THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

The dryas iulia genome supports multiple gains of a W chromosome from a B chromosome in butterflies

Citation for published version:
Lewis, JJ, Cicconardi, F, Martin, SH, Reed, RD, Danko, CG & Montgomery, SH 2021, 'The dryas iulia genome supports multiple gains of a W chromosome from a B chromosome in butterflies', *Genome Biology and Evolution*, vol. 13, no. 7, evab128. https://doi.org/10.1093/gbe/evab128

Digital Object Identifier (DOI):
10.1093/gbe/evab128

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Genome Biology and Evolution

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 17. Aug. 2021

The *Dryas iulia* Genome Supports Multiple Gains of a W Chromosome from a B Chromosome in Butterflies

James J. Lewis[1,2,*,†], Francesco Cicconardi[3,4,*,†], Simon H. Martin[5], Robert D. Reed[2], Charles G. Danko[1], and Stephen H. Montgomery[3,*]

[1]Baker Institute for Animal Health, Cornell University, Ithaca, New York, USA

[2]Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York, USA

[3]School of Biological Sciences, University of Bristol, United Kingdom

[4]Department of Zoology, University of Cambridge, United Kingdom

[5]Institute of Evolutionary Biology, University of Edinburgh, United Kingdom

[†]These authors contributed equally to this work.

*Corresponding authors: E-mails: jjl336@cornell.edu; f.cicconardi@bristol.ac.uk; s.montgomery@bristol.ac.uk.

Accepted: 29 May 2021

Abstract

In butterflies and moths, which exhibit highly variable sex determination mechanisms, the homogametic Z chromosome is deeply conserved and is featured in many genome assemblies. The evolution and origin of the female W sex chromosome, however, remains mostly unknown. Previous studies have proposed that a ZZ/Z0 sex determination system is ancestral to Lepidoptera, and that W chromosomes may originate from sex-linked B chromosomes. Here, we sequence and assemble the female *Dryas iulia* genome into 32 highly contiguous ordered and oriented chromosomes, including the Z and W sex chromosomes. We then use sex-specific Hi-C, ATAC-seq, PRO-seq, and whole-genome DNA sequence data sets to test if features of the *D. iulia* W chromosome are consistent with a hypothesized B chromosome origin. We show that the putative W chromosome displays female-associated DNA sequence, gene expression, and chromatin accessibility to confirm the sex-linked function of the W sequence. In contrast with expectations from studies of homologous sex chromosomes, highly repetitive DNA content on the W chromosome, the sole presence of domesticated repetitive elements in functional DNA, and lack of sequence homology with the Z chromosome or autosomes is most consistent with a B chromosome origin for the W, although it remains challenging to rule out extensive sequence divergence. Synteny analysis of the *D. iulia* W chromosome with other female lepidopteran genome assemblies shows no homology between W chromosomes and suggests multiple, independent origins of the W chromosome from a B chromosome likely occurred in butterflies.

Key words: sex chromosome evolution, butterfly genome, B chromosome.

Significance

Assembled sex chromosomes are important for studies on the evolution of sex, sex-linked traits, and sex-biased gene expression. Yet most of our knowledge of sex chromosome pairs comes from a few model organisms. ZZ/ZW sex determination, where females carry the heterogametic sex chromosome, is common outside of these models. The *Dryas iulia* genome provides an important case study on the origin and evolution of ZZ/ZW sex determination systems from a ZZ/Z0 ancestral state. Our findings support a B chromosome origin for the W chromosome and provide evidence that W chromosomes may have originated from B chromosomes multiple times in butterflies. These results indicate that sex determination mechanisms and sex chromosomes may turnover rapidly when sex chromosome pairs have nonhomologous origins.

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

Introduction

Lepidopteran sex chromosomes, and particularly Z chromosomes, are increasingly being used to gain novel insights into sex-chromosome evolution (Van Belleghem et al. 2018), mechanisms and systems of dosage compensation (Walters et al. 2015; Gu et al. 2019), and the role of neosex chromosomes (Gu et al. 2019; Martin et al. 2020). Few genome assemblies, however, have identified and validated clear candidates for the W chromosome in Lepidoptera and even fewer have been proposed for butterflies. Despite the overall lack of well-assembled W chromosomes, study of the Z and W chromosomes in moths has proposed a noncanonical origin for lepidopteran W chromosomes from a B chromosome (Dalíková et al. 2017; Fraïsse et al. 2017), which are generally nonessential chromosomes found variably within populations and species. Phylogenetic analysis of sex bodies in Trichoptera and Lepidoptera indicate that W chromosomes may have been absent from lepidopteran ancestors, whereas the Z chromosome is deeply conserved. Thus, the Z chromosome appears to have a single origin and W sequences are unlikely to be homologous to the Z (Lukhtanov 2000). The potential gain of a W chromosome from a B chromosome is a surprising contrast to the more commonly observed shared evolutionary origin with the homogametic sex chromosome (Furman et al. 2020). Evidence for a B chromosome origin of a lepidopteran W chromosome remains scarce, however, and additional study of DNA sequence and functional elements, such as W genes and cis-regulatory loci, is needed to test this hypothesis and determine the extent to which B chromosomes may drive W chromosome evolution in butterflies and moths.

High-quality genome assemblies characterized by completeness, contiguity, and well-ordered construction are essential for the study of sex chromosome evolution, given the technical challenges of sex chromosome biology. Sex chromosomes are often repetitive (Abe et al. 2005; Hobza et al. 2015), may display a high degree of copy number variation (Bachtrog 2013; Massaia and Xue 2017; Lucotte et al. 2018), and can be substantially degenerate relative to autosomes (Bachtrog 2013; Furman et al. 2020; Lenormand et al. 2020). There is, therefore, a critical need for additional high-quality lepidopteran genome assemblies to understand the processes of sex chromosome evolution in butterflies and moths. *Heliconius* butterflies have emerged as a key study system in lepidopteran genomics. Within the butterfly tribe Heliconiini, which includes the mimetic genera *Heliconius* and *Eueides*, reference quality assemblies exist for several *Heliconius* species (Davey et al. 2016; Van Belleghem et al. 2017). Although a number of additional *Heliconius* genomes have been assembled (Edelman et al. 2019; Massardo et al. 2020), there are currently few high-quality resources for members of the heliconiine genera that split at the base of the phylogeny and, to date, only the Z chromosome has been assembled, with no evidence of a W chromosome in any assembly. Multiple ecologically significant traits in Heliconiini butterflies show evidence of a sex-linked genetic architecture (Haag et al. 1993; Martin et al. 2014; Li et al. 2017; Lewis et al. 2020;). The lack of complete, high-quality assemblies, however, limits the ability to test macroevolutionary hypotheses on the origin and evolution of diversity within heliconiines, despite a continued interest in phylogenetic (Edelman et al. 2019), developmental (Martin et al. 2014; Lewis et al. 2019; 2020;), and population genetic (Supple et al. 2015; Martin et al. 2016) studies in this group.

To increase the breadth of phylogenetic coverage for heliconiine genomes, we sequenced and assembled a chromosome-level genome assembly for the butterfly *Dryas iulia*, which represents the sister clade to the remainder of the Heliconiini group (Kozak et al. 2015). *Dryas iulia* is a common member of the orange mimicry ring of heliconiine butterflies found throughout much of the neotropics and diverged from the remainder of the Heliconiini tribe approximately 26 Ma (fig. 1*A*) (Kozak et al. 2015). Here we identify and validate the Z and W sex chromosomes in the *D. iulia* genome assembly, then assess the potential role of B chromosomes as a source for the W chromosome in butterflies.

Materials and Methods

See supplementary methods, Supplementary Material online for detailed information on library preparation, genome assembly, genome annotation, and data analysis. Abbreviated methods are as follows:

Butterfly Stocks and Tissue Samples

Tissue samples for PacBio, Hi-C, RNA-seq, and PRO-seq assays were collected from a *D. iulia* laboratory colony derived from Costa Rican *D. iulia* butterflies. Larval individuals were reared in greenhouses on *Passiflora biflora* at approximately 28 °C. Adults were kept in cages at 28 °C with *Lantana camara* flowers for nectar and *P. biflora* for oviposition. Pupal wing tissues were dissected in cold PBS before following assay-specific protocols. Sample sex for sex-specific assays was determined as described in Beebe et al. (1960). All butterflies were reared following institutional guidelines.

Library Preparation

Library preparations for PacBio (Lewis et al. 2016), Hi-C (Lewis et al. 2019), and ATAC-seq (Lewis and Reed 2018) libraries were performed as described previously. RNA-seq libraries were prepared using the NEB Ultra II Directional RNA library preparation kit. PRO-seq libraries were performed following a low-cost protocol with minor modifications for insect cells (Chu et al. 2018). Libraries were sequenced at either 2 × 150 bp, 1 × 75 bp (PRO-seq), or 2 × 37 bp (ATAC-seq) on a HiSeq400 or NextSeq500 (ATAC-seq and PRO-seq).

