Manuscript Click here to download Manuscript Baughman et al.docx

Multiple factors influence population sex ratios in the Mojave Desert moss *Syntrichia caninervis*[1]

Jenna T. Baughman[2,4], Adam C. Payton[3], Amber E. Paasch[2], Kirsten M. Fisher[2], and Stuart F. McDaniel[3]

[2]Department of Biological Sciences, California State University, 5151 State University Drive, Los Angeles, CA 90032, USA, [3]Department of Biology, University of Florida, 876 Newell Drive, Gainesville, FL 32611, USA

[4]Author for correspondence: Department of Integrative Biology and University and Jepson Herbaria, University of California, 3040 Valley Life Sciences Building, Berkeley, CA 94720, USA; e-mail: jbaughman@berkeley.edu; phone: 219-393-0179.

Female biased *Syntrichia caninervis* genetic sex ratio

[1]Manuscript received _____; revision accepted _____.

Acknowledgements

This work was supported by the California State University Department of Biological Sciences Faculty-Student Achievement Fund to KMF, a Research Exchange Grant to JTB from the NSF EDEN RCN (IOS 0955517), the NSF Louis Stokes Alliances for Minority Participation (HRD 1363399) to AEP, the NIH-NIGMS Minority Biomedical Research Support-Research Initiative for Scientific Enhancement program (GM61331) to JTB, and an NSF grant (DEB 1541005) to SFM.

ABSTRACT

- *Premise of research:* Natural populations of many mosses appear highly female-biased based on the presence of reproductive structures. This bias could be caused by increased male mortality, lower male growth rate, or a higher threshold for achieving sexual maturity in males. Here we test these hypotheses using samples from two populations of the Mojave Desert moss *Syntrichia caninervis*.
- *Methods:* We used double digest restriction-site associated DNA (RAD) sequencing to identify candidate sex-associated loci in panel of sex-expressing plants. Next, we used putative sex-associated markers to identify the sex of individuals without sex structures.
- *Key results:* We found an 18:1 phenotypic female:male sex ratio in the higher elevation, moister site, and no sex expression at the drier, low elevation site. In contrast, based on genetic data we found a 2:1 female bias in the low stress site and only females in the high stress site. The area occupied by male and female genets was indistinguishable.
- *Conclusions:* These data suggest that both differential mortality and sexual dimorphism in thresholds for sex expression contribute to population sex ratio biases, and that sex-specific life history and survival interact with environmental stress to determine the frequency of sexual reproduction in *S. caninervis*.

Key words: bryophyte; Mojave Desert; moss; Pottiaceae; reproductive strategy; sex expression; sex ratio; *Syntrichia caninervis*

INTRODUCTION

Female biased sex ratios are a common yet unexplained phenomenon in bryophyte populations (Bisang and Hedenäs, 2005). Most of the female biases documented in mosses to date have been based on counting the number of sexually mature male and female ramets in a population (but see McLetchie et al., 2001; Korpelainen et al., 2008; Bisang et al., 2010; Hedenäs et al., 2010; Bisang and Hedenäs, 2013; Bisang et al., 2014; Bisang et al., 2015; Hedenäs et al., 2016). In dioicous bryophytes, sex determination happens at meiosis with the production of spores (Bachtrog et al., 2011). A spore carrying the female sex-determining locus (U) will form a gametophyte that produces archegonia and ultimately bears the offspring embryo (a sporophyte). Spores with the male sex-determining locus (V) form a gametophyte that produces antheridia. Because the diploid sporophyte is produced by the union of a U-bearing egg with a V-bearing sperm, the sporophyte is always heterozygous at the sex-determining locus. As a consequence of meiotic segregation, therefore, the null expectation is a 1:1 sex ratio. Identifying the underlying mechanisms that generate the persistently female-biased phenotypic sex ratios in moss populations is an essential first step toward understanding the evolutionary causes and consequences of sex ratio bias in this ecologically important group of plants.

A female bias in the production of sexually mature ramets in a natural population could be caused by at least three distinct processes. First, an apparently female biased population sex ratio could simply be a product of faster female growth, as has been found in several species (Shaw and Beer, 1999; Stark et al., 2005a; McDaniel et al., 2008). In this case, female genotypes on average would produce more ramets and occupy larger

areas than male genotypes. Second, males might exhibit a lower frequency of sex expression than females. If this were true, males would constitute a disproportionately large fraction of the sterile plants in a population (termed 'the shy male hypothesis'; Mishler and Oliver, 1991; Stark et al., 2005a). Finally, a female bias could be causd by elevated male mortality during spore production (McDaniel et al., 2007; Norrell et al., 2014), establishment, or at some later point in the life-cycle. Regardless of the proximate cause, elevated male mortality would decrease the amount of genetic diversity in males relative to females.

One of the most extreme cases of sex ratio variation in mosses is in Mojave Desert populations of *Syntrichia caninervis*. Previous data indicate that the phenotypic sex ratio in *S*. *caninervis* populations ranges from 7:1 (Paasch et al., 2015) to 14 F:1 M (Bowker et al., 2000; Bisang and Hedenäs, 2005), and that some populations lack male sex expression entirely (Stark et al., 2001). Mojave Desert *S*. *caninervis* is extremely desiccation tolerant and spends much of its life in an air-dried state, limiting all biological functions to infrequent post-rainfall periods, primarily in cool winter months (Stark, 1997; Stark et al., 1998). Environmental variation that corresponds with changes in timing and duration of this biologically active period appears to affect overall levels of sex expression in a population. A survey of 890 *S*. *caninervis* individuals from a 10 hectare elevation gradient (Bowker et al., 2000) found that total percentage of expressing individuals increased with elevation, and that male sex expression occured only at the higher elevations, while lower elevation populations contained only a few expressing females. In parallel with low levels of sex expression, sexual fertilization and production

111 of sporophytes is relatively rare, and established desert *S. caninervis* populations seem to
112 persist through vegetative cloning (Paasch et al., 2015).
113 Here we utilize restriction site-associated DNA genome sequencing (ddRADseq)
114 to identify the sex of sterile ramets and study the patterns of genetic variation in males
115 and females in two Mohave Desert populatoins of *S. caninervis*. We found extreme
116 female biased phenotypic sex ratios, consistent with previous studies, but less extreme
117 genetic biases, suggesting that males produce sex structures less frequently than females.
118 We also found lower genetic diversity in males than females, suggesting that the
119 remaining genetic sex ratio bias is best explained by elevated male mortality. Patterns of
120 genet size indicate that sexual dimorphism in growth rate is insufficient to explain the
121 female biased sex ratios in this species.
122
123 MATERIALS AND METHODS
124 ***Sample Collection***— We collected male, female, and sterile *S. caninervis* samples
125 at a low-stress and a high-stress site in the Mojave Desert. The low-stress site is located
126 at an elevation of 1800 m in Sheep Creek Wash near Wrightwood, CA (34° 22' 33.85" N,
127 117° 36' 34.59" W), at the west edge of the Mojave Desert and the northern base of the
128 San Gabriel Mountains. The average high and low annual temperature is 16.8 °C and 1.7
129 °C, with an average annual precipitation of 49.4 cm (2007-2011, Wrightwood Weather
130 Station, NOAA National Climatic Data Center). This site experiences little foot traffic.
131 The high-stress site is at an elevation of 1257 m near Phelan, CA (34° 25' 29.80" N, 117°
132 36' 30.91" W), about nine miles northeast of the low-stress site. The average high and
133 low annual temperature is 27 °C and 10 °C, while average annual precipitation is 28.2 cm

134 (2005-2009, Phelan, CA, NOAA National Weather Service). This site is also highly
135 disturbed by foot traffic and erosion.

136 To establish a panel of plants of known sex to use to identify sex-linked molecular
137 markers, we isolated 11 *S. caninervis* female and 10 male ramets from the low stress site
138 in March and April of 2013. Field collections of *S. caninervis* were identified by leaf
139 morphology, color, and hair points then later confirmed with leaf cross-sections. In the
140 laboratory under a dissecting microscope patches were hydrated and screened for
141 presence of current or past antheridia and archegonia.

142 To estimate the phenotypic sex ratio of the low stress population, in May 2014 we
143 collected in 3-5 cm patches in three parallel 20 m linear north-south (N-S) transects, 10 m
144 apart from one another, collecting one patch every 2 m from a variety of shaded and
145 exposed microhabitats. In June 2014 we made collections from the high stress site in a
146 similar manner in two parallel 20 m N-S transects. Additionally, due to the highly
147 irregular distribution of the species in the high stres site, a third mini-transect was
148 sampled (beginning at about the 23 m mark of transect 2 and extending approximately 2
149 m), more densely (ca. every 500 cm) selecting samples in a northwest-southeast (NW-
150 SE) direction from a variety of microhabitats. To estimate the genetic sex ratio and
151 perform diversity analyses, a maximum of three individual sterile ramets were isolated
152 from each patch using a dissecting microscope, resulting in a total of 99 ramets from the
153 low stress site and, due to lower frequency of this species at this site, 42 ramets from the
154 high stress site, for a total of n = 141 individual ramets.

