Masculinization of the X-chromosome in aphid soma and gonads

Jaquiéry J[*,1], Simon J-C[1], Robin S[1,2], Richard G[1], Peccoud J[3], Boulain H[4], Legeai F[1,2], Tanguy S[1], Prunier-Leterme N[1], Le Trionnaire G[1]

[1] INRAE, UMR 1349, Institute of Genetics, Environment and Plant Protection, Le Rheu, France

[2] University of Rennes, Inria, CNRS, IRISA F-35000 Rennes, France

[3] Laboratoire Ecologie et Biologie des Interactions, Equipe Ecologie Evolution Symbiose, Unité Mixte de Recherche 7267 Centre National de la Recherche Scientifique, Université de Poitiers, 86073 Poitiers CEDEX 9, France

[4] Department of Ecology and Evolution, University of Lausanne, 1015, Lausanne, Switzerland

***Correspondence:** Julie Jaquiéry, Julie.Jaquiery@inrae.fr

Keywords: Sex-biased gene expression, sexual conflict, sexual antagonism, dimorphism, sex chromosome, duplication

Summary

Males and females share essentially the same genome but differ in their optimal values for many phenotypic traits, which can result in intra-locus conflict between the sexes. Aphids display XX/X0 sex chromosomes and combine unusual X chromosome inheritance with cyclical parthenogenesis. Theoretical and empirical works support the hypothesis that the large excess of male-biased genes observed on the aphid X chromosome compared to autosomes has evolved in response to sexual conflicts, by restricting the product of a sexually antagonistic allele to the sex it benefits. However, whether such masculinization of the X affects all tissues (as expected if it evolved in response to sexual conflicts) or reflects tissue specificities (which would contradict the sexual conflict hypothesis) remains an open question. To address it, we measured gene expression in three different somatic and gonadic tissues of males, sexual females and parthenogenetic females of the pea aphid. We observed a masculinization of the X in each of the studied tissues, with male-biased genes being 2.5 to 3.5 more frequent on the X than expected. We also tested the hypothesis that gene duplication can facilitate the attenuation of conflicts by allowing gene copies to neo- or sub-functionalize and reach sex-specific optima. As predicted, X-linked copies of duplicated genes having their other copies on autosomes were more frequently male-biased (40.5% of the genes) than duplicated autosomal genes (6.6%) or X-linked single-copy genes (32.5%). These results highlight a peculiar pattern of expression of X-linked genes in aphids at the tissue level and provide further support for sex-biased expression as a mechanism to attenuate intra-locus sexual conflicts.

Introduction

Sexual dimorphism, the difference between males and females at any phenotypic trait such as behavior, morphology, physiology or life history, is widespread. These differences are pervasive among eukaryotes, from plants to nematodes, insects, birds and mammals, to name a few (Cox and Calsbeek 2009; Williams and Carroll 2009). Regardless of its amount, sexual dimorphism engages males and females in a constant tug-of-war because their reproductive interests (such as optimal mating rate, number of partners, parental investment…) never align, owing to constitutive investment differences in gametes and/or progeny (Bonduriansky and Chenoweth 2009).

Differences in optimal trait values between sexes may generate intra-locus sexual conflicts. Typically, a new allelic variant could be beneficial to a female but deleterious to a male or vice versa. Such a sexually antagonistic (SA) allele is predicted to increase in frequency as long as the cost/benefit balance is positive. This increase leads to a so-called gender load in the population, due to the transmission of SA alleles to both sons and daughters (Chippindale et al. 2001; Rice and Chippindale 2002; Bonduriansky and Chenoweth 2009).

Several mechanisms may alleviate gender load (Bonduriansky and Chenoweth 2009). One is the evolution of sex-biased or sex-specific gene expression through a modifier of expression (Rice 1984). Once a SA allele is frequent enough, the reduction of its expression in the sex where it is deleterious may allow this variant to further increase in frequency and to possibly reach fixation (Rice 1984; Ellegren and Parsch 2007; Bonduriansky and Chenoweth 2009). This implies that the reduction of expression of the SA allele is beneficial to individuals of this sex. For genes that must be expressed at a certain level, a gene duplication event could allow bringing a new gene copy to sub- or neo-functionalize toward the sex-specific optimum (Bonduriansky and Chenoweth 2009; Connallon and Clark 2011; Gallach and Betrán 2011). Interestingly, these two processes (the duplication and the change in expression) could occur simultaneously, when the duplicated copy inserts in a region of the genome that already shows specific expression pattern (e.g., Arthur et al. 2014).

The invasion of the population by a SA allele and the attenuation of gender load through duplication and/or evolution of sex-biased gene expression may take place at different timescales. Indeed, the increase in frequency of a SA allele can be as rapid as a few generations, depending on its effect on fitness (e.g., Dean et al. 2012 for an experimental demonstration). The attenuation of the conflict by expression change or gene duplication may take much longer as it relies on rare random events, themselves depending on effective population size and mutation rate (Rice 1984; Stewart et al. 2010; Connallon and Clark 2011; Collet et al. 2016).

Importantly, the conditions for invasion by a SA allele differ between autosomes and sex chromosomes (Rice 1984; Fry 2010). In XX/XY systems, any SA allele that benefits males can invade the Y without conflict assuming complete linkage between the SDR (sex-determining region) and the SA locus. The picture for the X is more complex (Vicoso and Charlesworth 2006). X-linked recessive alleles are exposed to selection in males, while the female-biased transmission of the X (X chromosomes are transmitted twice more often by females than by males) gives more importance to selection episodes occurring in females. As a result, the X should accumulate recessive male-beneficial alleles and dominant female-beneficial ones. Similar processes are expected to occur in ZZ/ZW systems (e.g., birds, lepidopterans…).

Aphids constitute an interesting model to study the evolution of SA alleles as they show an XX/X0 sex-determining system combined with cyclical parthenogenesis: the alternation between several parthenogenetic generations in spring and summer and a single sexual generation in autumn. As a result, three distinct reproductive morphs occur in aphids: males, sexual females and parthenogenetic (asexual) females. Sexual females are genetically identical to their parthenogenetic mother, while male production involves the random elimination of one of the X (Wilson et al. 1997). Furthermore, during spermatogenesis only sperm cells carrying an X chromosome develop (Blackman 1987), so that the fusion of a sperm cell (AX) and an ovum (AX) always produces a diploid individual at the X and autosomes, which develops into a parthenogenetic female.

Theoretical models (Jaquiéry et al. 2013) predict that the peculiar inheritance of the X in aphids, the alternation between sexual and asexual reproduction, and the presence of

100 three different morphs (sexual females, parthenogenetic females and males) have a major
101 influence on the genomic location of SA allelic variants. In particular, conditions for the
102 invasion of variants that are beneficial to males and deleterious to parthenogenetic females
103 are predicted to be less restrictive for the X than for autosomes. By contrast, the conditions
104 for the invasion of variants that are detrimental to males and beneficial to parthenogenetic
105 females are more restrictive for the X. These models thus predict the X to be optimized for
106 male functions. Genomic analyses on the pea aphid *Acyrthosiphon pisum* showed that the X
107 chromosome had a large excess of genes preferentially expressed in males (i.e., male-biased
108 genes) compared to autosomes, and a deficit of parthenogenetic female-biased genes,
109 resulting in a "masculinization" of this chromosome (Jaquiéry et al. 2013). This pattern
110 matched predictions made under the hypothesis that evolution of sex-biased gene
111 expression reduces sexual conflicts by decreasing the expression of a sexually antagonistic
112 allele to the sex it benefits (Rice 1984). Interestingly, masculinization of the X has also been
113 observed in another aphid species (*Myzus persicae*) that diverged from the pea aphid
114 lineage 40 MYA (million years ago) (Mathers et al. 2019), but not in psyllids (Li et al. 2020) –
115 an obligatory group of sexual species closely related to aphids. These studies provide further
116 support that the masculinization of the X evolved in response to intra-locus sexual conflicts
117 resulting from the peculiar life cycle (cyclical parthenogenesis) and X inheritance in aphids.
118 However, as previous studies on aphids analyzed whole-body transcriptomes (Jaquiéry et al.
119 2013; Jaquiéry et al. 2018; Mathers et al. 2019; Li et al. 2020), it remains unclear whether
120 the observed masculinization of the X systematically occurs within each type of tissue or is
121 driven by some specific tissue with unusual expression patterns. For example, a meta-
122 analysis of gene expression in *Drosophila* revealed that the X is enriched in male-biased
123 genes expressed in the brain, while for all other tissues there is either no significant excess
124 or a paucity of male-biased genes on this chromosome (Huylmans et al. 2015). Tissue-
125 specific patterns may thus originate from the interplay between sex-specific regulation of
126 gene expression (including sexual antagonism, a force presumably acting in all tissues), and
127 dosage compensation, which could be stronger in the brain (Nozawa et al. 2013, Vensko and
128 Stone 2015).

