1 **The genetics of mate preferences in hybrids between two young and sympatric Lake**

2 **Victoria cichlid species**

3

4 Ola Svensson12#, Katie Woodhouse13, Cock van Oosterhout14, Alan Smith1, George F

5 Turner15*, Ole Seehausen67*

6

7 #Corresponding author: ola.svensson@bioenv.gu.se

8 *Equal contribution

9

10 1 School of Biological, Biomedical and Environmental Sciences, University of Hull,

11 Hull HU67RX, UK

12 2 Present address: Department of Biological and Environmental Sciences, University of

13 Gothenburg, Box 463, Gothenburg 405 30, Sweden

14 3 Present address: Easton and Otley College, Easton Campus, Easton, Norwich NR4 7TJ, UK

15 4 Present address: School of Environmental Sciences, University of East Anglia, Norwich

16 Research Park, Norwich NR9 5DX, UK

17 5 Present address: School of Biological Sciences, Bangor University, Bangor, Gwynedd LL57

18 2UW, UK

19 6 Department of Fish Ecology and Evolution, EAWAG Swiss Federal Institute of Aquatic

20 Science and Technology, Center for Ecology, Evolution and Biogeochemistry, Seestrasse 79,

21 6047 Kastanienbaum, Switzerland

22 7 Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern,

23 Baltzerstrasse 6, 3012 Bern, Switzerland

24

25 Abstract

26   The genetic architecture of mate preferences are likely to affect significant evolutionary

27   processes, including speciation and hybridisation. Here, we investigate laboratory hybrids

28   between a pair of sympatric Lake Victoria cichlid fish species that appear to have recently

29   evolved from a hybrid population between in similar predecessor species. The species

30   demonstrate strong assortative mating in the lab associated with divergent male breeding

31   colouration (red dorsum vs blue). We show in a common garden experiment, using DNA-

32   based paternity testing, that the strong female mate preferences among males of the two

33   species are fully recovered in a large fraction of their F2 hybrid generation. Individual hybrid

34   females often demonstrated consistent preferences in multiple mate choice trials (≥ 5) across a

35   year or more. This result suggests that female mate preference is influenced by relatively few

36   major genes or genomic regions. These preferences were not changed by experience of a

37   successful spawning event with a male of the non-preferred species in a no-choice single-male

38   trial. We found no evidence for imprinting in the F2 hybrids, although the F1 hybrid females

39   may have been imprinted on their mothers. We discuss this nearly Mendelian inheritance of

40   consistent innate mate preferences in the context of speciation theory.

41

Introduction

Behavioural assortative mating is considered to play a significant role in the origin and maintenance of reproductive isolation among species [1, 2]. The rate of and constraints to the evolution of behavioural assortative mating is likely often influenced by the genetic architecture of mate preferences and the nature and strength of genetic and non-genetic influences, such as imprinting and experience. For example, modelling studies suggest that sympatric and parapatric speciation starting from a monomorphic population is more probable in cases where assortative mating or female preference among male courtship genotypes is influenced by relatively few genetic loci [3-5], although models starting from large standing variation may not have this constraint of preference architecture [6]. However, a small number of preference genes tends to facilitate speciation in many models of speciation with gene flow [7, 8]. Empirical studies of the genetics of species divergence in mating preferences are still rare. Some of the empirical results are consistent with few genes having a major effect on female assortative mating in cichlid fish and *Heliconius* butterflies [9-12]. In other systems, mostly insects, female choice appears to have a more quantitative genetic background [13-15].

The Lake Victoria rocky-shore cichlid fishes of the genus *Pundamilia* have emerged as a significant model system for the study of speciation, being representatives of a spectacular hyperdiverse, rapid adaptive radiation and being relatively tractable as a laboratory species for breeding and mate choice experiments [16, 17]. Following the completion of their genome sequence [18], the evolutionary history of focal populations in the SE part of the lake has been reconstructed [19]. Analysis of genome-wide sequence data indicates that the species with red dorsum (*P.* 'nyererei-like') and blue (*P.* 'pundamilia-like') males at Python Island have

70 recently diverged in situ, following a period of massive introgression with resident *P.*

71 *pundamilia* on the colonisation of the island by *P. nyererei* from elsewhere in the lake [19].

72

73 The *Pundamilia* species, like other haplochromine cichlid fishes, show strong sex role

74 differentiation and associated sexual dimorphism: the smaller, cryptic females are

75 mouthbrooders, caring for the offspring for several weeks, while the larger brightly coloured

76 males defend territories and display to attract females, but play no part in rearing the offspring

77 [20]. Such a breeding system is likely to generate strong sexual selection acting through

78 male-male competition and female preference for male courtship traits [21]. Closely-related

79 haplochromine species often differ markedly in male nuptial colour and it has been proposed

80 that this is associated with divergent female mate preferences [22], which have been

81 demonstrated in a number of experimental trials [23-25]. The resultant assortative mating

82 between females with a certain preference and males expressing the corresponding trait may

83 play a significant role in the maintenance and perhaps sometimes the origin of reproductive

84 isolation among sympatric species [16].

85

86 In the *Pundamilia* red/blue system, increasing water depth is associated with differentiation in

87 alleles at the long wavelength sensitive opsin gene (*LWS*), female preferences and male

88 nuptial colour, and it is likely that the sensory environment along this microhabitat gradient

89 has influenced divergence through a process of 'sensory drive' [26]. Of course, mating signals

90 are often multimodal and subject to multivariate selection [27-29] which is most likely also

91 the case in *Pundamilia* [16, 17, 30]. However, in the *Pundamilia* system, female preferences

92 for male nuptial colouration – itself likely to be oligogenic [31] – appear to be necessary and

93 sufficient for assortative mating [30, 32, 33].