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

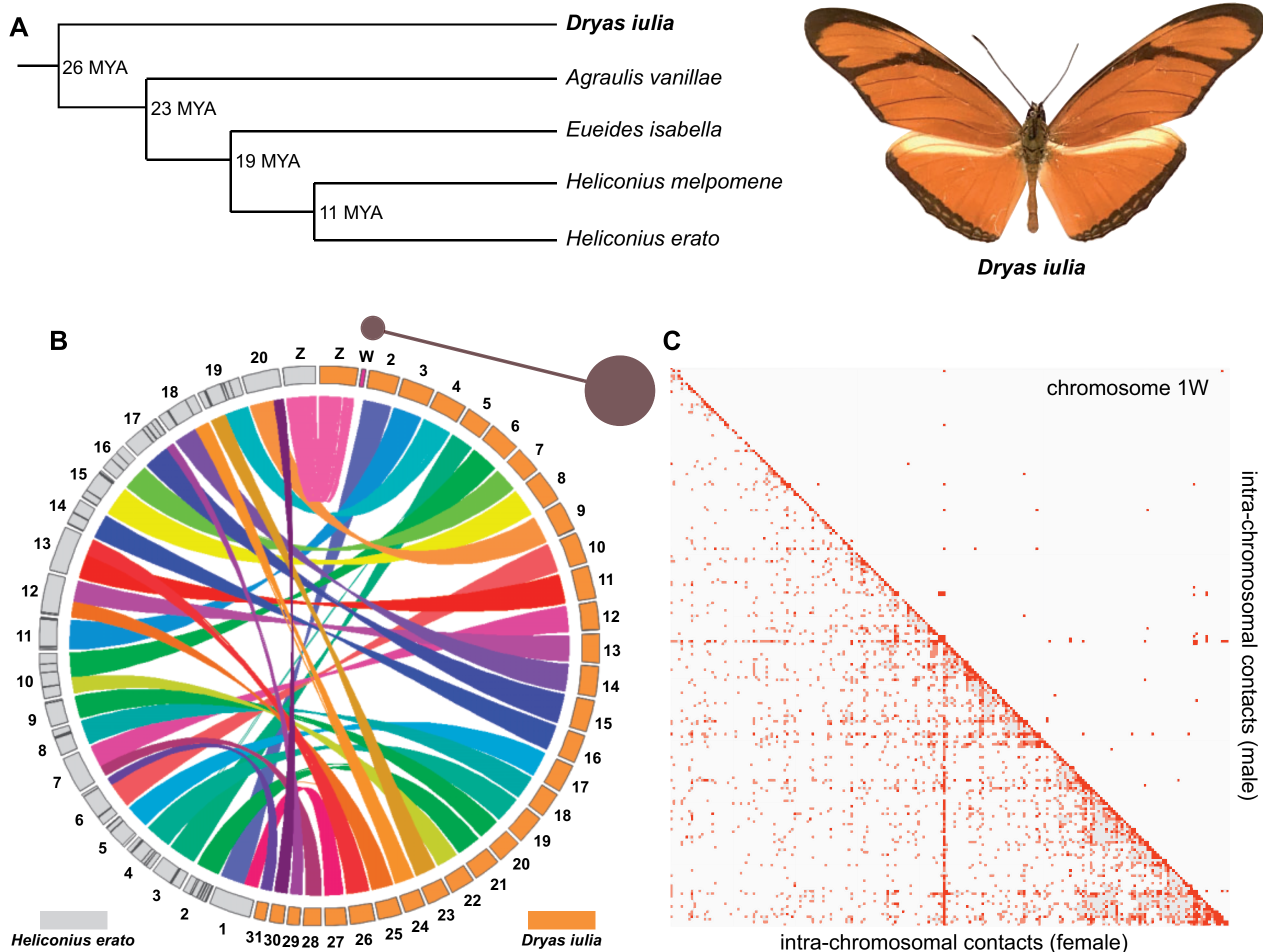


Fig. 1.—A putative W chromosome in the chromosome-level *D. iulia* genome assembly. (*A*) A phylogeny of Heliconiini butterflies shows *Dryas* as the sister taxa to the remainder of the Heliconiini genera (phylogeny and divergences times from Kozak et al. [2015]). (*B*) Synteny and chromosome composition of the *D. iulia* genome assembly shows the *D. iulia* assembly is composed of highly contiguous, ordered, and oriented chromosomes that map back to the expected *H. erato* chromosomes. (*C*) The putative W chromosome is validated by a substantial degree of female-specific chromatin contacts in sex-specific Hi-C data.

Genome Assembly and Annotation

Approximately 44× PacBio Sequel2 long reads from female *D. iulia* DNA were used for the initial genome assembly with CANU (Koren et al. 2017). The resulting contigs were cleaned and polished, then scaffolded using two 10× linked reads libraries. Synteny maps were used to delete putative duplicated regions from the scaffolded sequences. Chromosomes were assembled using Hi-C data aligned to the final scaffold sequences with the 3D-DNA pipeline (Dudchenko et al. 2017). The resulting chromosome sequences were manually curated using the Hi-C alignment maps to produce the frozen genome assembly. Sixteen strand-specific RNA-seq libraries were compiled for gene annotations. Gene predictions were then produced using RNA-seq alignments and protein predictions from multiple lepidopteran and dipteran genomes with BRAKER (Hoff et al. 2019). Assessment of completeness was performed with BUSCO (Seppey et al. 2019).

DNA Sequence Analysis

In addition to comparison of male- and female-specific Hi-C data sets, we also used short-read data to detect female-associated sequences. Two male and two female whole-genome DNA sequence data sets from adult *D. iulia* individuals collected in Key Largo, FL were aligned to the *D. iulia* genome assembly to test for female-associated DNA sequences on the W chromosome. Female-specific DNA sequences were identified as regions of the W chromosome for which female, but not male, reads mapped uniquely to the W sequence. To determine female bias in read alignments, the W chromosome was segmented into 100 bp windows and read depth for each window was calculated. Outlier windows for

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

female–male and female–female DNA alignments were classified using windows at or exceeding the fourth standard deviation of the studentized residuals from the least-squares regression of the male–male sample comparison, using the male–male regression equation. More simply, purple points in this analysis identify windows with <0.05% chance of occurring at random in the male–male sample comparison. Analysis of repetitive element content was performed with RepeatMasker (Tarailo-Graovac and Chen 2009) and the previously produced *D. iulia* repeat library (Ray et al. 2019).

W Chromosome Functional Annotations

Short-read alignments for ATAC-seq and PRO-seq data sets were performed as described previously (Chu et al. 2018; Lewis and Reed 2018). Uniquely mapped read files were then converted to BigWig format and functional elements on the W chromosome were manually curated. Paralogs to manually annotated genes were determined using BLAST at LepBase (Challi et al. 2016).

W Chromosome Synteny Analysis

Synteny analysis within the *D. iulia* genome assembly and between species was performed using Satsuma2 Synteny (Grabherr et al. 2010). A permissive synteny analysis was performed by testing for sequence similarity between only the W chromosome and the entirety of the target genome assembly, using Satsuma2 Synteny and considering a minimum synteny score of 0.7. This analysis will report secondary or tertiary hits from the W chromosome to the target genome assembly at the cost of a higher false-positive synteny rate. To further increase the sensitivity of our analyses, we used BLAST (blastn) to search for Wct transcript sequences in the *D. iulia* and *Kallima inachus* genome assemblies.

Results and Discussion

Assembly and Annotation of the *D. iulia* Genome

To assemble the *D. iulia* genome, we generated ∼44× coverage PacBio long-read sequences from a single female individual, two 10× linked-read data sets from separate individuals, and chromatin conformation capture libraries (Hi-C). The assembled genome had a total length of 440 Mb in 762 scaffolds. Scaffold N50 and N95 values were 15.23 Mb and 7.4 Mb, respectively, with 98.5% of the assembly ordered and oriented into single chromosomes (table 1). We then annotated the *D. iulia* assembly using 16 head and wing tissue mRNA-seq libraries from four developmental stages of larval and pupal development along with protein predictions from *Heliconius melpomene* and *Drosophila melanogaster*. We annotated approximately 18,000 coding genes with about 36,000 transcripts, which is similar to previous lepidopteran genomes (Heliconius Genome Consortium 2012; Davey et al. 2016; Lewis et al. 2016; van der Burg et al. 2019).

To verify the *D. iulia* genome assembly quality, we observed Hi-C signal within versus between assembled chromosomes, which showed clean clustering of chromosomes with greatly reduced signal between chromosomes (supplementary fig. S1, Supplementary Material online). Within chromosomes, we observed no signatures of misassembly due to misplacement or misorientation of contigs, whereas 45 large haplotype duplications (off diagonal horizontal lines in Hi-C data) were reduced to a single copy based on our Hi-C signal to produce the final genome assembly. In total, unsupervised assembly of the *D. iulia* scaffolds produced 32 chromosomes corresponding to the expected chromosome number in female individuals (assuming a ZW sex determination system) (Davey et al. 2016).

