Identification of sex chromosomes using genomic and cytogenetic methods in a range-expanding spider, *Argiope bruennichi* (Araneae: Araneidae)

Running title: Sex chromosomes in a spider

Monica M. Sheffer[1]*, Mathilde Cordellier[2], Martin Forman[3], Malte Grewoldt[2#], Katharina Hoffmann[2], Corinna Jensen[4,5], Matěj Kotz[3], Jiří Král[3], Andreas W. Kuss[4,5], Eva Líznarová[3,6], Gabriele Uhl[1]

[1]Zoological Institute and Museum, University of Greifswald, 17489 Greifswald, Germany
[2]Institute of Zoology, University of Hamburg, 20146 Hamburg, Germany
[3]Department of Genetics and Microbiology, Charles University, Viničná 5, 128 44 Prague 2, Czech Republic
[4]Center for Functional Genomics of Microbes, University of Greifswald, 17489 Greifswald, Germany
[5]Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, 17489 Greifswald, Germany
[6]Department of Botany and Zoology, Masaryk University, 611 37 Brno, Czech Republic

*Correspondence address: Monica M. Sheffer, Zoological Institute and Museum, University of Greifswald, Loitzer Strasse 26, 17489 Greifswald, Germany

Correspondence E-Mail: monicasheffer.research@gmail.com

[#]Current affiliation: Department of Molecular Biology and Genetics, Aarhus University, 8000 Aarhus, Denmark

Abstract:

Differences between sexes in growth, ecology and behavior strongly shape species biology. In some animal groups, such as spiders, it is difficult or impossible to identify the sex of juveniles. This information would be useful for field surveys, behavioral experiments, and ecological studies on e.g. sex ratios and dispersal. In species with sex chromosomes, sex can be determined based on the specific sex chromosome complement. Additionally, information on the sequence of sex chromosomes provides the basis for studying sex chromosome evolution. We combined cytogenetic and genomic data to identify the sex chromosomes in the sexually dimorphic spider *Argiope bruennichi*, and designed RT-qPCR sex markers. We found that genome size and GC content of this spider falls into the range reported for the majority of araneids. The male karyotype is formed by 24 acrocentric chromosomes with an X_1X_20 sex chromosome system, with little similarity between X chromosomes, suggesting origin of these chromosomes by X chromosome fission or early duplication of an X chromosome and subsequent independent differentiation of the copies. Our data suggest similarly sized X chromosomes in *A. bruennichi.* They are smaller chromosomes of the complement. Our findings open the door to new directions in spider evolutionary and ecological research.

Keywords: *Argiope bruennichi* – Araneae – Araneidae – genome size – karyotype – sex chromosome

Introduction

In the majority of animals, sex is genetically determined. Genotypic mechanisms of sex determination are very diverse, depending, for example, on the ratio of X chromosomes to autosomes, on sex-determining factor(s) localized on autosomes and/or sex chromosomes, on the presence or absence of a sex chromosome specific for the heterogametic sex (Y or W), or on ploidy level as in many Hymenoptera (Bachtrog *et al.*, 2014; Hamm, Meisel, & Scott, 2015; Vicoso, 2019).

In animal species with genotypic sex determination, females and males often differ in the type and/or number of sex chromosomes, with many variations. Knowledge on the karyotype and the specifics of the sex chromosomes provides the basis for studying a plethora of ecological and evolutionary questions. For ecological studies, it is desirable to know the sex of collected individuals. Offspring sex allocation, sex-specific growth and dispersal strategies, and sex differences in early developmental pathways and physiological plasticity could then be investigated (Cordellier *et al.*, 2020). However, in many animal groups, the sexes cannot easily be identified in adults and less so in juveniles. The sex of early stages is often only determinable based on their sex chromosome complement (e.g. Avilés & Maddison, 1991).

Spiders belong to one of the most diverse animal orders (Coddington & Levi, 1991). While they play an important role in terrestrial ecosystems, it is impossible to sex juvenile spiders in the field. Spiders possess complex sex chromosome systems. To date, karyotypes of more than 800 spider species have been characterized (Araujo *et al.*, 2021). In 67% of karyotyped spiders, and particularly in the entelegyne spiders, which are the most-studied, there are two X chromosomes,

X_1 and X_2 (X_1X_20 system), where the male exhibits one copy and the female two copies (♂X_1X_2/♀$X_1X_1X_2X_2$) (Araujo *et al.*, 2012). This pattern is probably ancestral for entelegynes (Král *et al.*, 2006). Systems with three X chromosomes or more also exist. These X chromosomes are supposedly non-homologous based on their achiasmatic pairing during male meiosis (Kořínková & Král, 2013). Some lineages of entelegyne spiders exhibit an X0 system, which has originated by X chromosome fusions, or neo-sex chromosomes, which arose by rearrangements between X chromosome(s) and autosome(s). Neo-sex chromosome systems contain both X and Y chromosome(s). Systems formed by two or more X chromosomes are termed multiple X chromosome systems (White, 1973).

As a consequence of the different number of X chromosome copies in males and females, sexing could be achieved in spiders once targeted molecular markers for X chromosomes are available: the relative copy number of sequences on the X chromosomes, with males having half the number of copies relative to females, could be assessed through quantitative PCR and sexing performed in minimally invasive way, at a throughput that is reasonable for ecological studies.

In recent years, several approaches to identify X chromosomal sequences have been introduced that rely on high throughput sequencing methods. These methods were reviewed by Palmer *et al.* (2019), and vary in their costs, required sample sizes, and on the availability of a reference genome (Al-Dous *et al.*, 2011; Picq *et al.*, 2014; Gautier, 2014; Hou *et al.*, 2015; Muyle *et al.*, 2016). In species with a chromosome-level genome assembly, comparing the sequencing coverage depth across chromosomes between individuals of each sex is a reliable method for identifying sex chromosomes. Females, as the homogametic sex, would show twice as much

82 coverage as males on X chromosomes (i.e. Vicoso & Bachtrog, 2011; Fraïsse, Picard, & Vicoso,
83 2017).

84 In 2021, Sheffer *et al.* published an annotated genome assembled to haploid chromosome level
85 of an entelegyne orb-weaving spider, *Argiope bruennichi* (Scopoli, 1772). With this chromosome-
86 level genome, we can employ the coverage depth approach to identify sex chromosome
87 sequences in this species. *A. bruennichi* is a sexually dimorphic species with a Palearctic
88 distribution, which has rapidly expanded its range in Europe over the course of the last century
89 (Krehenwinkel & Tautz, 2013; Krehenwinkel, Rödder, & Tautz, 2015; Wawer *et al.*, 2017). Sex
90 differences in early life stages of these spiders are particularly interesting since juveniles perform
91 aerial dispersal with silk (ballooning), a behavior which is restricted to the first few instars
92 (Krüger, 2014) and might be sex biased (Krehenwinkel *et al.*, 2016).