94

95   In haplochromine cichlids, trait segregation in F2 hybrids has been shown for female

96   preferences [9, 12], male nuptial colouration [12, 31, 34] and male attractiveness to parental

97   species [33, 35]. This includes the *Pundamilia* system, where, furthermore, studies suggest an

98   absence of physical linkage between male nuptial colour and female mate preference [36]. At

99   Python Island, gene flow between the species is estimated to be ongoing [19]. Therefore, the

100 observed strong linkage disequilibrium between male colour and female preference is likely

101 to be maintained by divergent selection. A behavioural study on the second generation (F2)

102 hybrid offspring of *P*. sp. "pundamilia-like" and *P*. sp. "nyererei-like" crosses by Haesler and

103 Seehausen [9] revealed that female mate preference has a genetic basis, and that there may be

104 as few as 1 to 5 major genes that contribute to the variation in this trait. That study, however,

105 used a behavioural assay to measure mate choice, which may not be entirely predictive of

106 actual mating decisions. Here, we used a 'common garden' approach with full-contact

107 spawnings to examine female mate choice decisions in first and second-generation hybrids

108 (F1 and F2). Wild-type females were included as a control. We used molecular paternity

109 determination to measure directly the mating decisions of females in the laboratory [24] and

110 examined the repeatability (≥5 spawning decisions) of female mate choice over a year or

111 more to estimate the segregation of mate preferences in the F2 hybrids of the sympatric sister

112 species of *Pundamilia* from Python island. In contrast to Haesler and Seehausen [9], we

113 examined if mate preferences are consistently maintained across many spawning events (the

114 fully cycle from egg maturation to spawning).

115

116 If female preference is a polygenic quantitative trait with an additive genetic basis, F2 hybrids

117 preferences is expected to be distributed in a Gaussian-like fashion with few individuals

118 expressing significant preferences in the tails of the distribution. In contrast, for a polygenic

119 trait with strong dominance effects, the preference distribution of the F2 will be skewed

120 towards either end of the distribution [37-39]. On the other hand, if preferences are not

121 genetically determined, the preference distribution of F2 females is predicted to be more

122 uniform across F2 females given that individuals shared the same common environment.

123 However, in the case of gene flow, linkage disequilibrium between alleles in a polygenic trait

124 will be broken up by recombination [40, 41] and polygenic mating preferences will be

125 difficult to maintain under such conditions. Because ongoing gene flow [17, 19, 26] and

126 recombination (J. Meier and O. Seehausen, unpublished data) have been shown in this young

127 [19] species pair, and because differentiation in polygenic mating preferences will be difficult

128 to maintain under such conditions, we predicted mate preferences to segregate as an

129 oligogenic trait in a nearly Mendelian fashion.

130

131 **Methods**

132 *The experimental fish*

133 We used the sympatric sister species *Pundamilia* sp. "pundamilia-like" and *Pundamilia* sp.

134 "nyererei-like" (sensu Meier et al. [19]). These taxa show a striking difference in male nuptial

135 colours: *P.* sp. "pundamilia-like" males are grey on the flanks between black vertical bars and

136 have a metallic blue spinous dorsal fin, whereas *P.* sp. "nyererei-like" are orange on the

137 dorsum, dorsal head surface and dorsal fin and yellow on the flanks between black vertical

138 bars. It is estimated that there is currently a low to moderate level of gene flow between the

139 taxa at Python Island (The effective number of haploid immigrants per generation [2Nm]

140 forward in time is 0.7 from *P*. sp. "pundamilia-like" to *P.* sp. "nyererei-like" and 7.2 in the

141 opposite direction [19]). Species differences in female mate choice and divergent alleles at the

142 *LWS* opsin gene are not completely fixed [17] and males with intermediate colouration are

143 found [26]. In contrast, at Makobe Island in the open lake the sympatric species pair *P.*

144 *pundamilia* and *P. nyererei* shows stronger genome-wide differentiation, is more ecologically

145   differentiated, intermediate phenotypes are not observed and no mismatches have been

146   reported between male colouration and *LWS* opsin allele [17, 19, 26]. Both species are diploid

147   and have 22 chromosomes (2n=44) [18].

148

149   Wild-type females and two F1 hybrid families (one in each cross direction) used in the mate

150   choice experiment were bred from wild-caught parents. The fry were raised in stock tanks

151   until large enough to be tagged with an integrated transponder (PIT tag), to enable individual

152   identification. Using microsatellite DNA parentage analyses, we concluded that the 15 *P*. sp.

153   "pundamilia-like" females originated from 3 wild mothers and 1 wild sire and the 6 *P*. sp.

154   "nyererei-like" females from 3-6 wild mothers and 5 wild sires.

155

156   The two F2 families used in the mate choice experiment were bred from a lab stock collected

157   in 1992 [42]. The F1 families were bred from the second to third lab generation. The F2

158   generations were bred by holding one F1 male (no replacement, N=3) together with not more

159   than 10 F1 females in the same aquarium. One F2 half-sib family (PN1-33) was bred from

160   fish from two F1 families bred in 1999 from a female *P*. sp. "pundamilia-like" x male *P*. sp.

161   "nyererei-like", and vice versa. This was the same F2 family used by Haesler and Seehausen

162   [9]. The F2 broods were kept separate and hence some broods in the electronic supplemental

163   material figure S3 may have had the same mother, whereas we know which of the two males

164   was the father. The other F2 family (PN34) was bred from fish from one F1 family bred in

165   2001 from a female *P*. sp. "pundamilia-like" x male *P*. sp. "nyererei-like". The offspring were

166   pooled into the same aquaria and hence the father is known but not the brood or mother.

167   When F2 offspring were large enough, they were PIT-tagged and pooled into the same

168   aquaria. The breeding set-up is given in the electronic supplementary material, figure S1.

169

170 All females had been brooded in the mouth of their mothers until independently feeding and

171 were then raised apart from their mothers. In the data analyses we have included all spawning

172 wild type and F1 females and the 69 F2 females which spawned ≥ 5 times in the experiment.

173 Spawning decisions of females with ≤ 4 spawning decisions are given in the electronic

174 supplementary material (figure S3 and table S1) and were also used in the calculations of

175 paternal and brood effects.

176

177 *Mate choice*

178 Mate choice was tested using a "partial partition" design [24]. An aquarium measuring L 600

179 cm x W 80 cm x H 40 cm was divided into ten equally-sized compartments by plastic grids, 8

180 containing one male each, 4 of each species. Identical halved flower pots (D = 270 mm, L =

181 220 mm) served as the focal point in male territories. Two chambers were accessible to

182 females only. We used several males of each species to decrease the chance that effects of

183 individual variation in male attractiveness could override female mating preferences for males

184 of one species or the other. The mesh size of the plastic grids was adjusted to confine males in

185 their compartments, but to allow the smaller females to pass through. The complement of

186 males was replaced every second month and the female-only compartments were relocated. In

187 total, 11 wild caught and 8 lab-bred *P.* sp. "pundamilia-like" males and 11 wild caught and 6

188 laboratory-bred *P.* sp. "nyererei-like" males were used in the experiment (Electronic

189 supplementary material table S3). Wild type females were tested as a control that species-

190 specific mating preferences would be expressed in this setup. All females were tested with

191 wild type males; hybrid males were not used in these experiments.

