping BAC sequence, for ORF counting, but the BAC contigs were used for all other analyses, because they allowed detection of subtle differences in their sequence. Table 1 shows a summary of the final assemblies used in downstream analysis, and the full annotated assembly is provided in Supplementary File S3.

**Table 1.** Bacterial artificial chromosome (BAC) contig information. Overlapping BACs are displayed in the same row, and their approximate cumulative single copy length is shown. Inference of potential candidate sex-determining genes is based on the description of the first blastp hit. The predicted gene, transcript ID in expression data and NCBI sequence ID precede the gene description.

BAC Contig	Length (Kb)	Male-Specific PCR Hit	ORFs	Additional Predicted Genes	Potential Candidate Genes for Sex Determination
1, 2, 9	170	g16339	3	3	
3	110	g17561	1	3	
4, 21	143	g22704	4	0	
5, 20	200	g9930, g9932	3	0	
6	103	g31948	3	1	g18/g22596: <a href="#">XP_021600590.1</a> Agamous-like MADS-box protein AGL66
8	146	NA	13	3	g21/g20779: <a href="#">XP_002519813.1</a> Auxin response factor
10	127	g2445	1	0	
11	209	g5453	2	0	
12	123	gm20440	2	2	g41/g29576: <a href="#">XP_002267312.1</a> Light-dependent short hypocotyls 6
13	99	g3639	2	0	g44/g27235: <a href="#">XP_021670876.1</a>
7, 14, 23, 24	215	g15326, g15327	5	0	Two-component response regulator-like PRR73

PCR: Polymerase chain reaction; ORF: open reading frame; NA: not available.

#### 2.4. Transcript and Transposable Element Annotation

RepeatMasker [34] was run on the genomic contigs and the BAC contigs, using the custom *M. annua* repeat library [30]. This allowed us to compare the repeat content of the BACs to that of the full genome, and to produce a masked BAC assembly. BAC 8, whose ORFs were not sex-linked, was treated separately. The *M. annua* transcriptome ORFs were aligned to the masked BAC assembly with BLAT, using option minIdentity 92 [35]. BLAT allows large gaps that are compatible with introns. The resulting gff file was used to identify the number of sex-linked ORFs that localize on the BACs. The gff file was also used as a hints file, along with the parameter ‘species = arabidopsis’ by the gene predictor Augustus v3.2.3 [36], which was also run on the masked BAC contigs, to identify sex-linked genes that were not expressed in the samples used for genetic map construction. We identified the best reciprocal blastn hit [37] (e-value cut-off of  $10^{-6}$ ) per predicted transcript to obtain information on gene expression from the RNAseq experiment that was used to identify genes with male-specific expression [32], and another that used three pooled samples per sex, with ten individuals in each pool,



to investigate gene expression in three developmental times before, and one after, flowering in apical meristems [32].

### 3. Results

#### 3.1. Identification of Male-Specific Genes Based on Expression and Genome Capture Data

Twelve of the 24 genes that had male-specific expression were consistently amplified in males only (and are thus putatively Y-linked), whereas the remaining 12 genes amplified in both sexes (and are thus X-linked or autosomal). We also identified six genomic regions present on the Y-chromosome using genome capture data [31], by identifying sequences that were entirely missing from all females, but present in all males; one of these coincided with a sequence found on the basis of gene expression data, so that, in total, we obtained 17 male-specific PCR products. Three groups of these PCR products, containing seven transcripts, were already expected to be physically closely linked, based on their localization on the same genomic contig (g9930/g9932 and g15325/g15326/g15327, g17561/g17562). We thus conservatively estimate that there are 13 independent parts of the Y chromosome that can be sampled with these PCR products. Details for male-specific PCR amplification primers are summarized in Supplementary Table S2. Sequences of the products are provided in Supplementary File S1.

#### 3.2. Bacterial Artificial Chromosome Assembly

Our screening approach allowed us to identify 17 BACs containing 11 independent Y-linked DNA sequences, which we confirmed via male-specific PCR. A further BAC (BAC 8) turned out to be a false positive, i.e., it did not contain Y-linked sequence. The BACs were aligned to each other and grouped into 11 non-overlapping sets. Four of these groups contained the same male-specific PCR product, indicating either sequence duplication or within-population variation (recall that the sequences were obtained from two different males from Switzerland). Finally, 10/11 independently localizing male-specific PCR products were found in these 11 non-overlapping BAC groups (Table 1). Overall, these 11 BAC groups cover a genomic region of 1.5 Mb, corresponding to about 0.47% of the haploid *M. annua* genome.

#### 3.3. Functional Annotation of Genes Located on the Bacterial Artificial Chromosomes

By mapping the *M. annua* ORFs to the BAC scaffolds, we identified 24 broad genomic regions that each contained one or more complete ORFs (i.e., ORFs that mapped over their full length with identity >90%). Most of these ORFs can be considered putatively functional genes in the non-recombining region of the Y chromosome, both because they appear to be full-length genes, as well as because all but three of them had a best reciprocal blastn hit that was expressed in one or both of the expression experiments (the three unexpressed ORFs are marked in grey in Figure 2; expression data are available in Supplementary File S3). The exact gene number matching the broad 24 BAC regions is difficult to estimate because some ORFs overlap and may represent alternatively spliced variants of the same gene. Nevertheless, we identified at most 51 ORFs that mapped across their full length to the BACs, 14 of which blasted against sequences in the NCBI's non-redundant (nr) nucleotide database. In addition, we found 87 ORFs that mapped over part of their sequence length to the BACs, and some of these may represent truncated genes on the Y-chromosome; these sequences were located in 53 broad BAC regions. Finally, we found one complete and one truncated copy of the same gene (a sulfate transporter) next to each other on BAC 3, suggesting that it might be the product of a localized, incomplete gene duplication. BAC localization of ORFs and gene annotation information is provided in Supplementary File S4.