129 Here, we predicted that if intra-locus sexual conflict is a strong driver of the
130 masculinization of the aphid X chromosome, masculinization would occur in all tissues. To

131 verify this prediction, we measured gene expression in different tissues from males, sexual
132 females and parthenogenetic females, including gonadic and somatic tissues. We found a
133 masculinization of the X in each type of tissue. Moreover, we confirmed that the X-linked
134 copy of a duplicated gene having another copy on autosomes is more likely to show a male-
135 biased expression than its autosomal copy or an X-linked single-copy gene. This result
136 suggests that duplications facilitate sub- or neo-functionalization toward the sex-specific
137 optimum.

138

139

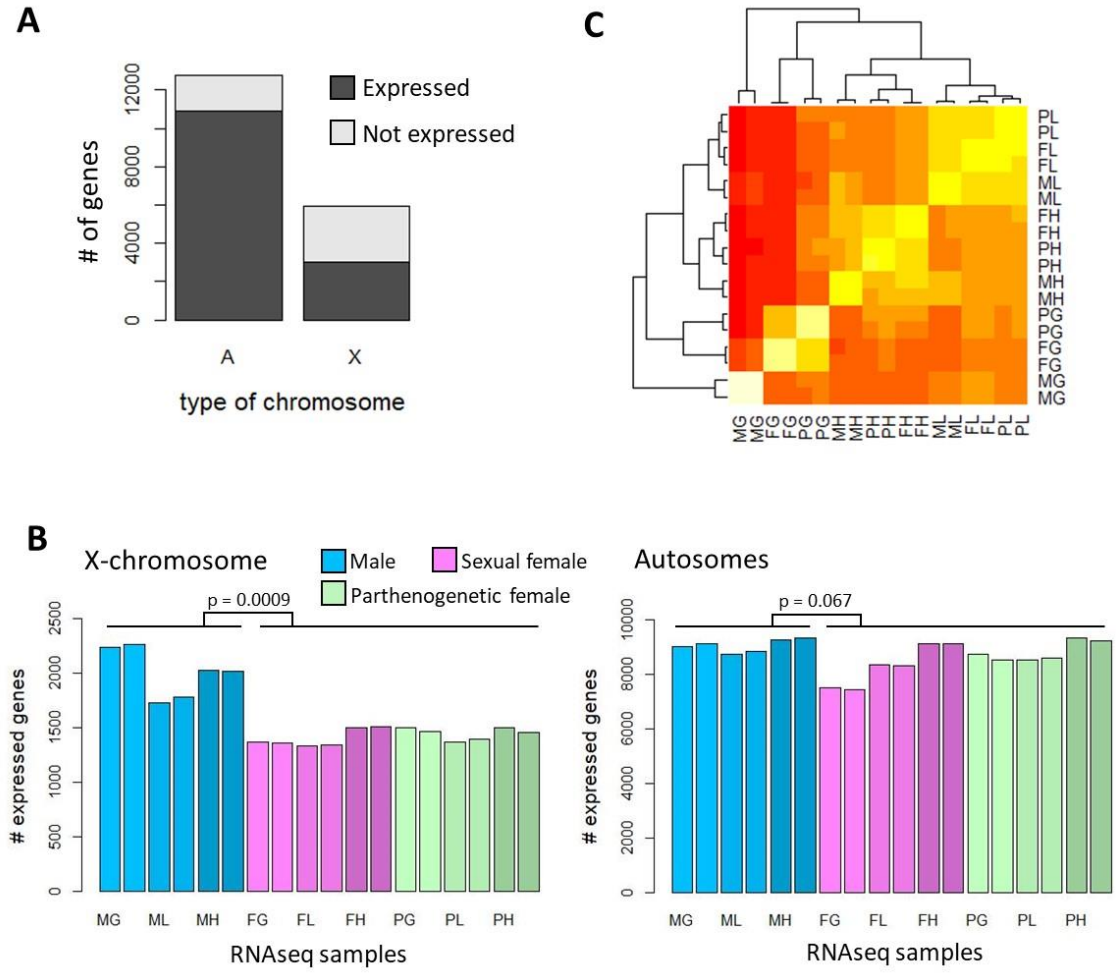
140 # Results

141 ## *Gene expression levels*

142 Gene expression levels in three different tissues (heads, legs and gonads) of the three
143 morphs (males, sexual females and parthenogenetic females) were measured from RNA-seq
144 counts on individuals produced by the same pea aphid clonal lineage (Supplementary table
145 S1). Overall, 14,605 genes out of the 20,639 predicted genes were expressed (> 1 count per
146 million reads [CPM] in at least two samples) in the 18 samples (3 morphs × 3 tissues × 2
147 replicates). We assigned 18,719 (90.7%) of the 20,639 predicted genes as autosomal or X-
148 linked, based on scaffold assignments from Jaquiéry et al. (2018). The genes that were not
149 assigned (9.3%) were located on scaffolds or part of scaffolds that were not clearly assigned
150 to X or autosomes in Jaquiéry et al. (2018) and they were thus not considered in the
151 subsequent analyses. Only 51% (3044/5961) of the X-linked genes were found to be
152 expressed, against 85% (10,890/12,758) of the autosomal genes (figure 1A). The genes
153 identified as not expressed in the 18 samples also generally showed no or low expression
154 support in whole-body RNAseq of males and females (Jaquiéry et al. 2013), especially for X-
155 linked genes (see supplementary text S1 for details). On average, more genes were
156 expressed in the samples from the different male tissues than in female samples, especially
157 for X-linked genes (X-chromosome: median number of expressed genes in males = 2020,
158 median in females = 1425, two-sided Mann-Whitney test, p = 0.0009; autosomes: median in

159 males = 9061, median in females = 8560, two-sided Mann-Whitney test, *p* = 0.067, figure
160 1B).

161 In the heatmap based on gene expression levels (figure 1C, supplementary figure S1),
162 samples grouped systematically by replicate of the same condition, and then by tissue for
163 leg and head samples. Within each of these tissues, the four female samples were always
164 more similar to each other than to the male samples. Gonad samples were the most
165 heterogeneous ones, samples from testes (MG) being highly different from all other
166 samples, and samples from parthenogenetic and sexual female gonads grouping together.



167

168 **Figure 1.** Gene expression in the 18 different RNA-seq samples. A) Number of genes considered as
169 expressed (CPM>1 in at least two samples) and not expressed on the X and on the autosomes. B)
170 Number of expressed genes per sample (expressed at more than 1 CPM) for X-linked and autosomal
171 genes (significance of differences between the X and autosomes was estimated with two-sided
172 Mann-Whitney tests). C) Heatmap of log(CPM+1). Samples group by tissue for head (H) and leg (L)

173 samples, with the parthenogenetic (P) and sexual female (F) samples being always more similar
174 compared to the male samples (M). Expression patterns of male gonad samples (MG) are the most
175 divergent. MG: male gonad, MH: male head, ML: male leg, FG: sexual female gonad, FH: sexual
176 female head, FL: sexual female leg, PG: parthenogenetic female gonad, PH: parthenogenetic female
177 head, PL: parthenogenetic female leg.

178 *The X chromosome is enriched in male-biased genes at the tissue-level*

179 To test whether some masculinization of the X was observed at the tissue-level, we first
180 categorized genes according to their relative expression patterns in the different conditions.
181 We defined a gene as "biased" toward, or preferentially expressed in, a set of samples
182 (which can be a particular tissue from a particular morph, all the tissues from a particular
183 morph or a particular tissue in all morphs) when at least 70% of all reads mapping to this
184 gene were observed in this set of samples (see methods). Note that increasing this
185 threshold to 80% or 90% or decreasing it to 60% or 50% did not qualitatively change the
186 results (Supplementary Figure S2). Testes showed the highest number of biased genes, with
187 1170 genes (referred to as MG+ genes) being preferentially expressed in this tissue. Then
188 came sexual female ovaries, with 375 FG+ genes, and male heads, with 203 MH+ genes
189 (figure 2). Tissue-biased genes (i.e., genes expressed mainly in a tissue of all morphs) were
190 common. Heads showed the highest number of tissue-biased genes (1169 H+ genes),
191 followed by legs (607 L+ genes) and gonads (511 G+ genes). Contrastingly, morph-biased
192 genes (i.e., genes expressed mainly in a morph in all tissues) were much less frequent for
193 females (only 58 F+ and 130 P+ genes) than for males (596 M+ genes) (figure 2). When
194 considering all genes preferentially expressed in a given morph, without considering tissues
195 (e.g., by summing MG+, MH+, ML+ and M+ genes for males), a total of 2001, 501 and 299
196 genes were biased toward males, sexual females and parthenogenetic females, respectively.