192

193 To test whether experience altered mating preferences, 16 F2 females that had spawned 6

194 broods each and whose preferences were hence known were isolated in a tank with a male of

195   the non-preferred species. The 5 *P.* sp. "pundamilia-like"-preferring females had spawned 90-

196   100% with *P.* sp. "pundamilia-like" (mean= 98 %), and the 11 *P.* sp. "nyererei-like"-

197   preferring females had spawned 83 –100% with *P.* sp. "nyererei-like" (mean= 96 %). The

198   females that subsequently spawned with a male of the species they had not preferred (N=9)

199   were allowed to brood fry until final release and potential independence of the fry. Thereafter,

200   they were released back into the experimental tank and allowed to spawn again with a choice

201   of males.

202

203   All experimental fish were marked with PIT tags and a small piece of the dorsal fin was cut to

204   provide a DNA sample. Females with eggs were placed in a separate aquarium until the eggs

205   hatched. All larvae/juveniles were euthanized using MS-222 (tricaine methanesulfonate) and

206   stored in 95% ethanol prior to paternity analyses. All females were released back into the

207   experimental tank after handling.

208

209   Paternity analyses

210   The experiment lasted 2.5 years. Five embryos from each brood were genotyped at 2-5

211   microsatellite loci, Ppun5, Ppun7, Pun17, Ppun21 and Ppun32. Methods for DNA extraction

212   and PCR reactions were as described previously [33] with additional optimizations for

213   multiplex analyses. The amplified DNA samples were genotyped on a Beckman Coulter CEQ

214   8000 capillary sequencer. Genotypes were received from the CEQ 8000 Series Genetic

215   Analysing System 8.0.52. Paternities were determined by direct inspection of the allele size

216   estimates on a spreadsheet, and males that possessed two alleles in a microsatellite locus that

217   were not present in the offspring were excluded as a possible father (electronic supplementary

218   material, table S1-S4). We used the number of spawning decisions in figures and statistical

219   calculations i.e. if a brood was confirmed to be fathered by more than one male each male was

220   considered to be a spawning decision. F2 females in the analysed data had 4-8 broods each

221   and 5-15 spawning decisions. The complete datasets of the wild type females, F1 hybrid

222   females, F2 hybrid females and the males used in the experiment are included in the

223   electronic supplementary material, figure S2-3 and table S1-S4. We also provide pictures of

224   the F2 males from PN1-33 in figure S4.

225

226   *Statistics*

227   When analysing between-group preferences (*P*. sp. "pundamilia-like vs. *P*. sp. "nyererei-

228   like"; F1 hubrid females with *P*. sp. "pundamilia-like" mother vs. F1 hybrid females with *P*.

229   sp. "nyererei-like" mother), we, for each female, subtracted the number of spawning decisions

230   with males of *P*. sp. "nyererei-like" from the number of spawning decisions with males of *P*.

231   sp. "pundamilia-like" and analysed the differences with Mann Whitney U-tests.

232

233   Within-group preferences were analysed with Wilcoxon signed ranks tests on the individual's

234   number of spawning decisions with *P*. sp. "pundamilia-like" and *P*. sp. "nyererei-like". In one

235   F1 family, a binomial test was used due to the low number of spawning decisions per female.

236   The preference of individual F2 hybrid females were also analysed with binomial tests. We

237   could not estimate individual female preferences of wild type and F1 hybrid females given the

238   small number of decisions obtained from each female.

239

240

241   To test whether the F2 female spawning patterns deviated significantly from random, we

242   simulated a distribution of spawning decisions of the 69 females that had ≥ 5 spawning

243   decisions with either a *P*. sp. "pundamilia-like" (*Pp*) or a *P*. sp. "nyererei-like" (*Pn*) male. To

244   express the level of deviation from randomness, we calculated the consistency of an

245   individual's mate choice and calculated the repeatability (R) of a female's spawning

246   decisions. In quantitative genetics, the repeatability can be used to determine the upper-bound

247   estimate of the broad sense heritability ($H^2 = V_G/V_P$) [p. 136-138, 37]. The broad   sense

248   heritability indicates the relative proportion of total phenotypic variation of a trait (V_P) that

249   has a genetic basis (V_G). Repeatability is an upper-bound estimate of this heritability, given

250   that similarity in a trait value (in this case, consistent preference for males of one of the two

251   species) can both have a genetic and an environmental basis (e.g. a given female may prefer

252   males of a given species due to previous experiences). The model assesses the extent to which

253   a female's first spawning decision can predict her subsequent decisions, as this informs us on

254   how strong a mate preference has been expressed. In other words, the model assesses how

255   significantly the pattern of spawning decisions deviates from a random pattern (i.e. no

256   preference) when analysed across all F2 females at the population level. In the simulations,

257   each female is given a probability of mating with a *Pp* or a *Pn* male equivalent to the

258   proportion of *P*. sp. "pundamilia-like" and *P*. sp. "nyererei-like" spawning decisions

259   observed. This probability determines her first spawning decision. However, once a female

260   has been allocated a mate preference, the strength with which this preference continues to

261   affect subsequent spawning decisions is given by the following formulae:

262

$$P(x_i = Pp) = Pp + R(1 - Pp)$$

$$P(x_i = Pn) = Pn + R(1 - Pn)$$

263

264   Here, $P(x_i = Pp)$ and $P(x_i = Pn)$ are the probabilities of spawning with a *Pp* and a *Pn*

265   male at the ith spawning decision ($i > 1$), and *Pp* and *Pn* are the observed proportions   of

266   spawning decisions (across the entire population) with a *P*. sp. "pundamilia-like" and a *P*. sp.