155 ***DNA Extraction and RADseq library preparation***— Total genomic DNA was
156 extracted and isolated from 162 total ramets (21 of known sex, and 141 of unknown sex)

157 using a modified cetyl trimethyl ammonium bromide-beta mercaptoethanol (CTAB)
158 procedure(McDaniel et al., 2007). Prior to extraction, samples were ground dry to a fine
159 powder using a GenoGrinder 2010 bead shaker (SPEX CertiPrep, Metuchen, NJ). DNA
160 quality was evaluated for each sample by electrophoresis and quantified using a Qubit
161 Fluorometer (Invitrogen, Carlsbad, CA, USA) and samples' DNAs were normalized prior
162 to library preparation.

163 Illumina libraries were prepared for sequencing using a modified version of the
164 Peterson et al. (2012) protocol using the endonucleases EcoRI and MseI (New England
165 Biolabs, Ipswich, MA). Following double enzyme digestion, unique barcoded adaptors
166 were ligated to the resulting EcoRI cutsites and a non barcoded universial adaptor to the
167 MseI cutsite. Variable length barcodes of 8, 9, and 10 basepairs (bp) were used with each
168 barcode, differing by at least 4 bp. Illumina flowcell binding and sequencing primer sites
169 were added to the adaptor ligated fragments throught 10 cycles of PCR using NEB Q5
170 PCR mastermix (New England Biolabs, Ipswich, MA). Success of library construction
171 was evaluated throught agarose gel electrophoresis, after which 5 µL of each sample's
172 final library was pooled into a single tube for the 141 sterile samples and a separate tube
173 for the 21 samples of known sex. Size selection and sequencing were performed at the
174 University of Florida's Interdisciplinary Center for Biotechnology Research. Pooled
175 libraries were fractioned using Pippin ELF precision electrophoresis (Sage Science,
176 Beverly, MA) with the resulting 250-500 bp fraction being used for Illumina sequencing.
177 The sterile ramet sample library was sequenced using an Illumina NextSeq 500 at mid
178 throughput, producing 150 bp single-end reads. The library consisting of the 21 known-

179 sex ramets was pooled with libraries constructed in an identical manner for a non-related
180 study and sequenced on one lane of a HiSeq 3000 producing 100 bp single end reads.
181 ***Data analysis***— Raw ddRAD sequence reads were assessed for quality with
182 fastQC (Gordon and Hannon, 2010). All manipulation of raw sequence reads was
183 performed using tools from the FASTX-Toolkit (Gordon and Hannon, 2010)
184 implemented in Galaxy on the University of Florida's Research Computing Cluster. To
185 eliminate low quality bases at the 3' end of reads, all sequences were trimed to 100 bp.
186 Reads were filterd for a minimum Phred quality score of 20 on at least 90% of the read.
187 All reads were then demultiplexed followed by the removal of the 5' barcode and EcoRI
188 enzyme cutsite.
189 RAD loci were assembled de novo using STACKS, version 1.24, (Catchen et al.,
190 2011; Catchen et al., 2013). Pertinent parameters for the STACKS pipeline were as
191 follows: a minimum of 2 reads required for each allele within an indivdial (-m option
192 ustacks), a maximum allowable distance of 2 nucleotides between alleles within a locus (-
193 M option ustacks), a maximum of 2 mismatches allowed when aligning secondary reads
194 to primary stacks (-N option ustacks), and a maximum nucleotide distance of 2 allowed
195 between loci of different inviduals (-n option cstacks). A subset of 30 samples, 15 from
196 each site, was used to create the master catalog of loci. Using the STACKS 'populations'
197 program, loci containing any SNP categorized as heterozygous for an individual were
198 identified. Since all individuals sampled were haploid gametophytes, all loci were
199 expected to be homozygous; therefore, those loci identified as heterozygous were
200 discarded as they likely represented paralogous or over-merged loci. Sequence data were
201 further filtered to retain only bi-allelic loci with 1 or fewer SNPs, found in a minimum of

202 74% of samples, with a minor allele frequency lower limit of 0.1, resulting in 2,234 loci
203 used for subsequent genetic analyses.

204 Twenty-one *S. caninervis* ramets of known sex (11 females expressing archegonia
205 and 10 males expressing antheridia) were used to identify potential sex-associated loci
206 (fragments) that would allow us to infer sex of sterile ramets. Any 80 bp locus that was
207 present in at least five individuals of one sex and absent in all individuals of the opposite
208 sex was considered a potential sex-associated locus. Using these criteria we identified
209 over 1,000 candidate sex-associated loci. From this list, the 100 male-associated and 100
210 female-associated loci that were present at the highest frequency in each sex (present in 9
211 of the 10 males and all of the 11 females) were selected for further analyses. A custom
212 Perl script was used to search for these sequences in the alignments from the final 131
213 sterile ramets used in our analyses (see Results). Of the 200 candidate sex-associated loci,
214 we identified loci that were present in the 131 sterile ramet samples and were in complete
215 linkage with one another. We discarded any loci that were not in complete linkage,
216 because this observation suggested that they were not truly sex-linked, along with any
217 loci that were not recovered in the sterile ramet data set. We counted the number of
218 ramets with only male-associated loci and the number with only female-associated loci
219 and tested for deviations from a 1:1 sex ratio with a Chi-square test and a 2-tailed p-
220 value.

221 In order to determine the number of unique genotypes among the 131 sterile
222 ramets, clonal assignments were made using genetic distance under the infinite allele
223 model with a pairwise distance threshold of 0 between individuals. Higher mismatch
224 thresholds, up to 8, were evaluated but all resulted in the same clonal assignments. Figure

225 2 shows the clonal decay distribution of number of unique genets identified at different
226 genetic distance threshold levels. Furthermore, all ramets that were identified as clones of
227 the same genet were also genetically assigned to the same sex. Clonal assignments were
228 used to fill in missing data, when possible, by finding a consensus sequence among
229 ramets per MLG using GENEIOUS version 8.1.4 (Kearse et al., 2012). To test whether
230 the female biased sex ratio was generated by greater female growth rates, we compared
231 the mean number of sites occupied by female and male MLGs, testing for significance
232 using a one-tailed t-test.

233 To compare levels of genetic variation between the two study populations, and
234 between males and females in the low stress population, we calculated five measures of
235 clonal diversity using the software GENODIVE version 2.0b23 (Meirmans and Van
236 Tienderen, 2004). Clonal diversity (P_d) is a measure of unique genets or multi-locus
237 genotypes (MLGs) relative to (divided by) the number of ramets sampled (Ellstrand and
238 Roose, 1987). The effective number of genotypes (N_{eff}) is an index that accounts for
239 frequencies of each genet and describes the number of MLGs that have equal frequencies
240 while minimizing low frequency MLGs. This index is analogous to effective number of
241 alleles but instead counts whole MLGs. Clonal evenness (effective number of genotypes
242 divided by number of genotypes) indicates how evenly the MLGs are divided over the
243 population and would be equal to 1.00 if all MLGs were represented equally. We also
244 calculated Simpson's diversity index, corrected for sample size, and the corrected
245 Shannon's index (Chao and Shen, 2003). The latter is a measure of clonal variation that
246 accounts for singletons (genets or MLGs sampled just once) in the population for sample
247 sizes greater than approximately 50 sampling units (Arnaud-Haond et al., 2007). We used

248 bootstrap tests with 1,000 permutations and subsampling to match population sizes to test
249 for differences in the latter three measures of clonal diversity.

250 To evaluate the degree to which differences between the two populations could be
251 due to fixed genetic differences, we first calculated the population differentiation statistic
252 F_{ST} among the 131 sterile ramets from the high and low stress sites using GENODIVE.
253 Next we performed a principal components analysis (PCA) of genetic distance
254 covariation, with the missing data filled in first with clones, where possible, then with the
255 mean allele frequencies of all ramets. The eigenvectors (axis loadings) from the first two
256 PCA axes were plotted with the ggplot2 R package (R Core Team, 2013). We also
257 estimated the population structure using the fastSTRUCTURE (Raj et al., 2014) inference
258 algorithm with a simple logistic prior and K = 1 through K = 4. The dataset used for this
259 contained 2,234 SNPs from 131 sterile ramets from the high and low stress sites where
260 missing data was filled in with clones, when possible. Membership coefficients for K = 2
261 were plotted using DISTRUCT version 2.2 from the fastSTRUCTURE software package.

