Sex chromosome dosage compensation in a sex reversing skink is not influenced by sexual phenotype

Authors: Benjamin J. Hanrahan[1], J King Chang[1], Ashley M. Milton[1], Nicholas C. Lister[1], Duminda S.B. Dissanayake[2], Jillian M. Hammond[3,4], Andre L.M. Reis[3,4,5], Ira W. Deveson[3,4,5], Aurora Ruiz-Herrera[6,7], Hardip R. Patel[8], Jennifer A. Marshall Graves[2,9], Arthur Georges[2], Paul D. Waters[1]*

Affiliations:

[1] Evolution & Ecology Research Centre, School of Biotechnology and Biomolecular Sciences, Faculty of Science, University of New South Wales, Sydney, NSW 2052, Australia.

[2] Institute for Applied Ecology, University of Canberra, Canberra, ACT 2601 Australia.

[3] Genomics Pillar, Garvan Institute of Medical Research, Sydney, NSW, Australia.

[4] Centre for Population Genomics, Garvan Institute of Medical Research and Murdoch Children's Research Institute, Australia.

[5] Faculty of Medicine, University of New South Wales, Sydney, NSW, Australia.

[6] Departament de Biologia Cel·lular, Fisiologia i Immunologia, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, 08193, Spain.

[7] Genome Integrity and Instability Group, Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, 08193, Spain.

[8] National Centre for Indigenous Genomics, Australian National University, ACT 2601, Australia.

[9] Department of Environment and Genetics, La Trobe University, Melbourne, Victoria 3068, Australia.

*corresponding author. Email: p.waters@unsw.edu.au

ABSTRACT

Background

Lizards have sex determination systems that can differ between even closely related species. These include XY and ZW systems, and thermolabile systems where genes and temperature interact to determine sex. The eastern three-lined skink (*Bassiana duperreyi*) has a differentiated XY sex determination system, in which low temperature incubation during development can cause female to male sex reversal, producing XX males. This provides a unique opportunity to investigate how genotype and sexual phenotype affect dosage compensation.

Results

Here, we present a draft genome assembly of the Eastern three-lined skink generated from nanopore sequencing. We also generated transcriptomes from brain and heart tissue of normal adult males and females, along with brain tissue of sex-reversed XX males. We observed partial dosage compensation between XX females and XY males in both brain and heart, with median gene expression from the X in normal males being 0.7 times that of normal females. In brain of sex reversed XX males the median X chromosome output matched that of the normal XX female level, and not that of normal XY males.

Conclusions

Partial dosage compensation in the Eastern three-lined skink is similar to several other species of lizard. However, here for the first time we describe dosage compensation

in a lizard with natural sex reversal, and show that in sex reversed individuals dosage compensation of the X chromosome follows genotypic sex and not phenotypic sex.

Key words: dosage compensation, sex chromosomes, *Bassiana duperreyi*, sex reversal

INTRODUCTION

Sex chromosome dosage compensation in mammals was proposed to have evolved in response to loss of gene function from the mammalian Y chromosome. Gene expression from the single X in males should need to be upregulated to restore ancestral autosomal levels. This X upregulation carried through to females resulting in disproportionately high X expression in XX females (1), and this was countered by X chromosome inactivation (XCI) to silence one X chromosome in the somatic cells of females (2). This classic model of dosage compensation appears to hold true for marsupial mammals, but transcriptional upregulation of the single X appears incomplete in eutherian mammals (3, 4). In monotremes, median gene expression from the X chromosome is increased compared to the autosomes in XY males, as it is for the bird Z in ZW females. However, global X (or Z) transcriptional output is not evenly balanced between the sexes (5).

Mammals and birds have ancient and relatively stable sex chromosomes and sex determination systems. Most mammals have a conserved XY system and birds have a conserved ZW system. In contrast, reptiles display a wide array of sex determination

and sex chromosome systems, even in closely related species (6, 7). These organisms exhibit not only XY and ZW genetic sex determination (GSD) systems, but also temperature-dependent sex determination (TSD), whereby the sex of offspring is determined by the temperature at which the egg is incubated during a thermosensitive window. GSD and TSD systems are often observed in closely related lizard species, and some species have a GSD system that can be overridden by temperature to cause sex reversal (8-11).

Studies of reptile dosage compensation have revealed a variety of non-canonical dosage compensation systems. In snakes, there are reports of partial dosage compensation of the Z chromosome by upregulation of Z-borne genes in females (12, 13). The green anole (*Anolis carolinensis*), a lizard with an XY sex chromosome system, has complete dosage compensation via 2-fold upregulation of the single X in males to match expression with the ancestral autosomal level and balancing expression with the two Xs in females (14), which is mediated by a lncRNA (15). The blacktail brush lizard also has two-fold upregulation of the single X in males (16), but with unknown mechanism. In contrast, the Komodo dragon (*Varanus komodoensis*), which has one of the oldest ZW sex chromosome systems, shows no evidence of dosage compensation of Z-specific genes (17). Partial dosage compensation has been observed in the brown basilisk (*Basiliscus vittatus*) such that the XY males had expression intermediate between full dosage compensation of the two X's in females, and half the expression (18). Incomplete dosage compensation was also observed in Gila monster (19), which has a ZW sex chromosome system.

100 The scincid lizard *Bassiana duperreyi* (Australian eastern three-lined skink) has a
101 differentiated XY sex chromosome system (20) that is thought to predate skink
102 radiation (21). However, the genetic sex determining switch can be overridden by
103 temperature (8, 9, 11). Reduced incubation temperature of eggs during the
104 thermosensitive period results in sex reversed XX males in the adult population and
105 the natural nests of both captive and wild populations (22, 23). This gives rise to three
106 different genotypic and phenotypic sexes: XX females, XY males and sex reversed
107 XX males. Therefore, *B. duperreyi* provides a unique opportunity not only to examine
108 dosage compensation in a reptile with differentiated sex chromosomes, but to also test
109 if dosage compensation in sex reversed XX males follows genotype or sexual
110 phenotype.
111
112 Here we present a chromosome level *B. duperreyi* draft genome assembly from a male
113 individual, in which we have identified both the X and Y chromosomes. We generated
114 mRNA sequence data from brain and heart tissue of genotypically normal XX females
115 and XY males, and brain tissue of XX male individuals to examine the effect of
116 genotype and phenotype on dosage compensation. In brain and heart of normal XY
117 male and XX female individuals, there was partial dosage compensation, such that the
118 single X of XY males was overexpressed. Sex reversed XX males were observed to
119 have the same median X gene expression as XX females, showing for the first time
120 that dosage of the X follows genotype in a sex reversed lizard.
121
122 **RESULTS**
123
124 **Genome assembly summary**