267   "nyererei-like" male, respectively. *R* is the repeatability coefficient ($0 \leq R \leq 1$). With *R*=0,

268   spawning is "random" and proportional to the observed proportion of *Pn* and *Pp* spawning

269   decisions. In this case, female choice will switch randomly between *Pp* and *Pn* males. With

270   *R*=1, however, spawning choice is fixed and all spawning decisions are for males of the same

271   species as the first choice. In this case, females will consistently choose either a *Pp* or a *Pn*

272   male. With intermediate values of *R*, there is a preference for a species of male, but this

273   preference will not completely determine a spawning decision.

274

275   Furthermore, we also calculated if the number of individuals with preference for one species

276   differed from random. When categorizing female preference for males of either one of the two

277   species we used binomial tests and $\alpha = 0.05$ for the data set that included females with ≥ 6

278   spawning decisions.

279

280   To address potential parental and brood effects, all 100 F2 females were divided into two

281   categories: majority of spawnings with *P.* sp. "pundamilia-like" and majority of spawnings

282   with *P.* sp. "nyererei-like". Four females were omitted because they spawned equally many

283   times with males of the two species leaving 96 females (see the electronic supplementary

284   material figure S3). We used Binomial tests to ask if the female offspring of each of the three

285   F1 fathers were biased in their spawning decisions towards one of the two species, and χ^2 to

286   test if there was a difference between F2 females fathered by different F1 males. When

287   analysing the brood effect we restricted the analyses to the 9 broods with ≥ 4 F2 females and

288   performed 36 pairwise Fisher exact test comparisons and Bonferroni correction to correct for

289   multiple comparisons.

290

291   Statistics were performed in SPSS v. 23. The individual-based model was constructed in

292   Minitab 12.1.

293

297

298 Results

299 *Wild type females spawned with their own species*

300 There was a significant difference in spawning decisions between females of the two species

301 (Mann Whitney U test, n = 20, U = 0.00, p <0.001, the electronic supplementary material

302 figure S2a). The *P.* sp. "pundamilia-like" females had 1-3 spawning decisions each (median

303 2), and 14 out of 15 spawned only with conspecific males. One female mated once with *P.* sp.

304 "nyererei-like" and twice with conspecific males (Wilcoxon signed ranks test T = 0, n = 15, p

305 <0.001). The *P.* sp. "nyererei-like" females also had 1-3 spawning decisions each (median 3),

306 and all 6 spawned only with conspecific males (Wilcoxon signed ranks test T = 0, n = 6, p =

307 0.024).

308

309 *F1 hybrid females generally spawned with the species of their mother*

310 There was a significant difference in spawning decisions between the two F1 families (Mann

311 Whitney U test, n = 16, U = 2.50, p = 0.002, the electronic supplementary material figure

312 S2b). This was caused by F1 hybrid females spawning more often with the species of their

313 mothers (*P.* sp. "pundamilia-like" mother, 2-3 spawning decisions per female, median 2; 2

314 females spawned with both species, 9 with *P.* sp. "pundamilia-like" only, N = 11, Wilcoxon

315 signed ranks test, z = 45, p = 0.004, *P.* sp. "nyererei-like" mother, 1 spawning decision each,

316 all spawned with *P.* sp. "nyererei-like", two tailed Binomial test, n = 5, p = 0.063).

317

F2 hybrid spawning consistency suggests innate mating preference

319 When including females with ≥6 spawning decisions and α=0.05, 41 out of 59 F2 females had

320 a significant preference for males of one of the two species, whereas we would have expected

321 <3 if females mated randomly (Fisher exact test, p<0.001; figure 1). The simulation model

322 showed that the pattern of spawning decisions significantly deviated from a random pattern

323 when analysed at the population level. Spawning preferences segregated in an almost

324 Mendelian fashion and the majority of the females repeatedly spawned with one of the two

325 species (figure 1). The model estimates a repeatability of spawning decisions of R=0.7 (figure

326 2), which indicates that in our F2 population, 70% of the variation in spawning decisions is

327 explained by actual female mate preference.

328

329 To address potential parental effects, all 100 F2 females (the electronic supplementary

330 material figure S3) were divided into two categories: majority of spawnings with *P*. sp.

331 "pundamilia-like" and majority of spawnings with *P*. sp. "nyererei-like". The female

332 offspring of the 3 F1 males were not significantly biased towards preferring either of the two

333 species (16:27, 12:16 and 11:14, Binomial tests p=0.072, p=0.57 and p=0.69) and there was

334 no difference in ratios between the offspring of the 3 males (χ^2=0.384, df=2, p=0.82). The

335 experimental design of the present study did not allow us to confidently rule out that females

336 from different broods differed in preferences, because most broods were small. However, the

337 data rule out a general maternal effect. When restricting the analyses to broods with ≥4

338 females, 4 out of 36 pairwise comparisons between broods yield p<0.05 with the lowest p

339 being p=0.015. All these are far from significant when correcting for multiple comparisons.

340 Furthermore, while their F2 brothers show considerable colour segregation within broods,

341 there is no indication of a strong correlation between a female's preference and the colour

342 phenotype of her brothers (electronic supplementary material figure S4).

343

344 *There is no sign of copying of previous choice*

345 Only 26 out of the 69 F2 females with ≥ 5 spawning decisions spawned with both species. Of

346 those females, 21 switched back and forth between species (figure 1). This demonstrates that

347 females do not simply copy their first mate choice or their most recent choice. In other words,

348 the high repeatability of mate choice decision is unlikely to be the result of copying a previous

349 choice.

350

351 Six of the 16 F2 females with a significant mating preference, which were enclosed with a

352 male of the non-preferred species, did not spawn at all, and one female that did spawn, did not

353 spawn again when reintroduced to the large choice experiment tank. The nine females that

354 had spawned in the no-choice situation against their preference and subsequently spawned

355 again in the choice experiment, all reverted to spawning with males of the previously

356 preferred species (*P*. sp. "pundamilia-like" preferring N=3, *P*. sp. "nyererei-like" preferring

357 N=6, Two tailed Binomial test p = 0.004) which highlights the innate strength of female mate

358 preference.

359

360 Discussion

361 The genetics of female mate preferences is likely to affect evolutionary processes, including

362 speciation and hybridisation between species. We report a long term common garden study

363 where we followed spawning decisions of F2 hybrid females between two sympatric sister

364 species throughout a large part of their reproductive lives as well as spawning decisions of F1

365 hybrid females. Wild type females of both species were used as a control.

