HAL
open science

Evidence for Dosage Compensation in Coccinia grandis, a Plant with a Highly Heteromorphic XY System

Cécile Fruchard, Hélène E Badouin, David Latrasse, Ravi S Devani, Aline Muyle, Bénédicte Rhoné, Susanne S Renner, Anjan K Banerjee, Abdelhafid Bendahmane, Gabriel Marais

To cite this version:

Cécile Fruchard, Hélène E Badouin, David Latrasse, Ravi S Devani, Aline Muyle, et al.. Evidence for Dosage Compensation in Coccinia grandis, a Plant with a Highly Heteromorphic XY System. Genes, 2020, 11 (7), pp.787. ⟨10.3390/genes11070787⟩. ⟨hal-03250668⟩

HAL Id: hal-03250668

https://hal.science/hal-03250668v1

Submitted on 4 Jun 2021

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.

Autorisation HAL

HAL Authorization

genes

MDPI

Article

Evidence for Dosage Compensation in *Coccinia grandis*, a Plant with a Highly Heteromorphic XY System

Cécile Fruchard [1], Hélène Badouin [1], David Latrasse [2], Ravi S. Devani [2,3], Aline Muyle [4], Bénédicte Rhoné [1,5], Susanne S. Renner [6], Anjan K. Banerjee [3], Abdelhafid Bendahmane [2] and Gabriel A. B. Marais [1,*]

1 Laboratoire de Biométrie et Biologie Evolutive (LBBE), UMR5558, Université Lyon 1, 69622 Villeurbanne, France; cecile.fruchard@gmail.com (C.F.); helene.badouin@univ-lyon1.fr (H.B.); benedicte.rhone@cirad.fr (B.R.)
2 Institute of Plant Sciences Paris Saclay (IPS2), University of Paris Saclay, 91405 Orsay, France; david.latrasse@ips2.universite-paris-saclay.fr (D.L.); ravi.devani@students.iiserpune.ac.in (R.S.D.); abdel.bendahmane@ips2.universite-paris-saclay.fr (A.B.)
3 Biology Division, Indian Institute of Science Education and Research (IISER), Pune 411008, Maharashtra, India; akb@iiserpune.ac.in
4 Department of Ecology and Evolutionary Biology, University of California Irvine, Irvine, CA 92697, USA; amuyle@uci.edu
5 Institut de Recherche pour le Développement (IRD), Université Montpellier, DIADE, F-34394 Montpellier, France
6 Systematic Botany and Mycology, University of Munich (LMU), Menzinger Str. 67, 80638 Munich, Germany; renner@lmu.de
* Correspondence: gabriel.marais@univ-lyon1.fr

Received: 26 May 2020; Accepted: 8 July 2020; Published: 13 July 2020

Abstract: About 15,000 angiosperms are dioecious, but the mechanisms of sex determination in plants remain poorly understood. In particular, how Y chromosomes evolve and degenerate, and whether dosage compensation evolves as a response, are matters of debate. Here, we focus on *Coccinia grandis*, a dioecious cucurbit with the highest level of X/Y heteromorphy recorded so far. We identified sex-linked genes using RNA sequences from a cross and a model-based method termed SEX-DETector. Parents and F1 individuals were genotyped, and the transmission patterns of SNPs were then analyzed. In the >1300 sex-linked genes studied, maximum X-Y divergence was 0.13–0.17, and substantial Y degeneration is implied by an average Y/X expression ratio of 0.63 and an inferred gene loss on the Y of ~40%. We also found reduced Y gene expression being compensated by elevated expression of corresponding genes on the X and an excess of sex-biased genes on the sex chromosomes. Molecular evolution of sex-linked genes in *C. grandis* is thus comparable to that in *Silene latifolia*, another dioecious plant with a strongly heteromorphic XY system, and cucurbits are the fourth plant family in which dosage compensation is described, suggesting it might be common in plants.

Keywords: dioecy; sex chromosomes; Y degeneration; sex-biased genes; cucurbits

1. Introduction

Some 5 or 6% of the angiosperms, depending on the assumed total species number, have male and female sporophytes, a sexual system termed dioecy [1]. Transitions from other sexual systems towards dioecy are estimated to have occurred between 871 and 5000 times independently [1]. Chromosomes and sex determination have been studied in few dioecious plants, and microscopically distinguishable

(heteromorphic) sex chromosomes have been reported in about 50 species only [2]. An important question is whether sex chromosomes evolve similarly in plants and animals. For example, the evolution towards heteromorphy might be common between both lineages, with an autosomal origin of the sex chromosomes, gradual recombination suppression between X and Y chromosomes, and genetic degeneration of the Y chromosome [2–5]. Animal Y chromosomes tend to shrink over time and can become tiny as in the old heteromorphic systems of mammals and *Drosophila* [3,4,6]. In plants, the size of the Y chromosomes might evolve in a non-linear way, with Y chromosomes of intermediate age being larger than their X counterparts, but very old Y chromosomes being smaller [2]. Systems with large Y chromosomes have not been found in animals (with the single exception of the large neo-Y of *Drosophila miranda*, [7]) and might be plant-specific; the reason for this is unknown. *Silene latifolia* has one such system and has been intensively studied (e.g., [8–18]). Studying more plant systems, in particular, those with larger-than-X Y chromosomes, is necessary to get a more precise picture of the evolution of heteromorphic sex chromosomes in plants.

An important aspect of the evolution of animal sex chromosomes is dosage compensation, which has been reported in old heteromorphic systems [19]. The genetic degeneration of the Y chromosome causes a global decrease of Y gene expression through gene loss and gene silencing [6]. Without compensation, this phenomenon would result in a dosage imbalance in males because many Y genes have reduced expression, or are not expressed at all, compared to females where both X copies are fully expressed [19]. Three such compensation mechanisms have been described: the mammalian system, the *Drosophila* system, and the *C. elegans* system [20]. The mechanism in the fruit fly appears to be straightforward, with a chromosome-wide doubling of X expression in males, re-establishing the proper dosage [20]. The mammal and *C. elegans* mechanisms are less straightforward, with an apparent doubling of expression on the X chromosome in both sexes and then a mechanism to correct expression in females/hermaphrodites (X-inactivation in mammals, downregulation of both Xs in *C. elegans*, [20]). In both lineages, dosage compensation seems to affect certain dosage-sensitive genes only [19–25]). In birds, dosage compensation also is local, affecting only a few dosage-sensitive genes [26,27].

In plants, dosage compensation was first documented in *S. latifolia* [15] although its existence was initially disputed as different patterns were observed for X/Y gene pairs and X-hemizygous genes (lacking Y copies) in RNA-seq-based studies [13,15,28]. However, a study comparing sex-linked genes found in BAC sequences and in RNA-seq data suggested that X-hemizygous genes were probably under-represented in RNA-seq-based studies of *S. latifolia* [29]. Partial genome sequencing of *S. latifolia* confirmed this finding and provided a large set of X-hemizygous genes [16]. This study also confirmed that partial dosage compensation does exist in *S. latifolia* and that a fraction of both the X/Y gene pairs and X-hemizygous genes are compensated [16]. Dosage compensation in *S. latifolia* seems to be achieved through an upregulation of the maternal X chromosome in both sexes, which is reminiscent of the scenario envisioned by Ohno [30] for the evolution of dosage compensation in mammals [17,31]. The precise molecular mechanism remains unknown [32], but epigenetic studies suggest that the two female X chromosomes bear different epigenetic marks, implying different expression regulation [33,34]. Additionally, evidence for dosage compensation has been found in *Silene otites*, which has a ZW system younger than the *S. latifolia* XY one [35], in *Rumex rothschildianus*, which has an XY_1Y_2 system that is 8–10 Myr old [36], and in *Cannabis sativa*, which has an XY system that is 20–30 Myr old [37]. An important question is whether dosage compensation is a general feature of plant sex chromosomes.

Here, we focus on *Coccinia grandis*, a species in the Cucurbitaceae with a large Y chromosome [38–41]. *C. grandis* is a perennial, tropical liana that can produce fruits in the first year but can reach stem diameters of >8 cm and ages of at least 20 years. The genus comprises 25 species, all of them dioecious. It belongs to the tribe Benincaseae, where its sister genus (*Diplocyclos*) consists of four monoecious species [42]. The *C. grandis* Y chromosome comprises about 200 Mb, making it four times larger than the X chromosome, mainly due to the accumulation of transposable elements (TEs) and satellite repeats, resulting in a 10% difference in the size of male and female genomes [39–41]. Plastid and mitochondrial-like sequences also have accumulated on the *C. grandis* Y chromosome. The repetitive

fraction of the male and female genomes of *C. grandis* is mainly composed of Ty1 copia and Ty3 gypsy LTR elements. Of these elements, five are found in much greater abundance on the Y than on other chromosomes [40]. Despite this heteromorphy, the species has been estimated to be only 3.1 Myr old based on a molecular-clock model applied to a phylogeny with all 25 species [42]. *C. grandis* males and females show no morphological differences except in their flowers, but between 2 and 8% of the genes appear to be differentially expressed between males and females [43,44]. The species is a promising system in which to study sex chromosome evolution because of its relatively small genome size (~1 Gb) and its phylogenetic proximity to *Cucumis* [45], which includes the fully sequenced *Cucumis sativus* and *Cucumis melo* genomes [46–49]. However, no sex-linked genes have been identified so far and no reference genome (with identified sex chromosomes) is currently available in *Coccinia*. Information about the extent of gene loss, the degradation of Y gene expression, the existence of dosage compensation, the genomic distribution of sex-biased genes is thus currently missing.