125 Genome length was 1.485 Gbp, assembled into 5,369 scaffolds. The six largest
126 scaffolds (288.7 - 75.5 Mbp) corresponded to autosome macrochromosomes (from 1
127 to 6), with the next six largest scaffolds (54.7 – 18.1 Mbp) representing the
128 microchromosomes (from 7 to 14) (Figure 1A, Table 1). Based on HiC contacts, contig
129 235 was found to contain merged microchromosomes, likely chromosomes 12, 13 and
130 14 (Figure 1A, inset). Including the sex chromosome scaffolds this comprised all the
131 chromosomes in *B. duperreyi* (2n = 30). The chromosome-scale scaffolds comprised
132 98.5% of the total assembled sequence. BUSCO analysis showed 96.0% complete
133 conserved orthologs in the tetrapod database and 96.6% complete conserved
134 orthologs in the vertebrate database (Table 2).

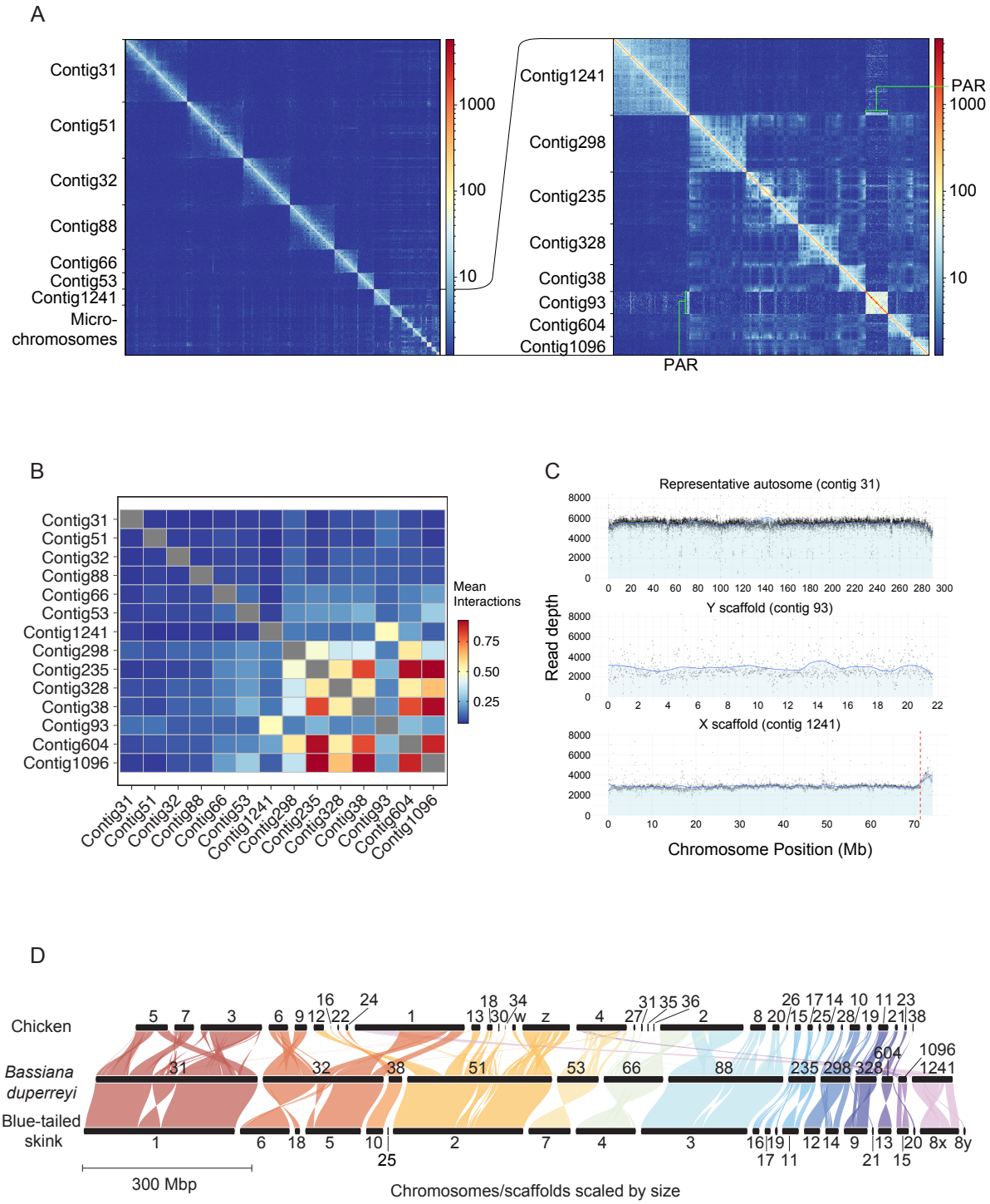

Figure 1. Genome summary of the *Bassiana duperreyi* assembly

137 **A)** Hi-C contact map generated with HiC explorer (53) showing scaffold (chromosome)
138 boundaries. Inset: magnified view of the HiC contact map for microchromosomes with
139 size 18.1 – 55.4 Mb, and the sex chromosomes. Number of contacts are between 100
140 kbp bins. Below: mean inter-chromosomal interactions for 100 kbp windows between
141 each chromosome. **B)** Number of reads mapped in 20kb windows of male Illumina
142 genome sequencing on a representative autosome (contig 31), the putative Y (contig
143 93) and X contigs (contig 1241). Vertical red line is the PAR boundary on the X contig.
144 **C)** Synteny across chicken, bassiana and blue tailed skink (*Cryptoblepharus egeriae*)
145 genomes, generated with GENESPACE (54). The bassiana Y scaffold (contig 93, not
146 shown) shares no synteny with the other two species.
147

148 **Table 1. Chromosome identity based on scaffold size**

Scaffold Name	Size (bp)	Chromosome Identity
BASDU_flyeontv1_Contig31	288676622	1
BASDU_flyeontv1_Contig51	258390987	2
BASDU_flyeontv1_Contig32	215829809	3
BASDU_flyeontv1_Contig88	205722031	4
BASDU_flyeontv1_Contig66	107145552	5
BASDU_flyeontv1_Contig53	75523483	6
BASDU_flyeontv1_Contig1241	74033703	X
BASDU_flyeontv1_Contig298	54678650	7
BASDU_flyeontv1_Contig235	50055479	12,13,14
BASDU_flyeontv1_Contig328	39280427	8
BASDU_flyeontv1_Contig38	25948193	9
BASDU_flyeontv1_Contig604	22020833	10
BASDU_flyeontv1_Contig93	21659042	Y