366

367   Using molecular paternity testing, our experiments indicated that wild-type females mostly

368   mated with conspecific males, although mating was not 100% assortative. This is consistent

369   with the results of previous studies on the same population using mating experiments [30] or

370   behavioural preference assays [9, 30, 42, 43], and indicates that either method can be used

371   reliably to estimate preferences. The occasional disassortative mating is also consistent with

372   modelling based on population genomic data suggesting ongoing gene flow between the same

373   sympatric species in nature, as well as between allopatric populations [19].

374

375   All F1 hybrid females mated with their maternal species, although a couple of them also

376   mated with the paternal species. This bias towards the maternal species is consistent with an

377   effect of imprinting, which had previously been demonstrated in Lake Victoria

378   haplochromines using controlled cross-fostering experiments with mate preferences assayed

379   with a behavioural choice test [44, 45]. Our results are, however, also consistent with the

380   possibility that genes influencing species-specific preferences were not entirely reciprocally

381   fixed between the wild-type individuals used to breed our F1 hybrids, e.g. as a result of

382   occasional introgression [19, 26]. It is not impossible that one of the parents of our two test F1

383   families may have been heterozygous at a mate preference locus, and that thus some of the F1

384   hybrid females were homozygous.

385

386   By contrast, the experimental design limited the potential for any imprinting of species-

387   specific preferences in F2 hybrids, since their mothers were all F1 hybrids. Furthermore, we

388   found that siblings in most families exhibited consistent preferences for males of different

389   species, which is inconsistent with imprinting. Likewise, our experimental test of the

390   preferences of females following a 'no-choice' mating with the non-preferred male species

391   indicated that females retained their original preferences in a subsequent free choice

392   experiment, suggesting that experience did not disrupt their innate preferences. In general,

393   many F2 hybrid females were consistent in choosing males of a particular species, with 41 out

394   of 59 females showing a significant preference, far more than the 3 expected if females had

395   mated by chance. This clear nearly Mendelian segregation in spawning preferences in the F2

396   generation is consistent with previous behavioural choice tests by Haesler and Seehausen [9].

397   The Mendelian segregation despite incomplete genetic isolation [17, 19, 26] and

398   recombination (J. Meier and O. Seehausen, unpublished data) in this species pair in the wild

399   implies that species-specific female mate choice among the *Pundamilia* sister species is

400   influenced by relatively few major genes or genomic regions containing several tightly linked

401   loci.

402

403   Repeatability and the heritability of mate choice

404   Our simulation indicated that the distribution of spawning decisions over F2 hybrid females

405   deviated significantly from expectations if mating was random when analysed at the

406   population level. A large excess of females showed a significant preference for males of either

407   one of the two species. Female choice of certain type of males within a species often has low

408   repeatability and is subject to change depending on e.g. experience, age, condition, mate

409   copying and the environment [46-48]. In our experiment, repeatability of spawning decisions

410   of F2 hybrid females was high (70 %) and preferences did not change over time and over

411   successive reproductive cycles of females, nor after the experience of a successful spawning

412   event with a male of the non-preferred species. Repeatability is also often used to determine

413   the upper-bound estimate of the broad sense heritability (H^2) in behavioural studies [46, 47].

414   The results from our simulation therefore suggest that up to 70% of the variation in spawning

415   decisions observed among the F2 hybrid females may have a heritable basis. However, the

416   remaining 30% could simply be due to lack of a consistent preference in the class of

417   preference heterozygote F2 hybrid females – these are expected to mate randomly [9].

418   Therefore, heritability may be higher than the estimated 70% [9, 49]. In the experimental

419   design, we aimed to minimize environmental variation introduced by differences in condition

420   between males by providing a choice among eight males, four of each species in each trial.

421   Differences in territory quality were unlikely in the standardised conditions of our experiment.

422   Thus, we conclude that the observed among-female variation in preference is likely to be due

423   to genetic factors.

424

425   **Sexual isolation by mate choice**

426   Behavioural reproductive isolation is of key importance to understanding the rapid evolution

427   of genetically differentiated sympatric species [1, 41, 50], such as those in African cichlid fish

428   radiations. The species pair that we studied here has been estimated to have arisen in just

429   slightly more than 150 generations, facilitated by hybridisation between the local *P.*

430   *pundamilia* and migrants of *P. nyererei* from around Makobe island [19].

431

432   The male trait (red dorsum vs. blue colour) that species-assortative female mating preferences

433   are based on [30] is likely oligogenic itself [31]. Theoretical work suggests that it is easier for

434   divergent selection to overcome homogenizing gene flow if traits under divergent selection

435   are due to relatively few genes, because the fewer genes that are responsible for a trait under

436   divergent selection, the higher are the selection coefficients for each locus [51-53]. Hence, the

437   genetic architecture of mate choice and mating traits in *Pundamilia* may facilitate speciation

438   in the face of gene flow, perhaps in combination with other selection pressures, as might be

439   generated by adaptation to divergent microhabitats, particularly water depths: field studies

440   have shown that red dorsum males tend to be found in deeper water than the blue males [26].

441

442 **Candidate genes for mate choice**

443 Candidate genes relating to species-specific mate preferences are likely to include those

444 affecting vision. Divergence has been shown in the long wavelength sensitive opsin gene

445 (*LWS*) [26]. In the red vs. blue species pair at Makobe Island, there is also divergence in the

446 short wavelength sensitive opsin gene (*SWS2A*) but this is not currently known in the species

447 pair of the present study [26]. At Makobe Island, there is also divergence in other putative

448 coding regions [18], some of which may be related to vision.

449

450 Many small genomic 'islands of differentiation' were found to differentiate *P. pundamilia* and

451 *P. nyererei* from Makobe Island [18]. However, the Python Island species pair having

452 recently (around 150 generations ago) re-emerged after a period of massive introgression

453 might be expected to be divergent at fewer regions, more directly related to divergent

454 selection pressures, which should make traits directly related to reproductive isolation easier

455 to detect. Malinsky et al. [54] identified several genomic regions with high differentiation in

456 two young ecomorphs of crater lake haplochromine cichlids (genus *Astatotilapia)* with partial

457 assortative mating. Candidate adaptive genes in these so called 'genomic islands of

458 differentiation' included rhodopsin and other twilight-vision-associated genes. Differentiation

459 in such 'islands' can resist ongoing gene flow, as shown in < 150 year old incipient

460 *Gasterosteus* stickleback species pairs in two Swiss lakes [55, 56].