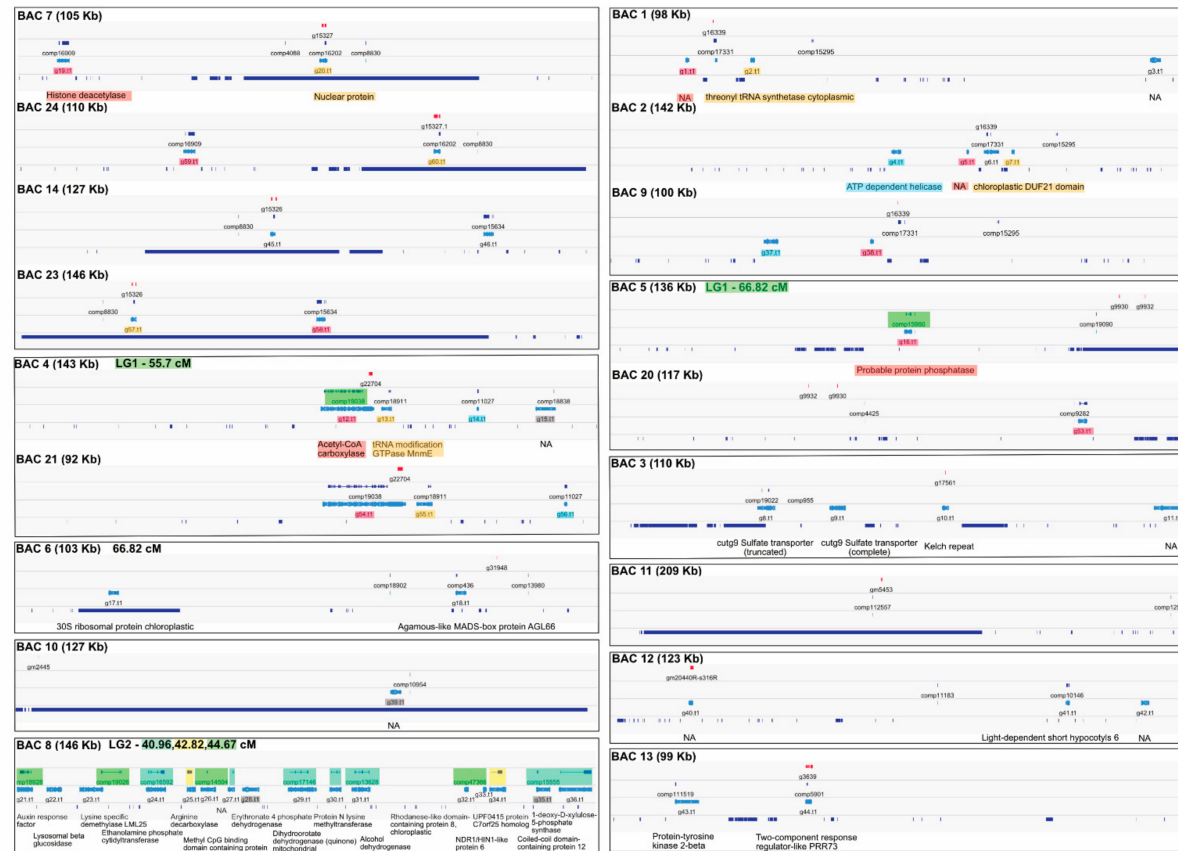
Gene prediction using Augustus v3.2.3 identified a further 12 putative genes on the BAC sequence (Table 1) that did not overlap with previously identified ORFs. These are likely candidates for further male-specific PCR (full sequences are provided in Supplementary File S5). Some of the ORFs mapping to the sequenced BACs might be involved in either sex determination or male-beneficial effects and

are good candidates for future study (see Table 1). They include a transcript in BAC 6 similar to the agamous-like MADS-box protein AGL66, which is required for pollen maturation and pollen-tube growth in *Arabidopsis* [38]. Another transcript, from BAC 12, is similar to light-dependent short hypocotyls 6, which is part of a family involved in response to light and organogenesis [39]. Finally, BAC 13 contained a transcript matching two-component response regulator-like PRR73, which controls photoperiodic flowering response [40]. Identification of such candidate genes is of course only the first step towards establishing involvement in sex determination and may be erroneous. For instance, BAC 8 revealed a strong candidate sex determiner (similarity to auxin response factor), but BAC 8 is probably not on the SDR, because its assembled sequence did not contain the male-specific PCR product used to identify the BAC, and most of its ORFs in fact mapped to an autosome (LG2; see Figure 2). Revealingly, BAC 8 also is very different from the other BACs in terms of both ORF density and repeat content (Figure 2).

Four groups of BACs contained the same ORFs (Figure 2). For the group containing BACs 1, 2 and 9, we found that three similar, but different, genomic regions had been sequenced. BAC 2 contains a predicted gene not found in the other BACs (possibly a chloroplastic insertion, Figure 2). However, BAC 1 and BAC 9 also differ, because BAC 9 is missing a predicted gene present in BAC 1 and BAC 2. As only two males were sequenced, each with a single Y chromosome, the results can be interpreted as a duplication of the whole BAC sequence, or assembly error. The finding of putatively duplicated sequences of the Y is consistent with the expected accumulation of male-beneficial (and possibly sexually antagonistic) variants, or points to ongoing degeneration through repeat proliferation. For example, a duplicated sequence linked to the Y chromosome and associated with the origin of males has been found in date palm [41]. The remaining three groups of BACs in *M. annua* appear to have sampled the same genomic region multiple times.

#### 3.4. Comparison of Transposable Element Density and Type Between the Bacterial Artificial Chromosomes and the Full Genome

Using RepeatMasker v 4.0.7 [34], we inferred that 76.9% of the BAC assembly comprised repetitive elements, substantially higher than the 47.9% repetitive content across the full genome. This was true for all categories of repetitive sequence, except for simple repeats. Long terminal repeats (LTRs) showed the highest enrichment on the BACs compared to the genomic contigs (25.33% vs. 8.45%; Supplementary Table S3). BAC 8 had a lower repeat content than that across the rest of the genome, increasing our confidence that it is indeed not sex-linked.



