1   **Evolutionary transition to XY sex chromosomes associated with Y-linked duplication of a**

2   **male hormone gene in a terrestrial isopod**

3

4   Aubrie Russell[1], Sevarin Borrelli[1], Rose Fontana[1], Joseph Laricchiuta[1], Jane Pascar[1,2], Thomas

5   Becking[3], Isabelle Giraud[3], Richard Cordaux[3], Christopher H. Chandler[1*]

6

7   [1]Department of Biological Sciences, State University of New York at Oswego, Oswego, NY,

8   United States

9   [2]Current address: Center for Reproductive Evolution, Department of Biology, Syracuse

10   University, Syracuse, NY, United States

11   [3]Laboratoire Ecologie et Biologie des Interactions, Equipe Ecologie Evolution Symbiose,

12   Université de Poitiers, UMR CNRS 7267 Poitiers, France

13

14   *Corresponding author:

15   Christopher H. Chandler

16   Department of Biological Sciences

17   State University of New York at Oswego

18   christopher.chandler@oswego.edu

19

20   Key words: *Wolbachia*, Oniscidea, *Trachelipus*, androgenic gland hormone, AGH

21

22 **Abstract**

23

24 Sex chromosomes are highly variable in some taxonomic groups, but the evolutionary

25 mechanisms underlying this diversity are not well understood. In terrestrial isopod crustaceans,

26 interactions with *Wolbachia*, a vertically transmitted endosymbiont causing male-to-female sex

27 reversal, are thought to drive rapid evolutionary turnovers in sex chromosomes. Here, we use

28 surgical manipulations and genetic crosses, plus genome sequencing, to examine sex

29 chromosomes in the terrestrial isopod *Trachelipus rathkei*. Although an earlier cytogenetics

30 study suggested a ZZ/ZW sex chromosome system in this species, we surprisingly find that in

31 our study population, sex is determined by an XX/XY system. Consistent with a recent

32 evolutionary origin for this XX/XY system, the putative male-specific region of the genome is

33 small. The genome shows evidence of sequences horizontally acquired from past *Wolbachia*

34 infections, as well as evidence of Y-linked duplications of the androgenic gland hormone gene,

35 thought to be a possible target for sex reversal by *Wolbachia*. Overall, these results are

36 consistent with the hypothesis that reproductive endosymbionts such as *Wolbachia* can promote

37 quick turnover of sex determination mechanisms in their hosts.

38

39

40 **Introduction**

41

42   Although sexual reproduction is shared by most eukaryotes, a variety of different cues can

43   trigger individuals to follow a male, female, or hermaphroditic developmental plan (Conover &

44   Kynard 1981; Janzen & Phillips 2006; Ospina-Álvarez & Piferrer 2008; Tingley & Anderson

45   1986; Verhulst et al. 2010). In many eukaryotes, sex is primarily determined genotypically, and

46   in most of those cases, sex chromosomes are the primary sex-determining factors, although

47   some exceptions, including polygenic systems and haplodiploidy, are also known (Vandeputte

48   et al. 2007; Heimpel & de Boer 2008). Sex chromosomes in animals are usually grouped into

49   two main classes: XY systems, in which males are heterogametic (XY) and females are

50   homogametic (XX); and ZW systems, in which females are heterogametic (ZW) and males are

51   homogametic (ZZ). However, non-genetic cues can also play an important role in some species.

52   For instance, environmental factors, such as temperature or population density, influence or

53   determine phenotypic sex in reptiles, fishes, and invertebrates (Conover & Kynard 1981; Janzen

54   & Phillips 2006; Tingley & Anderson 1986). In some cases, cytoplasmic factors, including sex-

55   reversing endosymbionts, such as *Wolbachia*, microsporidia, and paramyxids can serve as a

56   sex-determining signal (Bouchon et al. 1998; Pickup & Ironside 2018; Terry et al. 1998; Negri I

57   et al. 2006; Kageyama et al. 2002).

58

59   Evolutionary theory holds that the formation of sex chromosomes begins when an autosome

60   acquires a sex-determining locus (Rice 1996). Subsequently, recombination around the sex-

61   determining locus is selected against because of sexually antagonistic selection (Bergero &

62   Charlesworth 2009). For instance, selection should favor mutations that are beneficial in males

63   but deleterious in females when those alleles are linked to a dominant male-determining allele;

64   recombination, on the other hand, would break up this linkage and result in females that carry

65   these male-beneficial alleles. The non-recombining region is then expected to spread in the

66   presence of continued sexually antagonistic selection, and may eventually span the whole sex

67   chromosome, except for the usual presence of a small recombining pseudo-autosomal region

68   (Charlesworth et al. 2005). Once recombination has ceased, the non-recombining sex

69   chromosome, such as the Y chromosome in mammals or the W chromosome in birds, is

70   expected to degenerate. Non-recombining genes frequently undergo pseudogenization,

71   acquiring nonsense mutations or transposable element insertions (Charlesworth & Charlesworth

72   2000). At the same time, gene trafficking can occur when selection promotes the translocation

73   of formerly autosomal genes to the sex chromosomes (Emerson et al. 2004).

74

75   Different species appear to be at different stages of sex chromosome evolution. For instance,

76   the sex chromosomes of therian mammals are highly conserved, having originated ~160 million

77   years ago (Potrzebowski et al. 2008; Veyrunes et al. 2008). The highly degenerated,

78   heteromorphic Y chromosome represents an advanced stage of sex chromosome evolution. In

79   other taxonomic groups, on the other hand, sex chromosomes appear to undergo more frequent

80   evolutionary turnovers (Pennell et al. 2018; Cioffi et al. 2013; Myosho et al. 2015; Jeffries et al.

81   2018; Ross et al. 2009; Vicoso & Bachtrog 2015). Such young sex chromosomes may have little

82   or no recombination suppression, differentiation in gene content, or sex chromosome dosage

83   compensation, and may not be detectable by traditional cytogenetic methods because they are

84   visually indistinguishable (homomorphic) (Vicoso & Bachtrog 2015; Gamble et al. 2014). Sex

85   chromosomes may even be polymorphic within a species, with different sex-determining loci

86   segregating within or among populations (Traut 1994; Meisel et al. 2016; Ogata et al. 2008;

87   Orzack et al. 1980).

88

89   Unfortunately, we still have a limited understanding of why evolutionary turnovers of sex

90   chromosomes are rare in some groups but frequent in others. A variety of models have been

91   proposed to explain why these turnovers occur, including sexual antagonism, deleterious

92 mutations, and the 'hot potato' model (van Doorn & Kirkpatrick 2007; Blaser et al. 2013, 2014).

93 In some organisms, interactions with vertically transmitted reproductive endosymbionts are also

94 thought to influence the evolution of their hosts' sex determination mechanisms (Rigaud et al.

95 1997; Cordaux et al. 2011). However, many of these models have been difficult to test in nature.

96 This problem is exacerbated by the fact that, while sex chromosomes have been extensively

97 studied in model organisms like *Drosophila*, studies are more sparse in non-model organisms.

98

99 One group that has received relatively little attention is crustaceans. Different crustacean

100 species show a variety of distinct sex determining mechanisms, yet there are very few

101 crustacean species in which candidate master sex-determining genes have been identified

102 (Chandler et al. 2018, 2017). Within crustaceans, perhaps one of the best-studied groups in

103 terms of sex determination is the terrestrial isopods (Oniscidea). Terrestrial isopod species have

104 a mix of XY and ZW systems, along with reports of a few parthenogenic species and

105 populations (Rigaud et al. 1997; Johnson 1986; Fussey 1984). The bacterial endosymbiont

106 *Wolbachia* also influences sex determination by causing male-to-female sex reversal in some

107 isopod hosts (Bouchon et al. 1998; Cordaux et al. 2004). In fact, interactions with *Wolbachia* are

108 thought to drive rapid evolutionary turnover of the sex chromosomes in terrestrial isopods. This

109 hypothesis is supported by multiple lines of evidence. For instance, in the common pillbug

110 *Armadillidium vulgare*, a copy of the *Wolbachia* genome horizontally integrated into the host

111 genome (formerly known as the *f* element) led to the origin of a new W chromosome (Leclercq

112 et al. 2016). A recent phylogenetic analysis also identified, at minimum, several transitions in

113 heterogametic systems along the isopod phylogeny, including closely related species pairs with

114 different sex chromosome systems (Becking et al. 2017). Moreover, only a few species of

115 terrestrial isopods are known to have heteromorphic sex chromosomes, in which the X and Y, or

116 Z and W, chromosomes are visually distinguishable in cytogenetics experiments (Rigaud et al.

117 1997), and WW or YY individuals are often viable and fertile (Juchault & Rigaud 1995; Becking

118   et al. 2019), suggesting that the W and Y chromosomes have not lost any essential genes in

119   these species.

120

121   In this study, we examined sex determination in the widespread species *Trachelipus rathkei*.

122   This species was previously established by cytogenetic methods to have heteromorphic, albeit

123   slightly, Z and W sex chromosomes (Mittal & Pahwa 1980). We sought to confirm female

124   heterogamety by crossing females to sex-reversed neo-males (which have female genotypes

125   but male phenotypes), and assessing the sex ratio of the resultant progenies. Surprisingly, we

126   found that, at least in our focal population, sex is determined by an XX/XY system, suggesting a

127   recent sex chromosome turnover. To test this hypothesis, we performed whole-genome

128   sequencing. Consistent with a recent origin of an XX/XY sex determination system, we find

129   evidence that the putative male-specific region of the genome is small, and we identified a male-

130   specific partial duplication of the androgenic gland hormone (AGH) gene, a rare example of a

131   candidate sex-determining gene in a crustacean. In addition, although our study population

132   does not appear to harbor current *Wolbachia* infections, we find genomic evidence of past

133   infections. Overall, our results are consistent with the hypothesis that sex-reversing

134   endosymbionts like *Wolbachia* can drive rapid evolutionary turnover of sex chromosomes in

135   their hosts.

136

137   **Methods**

138

139   Animal collection and husbandry

140

141   We sampled wild isopods from Rice Creek Field Station (RCFS) at SUNY Oswego in Oswego,

142   NY. We captured animals using a combination of methods. First, we haphazardly searched

143   through leaf litter, logs, and rocks. We also used "potato traps", made by carving out a 1-2 cm

144   diameter core from a potato and placing it in the litter for 1-2 weeks. Finally, we constructed

145   pitfall traps from plastic cups buried in the ground with the rim of the cup flush with the ground.

146   The primary species captured were *Oniscus asellus* and *Trachelipus rathkei*, but we also

147   captured *Philoscia muscorum*, *Hyloniscus riparius*, *Trichoniscus pusillus*, and occasionally

148   *Cylisticus convexus*. Species identification was performed in the field and confirmed in the lab,

149   where we also determined the phenotypic sex of specimens.

150

151   Isopods were housed in plastic food storage containers with holes in the lids for air exchange,

152   on a substrate of moistened soil. Containers were checked twice weekly. Animals were fed

153   carrots and dried leaves *ad libitum*. The photoperiod was kept on a schedule of 18:6 light

154   hours:dark hours in the summer and and 14:10 in the winter. We isolated ovigerous females in

155   individual containers, and separated offspring from their mothers upon emergence from the

156   marsupium. We initially sexed offspring at six to eight weeks old, and separated males from

157   females to prevent sibling mating. We then double-checked offspring sex at roughly two week

158   intervals thereafter until four months of age to watch for individuals that might have shown late

159   signs of sexual differentiation. Terrestrial isopods are known to store sperm from a single mating

160   to fertilize future broods. Therefore, for experimental crosses we only used *T. rathkei* females

161   that were born in the lab, separated from brothers as soon as they could be sexed, and which

162   had not produced any offspring by 12 months of age.

