Chromosomal distribution of cyto-nuclear genes in a dioecious plant with sex chromosomes

Josh Hough[*†], J. Arvid Ågren[*§], Spencer C.H. Barrett, and Stephen I. Wright

Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada, M5S 3B2

[*] These authors contributed equally to this work; corresponding authors

[†] josh.hough@utoronto.ca

[§] arvid.agren@utoronto.ca

Running title: Cyto-nuclear interactions in a dioecious plant

Key words: Cyto-nuclear interactions, co-adaptation, sexual conflict, gene transfer

Word count: 2421

Abstract word count: 245

Number of figures: 1

Abstract

The coordination between nuclear and organellar genes is essential to many aspects of eukaryotic life, including basic metabolism, energy production, and ultimately, organismal fitness. Whereas nuclear genes are bi-parentally inherited, mitochondrial and chloroplast genes are almost exclusively maternally inherited, and this asymmetry may lead to a bias in the chromosomal distribution of nuclear genes whose products act in the mitochondria or chloroplasts. In particular, because X-linked genes have a higher probability of co-transmission with organellar genes (2/3) compared to autosomal genes (1/2), selection for co-adaptation has been predicted to lead to an over-representation of nuclear-mitochondrial (N-mt) or nuclear-chloroplast (N-cp) genes on the X chromosome relative to autosomes. In contrast, the occurrence of sexually antagonistic organellar mutations might lead to selection for movement of cyto-nuclear genes from the X chromosome to autosomes to reduce male mutation load. Recent broad-scale comparative studies of N-mt distributions in animals have found evidence for these hypotheses in some species, but not others. Here, we use transcriptome sequences to conduct the first study of the chromosomal distribution of cyto-nuclear interacting genes in a plant species with sex chromosomes (*Rumex hastatulus*; Polygonaceae). We found no evidence of under- or over-representation of either N-mt or N-cp genes on the X chromosome, and thus no support for either the co-adaptation or the sexual-conflict hypothesis. We discuss how our results from a species with recently evolved sex chromosomes fit into an emerging picture of the evolutionary forces governing the chromosomal distribution of N-mt and N-cp genes.

Introduction

The intimate relationships between nuclear and organellar genomes in eukaryotes represent some of the most striking examples of co-evolved mutualisms (Gillham 1994; Lane 2005; Aanen et al. 2014). The long co-evolutionary history of nuclear and mitochondrial genomes is perhaps best illustrated by the finding that the vast majority of mitochondrial genes in animals have been transferred to the nuclear genome (Adams and Palmer 2003; Rand et al. 2004; Burt and Trivers 2006). Indeed, animal mitochondria now encode only a few proteins after having lost the majority of their original genes (Berg and Kurland 2000; Ridley 2000; Bar-Yaacov et al. 2012). Moreover, almost one fifth of the *Arabidopsis thaliana* nuclear genome is of chloroplast origin (Martin 2003), suggesting that organellar-to-nuclear gene movement has played a crucial role in the evolution of plant genetic systems.

The evolution of cyto-nuclear interactions and the chromosomal distribution of the genes involved should be influenced by the contrasting modes of inheritance of organellar genes (maternal inheritance) and autosomal genes (bi-parental inheritance). This difference may, for example, result in conflict between nuclear and organellar genes over sex determination and sex ratio (Cosmides and Tooby 1981; Werren and Beukeboom 1998), and several mitochondrial genes in plants are known to cause male sterility (Burt and Trivers 2006; Touzet and Meyer 2014). In systems with XY sex determination, where males are the heterogametic (XY) and females the homogametic sex (XX), genes on the X chromosome spend 2/3 of their time in females (Rand et al. 2001) and therefore share a female-biased inheritance pattern relative to Y-linked or autosomal genes, which may result in inter-genomic co-adaptation and or conflict.