**Figure 2.** Graphical summary of the assembled BACs. Each assembly is annotated with the male-specific PCR product location (red), ORF *BLAT* hits (top dark blue), Augustus gene predictions on the repeat masked assembly (light blue), and *BLAT* hits from the *M. annua* repeat library (bottom dark blue). Each group of non-overlapping BACs is surrounded by a black box. Identical predicted gene models are indicated by overlaid colored boxes, for visual alignment of BACs containing the same genes. Green overlaid boxes indicate mapped ORFs that also map to the BACs, and their female recombination map position. Only genes with gray overlaid boxes were not expressed at all in the RNAseq experiments. The names displayed under the gene predictions are from their first blastp hit to the non-redundant (nr)protein database. Note the BAC contigs are not illustrated at the same scale.



## 4. Discussion

### 4.1. Identification of Y-Linked Markers

Using a combination of RNAseq and genome-capture data, we have identified 17 new single or low-copy Y-linked markers in diploid dioecious *M. annua*, which we confirmed through male-specific PCR amplification. We note that 12 of 24 transcripts with male-specific expression could only be amplified in males and are thus Y-linked. However, the best reciprocal blastn hit of these transcripts was expressed in some female samples of the RNAseq experiment, typically for pooled individuals (Supplementary File S3). Possible explanations for expression in females of genes identified in our expression experiment as having male-specific expression include: (1) contamination of the female pools with male RNA; (2) similar transcripts from other parts of the genome producing the best reciprocal blastn hit; (3) sequence divergence between the X and Y copy of the transcripts, so that PCR primers only amplify the Y-linked copy; and/or (4) duplication and sequence divergence of some transcripts on the Y chromosome. We nevertheless infer that these PCR products represent Y-linked genes because those that could be mapped are located close to the extreme ends of the SDR, and none mapped to a different genomic region. Our results confirm that the X and Y chromosomes of *M. annua* are differentiated at the sequence level, even though they appear homomorphic (see below). Moreover, the successful search for new sex-linked transcripts on the basis of sex-limited expression suggests that ignoring or filtering sex-limited genes in transcriptome analysis may overlook loci in the SDR Muyle [42].

The identification of 17 new potential single or low-copy sex-linked markers in *M. annua* represents a substantial advance on previous work on the species by Khadka et al. [43], who identified a single-sequence characterized amplified region (SCAR) marker that was male-specific. This SCAR marker was later found to correspond to a high-copy transposable element that is present in both sexes [30,43], and is thus of limited utility beyond the sexing of pre-reproductive individuals (see [32]). We used several of the new sex-linked markers to probe a newly constructed BAC library for sex-linked genomic regions, which were the main focus of the present study.

### 4.2. Size of the Sex-Determining Region

We may estimate the size of the SDR of *M. annua* using two different approaches that suggest rather different values. First, six ORFs contained in the BACs described here, which we independently inferred to be non-recombining based on the PCR result, were found in the sex-linked region of *M. annua* previously mapped by Ridout et al. [30]. These transcripts span a region from 52 to 66.82 cM in the female recombination map, corresponding to 441 ORFs, a length of 14.5 Mb and a proportion of 4.86% of the genome. This estimate is somewhat smaller than the one based on the mapping families alone (568 ORFs, equivalent to 19 Mb; [30]).

Second, given that only 6 of the 441 (1.3%) of sex-linked transcripts from Ridout et al. [30] mapped to the combined, non-overlapping, 1.5 Mb of the sequenced BACs, we might infer the SDR to be  $1.5 \text{ Mb} \times 441/6$ , or about 110 Mb (assuming that the sex-linked transcripts are distributed similarly on the rest of the SDR not sampled by our BACs). This second estimate corresponds to about 34% of the haploid genome of *M. annua* and is evidently much too large. It would seem likely, therefore, that our BACs substantially under-represent the average gene density across the SDR. Alternatively, this view might suggest an over-estimate of the number of sex-linked genes from [30]. We thus suggest that the most reliable estimate is still based on the number of transcripts not recombining in males, i.e., between 14.5 Mb (the length spanned by the sex-linked BACs) and 19 Mb (inferred from recombination in small crossing families).

Plant species with homomorphic sex chromosomes so far investigated have been found to have a SDR < 1% of the total length of the Y chromosome, e.g., *Vitis vinifera* [44,45], several *Populus* species [46,47], and *Fragaria chiloensis* [48]. *Mercurialis annua* thus appears to have a particularly large SDR for a species with homomorphic sex chromosomes. Ridout et al. [30] concluded that the sex

chromosomes of *M. annua* have been evolving independently of the rest of the genome for some time, though still without substantial transcript degeneration. This view of only mild Y-chromosome degeneration is consistent with the fact that YY males, which lack an X chromosome, are completely viable in *M. annua* [49].

#### 4.3. Content of the Sex-Determining Region

Although none of the de novo predicted genes on the BACs were obvious strong candidates for sex determination, three ORFs located on the BACs might function in sex determination or in the promotion of male function: a circadian gene involved in flowering (g44/g27235); a light response gene involved in organ and boundary differentiation (g41/g29576) [40]; and a transcription factor associated with pollen maturation (g18/g22596) [38]. These are interesting genes for future investigation, e.g., in surveys of population variation. Expression of these genes in two RNAseq datasets is presented in Supplementary Figure S1. None is significantly sex-biased. The pollen-maturation gene has higher expression in females early in development, but in males later in development. Similarly, the gene g27235, involved in flowering, was consistently more expressed in males than females in apical meristems, but showed an increased expression in mature leaves of males compared to females during flowering. A similar pattern was observed for the auxin response gene g21/g20779, which mapped to BAC 8 (i.e., it was not sex-linked), pointing to possible trans-regulation of the auxin response gene by the (unknown) Y-linked sex determiner.

Analysis of the content of the BAC contigs in *M. annua* has revealed typical features of non-recombining sex chromosomes, congruent with expectations of partial Y-chromosome degeneration. For instance, the sex-linked BACs contained mainly repetitive elements, which are expected to accumulate in non-recombining regions [50]. Comparison with the genomic scaffolds [30] revealed that the BACs are enriched in transposable elements (TEs), a finding similar to that reported for the *Carica papaya* Y<sup>h</sup> chromosome, where more than 80% of the non-recombining region on the Y chromosome comprises repetitive elements compared with only around 60% for the X chromosome [51]. We also found a clear case of partial gene duplication physically close to its complete gene copy, as well as multiple complete ORFs that did not completely map to the BACs. Although we have not been able to determine how many of these ORFs are bona fide disrupted genes (because no X-only contigs are yet available), their high number is consistent with early stages of Y degeneration in *M. annua*.