163

164   *Wolbachia* testing

165

166   We used PCR assays to detect *Wolbachia* in *T. rathkei* individuals. DNA was extracted from

167   one or two legs, depending on the size of the animal. We ruptured the leg tissue in 400 µL

168   deionized water along with a few 0.5 mm zirconia/silica beads (enough to cover the bottom of

169   the tube) using a bead beater machine. Samples were lysed following a protocol of 2500 RPM

170   for 10 seconds, followed by 4200 RPM for 10 seconds, and finally 4800 RPM for 10 seconds.

171   The tube was then visually inspected to confirm the leg was sufficiently pulverized. We then

172   transferred the lysate to a new tube, added 60 μL of a 5% Chelex® 100 molecular biology grade

173   resin suspension, and incubated for 15 minutes at 100° C. After incubation, we centrifuged the

174   extract at 16,000g for 3 minutes, and reserved 80 μL of supernatant for PCR testing. We

175   confirmed successful DNA extraction using the mitochondrial primers HCO2198/LCO1490

176   (Folmer et al. 1994). We performed PCRs in 10 μL reactions, using 4.95 μL of molecular biology

177   grade water, 2 μL NEB OneTaq Buffer, 1 μL of mixed dNTPS at a final concentration of 2mM for

178   each dNTP, 1 μL of a 5 μM solution of each primer, and 0.05 μL of NEB OneTaq. For the

179   mitochondrial primer set, PCR conditions included an initial denaturation of 94° C for 1 minute; 5

180   cycles of 94° C denaturation for 30s, 45° C annealing for 90s, and 68° C extension for 60s. The

181   samples then underwent 35 cycles of 94° C for 30s, 51° C for 90s, and 68° C for 60s. This was

182   followed by a final extension step of 68° C for 5 minutes. To test for *Wolbachia*, we performed

183   PCR using *Wolbachia*-specific primers targeting the *wsp* (81f/691r) and *ftsZ* (ftsZf1/ftsZr1)

184   genes (Braig et al. 1998; Werren John H. et al. 1995). We performed PCRs in 10 μL reactions,

185   using 4.95 μL of molecular biology grade water, 2 μL NEB OneTaq Buffer, 1 μL of mixed dNTPs

186   at a final concentration of 2mM for each dNTP, 1 μL of either *wsp* or *ftsZ* primers, and 0.05 μL

187   of NEB OneTaq. PCR conditions contained an initial denaturation of 95°C for 5 minutes,

188   followed by 36 cycles of 95° C for 60s, 54° C for 60s, and 68°C for 3 minutes. This was followed

189   by a final extension step at 68° C for 10 minutes.

190

191   Androgenic gland implantation and crosses

192

193   To test whether sex is determined by a ZZ/ZW or XX/XY system of sex determination in our

194   population of *T. rathkei*, we performed crosses between females and experimentally sex-

195   reversed neo-males. Juvenile female *T. rathkei* were implanted with live androgenic glands,

196   according to (Becking et al. 2017). Male donors and female recipients were selected from large

197   lab-reared broods with even (~1:1) sex ratios. An adult male was sacrificed by decapitation, and

198   live androgenic glands were dissected into Ringer solution (393 mM NaCl, 2 mM KCl, 2 mM

199    $CaCl_2 \cdot 2H_2O$, 2 mM $NaHCO_3$). Female recipients were between 5 and 8 weeks old, an age at

200   which males and females begin to become distinguishable, but at which sexual development is

201   not complete. The gland was injected using a pulled glass pipette into a hole pierced with a

202   dissecting needle in the 6th or 7th segment of the juvenile female's pereon. Recipients were

203   isolated in a small plastic container with a moist paper towel for recovery and observation.

204   Experimental animals were monitored for signs of male development. Any animal that failed to

205   develop male genitalia by 4 months post-implantation was considered to be a failed injection.

206   After maturation, adult neo-males were placed in individual containers with 1-3 previously

207   unmated females. Crosses were monitored twice weekly to check for signs of reproduction in

208   females. Gravid females were then isolated into their own containers until parturition.

209

210   Genome sequencing

211

212   All raw sequencing data and the draft genome assembly are available under NCBI BioProject

213   PRJNA633105. Data analysis scripts are available at

214   https://github.com/chandlerlab/trachelipus_genome.

215

216   We performed whole-genome sequencing using a combination of Illumina, PacBio, and Oxford

217   Nanopore sequencing, with multiple sequencing samples of each sex (Table 1). Because we

218   expected the *T. rathkei* genome to be large, repetitive, and highly polymorphic, and because we

219   expected to need to isolate DNA from multiple individuals, we established a partially inbred

220   laboratory line using offspring from a single female collected from RCFS. We mated brothers

221   and sisters from this female for two generations in the lab prior to collecting genetic samples

222   from the third generation for sequencing. DNA was collected for sequencing using the Qiagen

223   DNEasy Blood and Tissue Kit following the manufacturer's instructions. DNA was quantified

224   using the Qubit DNA Broad Range assay kit, and the A260/280 value was checked with a

225   Nanodrop spectrophotometer. Samples were stored at -80°C prior to being shipped to the

226   sequencing center. Illumina sequencing was performed at the State University of New York at

227   Buffalo Genomics and Bioinformatics Core Facility.

228

229   For PacBio sequencing, we had to pool DNA from multiple individuals to obtain sufficient

230   quantities of DNA for library preparation. We performed separate DNA extractions from three

231   individuals of each sex as above. Then, we pooled the DNA from the three individuals of each

232   sex and concentrated it using Ampure XP beads (Beckman-Coulter). Briefly, we washed the

233   beads three times in molecular biology grade water, once in Qiagen buffer EB, and finally re-

234   suspended the beads in their original buffer. We then added equal volumes of Ampure XP

235   suspension to the DNA samples, mixed them on a shaker for 15 minutes, placed the tubes on a

236   magnetic bead separator, and removed the supernatant. We washed the beads twice with 1.5

237   mL of 70% ethanol, and finally eluted the DNA samples in 30 μL of Qiagen buffer EB.

238   Sequencing libraries were prepared and sequenced at the University of Delaware Sequencing &

239   Genotyping Center on a PacBio RSII.

240

241   We also supplemented our PacBio dataset with Oxford Nanopore sequencing data. We isolated

242   DNA from a single *T. rathkei* female and two separate males using a Qiagen DNEasy Kit as

243   described above. We then performed sequencing on Oxford Nanopore Minion flowcells (R9.4)

244   with the Rapid Sequencing Kit (SQK-RAD004) following the manufacturer's instructions.

245

246   Genome assembly

247

248   We performed a hybrid assembly combining the short- and long-read sequence data. We first

249   removed adapters and trimmed the Illumina sequencing data using Trimmomatic v. 0.36 (Bolger

250   et al. 2014); we removed leading and trailing bases, as well as internal windows of at least 4bp,

251   with a quality score of 5 or lower, and discarded any reads shorter than 36bp after trimming. We

252   then used SparseAssembler (Ye et al. 2012) to assemble the cleaned Illumina data from sample

253   Mpool (to minimize the number of sequence polymorphisms that would be present in the data

254   with additional samples), using two different kmer sizes (k=51 and k=61). After performing

255   preliminary quality checks using Quast (Gurevich et al. 2013), we decided to proceed with the

256   k=61 assembly, which had the longer total length and N50 (Supplementary Table 1). However,

257   because we suspected the genome might still contain high levels of heterozygosity despite two

258   generations of inbreeding, we used Redudans (Pryszcz & Gabaldón 2016) to remove putative

259   allelic contigs from the Illumina-only assembly; we set identity and overlap thresholds of 95%.

260

261   Prior to performing hybrid assembly, we used the short reads to correct sequencing errors in the

262   long reads using FMLRC (Wang et al. 2018) with the default settings, except requiring a

263   minimum count of 3 to consider a path (-m 3). PacBio and Oxford Nanopore reads derived from

264   female isopods were corrected using Illumina sample Fpool, while long reads from male

265   samples were corrected using sample Mpool.

266

267   We next performed hybrid assembly using DBG2OLC (Ye et al. 2016), which accepts a short-

268   read assembly (rather than raw short-read sequence data) and long-read sequence data (in this

269   case, our PacBio and Oxford Nanopore reads) as input. We tested out a range of different

270   parameter values: from the Redundans-filtered assembly, we first removed contigs less than

271   100bp or 200bp; we tested kmer sizes of 17 and 19; for the kmer coverage threshold, we tried

272   values of 2 and 5; and for the minimum overlap, we tried values of 10 and 30. We used an

273   adaptive threshold of 0.01. These assemblies ranged in size from ~5.2 Gb to 8.5 Gb; we

274   selected three assemblies across the range of total sizes for further processing.

275

276   We next corrected errors in these assemblies, caused by the relatively high error rates in long-

277   read sequence data. In the standard DBG2OLC pipeline, the resulting contigs are corrected

278   using the contigs from the short-read assembly and from the long reads using Sparc (Ye & Ma

279   2016); however, in our initial attempts, large portions of the assemblies went uncorrected,

280   perhaps because we had relatively low-coverage long-read data. Therefore, instead we

281   performed three rounds of error correction using Pilon (Walker et al. 2014), by mapping the raw

282   Illumina sequence reads to each assembly using bbmap (first two rounds; (Bushnell et al.

283   2017)) and bwa mem (third round; with the parameters -A 1 -B 1 -O 1 -E 1 -k 11 -W 20 -d 0.5 -L

284   6 for mapping to an error-prone assembly; (Li 2013)).

285

286   Finally, we assessed the quality of each assembly using BUSCO v.3.0.2 (Simão et al. 2015),

287   with the arthropod reference gene set, and selected the assembly with the greatest number of

288   BUSCO reference genes present for further analysis.

289

290   To remove contaminants from the final assembly, we generated blob plots using Blobtools v.1.0

291   (Laetsch & Blaxter 2017). To accomplish this, we BLASTed all contigs against the NCBI

292   nucleotide (nt) database using megablast (Morgulis et al. 2008) and against Uniprot reference

293   proteomes using diamond blastx (Buchfink et al. 2015). We then removed any contigs that were

294   identified as coming from plants, fungi, viruses, or bacteria, except for those matching

295   *Wolbachia*.

296

297   Genome annotation

298

299   We used RepeatModeler v.1.0.10, which uses RECON (Bao & Eddy 2002), RepeatScout (Price

300   et al. 2005), and Tandem Repeat Finder (Benson 1999), to construct a custom repeat library for

301   *T. rathkei*. Because we were unable to run RepeatModeler successfully using the full assembly,

302   we ran RepeatModeler on a random subset 40% of the contigs; this should still successfully

303   identify most repetitive elements in the genome as long as all repeat families are still well

304   represented in the subset. We then masked the assembly using RepeatMasker 4.0.7 (Tarailo-

305   Graovac & Chen 2009).

306

307   We annotated coding sequences using the MAKER pipeline (Cantarel et al. 2008). We initially

308   ran MAKER using assembled transcript sequences (est2genome=1) from previously available

309   data from one wild-caught male and one wild-caught female *T. rathkei* from the same population

310   (Becking et al. 2017) (SRR5198727, SRR5198726), along with protein alignments against

311   Uniprot-Swissprot (version March 2020) and used the resulting output to train SNAP (Korf

312   2004). To train AUGUSTUS (Stanke et al. 2006) we used the output from the BUSCO quality

313   assessment described earlier. We then completed a final round of MAKER using the trained

314   gene models.