Here, we identify sex-linked genes in *C. grandis* using a model-based method that we developed and termed SEX-DETector [50] and which uses RNA-seq data to genotype the parents and F1 individuals from a cross. For each SNP, the transmission from parents to offspring of each allele is analyzed. Sex-linkage or autosomal segregation types have typical patterns that the SEX-DETector is able to differentiate even when there are genotyping errors. For example, a bi-allelic SNP in which one allele is transmitted exclusively from father to sons, while the other is transmitted from both parents to all progeny will be identified as an X/Y SNP (with the male-specific allele being the Y allele). The information of all SNPs in a gene is then combined into a probability for the gene to be sex-linked. RNA-seq-based segregation analysis is both relatively cheap and efficient, and has been applied in several plant systems in which sex-linked genes have been identified successfully, initially using empirical methods [13–15,51,52]. More recently, SEX-DETector has been applied in *Cannabis sativa*, *Mercurialis annua*, several *Silene* species, and *Vitis sylvestris* [17,35,37,50,53–55]. Based on the detected sex-linked genes, we aimed to estimate the age of the sex chromosomes and test for Y degeneration, dosage compensation, and sex-biased genes in *Coccinia grandis*.

2. Material and Methods

2.1. Plant Material

RNA sequencing data were obtained from the progeny and the parents of a cross between a male and a female individual of the dioecious plant *C. grandis*, both grown in the experimental fields of IISER Pune, India. Seeds from the cross were collected as soon as the fruit matured. The 24 seedlings raised from these seeds took four to seven months to begin flowering, which allowed sexing of the individuals. Flower buds at early developmental stages 3–4 (defined in [56]) were sampled from plants being grown in the experimental fields. RNAs were isolated from 5 males (sons) and 5 females (daughters) from the F1 generation as well as from their parents.

2.2. RNA Sequencing

The flower buds were sent to IPS2 Paris, France using RNA later ICE kits by Thermo Fisher. Total RNA was extracted from 12 flower bud samples using Agilent's spin column purification method, mRNA was isolated with Oligo-dT Beads from NEB, and RNA-seq libraries were constructed with the Directional Kit from NEB. Sequencing was performed at IPS2 Paris, France, with Illumina NextSeq500 following a paired-end protocol of library preparation (fragment lengths 100–150 bp, 75 bp sequenced from each end). RNA samples were checked for quality, individually tagged, and sequenced (see Supplementary Table S1 for library sizes). The sequence data are available in the ENA database under the study accession number PRJEB39318.

2.3. De Novo Transcriptome Assembly

A reference transcriptome was built for *C. grandis* using Trinity [57] on the combined libraries of the 12 individuals (the parents and their ten offspring). For each individual, 100% identical reads, assumed to be PCR duplicates, were filtered out using the ConDeTri v2.3 trimming software [58]. Reads were then filtered out for sequencing adapters and low quality using ea-utils FASTQ processing utilities v1.04.636 [59]. Cleaned reads from all male and female individuals were combined and assembled with Trinity version 2.4.0 with default settings [57] leading to 276,225 contigs. Poly-A tails were removed from contigs using PRINSEQ v0.20.4 [60] with parameters -trim_tail_left 5 -trim_tail_right 5. rRNA-like sequences were removed using riboPicker version 0.4.3 [61] with parameters—i 90 -c 50 -l 50 and the following databases: SILVA Large subunit reference database, SILVA Small subunit reference database, the GreenGenes database, and the Rfam database. To ensure that X and Y gametologs are assembled in consensus contigs (required for the SEX-DETector analysis, [50]), Trinity components were merged using Cap3 [62], with parameter -p 90 and custom perl scripts. Coding sequences were predicted using Trinity TransDecoder version 3.0.1 [57] and including Pfam domain searches as ORF retention criteria. This assembly included 128,904 ORFs. To avoid mapping X and Y reads on separate contigs of the same gene, we chose to work on the longest ORF predicted per Trinity isoforms, which resulted in a final set of 82,699 contigs (see Table 1). BUSCO v3.0.2 (Benchmarking Universal Single-Copy Orthologs) was used to assess the completeness of our transcriptome according to conserved gene content from the Plant Dataset [63]. Results are shown for full assembly with all ORFs and for the longest ORF per Trinity isoforms, hereafter referred to as our reference transcriptome (128,904 and 82,699 contigs respectively) in Supplementary Table S2.

Table 1. Transcriptome assembly statistics of *Coccinia grandis* flower buds. Statistics in the final Trinity transcriptome and the working transcriptome containing the longest ORF predicted per Trinity isoform.

	Full Transcriptome	Longest ORF per Isoform
Total contigs	128,904	82,699
Total assembled bases (bp)	103,275,123	27,290,670
Median contig length	552	836
Average contig length	801.18	836.83
Maximum contig length	16,296	16,296
Minimum contig length	297	297
N50	1029	1086
Total contigs longer than 1kb	30,795	21,587
GC content (%)	42.96	42.96

2.4. Functional Annotation and Gene Ontology Enrichment Analysis

De novo annotation of our transcriptome was performed using Trinnotate v3.1.0 [57] and resulted in 59,319 annotations for 82,699 contigs (71.73%). Gene Ontology (GO) was assessed using GOSeq [64] version 1.30.0 on R version 3.4.3 (2017-11-30) to identify over or under-represented GO terms (p-value cutoff = 0.05).

2.5. Inferring Sex-Linked Contigs

The raw Illumina reads were mapped on the reference transcriptome using BWA mapping [65] version 0.7.15 with the following commands: bwa aln -n 5 and bwa sampe. Mapping statistics are shown in Supplementary Table S3. Mapped reads were kept with Samtools version 1.3.1 and individual genotypes were predicted with reads2snps version 2.0.64 with paralog detection [66,67], option -aeb which allows alleles to have different expression levels, and -par 0 to avoid removal of paralogous positions by the paraclean program, which tends to filter out X/Y SNPs.

SEX-DETector version 1.0 (3rd September 2017) was used to infer contig segregation types using a stochastic expectation maximization (SEM) algorithm [50]. The detected SNPs were filtered using

Perl scripts to retrieve the autosomal or sex-linked SNPs, when their posterior probability to be either autosomal or sex-linked was higher than 0.8. A contig was then inferred as sex-linked if its global probability of being sex-linked was higher than the probability of it being autosomal and if it at least had one sex-linked SNP without genotyping error. Amongst sex-linked genes, X-linked contigs without a detectable homologous Y-linked copy are called X-hemizygous. Sex-linked contigs with no Y expression were considered as X-hemizygous, the rest as X/Y.

2.6. Correcting Mapping Bias

To avoid biases towards the reference allele in expression level estimates, a second mapping was done with GSNAP [68], a SNP-tolerant mapping software (see Supplementary Table S3). A SNP file of X/Y SNPs identified in the first run of SEX-DETector was produced with home-made perl scripts as described in [17]. Raw Illumina reads were mapped with GSNAP version 2017-11-15 and parameters -m 10 and -N 1. Only uniquely mapped and concordant paired reads were kept for expression analysis. SEX-DETector was run a second time on this new mapping, and the new inferences were used afterwards for all analyses. No significant difference in the number of sex-linked inferences was observed: 1196 X/Y and 168 X-hemizygous contigs were found as shown in Table 2.

Table 2. Results of the SEX-DETector pipeline on the *C. grandis* dataset. Number of contigs assigned by SEX-DETector to an autosomal, X/Y or X-hemizygous segregation type before and after SNP-tolerant mapping.

	SEX-DETector with BWA Mapping	SEX-DETector with GSNAP SNP-Tolerant Mapping
Contigs in final assembly	82,699	82,699
Contigs with enough coverage to be studied	82,689	70,298
Contigs with enough informative SNPs to compute a segregation probability	4320	3801
Contigs assigned to an autosomal segregation type	2889	3706
Contigs assigned to a X-Y segregation type	1239	1196
Contigs assigned to a X-hemizygous segregation type	192	168

2.7. Estimating the Age of Sex Chromosomes

The X and Y ORF sequences were produced by SEX-DETector using only X/Y segregating SNPs, and pairwise dS was estimated by the codeml program implemented in the PAML suite [69] version 4.8 (see Supplementary Figure S1). The age of the sex chromosomes can be obtained from the X/Y gene pairs with the highest synonymous divergence (the first to stop recombining, [3,4]). To get age estimates in millions years, we used three Brassicaceae molecular clocks: 1.5×10^{-8} substitution/synonymous site/year, derived from an assumed divergence time of *Barbarea* and *Cardamine* of 6.0 Myr [70], 7.1×10^{-9} substitutions/site/generation based on spontaneous mutations in *Arabidopsis thaliana* [71], and 4×10^{-9} substitutions/synonymous site/year, derived from a phylogeny calibrated with six Brassicales fossils [72]. We obtained the age estimates as follows: age (in years) = dSmax/rate, using the molecular clock of [70] and [72], and age (in number of generations) = dSmax/2μ, using the molecular clock of [71] and assuming a *C. grandis* generation time of 1 to 5.5 years.