70 A potential consequence of inter-genomic conflict and co-adaptation between
71 cyto-nuclear genes and other regions of the genome is a shift in the chromosomal location
72 of cyto-nuclear genes, either becoming more or less abundant on the X chromosome. This
73 idea has been explored by several recent studies (Drown et al. 2012; Hill and Johnson
74 2013; Dean et al. 2014; Rogell et al. 2014), and two main processes have been proposed
75 to account for the movement of genes to or from the X chromosome. The co-adaptation
76 hypothesis predicts that the co-transmission of X-linked and organellar genes should
77 result in selection for the co-adaptation of such genes, thus leading to an over-
78 representation of N-mt and N-cp genes on the X chromosome relative to autosomes
79 (Rand et al. 2004; Drown et al. 2012). In contrast, the sexual conflict hypothesis predicts
80 the opposite chromosomal distribution, with more cyto-nuclear genes occurring on
81 autosomes to alleviate mutation load in males. To date, empirical evidence for the above
82 hypotheses are mixed. Drown et al. (2012) used previously published reference genomes
83 to examine the chromosomal distribution of N-mt genes in 16 vertebrates and found a
84 strong under-representation of such genes on the X chromosomes relative to autosomes in
85 14 mammal species, but not in two avian species with ZW sex determining systems.
86 Dean et al. (2014) included seven additional species in their analysis with independently
87 derived sex chromosomes, and found that the under-representation of N-mt genes on the
88 X chromosome was restricted to therian mammals and *Caenorhabditis elegans*.
89 Here, we use sex-linked and autosomal transcriptome sequences to investigate the
90 chromosomal distributions of cyto-nuclear interactions in the dioecious annual plant
91 *Rumex hastatulus* (Polygonaceae). Examining cyto-nuclear interactions within a plant
92 species is of interest for several reasons (see Sloan 2014). First, plants carry an additional

93 maternally inherited organellar genome that is absent in animals, the chloroplast genome.
94 This provides an opportunity to compare the chromosomal distribution of two
95 independent kinds of cyto-nuclear interacting genes: nuclear-mitochondrial (N-mt) and
96 nuclear-chloroplast (N-cp). Second, whereas animal sex chromosomes evolved hundreds
97 of millions of years ago (180 MYA in mammals and 140 MYA in birds; Cortez et al.
98 2014), the origin of plant sex chromosomes is a more recent event (Charlesworth 2013).
99 In *R. hastatulus,* sex chromosomes are thought to have evolved approximately 15-16
100 MYA (Navajas-Perez et al. 2005) and genes on the Y chromosome show evidence of
101 degeneration, resulting in a considerable proportion of genes that are hemizygous on the
102 X chromosome (Hough et al. 2014). *Rumex hastatulus* therefore provides an opportunity
103 to test whether the early changes involved in sex chromosome evolution have been
104 associated with a concomitant shift in the chromosomal location of N-mt or N-cp genes.
105 Moreover, the presence in this system of X-linked genes that have recently become
106 hemizygous provides an opportunity to compare the chromosomal distributions of X-
107 linked genes that are hemizygous versus those that have retained Y-linked alleles (X/Y
108 genes). Hemizygous genes are particularly good candidates for evaluating evidence for
109 co-adaptation and/or sexual conflict because of their relatively older age (Hough et al.
110 2014), and because beneficial mutations in such genes are exposed to positive selection
111 regardless of dominance and may therefore spread more rapidly.
112
113
114
115