BASDU_flyeontv1_Contig1096	18146980	11

149

150 **Table 2. BUSCO assessment of the *B. duperreyi* genome**

BUSCO Assessment	Total	Complete and single-copy	Complete and duplicated	Fragmented	Missing	Groups searched
tetrapod_odb10	5156	5098 (96.0%)	58 (1.1%)	44 (0.8%)	110 (2.1%)	38
vertebrate_odb 10	3290	3240 (96.6%)	50 (1.5%)	32 (1.0%)	32 (0.9%)	67

151

152 **Sex chromosome summary**

153 Contig 93 was identified as the Y chromosome by BLASTn with male specific *B.*
154 *duperreyi* k-mer contigs (Supplemental Figure 1). Read depth analysis of short read
155 Illumina whole genome sequencing from a male individual confirmed contig 1241 as
156 the X (Figure 1C), and revealed the pseudoautosomal region (PAR) boundary location
157 at which read coverage increased compared to the X and Y specific regions. The X
158 and Y chromosomes were 71.2 Mbp and 21.7 Mbp, respectively, and shared a small
159 PAR of approximately 2.8 Mbp that was assembled onto the X scaffold (Figure 1C).
160 The Y chromosome had more interactions with the macrochromosomes, than the X
161 chromosome (Figure 1B). Fewer interactions of the X with autosomes is similar to
162 eutherian XY systems that also have fewer interactions between the X chromosome
163 and autosomes (24). As expected, microchromosomes interacted with each other
164 more than with the macrochromosomes (25).

165

166 Annotation identified a total of 13,979 genes in the genome. Of these, 667 were X-
167 specific genes, sharing homology with chicken chromosome 1 (Figure 1D). A total of
168 62 genes were annotated on the male specific region of the Y: 47 of these were X/Y
169 shared, with the remaining 15 having no detected X-borne partner. A total of 39 genes
170 were annotated in the PAR.
171
172 There were no XY shared or Y specific genes with a known role in the sex determining
173 pathway of other vertebrates (Supplemental Table 1). The only gene conserved on the
174 Y chromosome of the Christmas Island blue-tailed skink was *Ifit5*, which encodes for
175 an interferon that has not been implicated in sex determining pathways. Gene Onology
176 (GO) analysis of the 15 Y specific genes showed no enrichment of genes involved in
177 male related functions. In fact, these genes were not enriched for any biological
178 function. GO analysis of XY shared genes, as well as all genes on the Y (XY shared
179 and Y specific genes), revealed an enrichment of genes involved in histone H4
180 acetyltransferase activity (Supplemental Tables 2 and 3). In summary, there was no
181 candidate sex determining gene identified.
182
183 The karyotypic difference between the X and Y (20), and the disparity between their
184 gene contents, confirmed their differentiation. This suggested an unbalanced output
185 of X-specific genes in XX females and XY males that might be balanced by a sex
186 chromosome dosage compensation mechanism.
187
188 **Dosage compensation of the X chromosome**
189 In brain and heart, median gene expression from the autosome scaffolds and PAR
190 was equivalent between XY males and XX females (Figure 2A). However, median X

191 chromosome transcriptional output in XY males was reduced compared to XX females
192 (to 70% in brain and 65% in heart: Mood's median test, $p < 0.0001$).
193
194 We next tested if phenotypic sex influenced total transcriptional output from the X
195 versus the autosomes in brain. We measured the ratios of median X specific to
196 autosome gene expression levels within the three genotypic and phenotypic sexes. In
197 XX females and XX males there was no significant difference between X and
198 autosomal expression levels. In XY males, median transcriptional output from the X
199 was 72% that of the autosomal median (permutation test, $p < 0.0001$) (Figure 2B).
200
201 We then compared transcriptional output of the X specific region between the three
202 phenotypic and genotypic sexes. Median gene expression from autosome scaffolds of
203 XX females, XY males and sex reversed XX males were equivalent (Figure 2C),
204 individual autosome scaffolds were also equivalent (Supplemental Figure 2). PAR
205 region showed slightly higher expression in XY males, however this was only
206 significant in the XYm/XXf comparison. As observed for heart and brain above, median
207 gene expression from the X chromosome was significantly lower in normal XY males
208 compared to XX females (down to 69%: Mood's median test, $p < 0.0001$) (Figure 2C).
209 Median gene expression from the X chromosome was also lower in XY males than XX
210 males (down to 71%: Mood's median test, $p < 0.0001$). These ratios are both greater
211 than the 50% expected from a non-dosage compensated X chromosome, implying
212 significant upregulation of the single X in XY males relative to both XX females and
213 XX males (Figure 2C). This upregulation of genes on the X in XY males compared to
214 both XX females and XX males was not region specific on the X (Supplemental Figure
215 3), nor was it restricted to a subset of genes (Supplemental Figure 4).

We also compared expression from the X between sex reversed XX males and normal XX females. Both have two X chromosomes, so we could directly address if sex chromosome dosage compensation strictly follows genotype, or was also influenced by phenotype. We observed that XX males and XX females had equivalent transcriptional output from the X, with the same expression ratio as the autosomes (Mood's median test, $p = 0.886$) (Figure 2C), and that the median output from the X was not different compared to autosomes in either XX males (permutation test, $p = 0.333$) or XX females (permutation test, $p = 0.433$) (Figure 2B).

Having established that sex chromosome dosage compensation followed genotype, we investigated expression of X genes with a Y gametologue. These X genes (n=42) had significantly lower total transcriptional output in XY males than did XX females (Mood's median test, $p < 0.0001$) and XX males (Mood's median test, $p < 0.0001$) (Figure 2C). However, when counts for X and Y gametologues in XY males were summed their expression was equal to that of X gametologues in XX males and XX females (Figure 2D). Median expression of the XY shared genes was found to be higher than total X gene expression in all sexes (Mood's median test: XXf ,$p < 0.001$; XXm, $p < 0.001$; XYm, $p < 0.05$) (Figure 2D).