We next determined synteny between our assembled *D. iulia* chromosomes and chromosomes of *Heliconius erato*, which was assembled into chromosomes using a high-quality linkage map (Van Belleghem et al. 2017). The 30 *D. iulia* autosomal chromosomes and the Z chromosome showed greater contiguity than the *H. erato* genome assembly and a high degree of ordered synteny with the expected chromosomes in *H. erato*, including correctly placing the 10 chromosomal fusion events that occurred between the divergence of *Heliconius* and *Eueides* (15, Cicconardi et al. 2021) (fig. 1*B*). We then used BUSCO (Seppey et al. 2019) analysis of gene ortholog presence and copy number as an orthogonal test of the completeness and single-haplotype composition of the *D. iulia* genome assembly (table 1). Of the 1658 Insecta BUSCOs, 99.3% were found to be complete within our assembly, 98.7% were single copy sequences, and 0.6% were duplicated. An additional two BUSCO genes (0.1%) were present, but fragmented, and only ten BUSCO genes (0.6%) were missing. Taken together, the contiguity, completeness, and single-copy composition of the *D. iulia* genome assembly places it easily on par with the best lepidopteran genome assemblies to date (supplementary table S1, Supplementary Material online).

Identification of a Putative W Chromosome in *D. iulia*

In addition to the expected conserved 30 autosomes and Z chromosome, we also assembled a short 2.2-Mb sequence that displayed no robust evidence of physical connection to any of the 31 assembled chromosomes previously described within heliconiines—a result consistent with this sequence being a 32nd chromosome. Synteny analysis of this sequence showed no homology with any chromosome within either the *H. melpomene* or *H. erato* genome assemblies (fig. 1*B* and supplementary fig. S2, Supplementary Material online). This result, the female sex of our DNA donor, and the small assembled fragment size, which is similar to the expected W chromosome size of *Heliconius* (Jiggins et al. 2005), together

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

Table 1
Assembly and Annotation Statistics for the *Dryas iulia* Genome Assembly

Genome Assembly and Annotation Statistics					
Assembly length	**Scaffolds**	**Contigs**	**Scaffold N50**	**Scaffold N95**	**% Order/oriented**
440.278 Mb	762	927	15.23 Mb	7.4 Mb	98.50%
Complete BUSCOs	**Single copy**	**Duplicated**	**Fragmented**	**Missing**	
99.3% (1,646)	98.7% (1,636)	0.6% (10)	0.1% (2)	0.6% (10)	

suggest that this sequence may represent the *D. iulia* W chromosome. If so, a W chromosome should show a female-specific enrichment in genomic samples. To test the female specificity of the candidate W chromosome, we compared Hi-C data sets produced from male only and female only tissue samples. Alignment rates between male and female samples were very similar and indicated no significant loss of read alignment in female samples. Consistent with our hypothesis that this sequence represented the *D. iulia* female sex chromosome, sex-specific Hi-C alignments indeed showed a female-specific presence of chromatin interactions for the candidate W sequence (fig. 1*C*), leading us to label this the putative W.

High DNA Copy Number and Unique Repeat Content of the W Chromosome

We next sought to understand how the DNA sequence of the W chromosome may differ from the Z sex chromosome and autosomes. Previously assembled heterogametic sex chromosomes have sometimes shown unusual DNA sequence content, such as high copy number sequences (Lohe et al. 1993) and a greater accumulation of transposable element content than found on the autosomes or homogametic sex chromosome (Erlandsson et al. 2000). This atypical DNA sequence evolution is likely due to the combination of reduced effective population size (Van Belleghem et al. 2018) and the lack of recombination in much or all of the heterogametic sex chromosome (Charlesworth 1978; Furman et al. 2020). In butterflies, where recombination appears to be male limited (Jiggins et al. 2005), the Z chromosome should experience crossover events at roughly the same frequency per meiosis as autosomes (Martin et al. 2019), whereas the W should not recombine at all.

To understand the effects this might have on the putative W chromosome in *D. iulia*, we aligned two male and two female whole-genome short-read DNA sequence data sets to the female *D. iulia* genome assembly to test for differences in sequence and DNA copy number between sexes. Consistent with a highly repetitive DNA composition, very little (∼6%) of the W chromosome mapped uniquely in a completely female-specific manner in our short-read analysis. Nonetheless, DNA sequence alignments showed a moderate female-specific bias. Analysis of mean read depth for 100 bp windows along the W chromosome for male and female samples found that a significant fraction of windows were overrepresented in female DNA and a large number of windows were found only in female DNA (fig. 2). Much of this DNA was found to have high copy number, and the high DNA copy number on the putative W is a clear contrast with autosomes and the Z chromosome (fig. 2 and supplementary S3, Supplementary Material online). Realignment of our PacBio reads to the *D. iulia* genome showed a distinct peak of read enrichment for the W chromosome at one-half of the autosomal depth and similar to the Z. This suggests that the W sequence is mostly complete, and the long tail of the W read depth is consistent with a highly repetitive sequence (supplementary fig. S4, Supplementary Material online).

Consistent with evidence for high DNA copy number on the W chromosome, observation of alignment rates in female DNA found that the W chromosome is substantially less uniquely alignable with 74 bp reads than either of the other two chromosome classes (fig. 3*A*). Although nucleotide diversity is expected to decrease in heterogametic sex chromosomes due to their reduced recombination and lower effective population sizes, we would expect any nucleotide and repeat diversity to be specific to a W chromosome lineage, and thus show nonuniform dispersion between lineages. This is indeed what we observed. Pairwise comparisons of our male and female samples indicated that sequence copy number appeared more dispersed for female samples (fig. 2), likely due to the clonal nature of W chromosomes in *D. iulia* lineages and consistent with what we would expect for any nonrecombining heterogametic sex chromosome segment. The increased read depth for female windows suggests a high degree of divergence in simple and complex repeat number between lineages, though the possibility of additional unassembled W sequence would complicate this analysis. These results suggest that a lack of recombination of the W has been a primary driver of nucleotide composition on the *D. iulia* W chromosome.

We next aimed to better understand the sequence composition of the putative W chromosome. We ran RepeatMasker (Tarailo-Graovac and Chen 2009) on the *D. iulia* genome assembly to test whether repetitive element content could explain some of the sequence differences between the W and the Z or autosomal chromosomes. The results revealed a notable shift in relative repetitive element class