197 Interestingly, these different categories of genes differed in their chromosomal
198 locations (figure 2, Supplementary table S2). The proportions of X-linked genes among
199 genes expressed preferentially in testes (MG+), male heads (MH+), male legs (ML+) or
200 simply in males regardless of tissue (M+) varied from 54% to 73%, and significantly
201 exceeded (two-sided binomial tests, $p < 10^{-7}$ in all cases) the null expectation, which we took
202 as the proportion of X-linked genes among all expressed genes (22%) (if we consider all

203 predicted genes – supported by expression data or not – 31.8% locate on the X).
204 Contrastingly, genes that were preferentially expressed in parthenogenetic females were
205 less likely to locate on the X than expected (two-sided binomial tests, *p* ranging from 0.014
206 to 0.023 for PG+, PH+ and P+, not significant for PL+), with proportions of X-linked genes
207 ranging from 0% to 13% depending on tissues. The proportion of X-linked genes among
208 sexual female-biased genes were intermediate (13% to 32%), with only those preferentially
209 expressed in sexual female gonads being more frequent on the X (two-sided binomial test, p
210 $= 10^{-5}$). Genes that were preferentially expressed in gonads (G+) showed no deviation from
211 the null expectation, as 22% of them located on the X, while genes preferentially expressed
212 in heads (H+) and in legs (L+) were significantly less frequent on the X (15% to 16%) than
213 expected (two-sided binomial tests, $p = 10^{-8}$ and $p = 0.0006$, respectively). Baring the strong
214 difference between the X and autosomes, the distribution of biased genes within
215 chromosomes is rather homogeneous (Supplementary Figure S3).

216 We found that the breadth of gene expression, measured with τ (an index that
217 ranges from of 0 – indicating similar expression in all conditions – to 1 – indicating
218 expression in one condition only) was significantly narrower for X-linked genes than for
219 those on autosomes (median $\tau_x = 0.71$, $\tau_A = 0.33$, Mann-Whitney U-test, U = 23,201,000, $p <$
220 10^{-15}).

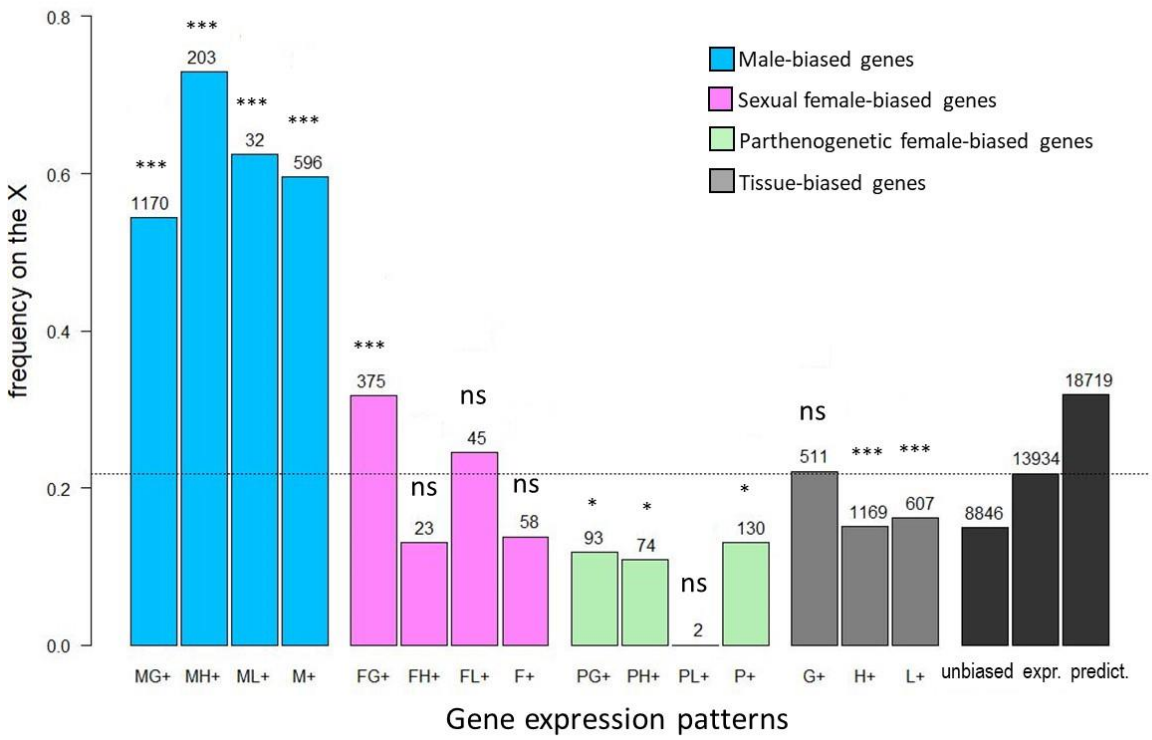


Figure 2. Proportions of X-linked genes among genes preferentially expressed in various morphs and/or tissues. Blue bars represent male-biased genes; MG+, MH+ and ML+: genes expressed preferentially in male gonads, heads and legs, respectively. M+: genes preferentially expressed in males when pooling all tissues, excluding the genes assigned to the previous categories. Pink bars represent sexual female-biased genes, with F standing for females and letters G, H and L having the same meaning as in males. Green bars represent parthenogenetic female-biased genes (P). Grey bars represent genes expressed preferentially in one of the tissues (gonads, heads or legs) and not limited to a particular morph. Black bars represent the frequency of X-linked genes among genes with unbiased expression ("unbiased"), genes expressed with CPM > 1 in at least two libraries ("expr.") or all predicted genes ("predict."). The horizontal dotted line represents the proportion of X-linked gene among expressed genes. The number of genes from each category is shown above bars, as well as the p-value (two-sided binomial tests against the expected frequency on the X chromosome estimated from expressed genes, which corresponds to the dotted horizontal black line). ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; ns: $p \geq 0.05$.

Expression patterns of single- and two-copy genes

To investigate the extent to which gene duplication facilitates the evolution of gene expression toward the sex-specific optimum, we compared the expression of autosomal and

239 X-linked genes that belong to single-copy and multicopy gene families. Multigenic gene
240 families were identified by Boulain et al. (2018) from orthoDB on 17 arthropod genomes.

241 We found that the X chromosome contained more genes that belong to multicopy
242 families than autosomes (38% of the genes on the X belong to multicopy families, against
243 28% for autosomal genes). When restricting our analyses to genes supported by expression
244 data, 1633 and 7428 single-copy genes locate on the X and on autosomes, respectively. We
245 also found 210 gene families composed of two expressed genes with one being on the X and
246 the other on autosomes. On autosomes, the percentages of genes with male-biased
247 expression (combining M+, MG+, ML+ and MH+ genes) were very similar between single-
248 and two-copy genes, at 6.5% and 6.6% respectively (figure 3, Chi-squared test, $\chi^2 \approx 0$, df = 1,
249 $p = 1$). On the X chromosome however, the proportion of male-biased genes was
250 significantly higher for two-copy genes (40.5%) than for single-copy genes (32.5%) (Chi-
251 squared test, $\chi^2 = 5$, df = 1, $p = 0.025$), these two proportions being much higher than their
252 equivalents on autosomes (Chi-squared tests, single-copy genes: $\chi^2 = 939.5$, df = 1, $p < 10^{-15}$,
253 two-copy genes: $\chi^2 = 64.8$, df = 1, $p < 10^{-15}$). Sexual female- and parthenogenetic female-
254 biased genes accounted only for a few percent of single- and two-copy genes. These female-
255 biased genes showed minor differences in proportion between chromosomes (significant for
256 single-copy genes only for parthenogenetic female [$\chi^2 = 4.9$, df = 1, $p = 0.027$] and for sexual
257 female [$\chi^2 = 8.9$, df = 1, $p = 0.003$], figure 3). They constituted similar proportions of the
258 single- and two-copy genes on a given chromosome type (Supplementary table S3).