93 The *A. bruennichi* genome assembly was based upon female specimens, and is comprised of 13
94 scaffolds, consistent with the known number of chromosome pairs for females of the species
95 (Zhang & Tong, 1990). With additional karyological and sequence data for males, it would be
96 possible to unequivocally ascertain the karyotype/sex determination system of the species and
97 identify the sex chromosomes in the published assembly.

98 We set out to (i) determine the karyotype and genome size of *A. bruennichi*, (ii) identify which
99 superscaffolds represent sex chromosomes in the genome assembly, and (iii) develop molecular
100 markers for sexing spiders using qPCR.

101 # Methods

102 ## Chromosome preparations and their evaluation

103 We collected eight subadult *A. bruennichi* males for chromosomal analysis at village Březová,
104 Czech Republic (49.901N, 13.887E) in late June 2021. The number of meiotic cells in the testes of
105 these males was lower than found in subadult or young adult males of other araneids (Kotz and
106 Forman, unpublished observation). We followed the protocol of Dolejš *et al.* (2011) for
107 preparation of chromosome slides and Giemsa staining. We inspected the slides under an
108 Olympus BX 50 microscope, and photographed selected plates using an Olympus DP 71 CCD
109 camera. We measured chromosome lengths using ImageJ (https://imagej.nih.gov/ij/index.html).
110 The length ratio of sex chromosomes was established using 20 randomly selected diakinesis and
111 metaphase I plates. Chromosome morphology was based on centromere position, which was
112 apparent during early metaphase II.

113 ## Determination of genome size and GC content

114 We determined the diploid genome size (2C value) and guanosine cytosine (GC) content of *A.*
115 *bruennichi* by flow cytometry (FCM) using five adult males and five females (collected in July
116 2020, Albrechtice, Czech Republic, 49.924N, 16.658E, stored at −80 °C). Sample preparation was
117 based on the protocol of Král *et al.* (2019) with different buffer. We chopped legs together with
118 an internal plant standard in GBP buffer (Loureiro *et al.*, 2007), with pH adjusted to 9.5 and 1.5%
119 (v/v) of Triton X-100 (Sigma-Aldrich) and 1.5% (v/v) of polyvinylpyrrolidone (Sigma-Aldrich). After

120 filtration through a nylon sifter, we added fluorescent dye. We performed two concurrent FCM
121 gauges, using: (i) Propidium iodide (PI, 50 µg/ml) - nonspecific dye to establish genome size of
122 the sample and (ii) 4',6-Diamidine-2'-phenylindole dihydrochloride (DAPI, 4 µg/ml) as a
123 fluorochrome with higher specificity to AT pairs. The resulting histograms were evaluated using
124 FloMax software (Partec). We calculated the 2C value as the ratio of PI fluorescence of sample
125 and standard. We compared PI and DAPI ratios to establish GC content as described in Šmarda
126 *et al.* (2008). We measured each individual twice for PI and twice for DAPI to account for machine
127 fluctuations. We used fresh leaves of *Pisum sativum* cultivar 'Ctirad' as internal standard. We
128 used the following genomic values of this plant for calculations: 2C=7841.27 Mbp, GC=41.77%
129 (Veselý *et al.*, 2012). Primary data used for calculation of 2C and base content are available in
130 Supplementary Table S1.

131 Whole-genome sequencing of male and female *Argiope bruennichi*

132 *DNA extraction and library preparation*

133 We will refer to chromosome-level scaffolds in the genome assembly as "superscaffolds" and
134 number all of them by length. For karyotype, we use the traditional approach and number
135 autosomes by length and refer to X chromosomes separately.
136 In order to identify the X chromosomes in the *A. bruennichi* genome assembly, we sequenced
137 one adult male and one adult female collected near Klausdorf, Germany at the end of July, 2020
138 (54.424N, 13.029E). We extracted DNA from eight legs of the male and four legs of the female,
139 as follows: we disrupted the leg tissue with a mortar, pestle, and liquid nitrogen. The powdered

140 tissue was transferred into tubes containing Proteinase K and cell lysis buffer (10mM Tris pH 8,
141 100mM NaCl, 10mM EDTA pH 8, 0.5% SDS, double-distilled water). We incubated the tubes
142 overnight at 55°C, cooled them to room temperature and added RNAse A. We used 5M NaCl to
143 precipitate proteins and Isopropanol to precipitate DNA. We cleaned the DNA extract using 70%
144 Ethanol, and eluted it in TE buffer. We prepared the libraries with the NEBNext Ultra II FS DNA
145 Library Prep Kit (New England Biolabs, Ipswich, Massachusetts, USA) according to the
146 manufacturer's protocol. We indexed and amplified the libraries with six PCR cycles, then pooled
147 them and sequenced the 2x75 bp paired-end libraries on one Flowcell of an Illumina NextSeq 550
148 with a High Output v2 kit (150 cycles) (Illumina Inc., San Diego, California, USA).

149 *Bioinformatic processing and coverage comparison*

150 The reads were demultiplexed using the Illumina bcl2fastq tool in BaseSpace. We removed
151 adapter sequences using AdapterRemoval (v.2.2.2) (Schubert, Lindgreen, & Orlando, 2016) with
152 the --identify-adapters flag. Read quality was assessed using FastQC (v0.11.8) (Andrews, 2010);
153 all reads were high quality and no further quality trimming was necessary. Adapter-trimmed
154 reads are available under BioProject PRJNA629526. We mapped the reads of each specimen onto
155 the *A. bruennichi* genome assembly using BWA-MEM (v.0.7.12-r1039) with default settings (Li,
156 2013), sorted and indexed the mapped reads using SAMtools sort and SAMtools index (SAMtools
157 v.1.3.1) (Li *et al.*, 2009), then calculated the coverage of each of the 13 superscaffolds using
158 QualiMap (v.2.2.1) (Okonechnikov, Conesa, & García-Alcalde, 2016).

159 # Development of sex markers using RT-qPCR

160 ## *DNA extraction*

161 We extracted genomic DNA from two males collected in Greifswald (54.093N, 13.366E) and four
162 females collected near Klausdorf (54.424N, 13.029E), Germany at the beginning of August, 2020
163 with the Promega ReliaPrep™ gDNA Tissue Miniprep System. We ground two legs from females
164 or four legs from males with a Dounce homogenizer; we added 200 µl of Tail Lysis Buffer and 30
165 µl of Proteinase K Solution to the sample, and extracted the DNA following the manufacturer's
166 instructions. We carried out the final elution step twice to a total elution volume of 100 µl. We
167 determined the DNA concentration for each sample using a NanoDrop spectrophotometer
168 (Thermo Scientific, Wilmington, DE) and a Qubit 4 Fluorometer with the Quant-iT™ Qubit™
169 dsDNA BR Assay Kit (Thermo Scientific, Wilmington, DE) (Table S2).