315

316   Development of sex-linked PCR markers

317

318   We used multiple approaches to develop male-specific, putatively Y-linked PCR markers. Initial

319   attempts to perform a genomic assembly with male samples and then identify contigs with zero

320   coverage in females were unsuccessful. We therefore developed a complementary approach by

321   looking for male-specific k-mers in the raw sequencing reads. We used kmc v.3.1.0 (Kokot et al.

322   2017) to count all the canonical 21-mers in each of the Illumina sequencing datasets (in other

323   words, each 21-mer and its reverse complement were considered to be the same k-mer during

324   counting). We then searched for k-mers that occurred at least 8 times in the Mpool Illumina

325   sequencing dataset and a total of at least 3 times combined across the lower coverage M2, M5,

326   M6, and wildM samples, but which were completely absent from all female samples. We then

327   extracted all sequence reads containing these candidate male-specific k-mers using mirabait

328   v.4.0.2 (Chevreux et al. 1999), and assembled them using Spades v.3.11.1 (Bankevich et al.

329   2012).

330

331   To test male-specificity of these contigs, we used polymerase chain reaction (PCR). We

332   developed PCR primers for a subset of candidate male-specific contigs. To identify the best

333   candidates, we first mapped raw sequencing reads from all male and female samples to the full

334   genome sequence plus the candidate male-specific contigs, and identified contigs that had

335   coverage in male samples but not female samples; we also avoided contigs that showed

336   evidence of containing repeat elements, after BLAST searches against the whole genome

337   assembly. We designed primers using PRIMER3 (Untergasser et al. 2012, 3). In these PCRs,

338   primers were initially screened using template DNA from two male samples and two female

339   samples; primers that showed evidence of sex specificity after this first PCR were re-tested

340   using a larger number of samples. PCR primers were initially tested using a cycle of 98°C for 3

341   minutes, followed by 40 cycles of 98°C for 15s, 50°C for 35s, and 68°C for 60s; this was

342   followed by a final extension step of 68°C for 10 minutes. For samples that did not amplify under

343   this program, a gradient PCR was run to determine optimal annealing temperature. All PCRs

344   were performed using the same recipe and reaction conditions as the *Wolbachia* PCRs

345   described above.

346

347   We also identified open reading frames (ORFs) in these candidate male-specific contigs using

348   Transdecoder v.4.0.0 (Haas & Papanicolaou 2016), and annotated the ORFs using Trinotate

349   v.3.1.1 (Bryant et al. 2017). Subsequently, we designed additional primers targeting these

350   ORFs.

351

352 **Results**

353

354 <u>No *Wolbachia* and balanced sex ratios in *T. rathkei*</u>

355

356 Among the 100+ individuals captured and tested between 2015 and 2017, no *T. rathkei* from

357 RCFS conclusively tested positive for *Wolbachia*. This was not due to inadequate testing

358 protocols; for instance, a captive population of *Porcellio laevis* housed in our lab shows nearly a

359 100% infection rate using the same methods (not shown). Approximately 150 *T. rathkei* broods

360 were raised in the lab from either mated, wild-caught females or first-generation crosses. The

361 mean and median brood sizes of this species in our lab were 27.1 and 22.5 offspring,

362 respectively, and the vast majority of these broods had a balanced sex ratio (Supplementary

363 Table 2). Thus, the prevalence of *Wolbachia* and other sex ratio distorters is at most very low in

364 this population of *T. rathkei*. In addition, some wild-caught females produced broods even after

365 several months to a year in isolation in the lab (Supplementary Table 2), confirming that this

366 species is capable of long-term sperm storage.

367

368 <u>Crossing sex-reversed individuals indicates an XY sex determination system</u>

369

370 Five juveniles implanted with androgenic glands survived to mature into males; they were

371 crossed with virgin females from families with normal sex ratios. Each putative neo-male was

372 paired with 2 to 3 females, and each female produced 1-3 broods of offspring. Two of these

373 males sired broods with balanced sex ratios (not significantly different from a 1:1 ratio of males

374 to females; Table 2); these were likely individuals that would have developed into males even

375 without the AG implantation, and thus are uninformative with respect to the sex determination

376 system (Becking et al. 2017). However, three other males produced only female offspring,

377   consistent with an XX/XY system (XX neo-male × XX female yields all XX and therefore 100%

378   female offspring) but not a ZZ/ZW system (ZW neo-male × ZW female expected to produce 1/4

379   ZZ, 1/2 ZW, and 1/4 WW offspring, thus 75% female or 66.7% female depending on whether

380   WW genotypes are viable).

381

382   <u>Genome assembly</u>

383

384   The draft genome assembly of *T. rathkei* is approximately 5.2 Gb in total length. The genome is

385   highly repetitive, consisting of approximately 70% repetitive elements. Transposable elements

386   constitute the largest repeat category, with LINEs, followed by DNA elements and LTRs, being

387   the most represented (Figure 1). All repeat families seem to have a single divergence peak of

388   around 7-10% (Figure 1).

389

390   Despite its large size, the draft assembly is likely only partially complete, with ~25% of arthropod

391   BUSCO genes missing (Table 3). There were 15,805 transcripts assembled from the previously

392   available transcriptome dataset whose best hits in blastn searches against the NCBI nt

393   database and diamond blastx searches against Uniprot-Swissprot were from other arthropods;

394   of those, only 53% had nearly full-length matches in the genome (≥ 90% of the transcript length

395   at ≥ 90% sequence identity), suggesting some missing data and/or remaining uncorrected

396   sequencing errors in the draft assembly as well.

397

398   We screened the *T. rathkei* genome for *Wolbachia* nuclear insertions by BLASTing the

399   assembled contigs against a collection of *Wolbachia* genome sequences, and then BLASTing

400   the matching regions against all representative bacterial genomes from RefSeq to rule out false

401   positives. After this filtering step, we were left with 1010 high confidence matches (best BLAST

402   hit in a *Wolbachia* genome, e-value < 1 x 10^{-6}) spread across 719 contigs, with a total length of

403   ~350 kb for the matching sequences (Supplementary Table 3), much smaller than a typical full

404   *Wolbachia* genome of about 1 - 1.6 Mb on average (Sun et al. 2001). These likely horizontally

405   acquired sequences clustered closely with *Wolbachia* strain *wCon* from the isopod *Cylisticus*

406   *convexus* (Badawi et al. 2018) with 100% bootstrap support, in a group sister to *wVulC* and the *f*

407   element of *Armadillidium vulgare* (Leclercq et al. 2016) (Figure 2). Moreover, there were

408   multiple insertions carrying sequences similar to the cytoplasmic incompatibility genes *cifA* and

409   *cifB* genes (LePage et al. 2017; Beckmann et al. 2017; Lindsey et al. 2018). These *cifA*- and

410   *cifB*-like sequences probably represent nonfunctional pseudogenes, however, as they show

411   evidence of being broken up by insertions (Supplementary Figure 1).

412

413   Searching for candidate sex-determining genes

414

415   We identified ~6.04 x 10^6 21-mers as potentially male-specific, suggesting there is

416   approximately 6 Mb of male-specific sequence content in the genome. However, when we

417   isolated the raw sequencing reads containing those 21-mers and assembled them, we obtained

418   89.4 Mb of assembled sequences, suggesting the male-specific region may be as large as ~90

419   Mb, but still shares significant similarity with the X chromosome. Even if up to 90 Mb of

420   sequence is partially sex-linked, this represents just 1.7% of the genome, consistent with an

421   evolutionarily young Y chromosome in this species.

422

423   Of the initial 16 candidate Y-linked PCR markers designed from anonymous sequences, none

424   showed the expected pattern of male-specific amplification in our early tests (Supplementary

425   Table 4). This may be due to the highly repetitive nature of the *T. rathkei* genome, despite our

426   best efforts to target primers to non-repetitive sequences.

427

428   Because the candidate male-specific contigs were assembled from Illumina data only and thus

429   short and fragmented, we were unable to screen them for annotated candidate sex-determining

430   genes using the typical MAKER pipeline. However, we were able to identify open reading

431   frames (ORFs) and annotate them like transcripts using Trinotate (Bryant et al. 2017). Two

432   contigs in the male-specific assembly showed homology to the androgenic gland hormone

433   (AGH) gene upon annotation, suggesting there may be a Y-linked duplication of the AGH gene.

434   Therefore, we designed PCR primers specifically targeting one of the Y-linked AGH-like

435   sequences (AGHY1 on NODE_44048_length_535; F: 5'-ATTCTTGACTCTCCCCACGA-3'; R:

436   5'-TCTCCAACTACGATTTCGTTAATT-3'). These primers resulted in a PCR product of the

437   expected size (195 bp) in all male samples screened (7/7), but not in any of the female samples

438   (0/7), all of which were unrelated wild-caught individuals, confirming the male-specificity of this

439   AGH allele.

440

441   This AGH sequence could be either a male-specific duplication of the AGH gene, or a Y-linked

442   allele that has diverged from an X-linked copy (in other words, gametologs). To distinguish

443   between these possibilities, we examined the sequencing depth of these genes and of other

444   putatively single-copy genes (identified in the BUSCO analysis) in male and female sequencing

445   samples. If the male-specific AGH sequence is a gametolog of an X-linked sequence, we would

446   expect the total sequencing depth of all AGH sequences (putative autosomal and putative Y-

447   linked) to be the same in both the pooled male and pooled female samples, with the female

448   sample having a higher average sequencing depth for the putative X-linked AGH sequences

449   (since they would be homozygous for the X-linked gametologs, while males would be

450   hemizygous for the X-linked gametologs). If, on the other hand, the male-specific AGH

451   sequences are Y-linked duplicates, and not allelic to the other AGH sequences in our assembly,

452   we would expect the shared autosomal AGH sequences to have similar sequencing depth in

453   both male and female samples, and the combined sequencing depth of all AGH sequences

454   (putative autosomal and putative Y-linked) would be higher in the male sample. Our results

455   were consistent with the latter scenario, suggesting these are Y-linked duplicates rather than

456   gametologs (Figure 3). Note that sequencing depth of AGHY1 and AGHY2, though much lower

457   in the female sample than in the male sample, is still non-zero in the female sample, probably

458   because of ambiguously mapped reads due to high similarity between the Y-linked and

459   autosomal copies.

460

461   Because the male-specific AGH sequences were found only in our Illumina data, we were

462   unable to assemble them into long contigs, even after repeated attempts to assemble them

463   individually with different assemblers and parameter values (not shown); all these contigs were

464   ~600bp or less in length. Thus we are unable to determine whether these are complete

465   duplicates of the whole gene, or fragments. Nevertheless, a phylogenetic analysis suggests that

466   one of the Y-linked duplicates is a copy of the other, rather than an independent duplication of

467   an autosomal copy, and based on branch lengths they are as divergent from one another as

468   AGH orthologs in different species (Figure 3). In addition, these Y-linked copies seem to lack an

469   intron that is present in the autosomal copies (Figure 4), suggesting they may have originated

470   via retrotransposition.