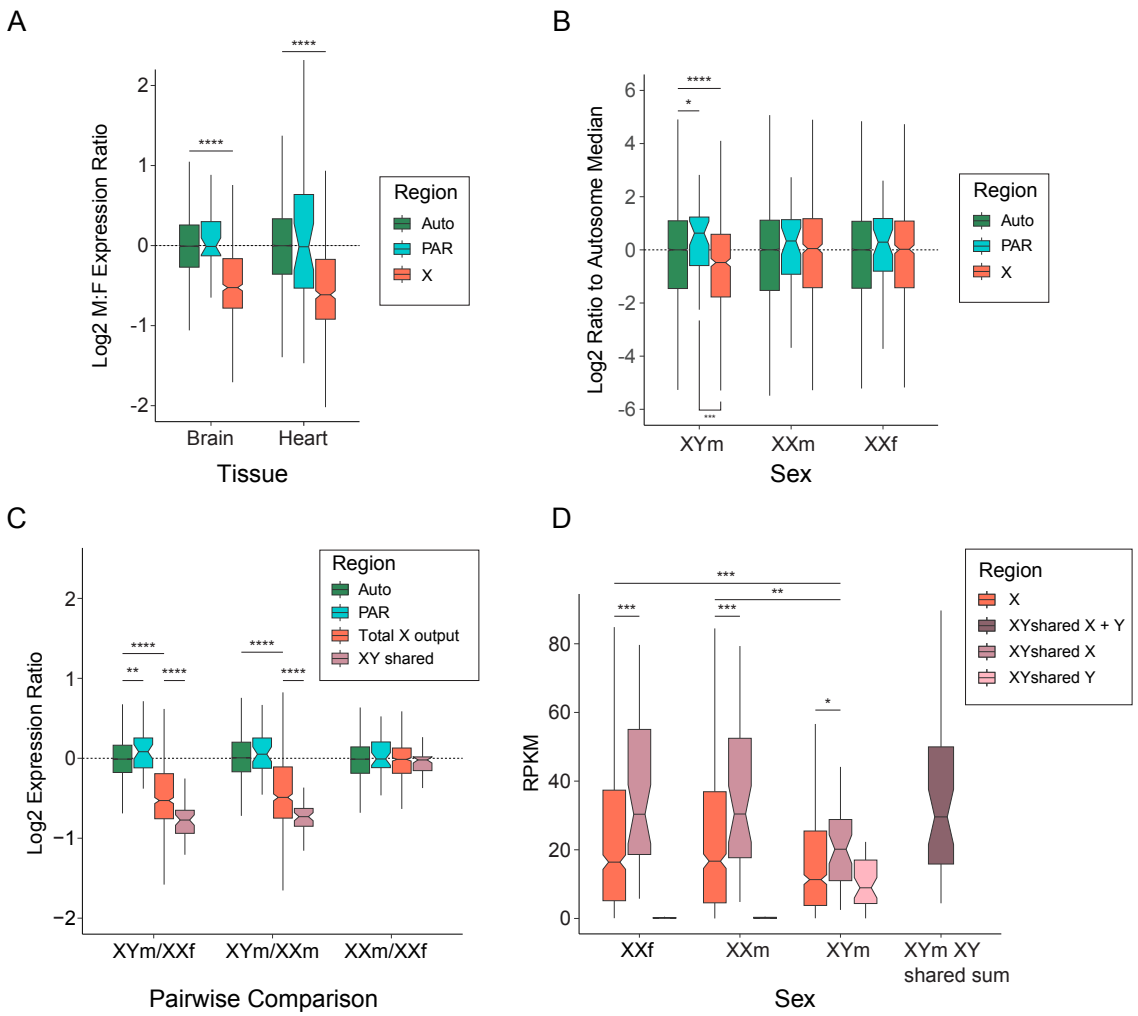

Figure 2. Dosage of gene expression between sexes

A) Median expression male (XY) to female (XX) ratios were calculated for adult brain and heart tissue (n=1 for both sexes and tissues). A ratio above zero indicates higher expression in male and a ratio below zero indicates higher expression in female. Median is plotted in the box with exact value above each median. Boxes represent the middle 50% of the data, and whiskers represent 1.5 times the interquartile range. Outliers not plotted. **B)** Brain transcriptomes were used to calculate mean RPKM for each sex condition (XYm, n=3; XXm, n=2; and XXf, n=3) and plotted for autosomal, X

245 specific and PAR genes as a ratio to autosomal median. Ratio to the autosomal
246 median is plotted in the box with exact value above each median. Permutation tests
247 were used to calculate whether median PAR and X specific RPKMs were statistically
248 different from the median autosome RPKM, as well as if the median PAR RPKM was
249 different from the median X specific RPKM (**** $p<0.0001$, *** $p<0.001$, * $p<0.05$). **C)**
250 Median expression ratios were calculated for pairwise comparisons of the three sexes
251 (XYm, XXm and XXf). Genomic regions were separated and plotted as follows;
252 autosomes (green), PAR (blue), total X output (red) and XY shared genes (pink).
253 Mood's median tests were used to calculate if the ratios for the autosomes and X, as
254 well as Xs between sexes, were statistically different (**** $p<0.0001$). **D)** As above
255 mean RPKM values were calculated for each sex conditions and represented as a
256 boxplot. Genomic regions on the X chromosome were separated as follows; X (red),
257 XY shared Y (light pink), XY shared X (pink) and sum of X and Y ametologues (dark
258 pink).
259

260 **DISCUSSION**

261 Here we present a draft assembly for *B. duperreyi* with chromosome level scaffolds
262 that included both X and Y chromosomes. Gene expression of the X chromosome
263 between XY males and XX females revealed partial dosage compensation of the X in
264 XY males. Dosage of the genes on the X chromosome in sex reversed XX males was
265 equal to the X in XX females, therefore sex dosage compensation followed genotype.
266

267 In normal XY male *B. duperreyi* there was partial dosage compensation of the X
268 chromosome (Figure 2C), with median X chromosome gene expression 69% that of
269 XX females. Therefore, the single X chromosome is being over-expressed by a factor