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

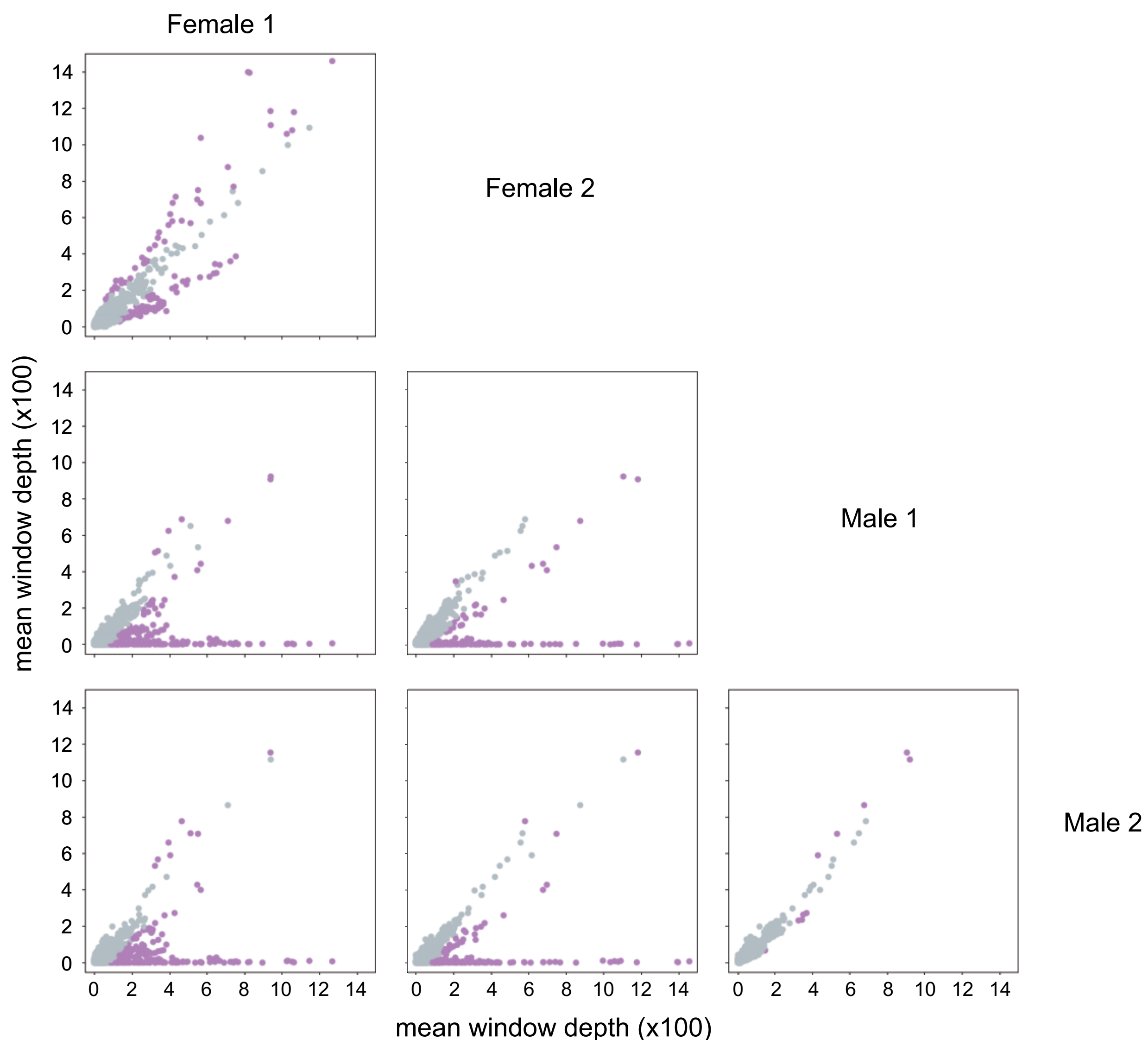


Fig. 2.—Sex-specific DNA sequence alignments show female-associated DNA on the W chromosome. Scatterplots show read depth in 100 bp windows along the W chromosome for two male and two female whole-genome DNA sequence samples. Purple dots represent windows whose studentized residuals fall at or above the fourth standard deviation of the studentized residuals of the male–male plot using the male–male regression equation and estimates of residual variances. The female–female plot shows increased variance compared with the male–male plot and female–male plots show a female-biased association. Both results are consistent with a clonal W chromosome lacking recombination.

frequency, with a strong reduction in shorter repetitive elements (e.g., SINEs) and a substantial increase in the presence of longer repeat classes (e.g., LINEs, LTRs) when compared with the Z and autosomal chromosomes (fig. 3*B* and supplementary S5, Supplementary Material online). Given the shared landscape of active repetitive elements between chromosomes, this suggests that longer repeats are actively removed from the Z and autosomal chromosomes, whereas selection less effectively purges these elements from the W. To test this hypothesis, we analyzed the distribution of repeat element similarity to the consensus sequence for both LINE and LTR elements (fig. 3*C* and *D*). For elements in both classes, W chromosome elements saw a shift in repeat sequence conservation away from the consensus. This could reflect either a greater age for W chromosome repeats or the rapid accumulation of nucleotide polymorphisms due to a reduced frequency of purifying selection on the W (Charlesworth et al. 1993). This suggests 1) that evolution of the W chromosome may have been heavily driven by repetitive elements, and 2) that purifying selection may be an important factor for genome evolution via reduction of long repeat content in autosomes (e.g., Lockton et al. 2008), but less effective in nonrecombining sex chromosomes. A reduced frequency of essential loci, such as genes, on a heterogametic sex

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

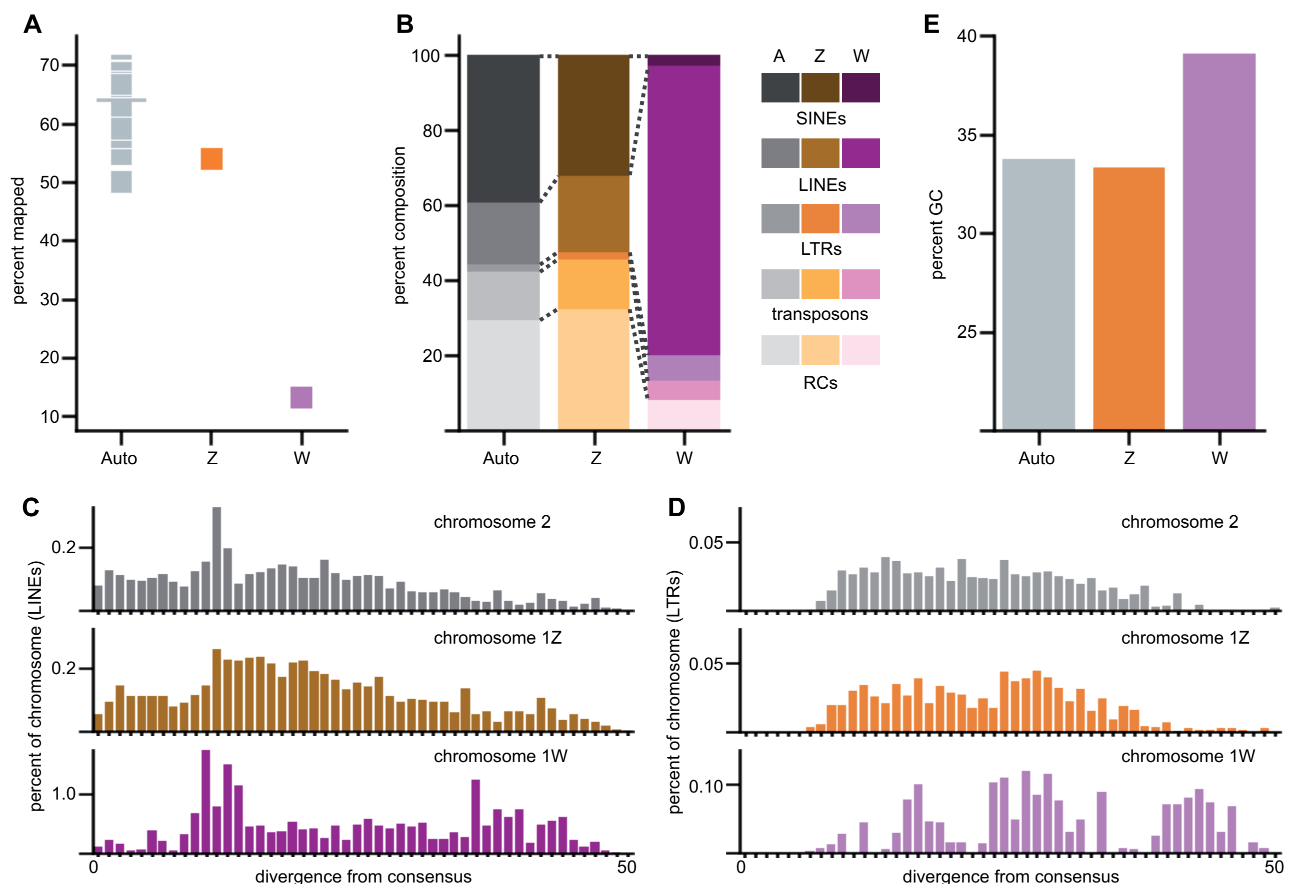


Fig. 3.—Repetitive elements underlie the unusual DNA sequence content of the W chromosome. (*A*) Percent of the genome uniquely mappable with 74 bp sequences shows that the W chromosome is much more repetitive than the autosomal or Z chromosomes. Gray bar marks the mean for autosomal chromosomes. (*B*) Stacked bar charts with the relative composition of repeat classes on the autosomal, Z, and W chromosomes. Longer repeat classes (LINEs and LTRs) are substantially increased on the W chromosome relative to the other chromosomes. LINEs (*C*) and LTRs (*D*) both show an increased divergence from the repeat consensus sequence, consistent with reduced purifying or background selection on the W chromosome. *E*, GC content on the W chromosome is much higher than expected from the autosomes or the Z chromosome.