471

472   We also find some evidence of additional autosomal duplicates of the androgenic gland

473   hormone (AGH) gene. Two contigs in the full assembly contained annotated transcripts with at

474   least partial homology to the expressed transcript identified as the AGH sequence, and a third

475   contained no annotated genes but still showed high sequence similarity to AGH in BLAST

476   searches. However, not all of the annotated exons in the first two copies matched the expressed

477   transcript, and there were unannotated portions of the same contigs that did show sequence

478   similarity to the transcript (Figure 4). Moreover, some of the matching portions of the assembled

479   contigs had less than 90% sequence identity to the expressed transcript, and analysis of the

480   sequencing depth of these regions reveals that one has very low coverage, suggesting it may

481   be an assembly artifact (see below). Thus, we cannot rule out the possibility that some of these

482   possible autosomal duplicates represent assembly and/or annotation artifacts. If they are real,

483   these autosomal duplicates appear to be specific to *Trachelipus*, occurring after its divergence

484   from *Porcellio* (Figure 3), but they may still be nonfunctional.

485

486   **Discussion**

487

488   Possible sex chromosome polymorphism and recent transition to XY sex chromosomes in

489   *Trachelipus rathkei*

490

491   We have shown that, at least in our upstate New York population, sex determination in the

492   terrestrial isopod *Trachelipus rathkei* is based on an XX/XY sex chromosome system. Two

493   independent lines of evidence support this finding: first, crosses between females and sex-

494   reversed neo-males yielded all female offspring (Table 2), consistent with an XX/XY system but

495   not a ZZ/ZW system (Becking et al. 2017); second, we have identified PCR primers that only

496   amplify a product in male samples, indicating the presence of a male-specific genomic region,

497   i.e., a Y chromosome.

498

499   Our findings run counter to a previously published study showing evidence of female

500   heterogamety in this species based on cytogenetics; in that study, female germ cells contained

501   one set of unpaired chromosomes (presumably, the Z and W sex chromosomes), while male

502   germ cells did not (Mittal & Pahwa 1980). There are multiple possible explanations for this

503   contradiction. First, it is possible that the previous study incorrectly identified the species of

504   study specimens, as no information on identification is given in the paper; however, *T. rathkei* is

505   relatively easy to distinguish from other cosmopolitan terrestrial isopod species by its five pairs

506   of pleopodal "lungs" (most superficially similar species such as *Porcellio scaber* have only two

507   pairs; (Hatchett 1947; Shultz 2018)). In addition, that study was published before feminizing

508   *Wolbachia* was widely recognized in terrestrial isopods. It is therefore theoretically possible that

509   the females used in that study carried an XY genotype but were feminized by *Wolbachia*, while

510   the males in that study might have carried a YY genotype, perhaps resulting from a cross

511   between an XY father and a sex-reversed XY or YY mother which failed to transmit *Wolbachia*

512   (Becking et al. 2019). However, we found no evidence of sex-reversing *Wolbachia* in our *T.*

513   *rathkei* population.

514

515   Perhaps the most likely explanation is sex chromosome polymorphism. Indeed, this would not

516   be unprecedented, as sex determination in terrestrial isopods is thought to evolve rapidly

517   (Rigaud et al. 1997; Cordaux et al. 2011; Becking et al. 2017), and within-species sex

518   chromosome polymorphisms are documented in a few other species. For instance, two

519   subspecies of *Porcellio dilatatus*, *P. dilatatus dilatatus* and *P. dilatatus petiti* have XX/XY and

520   ZZ/ZW systems, respectively (Juchault & Legrand 1964; Legrand et al. 1974; Becking et al.

521   2017). In addition, multiple sex determining elements segregate in populations of the common

522   pillbug *Armadillidium vulgare* (Juchault et al. 1992), including a novel W chromosome that

523   resulted from the integration of an almost entire *Wolbachia* genome into the host genome

524   (Leclercq et al. 2016). Outside terrestrial isopods, sex chromosome polymorphisms are also

525   documented in a range of other arthropods and vertebrates (Rodrigues et al. 2013; Orzack et al.

526   1980; Franco et al. 1982; Ogata et al. 2008). *T. rathkei* is probably non-native in North America

527   where this study was conducted (Jass & Klausmeier 2000), as well as perhaps in India where

528   the prior study on cytogenetics was done (Mittal & Pahwa 1980). Given its cosmopolitan

529   distribution, and the fact that other terrestrial isopods have moderate to high levels of genetic

530   diversity (Romiguier et al. 2014), it might not be especially surprising for *T. rathkei* to harbor

531   multiple polymorphic sex-determining loci. Hopefully future follow-up work can further

532   characterize geographic variation in sex determination in this species.

533

534   Regardless of whether or not sex determination is polymorphic in *T. rathkei*, the sex

535   chromosomes in this species are likely evolutionarily young because it is nested within a clade

536   that mostly consists of ZZ/ZW species (Becking et al. 2019). In addition, the putative male-

537   specific region of its genome is relatively small, displaying only moderate divergence from

538   candidate gametologous sequences, similar to other terrestrial isopods examined so far (Chebbi

539   et al. 2019; Becking et al. 2019). Given that we found genomic evidence of a past association

540   with *Wolbachia* in this species and that infection by *Wolbachia* has been found in other *T.*

541   *rathkei* populations (Cordaux et al. 2012), this observation is consistent with the hypothesis that

542   transitions in sex determination mechanisms may be triggered by *Wolbachia* and other

543   endosymbionts that manipulate host reproduction (Rigaud et al. 1997; Cordaux et al. 2011). If

544   other populations of *T. rathkei* with different sex determination mechanisms can be identified, it

545   may be possible to leverage this system to further study the mechanisms and selective forces

546   influencing transitions in sex determination mechanisms. In addition, studies of sex

547   determination in a phylogenetic context involving other members of the family Trachelipodidae

548   would shed further light on the origins of the X and Y chromosomes in *T. rathkei*.

549

550   <u>Genome size, structure, and repetitive elements</u>

551

552   The draft genome assembly of *T. rathkei* is especially large, at around 5.2 Gb, with

553   approximately 29% GC content. The actual genome is likely to be even larger, given that ~25%

554   of the BUSCO arthropod orthologs were missing in our assembly. By comparison, genomes of

555   pillbugs in the genus *Armadillidium* tend to be smaller at around 1.2 - 2 Gb in size (Chebbi et al.

556   2019; Becking et al. 2019), but other terrestrial isopods have genomes ranging to over 8 Gb

557   (Gregory 2020), and other crustacean relatives such as amphipods also have large genomes

558   (Rees et al. 2007; Rivarola-Duarte et al. 2014; Kao et al. 2016), so *T. rathkei* is not out of the

559   ordinary for this group.

560

561   The *T. rathkei* genome contains a large proportion of repetitive elements, in particular

562   transposable elements (Figure 1). The most common transposable element families are LINEs,

563   DNA elements, and LTRs, similar to *Armadillidium vulgare* and *A. nasatum* (Chebbi et al. 2019;

564   Becking et al. 2019). The distribution of divergence values, with a single mode around 7-10%

565   divergence, suggests that most repeat families expanded around the same time as in *A. vulgare*

566   and *A. nasatum*; however, unlike in *A. vulgare*, *T. rathkei* shows no evidence of a second more

567   recent burst in DNA element activity. Simple repeats also comprise a substantial portion of the

568   genome; even manually looking through the assembled contigs reveals a high abundance of

569   (TA)x repeats.

570

571   *Wolbachia* insertions

572

573   We found many contigs with high similarity to the *Wolbachia* genome (Supplementary Table 3),

574   even though we were unable to detect current *Wolbachia* infections in our population using

575   PCR. This is not surprising given that *Wolbachia* is relatively common in terrestrial isopods and

576   arthropods in general (Cordaux et al. 2012; Pascar & Chandler 2018; Medina et al. 2019), has

577   been found in other populations of *T. rathkei* (Cordaux et al. 2012), and that horizontal transfers

578   of *Wolbachia* DNA into host genomes is also common (Dunning Hotopp 2011). These

579   *Wolbachia* insertions seem to be most closely related to strain *wCon* from *C. convexus*, which

580   does not induce male-to-female sex reversal, but rather causes cytoplasmic incompatibility

581   (Moret et al. 2001; Badawi et al. 2018). Consistent with this, we find that several copies of the

582   cytoplasmic incompatibility genes *cifA* and *cifB* among these insertions in the *T. rathkei* genome

583   (Supplementary Figure 1). Thus, these past *T. rathkei* infections may have caused cytoplasmic

584   incompatibility rather than host sex reversal, but it is possible that the same *Wolbachia* strain

585   may have multiple effects on host phenotypes.

586

587   <u>Candidate sex-determining genes and repeated duplication of the AGH gene</u>

588

589   Male differentiation in terrestrial isopods is controlled by the androgenic gland hormone, AGH.

590   AGH is a peptide hormone similar in structure to insulin, and is secreted by the androgenic

591   gland (Martin et al. 1999). AGH expression is sufficient to transform juvenile female isopods into

592   fertile males (Martin et al. 1999). Presumably, in wild-type males, the primary sex-determining

593   signal triggers the differentiation of the androgenic glands during development, which then

594   secretes AGH. Interestingly, the draft genome of *T. rathkei* contains multiple AGH-like

595   sequences. While some of these may be assembly artifacts, there is evidence three at least

596   partial Y-linked sequences, of which one was confirmed by PCR to be male-specific. These

597   duplications seem to be specific to *T. rathkei* (Figure 5), though other members of the genus

598   *Trachelipus* or the family Trachelipodidae have yet to be examined. Consistent with this, a past

599   study found no evidence of any expressed AGH duplications in other terrestrial isopod species

600   except *Porcellio gallicus* (Cerveau et al. 2014).

601

602   In many other species, novel sex chromosomes have arisen via duplication of a sex-determining

603   gene. For instance, duplicates of the vertebrate gene *Dmrt1* have evolved into master sex-

604   determining signals on the W and Y chromosomes, respectively, in the frog *Xenopus laevis*

605   (Yoshimoto et al. 2008) and the medaka *Oryzias latipes* (Nanda et al. 2002; Matsuda et al.

606   2002, 2007), while a Y-linked duplicate of the anti-Müllerian hormone gene is a candidate

607   master sex-determining gene in the teleost fish *Odontesthes hatcheri* (Hattori et al. 2012). The

608   presence of Y-linked AGH copies in *T. rathkei*, and no other obvious open reading frames

609   homologous to known sex determination or sex differentiation genes, makes these genes

610   obvious candidates for the master male-determining signal in *T. rathkei*. Unfortunately, we were

611   unable to assemble full copies of these Y-linked AGH homologs because they only showed up

612   in our Illumina data, not in our low-coverage long read data. Future deep sequencing using long

613   reads should further clarify the molecular evolution of these genes. In addition, expression

614   studies should determine which of these genes are expressed, in what tissues, and at what

615   stages.