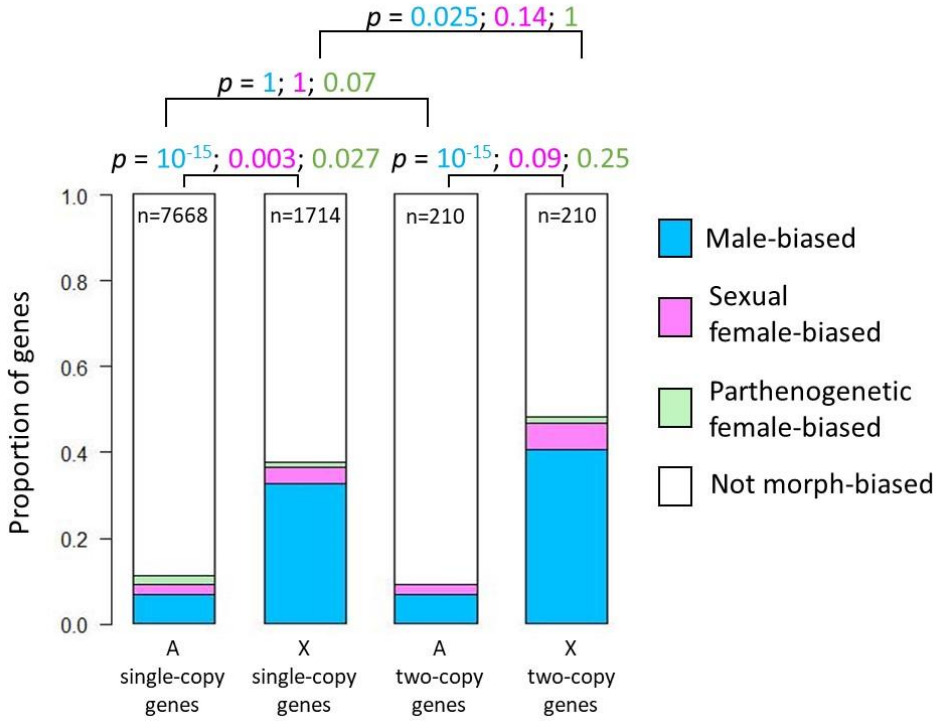
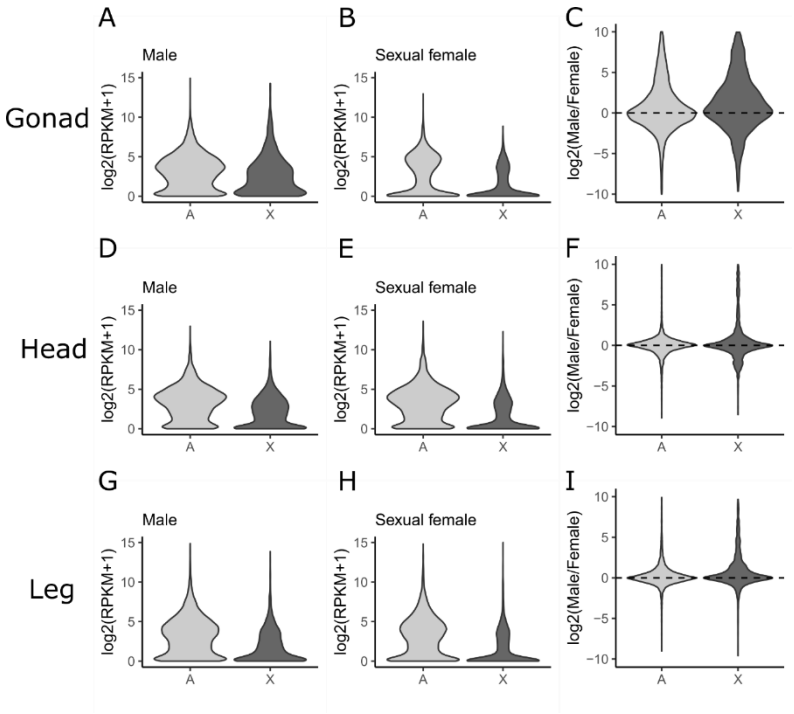


Figure 3. Proportion of genes showing preferential expression in different morphs, according to their number of copies (single or two copies) and chromosomal location (A: autosomes, X: X chromosome). Each two-copy gene has one copy on an autosome and the other on the X. P-values (Chi-squared tests) are shown, with font colors corresponding to the tested morph, according to the color of sectors. The number of genes composing each distribution is indicated on the plots.

Expression levels in morphs and tissues

The median expression levels of X-linked genes in somatic tissues and gonads from male and sexual female morphs were systematically lower than those of autosomal genes (figure 4, $p < 10^{-15}$ in all comparisons, two-sided Mann-Whitney tests), irrespective of the dose of X chromosomes per cell (two for sexual females and one for males). The same patterns were observed in parthenogenetic females ($p < 10^{-15}$ in all comparisons, supplementary figure S4). The mode of Log2 ratio of male-to-female RPKM (using sexual females in figure 4CFI and parthenogenetic females in Supplementary figure S4) lies close to 0 for both autosomal and X-linked genes, indicating dosage compensation for non-sex-biased genes in gonads and somatic tissues. Yet, we observed an excess of genes with high Log2 ratio of male to female expression, especially for the X chromosome in gonads and heads (figure 4CF). This indicates

277 an overexpression of some of the genes located on the single X chromosome of male cells,
278 which exceeds dosage compensation. This pattern was expected, given that male-biased
279 genes are significantly more frequent on the X than on autosomes (figure 2). When we
280 removed the genes characterized by a fold-change larger than 2 between males and sexual
281 females in at least one if the tissues (i.e. sex-biased genes), we still found strong evidence
282 for dosage compensation in the three tissues (Supplementary Figure S5). We also observed
283 that dosage compensation occurs throughout the X chromosome for all 3 tissues
284 (supplementary figure S6BDF), although the terminal portion of this chromosome appeared
285 to be particularly rich in sex-biased genes (supplementary figure S6ACE and S3).



286

287 **Figure 4.** Logarithm of gene expression (RPKM, reads per kilobase per million mapped reads) for X-
288 linked and autosomal genes for the different types of libraries (gonads, heads or legs of males and
289 sexual females). X-linked genes are significantly less expressed than autosomal genes in all cases
290 (two-sided Mann-Whitney tests, $p < 10^{-15}$). Logarithm of male-to-sexual female ratio of RPKM is also
291 shown for the three different tissues.

292

293 Discussion

294 Because of the peculiar inheritance of the X chromosome in aphids – males transmit
295 systematically their unique X to all sperm cells leading to the production of female-only

296 progeny –, the presence of three distinct morphs (sexual females, parthenogenetic females
297 and males) and the alternation between sexual and asexual reproduction, a specialization of
298 the X into male functions is expected (Jaquiéry et al. 2013). Indeed, models have shown that
299 the conditions for invasion by male-beneficial/parthenogenetic female-detrimental SA
300 alleles are less restrictive for the X than for autosomes, while the opposite is true for male-
301 detrimental/parthenogenetic female-beneficial alleles. SA alleles that are favorable to
302 sexual females should show little bias, the direction of which (i.e., the depletion or
303 enrichment of the X with genes carrying such variants) depending on the selective effect of
304 the allele on males and parthenogenetic females (table 1). A key finding of these predictions
305 is that they are not qualitatively affected by allele dominance level *h* (the aphid X
306 chromosome is a preferred location for male-beneficial alleles for all values of $h \neq 1$). This
307 contrasts with other X0 and XY species, where the X accumulates both recessive ($h<0.5$)
308 male-beneficial alleles and dominant ($h>0.5$) female-beneficial alleles (Vicoso and
309 Charlesworth 2006; Ellegren and Parsch 2007). Consequently, simpler predictions can be
310 made on aphids: the X should be enriched with male-beneficial alleles, parthenogenetic
311 female-beneficial alleles should be more common on autosomes (being counter selected on
312 the X), and sexual-female beneficial alleles should show no consistent bias (table 1).

313 **Table 1.** Preferred chromosomal location (X chromosome *versus* autosomes) of different types of
314 alleles with morph-antagonistic effects (s_m, s_f and s_p stand for the selective coefficient of the new
315 variant on males, sexual females and parthenogenetic females, respectively). Predictions originate
316 from analytical and simulation models developed in Jaquiéry et al. (2013).

Fitness effect of a SA mutation	**Preferred chromosomal location in aphids**
$s_m > 0$, $s_f < 0$, $s_p < 0$	Favored if on the X
$s_m > 0$, $s_f > 0$, $s_p < 0$	Favored if on the X
$s_m < 0$, $s_f < 0$, $s_p > 0$	Disfavored if on the X
$s_m < 0$, $s_f > 0$, $s_p > 0$	Disfavored if on the X
$s_m < 0$, $s_f > 0$, $s_p < 0$	Slightly disfavored if on the X
$s_m > 0$, $s_f < 0$, $s_p > 0$	Slightly favored if on the X

317

318 Although high-throughput approaches have been developed to pinpoint putative SA genes
319 (Innocenti and Morrow 2010; Lucotte et al. 2016; Ruzicka et al. 2019), identifying these
320 genes and estimating sex-specific selection and dominance coefficients are challenging tasks
321 that have been achieved in only a handful of studies (Barson et al. 2015; Husby et al. 2015).
322 However, since the evolution of a lower expression level in the sex that suffers from an
323 antagonistic allele could help resolving intra-locus sexual conflicts and allow the SA allele to
324 reach fixation (Rice 1984; Vicoso and Charlesworth 2006; Ellegren and Parsch 2007), the
325 chromosomal location of sex-biased genes can be used to indirectly test predictions.

326 Here, we analyzed gene expression in different tissues of distinct morphs in the pea
327 aphid and found that the chromosomal location of morph-biased genes followed the
328 predictions made under SA models from Jaquiéry et al. (2013). In the three tissues
329 considered, male-biased genes were largely overrepresented on the X, parthenogenetic
330 female-biased genes were underrepresented, and sexual female-biased genes showed no
331 consistent bias. These empirical data thus support the hypothesis that, in aphids, sexual
332 conflicts would be the key driver of the masculinization of the X and of the specialization of
333 autosomes for the parthenogenetic phase of the life cycle.