270 of approximately 1.45. This is similar to observations in non-therian mammal
271 vertebrates such as platypus (5). It is also similar to the dosage system seen in the
272 brown basilisk (18) and Gila monster (19). However, it is different from dosage
273 compensation in other lizard species; the absence of sex chromosome dosage
274 compensation in the ZW system of the Komodo dragon (17), and complete dosage
275 compensation in the green anole (14) and the blacktail brush lizard (16).
276
277 Thus, we demonstrate that *B. duperreyi* represents a partial dosage compensation
278 system in a lizard that is chromosome-wide with no region-specific effect. Thus
279 different dosage compensation systems have evolved repeatedly and independently
280 in different lizard lineages, perhaps depending on the dosage sensitivity of the
281 progenitor autosomes from which the sex chromosome systems evolved. Study of
282 more lizards will elucidate the breadth of reptile sex chromosome dosage
283 compensation and highlight different evolutionary strategies.
284
285 As well as comparing normal XY males and XX females, for the first time, we describe
286 dosage compensation in a model with naturally occurring sex reversal. We observed
287 that gene expression in the sex reversed XX males was equivalent to that in normal
288 XX females, with median X expression 99% that of XX females and 140% that of XY
289 males. Therefore, it is genotype, as opposed to sexual phenotype, that dictates sex
290 chromosome dosage compensation in *B. duperreyi*. This is the first example showing
291 dosage compensation in a sex reversed animal matches that of the sex with the same
292 genotype. It remains to be seen if the offspring of normal XX females and sex reversed
293 XX males also follow this trend of dosage compensation matching normal females, as

294 it is unknown if sex reversed XX males are reproductively viable (22, 26), though it is
295 likely this is the case.
296
297 Temperature-induced sex reversal in lizards is not unique to *B. duperreyi,* having been
298 observed in several species (8, 10). It is possible that sex reversal in lizards is more
299 widespread than previously believed, as it has been systematically investigated in only
300 a few species (27), and there may be numerous other species whose sex reversal
301 status is yet to be confirmed (see 28, 29, 30). Examining dosage compensation in
302 other sex reversing models will confirm whether our finding that phenotype does not
303 influence sex chromosome dosage is ubiquitous.
304
305 Recent evidence suggests that in systems with partial or incomplete dosage
306 compensation in the transcriptome, the abundance of proteins can remain balanced
307 (31). This should be investigated further in organisms such as *B. duperreyi,* brown
308 basilisk, Gila monster and komodo dragon that show partial or lack of dosage
309 compensation in the transcriptome, to observe if the sex chromosome output is
310 balanced in the proteome.
311
312 **CONCLUSIONS**
313 Here we analysed transcriptomes from brain and heart of adult *B. duperreyi* to show
314 that in canonical XY males there is partial dosage compensation of the X chromosome.
315 The brain transcriptome of hatchling brains also revealed that sex reversed XX males
316 have the same transcriptomic output from the X chromosome as canonical XX
317 females. This is the first description of X chromosome dosage in a naturally sex-

318 reversing system and demonstrates that, in *B. duperreyi*, dosage follows the
319 underlying genotype, regardless of an individual's sex.
320
321
322 **MATERIALS AND METHODS**
323
324 **Sample collection and sexing**
325 In December 2020, samples (eggs) of alpine *B. duperreyi* were collected from nests
326 in the field location within the Brindabella Range (Piccadilly Circus – 1240 m a.s.l.,
327 35°21'42.0"S 148°48'12.5"E) after ensuring that approximately 90% of the
328 development period had passed in natural conditions (9). The eggs were then
329 collected and transported to the University of Canberra, where they were buried in
330 moist vermiculite with a ratio of 4 parts water to 5 parts vermiculite by weight, and
331 placed in incubators (LabWit, ZXSDR1090) that maintained 23°C, which produces a
332 balanced sex ratio (9). Details egg collection methods (23) and a description of the
333 alpine study site can be found in (32).
334
335 To determine phenotypic sex of the hatchlings, tail bases of 7-day old hatchlings were
336 squeezed to evert the hemipenes (33) and sex was checked again by hemipene
337 transillumination after 5 weeks (23). Tail snips were collected to determine the
338 genotype of the lizards. Genotypic sex was determined for *B. duperreyi* using
339 polymerase chain reaction (PCR)-based molecular sex tests from extracted DNA
340 collected from tissue samples. DNA purity was determined using a NanoDrop 1000
341 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA) and
342 quantified using the Qubit 2.0 Fluorometric Quantitation (Invitrogen, Life technologies,

343 Sydney, N.S.W., Australia). The sex-reversal status was determined for *B. duperreyi*
344 using PCR as described by Dissanayake et al. (20), where the genotypic sex was
345 identified based on Y-specific markers allowing identification of XY males (XYm) and
346 XX males (XXm). The lizards were euthanised via intraperitoneal injection of sodium
347 pentobarbitone (100–150 µg/g body weight) as approved by the University of
348 Canberra Animal Ethics Committee (AEC 17–26).
349
350 **RNA extraction and sequencing**
351 Total RNA was extracted from the brain tissue of three XY males (XYm), three XX
352 females (XXfm), and two XX males (XXm). Tissue extracts were homogenized using
353 T10 Basic ULTRA-TURRAX® Homogenizer (IKA, Staufen im Breisgau, Germany) and
354 extracted using TRIzol reagent following the manufacturer's instructions, purifying with
355 an isopropanol precipitation. Seventy-five bp single-ended reads were generated on
356 the Illumina NextSeq 500 platform at the Ramaciotti Centre for Genomics (UNSW,
357 Sydney, Australlia).
358
359 **DNA extraction and sequencing for assembly**
360 Genomic DNA was extracted from 13 mg of ethanol-preserved muscle tissue from a
361 male (XY) *B. duperreyi*, using the Circulomics Nanobind tissue kit as per the
362 manufacturer's protocols, including the specified pre-treatment for ethanol removal.
363 Library preparation was performed with 3 µg of DNA as input, using the SQK-LSK109
364 kit from Oxford Nanopore Technologies and sequenced across two promethION (FLO-
365 PRO002) flow cells, with washes (EXP-WSH004) performed every 24 hours.
366
367 **Genome assembly**