262 For another means to visualize the patterns of genetic distance among genets in
263 the two populations, we constructed a midpoint-rooted neighbor joining genetic distance
264 tree using CLEARCUT. The dataset used for this contained 2,234 SNPs from 51 genets
265 from the high and low stress sites where missing data were filled in with clones, when
266 possible. Two genets with greater than 80% missing data were excluded. Because the
267 RAD loci are not completely linked, this tree does not represent the genealogical
268 relationships among these individuals but rather genetic similarity.

269

270 RESULTS

271 ***Phenotypic sex ratios***— Of the 49 patches collected in the low stress site, 31
272 contained no sex structures, 17 contained ramets expressing archegonia and were
273 classified as female, 1 contained ramets of both sexes, and no patches contained ramets
274 with only male gametangia. The resulting phenotypic sex ratio of 18 F:1 M differs
275 significantly from 1:1 meiotic expectations (Chi-squared test, two-tailed p-value <
276 0.0001).
277 ***Sequencing statistiscs***—Approximately 150 million total raw sequencing reads
278 were generated for the 141 ramet samples sequenced resulting in roughly 1.06 million
279 reads per barcode. About 40% of the reads were discarded due to low quality, resulting in
280 88 million reads that passed initial quality filters. Ten ramets with fewer than 4.5
281 thousand reads were discarded, leaving 98 sterile ramets in the high evelavtion, low stress
282 site, 33 sterile ramets in the low elevation, high stress site, and 21 ramets of known sex
283 from the low stress site, ranging from 51 thousand to 2.6 million reads per ramet. The
284 data matrix of 2,234 SNPs from 131 sterile ramet samples was 80% complete when using
285 original reads but increased to 94% complete when missing data were filled in with data
286 from identical clonal MLGs. The data matrix of concatenated loci used for the RAxML
287 analysis was 83.3% complete with a mean depth of about 5.6 reads per locus.
288 ***Identifying sex linked loci and genetic sex ratio***— Of the 100 candidate male-
289 associated loci, 63 co-segregated exclusively within 33 sterile ramet samples and were in
290 complete linkage with one another. These 63 loci had an average read depth of 6.3 and on
291 average 33.5 male-associated loci were found within each newly identified male ramet. In
292 parallel, 65 of the 100 candidate female-associated loci co-segregated exclusively in the
293 remaining 98 sterile ramets and were also in complete linkage. These loci had a mean

294 read depth of 5.7 and an average of 25.3 female-associated loci were detected in each
295 newly identified female ramet. Of the remaining loci, 52 that were found in the test set of
296 21 ramets of known sex but were absent in the experimental set of 131 sterile ramets.
297 These may have been lost due to subtle differences in the library size selection procedure,
298 or may have been missed due to stochastic sampling. The remaining 20 loci may have
299 been incorrectly categorized as sex-linked, based on the small sample size of the trial set,
300 but it is also possible that some females could have contained fertilized archegonia that
301 went unnoticed during selection of ramets for DNA extraction and, thus, male-associated
302 loci were detected in unobserved embryos (indeed, 13 putative male-associated markers
303 found in sterile females, while 7 putative female-associated markers found in sterile
304 males). Using these putative sex linked markers, we found 65 female ramets and 33 male
305 ramets in samples from the low stress site, equating to an approximate 2:1 female-to-male
306 ratio. This represents a significant deviation from both the 1:1 expectation (Chi-squared
307 test, two-tailed p-value = 0.0012) and from the observed phenotypic sex expression at the
308 patch level (Chi-squared test, two-tailed p-value < 0.0001). Genetic sex ratio on a clonal
309 lineage level in the low stress site was also approximately 2:1 (31 female clonal lineages
310 and 14 male clonal lineages). All 33 of the ramets from the high stress site were found to
311 be female.

312 ***Distribution of genotypic variation***— We found 53 unique MLGs among the 131
313 sterile ramets, based on analysis of 2,234 biallelic SNPs. The low stress site contained 45
314 unique MLGs from 98 ramets and the high stress site had eight MLGs from 33 ramets. A
315 total of 53 MLGs were distinguishable at genetic distance thresholds of 0 through 7, after

316 which the number of MLGs began to drop incrementally without distinct breaks,
317 indicating robust clonal assignments based on the SNP data (Figure 1).
318 Most unique MLGs in the low stress site were restricted to the 3-5cm patch they
319 were collected from. Of the 45 unique geneotypes in the low stress site, three were found
320 within adjacent patches (approximately 2 m apart) and only one MLG extended through
321 three patches for a total of about 4 m (Figure 2). Female and male MLGs did not occupy
322 significantly different numbers of patches (Figure 3A) nor did they contain significantly
323 different numbers of ramets per MLG (Figure 3B).
324 Both the Simpson's diversity and the corrected Shannon indices were
325 significantly higher in the low stress site than in the high stress site (p-value = 0.001).
326 Evenness was also slightly greater in the low stress site, but not significantly so (Table 1).
327 The Simpson's diversity and evenness values were not significantly different between the
328 two sexes in the low stress site, but the corrected Shannon Index was significantly higher
329 for females than for males (Table 2, p-value = 0.001).
330 ***Population differentiation and structure***— The F_{ST} for the high and low stress
331 sites was 0.102 when all ramets were included. When estimated with just unique genets,
332 however, F_{ST} dropped to 0.028. The neighbor joining tree recovered genets of the high
333 stress site nested within the low stress site (Figure 5). Population structure analysis in
334 fastSTRUCTURE estimated a minimum of one model component to explain structure in
335 the data. However, this estimate was not supported by the corresponding marginal
336 likelihoods, which increased with the K parameter with no apparent maximum.
337 Membership coefficients for K = 2, visualized on Figure 5 show some support for
338 structure among the two sites.

339 The first axis of the principle components of covariance analysis (PCA) explained
340 8.112% of the total variation in the 131 ramets and 6.051% was explained by the second
341 axis. A scatterplot of the first two PC axes (Figure 6) shows a tighter clustering of many
342 ramets from the low stress site and lower association of ramets from the high stress site,
343 mostly resolved along PC axis 1. Yet, a few ramets from the high stress site are
344 genetically similar, and therefore clustered near, many in the low stress site.
345
346 DISCUSSION
347 *S. caninervis* has long been a model for investigations regarding the evolution and
348 ecology of sex ratio variation in mosses (Stark et al., 1998; Stark et al., 2001; Stark et al.,
349 2005a; Stark et al., 2005b; Stark and McLetchie, 2006; Paasch et al., 2015). However, the
350 inferences regarding sex ratio variation in this species, as well as other bryophytes, has
351 been limited by the large number of sterile plants in most bryophyte populations. In
352 principle, this limitation is overcome by the use of sex linked molecular markers, which
353 provide a simple means of determining the sex of sterile plants. Identifying reliable sex-
354 linked markers is not trivial, though, requiring screening large numbers of polymorphic
355 loci in a large pedigree or mapping population. In addition, sex linkage ideally should be
356 confirmed using a physical map, or molecular evolutionary analyses, for example to test
357 for complete linkage disquilibrium (LD) among sex linked loci, as predicted based on the
358 observation that recombination does not occur on UV sex chromosomes (Bachtrog et al.,
359 2011). To date, within mosses this has only been accomplished in the model system
360 *Ceratodon purpureus* (McDaniel et al., 2007; McDaniel et al., 2013). A less rigorous
361 approach involves testing for an association between a molecular marker and sex in a

362 large panel of unrelated individuals (as has been done in three *Drepanocladus* species;
363 (Bisang et al., 2010; Bisang and Hedenäs, 2013; Bisang et al., 2015; Hedenäs et al.,
364 2016). Here, we have screened several thousand restriction-site associated sequenced
365 (RADseq) loci in a small number of individuals of known sex, but additionally retained
366 only those that exhibited complete LD among all loci in a larger sample. Beyond the
367 large numbers of polymorphisms, RADseq loci have the advantage of being defined by a
368 unique DNA sequence, which both allows us to be confident in the homology of our loci
369 (unlike gel band-length approaches), and ultimately provides a means of identifying sex-
370 linked genes in published transcriptomes (Gao et al., 2014; Wickett et al., 2014).
371 We use these putative sex-linked markers to show that the genetic sex ratios in
372 two Mojave Desert populations populations of *S. caninervis* are female biased (2 F:1 M
373 in the low stress site, and entirely female in the high stress site). Importantly, however,
374 the phenotypic sex ratios in this sample were far more biased (17 F:1 M in the low stress
375 site). While it is certainly possible that we have over-estimated the long-term female sex
376 expression rates, due to drought in study area in the years prior to collection (2013-2014),
377 collections from seasons with more typical winter weather patterns reported phenotypic
378 sex ratios of 7 F:1 M (Paasch et al., 2015), approximately three times more female biased
379 than the genetic sex ratio we found. These data indicate that males constitute a
380 disporportionately large fraction of the sterile plants, providing unequivocal support for
381 the shy male hypothesis.
382 The greater frequency of female ramets, however, indicates that factors beyond
383 thresholds for sex expression must also contribute to the population sex ratio variation in
384 *S. caninervis*. Experimental manipulations show that females regenerate more readily

385 from plant fragments than males do under both cool conditions and desiccation stress
386 (Stark et al., 2004; Stark et al., 2005b; Stark and McLetchie, 2006), and potentially may
387 exhibit greater clonal propagation. However, the fact that we observed no difference in
388 number of sites occupied between female and male MLGs (Figure 3A) indicates that the
389 genetic female bias in this species is unlikely to be due to faster female growth.