334 Genes showing a male-biased expression were much more frequent on the X
335 chromosome than expected under a random distribution, irrespective of the tissue
336 considered. Genes expressed mainly in testes, male heads and male legs were 2.5 to 3.5
337 times more frequent than expected on the X, so were genes that were male-biased in all the
338 tissues considered. As all the investigated tissues contribute to this pattern, these results
339 extend previous studies reporting a general enrichment in male-biased genes (measured
340 from whole-body samples) on the aphid X chromosome (Jaquiéry et al. 2013; Jaquiéry et al.
341 2018; Mathers et al. 2019). Overall, this and previous studies demonstrate strong and
342 consistent bias toward the X in aphids for genes expressed predominantly in males.
343 Interestingly, on the X chromosome, the percentage of male-biased genes was significantly
344 higher for two-copy genes (40.5%) than for single-copy genes (32.5%), while these
345 percentages were similar for autosomal genes (6.5% and 6.6%). These results match the
346 predictions that duplications may facilitate sub- or neo-functionalization toward sex-specific
347 optima, and that, for genes located on the X chromosome, the male optimal expression
348 levels are favored rather than the female optima.

349 Conversely, parthenogenetic female-biased genes were significantly less frequent on
350 the X, except for parthenogenetic female leg-biased (PL+) genes due to lack of power, as
351 only two such genes were found, both located on autosomes. This matches previous
352 observations on whole body transcriptomes (Jaquiéry et al. 2013; Mathers et al. 2019). This
353 trend also reveals a consistent reduction of the contribution of the X chromosome to
354 biological functions and processes occurring in parthenogenetic females, and corroborates
355 the observation that the X is depleted from functionally important genes for the
356 parthenogenetic phase (Li et al. 2020). Interestingly, the chromosomal location of sexual
357 female-biased genes did not significantly depart from random expectations, except for
358 genes specifically expressed in ovaries. These were significantly more frequent than
359 expected on the X, although their proportion (32%, Figure 2) was much lower than for male-
360 biased genes (54%-73%). According to the models (table 1), variants of X-linked genes that
361 are beneficial to sexual females should increase in frequency only if they are also beneficial
362 to males. This could be the case for genes controlling sexual reproduction-related functions
363 that were not sex-specific (e.g., meiosis). In this case however, we do not expect such genes
364 to be preferentially expressed in sexual females. These genes may be less/not expressed in

365 adult males because spermatogenesis is often already completed at this stage (Wieczorek et
366 al. 2019). Alternatively, it is possible that these genes had variants that were beneficial to
367 both sexes, which have since evolved functions that are more specific to sexual females.

368 Among hemipterans, the accumulation of male-biased genes on the X appears to be
369 specific of aphids. A deficit of male-biased genes on the X was observed in three hemipteran
370 species (two heteropteran bugs and a leafhopper) that only reproduce sexually and are very
371 distantly related to aphids, despite an apparent homologous origin of the X chromosome
372 (Pal and Vicoso 2015; Mathers et al. 2021). In psyllids, which diverged more recently (~200-
373 250 MYA) from aphids and are characterized by obligate sexual reproduction and X0 sex
374 determination, no enrichment of male-biased genes on the X chromosome was found (Li et
375 al. 2020). Hence, the accumulation of male-biased genes on the X chromosome in aphids
376 would have started after the divergence between aphids and psyllids. This scenario supports
377 a role for cyclical parthenogenesis (which appeared ~200 MYA in aphid ancestors, Davis
378 2012) in the masculinization of the X in aphids.

379 While the chromosomal location of the different types of sex-biased genes matches
380 the predicted evolution of gene expression as a mitigation mechanism of sexual conflicts,
381 other processes could contribute to the observed patterns. Genes located on different
382 chromosomes types could differ in other characteristics that could result in a non-random
383 genomic distribution of male-biased genes. For example, the X chromosome of *Drosophila* is
384 enriched in young genes (those that are present in only a restricted taxonomic group),
385 which are more likely to show sex-biased expression (Palmieri et al. 2014). In our analysis
386 however, the X-linked copy of a two-copy gene had a much greater probability of being
387 male-biased (0.405) than its autosomal counterpart (0.066, figure 3), demonstrating that the
388 masculinization of the X does not solely reflect differences in gene characteristics.
389 Moreover, the three *A. pisum* autosomes are systematically depleted in male-biased genes
390 and enriched in parthenogenetic female-biased genes (supplementary table S4,
391 supplementary figure S3). This reinforces the specific hemizygosity of the X in males as a
392 determinant factor for the observed patterns. Consequently, we do not see an alternative
393 hypothesis to explain the accumulation of male-biased genes on the X to the one based on
394 the resolution of sexual conflicts by an evolution of sex-specific or sex-biased gene
395 expression.

396 The number of morph-biased genes was highest for gonadic tissues, with 1170, 375
397 and 130 genes being specific of gonads of males, sexual females or parthenogenetic
398 females, respectively. This observation likely reflects the highly morph-specific functions of
399 gonads: the production of sperm in males, of yolky eggs in sexual females and of embryos in
400 viviparous parthenogenetic females (Michalik et al. 2013). In other species, testis also stands
401 out as the tissue showing the most specific gene expression patterns (Meiklejohn and
402 Presgraves 2012; Uhlén et al. 2015). Many genes also showed preferential expression in
403 male heads (202), against 23 and 93 for sexual and parthenogenetic female heads,
404 respectively. This could reflect sensorial and/or behavioral differences between males and
405 females: males have to actively search for females and initiate mating, while females spend
406 most of their time in feeding. Legs showed a low number of morph-biased genes, most of
407 which were overexpressed in sexual females (45 genes, against 32 for males and 2 for
408 parthenogenetic females). This result may reflect the existence of specific organs on sexual
409 female tibias (scent plaques), which are responsible for the secretion and release of sex
410 pheromones (Murano et al. 2018).

411 Interestingly, the strong enrichment of the X with male-biased genes is similar across
412 tissues (ranging from 54% to 73%). This result strikingly contrasts with *Drosophila*, where
413 male-biased genes from different tissues show opposing patterns: male-biased genes in
414 brains are strongly enriched on the X, while there is either no departure from random
415 expectation or a paucity of male-biased genes on the X for the other tissues (Huylmans and
416 Parsch 2015). These differences were interpreted as resulting from the interplay between
417 dosage compensation (the brain could be more sensible to gene dose, and thus would
418 require a tighter dosage control) and sex-specific regulation of gene expression (Huylmans
419 and Parsch 2015). In aphids, the consistent enrichment of the X with male-biased genes in
420 different tissues suggests that similar evolutionary forces apply in reproductive and somatic
421 tissues. Under the hypothesis that sex-biased expression evolved in response to sexual
422 conflicts, our results suggest that their attenuation occurs in all tissues. Our data also
423 indicate that sexual conflicts occur in a wide range of tissues and not only in gonads. This
424 pattern was observed by Innocenti and Morrow (2010), who found that transcripts showing
425 signature of sexual antagonism in *Drosophila* were frequent in soma.

426 The high expression of a substantial number of X-linked genes in males despite the
427 haploid state of this chromosome in this sex is intriguing. The X chromosome shows several
428 specific characteristics, among which a slightly larger amount of intra-chromosome
429 duplicates (449 for 132 Mb) compared to the largest autosomes (413 for 170 Mb, Li et al.
430 2019), and an enrichment with multi-copy orthologs (Li et al. 2020). Our analyses confirmed
431 this trend, as 38% of the genes on the X belongs to multicopy families, against 28% for
432 autosomal genes. Whether recent (undetected) duplications are more frequent on the X
433 and could help to achieve strong expression for some of the X-linked genes in males remains
434 an open question. However, epigenetic mechanisms probably play a more important role.
435 Indeed, the X-chromosome is hypermethylated in male aphids compared to autosomes, and
436 differential gene methylation between males and females positively correlates with
437 differential expression, especially for the X chromosome (Mathers et al. 2019). An increased
438 accessibility of the chromatin of the X in males was also documented (Richard et al. 2017).
439 Another interesting feature of the aphid X chromosome is its larger fraction of
440 unexpressed genes compared to autosomes, amounting to half of the X-linked genes based
441 on our tissue samples data (against 15% of the autosomal genes). This characteristic was
442 also underlined from whole body transcriptomes (Jaquiéry et al. 2013; 2018; Richard et al.
443 2017; Li et al. 2020; Mathers et al. 2019). Two factors may contribute to the large number of
444 X-linked genes classified as unexpressed. X-linked genes might show a narrower expression
445 breadth, being restricted to some (unsampled) morphs, tissues or stages. The exceptionally
446 high τ value for expressed X-linked genes indeed provides some support to this hypothesis.
447 Nevertheless, a large fraction of the X-linked genes classified as “unexpressed” here also
448 shows low expression support in whole body libraries (see supplementary text S1): 40% are
449 supported by 0 reads and 34% by 1-5 reads. It is thus probable that both effects (absence of
450 expression or narrower expression breadth) explain the large fraction of unexpressed X-
451 linked genes. RNA sequencing on a larger diversity of tissues (especially in males) and stages
452 may be required to resolve this point. Interestingly, Li et al. (2020) found that genes
453 considered as “functionally important” were less likely to locate on the X chromosome. The
454 evolutionary forces that drive these patterns remain to be identified, but they could be
455 linked to sexual antagonism or to the particular epigenetic state of the X.