170 ## *Primer design, PCR and RT-qPCR*

171 We designed primers for loci on the putative X chromosomes (superscaffolds 9 and 10 in the *A.*
172 *bruennichi* assembly; see results section) and 2 putative autosomes, superscaffolds 5 and 7, using
173 Primer3 (version 4.1.0; Untergasser *et al.*, 2012) with the following parameters: length between
174 20 and 25 nucleotides, melting temperature between 57 and 62°C, region amplified between 150
175 and 200 bp in length, with one GC clamp (Table S3). We used additional data to select the loci,
176 such as the location of homologous sequences in spiders *Trichonephila senegalensis* (Araneidae)
177 (Walckenaer, 1841) (Grewoldt *et al.,* unpublished data) and *Stegodyphus mimosarum* (Eresidae)
178 (Pavesi, 1883) (Bechsgaard *et al.*, 2019). We tested the amplification with conventional PCR,

179 checked on a 1.5% agarose gel with ethidium bromide staining and UV transillumination. We then
180 conducted RT-qPCR using the EvaGreen® qPCR-Mix II, on a ThermoFisher StepOne™ Real-Time
181 PCR System (Thermo Fisher Scientific, Waltham, USA). Primer efficiencies were measured by
182 diluting template DNA from a female in a stepwise manner to obtain a standard curve. After
183 ensuring comparable efficiencies for all primer pairs, we used DNA from the two females and
184 males as templates and simultaneously measured amplification in real time for each primer pair
185 in triplicate. One negative control (no DNA template) was used for each primer combination. Both
186 amplification protocols were conducted with the following reaction mix: 2 µL EvaGreen® qPCR-
187 Mix II 5x, 0.2 µL forward primer, 0.2 µL reverse primer, 1µL template DNA, and 6.6 µL nuclease
188 free water. We used the following program: initial denaturation for 15 minutes at 95°C, 40 cycles
189 of denaturation (15 seconds at 95°C), annealing (20 seconds at 60°C), and extension (20 seconds
190 at 72°C), with a final extension at 72°C for 10 minutes, and a melting curve (ramping from 65°C
191 to 95°C). Ct values were collected for further analysis (Table S2).

192 *RT-qPCR data analysis*

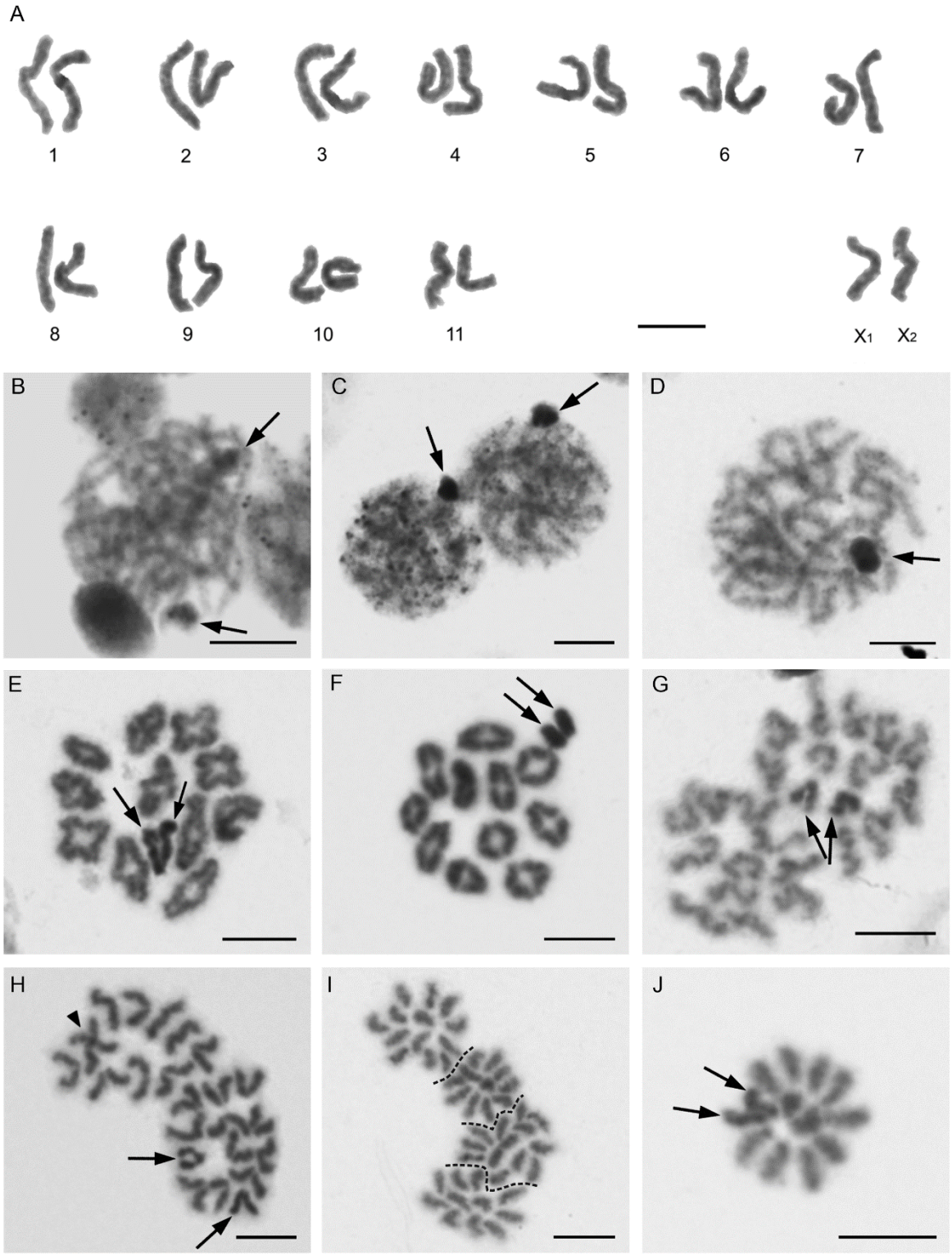
193 We applied the $2^{-\Delta\Delta Ct}$ method according to Rao *et al.*, 2013 to analyze the RT-qPCR data. To
194 account for different primer efficiencies (E_{real}), the Ct values were normalized to correspond to
195 an optimal efficiency of 100% (E_{opt}). We averaged the Ct values of the three technical replicates;
196 if any of the three replicates showed a standard deviation >1, it was excluded from the average.
197 The ΔCt value was calculated from the difference between reference sequences (autosomal loci)
198 and a target sequence of interest (X-linked locus). Further, the female data points were used as
199 the reference, as females possess two sex-linked and two autosomal gene copies. Consequently,

200 if the number of gene copies between females and males is the same, a value of 1 is expected,
201 whereas if the number of gene copies in males is half that in females, a value of 0.5 is expected.