Finally, the overall gene density in the sex-linked BACs of *M. annua* appears to be substantially lower than that observed in the rest of the genome [30], consistent with the much higher gene density and a much lower repeat composition found for BAC8, which turned out not to be sex-linked, compared with the sex-linked BACs or the sex-linked genomic contigs [30].

## 5. Conclusions

Dioecious plants offer tremendous scope for examining the evolution of sex chromosomes because separate sexes have evolved independently, often relatively recently. A particularly noteworthy feature of dioecious plants is the degree to which they vary in the relative sizes of their X and Y (or Z and W) chromosomes, with species that have strongly heteromorphic sex chromosomes often closely related to species whose sex chromosomes are homomorphic [6,7]. Species with homomorphic sex chromosomes might simply be young, or might be subject to processes that maintain relative uniformity between homologues and a small SDR, such as frequent turnover [52] or occasional recombination [53]. Our study of the Y chromosome of *M. annua* illustrates that the SDR in homomorphic sex chromosomes may also be relatively large. Indeed, our estimate of the SDR of *M. annua* is the largest for any plant species with homomorphic sex chromosomes studied so far, whether viewed in absolute terms, or relative to the size of the sex chromosomes or the rest of the genome. The BAC sequences analyzed here point to an SDR with low gene density and enriched for repeats, with often incomplete mapping of complete ORFs (see also [30]). In this sense, *M. annua* may represent a species at a

particularly interesting intermediate stage along the path towards Y-chromosome degeneration and sex-chromosome heteromorphism.

**Supplementary Materials:** The following materials are available online at <http://www.mdpi.com/2073-4425/9/6/277/s1>. Figure S1: Plots of transcript per million data of genes localizing to the BACs and potentially involved in sex determination, in the two RNAseq experiments. Table S1: Summary of information on size and divergence in systems with homomorphic and heteromorphic sex chromosomes. References [10,17–19,22,45–48,50,51,54–91] are cited in Supplementary Table S1. Table S2: Information on PCR primers. Table S3: Summary output from RepeatMasker using the *M. annua* repeat library. File S1: Sanger sequence of PCR products. File S2: BAC assembly and gene prediction gff. File S3: Expression data of the BAC ORFs. File S4: BLAST information on ORF hits to the BAC assembly. File S5: Aminoacid sequence of predicted genes on the BACs.

**Author Contributions:** J.R.P., G.C., P.V. conceived the project. G.C. and G.B. prepared the BAC library. E.B. assembled the BAC contigs. G.C. identified male-specific transcripts in RNAseq. P.V. designed and run the PCR verification along with D.S.B. S.C.G.-M. designed and provided the genome capture data, which were analysed by C.R. J.R.P., G.C. and P.V. wrote the manuscript with input from all authors. All authors have approved the final article.

**Funding:** This work was supported by a Sinergia and personal grant awarded to John R. Pannell from the Swiss National Science Foundation, and a Marie Skłodowska Curie Fellowship grant awarded to Santiago C. González-Martínez and John R. Pannell. The purchase of the Pacific Biosciences RSII instrument at the University of Lausanne was financed in part by the Loterie Romande through the Fondation pour la Recherche en Médecine Génétique.

**Acknowledgments:** We are grateful to Thomas Chassaing (Angers, France), Yusuf Kurt (Harran, Turkey) and Fernando del Caño (Madrid, Spain) for providing plant material, to Melanie Dupasquier and Lausanne Genomic Technologies Facility, Lausanne, Switzerland, for PacBio library production and sequencing, to Deborah Charlesworth (Edinburgh, UK) for helpful comments on an earlier version of the manuscript, and to Mathias Scharmann and other members of the Pannell lab (Lausanne, Switzerland) for valuable discussion.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Renner, S.S. The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *Am. J. Bot.* **2014**, *101*, 1588–1596. [[CrossRef](#)] [[PubMed](#)]
- Renner, S.S.; Ricklefs, R.E. Dioecy and its correlates in the flowering plants. *Am. J. Bot.* **1995**, *82*, 596–606. [[CrossRef](#)]
- Beukeboom, L.W.; Perrin, N. *The Evolution of Sex Determination*; Oxford University Press: Oxford, UK, 2014.
- Policansky, D. Sex change in plants and animals. *Annu. Rev. Ecol. Syst.* **1982**, *13*, 471–495. [[CrossRef](#)]
- Zimmerman, J.K. Ecological correlates of labile sex expression in the orchid *Catasetum viridiflavum*. *Ecology* **1991**, *72*, 597–608. [[CrossRef](#)]
- Ming, R.; Bendahmane, A.; Renner, S.S. Sex chromosomes in land plants. *Annu. Rev. Plant Biol.* **2011**, *62*, 485–514. [[CrossRef](#)] [[PubMed](#)]
- Charlesworth, D. Plant sex chromosomes. *Annu. Rev. Plant Biol.* **2016**, *67*, 397–420. [[CrossRef](#)] [[PubMed](#)]
- Sousa, A.; Fuchs, J.; Renner, S.S. Cytogenetic comparison of heteromorphic and homomorphic sex chromosomes in *Coccinia* (Cucurbitaceae) points to sex chromosome turnover. *Chromosome Res.* **2017**, *25*, 191–200. [[CrossRef](#)] [[PubMed](#)]
- Charlesworth, D. Plant contributions to our understanding of sex chromosome evolution. *New Phytol.* **2015**, *208*, 52–65. [[CrossRef](#)] [[PubMed](#)]
- Sousa, A.; Fuchs, J.; Renner, S.S. Molecular cytogenetics (FISH, GISH) of *Coccinia grandis*: A ca. 3 myr-old species of Cucurbitaceae with the largest Y/autosome divergence in flowering plants. *Cytogenet. Genome Res.* **2013**, *139*, 107–118. [[CrossRef](#)] [[PubMed](#)]
- Cherif, E.; Zehdi-Azouzi, S.; Crabos, A.; Castillo, K.; Chabrilange, N.; Pintaud, J.C.; Salhi-Hannachi, A.; Glemin, S.; Aberlenc-Bertossi, F. Evolution of sex chromosomes prior to speciation in the dioecious *Phoenix* species. *J. Evol. Biol.* **2016**, *29*, 1513–1522. [[CrossRef](#)] [[PubMed](#)]
- AbdAlla, M.M.; Abd El-Kawy, M. Cytological studies for date palm (*Phoenix dactylifera* L.) tissue culture derived plants. *Rep. Opin.* **2010**, *2*, 17–21.
- Hagelamin, O.; Alzahrani, M.M.A. Characterization and analysis of date palm (*Phoenix dactylifera* L.) karyotypes. *Asia Life Sci.* **2015**, *24*, 809–822.