456 Previous studies suggested dosage compensation in the pea aphid and the green
457 peach aphid from whole-body transcriptomes (Jaquiéry et al. 2013; Pal and Vicoso 2015;
458 Richard et al. 2017; Li et al. 2020; Mathers et al. 2019). Here, we show dosage compensation
459 in all investigated tissues, including testis. While single-cell transcriptomics will be essential
460 to demonstrate or refute that X-linked genes are dosage compensated in various cell types
461 during spermatogenesis, dosage compensation in testis would be another peculiarity of
462 aphids (but see Witt et al. 2021; Mahadevaraju et al. 2021). Indeed, sex chromosomes of
463 other dosage-compensated species seem generally not compensated in the gonads of the
464 heterogametic sex in diptera and lepidoptera, at the scale of entire organs (Vicoso and
465 Bachtrog 2015; Gu et al. 2017; Gu and Walters 2017). However, single-cell RNAseq analysis
466 of *Drosophila* testis has evidenced dosage compensation in pre-meiotic and somatic testis
467 cells (Witt et al. 2021, Mahadevaraju et al. 2021). In other groups (mammals, birds,
468 nematods, fungi), sex-linked genes are silenced by meiotic sex chromosome inactivation
469 MSCI (Shiu et al. 2001; Bean et al. 2004; Turner 2007; Schoenmakers et al. 2009; but see
470 Guioli et al. 2012; Daish et al. 2015), and recent studies suggest that MSCI could also occur
471 in *Drosophila* (Mahadevaraju et al. 2021; Witt et al. 2021). Different hypotheses have been
472 proposed to explain the silencing of sex-linked genes during gametogenesis. MSCI could be a
473 consequence of the mechanisms that protect against unwanted recombination between X
474 and Y and allow DNA repair in its absence (Lu et al. 2015). MSCI may also has evolved as a
475 mean to protect unsynapsed chromosomes from the invasion of transposons (Huynh et al.
476 2005) or segregation distorters that would bias sex-ratio (Meiklejohn and Tao 2010). If
477 protection from segregation distorters is an important driver of sex-linked gene silencing,
478 this could explain why the expression of X-linked genes in aphid testis has not been
479 repressed during evolution and thus can be subject to dosage compensation. In aphids,
480 segregation distortion is already maximal (all sperm cells that do not carry an X degenerate,
481 and those that are functional carry the single identical X chromosome), so no X-linked allele
482 can further increase its transmission during spermatogenesis. Hence, there is no possibility
483 for X-linked distorters to evolve, such that mechanisms to protect from distorters (i.e., sex-
484 linked gene silencing) may have been lost.
485 In conclusion, we document an atypical genome-wide pattern of gene expression in
486 aphids, with a high degree of masculinization of the X chromosome in both somatic and

487 gonadic tissues. Our study reinforces the hypothesis that this masculinization evolved in
488 response to sexual conflicts raised by the accumulation of male-beneficial alleles on the X
489 (Jaquiéry et al. 2013). To further support this hypothesis, masculinization of the X should be
490 assessed in distantly related aphid lineages (which diverged up to 200 MYA, Davis 2012) and
491 other species that show an "aphid-like" life cycle and X-inheritance (e.g. *Strongyloides*
492 nematods, Nemetschke et al. 2010; Streit 2017). Finally, understanding the functional
493 epigenetic or post-transcriptional mechanisms responsible for sex-biased gene expression in
494 aphids would help to understand how such a strong chromosomal specialization in gene
495 expression has been achieved.

496

497 # Methods

498 ## *Sex- and tissue-biased gene expression analysis*

499 Gene expression levels in three different tissues (head, legs, gonads) of three reproductive
500 morphs (male, sexual female, parthenogenetic female) were measured from RNA-Seq
501 collected on a single pea aphid genotype (clone LSR1, from alfalfa, IAGC 2010). Aphids were
502 reared on broad bean *Vicia faba* at less than five individuals per plant to prevent the
503 production of winged morphs. Parthenogenesis was maintained under a 16-hour light
504 regime and a temperature of 18°C. The production of sexual individuals was initiated by
505 transferring larvae (at stage 3) to a 12-hour light regime at the same temperature of 18°C.
506 Two generations later, sexual females and males were observed. A total of 100 adult
507 parthenogenetic females (produced under 16-hour light regime), 100 adult sexual females
508 and 100 adult males were immediately frozen into liquid nitrogen. Heads and legs were
509 scalpel-cut. Twenty additional individuals per morph were also dissected in a saline solution
510 with fine forceps to collect gonads. Gonads included testes and accessory glands in males
511 and ovarioles in sexual females. Embryos of stage >10 (according to Miura et al. 2003) were
512 removed from parthenogenetic female ovarioles (which already contain asexually
513 developing embryos) to avoid the contribution of developing and late embryos to RNA
514 production. All collected tissues were stored in RNA later (Qiagen) immediately after
515 collection and pooled in batches before RNA extraction (with two replicates by sex and

516 tissue). Hence, a total of 18 RNA extractions (3 morphs × 3 tissues × 2 biological replicates)
517 were performed using the SV Total RNA Isolation System (Promega) according to
518 manufacturer's instructions. RNA quality was checked on Bioanalyzer (Agilent) and
519 quantified on Nanodrop (Thermo Scientific). The 18 RNA samples were subsequently sent to
520 the GetPlage platform (Toulouse, France) for library preparation (TruSeq Stranded mRNA
521 Library Preparation kit) and 150 bp RNA paired-end sequencing (Illumina HiSeq3000).

522 After filtering for rRNA, reads from each library were mapped to the V2 assembly of
523 the pea aphid genome (Acyr 2.0, Genbank accession GCA_000142985.2) using STAR version
524 2.5.2a (Dobin et al. 2013) with default parameters, except: outFilterMultimapNmax = 5,
525 outFilterMismatchNmax = 3, alignIntronMin = 10, alignIntronMax = 50000 and
526 alignMatesGapMax = 50000. Then, we counted reads mapped on exons of each predicted
527 gene (NCBI Annotation release ID: 102) using FeatureCounts version 1.5.0-p3 (Liao et al.
528 2014) with default parameters except: -g gene -C -p -M --fraction. The numbers of mapped
529 reads per library ranged from 19.8 to 28.6 million (mean 24.3 million, supplementary table
530 S1).

531 We used the R package DESeq (Anders and Huber 2010) to normalize the libraries
532 (upperquartile method with p = 0.75) and calculate CPM. Only genes with CPM>1 in at least
533 two libraries (out of 18) were considered as expressed and retained in the analyses, unless
534 mentioned otherwise. To identify genes predominantly expressed in a specific tissue or a
535 morph, we imposed that a minimal percentage (70%) of the reads mapping to a given gene
536 was sequenced from that specific tissue/morph. Doing so avoided the reliance on p-values
537 from differential expression analyses (which in turn depend on the absolute expression level
538 of a gene). This 70% threshold allowed identifying genes showing considerable bias in
539 expression toward a tissue/morph, and to retrieve a large number of genes for more
540 powerful analyses. So, when >70% of the reads mapping to a gene were from a given tissue
541 from a given morph, the gene was classified as predominantly expressed in that tissue (e.g.,
542 PL+ for those predominantly expressed in legs of parthenogenetic females; etc). When >70%
543 of the reads were from a given morph (but not restricted to a single tissue), the gene was
544 classified as morph-specific (M+, F+ or P+, for males, sexual females, or parthenogenetic
545 females, respectively). Similarly, genes mainly expressed (>70% of reads) in a tissue (but not
546 restricted to a single morph) were classified as tissue-specific (G+, H+ or L+, for gonads,

547 heads or legs, respectively). Expressed genes were thus classified into 16 mutually exclusive
548 classes: MH+, ML+, MG+, FH+, FL+, FG+, PH+, PL+, PG+, M+, F+, P+, H+, L+, G+ or unbiased.
549 Additional analyses with threshold values of 50%, 60%, 80% and 90% to identify a gene as
550 predominantly expressed in a set of samples were also conducted to verify the robustness
551 of our results to this parameter.