616

617   If one of these AGH duplicates is indeed the master sex-determining signal in *T. rathkei*, this

618   could support the idea of a transition in sex determination mechanisms triggered by *Wolbachia*

619   and other reproductive endosymbionts (Rigaud et al. 1997; Cordaux et al. 2011). AGH may be a

620   primary molecular target by which *Wolbachia* causes male-to-female sex reversal in isopod

621   hosts, as injection with AGH does not cause female-to-male sex reversal in individuals infected

622   by *Wolbachia* (Juchault & Legrand 1985; Cordaux & Gilbert 2017). Thus, *Wolbachia* may

623   impose strong selection (via female-biased sex ratios) favoring duplication and divergence of

624   the AGH gene to escape this sex reversal. The finding that *T. rathkei* may have historically been

625   infected by a *Wolbachia* strain causing cytoplasmic incompatibility does not necessarily negate

626   this finding, as the same *Wolbachia* strain can have multiple effects on host phenotypes.

627

628   Conclusions

629

630   We have shown that the terrestrial isopod *Trachelipus rathkei* uses an XX/XY sex chromosome

631   system, at least in upstate New York, in contrast to a past cytogenetic study suggesting a

632   ZZ/ZW mechanism (Mittal & Pahwa 1980). In line with this, whole-genome sequencing and

633   follow-up PCRs demonstrate the existence of male-specific, Y-linked copies of the androgenic

634   gland hormone gene in this species. These findings highlight the role of gene duplication in the

635   evolution of sex chromosomes and support the possibility that reproductive endosymbionts like

636   *Wolbachia* may favor evolutionary transitions in host sex determination mechanisms.

637

638

639

640 **Acknowledgments**

641

643

644 **Table 1.** *Trachelipus rathkei* DNA samples used for genomic sequencing.

Sample name	Sex	Notes	Platform/ read length	Total data (Gb)	Ref.	Accession Number
M-pool	M	Pool of three brothers, lab-reared	Illumina, 2x100	191.6	This study	SRR11797365
F-pool	F	Pool of three sisters, lab-reared	Illumina, 2x100	191.9	This study	SRR11797364
M2	M	One male, lab-reared	Illumina, 2x250	22.3	This study	SRR11797360
M5	M	One male, lab-reared	Illumina, 2x250	20.6	This study	SRR11797359
M6	M	One male, lab-reared	Illumina, 2x250	23.6	This study	SRR11797358
F3	F	One female, lab-reared	Illumina, 2x250	131.5	This study	SRR11797357
F4	F	One female, lab-reared	Illumina, 2x250	26.8	This study	SRR11797356
Wild-M	M	One male, wild-caught	Illumina, 2x100	14.1	(Chandler et al. 2015)	SRR4000567
Wild-F	F	One female, wild-caught	Illumina, 2x100	14.4	(Chandler et al. 2015)	SRR4000573
PB-F1	F	Pool of three sisters, lab-reared	PacBio	22.2	This study; (Peccoud et al. 2017)	SRR11797355
PB-F2	F	Pool of three sisters, lab-reared	PacBio	1.3	This study; (Peccoud et al. 2017)	SRR11797354
PB-M	M	Pool of three brothers, lab-reared	PacBio	3.9	This study; (Peccoud et al. 2017)	SRR11797353
ONT-F	F	One female	ONT	0.4	This study	SRR11797363
ONT-M1	M	One male	ONT	3.9	This study	SRR11797362
ONT-M2	M	One male	ONT	3.3	This study	SRR11797361

645

646

647   **Table 2.** Sex ratios from crosses between putative neo-males (juvenile females implanted with

648   an androgenic gland) and females. *p* gives the probability of the observed or a more extreme

649   result (pooling the data for each neo-male) under a balanced sex ratio (i.e., assuming each

650   individual offspring has an equal probability of being male or female).

651

Neo-male	Female	Number of female offspring	Number of male offspring	*p*
D-4-7	1	23	0	5.7×10^{-14}
	2	14	0	
	3	7	0	
G-4-22	1	16	8	0.13
	2	17	14	
	3	13	13	
F-4-9	1	36	0	1.7×10^{-18}
	2	23	0	
10-8	1	24	35	0.19
	2	40	40	
AGS169-2	1	45	0	2.1×10^{-25}
	2	37	0	

652

653

654   **Table 3.** Assembly statistics for the *T. rathkei* draft genome.

Total length	5,181,251,014 bp
Number of contigs	421,784
N50	39,761 bp
GC content	29.0%
Complete BUSCO genes	533 single copy (51.9%); 39 duplicated (3.7%)
Fragmented BUSCO genes	203 (19.0%)
Missing BUSCO genes	271 (25.4%)

655

656

657   **Figure Legends**

658

659   **Figure 1.** Distribution of divergence levels for repetitive elements in the *Trachelipus rathkei*

660   genome.

661

662   **Figure 2.** Phylogenetic tree showing the position of candidate horizontally transferred

663   *Wolbachia* segments in the *Trachelipus rathkei* genome. Numbers by nodes indicate bootstrap

664   support. The tree was generated by concatenating all candidate *Wolbachia* insertions in *T.*

665   *rathkei* longer than 1000 bp, along with the best-matching regions in the reference *Wolbachia*

666   genomes, aligning with MUSCLE v. 3.8.31 (Edgar 2004), filtering alignments with trimal v.

667   1.2rev59 (Capella-Gutiérrez et al. 2009), selecting a model using ModelTest-NG v. 0.1.6

668   (Darriba et al. 2020), and running the analysis in RAxML-NG v. 0.9.0 (Stamatakis 2014) with

669   100 bootstrap replicates.

670

671   **Figure 3.** Distribution of sequencing depth for single-copy BUSCO genes in male and female

672   Illumina sequencing datasets (M-pool and F-pool). Labeled dots indicate the sequencing depth

673   for the different AGH copies in each sample.

674

675   **Figure 4.** Possible duplicates of the androgenic gland hormone gene in the *Trachelipus rathkei*

676   genome. The green bars represent the sequence of the expressed AGH sequence, assembled

677   from previously available transcriptome data. Gray bars represent contigs in the draft genome

678   assembly, and the pink bars on contigs represent annotated exons. Purple segments

679   connecting portions of the transcript to portions of contigs represent BLAST hits. The

680   incongruence between annotated exons and BLAST matches between the transcript and

681   contigs suggests the annotation still contains some errors.

682

683   **Figure 5.** Phylogenetic tree showing relationships among AGH sequences from terrestrial

684   isopods. AGH2 and AGHY3 are missing from this phylogeny because those sequences were

685   omitted because of their short length. The tree was generated using all AGH-like sequences

686   from *T. rathkei* of at least 100 bp, along with reference AGH nucleotide sequences from other

687   species, aligning them with MUSCLE v. 3.8.31 (Edgar 2004), selecting a model using

688   ModelTest-NG v. 0.1.6 (Darriba et al. 2020), and running the analysis in RAxML-NG v. 0.9.0

689   (Stamatakis 2014) with 100 bootstrap replicates.

690

691

692

693

694 **References**

695

696 Badawi M, Moumen B, Giraud I, Grève P, Cordaux R. 2018. Investigating the molecular genetic
697 basis of cytoplasmic sex determination caused by Wolbachia endosymbionts in terrestrial
698 isopods. Genes. 9:290. doi: 10.3390/genes9060290.

699 Bankevich A et al. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to
700 Single-Cell Sequencing. Journal of Computational Biology. 19:455–477. doi:
701 10.1089/cmb.2012.0021.

702 Bao Z, Eddy SR. 2002. Automated de novo identification of repeat sequence families in
703 sequenced genomes. Genome Res. 12:1269–1276. doi: 10.1101/gr.88502.

704 Becking T et al. 2017. Diversity and evolution of sex determination systems in terrestrial
705 isopods. Sci Rep. 7:1084. doi: 10.1038/s41598-017-01195-4.

706 Becking T et al. 2019. Sex chromosomes control vertical transmission of feminizing Wolbachia
707 symbionts in an isopod. PLOS Biology. 17:e3000438. doi: 10.1371/journal.pbio.3000438.

708 Beckmann JF, Ronau JA, Hochstrasser M. 2017. A Wolbachia deubiquitylating enzyme induces
709 cytoplasmic incompatibility. Nat Microbiol. 2:17007. doi: 10.1038/nmicrobiol.2017.7.

710 Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids
711 Res. 27:573–580. doi: 10.1093/nar/27.2.573.

712 Bergero R, Charlesworth D. 2009. The evolution of restricted recombination in sex
713 chromosomes. Trends in Ecology & Evolution. 24:94–102. doi: 10.1016/j.tree.2008.09.010.

714 Blaser O, Grossen C, Neuenschwander S, Perrin N. 2013. Sex-chromosome turnovers induced
715 by deleterious mutation load. Evolution. 67:635–645. doi: 10.1111/j.1558-5646.2012.01810.x.

716 Blaser O, Neuenschwander S, Perrin N. 2014. Sex-Chromosome Turnovers: The Hot-Potato
717 Model. The American Naturalist. 183:140–146. doi: 10.1086/674026.

718 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence
719 data. Bioinformatics. 30:2114–2120. doi: 10.1093/bioinformatics/btu170.

720 Bouchon D, Rigaud T, Juchault P. 1998. Evidence for widespread Wolbachia infection in isopod
721 crustaceans: molecular identification and host feminization. Proceedings of the Royal Society of
722 London. Series B: Biological Sciences. 265:1081–1090. doi: 10.1098/rspb.1998.0402.

723 Braig HR, Zhou W, Dobson SL, O'Neill SL. 1998. Cloning and Characterization of a Gene
724 Encoding the Major Surface Protein of the Bacterial Endosymbiont Wolbachia pipientis. J
725 Bacteriol. 180:2373–2378.

726 Bryant DM et al. 2017. A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification
727 of Limb Regeneration Factors. Cell Reports. 18:762–776. doi: 10.1016/j.celrep.2016.12.063.

728   Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND.
729   Nature Methods. 12:59–60. doi: 10.1038/nmeth.3176.

730   Bushnell B, Rood J, Singer E. 2017. BBMerge – Accurate paired shotgun read merging via
731   overlap. PLOS ONE. 12:e0185056. doi: 10.1371/journal.pone.0185056.

732   Cantarel BL et al. 2008. MAKER: An easy-to-use annotation pipeline designed for emerging
733   model organism genomes. Genome Res. 18:188–196. doi: 10.1101/gr.6743907.

734   Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated
735   alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25:1972–1973. doi:
736   10.1093/bioinformatics/btp348.

737   Cerveau N, Bouchon D, Bergès T, Grève P. 2014. Molecular evolution of the androgenic
738   hormone in terrestrial isopods. Gene. 540:71–77. doi: 10.1016/j.gene.2014.02.024.

739   Chandler CH, Badawi M, Moumen B, Grève P, Cordaux R. 2015. Multiple Conserved
740   Heteroplasmic Sites in tRNA Genes in the Mitochondrial Genomes of Terrestrial Isopods
741   (Oniscidea). G3: Genes, Genomes, Genetics. 5:1317–1322. doi: 10.1534/g3.115.018283.

742   Chandler JC, Elizur A, Ventura T. 2018. The decapod researcher's guide to the galaxy of sex
743   determination. Hydrobiologia. 825:61–80. doi: 10.1007/s10750-017-3452-4.

744   Chandler JC, Fitzgibbon QP, Smith G, Elizur A, Ventura T. 2017. Y-linked iDmrt1 paralogue
745   (iDMY) in the Eastern spiny lobster, Sagmariasus verreauxi: The first invertebrate sex-linked
746   Dmrt. Developmental Biology. 430:337–345. doi: 10.1016/j.ydbio.2017.08.031.