14. Bergero, R.; Charlesworth, D. The evolution of restricted recombination in sex chromosomes. *Trends Ecol. Evol.* **2009**, *24*, 94–102. [[CrossRef](#)] [[PubMed](#)]
15. Rice, W.R. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution* **1987**, *41*, 911–914. [[CrossRef](#)] [[PubMed](#)]
16. Hobza, R.; Cegan, R.; Jesionek, W.; Kejnovsky, E.; Vyskot, B.; Kubat, Z. Impact of repetitive elements on the Y chromosome formation in plants. *Genes* **2017**, *8*, 12. [[CrossRef](#)] [[PubMed](#)]
17. Qiu, S.; Bergero, R.; Guirao-Rico, S.; Campos, J.L.; Cezard, T.; Gharbi, K.; Charlesworth, D. RAD mapping reveals an evolving, polymorphic and fuzzy boundary of a plant pseudoautosomal region. *Mol. Ecol.* **2016**, *25*, 414–430. [[CrossRef](#)] [[PubMed](#)]
18. Fraser, L.G.; Tsang, G.K.; Datson, P.M.; De Silva, H.N.; Harvey, C.F.; Gill, G.P.; Crowhurst, R.N.; McNeilage, M.A. A gene-rich linkage map in the dioecious species *Actinidia chinensis* (kiwifruit) reveals putative X/Y sex-determining chromosomes. *BMC Genom.* **2009**, *10*, 15. [[CrossRef](#)] [[PubMed](#)]
19. Mathew, L.S.; Spannagl, M.; Al-Malki, A.; George, B.; Torres, M.F.; Al-Dous, E.K.; Al-Azwani, E.K.; Hussein, E.; Mathew, S.; Mayer, K.F.X.; et al. A first genetic map of date palm (*Phoenix dactylifera*) reveals long-range genome structure conservation in the palms. *BMC Genom.* **2014**, *15*, 285. [[CrossRef](#)] [[PubMed](#)]
20. Hobza, R.; Lengerova, M.; Svoboda, J.; Kubekova, H.; Kejnovsky, E.; Vyskot, B. An accumulation of tandem DNA repeats on the Y chromosome in *Silene latifolia* during early stages of sex chromosome evolution. *Chromosoma* **2006**, *115*, 376–382. [[CrossRef](#)] [[PubMed](#)]
21. Na, J.K.; Wang, J.P.; Murray, J.E.; Gschwend, A.R.; Zhang, W.L.; Yu, Q.Y.; Navajas-Perez, R.; Feltus, F.A.; Chen, C.X.; Kubat, Z.; et al. Construction of physical maps for the sex-specific regions of papaya sex chromosomes. *BMC Genom.* **2012**, *13*, 176. [[CrossRef](#)] [[PubMed](#)]
22. Telgmann-Rauber, A.; Jamsari, A.; Kinney, M.S.; Pires, J.C.; Jung, C. Genetic and physical maps around the sex-determining M-locus of the dioecious plant *Asparagus*. *Mol. Genet. Genom.* **2007**, *278*, 221–234. [[CrossRef](#)] [[PubMed](#)]
23. Pannell, J.R.; Dorken, M.E.; Pujol, B.; Berjano, R. Gender variation and transitions between sexual systems in *Mercurialis annua* (Euphorbiaceae). *Int. J. Plant Sci.* **2008**, *169*, 129–139. [[CrossRef](#)]
24. Russell, J.R.W.; Pannell, J.R. Sex determination in dioecious *Mercurialis annua* and its close diploid and polyploid relatives. *Heredity* **2015**, *114*, 262–271. [[CrossRef](#)] [[PubMed](#)]
25. Durand, B. Le complexe *Mercurialis annua* L. s.l.: Une étude biosystématique. *Ann. Sci. Nat. Bot. Paris* **1963**, *12*, 579–736.
26. Durand, R.; Durand, B. Dioecy, monoecy, polyploidy and speciation in the annual Mercuries. *Bull. Soc. Bot. France Lett. Bot.* **1992**, *139*, 377–399.
27. Obbard, D.J.; Harris, S.A.; Pannell, J. Sexual systems and population genetic structure in an annual plant: Testing the metapopulation model. *Am. Nat.* **2006**, *167*, 354–366. [[CrossRef](#)] [[PubMed](#)]
28. Eppley, S.M.; Pannell, J.R. Sexual systems and measures of occupancy and abundance in an annual plant: Testing the metapopulation model. *Am. Nat.* **2007**, *169*, 20–28. [[CrossRef](#)] [[PubMed](#)]
29. Pannell, J.R.; Eppley, S.M.; Dorken, M.E.; Berjano, R. Regional variation in sex ratios and sex allocation in androdioecious *Mercurialis annua*. *J. Evol. Biol.* **2014**, *27*, 1467–1477. [[CrossRef](#)] [[PubMed](#)]
30. Ridout, K.; Veltsos, P.; Muyle, A.; Emery, O.; Rastas, P.; Marais, G.; Filatov, D.; Pannell, J.R. Hallmarks of early sex-chromosome evolution in the dioecious plant *Mercurialis annua* revealed by de novo genome assembly, genetic mapping and transcriptome analysis. *bioRxiv* **2017**. [[CrossRef](#)]
31. Gonzalez-Martinez, S.C.; Ridout, K.; Pannell, J.R. Range expansion compromises adaptive evolution in an outcrossing plant. *Curr. Biol.* **2017**, *27*, 2544–2551. [[CrossRef](#)] [[PubMed](#)]
32. Cossard, G.; Pannell, J.R. Sexual dimorphism and rapid turnover in gene expression in pre-reproductive seedlings of a dioecious herb. **2018**, submitted.
33. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)] [[PubMed](#)]
34. Smit, A.F.A.; Hubley, R.; Green, P. RepeatMasker Open-4.0. Available online: <http://www.repeatmasker.org/>.
35. Kent, W.J. BLAT—The BLAST-like alignment tool. *Genome Res.* **2002**, *12*, 656–664. [[CrossRef](#)] [[PubMed](#)]