390 A female genotype bias could also be causd by elevated male mortality. The
391 hypothesis uniquely predicts a lower male haplotype diversity, relative to females, as we
392 found (corrected Shannon index, 1.596 in females, 1.208 in males, $p = 0.001$). Indeed,
393 when comparing the two populations, the high stress site is less clonally diverse
394 (Simpson's and corrected Shannon indices, $p = 0.001$; P_d, no significance test) and has a
395 a greater bias than the low stress site. Importantly, the relatively low F_{ST} indicates that
396 most surveyed polymorphisms are shared between the two populations, suggesting that
397 the difference between these two populations are unlikely to result from fixed genetic
398 differences between the two populations.

399 With our current data it is not possible to isolate when in the life cycle male and
400 female demography differ, nor whether it is locally adapted male genotypes that survive
401 or only those genotypes that by chance land in more permissive environmental
402 conditions. However, the association between stress and male mortality suggests that the
403 interaction between male physiology and extrinsic environmental factors are more likely
404 to govern population sex ratio than exclusively intrinsic factors like sex ratio distorters
405 (McDaniel et al., 2007; Norrell et al., 2014). Indeed, the available evidence suggests that
406 *S. caninervis* female biased sex ratios in Mojave Desert populations correlate with
407 precipitation and temperature (this study; Bowker et al., 2000). One potential cause of the

408 elevated male mortality in *S. caninervis* is differential resourse allocation to reproduction.
409 Although sexual reproduction is costly for both sexes, males experience a higher initial
410 cost in the production of antheridia, while fertilized females bear the cost of nurturing a
411 sporophyte (Rydgren et al., 2010). Although we observe sporophytes only rarely in *S.*
412 *caninervis*, the relatively high diversity evident in our sample, along with the limited
413 population structure and weak structure in the population genealogy all indicate that
414 sexual reproduction is relatively frequent. Thus, males may experience sexual selection to
415 produce more antheridia (McDaniel et al., submitted), although it may be opposed by
416 natural selection which favors investment in the maintenance of vegetative tissues.
417 Several bryophyte demographic models predict the eventual local extinction of
418 males, based on vegetative growth patterns similar to those found in *S. caninervis*
419 (McLetchie et al., 2001; Crowley et al., 2005; Rydgren et al., 2010), although these
420 models generally assume that migration is negligable. The lack of population structure
421 between our two study sites indicates that population sex ratios may be influenced by
422 migration of male spores from other populations in addition to local population
423 dynamics. The mixing of genotypes in the PCA and distance tree suggests that both sites
424 in Sheeps Creek Wash draw spores from the same metapopulation. Indeed, the overall
425 genetic structure of S. caninervis is likely to be governed by low rates of sexual
426 reproduction throughout the whole region – to which the low stress site contributes some
427 spores – and those spores that land where they can grow and reproduce end up
428 contributing to subsequent generations. Other sites, like the high stress site, end up being
429 genotype sinks. These inferences are consistent with phylogeographic studies in other
430 species that report limited population structure even at much larger spatial scales

431 (McDaniel and Shaw, 2005; Vanderpoorten et al., 2008; Korpelainen et al., 2012; Shaw
432 et al., 2014).

433 ***Conclusions***—This study demonstrates that the highly female-biased sex ratios
434 observed in Mojave Desert *S. caninervis* are congruent with both the shy male hypothesis
435 (Mishler and Oliver, 1991; Stark et al., 2005a) and increased male mortality. These
436 results highlight the importance of environmental conditions and demographic history for
437 shaping sex ratios in this species, and may have important consequences for the
438 persistence of local populations in the presence of long-term shifts in climate.
439 Importantly, both mechanisms are grounded in the disproportionate pre-zygotic energetic
440 cost of sexual reproduction experienced by males relative to females (Mishler and Oliver,
441 1991; Stark et al., 2000; Stark et al., 2005a). Genomic approaches like those we have
442 used here are likely to shed light on sex ratio variation in other moss species where other
443 processes may predominate (Cronberg, 2003; Bisang and Hedenäs, 2005; Stark et al.,
444 2010; Horsley et al., 2011; Bisang et al., 2014; Norrell et al., 2014). A future challenge is
445 to determine the contributions of migration and environmental factors like water
446 availability, through its effects on determining sex expression and mortality, on the
447 trajectory of population sex ratio change through time.
448

449

450

451

452 LITERATURE CITED

453 ARNAUD-HAOND S., C. M. DUARTE, F. ALBERTO, AND E. A. SERRÃO. 2007.
454 Standardizing methods to address clonality in population studies. *Molecular Ecology* 16:
455 5115-5139.

456 BACHTROG D., M. KIRKPATRICK, J. E. MANK, S. F. MCDANIEL, J. C. PIRES, W. R.
457 RICE, AND N. VALENZUELA. 2011. Are all sex chromosomes created equal? *Trends in*
458 *Genetics* 27: 350-357.

459 BISANG I., J. EARLEN, C. PERSSON, AND L. HEDENÄS. 2014. Family affiliation, sex ratio
460 and sporophyte frequency in unisexual mosses. *Botanical Journal of the Linnaean Society*
461 174: 163-172.

462 BISANG I., J. EHRLEN, H. KORPELAINEN, AND L. HEDENÄS. 2015. No evidence of sexual
463 niche partitioning in a dioecious moss with rare sexual reproduction. *Annals of Botany* 116:
464 771-779.

465 BISANG I. AND L. HEDENÄS. 2005. Sex ratio patterns in dioicous bryophytes re-visited.
466 *Journal of Bryology* 27: 207-219.

467 BISANG I. AND L. HEDENÄS. 2013. Males are not shy in the wetland moss *Drepanocladus*
468 *lycopodioides*. *International Journal of Plant Sciences* 174: 733-739.

469 BISANG I., H. KORPELAINEN, AND L. HEDENÄS. 2010. Can the sex- specific molecular
470 marker of Drepanocladus trifarius uncover gender in related species? *Journal of Bryology*
471 32: 305-308.

472 BOWKER M. A., L. R. STARK, D. N. MCLETCHIE, AND B. D. MISHLER. 2000. Sex
473 expression, skewed sex ratios, and microhabitat distribution in the dioecious desert moss
474 *Syntrichia caninervis* (Pottiaceae). *American Journal of Botany* 87: 517-526.

475 CATCHEN J., A. AMORES, P. HOHENLOHE, W. CRESKO, AND J. POSTLETHWAIT.
476 2011. Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes,*
477 *Genomes, Genetics* 1: 171-182.

478 CATCHEN J., P. HOHENLOHE, S. BASSHAM, A. AMORES, AND W. CRESKO. 2013.
479 Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22: 3124-3140.

480 CHAO A. AND T. SHEN. 2003. Nonparametric estimation of Shannon's index of diversity when
481 there are unseen species in sample. *Environmental and Ecological Statistics* 10: 429-443.

482 CRONBERG N. 2003. Clonal distribution, fertility and sex ratios of the moss *Plagiomnium affine*
483 (Bland.) T. Kop. in forests of contrasting age. *Journal of Bryology* 25: 155-162.

484 CROWLEY P. H., C. STIEHA, AND D. N. MCLETCHIE. 2005. Overgrowth competition,
485 fragmentation and sex ratio dynamics: a spatially explicit, sub-individual-based model.
486 *Journal of Theoretical Biology* 233: 25-42.

487 ELLSTRAND N. C. AND M. L. ROOSE. 1987. Patterns of genotypic diversity in clonal plant-
488 species. *American Journal of Botany* 74: 123-131.

489 GAO B., D. ZHANG, X. LI, H. YANG, AND A. J. WOOD. 2014. De novo assembly and
490 characterization of the transcriptome in the desiccation-tolerant moss *Syntrichia caninervis*.
491 *BMC Research Notes* 7: 1-12.

492 GORDON A. AND G. J. HANNON. 2010. Fastx-toolkit.
493 http://hannonlab.cshl.edu/fastx_toolkit/index.

494 HEDENÄS L., I. BISANG, H. KORPELAINEN, AND B. CRONHOLM. 2010. The true sex
495 ratio in European *Pseudocalliergon trifarium* (Bryophyta: Amblystegiaceae) revealed my a
496 novel molecular approach. *Biological Journal of the Linnean Society* 100: 132-140.