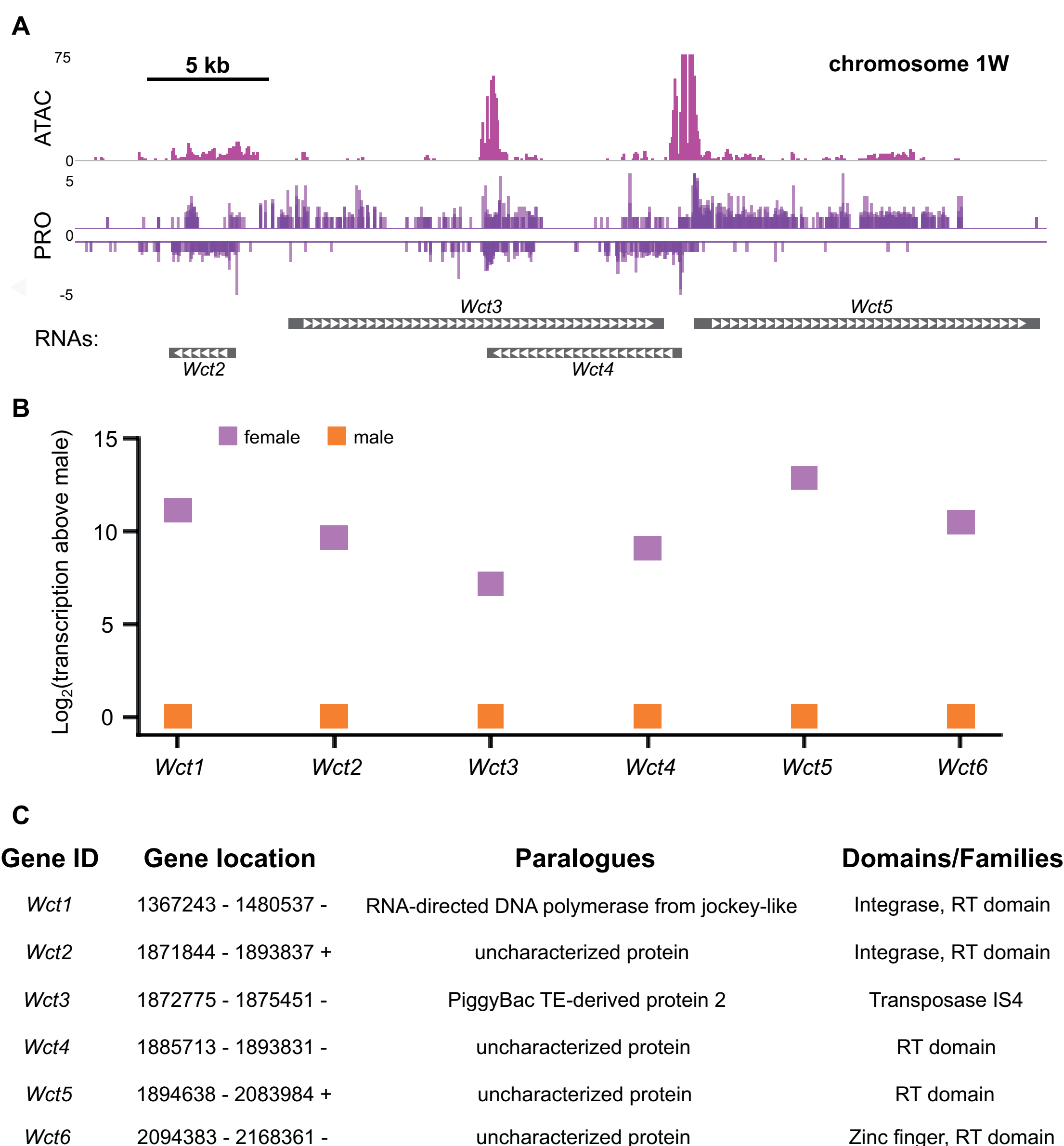
chromosome (see below) could further decrease the effect of selection on these chromosomes as fewer purifying events are likely in this case.

Our observation of substantially increased LINE insertions raises the question: what limits the expansion of the W chromosome sequence length in the absence of recombination and reduced purifying selection? Because there is no recombination on the W chromosome, extant W lineages that increase in sequence length may be unable to reduce this length without loss of the lineage—a process similar to Muller's Rachet. Indeed, W sequence variation within *D. iulia* is consistent with lineage-specific sequence accumulation and some Ws in *D. iulia* may be larger than the reference sequence (fig. 2). One potential mechanism to avoid runaway W chromosome expansion is reduced repetitive element insertion affinity, as previous reports have shown that LINE elements preferentially insert in regions with low GC content (Ovchinnikov et al. 2001). Analysis of GC content found that the *D. iulia* W chromosome sequence is about 15% more GC rich than autosomes or the Z (fig. 3*E*), a result more similar to vertebrate GC content than what is observed in Lepidoptera (Fujita et al. 2011). Although analysis of other lepidopteran W chromosomes will be necessary to validate this result, this suggests that an increased density of GC bases in the W might limit chromosome expansion without the need for selective size-limiting processes.

Repetitive Elements Drive Functional Evolution of the W Chromosome

Our finding that certain repeat classes are preferentially found in the putative W sequence suggests that repetitive elements may also contribute to the functional evolution of the *D. iulia* W chromosome. Most known heterogametic sex chromosomes originate from the same ancestral chromosome as the homogametic sex chromosome, and subsequent

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021



Gene ID	Gene location	Paralogues	Domains/Families
Wct1	1367243 - 1480537 -	RNA-directed DNA polymerase from jockey-like	Integrase, RT domain
Wct2	1871844 - 1893837 +	uncharacterized protein	Integrase, RT domain
Wct3	1872775 - 1875451 -	PiggyBac TE-derived protein 2	Transposase IS4
Wct4	1885713 - 1893831 -	uncharacterized protein	RT domain
Wct5	1894638 - 2083984 +	uncharacterized protein	RT domain
Wct6	2094383 - 2168361 -	uncharacterized protein	Zinc finger, RT domain

Fig. 4.—Functional loci on the W chromosome are derived from repetitive DNA sequences. (*A*) Genome browser screenshot of the W chromosome shows female ATAC-seq and PRO-seq data that mark active cis-regulatory elements and putative gene bodies, respectively. *Wtc2*, *Wtc4*, and *Wct5* show characteristic "pause" peaks of RNA polymerase near the transcription start site. (*B*) Comparison of male and female read alignments from PRO-seq data confirms putative coding regions on the W chromosome are female associated. (*C*) Gene paralogs to the *Wct* sequences of the W chromosome all display evidence of domesticated repetitive elements.

evolution after recombination suppression results in much of the novel functional DNA content of the heterogametic chromosome (Furman et al. 2020). In this case, counterparts to some genes on the homogametic chromosome may remain (gametologs). The proposed noncanonical origin of the lepidopteran W chromosomes from a B chromosome, however, predicts that most of the W chromosome's functional DNA content is gained via novel gene birth rather than modification of tandem gene duplications or gametologues.

Alignments of RNA-seq reads to the W chromosome were difficult to interpret due to the highly repetitive DNA composition. Instead, we used male and female PRO-seq (which assays nascent transcription over the entire gene body) and ATAC-seq (which assays accessible regulatory elements) data sets to identify putative genes and cis-regulatory elements to investigate the functional DNA content of the W chromosome and to test the hypothesis that these sequences represent de novo gain of function evolution. In general, actively

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

transcribed regions on the W chromosome were rare compared with the autosomes or the Z chromosome, consistent with a limited functional role of the W in *Dryas* development (Furman et al. 2020). Manual annotation of the W chromosome identified six regions with pupal wing PRO-seq alignment profiles characteristic of gene bodies, which we labeled *W chromosome transcript 1–6* (fig. 4*A*). Transcript length and sequence mappability suggest these transcripts represent coding or noncoding genes, rather than active transposable elements. Similarly, ATAC-seq profiles showed three loci defined by accessible chromatin, all of which were at the 5′ end of annotated transcripts, suggesting that limited or no use of distal regulatory elements occurs on the W chromosome. PRO-seq data at the 5′ end of several putative gene bodies displayed a characteristic "pause" peak where polymerase elongation is halted after initiation (Adelman and Lis 2012), confirming that some mechanisms that regulate transcription remain constant across the *D. iulia* sex chromosomes and autosomes (e.g., fig. 4*A*, *Wct4*/*Wct5*). Although nonspecific alignment of short-read data sets remains a challenge, all annotated transcripts and putative CREs showed a strong female bias in read count that supports these loci being female-specific genes (fig. 4*B*).

Analysis of sequence homology found that none of the putative W genes were homologous to known proteins in previously annotated Lepidoptera. In contrast, many Z chromosome genes have been found conserved between species. All *Wct*s, however, were potential paralogs to a series of predicted protein coding genes with integrase and reverse transcriptase domains, several of which were confirmed repeat-derived genes (fig. 4*C*). True paralogy via duplication is unlikely in this case, because sequence similarity is likely derived from the active repetitive element landscape shared across chromosomes. Instead, sequence similarity likely reflects the potential role of domesticated repetitive elements in driving gene birth on the W chromosome (Jangam et al. 2017).

A Potential B Chromosome Origin for the *D. iulia* W Sex Chromosome

If the *D. iulia* W chromosome is derived from the same ancestral autosomal pair as the Z chromosome, we would expect to find at least some sequence similarity between Z genes and transcripts on the W (Bachtrog 2013; Furman et al. 2020). Although alignment of our PRO-seq and ATAC-seq data sets was limited by the high copy number and repetitive element content of the W chromosome, our gene set is consistent with a noncanonical B chromosome origin for the *D. iulia* W chromosome. Indeed, we found no evidence of gene homology with genes or nonrepetitive DNA sequences from the Z or autosomal chromosomes, despite a substantial degree of gene expression on the Z chromosome in the same data sets—a result suggesting a nonautosomal origin for the W chromosome sequence (fig. 4C and supplementary fig. S6 and table S2, Supplementary Material online). Repetitive sequence alignments of *D. iulia* W chromosome transcripts to the Z chromosome did not cluster on the Z and no serial sequence homology was observed (supplementary table S2, Supplementary Material online). Multiple lines of evidence, from chromosome size and number to functional DNA content and lack of sequence homology within heliconiines, are therefore most consistent with a B chromosome origin for the *Dryas* W chromosome sequence, albeit providing indirect evidence.