368 The genome assembly pipeline relied on using a combination of whole-genome
369 sequencing ONT long reads, Illumina short reads and Hi-C reads. Firstly, adapters
370 were removed from Illumina and Hi-C reads using TrimGalore v0.6.6 with default
371 parameters. A primary assembly was generated using Flye v2.8.3 (34) and the ONT
372 long reads, with the following parameters "--trestle --iterations 2". Illumina short reads
373 were aligned to the primary assembly with bwa-mem v0.7.17-r1188 and ONT long
374 reads were aligned with minimap2 v2.24-r1122 (35), and both alignments were used
375 to polish the assembly with hypo v0.5.1. Homologous contigs were identified and
376 removed using Purge Haplotigs v1.1.1 (36). Chromosome-length scaffolding was
377 performed with Hi-C data using Juicer v1.6 (37) to generate a contact matrix of the
378 connections between contigs and 3d-dna v180922 to organise the contigs into larger
379 scaffolds. Gaps in the assembly were filled using PBjelly v15.8.24 (38) with the ONT
380 reads. The final assembly was assessed for completeness using BUSCO and
381 accuracy using Merqury v1.3 (39), by comparing the assembly k-mer spectrum to
382 those found in the Illumina reads.
383
384 **Annotation**
385 The annotation used in this study was generated using AUGUSTUS v3.4.0 (40) with
386 RNA sequencing data from brain, heart and gonad from male and female individuals.
387 Prior to running AUGUSTUS, a transcriptome assembly and a soft-masked genome
388 assembly were generated with Trinity v2.12.0 (41) and RepeatMasker v4.1.2-p1 (42),
389 respectively. AUGUSTUS parameters were optimised with a training set of 500 genes,
390 including 5 rounds of optimisation with exon and UTR parameters, before gene
391 prediction and stitching. Peptides were inferred from the final annotation using gffread
392 v0.12.7 (43) and blasted against the uniprot database with blast v2.11.0 (44) for gene

393 identification. A github repository of this pipeline can be found at
394 https://github.com/kango2/Annotation.
395
396 Y specific sequences (20) in *B. duperreyi* were blasted against the generated genome
397 assembly with blast v2.11.0 (44).
398
399 **Read depth analysis**
400 DNA was extracted from muscle samples of the individual animals using the Gentra
401 Puregene Tissue Kit (QIAGEN, Australia), following the manufacturer's protocols with
402 the modifications described below. The volume and reagent amounts were adjusted
403 according to the size of the tissue sample, using three times more reagent than
404 specified in the manufacturer's protocols. Additionally, we made modifications to the
405 DNA precipitation steps outlined in the manufacturer's protocol. The DNA thread was
406 spooled out using a galls rod and submerged in 300ul of 70% ethanol, then air dried
407 for one minute. Subsequently, we used TE buffer for DNA hydration and allowed it to
408 dissolve overnight at room temperature.
409
410 Raw read quality of the DNA sequencing was assessed using FastQC v0.11.9 (45)
411 and trimmed accordingly with trimmomatic v0.38 (46). Trimmed reads were aligned to
412 the newly generated *B. duperreyi* genome assembly with subread-align command in
413 subread v2.0.1 (47). Read depth in 20000 bp windows was calculated using the
414 bamCoverage command, from deeptools v3.5.1 (48), for the X and Y scaffolds as well
415 as a representative autosomal scaffold (contig 328). The generated BED files were
416 then used to plot read depth across these scaffolds using the ggplot2 package in R
417 v4.2.1.

418

419 **Bioinformatics analysis of RNA-seq**

420 Raw read quality of RNA sequencing was assessed using FastQC v0.11.9 (45) and
421 trimmed accordingly with trimmomatic v0.38 (46). Trimmed reads were aligned to the
422 newly generated *B. duperreyi* genome assembly with subread-align function in
423 subread v2.0.1 (47). The subread-featurecount function from subread v2.0.1 (49) was
424 used to count reads that overlapped genomic features using the settings "-0 -2 -t
425 CDS,five_prime_utr,three_prime_utr", the remaining settings were left as default.
426 Library size and read counts were normalised by counts per million and RPKM
427 respectively using the edgeR package in R v4.2.1 (50). Gene expression ratios were
428 calculated and plotted with the ggplot2 package. Gene expression values were
429 calculated from the RNA-seq data (hatchling brain tissue) of XX females (n=3), XY
430 males (n=3) and sex reversed XX males (n=2). Normalised expression values were
431 used to create male to female median expression ratios between the three genotypic
432 and phenotypic sexes for each gene. Genes were binned according to their location
433 on an autosome, the PAR, X as total output and X genes with Y gametologues (XY
434 shared).

435

436 **Gene ontology and gene function**

437 Gene ontology analysis was performed using g:Profiler (51) with gene lists identified
438 using the *B. duperreyi* annotation. Results were restricted to biological process and
439 molecular function and used the human database.

440

441 Protein functions of Y genes were obtained using the panther database v19.0 (52).

442

443 **DECLARATIONS**

444

445 **Ethics approval and consent to participate**

446

447 The animals involved in this study were collected from nests in the field under permits
448 issued by the Australian state government: ACT permit numbers (LT201826,
449 LT2017956) and NSW government (SL102002). The collection was carried out in
450 accordance with their respective guidance and protocols, as well as the University of
451 Canberra Animal Ethics Committee (Approval Number AEC 17–26). All husbandry
452 practices complied with the Australian Code for the Care and Use of Animals for
453 Scientific Purposes, 8th edition (2013), particularly sections 3.2.13–3.2.2.

454

455 **Consent for publication**

456 Not applicable

457

458 **Availability of data and materials**

459 All raw sequencing data generated in this study have been submitted to the NCBI
460 BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession
461 number PRJNA980841. The genome assembly presented in this study is available
462 from the NCBI Genome database under GenBank accession GCA_041722995.1.

463

464 **Competing interests**

465 The authors declare that they have no competing interests.

466

467 **Funding**

B.J.H. is supported by an Australian Government Research Training Program (RTP) Scholarship. P.D.W. and J.A.M. are supported by Australian Research Council Discovery Projects (DP210103512 and DP220101429). P.D.W. and H.R.P. are supported by an NHMRC Ideas Grant (2021172). A.R-H. acknowledges the Spanish Ministry of Science and Innovation (PID2020-112557GB-I00 founded by AEI/10.13039/501100011033), the Agència de Gestió d'Ajuts Universitaris i de Recerca, AGAUR (2021SGR00122) and the Catalan Institution for Research and Advanced Studies (ICREA). Fieldwork and initial laboratory work were supported by Australian Research Council Grants DP110104377 and DP170101147, awarded to A.G.

Authors' contributions

B.J.H wrote the original draft. P.D.W and A.G designed the study. D.S.B.D collected and harvested samples. B.J.H and J.K.C conducted extraction experiments to generate data. A.L.M.R, J.M.H and J.K.C generated the genome assembly and annotation. B.J.H and J.K.C conducted data analysis with input from A.M.M and N.C.L. D.S.B.D, I.W.D, A.R-H, H.R.P, J.A.M.G and A.G provided insight into data interpretation and writing. All authors reviewed and approved the submitted manuscript.