497 HEDENÄS L., H. KORPELAINEN, AND I. BISANG. 2016. Identifying sex in non-fertile
498 individuals of the moss *Drepanocladus turgescens* (Bryophyta: Amblystegiaceae) using a
499 novel molecular approach. *Journal of Plant Research* 129: 1005-1010.

500 HORSLEY K., L. R. STARK, AND D. N. MCLETCHIE. 2011. Does the silver moss Bryum
501 argenteum exhibit sex-specific patterns in vegetative growth rate, asexual fitness or
502 prezygotic reproductive investment? *Annals of Botany* 107: 897-907.

503 KEARSE M., R. MOIR, A. WILSON, S. STONES-HAVAS, M. CHEUNG, S. STURROCK, S.
504 BUXTON, A. COOPER, S. MARKOWITZ, C. DURAN, T. THIERER, B. ASHTON, P.
505 MENTJIES, AND A. DRUMMOND. 2012. Geneious Basic: an integrated and extendable
506 desktop software platform for the organization and analysis of sequence data. *Bioinformatics*
507 28: 1647-1649.

508 KORPELAINEN H., I. BISANG, L. HEDENÄS, AND J. KOLEHMAINEN. 2008. The first sex-
509 specific molecular marker discovered in the moss Pseudocalliergon trifarium. *Journal of*
510 *Heredity* 99: 581-587.

511 KORPELAINEN H., A. K. JÄGERBRAND, AND M. VON CRÄUTLEIN. 2012. Genetic
512 structure of mosses *Pleurozium schreberi* (Willd. ex Brid.) Mitt. and *Racomitrium*
513 *lanuginosum* (Hedw.) Brid. along altitude gradients in Hokkaido, Japan. *Journal of Bryology*
514 34: 309-312.

515 MCDANIEL S. F., J. ATWOOD, AND J. G. BURLEIGH. 2013. Recurring evolution of dioecy
516 in bryophytes. *Evolution* 67: 567-572.

517 MCDANIEL S. F., S. B. CAREY, C. HENRIQUEZ, A. P. POWELL, G. T. GENNIS, E. Y.
518 WALKER, K. S. JONES, AND A. C. PAYTON. Sexual selection on male maturation
519 driven by mating season precipitation. *Evolution*.

520 MCDANIEL S. F. AND A. J. SHAW. 2005. Selective sweeps and intercontinental migration in
521 the cosmopolitan moss *Ceratodon purpureus* (Hedw.) Brid. *Molecular Ecology* 14: 1121-
522 1132.

523 MCDANIEL S. F., J. H. WILLIS, AND A. J. SHAW. 2007. A linkage map reveals a complex
524 basis for segregation distortion in an interpopulational cross in the moss *Ceratodon*
525 *purpureus*. *Genetics* 176: 2489-2500.

526 MCDANIEL S. F., J. H. WILLIS, AND A. J. SHAW. 2008. The genetic basis of developmental
527 abnormalities in interpopulation hybrids of the moss *Ceratodon purpureus*. *Genetics Society*
528 *of America*1425-1435.

529 MCLETCHIE D. N., G. GARCÍA-RAMOS, AND P. H. CROWLEY. 2001. Local sex-ratio
530 dynamics: a model for the dioecious liverwort *Marchantia inflexa*. *Evolutionary Ecology* 15:
531 231-254.

532 MEIRMANS P. G. AND P. H. VAN TIENDEREN. 2004. GENOTYPE and GENODIVE: two
533 programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology*
534 *Notes* 4: 792-794.

535 MISHLER B. D. AND M. J. OLIVER. 1991. Gametophytic phenology of *Tortula ruralis,* a
536 desiccation-tolerant moss, in the Organ Mountains of southern New Mexico. *The Bryologist*
537 94: 143-153.

538 NORRELL T. E., K. S. JONES, A. C. PAYTON, AND S. F. MCDANIEL. 2014. Meiotic sex
539 ratio variation in natural populations of *Ceratodon purpureus* (Ditrichaceae). *American*
540 *Journal of Botany* 101: 1572-1576.

541 PAASCH A. E., B. D. MISHLER, S. NOSRATINIA, L. R. STARK, AND K. M. FISHER. 2015.
542 Decoupling of sexual reproduction and genetic diversity in the female-biased Mojave Desert
543 moss *Syntrichia caninervis* (Pottiaceae). *International Journal of Plant Sciences*.

544 PETERSON B. K., J. N. WEBER, E. H. KAY, H. S. FISHER, AND H. E. HOEKSTRA. 2012.
545 Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping
546 in model and non-model species. *PLOS One* 7.

547 R CORE TEAM. 2013. R: A language and environment for statistical computing. Vienna,
548 Austria.

549 RAJ A., M. STEPHENS, AND J. K. PRITCHARD. 2014. fastSTRUCTURE: Variational
550 Inference of Population Structure in Large SNP Data Sets. *Genetics* 197: 573-589.

551 RYDGREN K., R. HALVORSEN, AND N. CRONBERG. 2010. Infrequent sporophyte
552 production maintains a female-biased sex ratio in the unisexual clonal moss *Hylocomium*
553 *splendens*. *Journal of Ecology* 98: 1224-1231.

554 SHAW A. J. AND S. C. BEER. 1999. Life history variation in gametophyte populations of the
555 moss *Ceratodon purpureus* (Ditrichaceae). *American Journal of Botany* 86: 512-521.

556 SHAW A. J., G. K. GOLINSKI, E. G. CLARK, B. SHAW, H. K. STENØIEN, AND K. I.
557 FLATBERG. 2014. Intercontinental genetic structure in the amphi-Pacific peatmoss
558 *Sphagnum miyabeanum* (Bryophyta: Sphagnaceae). *Biological Journal of the Linnean*
559 *Society* 111: 17-37.

STARK L. R. 1997. Phenology and reproductive biology of *Syntrichia inermis* (Bryopsida, Pottiaceae) in the Mojave Desert. *The Bryologist* 100: 13-27.

STARK L. R., D. N. MCLETCHIE, AND S. M. EPPLEY. 2010. Sex ratios and the shy male hypothesis in the moss *Bryum argenteum* (Bryaceae). *The Bryologist* 113: 788-797.

STARK L. R., D. N. MCLETCHIE, AND B. D. MISHLER. 2005a. Sex expression, plant size, and spatial segregation of the sexes across a stress gradient in the desert moss *Syntrichia caninervis*. *American Bryological and Lichenological Society* 108: 183-193.

STARK L. R. AND D. N. MCLETCHIE. 2006. Gender-specific heat-shock tolerance of hydrated leaves in the desert moss *Syntrichia caninervis*. *Physiologia Plantarum* 126: 187-195.

STARK L. R., N. MCLETCHIE, AND B. D. MISHLER. 2001. Sex expression and sex dimorphism in sporophytic populations of the desert moss *Syntrichia caninervis*. *Plant Ecology* 157: 183-196.

STARK L. R., B. D. MISHLER, AND D. N. MCLETCHIE. 1998. Sex expression and growth rates in natural populations of the desert soil crustal moss *Syntrichia caninervis*. *Journal of Arid Environments* 40: 401-416.

STARK L. R., B. D. MISHLER, AND D. N. MCLETCHIE. 2000. The cost of realized sexual reproduction: assessing patterns of reproductive allocation and sporophyte abortion in a desert moss. *American Journal of Botany* 87: 1599-1608.

STARK L. R., L. NICHOLS II, D. N. MCLETCHIE, S. D. SMITH, AND C. ZUNDEL. 2004. Age and sex-specific rates of leaf regeneration in the Mojave Desert moss *Syntrichia caninervis*. *American Journal of Botany* 91: 1-9.

STARK L. R., L. I. NICHOLS, D. N. MCLETCHIE, AND M. L. BONINE. 2005b. Do the sexes of the desert moss *Syntrichia caninervis* differ in desiccation tolerance? A leaf regeneration assay. *International Journal of Plant Sciences* 166: 21-29.

VANDERPOORTEN A., N. DEVOS, B. GOFFINET, O. J. HARDY, AND A. J. SHAW. 2008. The barriers to oceanic island radiation in bryophytes: insights from the phylogeography of the moss *Grimmia montana*. *Journal of Biogeography* 35: 654-663.

WICKETT N. J., S. MIRARAB, N. NGUYEN, T. WARNOW, E. CARPENTER, N. MATASCI, S. AYYAMPALAYAM, M. S. BARKER, J. G. BURLEIGH, M. A. GITZENDANNER, B. R. RUHFEL, E. WAFULA, J. P. DER, S. W. GRAHAM, S. MATHEWS, M. MELKONIAN, D. E. SOLTIS, P. S. SOLTIS, N. W. MILES, C. J. ROTHFELS, L. POKORNY, A. J. SHAW, L. DEGIRONIMO, D. W. STEVENSON, B. SUREK, J. C. VILLARREAL, B. ROURE, H. PHILIPPE, C. W. DEPAMPHILIS, T. CHEN, M. K. DEYHOLOS, R. S. BAUCOM, T. M. KUTCHAN, M. M. AUGUSTIN, J. WANG, Y. ZHANG, Z. TIAN, Z. YAN, X. WU, X. SUN, G. K. WONG, AND J. LEEBENS-MACK. 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences* 111: E4859-E4868.