116 **Methods**

117 *Gene identification and functional annotation*

118 We used sex-linked and autosomal transcriptome sequence data for *R. hastatulus* reported
119 in Hough et al. (2014; GenBank Sequence Read Archive accession no. SRP041588), and
120 obtained three sets of genes with which to test for an over- or under-representation of
121 nuclear-mitochondrial (N-mt) or nuclear-chloroplast (N-cp) genes. In total our analyses
122 included 1167 autosomal genes, 624 X-linked genes, and 107 hemizygous X-linked genes
123 (see Hough et al. 2014; *Methods* and *SI Appendix* for full details regarding the
124 identification of such genes from transcriptome sequence data). For each gene set, we
125 queried the sequences translated in all reading frames against the *A. thaliana* protein
126 database using the BLASTx homology search implemented in Blast2GO (Conesa et al.
127 2005), with a significance threshold (BLAST ExpectValue) of $1x10^{-3}$, above which
128 matches were not reported. We limited our searches to the *A. thaliana* protein database
129 because sequence matches to this database returned more detailed functional information
130 than is available for most other species in the NCBI plant database. We obtained
131 BLASTx results for 1073 autosomal genes (90%), 567 X-linked genes (90%), and 95
132 hemizygous genes (89%). Gene Ontology (GO) terms associated with the hits from
133 BLASTx queries were then retrieved using the 'Mapping' function in Blast2GO, which
134 used BLAST accessions to link the queried sequences to functional information stored in
135 the GO database (The Gene Onotology Consortium 2008). Gene names were retrieved
136 using NCBI mapping files 'gene info' and 'gene2accession', and GO terms were assigned
137 to query sequences using the 'Annotation' function with an E-Value-Hit-Filter of $1x10^{-6}$
138 and an annotation cut off of 55 (default parameters). Finally, we ran InterProScan

139 (Quevillon et al. 2005) to retrieve sequence domain/motif information and merged the
140 corresponding annotations with previously identified GO terms. This procedure generated
141 output files containing GO ID's and functional descriptions for each gene in our data set
142 (files will be uploaded to GitHub).
143
144 *Statistical analyses*
145 We used a similar approach to Drown et al. (2012) and Dean et al. (2014) and estimated
146 the number of N-mt and N-cp genes on the X chromosome and autosomes, and then
147 compared each of these estimates to an expected number. The expected number of N-mt
148 genes was obtained by calculating the product of the proportion of all genes in the data
149 set with mitochondrial annotations (matching GO:0005739) and the number of annotated
150 genes in a given gene set. The expected numbers of N-cp genes were calculated similarly,
151 using GO:0009507. We then calculated the ratios of the observed-to-expected numbers
152 for both N-mt and N-cp genes in each gene set. The observed-to-expected ratio is
153 expected to equal one when there is no under- or over-representation, and greater than
154 one when there is an over-representation. We note that, unlike for X-linked genes, we did
155 not have information regarding the particular chromosome locations for autosomal genes,
156 and therefore could not obtain the expected numbers of N-mt and N-cp genes per-
157 autosome as in previous studies (Drown et al. 2012; Dean et al. 2014). The expected
158 numbers were thus calculated assuming that the set of autosomal genes represented a
159 random sample of the autosomal chromosomes in this species, which is likely a valid
160 assumption given that the sequences were obtained using whole transcriptome shotgun
161 sequencing (Hough et al. 2014). Calculating the expected-to-observed ratios across X-

162 linked, autosomal, and X-hemizygous genes thus allowed us to determine whether any of
163 these gene sets contained an under- or over-representation of N-mt and N-cp genes
164 compared to the expectation based on the proportion of such genes in the full data set. We
165 tested the significance of over- or under-representation using Fisher's exact tests, and
166 calculated 95% confidence intervals for the numbers of N-mt or N-cp genes using 10,000
167 replicate bootstrapped samples. All data analysis was done in R (R Development Core
168 Team 2013; scripts will be available for download from GitHub).
169

170 **Results and Discussion**

171 It has been suggested that cyto-nuclear genes may be either over- or under-represented on
172 the X chromosome compared to autosomes, depending on whether their interactions are
173 driven by co-adaptation or sexual conflict (Rand et al. 2001; Drown et al. 2012; Hill and
174 Johnson 2013; Dean et al. 2014; Rogell et al. 2014). We annotated sex-linked and
175 autosomal transcriptome sequences to test these predictions in the dioecious plant *R.*
176 *hastatulus*. We found that neither mitochondria- or chloroplast-interacting nuclear genes
177 were under- or over-represented on the X chromosome (Fisher's exact test, $P = 0.4947$
178 and $P = 0.3074$, respectively; Figure 1). This pattern indicates that neither the co-
179 adaptation nor the sexual conflict hypothesis alone is sufficient to explain the
180 chromosomal distribution of cyto-nuclear genes in *R. hastatulus.*