202 Results

203 Chromosomes

204 The male karyotype of *A. bruennichi* was formed by 24 acrocentric chromosomes; the size
205 difference between chromosomes was gradual from the longest to the shortest (Figure 1A).
206 Analysis of male meiotic division revealed an X_1X_20 sex chromosome system (Figure 1B, E-J).
207 Chromosomes X_1 and X_2 were similar in length; X_2 length corresponded to 91.3% (±5.3) of X_1. We
208 observed nearly the complete course of male meiotic division except anaphase I, telophase I, and
209 telophase II. X chromosomes were positively heteropycnotic (i.e. they stained more intensively)
210 from the beginning of meiosis until diakinesis (Figure 1B-E). The intensity of heteropycnosis
211 decreased during metaphase I; X chromosomes were isopycnotic or slightly positively
212 heteropycnotic during this period (Figure 1F). Positive heteropycnosis reappeared during
213 prophase II (Figure 1G). During metaphase II (Fig. 1H), X chromosomes exhibited a slight positive
214 heteropycnosis only. This pattern was also displayed at late anaphase II (Figure 1J). Tight pairing
215 of X_1 and X_2 chromosomes was initiated at zygotene (Figure 1C). During zygotene and pachytene,
216 sex chromosomes formed a body on the periphery of the plate (Figure 1C, D). During late
217 prophase I (diplotene-diakinesis), X chromosome pairing was restricted to centromeric regions
218 (Figure 1E). From metaphase I onwards, X chromosomes were arranged in parallel (Figure 1F, G,
219 J).



221 *Figure 1.* A. bruennichi *male, 2n=24, X_1X_20, Giemsa stained mitotic (A) or meiotic (B-J) chromosomes. Arrows points to sex*
222 *chromosomes. (A) Karyotype, based on spermatogonial prometaphase, (B) leptotene. Chromosomes X_1 and X_2 are not associated,*

(C) two zygotene plates. Sex chromosomes form the body on the periphery of the nucleus, (D) pachytene. The sex chromosome body persists on the periphery of the nucleus, (E) diplotene. Heteropycnotic X_1 and X_2 pair by their centromeric regions, (F) metaphase I. Note a slight positive heteropycnosis of X chromosomes, which are arranged in parallel on the periphery of the plate, (G) two sister prophases II. The left plate contains 11 chromosomes and the right plate 13 chromosomes, (H) two sister metaphases II. Sex chromosomes are slightly positively heteropycnotic (arrowhead - two overlapping chromosomes), (I) two sister anaphases II. Two half plates with 11 chromosomes and two half plates with 13 chromosomes (half plates separated by dashed line). X chromosomes indistinguishable from the other chromosomes, (J) late anaphase II, half plate. Note a slight positive heteropycnosis of X chromosomes. Scale bar = 5 µm.

Genome size and GC content

Female 2C was 4,079.809 Mbp (standard error (SE) ±256.265) and male 2C 3,921.083 Mbp (SE ±207.799). The average chromosome size (2C/2n) was 160.147 Mbp (156.916 Mbp based on female and 163.378 Mbp based on male data, respectively). X chromosomes contained 158.727 Mbp of DNA, based on the difference in female and male 2C. Taking in account the length ratio of X chromosomes, size of X_1 chromosome can be estimated as 82.973 Mbp and size of X_2 as 75.754 Mbp. The size of the X chromosomes was below the average chromosome size. Therefore, X chromosomes belonged to small chromosomes of the karyotype. The genome contained a low GC proportion, 35.500 % of GC (SE ±1.289) in females and 36.730% (SE ±1.189) in males, respectively.

Coverage of male and female

The mean coverage (± standard deviation) across the 13 superscaffolds was 15.34 ±2.57 fold for the male and 16.50 ±0.76 fold for the female. We divided the coverage of the male for each superscaffold by the coverage of the female, to get the male's relative coverage. 11/13 superscaffolds have nearly equal coverage of male relative to female, while two superscaffolds (9 and 10) have half coverage in the male (Figure 2A).

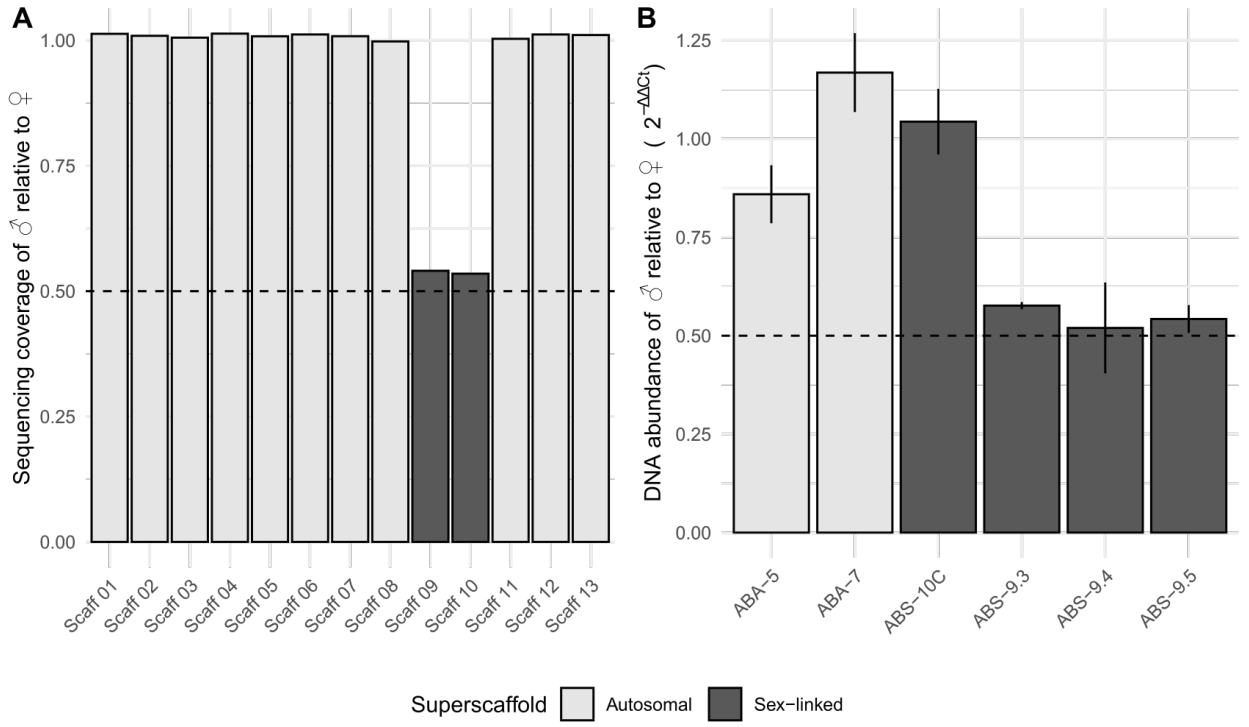


Figure 2. Identification of the sex chromosomes in the Argiope bruennichi *genome assembly. The dashed horizontal line at 0.5 represents (A) the expected coverage and (B) expected relative DNA abundance for sex chromosomes in males compared to females. (A) Comparing the relative sequencing coverage of males and females across the 13 superscaffolds (Scaff) in the genome assembly reveals approximately half coverage for males for Scaff 9 and 10. (B) RT-qPCR results for autosomal loci show approximately equal DNA abundance for males and females. Sex-linked loci on Scaff 9 (ABS-9.3, ABS-9.4, ABS-9.5) show half DNA abundance in males. For the putatively sex-linked locus on Scaff 10 (ABS-10C), males and females have approximately equal DNA abundance, not fitting expectations. Black vertical lines represent standard errors.*