598 TABLES

Site	N	G	P_d	N_{eff}	Simpson's	Evenness	Shannon (corrected)
Low stress	98	45	0.459	36.379	0.983*	0.808	1.753*
High stress	33	8	0.242	6.368	0.869*	0.796	0.874*

599

600 Table 1. Clonal diversity indices for high and low stress Sheep Creek Wash sites.
601 Note: *Statistical difference between high and low, p-value = 0.001, based on 1,000
602 permutations with subsampling to match population sizes. N = number of ramets
603 sampled, G = number of genets, P_d = clonal diversity, N_{eff} = effective number of
604 genotypes.

605

606

Sex	N	G	P_d	N_{eff}	Simpson's	Evenness	Shannon (corrected)
Female	65	31	0.477	25.000	0.975	0.806	1.596*
Male	33	14	0.424	11.463	0.941	0.819	1.208*

607

608 Table 2. Clonal diversity indices for the sexes in low stress Sheep Creek Wash site.
609 Note: *Statistical difference between the sexes in low stress site, p-value = 0.001, based
610 on 1,000 permutations with subsampling to match population sizes. N = number of
611 ramets sampled, G = number of genets, P_d = clonal diversity, N_{eff} = effective number of
612 genotypes.

613

614 APPENDIX 1

615 Specimen vouchers were deposited in the California State University, Los Angeles

616 Herbarium (CSLA).

617

618 FIGURE LEGENDS

619 Figure 1. Clonal lineage decay at different genetic distance thresholds. This figure shows

620 the number of genets that would be identified at each genetic distance parameter setting.

621

622 Figure 2. Spatial distribution of genets. The low stress site contained 53 unique genets in

623 98 ramets from 33 patches along three transects (T1, T2, and T3). Females are

624 represented as circles and males as squares. Genets or multi-locus genotypes (MLGs) are

625 outlined in dashed lines. In cases that genets are in different patches, a dashed line

626 connects different ramets of the same MLG.

627

628 Figure 3. Number of patches occupied by (A) and (B) number of ramets per male and

629 female MLG. The average number of patches occupied by and the number of ramets

630 male and female MLGs in the Sheep Creek Wash low stress site do not differ

631 significantly. Error bars are standard deviation.

632

633 Figure 4. Proportion of genets with 1-5 ramets in the high and low stress sites. Proportion

634 of each vertical bar represents proportion of total genets or multi-locus genotypes

635 (MLGs) that were represented by each number of ramets.

636

637 Figure 5. fastSTRUCTURE membership coefficients for K = 2. Vertical bars represents
638 an individual ramets while colors represent genetic clusters detected (K = 2). Bars on the
639 left represent ramets from the high stress site (n = 33) and bars on the right represent
640 ramets from the low stress site (n = 98). Membership coefficients are plotted for each
641 ramet and colored corresponding to the proportion of its MLG that most closely aligns
642 each of the two clusters. The dataset used for this contained 2,234 SNPs from 131 sterile
643 ramets from the high and low stress sites where missing data was filled in with clones,
644 when possible.

645

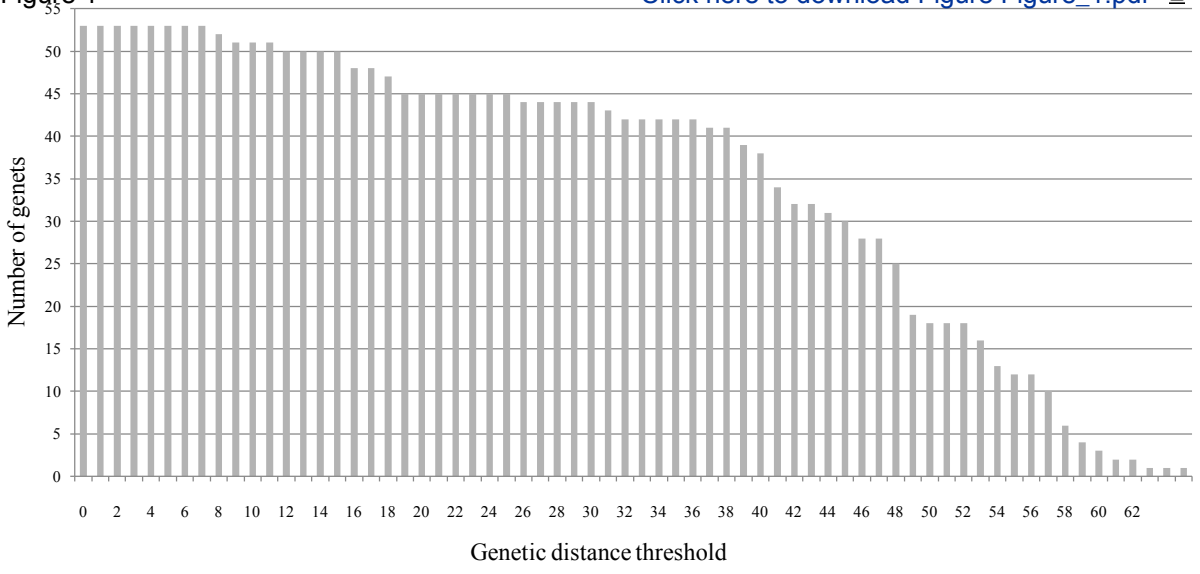
646 Figure 6. CLEARCUT neighbor joining genetic distance tree of ramets from the low and
647 the high stress site of Sheep Creek. Branches that lead to the low stress ('L') site are blue
648 and those that lead to the high stress site ('H') are red. Branches that lead to mixed-
649 population groups are purple. Symbols indicate inferred sex.

650

651 Figure 7. Genetic distance PCA of all MLGs in the high and low stress sites Sheep Creek
652 Wash. Ramets from the high stress site are red and from the low stress site are blue.
653 Males are indicated with triangles and females with circles, and shape size is scaled to the
654 number of ramets sharing the same MLG ('Clones'). Missing alleles were filled in first
655 with clones, when possible, then with the mean allele frequencies of all ramets. PC1
656 explains 8.112% of the total variation in the samples and 6.051% is explained by PC2
657 (Cumulative 14.164%).