2.8. Estimating Gene Loss

X-hemizygous genes (X-linked genes without detectable Y copies) have been used to infer the extent of gene loss on Y chromosomes. This only gives a rough idea of gene loss as X-hemizygous genes inferred by SEX-DETector comprise both genes with deleted or silenced Y copies (true lost Y genes) and genes with Y copies that are expressed in some tissues but not in the one used for RNA-seq (false lost Y genes). In addition, X-hemizygous contigs are inferred by SEX-DETector from X polymorphism, as explained in [50], whereas X/Y contig inference relies on fixed mutations. X-hemizygous contigs can therefore only be detected in contigs with X polymorphism, resulting in their underestimation [14]. We corrected for this by using the number of X-hemizygous contigs (168) relative to X/Y contigs with

X polymorphism (424) that were listed in the output of SEX-DETector. Premature stop codons were detected using a custom script on X and Y alleles.

2.9. Analysis of Expression Level Differences between X and Y Alleles

2.9.1. Allelic Expression Measurement

Normalized allelic expression of sex-linked contigs was computed as in [17] from SEX-DETector output. Expression of X and Y alleles was computed using reads spanning diagnostic X/Y SNPs only and were normalized using the library size and the number of studied SNPs in the contig. Normalized expression levels were lower in males compared to females in autosomal contigs, as seen in Supplementary Figure S2. This may be because a small subset of genes are very highly expressed in developing male organs, as observed in other plants (e.g., [73]), resulting in an apparent lower expression of housekeeping genes after normalization for total library size. We applied a correction coefficient to male expression using the ratio of median male autosomal expression over median female autosomal expression, to have a comparable median expression in males and females for autosomal contigs. The same ratio was then applied to sex-linked contigs. The results of this correction are shown in Supplementary Figures S2–S4. We used corrected expression levels to prepare Figure 1 and to perform all downstream analyses.

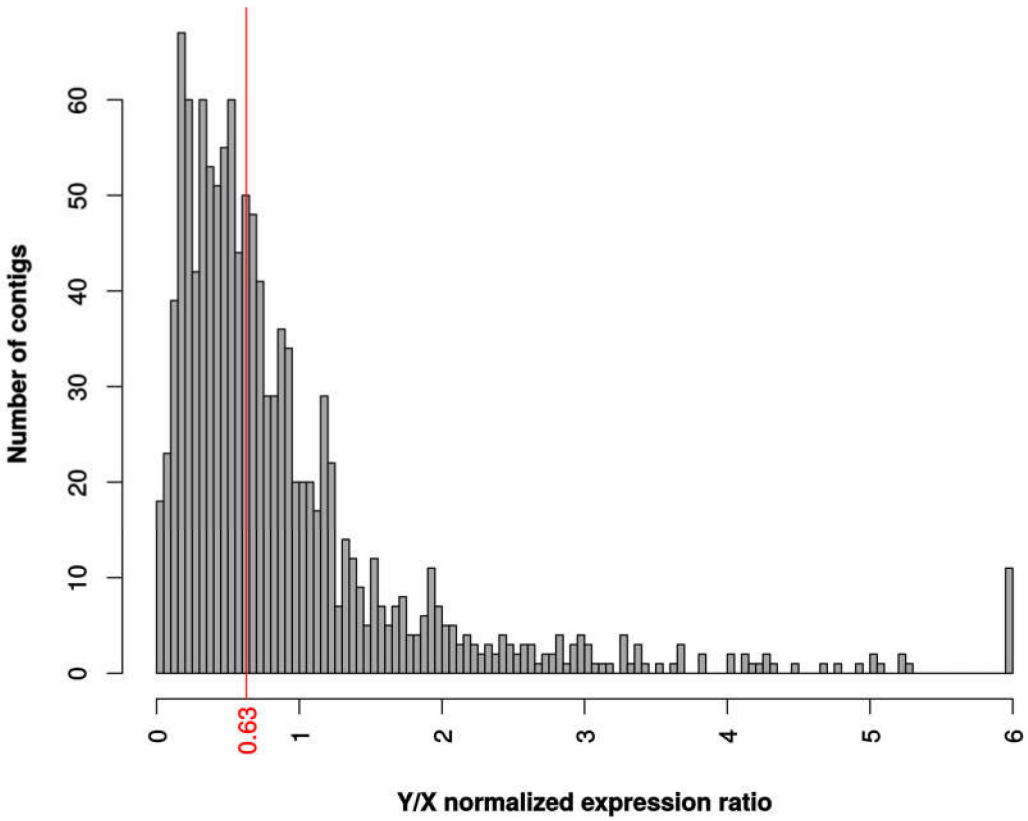

Figure 1. Y/X expression ratio in *C. grandis*. Distribution of normalized expression ratio between X and Y alleles. Total Y and X read numbers were summed at sex-linked SNP locations for each contig and normalized for each male separately, then averaged across males to obtain the Y/X ratio. The median is shown in red.

2.9.2. Analysis of Dosage Compensation in X-Hemizygous Contigs

Contigs with Xmale/2Xfemale ratios above 8 or under 0.125 were excluded as in [15] because we do not expect dosage compensation to occur in these sex-biased genes. To test for dosage

compensation, we filtered genes based on their $\log_2$ male-to-female expression ratio following [74]. X-hemizygous contigs that showed the same expression level in males and in females, i.e., with of $\log_2$(male/female) of 0 ± 0.2 were considered as compensated and X-hemizygous contigs with twice as much expression in females compared to males, i.e., with of $\log_2$(male/female) of −1 ± 0.2 were considered as non-compensated (see Figure 2). A wider range of $\log_2$(male/female) of ±0.5, as used in [27] for dosage-compensated contigs, had the same GO enrichment categories. A smaller range of ±0.1 did not have enough contigs per category to allow for a significant GO enrichment analysis.

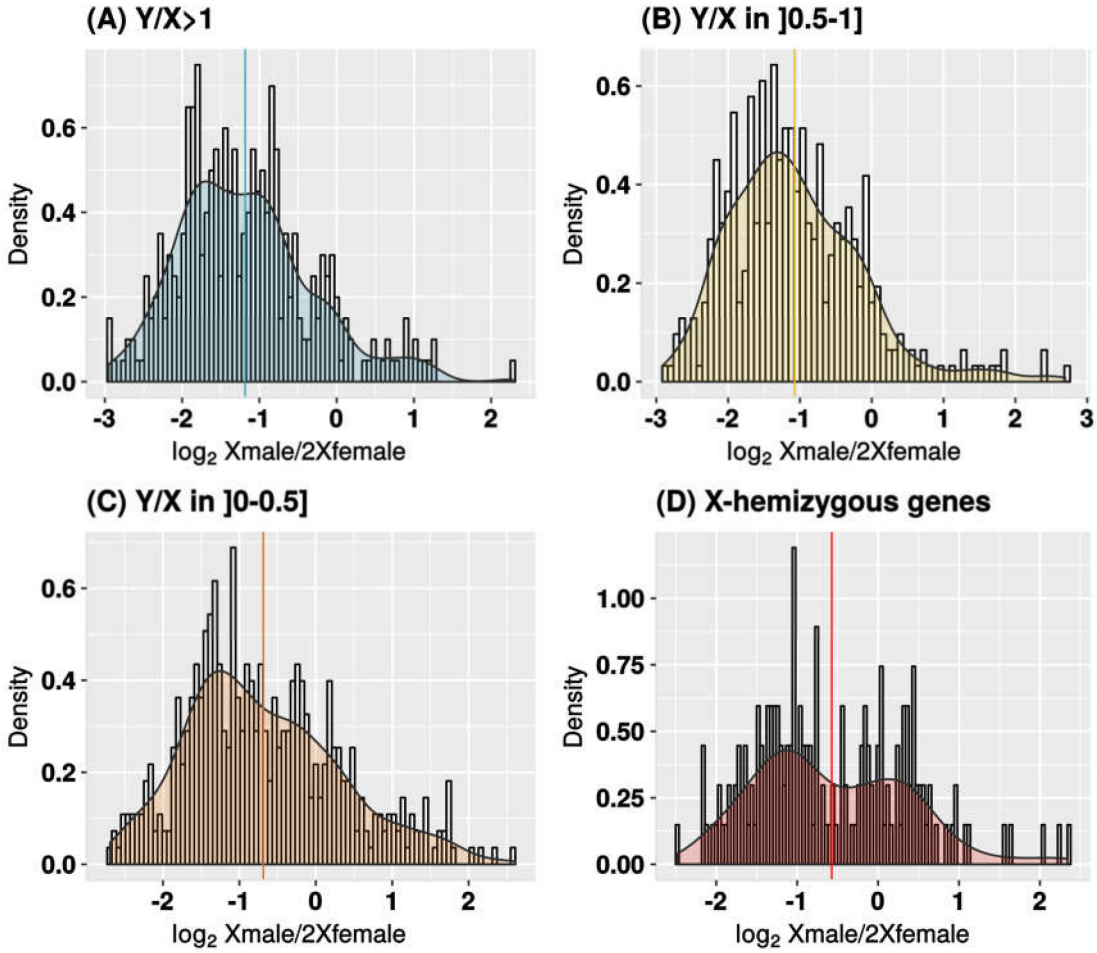

Figure 2. X expression in males versus females in *C. grandis*. Distribution of the ratio between the expression of the single X in males and the two X copies in females ($\log_2$ Xmale/2Xfemale) for all sex-linked contigs. Distributions are shown for Y/X expression ratio categories in males: (**A**) Y/X>1; (**B**) X/Y in]0.5–1]; (**C**) X/Y in]0–0.5]; (**D**) X-hemizygous genes (no Y copy expression). Total X read numbers were summed at sex-linked SNP locations in each contig and normalized for each individual separately, then averaged among males and females to get the Xmale/2Xfemale ratio. Distribution is shown in $\log_2$ scale with its density curve. Contigs with Xmale/2Xfemale ratios above 8 or under 0.125 were excluded, which reduced the dataset to 1351 sex-linked contigs. Sample sizes are: 0, 168; 0–0.5, 460 (8 outliers); 0.5–1, 389 (2 outliers); >1, 334 (3 outliers). Medians are indicated for each Y/X ratio category.