461

462 **To conclude**

463 We show in a common garden long term mating experiment that strong female mating

464 preferences for males of either one of two sister species are recovered in large fractions of the

465 F2 hybrid generation. The genetic assays of mate choice in F2 hybrids between *P.* sp.

466 "pundamilia-like" and *P.* sp. "nyererei-like" show high repeatability and consistency in

467   female choice across many reproductive cycles, and we argue that the variation is influenced

468   by the segregation of a few genes with large effects. We propose that a simple genetic basis

469   could help facilitate stable phenotypic differentiation in sympatry in the face of gene flow.

470

471 Data accessibility

472   The complete datasets of the wild type, F1 and F2 females, and the males used in the

473   experiment are included in figure 1 and the electronic supplementary material, figure S2-3

474   and table S1-S4. The raw data in the electronic supplementary material, table S1-S4 are also

475   available from the Dryad Digital Repository http://dx.doi.org/10.5061/dryad.q58hr [57]. The

476   Minitab 12.1 macro to test the repeatability of mate choice is deposited at GitHub

477   https://github.com/Ward9250/FishSpawn

478

479

480 Authors' contribution

481   G.F.T. and O.Se. conceived the project, O.Sv., G.F.T. and O.Se. designed the experiments,

482   K.W. and A.S. carried out the crosses, O.Sv., K.W. and A.S. performed the experiments,

483   O.Sv. carried out microsatellite paternity analyses and processed the data, C.v.O. wrote the

484   simulation model, O.Sv. and C.v.O. analysed the data, O.Sv. wrote the manuscript with

485   important contribution from C.v.O, G.F.T. and O.Se. All authors have provided critical

486   revision of the manuscript and approved the final version.

487

488 Competing interest

489   We declare no competing interest

490

491 Funding

504   References

505

506   1.   Coyne J.A., Orr H.A. 2004 *Speciation*. Sunderland, Sinauer.

507   2.   Nosil P. 2012 *Ecological Speciation*. Oxford, Oxford University Press; 280 p.

508   3.   Arnegard M.E., Kondrashov A.S. 2004 Sympatric speciation by sexual selection

509   alone is unlikely. *Evolution* **58**, 222-237. (doi:10.1111/j.0014-3820.2004.tb01640.x).

510   4.   Doebeli M. 2005 Adaptive speciation when assortative mating is based on

511   female preference for male marker traits. *J Evol Biol* **18**, 1587-1600. (doi:10.1111/j.1420-

512   9101.2005.00897.x).

513   5.   Lande R. 1981 Models of speciation by sexual selection on polygenic traits.

514   *Proc Natl Acad Sci USA* **78**, 3721-3725. (doi:10.1073/pnas.78.6.3721).

515   6.   Higashi M., Takimoto G., Yamamura N. 1999 Sympatric speciation by sexual

516   selection. *Nature* **402**, 523-526. (doi:10.1038/990087).

517   7.   Kawata M., Shoji A., Kawamura S., Seehausen O. 2007 A genetically explicit

518   model of speciation by sensory drive within a continuous population in aquatic environments.

519   *BMC Evol Biol* **7**, 99. (doi:10.1186/1471-2148-7-99).

520   8.   Thibert-Plante X., Gavrilets S. 2013 Evolution of mate choice and the so-called

521   magic traits in ecological speciation. *Ecol Lett* **16**, 1004-1013. (doi:10.1111/ele.12131).

522   9.   Haesler M.P., Seehausen O. 2005 Inheritance of female mating preference in a

523   sympatric sibling species pair of Lake Victoria cichlids: implications for speciation. *Proc R*

524   *Soc B* **272**, 237-245.

525   10.   Kronforst M.R., Young L.G., Kapan D.D., McNeely C., O'Neill R.J., Gilbert

526   L.E. 2006 Linkage of butterfly mate preference and wing color preference cue at the genomic

527   location of wingless. *Proc Natl Acad Sci USA* **103**, 6575-6580.

528   (doi:10.1073/pnas.0509685103).

529   11.   Merrill R.M., Van Schooten B., Scott J.A., Jiggins C.D. 2011 Pervasive genetic

530   associations between traits causing reproductive isolation in *Heliconius* butterflies. *Proc R*

531   *Soc Lond, Ser B: Biol Sci* **278**, 511-518. (doi:10.1098/rspb.2010.1493).

532   12.   Ding B., Daugherty D.W., Husemann M., Chen M., Howe A.E., Danley P.D.

533   2014 Quantitative genetic analyses of male color pattern and female mate choice in a pair of

534   cichlid fishes of Lake Malawi, East Africa. *Plos One* **9**. (doi:e114798

535   10.1371/journal.pone.0114798).

536   13.   Diao W.W., Mousset M., Horsburgh G.J., Vermeulen C.J., Johannes F., van de

537   Zande L., Ritchie M.G., Schmitt T., Beukeboom L.W. 2016 Quantitative trait locus analysis

538   of mating behavior and male sex pheromones in *Nasonia* wasps. *G3-Genes Genomes Genet* **6**,

539   1549-1562.   (doi:10.1534/g3.116.029074).

540   14.   Laturney M., Moehring A.J. 2012 Fine-scale genetic analysis of species-specific

541   female preference in *Drosophila simulans*. *J Evol Biol* **25**, 1718-1731. (doi:10.1111/j.1420-

542   9101.2012.02550.x).

543   15.   Oh K.P., Fergus D.J., Grace J.L., Shaw K.L. 2012 Interspecific genetics of

544   speciation phenotypes: song and preference coevolution in Hawaiian crickets. *J Evol Biol* **25**,

545   1500-1512.   (doi:10.1111/j.1420-9101.2012.02531.x).

546   16.   Maan M.E., Seehausen O. 2010 Mechanisms of species divergence through

547   visual adaptation and sexual selection: Perspectives from a cichlid model system. *Curr Zool*

548   **56**, 285-299.