747   Charlesworth B, Charlesworth D. 2000. The degeneration of Y chromosomes Charlesworth, B.
748   & Harvey, PH, editors. Phil. Trans. R. Soc. Lond. B. 355:1563–1572. doi:
749   10.1098/rstb.2000.0717.

750   Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex
751   chromosomes. Heredity. 95:118–128. doi: 10.1038/sj.hdy.6800697.

752   Chebbi MA et al. 2019. The Genome of Armadillidium vulgare (Crustacea, Isopoda) Provides
753   Insights into Sex Chromosome Evolution in the Context of Cytoplasmic Sex Determination. Mol
754   Biol Evol. 36:727–741. doi: 10.1093/molbev/msz010.

755   Chevreux B, Wetter T, Suhai S, others. 1999. Genome sequence assembly using trace signals
756   and additional sequence information. In: German conference on bioinformatics.Vol. 99 Citeseer
757   pp. 45–56.

758   Cioffi MB, Liehr T, Trifonov V, Molina WF, Bertollo L a. C. 2013. Independent Sex Chromosome
759   Evolution in Lower Vertebrates: A Molecular Cytogenetic Overview in the Erythrinidae Fish
760   Family. CGR. 141:186–194. doi: 10.1159/000354039.

761   Conover DO, Kynard BE. 1981. Environmental Sex Determination: Interaction of Temperature
762   and Genotype in a Fish. Science. 213:577–579. doi: 10.1126/science.213.4507.577.

763   Cordaux R et al. 2012. Widespread Wolbachia infection in terrestrial isopods and other
764   crustaceans. Zookeys. 123–131. doi: 10.3897/zookeys.176.2284.

765 Cordaux R, Bouchon D, Grève P. 2011. The impact of endosymbionts on the evolution of host
766 sex-determination mechanisms. Trends in Genetics. 27:332–341. doi:
767 10.1016/j.tig.2011.05.002.

768 Cordaux R, Gilbert C. 2017. Evolutionary Significance of Wolbachia-to-Animal Horizontal Gene
769 Transfer: Female Sex Determination and the f Element in the Isopod Armadillidium vulgare.
770 Genes (Basel). 8:186. doi: 10.3390/genes8070186.

771 Cordaux R, Michel-Salzat A, Frelon-Raimond M, Rigaud T, Bouchon D. 2004. Evidence for a
772 new feminizing Wolbachia strain in the isopod Armadillidium vulgare: evolutionary implications.
773 Heredity. 93:78–84. doi: 10.1038/sj.hdy.6800482.

774 Darriba D et al. 2020. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and
775 Protein Evolutionary Models Crandall, K, editor. Molecular Biology and Evolution. 37:291–294.
776 doi: 10.1093/molbev/msz189.

777 van Doorn GS, Kirkpatrick M. 2007. Turnover of sex chromosomes induced by sexual conflict.
778 Nature. 449:909–912. doi: 10.1038/nature06178.

779 Dunning Hotopp JC. 2011. Horizontal gene transfer between bacteria and animals. Trends in
780 Genetics. 27:157–163. doi: 10.1016/j.tig.2011.01.005.

781 Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and
782 space complexity. BMC Bioinformatics. 5:113. doi: 10.1186/1471-2105-5-113.

783 Emerson JJ, Kaessmann H, Betrán E, Long M. 2004. Extensive Gene Traffic on the Mammalian
784 X Chromosome. Science. 303:537–540. doi: 10.1126/science.1090042.

785 Folmer O, Black MB, Hoeh WR, Lutz R, Vrijenhoek RC. 1994. DNA primers for amplification of
786 mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular
787 marine biology and biotechnology. 3:294–299.

788 Franco MG, Rubini PG, Vecchi M. 1982. Sex-determinants and their distribution in various
789 populations of Musca domestica L. of Western Europe. Genetics Research. 40:279–293. doi:
790 10.1017/S0016672300019157.

791 Fussey GD. 1984. The distribution of the two forms of the woodlouse Trichoniscus pusillus
792 Brandt (Isopoda: Oniscoidea) in the British Isles: a reassessment of geographic
793 parthenogenesis. Biol J Linn Soc. 22:309–321. doi: 10.1111/j.1095-8312.1984.tb01681.x.

794 Gamble T, Geneva AJ, Glor RE, Zarkower D. 2014. Anolis Sex Chromosomes Are Derived from
795 a Single Ancestral Pair. Evolution. 68:1027–1041. doi: 10.1111/evo.12328.

796 Gregory TR. 2020. Animal Genome Size Database. http://www.genomesize.com/ (Accessed
797 January 21, 2020).

798 Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome
799 assemblies. Bioinformatics. 29:1072–1075. doi: 10.1093/bioinformatics/btt086.

800 Haas B, Papanicolaou A. 2016. TransDecoder (find coding regions within transcripts).
801 http://transdecoder.github.io.

802   Hatchett SP. 1947. Biology of the Isopoda of Michigan. Ecological Monographs. 17:47–80. doi:
803   10.2307/1948613.

804   Hattori RS et al. 2012. A Y-linked anti-Müllerian hormone duplication takes over a critical role in
805   sex determination. PNAS. 109:2955–2959. doi: 10.1073/pnas.1018392109.

806   Heimpel GE, de Boer JG. 2008. Sex Determination in the Hymenoptera. Annual Review of
807   Entomology. 53:209–230. doi: 10.1146/annurev.ento.53.103106.093441.

808   Janzen FJ, Phillips PC. 2006. Exploring the evolution of environmental sex determination,
809   especially in reptiles. Journal of Evolutionary Biology. 19:1775–1784. doi: 10.1111/j.1420-
810   9101.2006.01138.x.

811   Jass J, Klausmeier B. 2000. Endemics and Immigrants: North American Terrestrial Isopods
812   (Isopoda, Oniscidea) North of Mexico. Crustaceana. 73:771–799.

813   Jeffries DL et al. 2018. A rapid rate of sex-chromosome turnover and non-random transitions in
814   true frogs. Nat Commun. 9:4088. doi: 10.1038/s41467-018-06517-2.

815   Johnson C. 1986. Parthenogenetic Reproduction in the Philosciid Isopod, Ocelloscia floridana
816   (Van Name, 1940). Crustaceana. 51:123–132.

817   Juchault P, Legrand J. 1964. Démonstration de l'homogamétie femelle par croisement de deux
818   femelles génétiques chez les Oniscoïdes Porcellio dilatatus et Helleria brevi-cornis. Comptes
819   Rendus de l'Académie des Sciences de Paris. 258:2685–2686.

820   Juchault P, Legrand JJ. 1985. Contribution à l'étude du mécanisme de l'état réfractaire à
821   l'hormone androgène chez les Armadillidium vulgare Latr, (crustacé, isopode, oniscoïde)
822   hébergeant une bactérie féminisante. General and Comparative Endocrinology. 60:463–467.
823   doi: 10.1016/0016-6480(85)90082-6.

824   Juchault P, Rigaud T. 1995. Evidence for female heterogamety in two terrestrial crustaceans
825   and the problem of sex chromosome evolution in isopods. Heredity. 75:466–471. doi:
826   10.1038/hdy.1995.163.

827   Juchault P, Rigaud T, Mocquard J-P. 1992. Evolution of sex-determining mechanisms in a wild
828   population of Armadillidium vulgare Latr. (Crustacea, Isopoda): competition between two
829   feminizing parasitic sex factors. Heredity. 69:382–390. doi: 10.1038/hdy.1992.138.

830   Kageyama D, Nishimura G, Hoshizaki S, Ishikawa Y. 2002. Feminizing Wolbachia in an insect,
831   Ostrinia furnacalis (Lepidoptera: Crambidae). Heredity. 88:444–449. doi:
832   10.1038/sj.hdy.6800077.

833   Kao D et al. 2016. The genome of the crustacean Parhyale hawaiensis, a model for animal
834   development, regeneration, immunity and lignocellulose digestion VijayRaghavan, K, editor.
835   eLife. 5:e20062. doi: 10.7554/eLife.20062.

836   Kokot M, Długosz M, Deorowicz S. 2017. KMC 3: counting and manipulating k-mer statistics.
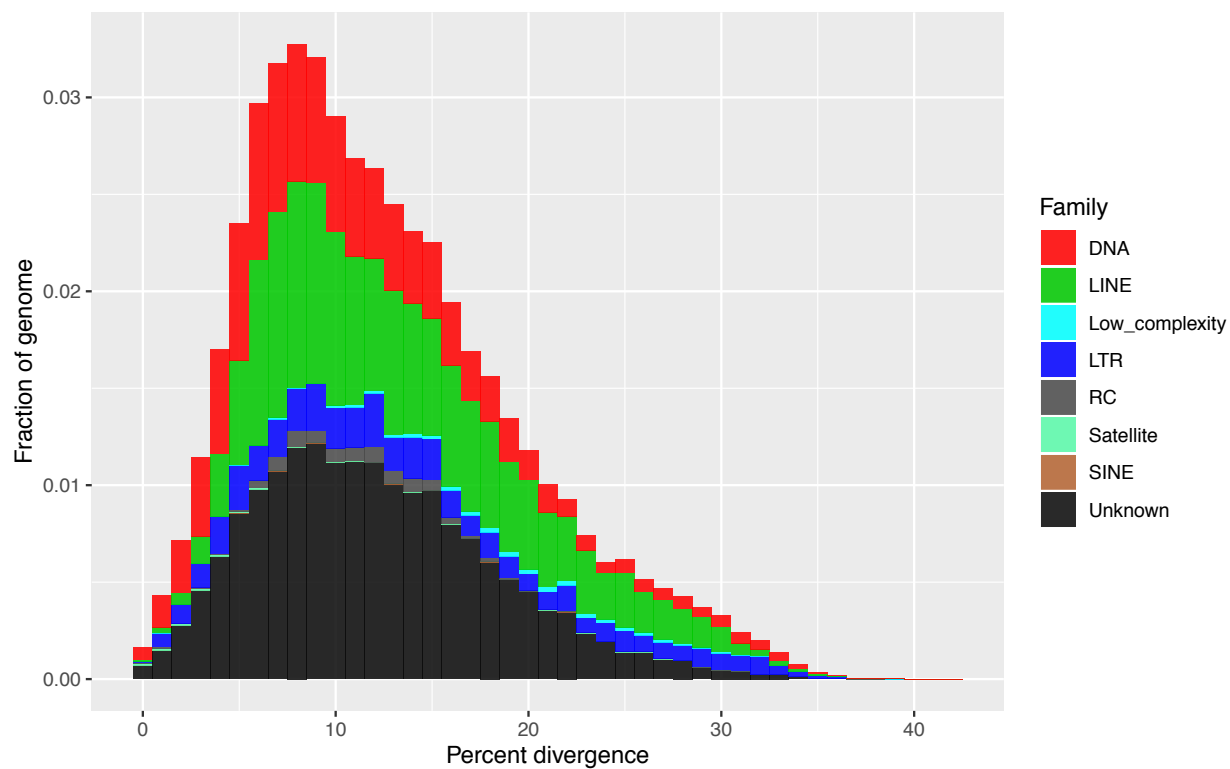837   Bioinformatics. 33:2759–2761. doi: 10.1093/bioinformatics/btx304.

838   Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics. 5:59. doi: 10.1186/1471-
839   2105-5-59.

840   Laetsch DR, Blaxter ML. 2017. BlobTools: Interrogation of genome assemblies. F1000Res.
841   6:1287. doi: 10.12688/f1000research.12232.1.