2.10. Identifying Contigs with Sex-Biased Expression

We used the R packages DESeq2 [75] version 1.18.1, edgeR [76,77] version 3.20.8, and Limma-Voom [78] version 3.34.8 in R version 3.4.3 (2017-11-30) to perform biased gene expression analysis between males and females. DGE analysis was performed on the raw read counts (untransformed, not normalized for sequencing depth). Contigs with a CPM (count per million) lower

than 0.5 (corresponding to a count of 10) were filtered out. Regularized log-transformation of the DESeq2 package was used to reduce variance of low read counts. Normalization with edgeR was made with a tagwise dispersion and GLM normalization method (calcNormFactors, estimateTagwiseDisp, and glmLRT functions). With LimmaVoom, counts were fitted to a linear model and differential expression was computed by empirical Bayes (lmFit and eBayes functions).

Contigs identified with at least two methods with an FDR cutoff of 0.0001 were retained as differentially expressed (see Supplementary Figure S5 and Supplementary Table S4). The three methods do not have the same characteristics, and keeping contigs identified with at least two methods as differentially expressed is more robust. DESeq2 and edgeR are based on the same method, but edgeR has fewer false positives. LimmaVoom is very specific (see Supplementary Figure S5), and therefore also has few false positives. Keeping contigs identified with at least two methods is a way to both remove false negatives of DESeq2 and EdgeR and retain true positives that LimmaVoom tends to discard.

2.11. Statistics

Unless stated otherwise in the relevant sections, all statistical analyses and graphs were done with R version 3.4.3 [79]. Fisher's exact tests were two-tailed and *p*-values were adjusted with the FDR method [80]. Exact adjusted *p*-values are provided for each test.

3. Results

3.1. Sex-Linked Genes Identified by SEX-DETector

We assembled a de novo transcriptome for *C. grandis* with male and female reads (82,699 contigs, see Table 1). We used RNA-seq data from a *C. grandis* F1 cross mapped to our reference transcriptome to identify genes located on the sex chromosomes (Supplementary Table S1). The raw reads were mapped on open reading frames (ORF). A total of 45.76% reads were mapped with standard mapping and 49.24% with SNP-tolerant mapping (see Supplementary Table S3). We divided the contigs expressed in buds into autosomal, sex-linked X/Y (defined as contigs having both X- and Y-linked alleles), and X-hemizygous contigs (sex-linked, but with no Y-copy expression). These categories were inferred from single nucleotide polymorphisms (SNPs) segregating in a family, using a probabilistic model [50]. Out of the 82,699 contigs, 5070 had enough informative SNPs to be assigned to a segregation type, 3706 were inferred as autosomal (73.10% of contigs with enough informative SNPs), 1196 as X/Y (23.59%), and 168 as X-hemizygous (3.31%) (see Table 2).

3.2. Age of the C. grandis XY System

Age estimates are based on the divergence between X and Y copies and three Brassicaceae molecular clocks (see Table 3). We obtained a maximum dS of 0.17 in all contigs and of 0.13 in contigs longer than 1kb. The molecular clocks obtained from fossil-calibrated Brassicaceae phylogenies [70,72] returned age estimates between 8.7 and 34.7 Myr old. The molecular clock obtained from *Arabidopsis thaliana* substitution rates [71] returned estimates ranging from 9.3 to 12.1 Myr, when assuming a generation time of 1 year. The true generation time of *C. grandis*, however, is unknown. One year is the onset of sexual maturity in this plant and corresponds to a lower bond estimate of the generation time. *Coccinia grandis* can reach 20 years in the wild, and individuals of 10 years are known from botanical gardens (S.S. Renner, pers. com). Assuming an average generation time of 5.5 years (from the minimum value of 1 year and the conservative maximum value of 10 years), yields much higher age estimates for XY divergence (see Table 3).

Table 3. Age estimates of the *C. grandis* XY system. These estimates were obtained using the maximum synonymous divergence between X and Y chromosomes and several molecular clocks from Brassicaceae. Estimates are shown in increasing order.

Molecular Clocks	Age Estimates of the Sex Chromosomes, with dS max = 0.17	Age Estimates of the Sex Chromosomes, with dS max = 0.13
From [70], calibrated with an assumed divergence time of *Barbarea* and *Cardamine* of 6.0 My	11.3	8.7
From [71], generation time = 1 year, assumed for *Arabidopsis thaliana*	12.1	9.3
From [71], generation time = 1.5 year	18.2	13.9
From [72], calibrated with six Brassicales fossils	34.7	26.5
From [71], generation time = 5.5 year	66.8	51.1

3.3. Patterns of Y Degeneration

We looked for patterns of degeneration in our data. Males showed lower gene expression than females for most of the genes, because a small subset of genes are very highly expressed in developing male flower buds, resulting in apparent lower mean expression after normalization for total library size (Supplementary Figure S2). After correcting for this expression bias between males and females, we found that sex-linked genes are less expressed in males than in females (see Supplementary Figures S2 and S4), Wilcoxon ranked test p-value = 2.57×10^{-8}). To refine our analysis of Y chromosome degeneration in *C. grandis*, we analyzed the allelic expression of genes inferred as sex-linked, which showed that Y-linked alleles were significantly less expressed than X-linked alleles in males (see Figure 1, Wilcoxon ranked test p-value = 3.06×10^{-13}). Lost Y genes can be detected by SEX-DETector when the Y copy is absent or unexpressed and are assigned as X-hemizygous. But given that X-hemizygous contigs can be inferred from segregation patterns only if there is polymorphism on the X chromosome [14,50], their number may be underestimated when segregation patterns are analyzed. We corrected for this as described in Materials and Methods, which resulted in a corrected rate of gene loss of 39.62%. Another hallmark of degeneration is the presence of premature stop codons. We detected 17 X (1.4%) and 56 Y (4.7%) alleles with a premature stop, implying that premature stop codons are more abundant in Y alleles (Fisher's exact test p-value = 6.9×10^{-6}). These observations clearly point to a significant level of genetic degeneration on the Y chromosome in *C. grandis*.

3.4. Patterns of Dosage Compensation

To determine whether some genes are dosage-compensated, we first studied the $\log_2$ fold change between male and female expressions. In the absence of dosage compensation, the Xmale/2Xfemale expression ratio is expected to be 0.5, so the $\log_2$ of the ratio is expected to be −1, because males (XY) have one X-linked copy and females (XX) have two. This is what we observed for contigs that do not show reduced expression of the Y-linked allele relative to the X-linked allele, i.e., that have a Y/X expression ratio close to 1 (median of $\log_2$ Xmale/2Xfemale ratio is −1.29 for contigs with Y/X > 1; see Figure 2). For contigs with reduced Y expression (low Y/X ratios), we observed a higher Xmale/2Xfemale expression ratio, which suggests that dosage compensation occurs for some genes (median of contigs with Y/X ≤ 0.5 is −0.85; see Figure 2). Finally, in X-hemizygous contigs, the distribution of Xmale/2Xfemale expression ratio was bimodal, with a set of 31 contigs centered on a $\log_2$ Xmale/2Xfemale ratio of −1 (no compensation), and 23 contigs centered on 0 (total compensation). This suggests that these two sets of genes exhibited respectively no compensation and total compensation resulting in equal expression in males and females. This trend is also present in a less visible pattern for X/Y contigs with low Y expression (see Figure 2). To investigate dosage compensation further, we compared expression of X-linked and Y-linked alleles in males and females for different Y/X expression ratio categories (Figure 3), using female expression as a reference. We excluded 1% of the sex-linked contigs that showed either an elevated Y expression (high Y/X ratios) or sex-biased X expression (very high or very low Xmale/2Xfemale ratios). The Y/X ratio was computed in *C. grandis* males and averaged between individuals, and used as a proxy for Y degeneration. In the absence of dosage compensation,

Xmale/2Xfemale expression ratio is expected to be 0.5. Instead, we found that X expression in males increases with decreasing Y expression, which results in similar expression levels of sex-linked contigs in both sexes and provides further evidence for dosage compensation in *C. grandis*.