552 *Comparisons between chromosome types*

553 The chromosomal location (X vs autosomes) of each gene was that of Jaquiéry et al. (2018).
554 This represented the only assignments available at the start of the analyses. Since then,
555 chromosome-level genome assemblies have been produced for the pea aphid (Li et al. 2019;
556 Mathers et al. 2021). We found that our assignments to chromosomes were consistent for
557 97% to 98% of the genes located on the four mega scaffolds corresponding to the four pea
558 aphid chromosomes, depending on the assembly. We therefore considered our initial
559 assignment reliable. Furthermore, the number of NCBI predicted genes that we assigned to
560 chromosome (18719) was larger than for the Li et al. (2019) assembly (17315). For the
561 genome assembly from Mathers et al. (2021), DNA was obtained from a lineage sampled on
562 *Lathyrus*, a host plant genus that harbors a cryptic species of the pea aphid complex
563 (Peccoud et al. 2009; 2014) that is quite divergent from the LSR1 clone we used (J. Peccoud,
564 personal communication). Significant deviation from a random chromosomal distribution
565 for each class of genes with a specific expression pattern was tested with two-sided
566 binomial tests, the expected proportion was computed as the proportion of expressed
567 genes on the X. To compare expression patterns between the three autosomes
568 (supplementary table S4, supplementary figure S3), we used the assembly from Li et al.
569 (2019). Expression breadth for X-linked and autosomal genes was estimated with τ (Yanai
570 2005) on log+1 transformed data.

571 *Investigation of gene families*

572 To identify single-copy and multicopy genes in *A. pisum* genome, we used orthologs identified
573 among 17 arthropod genomes by Boulain et al. (2018). Briefly, the longest protein isoform
574 from each arthropod species was used to run OrthoDB_soft_1.6 (Kriventseva, et al. 2015) and
575 the levels of orthology were assigned by referring to the species phylogeny established in

576 Boulain et al. (2018). The groups of orthologs generated by OrthoDB were then used to
577 identify *A. pisum* unique (single-copy) or duplicated (multicopy) genes. To control for possible
578 differences in gene content between the X and autosomes, which could account for different
579 expression patterns between chromosomes, we searched for gene families composed of two
580 genes, one being on the X and the other on autosomes. To statistically compare frequencies
581 of the various classes of expression between the X and autosomal copies with sufficient power
582 given the limited sample size (n = 210 pairs of genes), the gene classes with male-biased
583 expression (i.e., M+, MG+, ML+ and MH+ genes) were grouped into a single category of male-
584 biased genes. We proceeded similarly for sexual female-biased genes and parthenogenetic
585 female-biased genes. We also created a new class encompassing all genes that were not
586 morph-biased (i.e, G+, H+, L+ and unbiased genes). For each aggregated gene class, we
587 compared its frequency among chromosomes for single and duplicated copies and then
588 among single and duplicated copies within chromosomes (chi-squared tests).

589 *Dosage compensation*

590 To investigate dosage compensation, RPKM (reads per kilobase per million mapped reads)
591 were calculated with EdgeR (Robinson et al. 2010), and only genes with RPKM > 1 in at least
592 two of the 18 libraries were kept. After log-transformation of RPKM, differences in
593 expression levels between chromosomes were examined with two-sided Mann-Whitney
594 tests for each sex and each tissue. Then, the logarithm of the male to female (sexual or
595 parthenogenetic) ratio of RPKM were estimated for X and autosomes for each tissue. As the
596 uneven frequency of biased genes between chromosomes could interfere with dosage
597 compensation patterns, the same analyses were performed by eliminating genes that
598 showed a fold change in expression greater than 2 between males and females in at least
599 one tissue.

600

601 **Acknowledgments**

602 This work was supported by a grant from the University of Rennes 1 ("Défi émergent") to JJ,
603 as well as grants from the French Research Agency (ANR) SexAphid (ANR-09-GENM-017-
604 001) and Mecadapt (ANR-11-BSV7-005-01). The authors of this preprint declare that they

605 have no financial conflict of interest with the content of this article. The raw data are
606 publicly available on NCBI (projects PRJNA547535 and PRJNA748848. The bioinformatic and
607 R scripts, as well as the R input files can be found on Zenodo:
608 https://doi.org/10.5281/zenodo.6242803. We thank the editor Tanja Schwander, Ann
609 Kathrin Huylmans and another anonymous referee who provided constructive comments to
610 a previous draft of this manuscript.

611

612

613 # References

614 Anders S, Huber W. 2010. Differential expression analysis for sequence count data. **Genome Biology**
615 11:R106.

616 Arthur RK, Ma L, Slattery M, Spokony RF, Ostapenko A, Nègre N, White KP. 2014. Evolution of
617 H3K27me3-marked chromatin is linked to gene expression evolution and to patterns of gene
618 duplication and diversification. **Genome research** 24:1115-1124.

619 Barson NJ, Aykanat T, Hindar K, Baranski M, Bolstad GH, Fiske P, Jacq C, Jensen AJ, Johnston SE,
620 Karlsson S, et al. 2015. Sex-dependent dominance at a single locus maintains variation in age at
621 maturity in salmon. **Nature** 528:405-408.

622 Bean CJ, Schaner CE, Kelly WG. 2004. Meiotic pairing and imprinted X chromatin assembly in
623 *Caenorhabditis elegans.* **Nature Genetics** 36:100-105.

624 Blackman RL. 1987. Reproduction, cytogenetics and development. In: Minks A, P H, editors. Aphids:
625 their biology, natural enemies and control. Amsterdam: Elsevier. p. 163-195.

626 Bonduriansky R, Chenoweth SF. 2009. Intralocus sexual conflict. **Trends in Ecology and Evolution**
627 24:280-288.

628 Boulain H, Legeai F, Guy E, Morliere S, Douglas NE, Oh J, Murugan M, Smith M, Jaquiéry J, Peccoud J,
629 et al. 2018. Fast evolution and lineage-specific gene family expansions of aphid salivary effectors
630 driven by interactions with host-plants. **Genome Biology and Evolution** 10:1554-1572.

631 Chippindale AK, Gibson JR, Rice WR. 2001. Negative genetic correlation for adult fitness between
632 sexes reveals ontogenetic conflict in *Drosophila*. **Proceedings of the National Academy of Sciences**
633 98:1671-1675.

634 Collet JM, Fuentes S, Hesketh J, Hill MS, Innocenti P, Morrow EH, Fowler K, Reuter M. 2016. Rapid
635 evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*. **Evolution**
636 70:781-795.

637 Connallon T, Clark AG. 2011. The resolution of sexual antagonism by gene duplication. **Genetics**
638 187:919-937.

639 Cox R, Calsbeek R. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of
640 intralocus sexual conflict. **The American Naturalist** 173:176-187.

641 Daish TJ, Casey AE, Grutzner F. 2015. Lack of sex chromosome specific meiotic silencing in platypus
642 reveals origin of MSCI in therian mammals. **BMC Biology** 13:106.

643 Davis GK. 2012. Cyclical parthenogenesis and viviparity in aphids as evolutionary novelties. **Journal**
644 **of Experimental Zoology Part B: Molecular and Developmental Evolution** 318:448-459.

645 Dean R, Perry JC, Pizzari T, Mank JE, Wigby S. 2012. Experimental evolution of a novel sexually
646 antagonistic allele. **Plos Genetics** 8:e1002917.

647 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013.
648 STAR: ultrafast universal RNA-seq aligner. **Bioinformatics** 29:15-21.

649 Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. **Nature**
650 **Reviews Genetics** 8:689-698.

651 Fry JD. 2010. The genomic location of sexually antagonistic variation: some cautionary comments.
652 **Evolution** 64:1510-1516.

653 Gallach M, Betrán E. 2011. Intralocus sexual conflict resolved through gene duplication. **Trends in**
654 **Ecology and Evolution** 26:222-228.

655 Gu L, Walters JR. 2017. Evolution of sex chromosome dosage compensation in animals: a beautiful
656 theory, undermined by facts and bedeviled by details. **Genome Biology and Evolution** 9:2461-2476.

657 Gu L, Walters JR, Knipple DC. 2017. Conserved patterns of sex chromosome dosage compensation in
658 the lepidoptera (WZ/ZZ): insights from a moth neo-Z chromosome. **Genome Biology and Evolution**
659 9:802-16.

660 Guioli S, Lovell-Badge R, Turner JMA. 2012. Error-prone ZW pairing and no evidence for meiotic sex
661 chromosome inactivation in the chicken germ line. **Plos Genetics** 8:e1002560.

662 Husby A, Kawakami T, Rönnegård L, Smeds L, Ellegren H, Qvarnström A. 2015. Genome-wide
663 association mapping in a wild avian population identifies a link between genetic and phenotypic
664 variation in a life-history trait. **Proceedings of the Royal Society B: Biological Sciences**
665 282:20150156.

666 Huylmans AK, Parsch J. 2015. Variation in the X:autosome distribution of male-biased genes among
667 *Drosophila melanogaster* tissues and its relationship with dosage compensation. **Genome Biology**
668 **and Evolution** 7:1960-1971.

669 Huynh KD, Lee JT. 2005. X-chromosome inactivation: a hypothesis linking ontogeny and phylogeny.
670 **Nature Reviews Genetics** 6:410-418.

671 Innocenti P, Morrow EH. 2010. The sexually antagonistic genes of *Drosophila melanogaster*. **Plos**
672 **Biology** 8:e1000335.

673 International Aphid Genomics Consortium. 2010. Genome sequence of the pea aphid *Acyrthosiphon*
674 *pisum*. **PLoS Biology**. 8:e1000313.