RT-qPCR

To compare sex-linked with autosomal loci, the efficiency of the primers for each locus were evaluated. Primer efficiencies ranged from 85.2- 104.8%, for the chosen loci (Table S3). Likewise, a melting curve was generated for each locus, confirming a specific amplification during RT-qPCR. X-linkage was validated using RT-qPCR on DNA extracted from two males and two females (Figure 2B). The respective loci, sex-linked as well as autosomal, were normalized against the average of all autosomal loci used and the relative DNA abundance in comparison to females was determined using the $2^{-(\Delta\Delta Ct)}$ method. The sex-linked loci on superscaffold 9 exhibited on average

263 a 0.51-0.58 fold-change in DNA abundance, while the locus on superscaffold 10 had an average
264 fold-change of 1.04. Fold-change was between 0.87-1.17 on average for the autosomal loci on
265 superscaffolds 5 and 7 (Figure 2B). Sex-linked loci on superscaffold 9 thus show about half the
266 DNA abundance in males with an average of 0.55 compared to females. For autosomal loci, the
267 average value of 1.01 corresponds to approximately the same DNA abundance.

268 Discussion

269 Our study determined the male karyotype of *A. bruennichi*. It is composed of 24 chromosomes
270 with an X_1X_20 system. The same karyotype was found in other two *Argiope* species, namely *A.*
271 *minuta* (Datta & Chatterjee, 1988), and *A. pulchella* (Bole-Gowda, 1958). This karyotype is the
272 most common in analyzed species of the superfamily Araneoidea, and is probably an ancestral
273 feature of this clade (Araujo *et al.*, 2015). As in most other entelegyne spiders, the karyotype of
274 *A. bruennichi* is formed by acrocentric chromosomes (Kořínková & Král, 2013); this karyotype
275 stucture can be considered ancestral for entelegynes (Král *et al.*, 2006). Data on three other
276 *Argiope* species are incomplete or different from *A. bruennichi*: the male karyotype of *A. amoena*
277 is also formed by 24 chromosomes, X_1X_20 system, but the chromosome morphology was not
278 determined (Suzuki, 1951). The same diploid number was found in *A. catenulata* (Amalin, Barrion,
279 & Rueda, 1992) and *A. luzona* (Carandang & Barrion, 1994), but the chromosome complement of
280 *A. luzona* is reported to consist of both monoarmed (acrocentric and subtelocentric) and biarmed
281 (metacentric and submetacentric) chromosomes (Carandang & Barrion, 1994). This karyotype
282 structure is exceptional in entelegynes, and should be revisited by analysis of metaphase II, when
283 centromeres are unequivocally identifiable. Biarmed chromosomes of *A. luzona* could arise from

284 acrocentric chromosomes, such as those found in *A. bruennichi*, by pericentric inversions. The
285 karyotype of *A. catenulata* is reported to be composed of holocentric chromosomes with an XY
286 system (Amalin *et al.*, 1992). In this case, the authors have most likely mistaken acrocentric
287 chromosomes for holocentric ones, as they do not report any features which support holocentric
288 structure of chromosomes. Moreover, other studied *Argiope* species have monocentric
289 chromosomes (Kořínková & Král, 2013). Furthermore, the determination of the XY system by
290 Amalin *et al.* (1992) is probably also incorrect, as they do not provide any supporting information
291 on the mode of sex chromosome pairing or sex chromosome segregation during male meiosis.
292 As mentioned above, our karyotype analysis revealed 2n=24, X_1X_20 in male of *A. bruennichi*.
293 Given the X_1X_20 system, where males are hemizygous for two X chromosomes, this result fits with
294 the 13 superscaffolds in the haploid genome assembly of a female *A. bruennichi* published by
295 Sheffer *et al.* (2021): 11 autosomes + X_1 + X_2. The result showing half coverage for superscaffolds
296 9 and 10 in the male indicates that these represent the X chromosomes in the *A. bruennichi*
297 genome assembly. These superscaffolds are very similar in length (123.24 Mbp and 122.82 Mbp),
298 which does not match the lengths of X_1 and X_2 (82.973 Mbp and 75.754 Mbp) derived from the
299 comparison of male and female genome sizes, with the X_1:X_2 ratio from image-based
300 measurements. However, the genome size measurements had a relatively high standard error,
301 and the superscaffold lengths fall within this error margin. The fact that we identified two
302 superscaffolds as putative X chromosomes supports the hypothesis that X chromosomes are non-
303 homologous or highly divergent (Kořínková & Král, 2013). If they were homologous, the process
304 of genome assembly would have likely combined them into a single superscaffold, as for
305 homologous autosomal pairs.

306 In addition to the multiple X chromosomes, an XY sex chromosome pair has been discovered in
307 two distantly related entelegyne families, Agelenidae and Lycosidae (Král, 2007; Král *et al.*, 2011).
308 The pair associates at the centromeres with the multiple X chromosomes during male meiosis
309 (Král, 2007). This pair has been termed "cryptic sex chromosome pair" (CSCP) (Sember *et al.*,
310 2020) since the chromosomes are homomorphic (i.e., exhibit no morphological differentiation),
311 and can only be detected under optimal conditions due to fragile meiotic association with the X
312 chromosomes (Král, 2007; Král *et al.*, 2011). Due to its discovery in some entelegyne lineages, as
313 well as in more basally branching spider groups, it has been hypothesized that the CSCP represent
314 ancestral spider sex chromosomes (Král, 2007; Král *et al.*, 2011, 2013; Sember *et al.*, 2020).
315 Whether or not *A. bruennichi* possesses the CSCP is beyond the scope of this study. However, if
316 it does exist in *A. bruennichi*, then the differentiation of X and Y must be very low, as we could
317 not detect them with the given data, and thus the male Y sequences map to the X of the CSCP in
318 the genome assembly. A very low differentiation between chromosomes X and Y of CSCP is
319 supported by identical morphology of these chromosomes (Král, 2007; Král *et al.*, 2011, 2013),
320 distribution of chiasmata, which can be formed along the entire or almost entire chromosome
321 length, as well as by lack of differential hybridization between chromosomes of the CSCP, probed
322 by comparative genomic hybridization (Sember *et al.*, 2020).
323 The X_1X_20 system of entelegyne spiders is thought to have arisen by X chromosome fission (e.g.,
324 Pätau, 1948), nondisjunction of the X chromosome of an X0 system (e.g., Postiglioni & Brum-
325 Zorrilla, 1981), or nondisjunction of the X chromosome of the CSCP (Král, 2007). The non-
326 homologous nature of X_1 and X_2 chromosomes found in *A. bruennichi* could reflect their origin by
327 X chromosome fission. Alternatively, lack of sequence homology could reflect a homologous

origin via early duplication of the X chromosome in an X0 system (or early duplications of the X chromosome of CSCP) followed by a long period of structural differentiation of X_1 and X_2 chromosomes by gene mutations and chromosomal rearrangements. Recent analyses of karyotype evolution of araneomorph spiders suggest a considerable age of X_1 and X_2 chromosomes of the entelegyne X_1X_20 system (Král *et al.*, 2006; Ávila Herrera *et al.*, 2021). Differentiation of X chromosome copies could be accelerated by the inactivation of X chromosome bivalents during meiosis of spider females (Král, 2007).