Acknowledgements

Sequencing costs for this project were met by grants from Bioplatforms Australia (AusARG initiative) and the Australian Research Council (DP170101147). We acknowledge the Garvan Sequencing Platform.

493 **REFERENCES**

494

495 1. Ohno S. Sex chromosomes and sex-linked genes. Berlin, New York etc.:
496 Springer-Verlag; 1967. x, 192 p. p.
497 2. Lyon MF. X-chromosome inactivation and developmental patterns in
498 mammals. Biol Rev Camb Philos Soc. 1972;47(1):1-35.
499 3. Xiong Y, Chen X, Chen Z, Wang X, Shi S, Wang X, et al. RNA sequencing
500 shows no dosage compensation of the active X-chromosome. Nat Genet.
501 2010;42(12):1043-7.
502 4. Lin F, Xing K, Zhang J, He X. Expression reduction in mammalian X
503 chromosome evolution refutes Ohno's hypothesis of dosage compensation. Proc
504 Natl Acad Sci U S A. 2012;109(29):11752-7.
505 5. Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schutz F, et al.
506 Mechanisms and evolutionary patterns of mammalian and avian dosage
507 compensation. PLoS Biol. 2012;10(5):e1001328.
508 6. Sarre SD, Georges A, Quinn A. The ends of a continuum: genetic and
509 temperature-dependent sex determination in reptiles. Bioessays. 2004;26(6):639-45.
510 7. Ezaz T, Sarre SD, O'Meally D, Graves JAM, Georges A. Sex chromosome
511 evolution in lizards: independent origins and rapid transitions. Cytogenet Genome
512 Res. 2009;127(2-4):249-60.
513 8. Quinn AE, Georges A, Sarre SD, Guarino F, Ezaz T, Graves JA. Temperature
514 sex reversal implies sex gene dosage in a reptile. Science. 2007;316(5823):411.
515 9. Shine R, Elphick MJ, Donnellan S. Co-occurrence of multiple, supposedly
516 incompatible modes of sex determination in a lizard population. Ecol Lett.
517 2002;5(4):486-9.
518 10. Holleley CE, O'Meally D, Sarre SD, Graves JAM, Ezaz T, Matsubara K, et al.
519 Sex reversal triggers the rapid transition from genetic to temperature-dependent sex.
520 Nature. 2015;523(7558):79-82.
521 11. Radder RS, Quinn AE, Georges A, Sarre SD, Shine R. Genetic evidence for
522 co-occurrence of chromosomal and thermal sex-determining systems in a lizard. Biol
523 Lett. 2008;4(2):176-8.

524 12. Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D. Comparative sex
525 chromosome genomics in snakes: differentiation, evolutionary strata, and lack of
526 global dosage compensation. PLoS Biol. 2013;11(8):e1001643.
527 13. Schield DR, Card DC, Hales NR, Perry BW, Pasquesi GM, Blackmon H, et al.
528 The origins and evolution of chromosomes, dosage compensation, and mechanisms
529 underlying venom regulation in snakes. Genome Res. 2019;29(4):590-601.
530 14. Marin R, Cortez D, Lamanna F, Pradeepa MM, Leushkin E, Julien P, et al.
531 Convergent origination of a Drosophila-like dosage compensation mechanism in a
532 reptile lineage. Genome Res. 2017;27(12):1974-87.
533 15. Tenorio M, Cruz-Ruiz S, Encarnacion-Guevara S, Hernandez M, Corona-
534 Gomez JA, Sheccid-Santiago F, et al. MAYEX is an old long noncoding RNA
535 recruited for X chromosome dosage compensation in a reptile. Science.
536 2024;385(6715):1347-54.
537 16. Davalos-Dehullu E, Baty SM, Fisher RN, Scott PA, Dolby GA, Munguia-Vega
538 A, et al. Chromosome-Level Genome Assembly of the Blacktail Brush Lizard,
539 Urosaurus nigricaudus, Reveals Dosage Compensation in an Endemic Lizard.
540 Genome Biol Evol. 2023;15(12).
541 17. Rovatsos M, Rehak I, Velenský P, Kratochvil L. Shared Ancient Sex
542 Chromosomes in Varanids, Beaded Lizards, and Alligator Lizards. Mol Biol Evol.
543 2019;36(6):1113-20.
544 18. Acosta A, Suarez-Varon G, Rodriguez-Miranda LA, Lira-Noriega A, Aguilar-
545 Gomez D, Gutierrez-Mariscal M, et al. Corytophanids Replaced the Pleurodont XY
546 System with a New Pair of XY Chromosomes. Genome Biol Evol. 2019;11(9):2666-
547 77.
548 19. Webster TH, Vannan A, Pinto BJ, Denbrock G, Morales M, Dolby GA, et al.
549 Lack of Dosage Balance and Incomplete Dosage Compensation in the ZZ/ZW Gila
550 Monster (Heloderma suspectum) Revealed by De Novo Genome Assembly.
551 Genome Biol Evol. 2024;16(3).
552 20. Dissanayake DSB, Holleley CE, Hill LK, O'Meally D, Deakin JE, Georges A.
553 Identification of Y chromosome markers in the eastern three-lined skink (Bassiana
554 duperreyi) using in silico whole genome subtraction. BMC Genomics.
555 2020;21(1):667.