Although a few examples of B chromosome-derived heterogametic sex chromosomes are known outside of Lepidoptera (e.g., Yoshida et al. 2011), it has been commonly found that a pair of autosomal chromosomes differentiate into the two sex chromosomes, as determined by shared sequence similarity between the sex chromosomes. To explicitly test the hypothesis that the *Dryas* W chromosome is derived from a B chromosome, we used synteny analysis to test for sequence homology between the assembled W chromosome and all 31 additional chromosomes within the *D. iulia* genome assembly. Although some similarity was found between the W and the remaining chromosomes due to a shared landscape of active repetitive elements, no chromosome showed substantially more homology to the W than any other. Importantly, this was especially true for the Z chromosome, which was found to have almost exactly the mean degree of sequence similarity with the W chromosome when normalized by target sequence length (fig. 5*A*). This result is consistent with similar findings in *K. inachus* (Yang et al. 2020), *Danaus plexippus* (Mongue et al. 2017), and *Spodoptera exigua* (Zhang et al. 2019), which all possess a W chromosome that lacks sequence homology to the Z or autosomes. To our knowledge, cytogenic and PCR-based studies have only shown clear homology between derived or neosex chromosomes, and never to putatively ancestral sex chromosomes in Lepidoptera (e.g., Yoshido et al. 2020).

W chromosomes have, however, been shown to evolve rapidly in Lepidoptera (e.g., Vítková et al. 2007; Yoshido et al. 2013;), and we cannot formally rule out the possibility that the W has simply diverged to the point that ancestral sequence homology has been lost. However, several examples of sex chromosome divergence shows that some homology is expected across the time scales we are considering. For example, gene sequence homology between the X and Y chromosomes of mammals has been maintained for more than 180 Myr (Hughes and Page 2015), and sufficient synteny remains between the autosomally derived segments of two neosex chromosomes in tortricid moths to easily determine sequence homology, even after 75 Myr of divergence (Šíchová et al. 2013), which exceeds divergence estimates between *Dryas* and *Kallima* (Kawahara et al. 2019). Similarly, the lepidopteran Z chromosome is deeply conserved, and the ancestral Z0 sex determination remains present in many clades (Sahara et al. 2012; Dalíková et al. 2017). The ultra-rapid degradation of a Z-homologous W chromosome that originated at the base of the lepidopteran phylogeny and

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

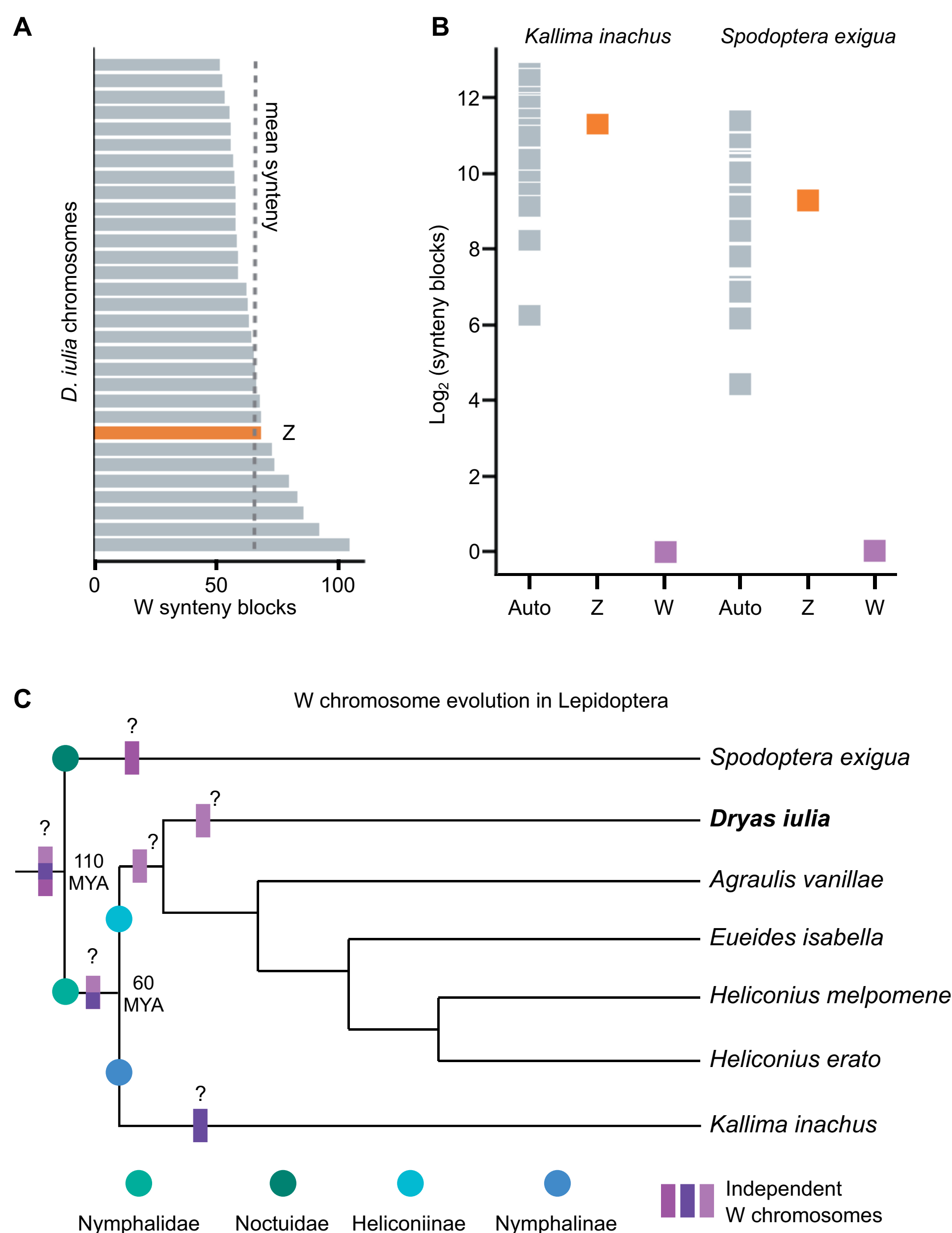


Fig. 5.—Chromosome synteny supports multiple gains of a W chromosome from a B chromosome in butterflies and moths. (*A*) Synteny analysis of the *D. iulia* W chromosome against the remaining 31 chromosomes (autosomes in gray, Z chromosome shown in orange) indicates a lack of shared sequence homology between the W and Z chromosomes. (*B*) Synteny between the *D. iulia* W chromosome and the *K. inachus* and *S. exigua* genome assemblies finds no synteny between the W chromosomes despite a high degree of synteny between the autosomes and Z chromosome. (*C*) A partial phylogeny of Lepidoptera shows our current hypothesis of multiple independent gains of a W chromosome in butterflies and moths. Question marks reflect uncertainty in those branches about the exact time and relationship to more basal clades of W sequence. Divergence times are from Kawahara et al. (2019).

underwent incomplete lineage sorting, or was convergently lost in multiple lineages, remains formally possible. Our findings, however, appear more consistent with an independently evolved W chromosome from a nonhomologous sequence, and therefore support a noncanonical origin of the W chromosome in butterflies as has been previously hypothesized (Fraïsse et al. 2017).

Synteny and Gene Homology Support Multiple Independent Origins of a W Chromosome in Butterflies

Evidence supporting the unusual origin of lepidopteran W chromosomes led us to investigate the degree of sequence turnover between the *D. iulia* W chromosome and those of other lepidopteran genome assemblies. Sex chromosomes, and specifically heterogametic sex chromosomes, are known

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

to evolve more quickly than autosomes due to a variety of population genetic processes including reduced population size, limited recombination, and various forms of selection (Charlesworth et al. 1987). Nonetheless, we would expect some degree of sequence similarity between homologous sex chromosomes despite this increased rate of evolution (Šíchová et al. 2013; Hughes and Page 2015).