549   17.   Seehausen O. 2009 Progressive levels of trait divergence along a 'speciation

550   transect' in the Lake Victoria cichlid fish *Pundamilia*. In *Ecological Reviews: Speciation and

551   *Patterns of Diversity* (eds. Butlin R., Bridle J., Schluter D.), pp. 155-176. Cambridge,

552   Cambridge University Press.

553   18.   Brawand D., Wagner C.E., Li Y.I., Malinsky M., Keller I., Fan S., Simakov O.,

554   Ng A.Y., Lim Z.W., Bezault E., et al. 2014 The genomic substrate for adaptive radiation in

555   African cichlid fish. *Nature* **513**, 375-381. (doi:10.1038/nature13726).

556   19.   Meier J.I., Sousa V.C., Marques D.A., Selz O.M., Excoffier L., Seehausen O.

557   2016 Demographic modeling of whole genome data reveals parallel origin of similar

558   *Pundamilia* cichlid species after hybridization. *Mol Ecol* **Epub ahead of print**.

559   (doi:10.1111/mec.13838. Epub ahead of print).

560   20.   Seehausen O. 1996 *Lake Victoria rock cichlids - taxonomy, ecology and

561   distribution*. Zevenhuizen, Verduijn Cichlids; 304 p.

562   21.   Andersson M. 1994 *Sexual selection*. Princeton, New Jersey, Princeton

563   University Press.

564 22.   Seehausen O. 2006 African cichlid fish: a model system in adaptive radiation

565 research. *Proc R Soc Lond, Ser B: Biol Sci* **273**, 1987-1998. (doi:10.1098/rspb.2006.3539).

566 23.   Selz O.M., Thommen R., Pierotti M.E.R., Anaya-Rojas J.M., Seehausen O.

567 2016 Differences in male coloration are predicted by divergent sexual selection between

568 populations of a cichlid fish. *Proc R Soc Lond, Ser B: Biol Sci* **283**.

569 (doi:10.1098/rspb.2016.0172).

570 24.   Knight M.E., Turner G.F. 2004 Laboratory mating trials indicate incipient

571 speciation by sexual selection among populations of the cichlid fish *Pseudotropheus zebra*

572 from Lake Malawi. *Proc R Soc B* **271**, 675-680. (doi:10.1098/rspb.2003.2639).

573 25.   Kidd M.R., Danley P.D., Kocher T.D. 2006 A direct assay of female choice in

574 cichlids: all the eggs in one basket. *J Fish Biol* **68**, 373-384. (doi:10.1111/j.1095-

575 8649.2005.00896.x).

576 26.   Seehausen O., Terai Y., Magalhaes I.S., Carleton K.L., Mrosso H.D.J., Miyagi

577 R., van der Sluijs I., Schneider M.V., Maan M.E., Tachida H., et al. 2008 Speciation through

578 sensory drive in cichlid fish. *Nature* **455**, 620-U623. (doi:10.1038/nature07285).

579 27.   Blows M.W., Brooks R., Kraft P.G. 2003 Exploring complex fitness surfaces:

580 Multiple ornamentation and polymorphism in male guppies. *Evolution* **57**, 1622-1630.

581 28.   Brooks R., Hunt J., Blows M.W., Smith M.J., Bussiere L.F., Jennions M.D.

582 2005 Experimental evidence for multivariate stabilizing sexual selection. *Evolution* **59**, 871-

583 880.   (doi:10.1111/j.0014-3820.2005.tb01760.x).

584 29.   Candolin U. 2003 The use of multiple cues in mate choice. *Biol Rev* **78**, 575-

585 595.

586 30.   Selz O.M., Pierotti M.E.R., Maan M.E., Schmid C., Seehausen O. 2014 Female

587 preference for male color is necessary and sufficient for assortative mating in 2 cichlid sister

588 species *Behav Ecol* **25**, 612-626. (doi:10.1093/beheco/aru024).

589   31.   Magalhães I.S., Seehausen O. 2010 Genetics of male nuptial colour divergence

590   between sympatric sister species of a Lake Victoria cichlid fish. *J Evol Biol* **23**, 914-924.

591   32.   Seehausen O., vanAlphen J.J.M., Witte F. 1997 Cichlid fish diversity threatened

592   by eutrophication that curbs sexual selection. *Science* **277**, 1808-1811.

593   33.   Stelkens R.B., Pierotti M.E.R., Joyce D.A., Smith A.M., van der Sluijs I.,

594   Seehausen O. 2008 Disruptive sexual selection on male nuptial coloration in an experimental

595   hybrid population of cichlid fish. *Philos Trans R Soc Lond, B, Biol Sci* **363**, 2861-2870.

596   (doi:10.1098/rstb.2008.0049).

597   34.   Barson N.J., Knight M.E., Turner G.F. 2007 The genetic architecture of male

598   colour differences between a sympatric Lake Malawi cichlid species pair. *J Evol Biol* **20**, 45-

599   53.   (doi:10.1111/j.1420-9101.2006.01228.x).

600   35.   Svensson O., Egger B., Gricar B., Woodhouse K., van Oosterhout C.,

601   Salzburger W., Seehausen O., Turner G. 2011 Segregation of species-specific male

602   attractiveness in F2 hybrid Lake Malawi cichlid fish. *Int J Evol Biol* **2011**, 1-7.

603   (doi:10.4061/2011/426179).

604   36.   van der Sluijs I., Dijkstra P.D., Lindeyer C.M., Visser B., Smith A.M.,

605   Groothuis T.G.G., van Alphen J.J.M., Seehausen O. 2013 A test of genetic association among

606   male nuptial coloration, female mating preference, and male aggression bias within a

607   polymorphic population of cichlid fish. *Curr Zool* **59**, 221-229.

608   37.   Falconer D.S., Mackay T.F.C. 1996 *Introduction to Quantitative Genetics*. 4 ed.

609   Harlow, Essex, UK, Pearson Education Limited; 480 p.

610   38.   Lande R. 1981 The minimum number of genes contributing to quantitative

611   variation between and within populations. *Genetics* **99**, 541-553.

612   39.   Turelli M., Barton N.H. 1994 Genetic and statistical analyses of strong selection

613   on polygenic traits: what, me normal? *Genetics* **138**, 913-941.

614   40.   Felsenstein J. 1981 Skepticism towards Santa Rosalia, or why are there so few

615   kinds of animals? *Evolution* **35**, 124-138. (doi:10.2307/2407946).

616   41.   Smadja C.M., Butlin R.K. 2011 A framework for comparing processes of

617   speciation in the presence of gene flow. *Mol Ecol* **20**, 5123-5140. (doi:10.1111/j.1365-

618   294X.2011.05350.x).