36. Stanke, M.; Schoffmann, O.; Morgenstern, B.; Waack, S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinform.* **2006**, *7*, 62. [[CrossRef](#)] [[PubMed](#)]
37. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
38. Adamczyk, B.J.; Fernandez, D.E. MIKC\* MADS Domain Heterodimers Are Required for Pollen Maturation and Tube Growth in *Arabidopsis*. *Plant Physiol.* **2009**, *149*, 1713–1723. [[CrossRef](#)] [[PubMed](#)]
39. Cho, E.N.; Zambryski, P.C. ORGAN BOUNDARY1 defines a gene expressed at the junction between the shoot apical meristem and lateral organs. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2154–2159. [[CrossRef](#)] [[PubMed](#)]
40. Murakami, M.; Matsushika, A.; Ashikari, M.; Yamashino, T.; Mizuno, T. Circadian-associated rice pseudo response regulators (OsPRRs): Insight into the control of flowering time. *Biosci. Biotechnol. Biochem.* **2005**, *69*, 410–414. [[CrossRef](#)] [[PubMed](#)]
41. Cherif, E.; Zehdi, S.; Castillo, K.; Chabrilange, N.; Abdoukader, S.; Pintaud, J.C.; Santoni, S.; Salhi-Hannachi, A.; Glémin, S.; Aberlenc-Bertossi, F. Male-specific DNA markers provide genetic evidence of an XY chromosome system, a recombination arrest and allow the tracing of paternal lineages in date palm. *New Phytol.* **2013**, *197*, 409–415. [[CrossRef](#)] [[PubMed](#)]
42. Muyle, A.; Zemp, N.; Deschamps, C.; Mousset, S.; Widmer, A.; Marais, G.A.B. Rapid de novo evolution of X chromosome dosage compensation in *Silene latifolia*, a plant with young sex chromosomes. *PLoS Biol.* **2012**, *10*, e1001308. [[CrossRef](#)] [[PubMed](#)]
43. Khadka, D.K.; Nejidat, A.; Tal, M.; Golan-Goldhirsh, A. DNA markers for sex: Molecular evidence for gender dimorphism in dioecious *Mercurialis annua* L. *Mol. Breed.* **2002**, *9*, 251–257. [[CrossRef](#)]
44. Fechter, I.; Hausmann, L.; Daum, M.; Sorensen, T.R.; Viehover, P.; Weisshaar, B.; Topfer, R. Candidate genes within a 143 kb region of the flower sex locus in *Vitis*. *Mol. Genet. Genom.* **2012**, *287*, 247–259. [[CrossRef](#)] [[PubMed](#)]
45. Picq, S.; Santoni, S.; Lacombe, T.; Latreille, M.; Weber, A.; Ardisson, M.; Ivorra, S.; Maghradze, D.; Arroyo-Garcia, R.; Chatelet, P.; et al. A small XY chromosomal region explains sex determination in wild dioecious *V. vinifera* and the reversal to hermaphroditism in domesticated grapevines. *BMC Plant Biol.* **2014**, *14*, 17. [[CrossRef](#)] [[PubMed](#)]
46. Paolucci, I.; Gaudet, M.; Jorge, V.; Beritognolo, I.; Terzoli, S.; Kuzminsky, E.; Muleo, R.; Mugnozza, G.S.; Sabatti, M. Genetic linkage maps of *Populus alba* L. and comparative mapping analysis of sex determination across *Populus* species. *Tree Genet. Genomes* **2010**, *6*, 863–875. [[CrossRef](#)]
47. Geraldès, A.; Hefer, C.A.; Capron, A.; Kolosova, N.; Martínez-Núñez, F.; Soolanayakanahally, R.Y.; Stanton, B.; Guy, R.D.; Mansfield, S.D.; Douglas, C.J.; et al. Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). *Mol. Ecol.* **2015**, *24*, 3243–3256. [[CrossRef](#)] [[PubMed](#)]
48. Tennesen, J.A.; Govindarajulu, R.; Liston, A.; Ashman, T.L. Homomorphic ZW chromosomes in a wild strawberry show distinctive recombination heterogeneity but a small sex-determining region. *New Phytol.* **2016**, *211*, 1412–1423. [[CrossRef](#)] [[PubMed](#)]
49. Kuhn, E. Selbstbestäubungen subdiöischer Blütenpflanzen, ein neuer Beweis für die genetische Theorie der Geschlechtsbestimmung. *Planta* **1939**, *30*, 457–470. [[CrossRef](#)]
50. Bachtrog, D. Y-chromosome evolution: Emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.* **2013**, *14*, 113–124. [[CrossRef](#)] [[PubMed](#)]
51. Wang, J.P.; Na, J.K.; Yu, Q.Y.; Gschwend, A.R.; Han, J.; Zeng, F.C.; Aryal, R.; VanBuren, R.; Murray, J.E.; Zhang, W.L.; et al. Sequencing papaya X and Y-h chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13710–13715. [[CrossRef](#)] [[PubMed](#)]
52. Van Doorn, G.S.; Kirkpatrick, M. Turnover of sex chromosomes induced by sexual conflict. *Nature* **2007**, *449*, 909–912. [[CrossRef](#)] [[PubMed](#)]
53. Perrin, N. Sex reversal: A fountain of youth for sex chromosomes? *Evolution* **2009**, *63*, 3043–3049. [[CrossRef](#)] [[PubMed](#)]
54. Sakamoto, K.; Akiyama, Y.; Fukui, K.; Kamada, H.; Satoh, S. Characterization, genome sizes and morphology of sex chromosomes in hemp (*Cannabis sativa* L.). *Cytologia* **1998**, *63*, 459–464. [[CrossRef](#)]
55. Van Bakel, H.; Stout, J.M.; Cote, A.G.; Tallon, C.M.; Sharpe, A.G.; Hughes, T.R.; Page, J.E. The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* **2011**, *12*, 17. [[CrossRef](#)] [[PubMed](#)]