181 There are several factors that are expected to be important in determining cyto-
182 nuclear gene distributions, and these may explain the lack of bias in *R. hastatulus*. For
183 example, under both the co-adaptation and sexual conflict hypotheses, the age of the sex
184 chromosomes will determine the extent to which selection (either for co-adaptation, or

185 sexual antagonism) has had time to operate, which depends on the rate of gene movement
186 onto and off of the sex chromosomes. Whereas previous studies of cyto-nuclear genes in
187 animals have focused almost exclusively on ancient sex chromosome systems (Drown et
188 al. 2012; Dean et al. 2014; Rogell et al. 2014), our study focused on a dioecious plant
189 species in which sex chromosomes evolved more recently (~15 MYA; Navajas-Perez et
190 al. 2005), and many genes likely stopped recombining much more recently (Hough et al.,
191 2014). The lack of bias in the chromosomal distribution of cyto-nuclear genes may
192 therefore reflect the recent time scale of sex chromosome evolution rather than the
193 absence of biased gene movement. The relatively young age of sex chromosomes may
194 also have played a role in the lack of bias reported in the sex and neo-sex chromosomes
195 in three-spined stickleback, which evolved ~10 MYA (Kondo et al. 2004) and ~2 MYA,
196 respectively (Natri et al. 2013). Comparative studies of sex chromosomes of different age
197 will be central for understanding the rate at which organellar gene movement occurs.
198 In addition to being evolutionarily older, X-linked hemizygous genes are expected
199 to show a greater effect of over-or under-representation than genes with both X- and Y-
200 alleles because recessive mutations (involved in either co-adaptation or sexual conflict)
201 will be exposed to selection instead of masked by an alternate allele in a heterozygous
202 genotype. We detected a slightly greater under-representation of X-hemizygous N-mt
203 genes compared to autosomes or X-genes with retained Y-alleles, but the effect was not
204 statistically significant ($P = 0.4947$). The opposite pattern was evident for N-cp genes,
205 which were slightly over-represented on hemizygous genes, but again this effect was not
206 significant ($P = 0.3074$). A larger sample of hemizygous genes would be required to more
207 confidently assess whether such genes are in fact more often involved in cyto-nuclear

208 interactions than other genes on the X chromosome, and to test whether the opposite
209 pattern for N-mt and N-cp hemizygous genes is a result of a different rate of nuclear gene
210 transfer between mitochondrial and chloroplast genomes.

211 Another factor that will affect the chromosomal distribution of cyto-nuclear genes
212 is the number of N-mt and N-cp genes that were located on the autosome from which the
213 sex chromosomes evolved. Since the origins of mitochondria and chloroplasts both vastly
214 predate that of sex chromosomes (1.5-2 BYA compared to < 200 MYA; Dyall et al.
215 2004; Timmis et al. 2004; Cortez et al. 2014) gene transfer from organellar genomes to
216 the nuclear genome began long before the evolution of sex chromosomes. A bias in the
217 chromosomal distribution of cyto-nuclear genes in either direction may therefore arise if
218 the ancestral autosome was particularly rich or poor in cyto-nuclear genes. Indeed, it is
219 striking that autosomes in the animal species previously examined exhibited extensive
220 variation in the relative number of N-mt genes (see Drown et al. 2012 Figure 1 and Dean
221 et al. 2014 Figure 1 and Figure 2). That the ancestral number of N-mt and N-cp genes is
222 likely to be important is highlighted by the fact that the majority of genes involved in
223 mitochondrial DNA and RNA metabolism in *A. thaliana* are found on chromosome III
224 (Elo et al. 2003). If such a biased autosomal distribution of organellar variation is
225 representative of the ancestral sex chromosomes, the X chromosome could carry
226 significantly more N-mt or N-cp genes because of this ancestral gene number rather than
227 a biased rate of gene movement. Genetic mapping and comparative genomic studies of
228 genes that have transferred from organellar genomes after the origin of sex chromosomes
229 may provide a means to control for ancestral differences in gene number and provide a
230 better test of biases in organellar-nuclear gene movement.