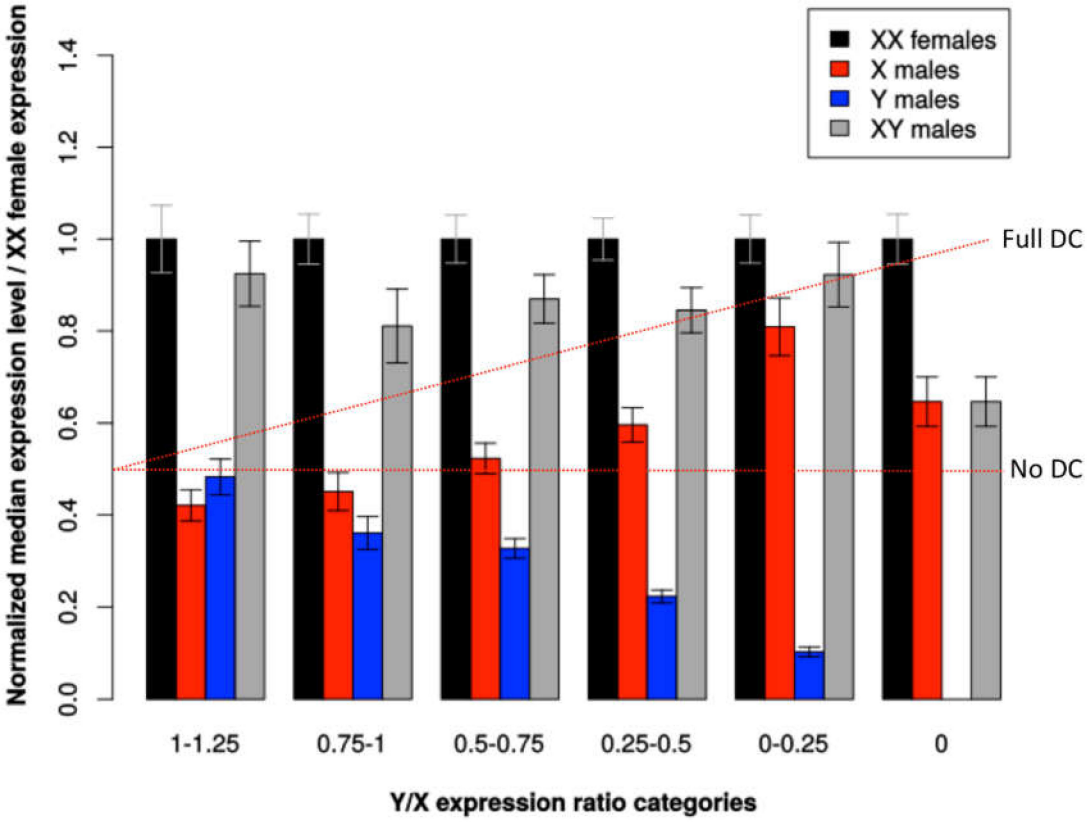

Figure 3. Allele-specific expression of sex-linked genes in both sexes in *C. grandis*. Expression levels of sex-linked contigs in both sexes are shown for different Y/X expression ratio categories. Total read numbers were summed at sex-linked SNP locations in each contig and normalized for each individual separately; medians for all contigs and individuals of the same sex were then obtained. XX females, median expression level of both X-linked alleles in females; X males, median expression level of the single X-linked allele in males; Y males, median expression level of the Y-linked allele in males; XY males, median expression level of the X-linked plus Y-linked alleles in males. To compare different Y/X expression ratio categories, medians were normalized using the XX expression levels in females. Two red lines indicate predictions for X males: full dosage compensation (X males = XX females − Y males) and no dosage compensation (X males = 0.5 XX females). With increasing Y degeneration (measured by the Y/X ratio), Y expression (blue bars) declines. We observed that X males (red bars) does not follow the prediction of no dosage compensation and instead tends to increase with increasing Y degeneration, following the 'full dosage compensation' line. The total expression in males (X males + Y males, grey bars) is mostly maintained across the different categories of Y degeneration and is comparable to the expression in females (black bars). These patterns suggest dosage compensation is taking place in *C. grandis*. The case of the X-hemizygous genes (Y/X = 0) is discussed in the text. Sample sizes are: 0, 168; 0–0.25, 207; 0.25–0.5, 261; 0.5–0.75, 243; 0.75–1, 148; 1–1.25, 108. Error bars indicate 95% confidence intervals of the median.

X/Y genes appeared to be depleted in hormone-related functions, such as response to ethylene and negative regulation of ethylene biosynthetic process (GO:0009723 and GO:0010366). X-hemizygous genes were depleted in dosage-sensitive functions such as macromolecular complex, intracellular ribonucleoprotein complex, ribonucleoprotein complex, transferase complex, and membrane protein

complex (GO:0032991, GO:0030529, GO:1990904, GO:1990234, and GO:0098796). However, these same dosage-sensitive functions were enriched in X-hemizygous genes that show dosage compensation when compared to all sex-linked genes, which suggests compensation targets dosage-sensitive genes.

3.5. Genomic Distribution of Sex-Biased Genes

We detected 3453 sex-biased genes with edgeR, 4881 with DESeq, and 538 with LimmaVoom (Supplementary Figure S5). To establish a robust set of sex-biased genes, we retained genes that were identified as sex-biased with at least two methods (3273 genes). Among these, 2682 (81.94%) were male-biased and 591 (18.06%) female-biased. Genes with sex-biased expression were significantly over-represented among sex-linked genes (see Supplementary Table S4, Fisher's exact test p-value $< 2.2 \times 10^{-16}$), with 241 out of 1364 sex-linked genes being sex-biased (17.67% of sex-linked genes, with respectively 13.42% and 4.25% having male and female-biased expression), and 206 out of 3706 autosomal genes being sex-biased (5.56% of them, 4.45% with male and 1.11% with female-biased expression). Out of the sex-biased genes that were localized on the sex chromosomes, 228 (94.61%) had a X/Y segregation type (181 male-biased and 47 female-biased), and only 11 male-biased genes and 2 female-biased were X-hemizygous. X-hemizygous contigs were not enriched for differentially expressed genes when compared to autosomal contigs (see Supplementary Table S4, Fisher's exact test p-value = 0.2302), which might be due to the small sample of X-hemizygous genes. Gene ontology analysis revealed several biological processes that are significantly over-represented among female-biased genes and under-represented among male-biased genes, or vice versa. GO categories related to pollen production were enriched in male-biased genes, such as pectin catabolic process, pollen wall assembly, sporopollenin biosynthetic process, pollen exine formation, pectin metabolic process, pollination, anther wall tapetum development, pollen sperm cell differentiation, anther development, rejection of self pollen, and regulation of pollen tube growth (GO:0045490, GO:0010208, GO:0080110, GO:0010584, GO:0045488, GO:0009856, GO:0048658, GO:0048235, GO:0048653, GO:0060320, and GO:0080092). Functions related to hormone signaling were enriched in female-biased genes, such as response to auxin, auxin-activated signaling pathway, regulation of ethylene-activated signaling pathway, brassinosteroid mediated signaling pathway, auxin polar transport, auxin mediated signaling pathway involved in phyllotactic patterning, ethylene receptor activity, ethylene binding, auxin transport, jasmonic acid and ethylene-dependent systemic resistance, and ethylene mediated signaling pathway (GO:0009733, GO:0009734, GO:0010104, GO:0009742, GO:0009926, GO:0060774, GO:0038199, GO:0051740, GO:0038199, GO:0051740, GO:0060918, and GO:0009871).

4. Discussion

4.1. Coccinia grandis XY are of Intermediate Age, Similarly to other Highly Heteromorphic Plant Systems

The divergence between X and Y copies in *C. grandis* reaches 0.13 to 0.17, and Brassicaceae-derived molecular clocks returned age estimates from 8.7 to 34.7 Myr for their divergence (and up to >50 Myr old assuming a generation time of 5.5 years, see Table 3). By contrast, a phylogeny that included 24 species of *Coccinia* and six outgroup taxa, calibrated with the divergence time of *Coccinia* and *Diplocyclos* of 15 ± 2.6 Myr, yielded a divergence time of *Coccinia grandis* from its sister species (whose sex chromosomes have not been studied) of 3.1 Myr [42]. In the other plant species with a huge Y chromosome, *S. latifolia*, sex chromosomes have been dated to 11.5 Myr old based on mutation rate measurements [81] or instead <1 to 10 Myr using molecular clocks ([9]: Brassicaceae and *Ipomoea* rates yield ~10 Myr for the sex chromosomes of *S. latifolia*; [82]: fossil-calibrated Caryophyllaceae phylogeny yields a divergence time of <1 – 3.75 Myr for *S. latifolia* and its sister clade). *Silene latifolia* has a generation time of 1 year, and it is thus tempting to think that *C. grandis* sex chromosomes should be as old or older than the ones of *S. latifolia*, unless mutation rate is a lot higher in *C. grandis*, which would be unexpected [81].

4.2. Coccinia grandis Y Chromosome Degeneration is Moderate, with an Unusually Reduced Y Expression

Our results suggest that not only is the Y chromosome of *C. grandis* accumulating repeats as shown previously, but it is also losing genes and becoming silent. Out of the 1364 sex-linked genes, 168 did not have a Y-linked homolog, and we estimated a total gene loss of about 40%, which is similar to what has been found in *S. latifolia* using a methodology similar to this study as well as other methods [16,17,29,50]. The Y/X expression ratio, however, was lower than what has been found in *S. latifolia* (0.63 vs. 0.77 respectively, [15]). The difference in size between the sex chromosomes is larger in *C. grandis* compared to *S. latifolia* [39], which might indicate that the TE load is larger in the former. TEs trigger host defense mechanisms, such as DNA methylation to silence them. This can affect genes close to TE insertions and reduce their expression level [83,84]. The possibly higher TE abundance on the *C. grandis* Y chromosome might trigger a higher level of gene silencing and thus explain its unusually low Y/X expression ratio. To test this idea further, the full sequences of the sex chromosomes in *C. grandis* and a DNA methylation study of TEs and neighboring genes would be needed.