842   Leclercq S et al. 2016. Birth of a W sex chromosome by horizontal transfer of Wolbachia
843   bacterial symbiont genome. Proc Natl Acad Sci U S A. 113:15036–15041. doi:
844   10.1073/pnas.1608979113.

845   Legrand J, Juchault P, Artault J, Mocquard J, Picaud J. 1974. Le statut systématique de la
846   «forme» petiti Vandel de Porcellio dilatatus Brandt, récoltée à l'île Saint-Honorat (Alpes-
847   Maritimes). Critères morphologiques, génétiques et physiologiques. Bull. Soc. Zool. Fr. 99:461–
848   471.

849   LePage DP et al. 2017. Prophage WO genes recapitulate and enhance Wolbachia -induced
850   cytoplasmic incompatibility. Nature. 543:243–247. doi: 10.1038/nature21391.

851   Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
852   arXiv:1303.3997 [q-bio]. http://arxiv.org/abs/1303.3997 (Accessed April 27, 2020).

853   Lindsey ARI et al. 2018. Evolutionary Genetics of Cytoplasmic Incompatibility Genes cifA and
854   cifB in Prophage WO of Wolbachia. Genome Biol Evol. 10:434–451. doi: 10.1093/gbe/evy012.

855   Martin G et al. 1999. The structure of a glycosylated protein hormone responsible for sex
856   determination in the isopod, Armadillidium vulgare. European Journal of Biochemistry. 262:727–
857   736. doi: 10.1046/j.1432-1327.1999.00442.x.

858   Matsuda M et al. 2007. DMY gene induces male development in genetically female (XX)
859   medaka fish. PNAS. 104:3865–3870. doi: 10.1073/pnas.0611707104.

860   Matsuda M et al. 2002. DMY is a Y-specific DM-domain gene required for male development in
861   the medaka fish. Nature. 417:559–563. doi: 10.1038/nature751.

862   Medina P, Russell SL, Corbett-Detig R. 2019. Deep data mining reveals variable abundance
863   and distribution of microbial reproductive manipulators within and among diverse host species.
864   bioRxiv. 679837. doi: 10.1101/679837.

865   Meisel RP et al. 2016. Is Multifactorial Sex Determination in the House Fly, Musca domestica
866   (L.), Stable Over Time? J Hered. 107:615–625. doi: 10.1093/jhered/esw051.

867   Mittal OP, Pahwa S. 1980. Cytogenetical studies on two terrestrial isopods (Crustacea),
868   Philoscia lodnensis Ramakrishna and Porcellio rathkei Brandt. Cytobios. 29:17–24.

869   Moret Y, Juchault P, Rigaud T. 2001. Wolbachia endosymbiont responsible for cytoplasmic
870   incompatibility in a terrestrial crustacean: effects in natural and foreign hosts. Heredity. 86:325–
871   332. doi: 10.1046/j.1365-2540.2001.00831.x.

872   Morgulis A et al. 2008. Database indexing for production MegaBLAST searches. Bioinformatics.
873   24:1757–1764. doi: 10.1093/bioinformatics/btn322.

874 Myosho T, Takehana Y, Hamaguchi S, Sakaizumi M. 2015. Turnover of Sex Chromosomes in
875 Celebensis Group Medaka Fishes. G3: Genes, Genomes, Genetics. 5:2685–2691. doi:
876 10.1534/g3.115.021543.

877 Nanda I et al. 2002. A duplicated copy of DMRT1 in the sex-determining region of the Y
878 chromosome of the medaka, Oryzias latipes. PNAS. 99:11778–11783. doi:
879 10.1073/pnas.182314699.

880 Negri I, Pellecchia M, Mazzoglio P.J, Patetta A, Alma A. 2006. Feminizing Wolbachia in
881 Zyginidia pullula (Insecta, Hemiptera), a leafhopper with an XX/X0 sex-determination system.
882 Proceedings of the Royal Society B: Biological Sciences. 273:2409–2416. doi:
883 10.1098/rspb.2006.3592.

884 Ogata M, Hasegawa Y, Ohtani H, Mineyama M, Miura I. 2008. The ZZ/ZW sex-determining
885 mechanism originated twice and independently during evolution of the frog, *Rana rugosa*.
886 Heredity. 100:92–99. doi: 10.1038/sj.hdy.6801068.

887 Orzack SH, Sohn JJ, Kallman KD, Levin SA, Johnston R. 1980. Maintenance of the Three Sex
888 Chromosome Polymorphism in the Platyfish, Xiphophorus maculatus. Evolution. 34:663–672.
889 doi: 10.2307/2408020.

890 Ospina-Álvarez N, Piferrer F. 2008. Temperature-Dependent Sex Determination in Fish
891 Revisited: Prevalence, a Single Sex Ratio Response Pattern, and Possible Effects of Climate
892 Change. PLOS ONE. 3:e2837. doi: 10.1371/journal.pone.0002837.

893 Pascar J, Chandler CH. 2018. A bioinformatics approach to identifying Wolbachia infections in
894 arthropods. PeerJ. 6:e5486. doi: 10.7717/peerj.5486.

895 Peccoud J et al. 2017. Untangling Heteroplasmy, Structure, and Evolution of an Atypical
896 Mitochondrial Genome by PacBio Sequencing. Genetics. 207:269–280. doi:
897 10.1534/genetics.117.203380.

898 Pennell MW, Mank JE, Peichel CL. 2018. Transitions in sex determination and sex
899 chromosomes across vertebrate species. Mol Ecol. 27:3950–3963. doi: 10.1111/mec.14540.

900 Pickup J, Ironside JE. 2018. Multiple origins of parasitic feminization: thelygeny and
901 intersexuality in beach-hoppers are caused by paramyxid parasites, not microsporidia.
902 Parasitology. 145:408–415. doi: 10.1017/S0031182017001597.

903 Potrzebowski L et al. 2008. Chromosomal Gene Movements Reflect the Recent Origin and
904 Biology of Therian Sex Chromosomes. PLOS Biology. 6:e80. doi:
905 10.1371/journal.pbio.0060080.

906 Price AL, Jones NC, Pevzner PA. 2005. De novo identification of repeat families in large
907 genomes. Bioinformatics. 21:i351–i358. doi: 10.1093/bioinformatics/bti1018.

908 Pryszcz LP, Gabaldón T. 2016. Redundans: an assembly pipeline for highly heterozygous
909 genomes. Nucleic Acids Res. 44:e113–e113. doi: 10.1093/nar/gkw294.

910 Rees DJ, Dufresne F, Glémet H, Belzile C. 2007. Amphipod genome sizes: first estimates for
911 Arctic species reveal genomic giants. Genome. 50:151–158. doi: 10.1139/G06-155.

912   Rice WR. 1996. Evolution of the Y Sex Chromosome in Animals. BioScience. 46:331–343. doi:
913   10.2307/1312947.

914   Rigaud T, Juchault P, Mocquard J-P. 1997. The evolution of sex determination in isopod
915   crustaceans. BioEssays. 19:409–416. doi: 10.1002/bies.950190508.

916   Rivarola-Duarte L et al. 2014. A first Glimpse at the genome of the Baikalian amphipod
917   Eulimnogammarus verrucosus. Journal of Experimental Zoology Part B: Molecular and
918   Developmental Evolution. 322:177–189. doi: 10.1002/jez.b.22560.

919   Rodrigues N, Betto-Colliard C, Jourdan-Pineau H, Perrin N. 2013. Within-population
920   polymorphism of sex-determination systems in the common frog (Rana temporaria). Journal of
921   Evolutionary Biology. 26:1569–1577. doi: 10.1111/jeb.12163.

922   Romiguier J et al. 2014. Comparative population genomics in animals uncovers the
923   determinants of genetic diversity. Nature. 515:261–263. doi: 10.1038/nature13685.

924   Ross JA, Urton JR, Boland J, Shapiro MD, Peichel CL. 2009. Turnover of Sex Chromosomes in
925   the Stickleback Fishes (Gasterosteidae). PLOS Genetics. 5:e1000391. doi:
926   10.1371/journal.pgen.1000391.

927   Shultz JW. 2018. A guide to the identification of the terrestrial Isopoda of Maryland, U.S.A.
928   (Crustacea). ZooKeys. 801:207–228. doi: 10.3897/zookeys.801.24146.

929   Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO:
930   assessing genome assembly and annotation completeness with single-copy orthologs.
931   Bioinformatics. 31:3210–3212. doi: 10.1093/bioinformatics/btv351.

932   Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
933   large phylogenies. Bioinformatics. 30:1312–1313. doi: 10.1093/bioinformatics/btu033.

934   Stanke M et al. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids
935   Res. 34:W435–W439. doi: 10.1093/nar/gkl200.

936   Sun LV et al. 2001. Determination of Wolbachia Genome Size by Pulsed-Field Gel
937   Electrophoresis. Journal of Bacteriology. 183:2219–2225. doi: 10.1128/JB.183.7.2219-
938   2225.2001.

939   Tarailo-Graovac M, Chen N. 2009. Using RepeatMasker to Identify Repetitive Elements in
940   Genomic Sequences. Current Protocols in Bioinformatics. 25:4–10. doi:
941   10.1002/0471250953.bi0410s25.

942   Terry RS, Smith JE, Dunn AM. 1998. Impact of a Novel, Feminising Microsporidium on its
943   Crustacean Host. Journal of Eukaryotic Microbiology. 45:497–501. doi: 10.1111/j.1550-
944   7408.1998.tb05106.x.

945   Tingley GA, Anderson RM. 1986. Environmental sex determination and density-dependent
946   population regulation in the entomogenous nematode Romanomermis culicivorax. Parasitology.
947   92:431–449. doi: 10.1017/S0031182000064192.

948 Traut W. 1994. Sex determination in the fly Megaselia scalaris, a model system for primary
949 steps of sex chromosome evolution. Genetics. 136:1097–1104.

950 Untergasser A et al. 2012. Primer3—new capabilities and interfaces. Nucleic Acids Res.
951 40:e115–e115. doi: 10.1093/nar/gks596.

952 Vandeputte M, Dupont-Nivet M, Chavanne H, Chatain B. 2007. A Polygenic Hypothesis for Sex
953 Determination in the European Sea Bass Dicentrarchus labrax. Genetics. 176:1049–1057. doi:
954 10.1534/genetics.107.072140.

955 Verhulst EC, Beukeboom LW, Zande L van de. 2010. Maternal Control of Haplodiploid Sex
956 Determination in the Wasp Nasonia. Science. 328:620–623. doi: 10.1126/science.1185805.

957 Veyrunes F et al. 2008. Bird-like sex chromosomes of platypus imply recent origin of mammal
958 sex chromosomes. Genome Res. 18:965–973. doi: 10.1101/gr.7101908.

959 Vicoso B, Bachtrog D. 2015. Numerous Transitions of Sex Chromosomes in Diptera. PLOS
960 Biology. 13:e1002078. doi: 10.1371/journal.pbio.1002078.

961 Walker BJ et al. 2014. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection
962 and Genome Assembly Improvement. PLOS ONE. 9:e112963. doi:
963 10.1371/journal.pone.0112963.

964 Wang JR, Holt J, McMillan L, Jones CD. 2018. FMLRC: Hybrid long read error correction using
965 an FM-index. BMC Bioinformatics. 19:50. doi: 10.1186/s12859-018-2051-3.