Figure 1

Click here to download Figure Figure_1.pdf



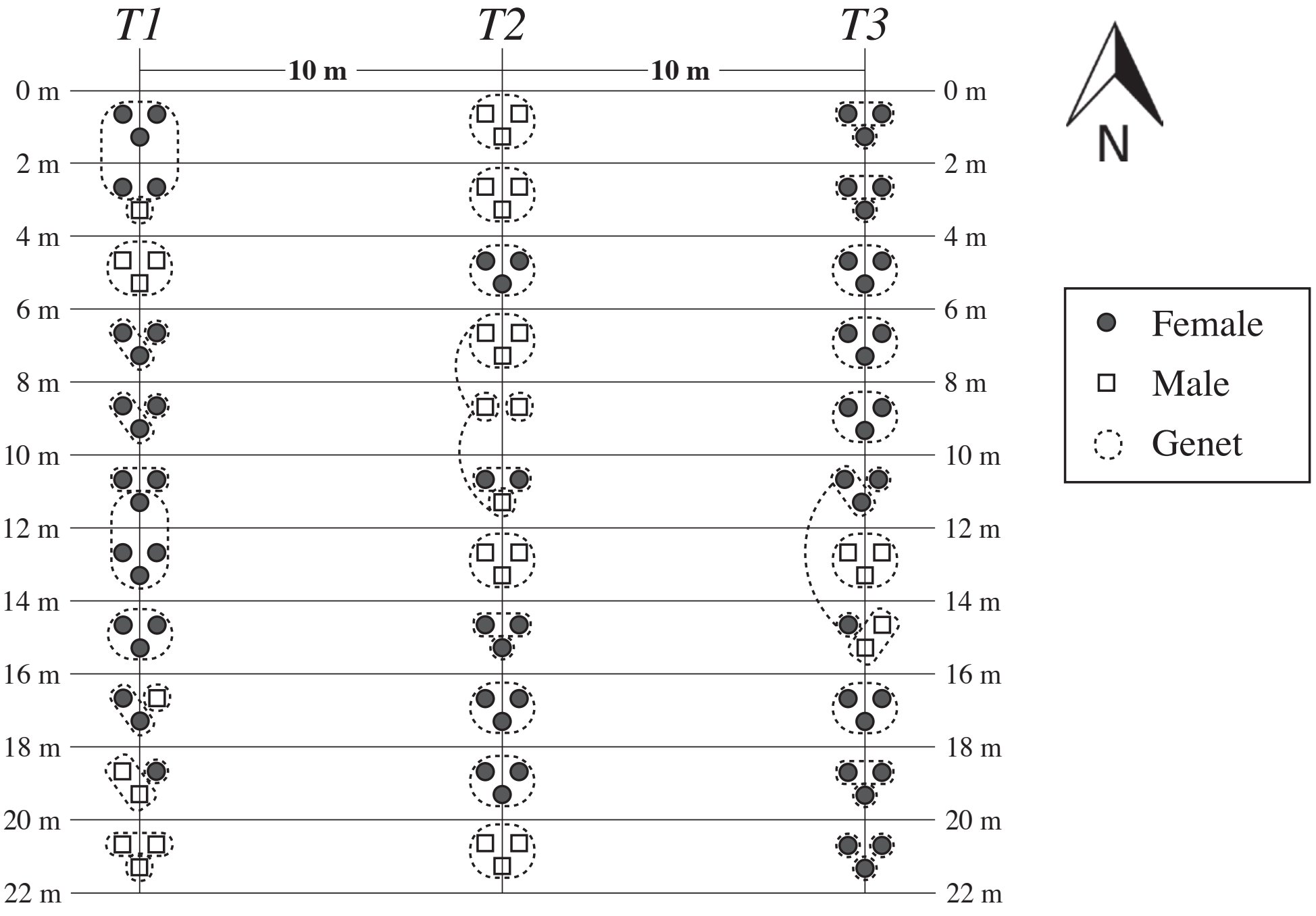
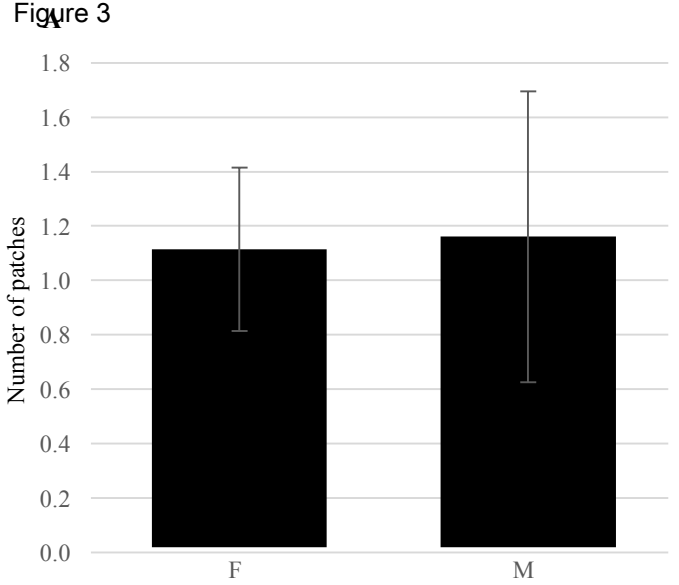
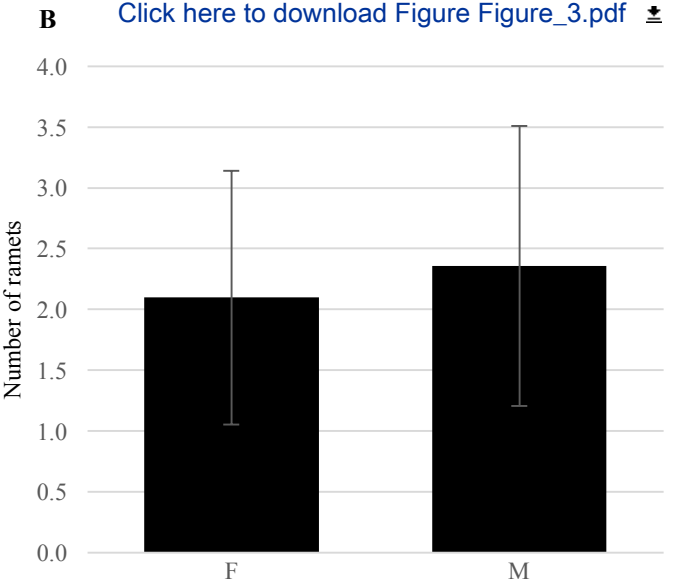

T1
T2
T3
10 m
10 m
0 m
2 m
4 m
6 m
8 m
10 m
12 m
14 m
16 m
18 m
20 m
22 m
N
Female
Male
Genet