556 21. Kostmann A, Kratochvil L, Rovatsos M. Poorly differentiated XX/XY sex
557 chromosomes are widely shared across skink radiation. Proc Biol Sci.
558 2021;288(1943):20202139.
559 22. Dissanayake DSB, Holleley CE, Deakin JE, Georges A. High elevation
560 increases the risk of Y chromosome loss in Alpine skink populations with sex
561 reversal. Heredity (Edinb). 2021;126(5):805-16.
562 23. Dissanayake DSB, Holleley CE, Georges A. Effects of natural nest
563 temperatures on sex reversal and sex ratios in an Australian alpine skink. Sci Rep.
564 2021;11(1):20093.
565 24. Álvarez-González L, Arias-Sardá C, Montes-Espuña L, Marín-Gual L, Vara C,
566 Lister NC, et al. Principles of 3D chromosome folding and evolutionary genome
567 reshuffling in mammals. Cell Rep. 2022;41(12):111839.
568 25. Waters PD, Patel HR, Ruiz-Herrera A, Álvarez-González L, Lister NC,
569 Simakov O, et al. Microchromosomes are building blocks of bird, reptile, and
570 mammal chromosomes. Proc Natl Acad Sci U S A. 2021;118(45).
571 26. Dissanayake DSB. Sex reversal in the alpine skink Bassiana duperreyi –
572 response to natural environment [Doctoral Thesis]: University of Canberra; 2022.
573 27. Holleley CE, Sarre SD, O'Meally D, Georges A. Sex Reversal in Reptiles:
574 Reproductive Oddity or Powerful Driver of Evolutionary Change? Sex Dev.
575 2016;10(5-6):279-87.
576 28. Whiteley SL, Castelli MA, Dissanayake DSB, Holleley CE, Georges A.
577 Temperature-Induced Sex Reversal in Reptiles: Prevalence, Discovery, and
578 Evolutionary Implications. Sex Dev. 2021;15(1-3):148-56.
579 29. Cornejo-Páramo P, Dissanayake DSB, Lira-Noriega A, Martínez-Pacheco ML,
580 Acosta A, Ramírez-Suástegui C, et al. Viviparous Reptile Regarded to Have
581 Temperature-Dependent Sex Determination Has Old XY Chromosomes. Genome
582 Biol Evol. 2020;12(6):924-30.
583 30. Van Dyke JU, Thompson MB, Burridge CP, Castelli MA, Clulow S,
584 Dissanayake DSB, et al. Australian lizards are outstanding models for reproductive
585 biology research. Australian Journal of Zoology. 2020;68(4):168-99.
586 31. Lister NC, Milton AM, Patel HR, Waters SA, Hanrahan BJ, McIntyre KL, et al.
587 Incomplete transcriptional dosage compensation of chicken and platypus sex
588 chromosomes is balanced by post-transcriptional compensation. Proc Natl Acad Sci
589 U S A. 2024;121(32):e2322360121.

590 32. Dissanayake DSB, Holleley CE, Sumner J, Melville J, Georges A. Lineage
591 diversity within a widespread endemic Australian skink to better inform conservation
592 in response to regional-scale disturbance. Ecol Evol. 2022;12(3):e8627.
593 33. Harlow PS. A Harmless Technique for Sexing Hatchling Lizards. Herpetol
594 Rev. 1996;27(2):71-.
595 34. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone
596 reads using repeat graphs. Nat Biotechnol. 2019;37(5):540-6.
597 35. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics.
598 2018;34(18):3094-100.
599 36. Roach MJ, Schmidt SA, Borneman AR. Purge Haplotigs: allelic contig
600 reassignment for third-gen diploid genome assemblies. BMC Bioinformatics.
601 2018;19(1):460.
602 37. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, et al.
603 Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C
604 Experiments. Cell Syst. 2016;3(1):95-8.
605 38. English AC, Richards S, Han Y, Wang M, Vee V, Qu J, et al. Mind the gap:
606 upgrading genomes with Pacific Biosciences RS long-read sequencing technology.
607 PLoS One. 2012;7(11):e47768.
608 39. Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality,
609 completeness, and phasing assessment for genome assemblies. Genome Biol.
610 2020;21(1):245.
611 40. Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically
612 mapped cDNA alignments to improve de novo gene finding. Bioinformatics.
613 2008;24(5):637-44.
614 41. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al.
615 Full-length transcriptome assembly from RNA-Seq data without a reference genome.
616 Nat Biotechnol. 2011;29(7):644-52.
617 42. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. 2013-2015.
618 43. Pertea G, Pertea M. GFF Utilities: GffRead and GffCompare. F1000Res.
619 2020;9.
620 44. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al.
621 BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.
622 45. Andrews S. FastQC: a quality control tool for high throughput sequence data.
623 2010.

624 46. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
625 sequence data. Bioinform. 2014;30(15):2114-20.
626 47. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable
627 read mapping by seed-and-vote. Nucleic Acids Res. 2013;41(10):e108.
628 48. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al.
629 deepTools2: a next generation web server for deep-sequencing data analysis.
630 Nucleic Acids Res. 2016;44(W1):W160-5.
631 49. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose
632 program for assigning sequence reads to genomic features. Bioinformatics.
633 2014;30(7):923-30.
634 50. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for
635 differential expression analysis of digital gene expression data. Bioinformatics.
636 2010;26(1):139-40.
637 51. Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H. g:Profiler-
638 interoperable web service for functional enrichment analysis and gene identifier
639 mapping (2023 update). Nucleic Acids Res. 2023;51(W1):W207-W12.
640 52. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou LP, Mi H.
641 PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci.
642 2022;31(1):8-22.
643 53. Ramirez F, Bhardwaj V, Arrigoni L, Lam KC, Gruning BA, Villaveces J, et al.
644 High-resolution TADs reveal DNA sequences underlying genome organization in
645 flies. Nat Commun. 2018;9(1):189.
646 54. Lovell JT, Sreedasyam A, Schranz ME, Wilson M, Carlson JW, Harkess A, et
647 al. GENESPACE tracks regions of interest and gene copy number variation across
648 multiple genomes. Elife. 2022;11.
649

650 **SUPPLEMENTAL FIGURE LEGENDS**

651

652 **Supplemental Figure 1** Blastn hits to the the largest 15 scaffolds representing the *B.*
653 *duperreyi* chromosomes with males specific kmer contigs.

654

655 **Supplemental Figure 2.** Expression ratios for each autosomal scaffold for the three
656 pairwise sex/ genotype comparisons presented as boxplots. Number of genes (n) is
657 show in the key for each contig.
658
659 **Supplemental Figure 3.** Expression ratios were calculated for each gene on **A**) a
660 representative autosome (contig 31) and **B)** the X chromosome (contig 1241), in a
661 pairwise fashion for each sex (XYm, XXm and XXf). Values are plotted on a log2 scale
662 and median values for each comparison are plotted as a red line. For panel B median
663 for the X specific region is a red line and PAR is a blue line.
664
665 **Supplemental Figure 4** Scatterplot of counts per million (CPM) for each gene for each
666 pairwise comparison of XXf to XYm, XXm to XYm and XXf to XXm. X genes are plotted
667 in black with slope plotted as a red line. Autosomal genes are plotted in blue with slope
668 as a green line. Slope and R-squared of each trendline is shown above each plot.
669
670 **Supplemental Table 1** List of genes on the Y chromosome (contig 93) including
671 protein function
672
673 **Supplemental Table 2** Gene ontology results for XY shared genes
674
675 **Supplemental Table 3** Gene ontology results for all Y genes
676