To assess W chromosome homology between species, we used synteny analysis to identify the degree of sequence similarity between the W chromosome of *D. iulia* and the genome assembly of the nymphaline butterfly *K. inachus* (Yang et al. 2020)—the only additional high-quality nymphalid genome with a proposed W chromosome, based on synteny and short-read alignments (fig. 5*B* and *C*). Unexpectedly, we found no similarity between the *D. iulia* W chromosome and the *K. inachus* chromosome assembly. In contrast, the autosomes were reasonably homologous between the two assemblies, as was the Z chromosome, which showed a degree of synteny similar to that of the autosomes despite a known increased rate of evolution (fig. 5*B*). A more permissive analysis, where we tested for synteny between only the *D. iulia* W chromosome and the remainder of the *K. inachus* assembly, also failed to identify syntenous sequences between the two W chromosomes (supplementary fig. S5, Supplementary Material online). *Danaus iulia* W transcripts were similarly absent from the *K. inachus* W chromosome when aligned with BLAST (supplementary table S2, Supplementary Material online). Taken together, these results suggest that *D. iulia* and *K. inachus* do not share an ancestral W chromosome.

To provide additional support for the rapid evolution of W chromosomes in nymphalid butterflies, we replicated our synteny analysis with W-associated scaffolds from the *D. plexippus* genome assembly (Mongue et al. 2017), as well as the complete *S. exigua* genome assembly (Zhang et al. 2019). Our additional synteny analyses mirrored our results against *K. inachus*, with highly conserved autosomes and Z chromosome, but no synteny between the *D. iulia* W sequence and *Spodoptera* or *Danaus* W sequences (fig. 5*B* and supplementary fig. S5, Supplementary Material online). No synteny was identified between the *D. iulia* and *S. exigua* W chromosomes in whole-genome comparisons, and permissive analyses indicated that the *D. iulia* W is no more homologous with other nonnymphalid or noctuid W chromosomes than with the autosomes. In contrast, W synteny between two noctuid moths (*Trichoplusia ni* [Fu et al. 2018] and *S. exigua*) indicated that the noctuid W chromosome has been conserved over 60 Myr of evolution (supplementary fig. S5, Supplementary Material online). Again, we cannot formally reject the possibility of rapid W sequence evolution such that W homology cannot be detected within nymphalids as an explanation for this lack of synteny. This scenario would, however, be a stark contrast to results from noctuid and tortricid moths (Šíchová et al. 2013). Thus, despite a common sex determination system, our results suggest an evolutionary scenario where W chromosomes have independently evolved from a B chromosome multiple times in the lepidopteran phylogeny (fig. 5*C* and supplementary S6, Supplementary Material online).

Conclusions

Our findings are most consistent with a B chromosome origin for the *D. iulia* W chromosome, as has been proposed for some moths (Fraïsse et al. 2017). Although conclusive data for a negative claim is unfeasible, we find no evidence that favors the alternative hypothesis—that the Z and W chromosomes share a homologous origin, even when using permissive homology detection methods. We considered two alternative origins for the *D. iulia* W. It remains possible that the *D. iulia* W shows no sequence homology with the Z, or any other chromosome, due to rapid divergence of the W sequence such that DNA sequence similarity cannot be detected. The lack of sequence homology between *D. iulia* sex chromosomes suggests evolution from a shared ancestral chromosome would have to have occurred more than 60–80 Ma. This is the range at which W or neo-W sex chromosomes have been shown to be homologous and is well beyond the period during which the *D. iulia* W likely evolved based on our phylogenetic analyses within nymphalids. It is also possible for a homogametic sex chromosome to gain heterogametic function (Meisel et al. 2017). We would expect in this case that an Z origin for the W sequence would be common to most nymphalid butterflies, yet we see no evidence of W and Z sequence homology within or between species. Evidence for a B chromosome origin for the *D. iulia* W chromosome certainly raises questions about the mechanisms and processes that gave rise to the W. Yet despite the incomplete evolutionary history of the *D. iulia* W, our evidence matches many of the expectations of a B chromosome-derived W sequence and, importantly, none of those for an alternative origin.

Synteny analysis between species also suggests that W chromosomes may have independently evolved multiple times through co-option of a B chromosome in Lepidoptera. This raises the question, how rapidly do W chromosomes turnover in butterflies? Sequence comparison of the *D. iulia* W chromosome against multiple *Heliconius* genomes (Edelman et al. 2019) with BLAST found no evidence of substantial sequence homology within other heliconiines. Although this could be due to missing W chromosomes in these genome assemblies, our results also raise the possibility that sex determination is not uniform within the Heliconiini tribe. In addition to the ZZ/ZW sex determination system, a ZZ/Z0 system where dosage of a Z allele determines sex has been proposed for multiple lepidopteran species and this could be the case in some Heliconiini species as well (Traut et al. 2007). The results here could provide an important set of functional elements to test for sequence homology of the W

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

chromosome in other heliconiines to address this question in the future.

Much of our knowledge of sex determination within moths and butterflies is dependent on only a few case studies, and more high-quality assemblies will be needed to accurately estimate the frequency of sex determination mechanisms and sex chromosome turnover in Lepidoptera. Yet despite limited genome assemblies for comparison, our results suggest that entirely new sex chromosomes may evolve more often or rapidly than previously thought—a result at odds with many vertebrate models, though sex determination systems have been known to evolve rapidly in some fish species (Gammerdinger and Kocher 2018; Clark and Kocher 2019; Darolti et al. 2019). If this is supported in future studies, it will ultimately become important to also determine the precise mechanisms of sex determination and how these mechanisms have convergently evolved.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank the Mank Lab and Deborah Charlesworth for their helpful comments on this manuscript. F.C. and S.H.M. are grateful to the High-Performance Computing team at the Advanced Computing Research Centre, University of Bristol for support. We thank two anonymous reviewers for their helpful comments. This research was supported by NASA 17-EXO-17-2-0112 to C.G.D., NSF IOS-1656514 and DEB-1546049 to R.D.R., and a NERC IRF (NE/N014936/1) and ERC Starter Grant (758508) to S. Montgomery. S. Martin was supported by the Royal Society URF\R1\180682.

Author Contributions

J.J.L., F.C., and S.H. Montgomery designed the study; J.J.L., F.C., and S.H. Martin performed research; R.D.R, C.G.D, and S.H. Montgomery provided resources and reagents; J.J.L., F.C., S.H. Martin, and S.H. Montgomery wrote the manuscript.

Data Availability

Genome assembly, genome annotations files, as well as data sets for this study are available at NCBI BioProjects PRJNA691346, PRJNA691348, and PRJNA732066 and on LepBase.

Literature Cited

Abe H, Mita K, Yasukochi Y, Oshiki T, Shimada T. 2005. Retrotransposable elements on the W chromosome of the silkworm, *Bombyx mori*. Cytogenet Genome Res. 110(1–4):144–151.

Adelman K, Lis JT. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet. 13(10):720–731.

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat Rev Genet. 14(2):113–124.

Beebe W, Jocelyn C, Henry F. 1960. A comparison of eggs, larvae and pupae in fourteen species of heliconiine butterflies from Trinidad, W.I. Zoologica 45(9):111–154.

Challi RJ, et al. 2016. Lepbase: the lepidopteran genome database. BioRxiv 056994. 10.1101/056994.

Charlesworth B. 1978. Model for evolution of Y chromosomes and dosage compensation. Proc Natl Acad Sci U S A. 75(11):5618–5622.

Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. Am Nat. 130(1):113–146.

Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134(4):1289–1303.

Chu T, et al. 2018. Chromatin run-on and sequencing maps the transcriptional regulatory landscape of glioblastoma multiforme. Nat Genet. 50(11):1553–1564.

Cicconardi F, et al. 2021. Chromosome fusion affects genetic diversity and evolutionary turnover of functional loci, but consistently depends on chromosome size. bioRxiv. 2021.01.06.425547.

Clark FE, Kocher TD. 2019. Changing sex for selfish gain: B chromosomes of Lake Malawi cichlid fish. Sci Rep. 9(1):20213.

Daliková Martina M, Zrzavá, et al. 2017. New insights into the evolution of the W chromosome in Lepidoptera. J Hered. 108(7):709–719.

Darolti I, et al. 2019. Extreme heterogeneity in sex chromosome differentiation and dosage compensation in livebearers. Proc Natl Acad Sci U S A. 116(38):19031–19036.

Davey JW, et al. 2016. Major improvements to the *Heliconius melpomene* genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution. G3 6(3):695–708.

Dudchenko O, et al. 2017. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science 356(6333):92–95.

Edelman NB, et al. 2019. Genomic architecture and introgression shape a butterfly radiation. Science 366(6465):594–599.

Erlandsson R, Wilson JF, Pääbo S. 2000. Sex chromosomal transposable element accumulation and male-driven substitutional evolution in humans. Mol Biol Evol. 17(5):804–812.

Fraïsse C, Picard MAL, Vicoso B. 2017. The deep conservation of the Lepidoptera Z chromosome suggests a non-canonical origin of the W. Nat Commun. 8(1):1486.

Fu, YY, et al. 2018. The genome of the Hi5 Germ Cell Line from *Trichoplusia ni*, an agricultural pest and novel model for small RNA biology. ELife 7: e31628.