To conclude, we conducted the first investigation of the extent to which co-adaptation and sexual conflict have shaped the chromosomal distribution cyto-nuclear genes in a plant species with sex chromosomes. We found no sign of under- or over-representation of either N-mt or N-cp genes on the X chromosome, implying that neither co-adaptation nor sexual conflict alone can explain the chromosomal distributions of these genes. Instead, we suggest that additional factors, including the age of sex chromosomes and the time since X-Y recombination became suppressed, are likely to have been important determinants of the patterns we observed. To determine whether the under-representation of mito-nuclear genes on the X chromosome previously reported in therian mammals and *C. elegans* (Drown et al. 2012; Dean et al. 2014) is indeed due to sexual conflict, rather than neutral processes such as the number of mito-nuclear genes on ancestral sex chromosomes, future studies should focus on quantifying rates of gene movement after sex chromosome origination. Cyto-nuclear conflict and co-evolution has undoubtedly played a major role in many aspects of genome evolution in both plant and animal systems; however, it remains unclear whether these processes have also shaped the chromosomal distribution cyto-nuclear genes.

Acknowledgements

We thank Rebecca Dean and Devin M Drown for comments on the manuscript. This research was supported by Discovery Grants to SCHB and SIW from the Natural Sciences and Engineering Council of Canada. JH was supported by an Ontario Graduate Fellowship and JAÅ by a Junior Fellowship from Massey College.

254 **References**

255 Aanen DK, Spelbrink JN, Beekman M. 2014. What cost mitochondria? The maintenance
256 of functional mitochondrial DNA within and across generations. Philos Trans R
257 Soc Lond B Biol Sci. 369.

258 Adams KL, Palmer JD. 2003. Evolution of mitochondrial gene content: gene loss and
259 transfer to the nucleus. Mol Phylogenet Evol. 29:380–395.

260 Bar-Yaacov D, Blumberg A, Mishmar D. 2012. Mitochondrial-nuclear co-evolution and
261 its effects on OXPHOS activity and regulation. Biochim Biophys Acta 1819:1107–
262 1111.

263 Berg OG, Kurland CG. 2000. Why mitochondrial genes are most often found in nuclei.
264 Mol Biol Evol. 17:951–961.

265 Burt A, Trivers R. 2006. Genes in conflict: the biology of selfish genetic elements.
266 Cambridge, MA: Belknap Press of Harvard University Press.

267 Charlesworth D. 2013. Plant sex chromosome evolution. J Exp Bot. 64:405–420.

268 Conesa A, et al. 2005. Blast2GO: a universal tool for annotation, visualization and
269 analysis in functional genomics research. Bioinformatics 21:3674–3676.

270 Cortez D, et al. 2014. Origins and functional evolution of Y chromosomes across
271 mammals. Nature 508:488–493.

272 Cosmides LM, Tooby J. 1981. Cytoplasmic inheritance and intragenomic conflict. J
273 Theor Biol. 89:83–129.

274 Dean R, Zimmer F, Mank JE. 2014. The potential role of sexual conflict and sexual
275 selection in shaping the genomic distribution of mito-nuclear genes. Genome Biol
276 Evol. 6:1096–1104.

277 Drown DM, Preuss KM, Wade MJ. 2012. Evidence of a paucity of genes that interact
278 with the mitochondrion on the X in mammals. Genome Biol Evol. 4:763–768.

279 Dyall SD, Brown MT, Johnson PJ. 2004. Ancient invasions: from endosymbionts to
280 organelles. Science 304:253–257.

281 Elo A, Lyznik A, Gonzalez DO, Kachman SD, Mackenzie SA. 2003. Nuclear genes that
282 encode mitochondrial proteins for DNA and RNA metabolism are clustered in the
283 *Arabidopsis* genome. Plant Cell 15:1619–1631.