675 Jaquiéry J, Peccoud J, Ouisse T, Legeai F, Prunier-Leterme N, Gouin A, Nouhaud P, Brisson JA, Bickel
676 R, Purandare S, et al. 2018. Disentangling the causes for faster-X evolution in aphids. **Genome**
677 **Biology and Evolution** 10:507-520.

678 Jaquiéry J, Rispe C, Roze D, Legeai F, Le Trionnaire G, Stoeckel S, Mieuzet L, Da Silva C, Poulain J,
679 Prunier-Leterme N, et al. 2013. Masculinization of the X chromosome in the pea aphid. **Plos Genetics**
680 9.

681 Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simão FA, Pozdnyakov IA, Ioannidis P,
682 Zdobnov EM. 2015. OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying
683 free software. **Nucleic Acids Research** 43:D250-256.

684 Li Y, Zhang B, Moran NA. 2020. The aphid X chromosome is a dangerous place for functionally
685 important genes: diverse evolution of Hemipteran genomes based on chromosome-level assemblies.
686 **Molecular Biology and Evolution** 37:2357-2368.

687 Li YY, Park H, Smith TE, Moran NA. 2019. Gene family evolution in the pea aphid based on
688 chromosome-level genome assembly. **Molecular Biology and Evolution** 36:2143-2156.

689 Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning
690 sequence reads to genomic features. **Bioinformatics** 30:923-930.

691 Lu LY, Yu X. 2015. Double-strand break repair on sex chromosomes: challenges during male meiotic
692 prophase. **Cell cycle** 14:516-525.

693 Lucotte EA, Laurent R, Heyer E, Ségurel L, Toupance B. 2016. Detection of allelic frequency
694 differences between the sexes in humans: a signature of sexually antagonistic selection. **Genome**
695 **Biology and Evolution** 8:1489-1500.

696 Mahadevaraju S, Fear JM, Akeju M, Galletta BJ, Pinheiro MMLS, Avelino CC, Cabral-de-Mello DC,
697 Conlon K, Dell'Orso S, Demere Z, et al. 2021. Dynamic sex chromosome expression in *Drosophila*
698 male germ cells. **Nature Communications** 12:892.

699 Mathers TC, Mugford ST, Percival-Alwyn L, Chen Y, Kaithakottil G, Swarbreck D, Hogenhout SA, van
700 Oosterhout C. 2019. Sex-specific changes in the aphid DNA methylation landscape. **Molecular**
701 **Ecology** 28:4228-4241.

702 Mathers TC, Wouters RHM, Mugford ST, Swarbreck D, van Oosterhout C, Hogenhout SA. 2021.
703 Chromosome-scale genome assemblies of aphids reveal extensively rearranged autosomes and long-
704 term conservation of the X chromosome. **Molecular Biology and Evolution**. 38:856-875.

705 Meiklejohn CD, Presgraves DC. 2012. Little evidence for demasculinization of the *Drosophila* X
706 chromosome among genes expressed in the male germline. **Genome Biology and Evolution** 4:1007-
707 1016.

708 Meiklejohn CD, Tao Y. 2010. Genetic conflict and sex chromosome evolution. **Trends in Ecology and**
709 **Evolution** 25:215-223.

710 Michalik A, Szklarzewicz T, Wegierek P, Wieczorek K. 2013. The ovaries of aphids (Hemiptera,
711 Sternorrhyncha, Aphidoidea): morphology and phylogenetic implications. **Invertebrate Biology**
712 132:226-240.

713 Miura T, Braendle C, Shingleton A, Sisk G, Kambhampati S, Stern DL. 2003. A comparison of
714 parthenogenetic and sexual embryogenesis of the pea aphid *Acyrthosiphon pisum* (Hemiptera:
715 Aphidoidea). **Journal of Experimental Zoology part B: Molecular and Developmental Evolution**
716 295:59-81.

717 Murano K, Ogawa K, Kaji T, Miura T. 2018. Pheromone gland development and monoterpenoid
718 synthesis specific to oviparous females in the pea aphid. **Zoological Letters** 4:9.

719 Nemetschke L, Eberhardt AG, Hertzberg H, Streit A. 2010. Genetics, chromatin diminution, and sex
720 chromosome evolution in the parasitic nematode genus *Strongyloides*. **Current Biology** 20:1687-
721 1696.

722 Nozawa M, Fukuda N, Ikeo K, Gojobori T. 2014. Tissue- and stage-dependent dosage compensation
723 on the neo-X chromosome in *Drosophila pseudoobscura*. **Molecular Biology and Evolution** 31:614-
724 24.

725 Pal A, Vicoso B. 2015. The X chromosome of Hemipteran insects: conservation, dosage
726 compensation and sex-biased expression. **Genome Biology and Evolution** 7:3259-3268.

727 Palmieri N, Kosiol C, Schlötterer C. 2014. The life cycle of *Drosophila* orphan genes. **eLife** 3:e01311.

728 Peccoud J, de la Huerta M, Bonhomme J, Laurence C, Outreman Y, Smadja CM, Simon J-C. 2014.
729 Widespread host-dependent hybrid unfitness in the pea aphid species complex. **Evolution** 68:2983-
730 2995.

731 Peccoud J, Ollivier A, Plantegenest M, Simon JC. 2009. A continuum of genetic divergence from
732 sympatric host races to species in the pea aphid complex. **Proceedings of the National Academy of**
733 **Sciences of the United States of America** 106:7495-7500.

734 Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. **Evolution** 38:735-742.

735 Rice WR, Chippindale AK. 2002. The evolution of hybrid infertility: perpetual coevolution between
736 gender-specific and sexually antagonistic genes. **Genetica** 116:179-188.

737 Richard G, Legeai F, Prunier-Leterme N, Bretaudeau A, Tagu D, Jaquiery J, Le Trionnaire G. 2017.
738 Dosage compensation and sex-specific epigenetic landscape of the X chromosome in the pea aphid.
739 **Epigenetics & Chromatin** 10:30.

740 Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential
741 expression analysis of digital gene expression data. **Bioinformatics** 26:139-140.

742 Ruzicka F, Hill MS, Pennell TM, Flis I, Ingleby FC, Mott R, Fowler K, Morrow EH, Reuter M. 2019.
743 Genome-wide sexually antagonistic variants reveal long-standing constraints on sexual dimorphism
744 in fruit flies. **Plos Biology** 17:e3000244.

745 Schoenmakers S, Wassenaar E, Hoogerbrugge JW, Laven JSE, Grootegoed JA, Baarends WM. 2009.
746 Female Meiotic Sex Chromosome Inactivation in Chicken. **Plos Genetics** 5:e1000466.

747 Shiu PK, Raju NB, Zickler D, Metzenberg RL. 2001. Meiotic silencing by unpaired DNA. **Cell** 107:905-
748 916.

749 Stewart AD, Pischedda A, Rice WR. 2010. Resolving intralocus sexual conflict: genetic mechanisms
750 and time frame. **Journal of Heredity** 101:S94-S99.

751 Streit A. 2017. Genetics: modes of reproduction and genetic analysis. **Parasitology** 144:316-326.

752 Turner JMA. 2007. Meiotic sex chromosome inactivation. **Development** 134:1823-1831.

753 Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C,
754 Sjöstedt E, Asplund A, et al. 2015. Tissue-based map of the human proteome. **Science** 347:1260419.

755 Vensko SP, Stone EA. 2015. X-to-autosome expression and msl-2 transcript abundance correlate
756 among *Drosophila melanogaster* somatic tissues. **PeerJ**. 3:e771-e

757 Vicoso B, Bachtrog D. 2015. Numerous transitions of sex chromosomes in Diptera. **Plos Biology**
758 13:e1002078.

759 Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes.
760 **Nature Reviews Genetics** 7:645-653.

761 Wieczorek K, Kanturski M, Sempruch C, Świątek P. 2019. The reproductive system of the male and
762 oviparous female of a model organism-the pea aphid, *Acyrthosiphon pisum* (Hemiptera, Aphididae).
763 **PeerJ** 7:e7573.

764 Williams TM, Carroll SB. 2009. Genetic and molecular insights into the development and evolution of
765 sexual dimorphism. **Nature Reviews Genetics** 10:797-804.

766 Wilson ACC, Sunnucks P, Hales DF. 1997. Random loss of X chromosome at male determination in an
767 aphid, *Sitobion* near *fragariae*, detected using an X-linked polymorphic microsatellite marker.
768 **Genetical Research** 69:233-236.

769 Witt E, Shao Z, Hu C, Krause HM, Zhao L. 2021. Single-cell RNA-sequencing reveals pre-meiotic X-
770 chromosome dosage compensation in *Drosophila* testis. **Plos Genetics** 17:e1009728.

771 Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safran
772 M, Domany E, et al. 2005. Genome-wide midrange transcription profiles reveal expression level
773 relationships in human tissue specification. **Bioinformatics** 21:650-659.