56. Grabowska-Joachimciak, A.; Sliwinska, E.; Pigula, M.; Skomra, U.; Joachimciak, A.J. Genome size in *Humulus lupulus* L. and *H. japonicus* Siebold & Zucc. (Cannabaceae). *Acta Soc. Bot. Pol.* **2006**, *75*, 207–214.
57. Divashuk, M.G.; Alexandrov, O.S.; Kroupin, P.Y.; Karlov, G.I. Molecular cytogenetic mapping of *Humulus lupulus* sex chromosomes. *Cytogenet. Genome Res.* **2011**, *134*, 213–219. [[CrossRef](#)] [[PubMed](#)]
58. Armstrong, S.J.; Filatov, D.A. A cytogenetic view of sex chromosome evolution in plants. *Cytogenet. Genome Res.* **2008**, *120*, 241–246. [[CrossRef](#)] [[PubMed](#)]
59. Błocka-Wandas, M.; Sliwinska, E.; Grabowska-Joachimciak, A.; Musial, K.; Joachimciak, A.J. Male gametophyte development and two different DNA classes of pollen grains in *Rumex acetosa* L., a plant with an XX/XY1Y2 sex chromosome system and a female-biased sex ratio. *Sex. Plant Reprod.* **2007**, *20*, 171–180. [[CrossRef](#)]
60. Bennett, M.D.; Leitch, I.J. Plant DNA C-values Database. Available online: <http://data.kew.org/cvalues/> (accessed on 16 April 2018).
61. Shibata, F.; Hizume, M.; Kuroki, Y. Differentiation and the polymorphic nature of the Y chromosomes revealed by repetitive sequences in the dioecious plant, *Rumex acetosa*. *Chromosome Res.* **2000**, *8*, 229–236. [[CrossRef](#)] [[PubMed](#)]
62. Hough, J.; Hollister, J.D.; Wang, W.; Barrett, S.C.H.; Wright, S.I. Genetic degeneration of old and young Y chromosomes in the flowering plant *Rumex hastatulus*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 7713–7718. [[CrossRef](#)] [[PubMed](#)]
63. Grabowska-Joachimciak, A.; Kula, A.; Ksiazczyk, T.; Chojnicka, J.; Sliwinska, E.; Joachimciak, A.J. Chromosome landmarks and autosome-sex chromosome translocations in *Rumex hastatulus*, a plant with XX/XY1Y2 sex chromosome system. *Chromosome Res.* **2015**, *23*, 187–197. [[CrossRef](#)] [[PubMed](#)]
64. Huang, S.X.; Ding, J.; Deng, D.J.; Tang, W.; Sun, H.H.; Liu, D.Y.; Zhang, L.; Niu, X.L.; Zhang, X.; Meng, M.; et al. Draft genome of the kiwifruit *Actinidia chinensis*. *Nat. Commun.* **2013**, *4*, 9. [[CrossRef](#)] [[PubMed](#)]
65. Zhang, Q.; Liu, C.Y.; Liu, Y.F.; VanBuren, R.; Yao, X.H.; Zhong, C.H.; Huang, H.W. High-density interspecific genetic maps of kiwifruit and the identification of sex-specific markers. *DNA Res.* **2015**, *22*, 367–375. [[CrossRef](#)] [[PubMed](#)]
66. Al-Dous, E.K.; George, B.; Al-Mahmoud, M.E.; Al-Jaber, M.Y.; Wang, H.; Salameh, Y.M.; Al-Azwani, E.K.; Chaluvadi, S.; Pontaroli, A.C.; DeBarry, J.; et al. De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nat. Biotechnol.* **2011**, *29*, 521–527. [[CrossRef](#)] [[PubMed](#)]
67. Al-Mssallem, I.S.; Hu, S.; Zhang, X.; Lin, Q.; Liu, W.; Tan, J.; Yu, X.; Liu, J.; Pan, L.; Zhang, T.; et al. Genome sequence of the date palm *Phoenix dactylifera* L. *Nat. Commun.* **2013**, *4*, 2274. [[CrossRef](#)] [[PubMed](#)]
68. Kafkas, S.; Khodaeiaminjan, M.; Guney, M.; Kafkas, E. Identification of sex-linked SNP markers using RAD sequencing suggests ZW/ZZ sex determination in *Pistacia vera* L. *BMC Genom.* **2015**, *16*, 98. [[CrossRef](#)] [[PubMed](#)]
69. Ming, R.; Hou, S.B.; Feng, Y.; Yu, Q.Y.; Dionne-Laporte, A.; Saw, J.H.; Senin, P.; Wang, W.; Ly, B.V.; Lewis, K.L.T.; et al. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* **2008**, *452*, 991. [[CrossRef](#)] [[PubMed](#)]
70. Spigler, R.B.; Lewers, K.S.; Main, D.S.; Ashman, T.L. Genetic mapping of sex determination in a wild strawberry, *Fragaria virginiana*, reveals earliest form of sex chromosome. *Heredity* **2008**, *101*, 507–517. [[CrossRef](#)] [[PubMed](#)]
71. Tamura, M.; Tao, R.; Yonemori, K.; Utsunomiya, N.; Sugiura, A. Ploidy level and genome size of several *Diospyros* species. *J. Jpn. Soc. Hortic. Sci.* **1998**, *67*, 306–312. [[CrossRef](#)]
72. Akagi, T.; Henry, I.M.; Tao, R.; Comai, L. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. *Science* **2014**, *346*, 646–650. [[CrossRef](#)] [[PubMed](#)]
73. Yin, T.; DiFazio, S.P.; Gunter, L.E.; Zhang, X.; Sewell, M.M.; Woolbright, S.A.; Allan, G.J.; Kelleher, C.T.; Douglas, C.J.; Wang, M.; et al. Genome structure and emerging evidence of an incipient sex chromosome in *Populus*. *Genome Res.* **2008**, *18*, 422–430. [[CrossRef](#)] [[PubMed](#)]
74. Pakull, B.; Groppe, K.; Meyer, M.; Markussen, T.; Fladung, M. Genetic linkage mapping in aspen (*Populus tremula* L. and *Populus tremuloides* Michx.). *Tree Genet. Genomes* **2009**, *5*, 505–515. [[CrossRef](#)]
75. Pucholt, P.; Ronnberg-Wastljung, A.C.; Berlin, S. Single locus sex determination and female heterogamety in the basket willow (*Salix viminalis* L.). *Heredity* **2015**, *114*, 575–583. [[CrossRef](#)] [[PubMed](#)]
76. Jaillon, O.; Aury, J.M.; Noel, B.; Policriti, A.; Clepet, C.; Casagrande, A.; Choisne, N.; Aubourg, S.; Vitulo, N.; Jubin, C.; et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **2007**, *449*, 463. [[PubMed](#)]