Interestingly, superscaffold 9 (i.e. chromosome X_1) contains Hox cluster "A" in the *A. bruennichi* genome assembly (Sheffer *et al.*, 2021). Hox genes are a set of highly conserved genes that regulate the body plan organization of animals (Pearson, Lemons, & McGinnis, 2005). This result should be tested with fluorescence *in situ* hybridization, and compared with other spiders, when chromosome-level assemblies and sequences for X chromosomes become available for more species. The localization of the cluster has potential implications of dosage differences in Hox genes for males and females, and the suspected compensation mechanisms should be investigated further.

Our measurements of genome size and GC content also roughly matched data from the published genome assembly (Sheffer *et al.*, 2021). The female measurement of 2C=4,079.809 Mbp ±256.265 corresponds to a haploid (1C) genome size of 2,039.904 Mbp ±128.132 (this study). The genome assembly has a size of 1,670.286 Mbp (Sheffer *et al.*, 2021). Likewise, the average chromosome size for females was 156.916 Mbp based on flow cytometry (this study), and 126.475 Mbp in the genome assembly (Sheffer *et al.*, 2021). The process of genome assembly may collapse repetitive regions, and usually results in a smaller assembly size than measured via

350 flow cytometry (see Pflug *et al.*, 2020). The GC content in flow cytometry measurements was
351 35.5% (this study), while the genome assembly has a GC content of 29.3%. Reduced GC content
352 is a common trait in spiders, and the estimates reported here fall within the range reported for
353 entelegyne spiders (27.2-35%, see Krehenwinkel *et al.*, 2019; Sheffer *et al.*, 2021), and indeed all
354 other investigated spider species (Král *et al.*, 2019). *A. bruennichi* exhibits a larger genome than
355 the two other studied *Argiope* species established using Feulgen densitometry, *A. aurantia* (1,584
356 Mbp) and *A. trifasciata* (1,653 Mbp) (Gregory & Shorthouse, 2003), but falls into the range of 1C
357 reported for the majority of araneids. Excluding the large genome of *Hyposinga* sp. (4,000 Mbp),
358 araneids exhibit a low diversity of genome sizes (1,438 – 2,680 Mbp) (Gregory & Shorthouse,
359 2003).
360 The DNA abundance values obtained by RT-qPCR in males for loci on superscaffold 9 show the
361 expected pattern of approximately half abundance relative to autosomal markers, and can thus
362 be reliably used for sexing spiders. The values obtained for ABS-10C, a locus on superscaffold 10,
363 and thus putatively sex-linked, on the other hand, resemble values obtained for autosomal loci.
364 Given the compelling evidence from the coverage comparison that supports superscaffold 10 as
365 being one of the two X chromosomes, this is likely due to technical errors during RT-qPCR
366 validation and requires further testing. The method tested here is applicable to very small DNA
367 quantities and would allow sexing juvenile spiders as well as early instar spiderlings inside the
368 egg sac in which they overwinter or right after their emergence.
369 Having identified the sex chromosomes in the genome assembly, future work can investigate sex
370 chromosome evolution in *A. bruennichi*. In species in which males are hemizygous for the sex
371 chromosomes, which we have demonstrated is the case for *A. bruennichi*, the X chromosome is

372 expected to evolve faster at the molecular level, relative to autosomes. This is due to a number
373 of factors, such as fewer recombination events, weaker purifying selection or higher positive
374 selection on recessive alleles relative to autosomes. These processes may lead to the acquisition
375 of specific gene content on X chromosomes and to an X chromosome to autosome divergence
376 (Charlesworth, Coyne, & Barton, 1987; Ellegren, 2009). Understanding the forces that drive this
377 specialization, which genes are involved and whether they are beneficial to females, would help
378 to increase our understanding of sex chromosome evolution, monogenic versus polygenic sex
379 determination, and eventually of the evolution of sexual differentiation (Rice, 1984).

380 Conclusions

381 We determined the *Argiope bruennichi* karyotype, genome size and GC content, which is
382 consistent with the existing genome assembly. By comparing sequencing coverage of males and
383 females, we were able to identify which of the superscaffolds in the genome assembly represent
384 the multiple X chromosomes. This result enabled us to develop the first RT-qPCR markers
385 allowing for a molecular approach to sexing spiders early on. Furthermore, having knowledge on
386 the sex chromosomes in the genome assembly will open doors to studies on sex determination,
387 sex allocation, sex specific dispersal and growth, and sex dependent physiological and
388 developmental trajectories.

389 Data Availability

390 Sequencing reads from the male and female specimens are available on NCBI's Short Read

Archive, under BioProject PRJNA629526. Raw data for the flow cytometry data are available in the supplementary material, as is the information on RT-qPCR primers and the Ct values from the RT-qPCR experiments.

Acknowledgements

We would like to thank Henrik Krehenwinkel and Stefan Prost for discussions and brainstorming on this project early on. We also thank Lars Jensen for his input on library preparation and sequencing and Douglas Araujo for access to the paper on the karyotype of *Argiope luzona*. We are grateful to Shou-Wang Lin, who provided a translation of the original *Argiope bruennichi* karyotype paper, which started us on this path to generate new data. We thank the Deutsche Forschungsgemeinschaft (DFG) for the funding of this study as part of the Research Training Group 2010 RESPONSE (GRK 2010) awarded to GU. The cytogenetic part of the study was supported by the Grant Agency of Charles University (project 1000119) and the Ministry of Education, Youth, and Sports of the Czech Republic (project LTAUSA 19142).

Author contributions

Other than the first author position, the authors are listed in alphabetical order by last name, not relative to contributions or institutional hierarchy. MMS, MC, GU, JK and MF conceived of the study. MMS, MF, and GU collected spiders. MMS and CJ isolated DNA, prepared libraries for the whole-genome sequencing and performed the sequencing, with input and infrastructure from GU and AWK. KH, MG, and MC selected loci, designed primers, and performed the RT-qPCR

410 analysis. MK and MF performed the karyological analysis of chromosomes, with input and
411 infrastructure from JK. EL performed the flow cytometry measurements of genome size and GC
412 content. All authors contributed to manuscript writing and revision.