284 Gillham NW. 1994. Organelle genes and genomes. New York: Oxford University Press.

285 Hill GE, Johnson JD. 2013. The mitonuclear compatibility hypothesis of sexual selection.
286 Proc R Soc B. 280:20131314.

287 Hough J, Hollister JD, Wang W, Barrett SCH, Wright SI. 2014. Genetic degeneration of
288 old and young Y chromosomes in the flowering plant *Rumex hastatulus*. Proc Natl
289 Acad Sci U S A. 111:7713–7718

290 Kondo M, Nanda I, Hornung U, Schmid M, Schartl M. 2004. Evolutionary origin of the
291 medaka Y chromosome. Curr Biol. 14:1664–1669.

292 Lane N. 2005. Power, sex, suicide: mitochondria and the meaning of life. Oxford, UK:
293 Oxford University Press.

294 Martin W. 2003. Gene transfer from organelles to the nucleus: frequent and in big chunks.
295 Proc Natl Acad Sci U S A. 100:8612–8614.

296 Natri HM, Shikano T, Merilä J. 2013. Progressive recombination suppression and
297 differentiation in recently evolved neo-sex chromosomes. Mol Biol Evol. 30:1131–
298 1144.

299 Navajas-Perez R, et al. 2005. The evolution of reproductive systems and sex-determining
300 mechanisms within *Rumex* (Polygonaceae) inferred from nuclear and
301 chloroplastidial sequence data. Mol Biol Evol. 22:1929–1939.

302 Quevillon E, et al. 2005. InterProScan: protein domains identifier. Nucleic Acids Res. 33:
303 W116–120.

304 R Development Core Team. 2013. R: A language and environment for statistical
305 computing. Vienna, Austria: R Foundation for Statistical Computing.

306 Rand DM, Clark AG, Kann LM. 2001. Sexually antagonistic cytonuclear fitness
307 interactions in *Drosophila melanogaster*. Genetics 159:173–187.

308 Rand DM, Haney RA, Fry AJ. 2004. Cytonuclear coevolution: the genomics of
309 cooperation. Trends Ecol Evol. 19:645–653.

310 Ridley M. 2000. Mendel's demon: gene justice and the complexity of life. London:
311 Weidenfeld & Nicolson.

312 Rogell B, Dean R, Lemos B, Dowling DK. 2014. Mito-nuclear interactions as drivers of
313 gene movement on and off the X chromosome. BMC Genomics 15:330.

314 Sloan DB. 2014. Using plants to elucidate the mechanisms of cytonuclear co-evolution.
315 New Phytol. doi: 10.1111/nph.12835

316 The Gene Ontology Consortiun. 2008. The Gene Ontology project in 2008. Nucleic
317 Acids Res. 36:D440–444.

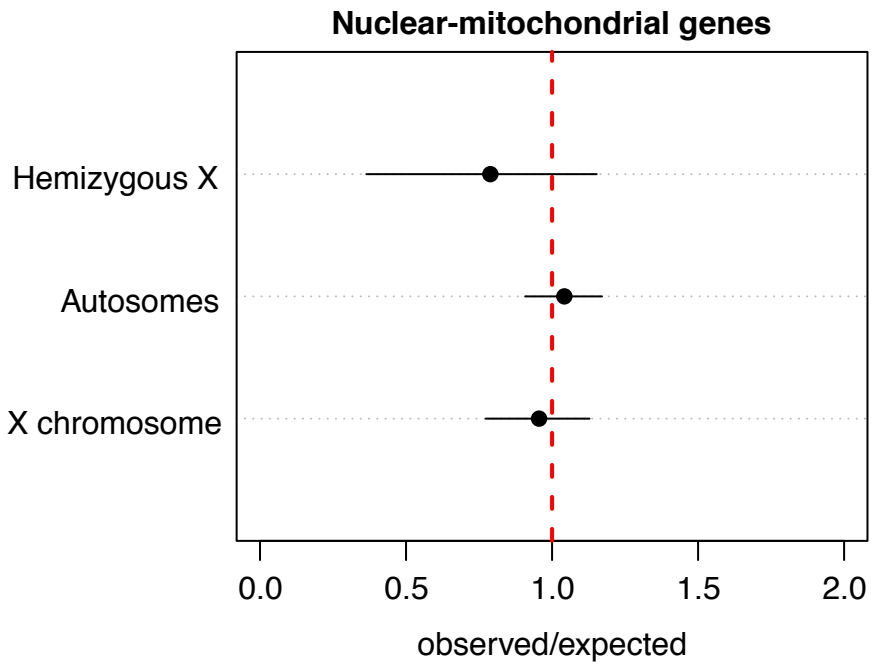
318 Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer:
319 organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 5:123–135.