4.3. Coccinia grandis Exhibit Sex Chromosome Dosage Compensation, a Phenomenon Observed in Several Plant Systems

When analyzing the expression of X and Y copies in both sexes, we found that the reduction of the Y copy was compensated by increasing expression of the X copy to maintain similar expression of the pair in both sexes (Figure 3). This suggests that dosage compensation has evolved in *C. grandis*. Again, a very similar pattern has been found in *S. latifolia* [15]. Compensation is probably partial, and not all the sex-linked genes are compensated (Figure 2), something also observed in *S. latifolia* [16]. We mentioned above that X-hemizygous genes are underrepresented in RNA-seq-based studies such as this one, and conclusions about dosage compensation are therefore more difficult to draw for this category of genes (discussed in [29]). Here, we found that the signal of dosage compensation was weak in X-hemizygous genes when taken together (Figure 3) as observed in *S. latifolia* with the same approach [15]. However, when looking at genes individually as in [16], we found evidence that some X-hemizygous genes are fully compensated (see bimodal distribution in Figure 2D). Strikingly, those genes are enriched in dosage-sensitive functions in agreement with findings in animals [21–23]; [19,20,24,25,27]. Evidence for dosage compensation has been found so far in Caryophyllaceae, Cannabaceae, and Polygonaceae [15–17,35–37]. Cucurbitaceae are thus the fourth plant family in which dosage compensation is documented.

This study did not allow us to identify a mechanism explaining the patterns of dosage compensation that we observed. Studies of animal and plant aneuploids have revealed that immediate adjustments of the genetic networks with expression levels changing at both sex-linked and autosomal genes are possible [85–92]. Other studies have revealed that several specific mechanisms of dosage compensation can evolve in parallel. In humans, for example, strong selection has maintained some essential dosage-sensitive genes on the Y, despite this chromosome being extremely degenerated [93]. For some other dosage-sensitive genes, the Y copy has been lost but this has been compensated by up-regulation of the X copy in both sexes, and inactivation of one of the X in females [21,22]. Finally, for a number of highly expressed tissue-specific genes that have lost their Y copy, dosage compensation took place through a process of duplication of the X copy and translocation of the duplicate to an autosome [94]. In plants, this remains to be investigated. In *S. latifolia*, the best-studied plant for dosage compensation, there are still debates about how dosage compensation is achieved [17,32]. In the present study, we assumed that autosomal expression was unchanged by Y degeneration, and used autosomal expression to further normalize male and female expression of sex-linked genes. Future studies should check this assumption, and could include an outgroup to study the evolution of both autosomal and sex-linked expression in *C. grandis* as in [17,53].

4.4. Coccinia grandis Sex Chromosomes are Enriched in Sex-Biased Genes

We found that 4% of the genes expressed in *C. grandis* floral buds are sex-biased (Supplementary Table S4), with a total of 3273 sex-biased genes identified (2682 male-biased and 591 female-biased), in agreement with prior studies on sex-biased genes in *C. grandis* [43,44]. Our results support enrichment for pollen production-related functions in male-biased genes as previously found in *C. grandis* [43,44] and in other plant systems (reviewed in [5]). Female-biased genes, on the other hand, were significantly enriched for hormone-signaling functions. These observations suggest that sex-biased expression may have evolved to support contrasting biological functions in *C. grandis* females and males. Male-biased genes are also significantly more numerous than female-biased genes, a pattern that is common in dioecious plants (e.g., [5,53,95,96] but see [97–99]). Lastly, sex-biased genes were found both on autosomes and on sex chromosomes but the latter were significantly enriched in such genes, again a common pattern in dioecious plants (for review, [5]; see also [97,98]). The most probable hypothesis is that sexually antagonistic selection favors sex linkage of sex-biased genes involved in sexual dimorphism, but further work to identify the footprints of sexually antagonistic selection (such as in [53]) will be needed to test this idea in *C. grandis*.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/11/7/787/s1, Figure S1: Distribution of pairwise synonymous divergence (dS) between X and Y alleles in X/Y contigs, Figure S2: Expression levels of autosomal and sex-linked contigs in both sexes before and after correction, Figure S3: Male vs. female autosomal gene expression ratio after correction, Figure S4: Male vs. female sex-linked gene expression ratio after correction, Figure S5: Venn diagram of Differentially Expressed Genes identified by DESeq2, edgeR, and LimmaVoom, Table S1: Library sizes and statistics for RNA sequencing from *C. grandis* male and female flower buds, Table S2: BUSCO results for *C. grandis* flower bud transcriptomes, Table S3: Mapping statistics of all the samples, Table S4: Number of Differentially Expressed Genes (DEG) per gene categories.

Author Contributions: Conceptualization, G.A.B.M., A.B., A.K.B.; Formal analysis, C.F., H.B., A.M.; Funding acquisition, G.A.B.M., A.B., S.S.R.; Investigation, C.F., H.B., D.L., R.S.D., A.M., B.R., A.K.B., A.B., G.A.B.M.; Project administration, G.A.B.M.; Resources, A.B., A.K.B., B.R.; Supervision, G.A.B.M., A.B., A.K.B., H.B.; Visualization, C.F.; Writing—Original Draft, C.F., G.A.B.M., S.S.R.; Writing—Review & Editing, G.A.B.M., S.S.R., C.F., A.M., H.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Agence Nationale de la Recherche (ANR), grant ANR-14-CE19-0021-01 to G.A.B.M. and A.B., D.L. and A.B. also acknowledge support from the European Research Council (ERC), ERC Advanced Grant SEXYPARTH.

Acknowledgments: All the bioinformatics was performed using the computing facilities of the CC LBBE/PRABI. We thank Stépahne Delmotte and Bruno Spataro for help in using the cluster, and for its maintenance. We thank Adnane Boualem for help with the RNA-seq data preparation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. 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]
2. Ming, R.; Bendahmane, A.; Renner, S.S. Sex chromosomes in land plants. *Annu. Rev. Plant Biol.* **2011**, *62*, 485–514. [CrossRef] [PubMed]
3. Charlesworth, D.; Charlesworth, B.; Marais, G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **2005**, *95*, 118–128. [CrossRef] [PubMed]
4. Bergero, R.; Charlesworth, D. The evolution of restricted recombination in sex chromosomes. *Trends Ecol. Evol.* **2009**, *24*, 94–102. [CrossRef]
5. Muyle, A.; Shearn, R.; Marais, G.A. The evolution of sex chromosomes and dosage compensation in plants. *Genome Biol. Evol.* **2017**, *9*, 627–645. [CrossRef]
6. Bachtrog, D. Y-chromosome evolution: Emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genome* **2013**, *14*, 113–124. [CrossRef]
7. Mahajan, S.; Wei, K.H.-C.; Nalley, M.J.; Gibilisco, L.; Bachtrog, D. De novo assembly of a young Drosophila Y chromosome using single-molecule sequencing and chromatin conformation capture. *PLoS Biol.* **2018**, *16*, e2006348. [CrossRef]