966 Werren John H., Zhang Wan, Guo Li Rong. 1995. Evolution and phylogeny of Wolbachia:
967 reproductive parasites of arthropods. Proceedings of the Royal Society of London. Series B:
968 Biological Sciences. 261:55–63. doi: 10.1098/rspb.1995.0117.

969 Ye C, Hill CM, Wu S, Ruan J, Ma Z. 2016. DBG2OLC: Efficient Assembly of Large Genomes
970 Using Long Erroneous Reads of the Third Generation Sequencing Technologies. Sci Rep.
971 6:31900. doi: 10.1038/srep31900.

972 Ye C, Ma Z (Sam). 2016. Sparc: a sparsity-based consensus algorithm for long erroneous
973 sequencing reads. PeerJ. 4:e2016. doi: 10.7717/peerj.2016.

974 Ye C, Ma ZS, Cannon CH, Pop M, Yu DW. 2012. Exploiting sparseness in de novo genome
975 assembly. BMC Bioinformatics. 13:S1. doi: 10.1186/1471-2105-13-S6-S1.

976 Yoshimoto S et al. 2008. A W-linked DM-domain gene, DM-W, participates in primary ovary
977 development in Xenopus laevis. Proceedings of the National Academy of Sciences. 105:2469–
978 2474. doi: 10.1073/pnas.0712244105.

979

980

981

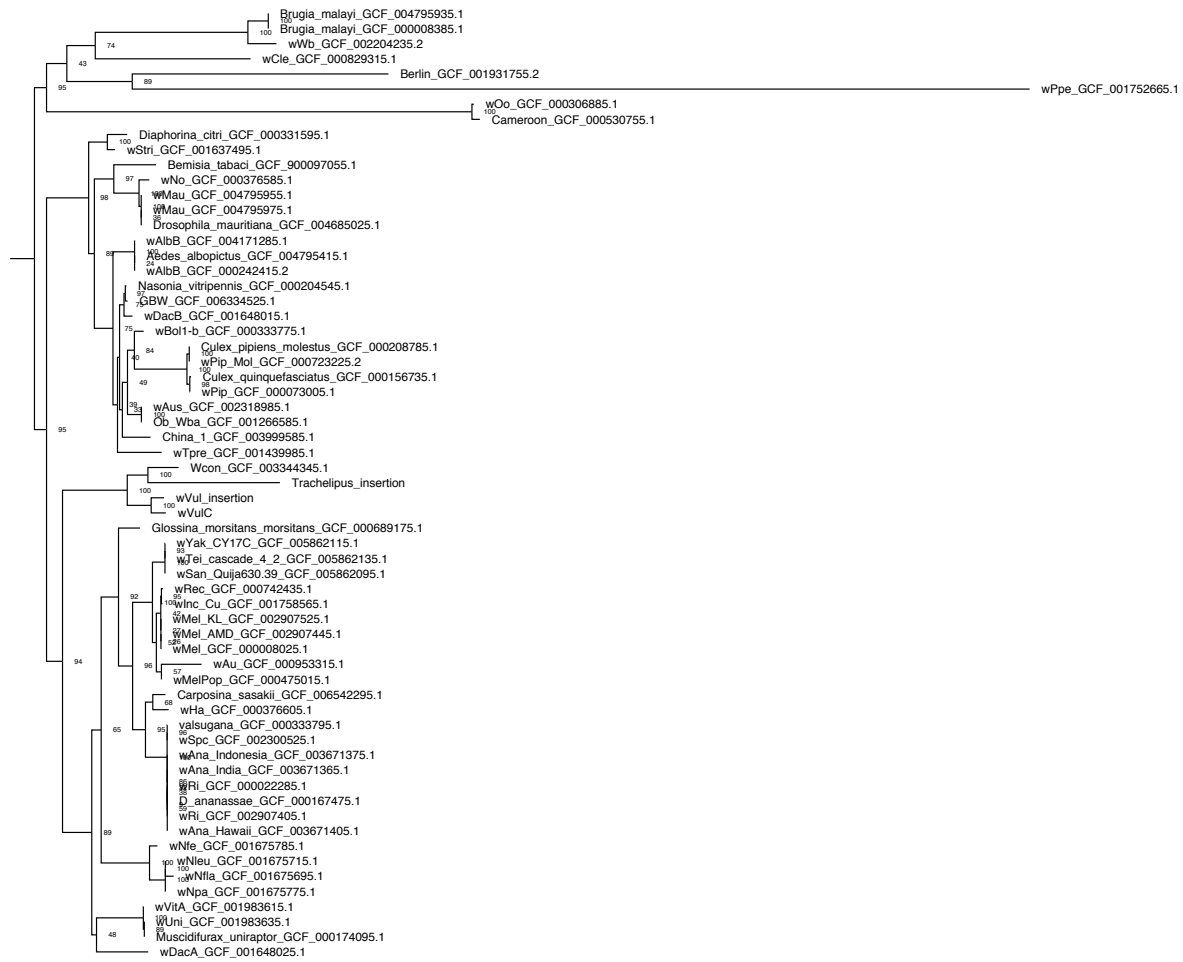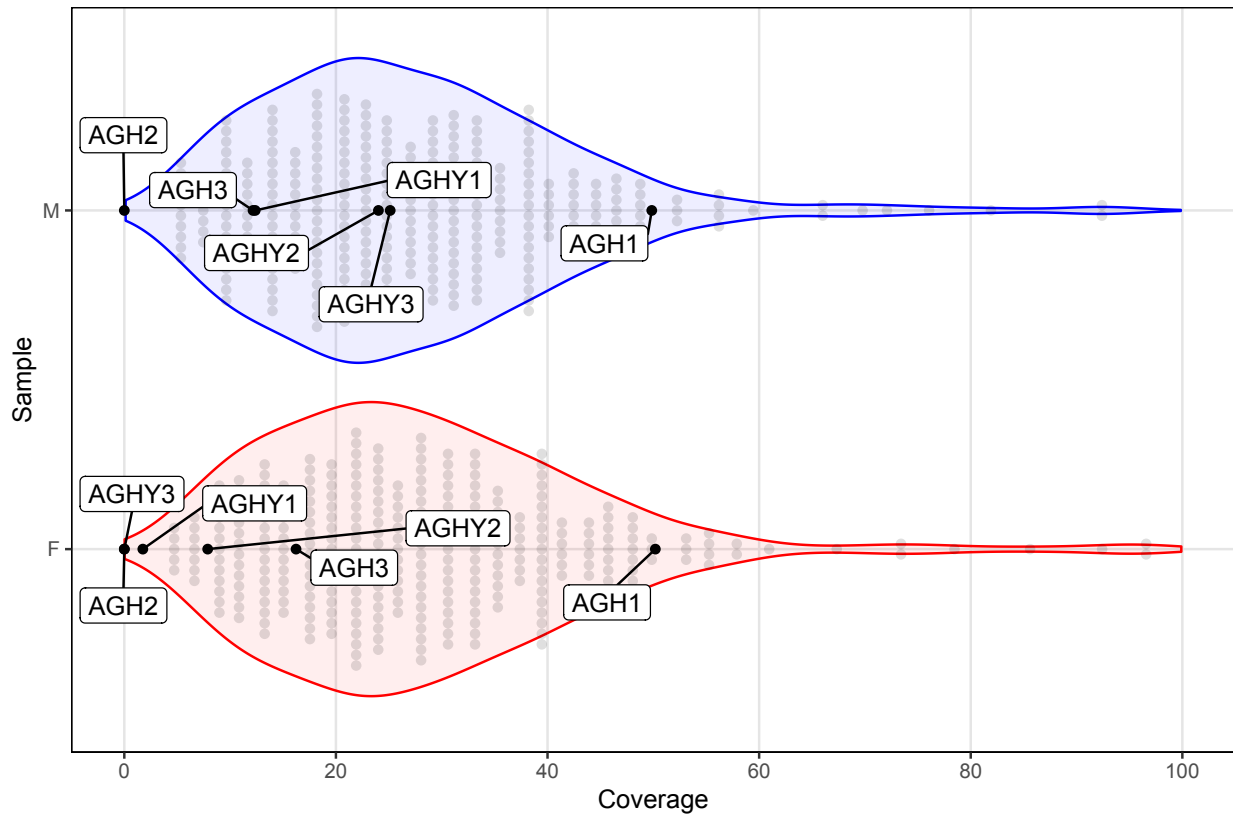982 **Figure 1**

983

984

985   **Figure 2**

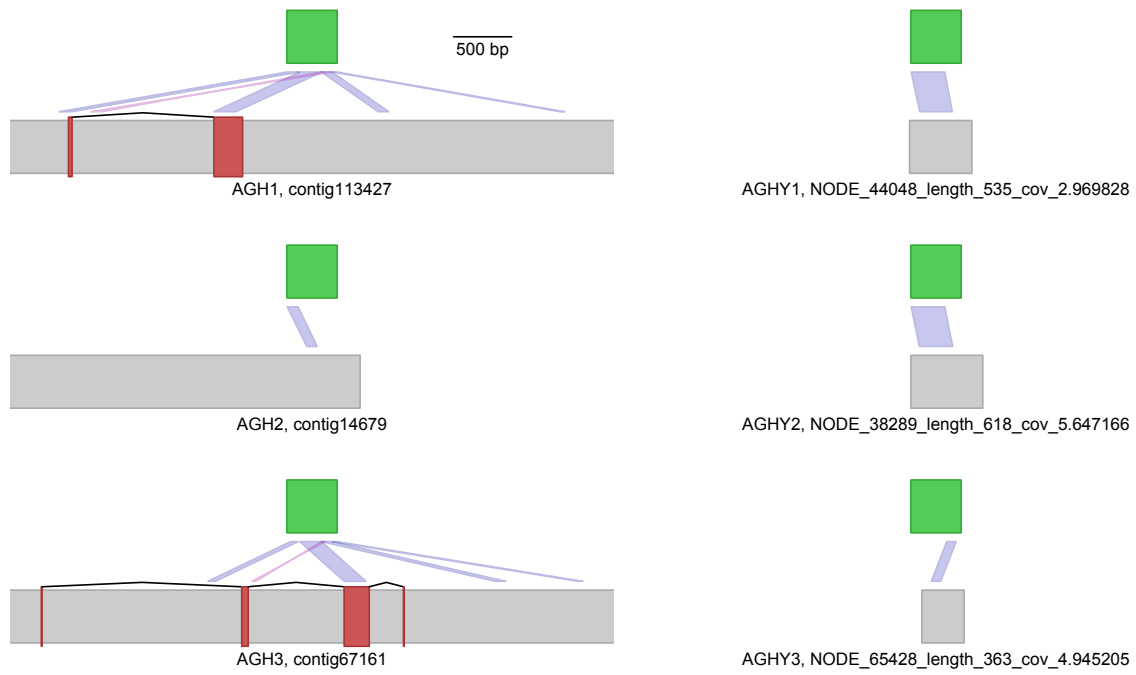
986

987

988 **Figure 3**

989

990

991　**Figure 4**

AGH1, contig113427

AGHY1, NODE_44048_length_535_cov_2.969828

AGH2, contig14679

AGHY2, NODE_38289_length_618_cov_5.647166

AGH3, contig67161

AGHY3, NODE_65428_length_363_cov_4.945205

992

993

994   **Figure 5**

995

                                                    0.2