Fujita MK, Edwards SV, Ponting CP. 2011. The *Anolis* lizard genome: an amniote genome without isochores. Genome Biol Evol. 3:974–984.

Furman BLS, et al. 2020. Sex chromosome evolution: so many exceptions to the rules. Genome Biol Evol. 12(6):750–763.

Gammerdinger WJ, Kocher TD. 2018. Unusual diversity of sex chromosomes in African cichlid fishes. Genes 9(10):480.

Grabherr MG, et al. 2010. Genome-wide synteny through highly sensitive sequence alignment: satsuma. Bioinformatics 26(9):1145–1151.

Gu L, et al. 2019. Dichotomy of dosage compensation along the neo Z chromosome of the monarch butterfly. Curr Biol. 29(23):4071–4077.e3.

Haag KL, de Araújo AM, Zaha A. 1993. Genetic structure of natural populations of *Dryas iulia* (Lepidoptera: Nymphalidae) revealed by enzyme polymorphism and mitochondrial DNA (mtDNA) restriction fragment length polymorphism (RFLP). Biochem Genet. 31(9–10):449–460.

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021

Heliconius Genome Consortium. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487(7405):94–98.

Hobza R, et al. 2015. Impact of repetitive DNA on sex chromosome evolution in plants. Chromosome Res. 23(3):561–570.

Hoff KJ, Alexandre L, Mark B, Mario S. 2019. Whole-genome annotation with BRAKER. Methods Mol Biol. 1962:65–95.

Hughes JF, Page DC. 2015. The biology and evolution of mammalian Y chromosomes. Annu Rev Genet. 49(1):507–527.

Jangam D, Feschotte C, Betrán E. 2017. Transposable element domestication as an adaptation to evolutionary conflicts. Trends Genet. 33(11):817–831.

Jiggins CD, et al. 2005. A genetic linkage map of the mimetic butterfly *Heliconius melpomene*. Genetics 171(2):557–570.

Kawahara AY, et al. 2019. Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. Proc Natl Acad Sci U S A. 116(45):22657–22663.

Koren S, et al. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27(5):722–736.

Kozak KM, et al. 2015. Multilocus species trees show the recent adaptive radiation of the mimetic *Heliconius* butterflies. Syst Biol. 64(3):505–524.

Lenormand T, Fyon F, Sun E, Roze D. 2020. Sex chromosome degeneration by regulatory evolution. Curr Biol. 30(15):3001–3006.e5.

Lewis JJ, et al. 2019. Parallel evolution of ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry. Proc Natl Acad Sci U S A. 116(48):24174–24183.

Lewis JJ, Reed RD. 2018. Genome-wide regulatory adaptation shapes population-level genomic landscapes in *Heliconius*. Mol Biol Evol. 36 (1):159–173.

Lewis JJ, Van Belleghem SM, Papa R, Danko CG, Reed RD. 2020. Many functionally connected loci foster adaptive diversification along a neotropical hybrid zone. Sci Adv. 6(39):eabb8617.

Lewis JJ, van der Burg KRL, Mazo-Vargas A, Reed RD. 2016. Chip-seq-annotated *Heliconius erato* genome highlights patterns of cis-regulatory evolution in Lepidoptera. Cell Rep. 16(11):2855–2863.

Li C, et al. 2017. Role of visual and olfactory cues in sex recognition in butterfly *Cethosia cyane cyane*. Sci Rep. 7(1):5033.

Lockton S, Ross-Ibarra J, Gaut BS. 2008. Demography and weak selection drive patterns of transposable element diversity in natural populations of *Arabidopsis lyrata*. Proc Natl Acad Sci U S A. 105(37):13965–13970.

Lohe AR, Hilliker AJ, Roberts PA. 1993. Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. Genetics 134(4):1149–1174.

Lucotte EA, et al. 2018. Dynamic copy number evolution of X- and Y-linked ampliconic genes in human populations. Genetics 209(3):907–920.

Lukhtanov VA. 2000. Sex chromatin and sex chromosome systems in nonditrysian Lepidoptera (Insecta). J Zool Syst. 38(2):73–79.

Martin A, et al. 2014. Multiple recent co-options of optix associated with novel traits in adaptive butterfly wing radiations. Evodevo 5(1):7.

Martin SH, et al. 2016. Natural selection and genetic diversity in the butterfly *Heliconius melpomene*. Genetics 203(1):525–541.

Martin SH, et al. 2020. Whole-chromosome hitchhiking driven by a male-killing endosymbiont. PLoS Biol. 18(2):e3000610.

Martin SH, Davey JW, Salazar C, Jiggins CD. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biol. 17(2):e2006288.

Massaia A, Xue Y. 2017. Human Y chromosome copy number variation in the next generation sequencing era and beyond. Hum Genet. 136(5):591–603.

Massardo D, et al. 2020. The roles of hybridization and habitat fragmentation in the evolution of Brazil's enigmatic longwing butterflies, *Heliconius nattereri* and *H. hermathena*. BMC Biol. 18(1):84.

Meisel RP, Gonzales CA, Luu H. 2017. The house fly Y chromosome is young and minimally differentiated from its ancient X chromosome partner. Genome Res. 27(8):1417–1426.

Mongue AJ, Nguyen P, Voleníková A, Walters JR. 2017. Neo-sex chromosomes in the monarch butterfly, *Danaus plexippus*. G3 7(10):3281–3294.

Ovchinnikov I, Troxel AB, Swergold GD. 2001. Genomic characterization of recent human LINE-1 insertions: evidence supporting random insertion. Genome Res. 11(12):2050–2058.

Ray DA, et al. 2019. Simultaneous TE analysis of 19 heliconiine butterflies yields novel insights into rapid TE-based genome diversification and multiple sine births and deaths. Genome Biol Evol. 11(8):2162–2177.

Sahara K, Yoshido A, Traut W. 2012. Sex chromosome evolution in moths and butterflies. Chromosome Res. 20(1):83–94.

Seppey M, Manni, M, Zdobnov EM, 2019. BUSCO: assessing genome assembly and annotation completeness. In: Kollmar M., editor. Gene Prediction. Methods in Molecular Biology. New York (NY): Springer New York. p. 227–245. 10.1007/978-1-4939-9173-0_14.

Šíchová J, Nguyen P, Dalíková M, Marec F. 2013. Chromosomal evolution in tortricid moths: conserved karyotypes with diverged features. PLoS One. 8(5):e64520.

Supple MA, Papa R, Hines HM, McMillan WO, Counterman BA. 2015. Divergence with gene flow across a speciation continuum of *Heliconius* butterflies. BMC Evol Biol. 15:204.

Tarailo-Graovac M, Chen N. 2009. Using repeatmasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics. 25(1):4.10.1–4.10.14.

Traut W, Sahara K, Marec F. 2007. Sex chromosomes and sex determination in Lepidoptera. Sex Dev. 1(6):332–346.

Van Belleghem SMV, et al. 2017. Complex modular architecture around a simple toolkit of wing pattern genes. Nat Ecol Evol. 1(3):0052.

Van Belleghem SM, et al. 2018. Patterns of Z chromosome divergence among *Heliconius* species highlight the importance of historical demography. Mol Ecol. 27(19):3852–3872.

van der Burg KRL, et al. 2019. Contrasting roles of transcription factors Spineless and EcR in the highly dynamic chromatin landscape of butterfly wing metamorphosis. Cell Rep. 27(4):1027–1038.e3.

Vítková M, Fuková I, Kubícková S, Marec F. 2007. Molecular divergence of the W chromosomes in pyralid moths (Lepidoptera). Chromosome Res. 15(7):917–930.

Walters JR, Hardcastle TJ, Jiggins CD. 2015. Sex chromosome dosage compensation in *Heliconius* butterflies: global yet still incomplete? Genome Biol Evol. 7(9):2545–2559.

Yang J, et al. 2020. Chromosome-level reference genome assembly and gene editing of the dead-leaf butterfly *Kallima inachus*. Mol Ecol Resour. 20(4):1080–1092.

Yoshida K, et al. 2011. B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. PLoS Genet. 7(8):e1002203.

Yoshido A, et al. 2020. Evolution of multiple sex-chromosomes associated with dynamic genome reshuffling in Leptidea wood-white butterflies. Heredity 125(3):138–154.

Yoshido A, Síchová J, Kubíčková S, Marec F, Sahara K. 2013. Rapid turnover of the W chromosome in geographical populations of wild silkmoths, *Samia cynthia ssp*. Chromosome Res. 21(2):149–164.

Zhang F, Zhang J, Yang Y, Wu Y. 2019. A chromosome-level genome assembly for the beet armyworm (*Spodoptera exigua*) using pacbio and Hi-C sequencing. BioRxiv 2019.12.26.889121. 10.1101/2019.12.26.889121.

Associate editor: Andrea Betancourt

Downloaded from https://academic.oup.com/gbe/article/13/7/evab128/6297255 by guest on 02 August 2021