8. Filatov, D.A.; Moneger, F.; Negrutiu, I.; Charlesworth, D. Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution. *Nature* **2000**, *404*, 388–390. [CrossRef]
9. Nicolas, M.; Marais, G.; Hykelova, V.; Janousek, B.; Laporte, V.; Vyskot, B.; Mouchiroud, D.; Negrutiu, I.; Charlesworth, D.; Moneger, F. A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants. *PLoS Biol.* **2005**, *3*, e4. [CrossRef]
10. Bergero, R.; Charlesworth, D.; Filatov, D.A.; Moore, R.C. Defining regions and rearrangements of the Silene latifolia Y chromosome. *Genetics* **2008**, *178*, 2045–2053. [CrossRef]
11. Cermak, T.; Kubat, Z.; Hobza, R.; Koblizkova, A.; Widmer, A.; Macas, J.; Vyskot, B.; Kejnovsky, E. Survey of repetitive sequences in Silene latifolia with respect to their distribution on sex chromosomes. *Chromosome Res.* **2008**, *16*, 961–976. [CrossRef] [PubMed]
12. Marais, G.; Nicolas, M.; Bergero, R.; Chambrier, P.; Kejnovsky, E.; Monéger, F.; Hobza, R.; Widmer, A.; Charlesworth, D. Evidence for degeneration of the Y chromosome in the dioecious plant *Silene Latifolia*. *Curr. Biol.* **2008**, *18*, 545–549. [CrossRef] [PubMed]
13. Chibalina, M.V.; Filatov, D.A. Plant Y chromosome degeneration is retarded by haploid purifying selection. *Curr. Biol.* **2011**, *21*, 1475–1479. [CrossRef] [PubMed]
14. Bergero, R.; Charlesworth, D. Preservation of the Y transcriptome in a 10-million-year-old plant sex chromosome system. *Curr. Biol.* **2011**, *21*, 1470–1474. [CrossRef]
15. Muyle, A.; Zemp, N.; Deschamps, C.; Mousset, S.; Widmer, A.; Marais, G.A. Rapid de novo evolution of X chromosome dosage compensation in *Silene latifolia*, a plant with young sex chromosomes. *PLoS Biol.* **2012**, *10*, e1001308. [CrossRef]
16. Papadopulos, A.S.; Chester, M.; Ridout, K.; Filatov, D.A. Rapid Y degeneration and dosage compensation in plant sex chromosomes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 13021–13026. [CrossRef]
17. Muyle, A.; Zemp, N.; Fruchard, C.; Cegan, R.; Vrana, J.; Deschamps, C.; Tavares, R.; Hobza, R.; Picard, F.; Widmer, A.; et al. Genomic imprinting mediates dosage compensation in a young plant XY system. *Nat. Plants* **2018**, *4*, 677–680. [CrossRef]
18. Lorenzo, J.L.R.; Hobza, R.; Vyskot, B. DNA methylation and genetic degeneration of the Y chromosome in the dioecious plant Silene latifolia. *BMC Gen.* **2018**, *19*, 540.
19. Gu, L.; Walters, J.R. Evolution of sex chromosome dosage compensation in animals: A beautiful theory, undermined by facts and bedeviled by details. *Genome Biol. Evol.* **2017**, *9*, 2461–2476. [CrossRef]
20. Ercan, S. Mechanisms of X chromosome dosage compensation. *J. Gen.* **2015**, *3*, 1. [CrossRef]
21. Pessia, E.; Makino, T.; Bailly-Bechet, M.; McLysaght, A.; Marais, G.A. Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5346–5351. [CrossRef]
22. Pessia, E.; Engelstädter, J.; Marais, G.A. The evolution of X chromosome inactivation in mammals: The demise of Ohno's hypothesis? *Cell. Mol. Life Sci.* **2014**, *71*, 1383–1394. [CrossRef] [PubMed]
23. Mank, J.E. Sex chromosome dosage compensation: Definitely not for everyone. *Trends Gen.* **2013**, *29*, 677–683. [CrossRef] [PubMed]
24. Albritton, S.E.; Kranz, A.-L.; Rao, P.; Kramer, M.; Dieterich, C.; Ercan, S. Sex-biased gene expression and evolution of the x chromosome in nematodes. *Genetics* **2014**, *197*, 865–883. [CrossRef] [PubMed]
25. Veitia, R.A.; Veyrunes, F.; Bottani, S.; Birchler, J.A. X chromosome inactivation and active X upregulation in therian mammals: Facts, questions, and hypotheses. *J. Mol. Cell Biol.* **2015**, *7*, 2–11. [CrossRef]
26. Arnold, A.P.; Itoh, Y.; Melamed, E. A bird's-eye view of sex chromosome dosage compensation. *Annu. Rev. Gen. Hum. Genet.* **2008**, *9*, 109–127. [CrossRef]
27. Zimmer, F.; Harrison, P.W.; Dessimoz, C.; Mank, J.E. Compensation of dosage-sensitive genes on the chicken Z chromosome. *Gen. Biol. Evol.* **2016**, *8*, 1233–1242. [CrossRef]
28. Bergero, R.; Qiu, S.; Charlesworth, D. Gene loss from a plant sex chromosome system. *Curr. Biol.* **2015**, *25*, 1234–1240. [CrossRef]
29. Blavet, N.; Blavet, H.; Muyle, A.; Kafer, J.; Cegan, R.; Deschamps, C.; Zemp, N.; Mousset, S.; Aubourg, S.; Bergero, R.; et al. Identifying new sex-linked genes through BAC sequencing in the dioecious plant Silene latifolia. *BMC Gen.* **2015**, *16*, 546. [CrossRef]
30. Ohno, S. *Sex. Chromosomes and Sex. Linked Genes*; Springer: Berlin, Germany, 1967.
31. Bergero, R. How to solve the gender expression gap. *Nat. Plants* **2018**, *4*, 637–638. [CrossRef]

32. Krasovec, M.; Kazama, Y.; Ishii, K.; Abe, T.; Filatov, D.A. Immediate dosage Compensation is triggered by the deletion of Y-linked genes in Silene latifolia. *Curr. Biol.* **2019**, *29*, 2214–2221. [CrossRef] [PubMed]
33. Siroky, J.; Ruffini Castiglione, M.; Vyskot, B. DNA methylation patterns of Melandrium album chromosomes. *Chromosome Res.* **1998**, *6*, 441–446. [CrossRef]
34. Bačovský, V.; Houben, A.; Kumke, K.; Hobza, R. The distribution of epigenetic histone marks differs between the X and Y chromosomes in Silene latifolia. *Planta* **2019**, *250*, 487–494. [CrossRef] [PubMed]
35. Martin, H.; Carpentier, F.; Gallina, S.; Gode, C.; Schmitt, E.; Muyle, A.; Marais, G.A.B.; Touzet, P. Evolution of young sex chromosomes in two dioecious sister plant species with distinct sex determination systems. *Genome Biol. Evol.* **2019**, *11*, 350–361. [CrossRef] [PubMed]
36. Crowson, D.; Barrett, S.C.H.; Wright, S.I. Purifying and positive selection influence patterns of gene loss and gene expression in the evolution of a plant sex chromosome system. *Mol. Biol. Evol.* **2017**, *34*, 1140–1154. [CrossRef] [PubMed]
37. Prentout, D.; Razumova, O.; Rhoné, B.; Badouin, H.; Henri, H.; Feng, C.; Käfer, J.; Karlov, G.; Marais, G.A. An efficient RNA-seq-based segregation analysis identifies the sex chromosomes of Cannabis sativa. *Genome Res.* **2020**, *30*, 164–172. [CrossRef] [PubMed]
38. Kumar, L.S.S.; Vishveshwaraiah, S. Sex mechanism in Coccinia indica Wight and Arn. *Nature* **1952**, *170*, 330–331. [CrossRef]
39. 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]
40. Sousa, A.; Bellot, S.; Fuchs, J.; Houben, A.; Renner, S.S. Analysis of transposable elements and organellar DNA in male and female genomes of a species with a huge Y chromosome reveals distinct Y centromeres. *Plant J.* **2016**, *88*, 387–396. [CrossRef]
41. 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]
42. Holstein, N.; Renner, S.S. A dated phylogeny and collection records reveal repeated biome shifts in the African genus Coccinia (Cucurbitaceae). *BMC Evol. Biol.* **2011**, *11*, 28. [CrossRef] [PubMed]
43. Devani, R.S.; Sinha, S.; Banerjee, J.; Sinha, R.K.; Bendahmane, A.; Banerjee, A.K. De novo transcriptome assembly from flower buds of dioecious, gynomonoecious and chemically masculinized female *Coccinia grandis* reveals genes associated with sex expression and modification. *BMC Plant Biol.* **2017**, *17*, 241. [CrossRef]
44. Mohanty, J.N.; Nayak, S.; Jha, S.; Joshi, R.K. Transcriptome profiling of the floral buds and discovery of genes related to sex-differentiation in the dioecious cucurbit *Coccinia grandis* (L.) Voigt. *Gene* **2017**, *626*, 395–406. [CrossRef] [PubMed]
45. Schaefer, H.; Heibl, C.; Renner, S.S. Gourds afloat: A dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proc. R. Soc. B Biol. Sci.* **2009**, *276*, 843–851. [CrossRef] [PubMed]
46. Huang, S.; Li, R.; Zhang, Z.; Li, L.; Gu, X.; Fan, W.; Lucas, W.J.; Wang, X.; Xie, B.; Ni, P.; et al. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* **2009**, *41*, 1275–1281. [CrossRef]
47. Garcia-Mas, J.; Benjak, A.; Sanseverino, W.; Bourgeois, M.; Mir, G.; González, V.M.; Hénaff, E.; Câmara, F.; Cozzuto, L.; Lowy, E.; et al. The genome of melon (*Cucumis melo* L.). *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11872–11877. [CrossRef]
48. Ruggieri, V.; Alexiou, K.G.; Morata, J.; Argyris, J.; Pujol, M.; Yano, R.; Nonaka, S.; Ezura, H.; Latrasse, D.; Boualem, A.; et al. An improved assembly and annotation of the melon (*Cucumis melo* L.) reference genome. *Sci Rep.* **2018**, *8*, 8088. [CrossRef]
49. Li, Q.; Li, H.; Huang, W.; Xu, Y.; Zhou, Q.; Wang, S.; Ruan, J.; Huang, S.; Zhang, Z. A chromosome-scale genome assembly of cucumber (*Cucumis sativus* L.). *GigaScience* **2019**, *8*. [CrossRef]
50. Muyle, A.; Kafer, J.; Zemp, N.; Mousset, S.; Picard, F.; Marais, G.A. SEX-DETector: A probabilistic approach to study sex chromosomes in non-model organisms. *Genome Biol. Evol.* **2016**, *8*, 2530–2543. [CrossRef]
51. Hough, J.; Hollister, J.D.; Wang, W.; Barrett, S.C.; Wright, S.I. Genetic degeneration of old and young Y chromosomes in the flowering plant Rumex hastatulus. *Proc. Natl. Acad. Sci. USA* **2014**. [CrossRef]