619   42.   Seehausen O., van Alphen J.J.M. 1998 The effect of male coloration on female

620   mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex).

621   *Behav Ecol Sociobiol* **42**, 1-8.

622   43.   van der Sluijs I., Van Dooren T.J.M., Hofker K.D., van Alphen J.J.M., Stelkens

623   R.B., Seehausen O. 2008 Female mating preference functions predict sexual selection against

624   hybrids between sibling species of cichlid fish. *Philos Trans R Soc Lond, Ser B: Biol Sci* **363**,

625   2871-2877.   (doi:10.1098/rstb.2008.0045).

626   44.   Verzijden M.N., Korthof R.E.M., ten Cate C. 2008 Females learn from mothers

627   and males learn from others. The effect of mother and siblings on the development of female

628   mate preferences and male aggression biases in Lake Victoria cichlids, genus *Mbipia*. *Behav*

629   *Ecol Sociobiol* **62**, 1359-1368. (doi:10.1007/s00265-008-0564-x).

630   45.   Verzijden M.N., ten Cate C. 2007 Early learning influences species assortative

631   mating preferences in Lake Victoria cichlid fish. *Biol Lett* **3**, 134-136.

632   46.   Jennions M.D., Petrie M. 1997 Variation in mate choice and mating preferences:

633   A review of causes and consequences. *Biol Rev* **72**, 283-327.

634   (doi:10.1017/s0006323196005014).

635   47.   Bell A.M., Hankison S.J., Laskowski K.L. 2009 The repeatability of behaviour:

636   a meta-analysis. *Animal Behaviour* **77**, 771-783. (doi:10.1016/j.anbehav.2008.12.022).

637   48.   Brooks R. 1996 Copying and the repeatability of mate choice. *Behav Ecol*

638   *Sociobiol* **39**, 323-329. (doi:10.1007/s002650050296).

639   49.   Dohm M.R. 2002 Repeatability estimates do not always set an upper limit to

640   heritability. *Funct Ecol* **16**, 273-280.

641   50.   Servedio M.R. 2016 Geography, assortative mating, and the effects of sexual

642   selection on speciation with gene flow. *Evolutionary Applications* **9**, 91-102.

643   (doi:10.1111/eva.12296).

644   51.   Gavrilets S., Vose A. 2005 Dynamic patterns of adaptive radiation. *Proc Natl*

645   *Acad Sci USA* **102**, 18040-18045. (doi:10.1073/pnas.0506330102).

646   52.   Yeaman S., Otto S.P. 2011 Establishment and maintenance of adaptive genetic

647   divergence under migration, selection, and drift. *Evolution* **65**, 2123-2129.

648   (doi:10.1111/j.1558-5646.2011.01277.x).

649   53.   Yeaman S., Whitlock M.C. 2011 The genetic architecture of adaptation under

650   migration-selection balance. *Evolution* **65**, 1897-1911. (doi:10.1111/j.1558-

651   5646.2011.01269.x).

652   54.   Malinsky M., Challis R.J., Tyers A.M., Schiffels S., Terai Y., Ngatunga B.P.,

653   Miska E.A., Durbin R., Genner M.J., Turner G.F. 2015 Genomic islands of speciation

654   separate cichlid ecomorphs in an East African crater lake. *Science* **350**, 1493-1498.

655   (doi:10.1126/science.aac9927).

656   55.   Marques D.A., Lucek K., Meier J.I., Mwaiko S., Wagner C.E., Excoffier L.,

657   Seehausen O. 2016 Genomics of rapid incipient speciation in sympatric threespine

658   stickleback. *PLoS Genetics* **12**, e1005887.   (doi:10.1371/journal.pgen.1005887).

659   56.   Marques D.A., Lucek K., Haesler M.P., Feller A.F., Meier J.I., Wagner C.E.,

660   Excoffier L., Seehausen O. 2016 Genomic landscape of early ecological speciation initiated

661   by selection on nuptial colour. *Mol Ecol* **Epub ahead of print**. (doi:10.1111/mec.13774).

662   57.   Svensson O., Woodhouse K., van Oosterhout C., Smith A., Turner G.F.,

663   Seehausen O. 2017 The genetics of mate preferences in hybrids between two young and

664   sympatric Lake Victoria cichlid species. *Dryad Digital Repository*.

665   (doi:doi:10.5061/dryad.q58hr).

666

667

668

669 **Figure legends**

670

671

672 **Figure 1.** Individual spawning decisions by the 69 F2 hybrid females. Spawning decisions

673 were determined by microsatellite DNA paternity analyses. Above the line y=0 is the number

674 of spawning decisions with *P*. sp. "pundamilia-like", and below the line is the number of

675 spawning decisions with *P*. sp. "nyererei-like". The order of spawning decisions is arranged

676 with the first spawning on the top, and the last on the bottom with a spawning decision with

677 *P*. sp. "pundamilia-like", marked in blue and a spawning decision with *P*. sp. "nyererei-like"

678 marked in red. *$p<0.05$, [a]$0.05<p<0.1$.

679

680

Figure 2. (A) Simulated (means and 5-95% error bars) spawning decisions of F2 females with

P. sp. "pundamilia-like" (blue dots), and with *P*. sp. "nyererei-like" (red dots) based on a

repeatability of an individual's spawning decision of *R*=0.7. Observed ratio of spawning

decisions is shown by the solid black lines. (B) The best fit of the model is with *R*=0.7, which

minimises the mean squares (MS) between the observed and simulated spawning distribution.

Lower values of *R* produce a more random spawning pattern, whilst higher values of *R*

increase the consistency of a females' spawning choices above those observed, which reduced

the fit of the model by inflating the MS.

Number of spawning decisions with *P.* sp. "pundamilia-like"

Number of spawning decisions with *P.* sp. "nyererei-like"

Order of spawning decisions

One spawning decision with *P.* sp. "pundamilia-like"

One spawning decision with *P.* sp. "nyererei-like"

Number of spawning decisions with *P.* sp. "pundamilia-like" / with *P.* sp. "nyererei-like"

- ● Simulated spawning decisions with *P.* sp. "nyererei python"
- ● Simulated spawning decisions with *P.* sp. "pundamilia python"

F2-female