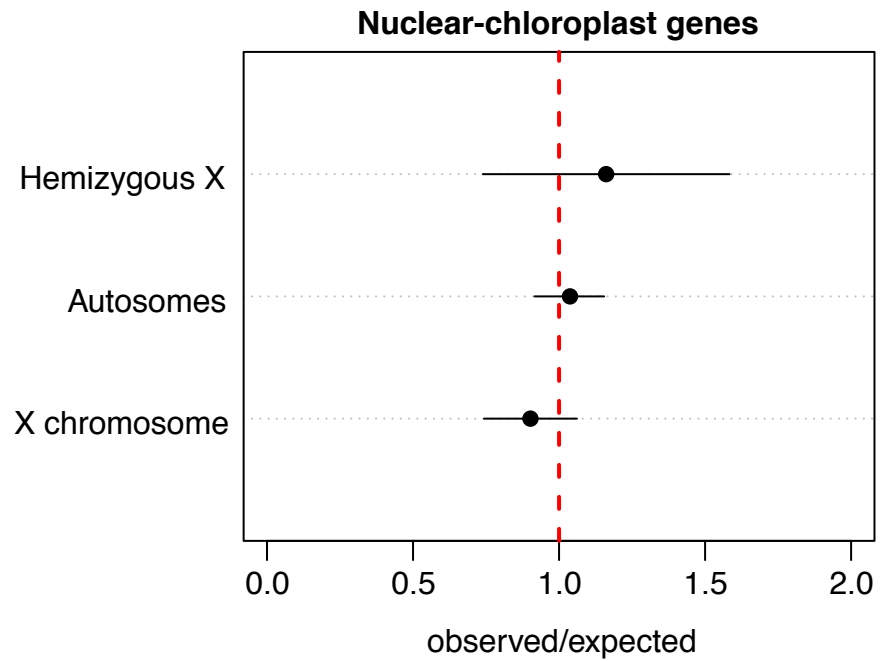
320 Touzet P, Meyer EH. 2014. Cytoplasmic male sterility and mitochondrial metabolism in
321 plants. Mitochondrion doi: 10.1016/j.mito.2014.04.009

322 Werren JH, Beukeboom LW. 1998. Sex determination, sex ratios, and genetic conflict.
323 Annu Rev Ecol Syst. 29:233–261.

324

325 **Figure 1.** Representation of the chromosomal location of cyto-nuclear genes in *Rumex*
326 *hastatulus*. Dots represent the observed to expected ratio of mito-nuclear (N-mt) and
327 chloro-nuclear (N-cp) genes on autosomes, the X chromosome, and hemizygous X genes,
328 with the 95% confidence intervals estimated by bootstrapping (10,000 replicates). The
329 vertical dotted line at 1 represents no over- or under-representation.


Nuclear-mitochondrial genes
Hemizygous X
Autosomes
X chromosome
0.0
0.5
1.0
1.5
2.0
observed/expected


Nuclear-chloroplast genes
Hemizygous X
Autosomes
X chromosome
0.0
0.5
1.0
1.5
2.0
observed/expected