Figure 3

Click here to download Figure Figure_3.pdf



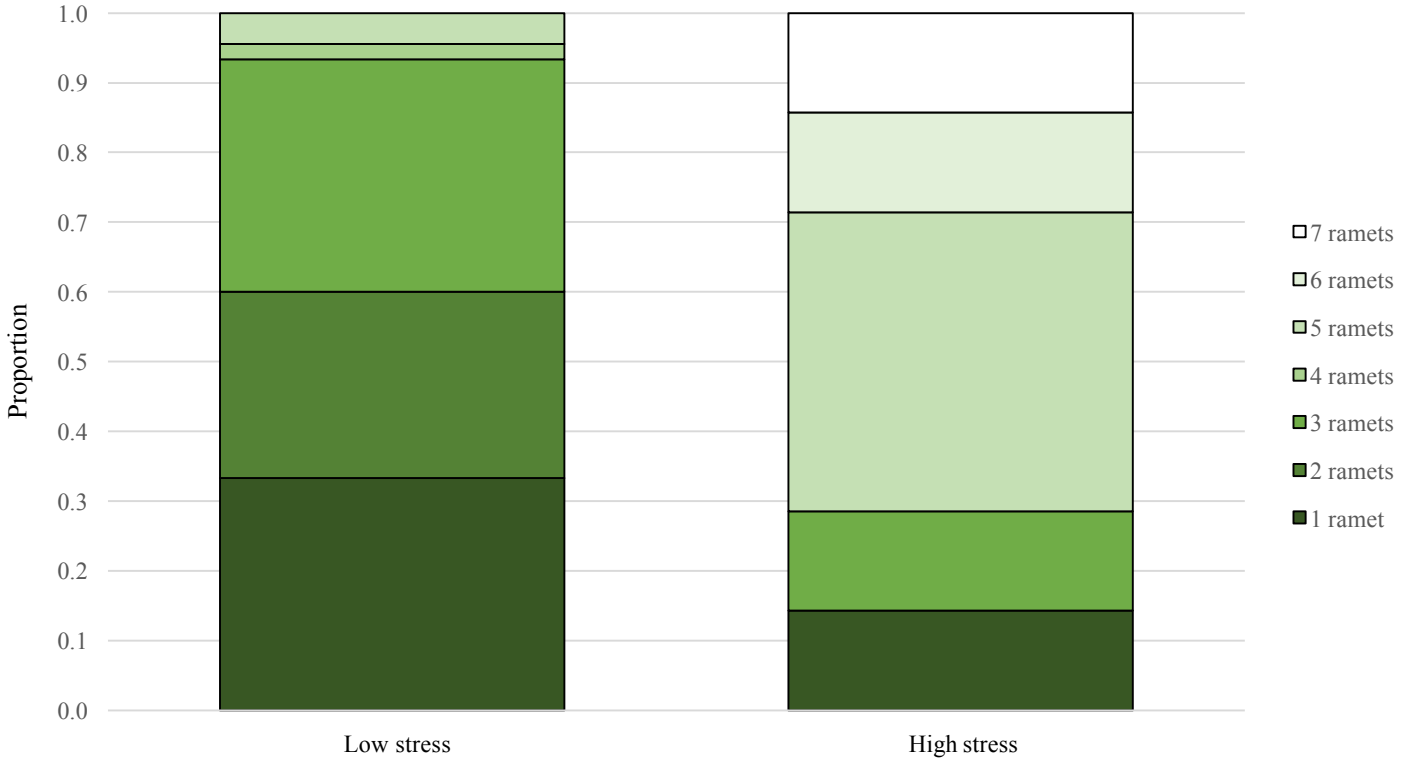
	F	M
Mean	1.097	1.143



	F	M
Mean	2.097	2.357

Figure 4

Click here to download Figure Figure_4.pdf



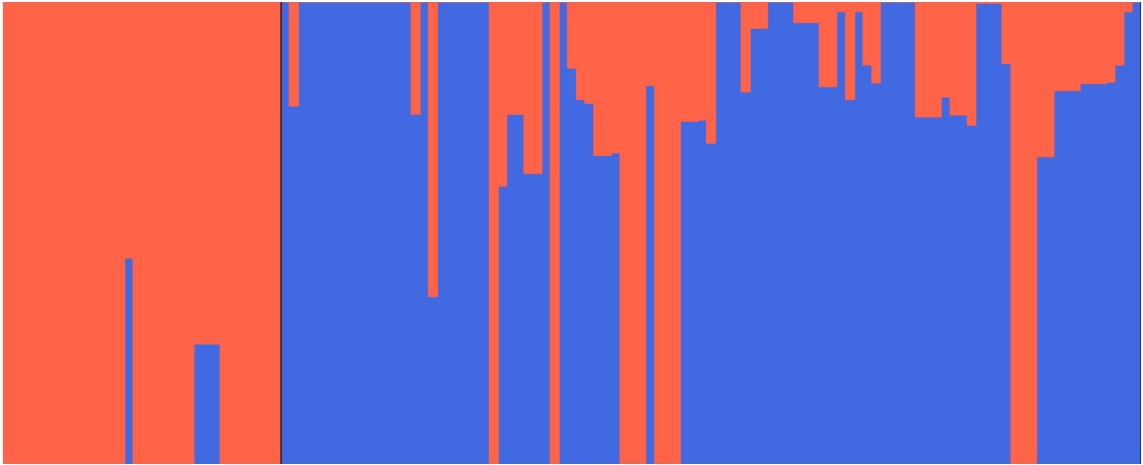

High stress
Low stress

Figure 6

Click here to download Figure Figure_6.pdf

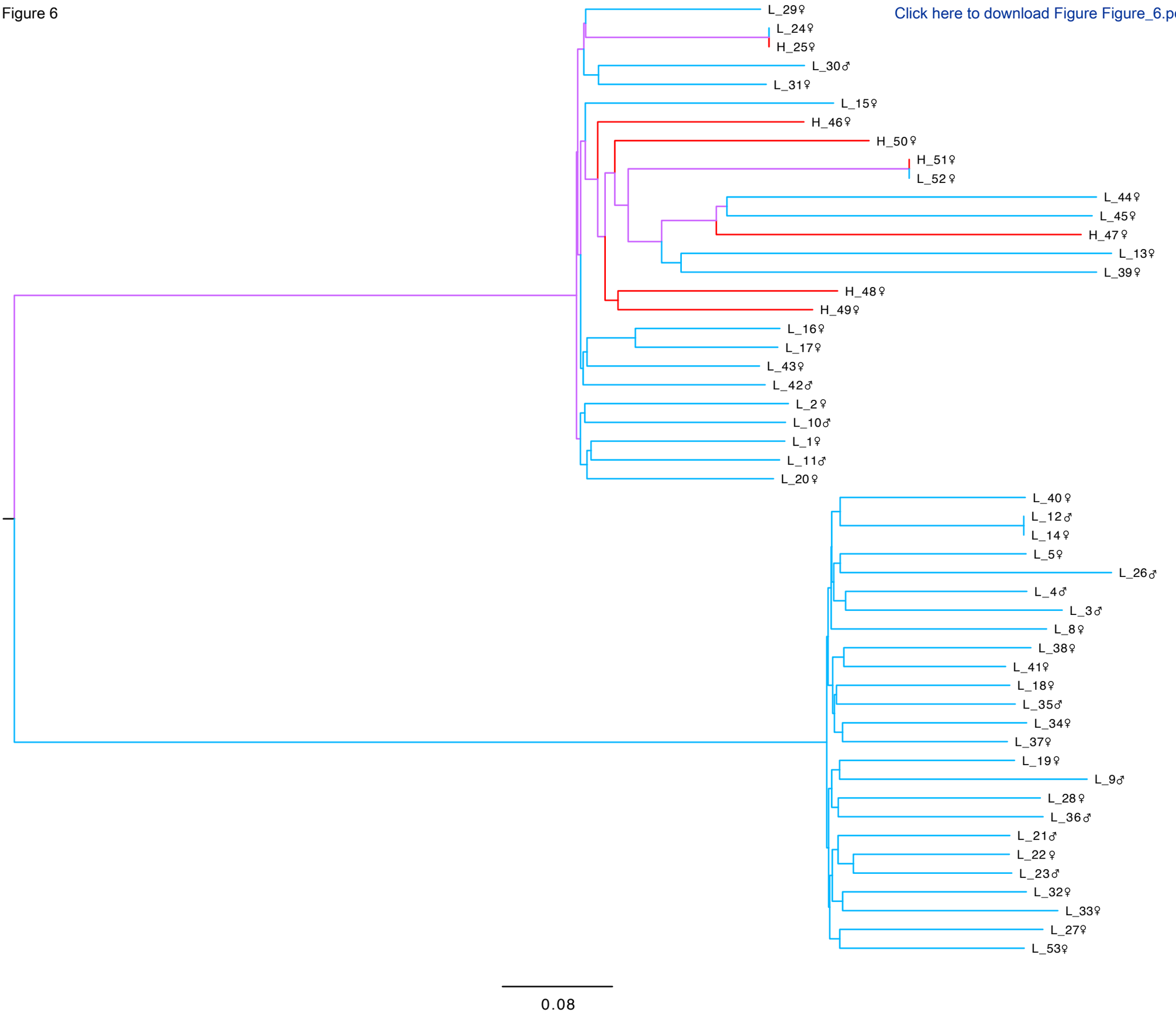
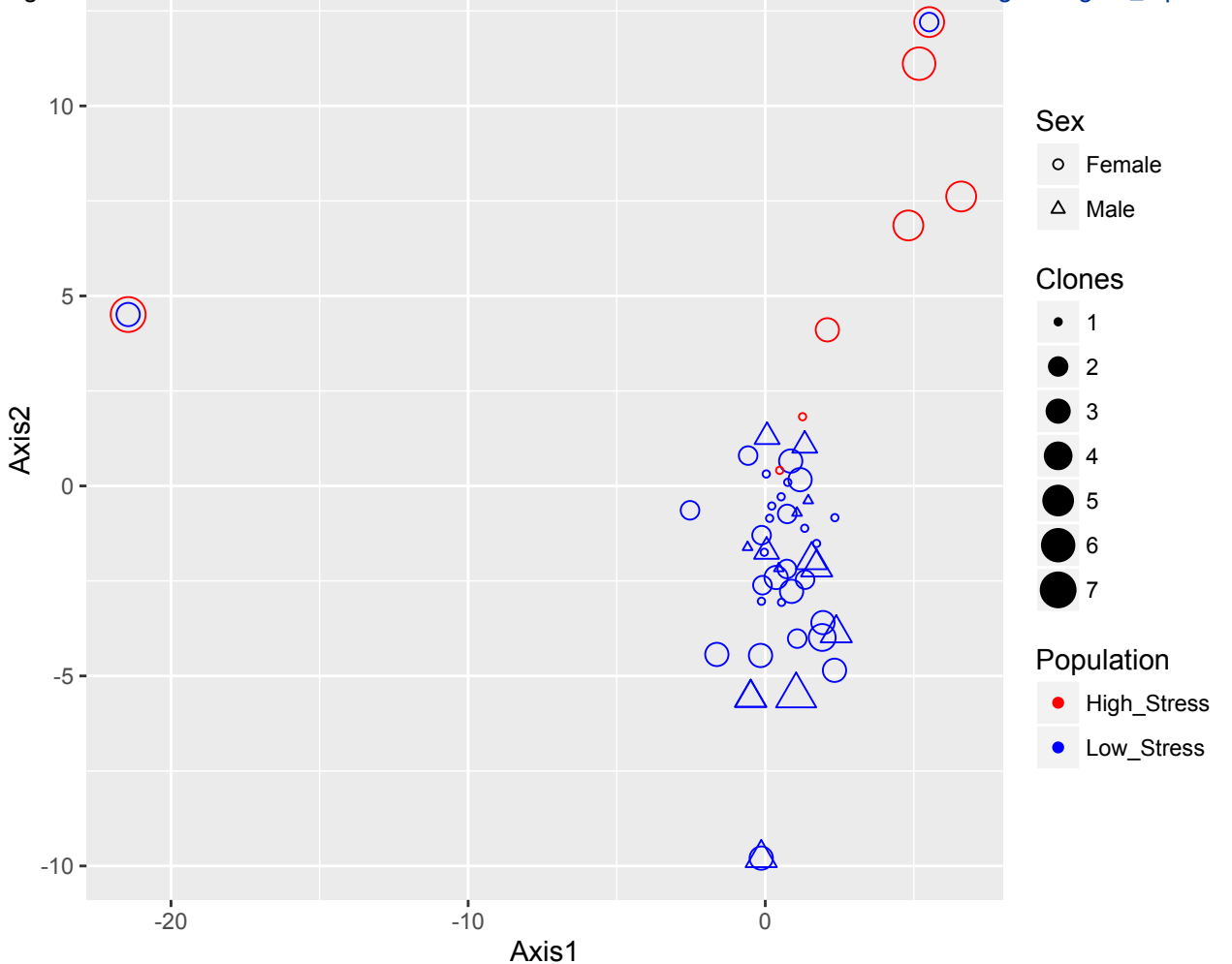


Figure 7

Click here to download Figure Figure_7.pdf