413 References

414 **Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, Salameh YM, Al-Azwani EK,**
415 **Chaluvadi S, Pontaroli AC, Debarry J, Arondel V, Ohlrogge J, Saie IJ, Suliman-Elmeer KM,**
416 **Bennetzen JL, Kruegger RR & Malek JA**. **2011**. De novo genome sequencing and comparative
417 genomics of date palm (*Phoenix dactylifera*). *Nature Biotechnology* **29**: 521–527.

418 **Amalin DM, Barrion AA & Rueda LM**. **1992**. Morphology and cytology of *Argiope catenulata*
419 (Doleschall) (Araneae: Araneidae). *Asia Life Sciences* **1**: 35–44.

420 **Andrews S**. **2010**. FastQC: A quality control tool for high throughput sequence data.

421 **Araujo D, Schneider MC, Paula-Neto E & Cella DM**. **2012**. Sex chromosomes and meiosis in
422 spiders: a review. In: Swan A, ed. *Meiosis: molecular mechanisms and cytogenetic diversity*.
423 Rijeka: InTechOpen, 87–108.

424 **Araujo D, Paula-Neto E, Brescovit AD, Cella DM & Schneider MC**. **2015**. Chromosomal
425 similarities between Nephilidae and Tetragnathidae indicate unique evolutionary traits among
426 Araneoidea. *Italian Journal of Zoology* **82**: 513–520.

427 **Araujo D, Schneider MC, Paula-Neto E & Cella DM**. **2021**. The spider cytogenetic database.

428 **Ávila Herrera IM, Král J, Pastuchová M, Forman M, Musilová J, Kořínková T, Šťáhlavský F,**
429 **Zrzavá M, Nguyen P, Just P, Haddad CR, Hiřman M, Koubová M, Sadílek D & Huber BA**. **2021**.
430 Evolutionary pattern of karyotypes and meiosis in pholcid spiders (Araneae: Pholcidae):
431 implications for reconstructing chromosome evolution of araneomorph spiders. *BMC Ecology*
432 *and Evolution* **21**: 75.

433 **Avilés L & Maddison W**. **1991**. When is the sex ratio biased in social spiders?: Chromosome
434 studies of embryos and male meiosis in *Anelosimus* species (Araneae, Theridiidae). *Journal of*
435 *Arachnology* **19**: 126–135.

436 **Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman TL, Hahn MW, Kitano J,**
437 **Mayrose I, Ming R, Perrin N, Ross L, Valenzuela N, Vamosi JC, Mank JE, Peichel CL, Ashman TL,**
438 **Blackmon H, Goldberg EE, Hahn MW, Kirkpatrick M, Kitano J, Mayrose I, Ming R, Pennell MW,**
439 **Perrin N, Valenzuela N & Vamosi JC**. **2014**. Sex determination: Why so many ways of doing it?
440 *PLoS Biology* **12**: 1–13.

441 **Bechsgaard J, Schou MF, Vanthournout B, Hendrickx F, Knudsen B, Settepani V, Schierup MH**
442 **& Bilde T**. **2019**. Evidence for faster X chromosome evolution in spiders. *Molecular Biology and*
443 *Evolution* **36**: 1281–1293.

Bole-Gowda BN. **1958**. A study of the chromosomes during meiosis in twenty-two species of Indian spiders. *Proceedings of the Zoological Society of Bengal* **11**: 69–108.

Carandang RB & Barrion AA. **1994**. Karyotype of the egg chromosomes of *Argiope luzona* (Walck), an orb-weaving spider (Araneae, Araneidae). *Philippine entomologist* **9**: 443–447.

Charlesworth B, Coyne JA & Barton NH. **1987**. The relative rates of evolution of sex chromosomes and autosomes. *American Naturalist* **130**: 113–146.

Coddington JA & Levi HW. **1991**. Systematics and evolution of spiders (Araneae). *Annual Review of Ecology and Systematics* **22**: 565–592.

Cordellier M, Schneider JM, Uhl G & Posnien N. **2020**. Sex differences in spiders: From phenotype to genomics. *Development Genes and Evolution* **230**: 155–172.

Datta SN & Chatterjee K. **1988**. Chromosomes and sex determination in 13 araneid spiders of North-Eastern India. *Genetica* **76**: 91–99.

Dolejš P, Kořínková T, Musilová J, Opatová V, Kubcová L, Buchar J & Král J. **2011**. Karyotypes of central European spiders of the genera *Arctosa, Tricca*, and *Xerolycosa* (Araneae: Lycosidae). *European Journal of Entomology* **108**: 1–16.

Ellegren H. **2009**. The different levels of genetic diversity in sex chromosomes and autosomes. *Trends in Genetics* **25**: 278–284.

Fraïsse C, Picard MAL & Vicoso B. **2017**. The deep conservation of the Lepidoptera Z chromosome suggests a non-canonical origin of the W. *Nature Communications* **8**: 1–9.

Gautier M. **2014**. Using genotyping data to assign markers to their chromosome type and to infer the sex of individuals: a Bayesian model-based classifier. *Molecular Ecology Resources* **14**: 1141–1159.

Gregory TR & Shorthouse DP. **2003**. Genome sizes of spiders. *Journal of Heredity* **94**: 285–290.

Hamm RL, Meisel RP & Scott JG. **2015**. The evolving puzzle of autosomal versus Y-linked male determination in *Musca domestica*. *G3: Genes, Genomes, Genetics* **5**: 371–384.

Hou J, Ye N, Zhang D, Chen Y, Fang L, Dai X & Yin T. **2015**. Different autosomes evolved into sex chromosomes in the sister genera of *Salix* and *Populus*. *Scientific Reports* **5**: 9076.

Kořínková T & Král J. **2013**. Karyotypes, sex chromosomes, and meiotic division in spiders. In: Nentwig W, ed. *Spider Ecophysiology*. Berlin: Springer, 159–171.

Král J, Musilová J, Šťáhlavský F, Řezáč M, Akan Z, Edwards RL, Coyle FA & Almerje CR. **2006**. Evolution of the karyotype and sex chromosome systems in basal clades of araneomorph spiders (Araneae: Araneomorphae). *Chromosome Research* **14**: 859–880.

Král J. **2007**. Evolution of multiple sex chromosomes in the spider genus *Malthonica* (Araneae: Agelenidae) indicates unique structure of the spider sex chromosome systems. *Chromosome Research* **15**: 863–879.

Král J, Kořínková T, Forman M & Krkavcová L. **2011**. Insights into the meiotic behavior and evolution of multiple sex chromosome systems in spiders. *Cytogenetic and Genome Research* **133**: 43–66.

Král J, Kořínková T, Krkavcová L, Musilová J, Forman M, Herrera IMÁ, Haddad CR, Vítková M, Henriques S, Vargas JGP & Hedin M. **2013**. Evolution of karyotype, sex chromosomes, and meiosis in mygalomorph spiders (Araneae: Mygalomorphae). *Biological Journal of the Linnean Society* **109**: 377–408.