52. Michalovova, M.; Kubat, Z.; Hobza, R.; Vyskot, B.; Kejnovsky, E. Fully automated pipeline for detection of sex linked genes using RNA-Seq data. *BMC Bioinform.* **2015**, *16*, 78. [CrossRef] [PubMed]
53. Zemp, N.; Tavares, R.; Muyle, A.; Charlesworth, D.; Marais, G.A.; Widmer, A. Evolution of sex-biased gene expression in a dioecious plant. *Nat. Plants* **2016**, *2*, 16168. [CrossRef] [PubMed]
54. Veltsos, P.; Ridout, K.E.; Toups, M.A.; Gonzalez-Martinez, S.C.; Muyle, A.; Emery, O.; Rastas, P.; Hudzieczek, V.; Hobza, R.; Vyskot, B.; et al. Early sex-chromosome evolution in the diploid dioecious plant Mercurialis annua. *Genetics* **2019**. [CrossRef] [PubMed]
55. Badouin, H.; Velt, A.; Gindraud, F.; Flutre, T.; Dumas, V.; Vautrin, S.; Marande, W.; Corbi, J.; Sallet, E.; Ganofsky, J.; et al. The wild grape genome sequence provides insights into the transition from dioecy to hermaphroditism during grape domestication. *bioRxiv* **2020**. [CrossRef]
56. Ghadge, A.G.; Karmakar, K.; Devani, R.S.; Banerjee, J.; Mohanasundaram, B.; Sinha, R.K.; Sinha, S.; Banerjee, A.K. Flower development, pollen fertility and sex expression analyses of three sexual phenotypes of Coccinia grandis. *BMC Plant Biol.* **2014**, *14*, 325. [CrossRef] [PubMed]
57. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [CrossRef]
58. Smeds, L.; Künstner, A. ConDeTri-a content dependent read trimmer for Illumina data. *PLoS ONE* **2011**, *6*, e26314. [CrossRef]
59. Aronesty, E. Ea-utils: "Command-Line Tools for Processing Biological Sequencing Data". 2011. Available online: https://github.com/ExpressionAnalysis/ea-utils (accessed on 12 July 2017).
60. Schmieder, R.; Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **2011**, *27*, 863–864. [CrossRef]
61. Schmieder, R.; Lim, Y.W.; Edwards, R. Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* **2012**, *28*, 433–435. [CrossRef]
62. Huang, X.; Madan, A. CAP3: A DNA sequence assembly program. *Genome Res.* **1999**, *9*, 868–877. [CrossRef]
63. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [CrossRef] [PubMed]
64. Young, M.D.; Wakefield, M.J.; Smyth, G.K.; Oshlack, A. goseq: Gene Ontology testing for RNA-seq datasets. *R Bioconductor* **2012**, *8*, 1–25.
65. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]
66. Gayral, P.; Melo-Ferreira, J.; Glemin, S.; Bierne, N.; Carneiro, M.; Nabholz, B.; Lourenco, J.M.; Alves, P.C.; Ballenghien, M.; Faivre, N.; et al. Reference-free population genomics from next-generation transcriptome data and the vertebrate-invertebrate gap. *PLoS Genet.* **2013**, *9*, e1003457. [CrossRef]
67. Tsagkogeorga, G.; Cahais, V.; Galtier, N. The population genomics of a fast evolver: High levels of diversity, functional constraint, and molecular adaptation in the tunicate Ciona intestinalis. *Genome Biol. Evol.* **2012**, *4*, 740–749. [CrossRef]
68. Wu, T.D.; Nacu, S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **2010**, *26*, 873–881. [CrossRef]
69. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [CrossRef]
70. Koch, M.; Haubold, B.; Mitchell-Olds, T. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis* and related genera (Brassicaceae). *Mol. Biol. Evol.* **2000**, *17*, 1483–1498. [CrossRef]
71. Ossowski, S.; Schneeberger, K.; Lucas-Lledó, J.I.; Warthmann, N.; Clark, R.M.; Shaw, R.G.; Lynch, M. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis Thaliana*. *Science* **2010**, *327*, 92–94. [CrossRef]
72. Beilstein, M.A.; Nagalingum, N.S.; Clements, M.D.; Manchester, S.R.; Mathews, S. Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18724–18728. [CrossRef]

73. Badouin, H.; Gouzy, J.; Grassa, C.J.; Murat, F.; Staton, S.E.; Cottret, L.; Lelandais-Brière, C.; Owens, G.L.; Carrère, S.; Mayjonade, B.; et al. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* **2017**, *546*, 148–152. [CrossRef] [PubMed]
74. Mullon, C.; Wright, A.E.; Reuter, M.; Pomiankowski, A.; Mank, J.E. Evolution of dosage compensation under sexual selection differs between X and Z chromosomes. *Nat. Commun.* **2015**, *6*, 7720. [CrossRef] [PubMed]
75. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]
76. McCarthy, D.J.; Chen, Y.; Smyth, G.K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucl. Acids Res.* **2012**, *40*, 4288–4297. [CrossRef] [PubMed]
77. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef]
78. Law, C.W.; Chen, Y.; Shi, W.; Smyth, G.K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **2014**, *15*, R29. [CrossRef] [PubMed]
79. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015; Available online: http/www.R-project.org (accessed on 30 November 2017).
80. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **1995**, *57*, 289–300. [CrossRef]
81. Krasovec, M.; Chester, M.; Ridout, K.; Filatov, D.A. The Mutation Rate and the Age of the Sex Chromosomes in Silene latifolia. *Curr. Biol.* **2018**, *28*, 1832–1838. [CrossRef]
82. Rautenberg, A.; Sloan, D.B.; Alden, V.; Oxelman, B. Phylogenetic relationships of Silene multinervia and Silene section conoimorpha (Caryophyllaceae). *Syst. Bot.* **2012**, *37*, 226–237. [CrossRef]
83. Hollister, J.D.; Smith, L.M.; Guo, Y.L.; Ott, F.; Weigel, D.; Gaut, B.S. Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2322–2327. [CrossRef]
84. Gorelick, R. Evolution of dioecy and sex chromosomes via methylation driving Muller's ratchet. *Biol. J. Linn. Soc.* **2003**, *80*, 353–368. [CrossRef]
85. Birchler, J.A. A study of enzyme activities in a dosage series of the long arm of chromosome one in maize. *Genetics* **1979**, *92*, 1211–1229. [PubMed]
86. Birchler, J.A. The genetic basis of dosage compensation of alcohol dehydrogenase-1 in maize. *Genetics* **1981**, *97*, 625–637.
87. Birchler, J.A.; Newton, K.J. Modulation of protein levels in chromosomal dosage series of maize: The biochemical basis of aneuploid syndromes. *Genetics* **1981**, *99*, 247–266. [PubMed]
88. Birchler, J.A.; Hiebert, J.C.; Paigen, K. Analysis of autosomal dosage compensation involving the alcohol dehydrogenase locus in Drosophila melanogaster. *Genetics* **1990**, *124*, 679–686. [PubMed]
89. Guo, M.; Birchler, J.A. Tran-acting dosage effects on the expression of model gene systems in maize aneuploids. *Science* **1994**, *266*, 1999–2002. [CrossRef] [PubMed]
90. Sun, L.; Johnson, A.F.; Donohue, R.C.; Li, J.; Cheng, J.; Birchler, J.A. Dosage compensation and inverse effects in triple X metafemales of Drosophila. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7383–7388. [CrossRef]
91. Sun, L.; Johnson, A.F.; Li, J.; Lambdin, A.S.; Cheng, J.; Birchler, J.A. Differential effect of aneuploidy on the X chromosome and genes with sex-biased expression in Drosophila. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16514–16519. [CrossRef]
92. Hou, J.; Shi, X.; Chen, C.; Islam, M.S.; Johnson, A.F.; Kanno, T.; Huettel, B.; Yen, M.R.; Hsu, F.M.; Ji, T.; et al. Global impacts of chromosomal imbalance on gene expression in Arabidopsis and other taxa. *Proc. Natl Acad. Sci. USA* **2018**, *115*, E11321–E11330. [CrossRef]
93. Bellott, D.W.; Hughes, J.F.; Skaletsky, H.; Brown, L.G.; Pyntikova, T.; Cho, T.J.; Koutseva, N.; Zaghlul, S.; Graves, T.; Rock, S.; et al. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* **2014**, *508*, 494–499. [CrossRef]
94. Hurst, L.D.; Ghanbarian, A.T.; Forrest, A.R.; Huminiecki, L. The constrained maximal expression level owing to haploidy shapes gene content on the mammalian X chromosome. *PLoS Biol.* **2015**, *13*, e1002315. [CrossRef] [PubMed]
95. Harkess, A.; Mercati, F.; Shan, H.Y.; Sunseri, F.; Falavigna, A.; Leebens-Mack, J. Sex-biased gene expression in dioecious garden asparagus (Asparagus officinalis). *New Phytol.* **2015**, *207*, 883–892. [CrossRef] [PubMed]

96. Cossard, G.G.; Toups, M.A.; Pannell, J.R. Sexual dimorphism and rapid turnover in gene expression in pre-reproductive seedlings of a dioecious herb. *Ann. Bot.* **2019**, *123*, 1119–1131. [CrossRef] [PubMed]
97. Darolti, I.; Wright, A.E.; Pucholt, P.; Berlin, S.; Mank, J.E. Slow evolution of sex-biased genes in the reproductive tissue of the dioecious plant Salix viminalis. *Mol. Ecol.* **2018**, *27*, 694–708. [CrossRef] [PubMed]
98. Sanderson, B.J.; Wang, L.; Tiffin, P.; Wu, Z.; Olson, M.S. Sex-biased gene expression in flowers, but not leaves, reveals secondary sexual dimorphism in Populus balsamifera. *New Phytol.* **2019**, *221*, 527–539. [CrossRef] [PubMed]
99. Muyle, A. How different is the evolution of sex-biased gene expression between plants and animals? A commentary on: 'Sexual dimorphism and rapid turnover in gene expression in pre-reproductive seedlings of a dioecious herb'. *Ann. Bot.* **2019**, *123*, iv–v. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).