77. Obbard, D.J.; Harris, S.A.; Buggs, R.J.A.; Pannell, J.R. Hybridization, polyploidy, and the evolution of sexual systems in *Mercurialis* (Euphorbiaceae). *Evolution* **2006**, *60*, 1801–1815. [[CrossRef](#)] [[PubMed](#)]
78. Skaletsky, H.; Kuroda-Kawaguchi, T.; Minx, P.J.; Cordum, H.S.; Hillier, L.; Brown, L.G.; Repping, S.; Pyntikova, T.; Ali, J.; Bieri, T.; et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **2003**, *423*, 825. [[CrossRef](#)] [[PubMed](#)]
79. Hughes, J.F.; Skaletsky, H.; Pyntikova, T.; Graves, T.A.; van Daalen, S.K.M.; Minx, P.J.; Fulton, R.S.; McGrath, S.D.; Locke, D.P.; Friedman, C.; et al. Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. *Nature* **2010**, *463*, 536–539. [[CrossRef](#)] [[PubMed](#)]
80. Ellegren, H. Evolution of the avian sex chromosomes and their role in sex determination. *Trends Ecol. Evol.* **2000**, *15*, 188–192. [[CrossRef](#)]
81. Ayers, K.L.; Davidson, N.M.; Demiyah, D.; Roeszler, K.N.; Grutzner, F.; Sinclair, A.H.; Oshlack, A.; Smith, C.A. RNA sequencing reveals sexually dimorphic gene expression before gonadal differentiation in chicken and allows comprehensive annotation of the W-chromosome. *Genome Biol.* **2013**, *14*, 16. [[CrossRef](#)] [[PubMed](#)]
82. Ogawa, A.; Murata, K.; Mizuno, S. The location of Z- and W-linked marker genes and sequence on the homomorphic sex chromosomes of the ostrich and the emu. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4415–4418. [[CrossRef](#)] [[PubMed](#)]
83. Zhou, Q.; Zhang, J.L.; Bachtrog, D.; An, N.; Huang, Q.F.; Jarvis, E.D.; Gilbert, M.T.P.; Zhang, G.J. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* **2014**, *346*, 1246338. [[CrossRef](#)] [[PubMed](#)]
84. Jones, F.C.; Grabherr, M.G.; Chan, Y.F.; Russell, P.; Mauceli, E.; Johnson, J.; Swofford, R.; Pirun, M.; Zody, M.C.; White, S.; et al. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **2012**, *484*, 55–61. [[CrossRef](#)] [[PubMed](#)]
85. Kasahara, M.; Naruse, K.; Sasaki, S.; Nakatani, Y.; Qu, W.; Ahsan, B.; Yamada, T.; Nagayasu, Y.; Doi, K.; Kasai, Y.; et al. The medaka draft genome and insights into vertebrate genome evolution. *Nature* **2007**, *447*, 714–719. [[CrossRef](#)] [[PubMed](#)]
86. Adams, M.D.; Celniker, S.E.; Holt, R.A.; Evans, C.A.; Gocayne, J.D.; Amanatides, P.G.; Scherer, S.E.; Li, P.W.; Hoskins, R.A.; Galle, R.F.; et al. The genome sequence of *Drosophila melanogaster*. *Science* **2000**, *287*, 2185–2195. [[CrossRef](#)] [[PubMed](#)]
87. Zhou, Q.; Zhu, H.M.; Huang, Q.F.; Zhao, L.; Zhang, G.J.; Roy, S.W.; Vicoso, B.; Xuan, Z.L.; Ruan, J.; Zhang, Y.; et al. Deciphering neo-sex and B chromosome evolution by the draft genome of *Drosophila albomicans*. *BMC Genom.* **2012**, *13*, 12. [[CrossRef](#)] [[PubMed](#)]
88. Richards, S.; Liu, Y.; Bettencourt, B.R.; Hradecky, P.; Letovsky, S.; Nielsen, R.; Thornton, K.; Hubisz, M.J.; Chen, R.; Meisel, R.P.; et al. Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome Res.* **2005**, *15*, 1–18. [[CrossRef](#)] [[PubMed](#)]
89. Steinemann, M.; Steinemann, S. Enigma of Y chromosome degeneration: Neo-Y and Neo-X chromosomes of *Drosophila miranda* a model for sex chromosome evolution. *Genetica* **1998**, *102*, 409–420. [[CrossRef](#)] [[PubMed](#)]
90. Kaiser, V.B.; Charlesworth, B. Muller's ratchet and the degeneration of the *Drosophila miranda* neo-Y chromosome. *Genetics* **2010**, *185*, 339. [[CrossRef](#)] [[PubMed](#)]
91. Criscione, C.D.; Valentim, C.L.L.; Hirai, H.; LoVerde, P.T.; Anderson, T.J.C. Genomic linkage map of the human blood fluke *Schistosoma mansoni*. *Genome Biol.* **2009**, *10*, R71. [[CrossRef](#)] [[PubMed](#)]