Král J, Forman M, Kořínková T, Lerma ACR, Haddad CR, Musilová J, Řezáč M, Herrera IMÁ, Thakur S, Dippenaar-Schoeman AS, Marec F, Horová L & Bureš P. **2019**. Insights into the karyotype and genome evolution of haplogyne spiders indicate a polyploid origin of lineage with holokinetic chromosomes. *Scientific Reports* **9**: 3001.

Krehenwinkel H, Graze M, Rödder D, Tanaka K, Baba YG, Muster C & Uhl G. **2016**. A phylogeographical survey of a highly dispersive spider reveals eastern Asia as a major glacial refugium for Palaearctic fauna. *Journal of Biogeography* **43**: 1583–1594.

Krehenwinkel H, Meese S, Mayer C, Ruch J, Schneider J, Bilde T, Künzel S, Henderson JB, Russack J, Simison WB, Gillespie R & Uhl G. **2019**. Cost effective microsatellite isolation and genotyping by high throughput sequencing. *Journal of Arachnology* **47**: 190–201.

Krehenwinkel H, Rödder D & Tautz D. **2015**. Eco-genomic analysis of the poleward range expansion of the wasp spider *Argiope bruennichi* shows rapid adaptation and genomic admixture. *Global Change Biology* **21**: 4320–4332.

Krehenwinkel H & Tautz D. **2013**. Northern range expansion of European populations of the wasp spider *Argiope bruennichi* is associated with global warming-correlated genetic admixture and population-specific temperature adaptations. *Molecular Ecology* **22**: 2232–2248.

Krüger J. **2014**. Präballooning-Verhalten bei der Wespenspinne *Argiope bruennichi*.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G & Durbin R. **2009**. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li H. **2013**. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*.

Loureiro J, Rodriguez E, Doležel J & Santos C. **2007**. Two new nuclear isolation buffers for plant DNA flow cytometry: A test with 37 species. *Annals of Botany* **100**: 875–888.

Muyle A, Käfer J, Zemp N, Mousset S, Picard F & Marais GA. **2016**. SEX-DETector: A probabilistic approach to study sex chromosomes in non-model organisms. *Genome Biology and Evolution* **8**: 2530–2543.

Okonechnikov K, Conesa A & García-Alcalde F. **2016**. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**: 292–294.

Palmer DH, Rogers TF, Dean R & Wright AE. **2019**. How to identify sex chromosomes and their turnover. *Molecular Ecology* **28**: 4709–4724.

Pätau K. **1948**. X-segregation and heterochromasy in the spider *Aranea reaumuri*. *Heredity* **2**: 77–100.

Pearson JC, Lemons D & McGinnis W. **2005**. Modulating Hox gene functions during animal body patterning. *Nature Reviews Genetics* **6**: 893–904.

Pflug JM, Holmes VR, Burrus C, Spencer Johnston J & Maddison DR. **2020**. Measuring genome

521 sizes using read-depth, k-mers, and flow cytometry: Methodological comparisons in beetles
522 (Coleoptera). *G3: Genes, Genomes, Genetics* **10**: 3047–3060.

523 **Picq S, Santoni S, Lacombe T, Latreille M, Weber A, Ardisson M, Ivorra S, Maghradze D,**
524 **Arroyo-Garcia R, Chatelet P, This P, Terral JF & Bacilieri R**. **2014**. A small XY chromosomal
525 region explains sex determination in wild dioecious *V. vinifera* and the reversal to
526 hermaphroditism in domesticated grapevines. *BMC Plant Biology* **14**: 229.

527 **Postiglioni A & Brum-Zorrilla N**. **1981**. Karyological studies on Uruguayan spiders II. Sex
528 chromosomes in spiders of the genus *Lycosa* (Araneae-Lycosidae). *Genetica* **56**: 47–53.

529 **Rao X, Huang X, Zhou Z & Lin X**. **2013**. An improvement of the 2ˆ(-delta delta CT) method for
530 quantitative real-time polymerase chain reaction data analysis. *Biostatistics, bioinformatics and*
531 *biomathematics* **3**: 71–85.

532 **Rice WR**. **1984**. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38**: 735–
533 742.

534 **Schubert M, Lindgreen S & Orlando L**. **2016**. AdapterRemoval v2: Rapid adapter trimming,
535 identification, and read merging. *BMC Research Notes* **9**: 1–7.

536 **Sember A, Pappová M, Forman M, Nguyen P, Marec F, Dalíková M, Divišová K, Doležálková-**
537 **Kaštánková M, Zrzavá M, Sadílek D, Hrubá B & Král J**. **2020**. Patterns of sex chromosome
538 differentiation in spiders: Insights from comparative genomic hybridisation. *Genes* **11**: 849.

539 **Sheffer MM, Hoppe A, Krehenwinkel H, Uhl G, Kuss AW, Jensen L, Jensen C, Gillespie RG, Hoff**
540 **KJ & Prost S**. **2021**. Chromosome-level reference genome of the European wasp spider *Argiope*
541 *bruennichi*: A resource for studies on range expansion and evolutionary adaptation. *GigaScience*
542 **10**: 1–12.

543 **Šmarda P, Bureš P, Horová L, Foggi B & Rossi G**. **2008**. Genome size and GC content evolution
544 of *Festuca*: Ancestral expansion and subsequent reduction. *Annals of Botany* **101**: 421–433.

545 **Suzuki S**. **1951**. A comparative study of the chromosomes in the family Argiopidae. *Journal of*
546 *Science of the Hiroshima University, Series B, Division 1* **12**: 67–98.

547 **Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M & Rozen SG**. **2012**.
548 Primer3—new capabilities and interfaces. *Nucleic Acids Research* **40**: e115–e115.

549 **Veselý P, Bureš P, Šmarda P & Pavlíček T**. **2012**. Genome size and DNA base composition of
550 geophytes: the mirror of phenology and ecology? *Annals of Botany* **109**: 65–75.

551 **Vicoso B**. **2019**. Molecular and evolutionary dynamics of animal sex-chromosome turnover.
552 *Nature Ecology & Evolution 2019 3:12* **3**: 1632–1641.

553 **Vicoso B & Bachtrog D**. **2011**. Lack of global dosage compensation in *Schistosoma mansoni*, a
554 female-heterogametic parasite. *Genome Biology and Evolution* **3**: 230–235.

555 **Wawer W, Rutkowski R, Krehenwinkel H, Lutyk D & Pusz- K**. **2017**. Population structure of the
556 expansive wasp spider (*Argiope bruennichi*) at the edge of its range. *Journal of Arachnology* **45**:
557 361–369.

558 **White MJD**. **1973**. *Animal Cytology and Evolution*. London, UK: Cambridge University Press.

Zhang YJ & Tong SJ. **1990**. The routine method for preparing the chromosomes in spiders. *Chinese Journal of Zoology* **25**: 30–31.