

Widespread cryptic variation in genetic architecture between the sexes

Wouter van der Bijl¹ and Judith E. Mank^{1,2}

¹ Department of Zoology and Biodiversity Research Centre, University of British Columbia, Canada

² Department of Genetics, Evolution and Environment, University College London, UK

Abstract

The majority of the genome is shared between the sexes, and it is expected that the genetic architecture of most traits is shared as well. This common architecture has been viewed as a major source of constraint on the evolution of sexual dimorphism (SD). SD is nonetheless common in nature, leading to assumptions that it results from differential regulation of shared genetic architecture. Here, we study the effect of thousands of gene knock-out mutations on 202 mouse phenotypes to explore how regulatory variation affects SD. We show that many traits are dimorphic to some extent, and that a surprising proportion of knock-outs have sex-specific phenotypic effects. Many traits, regardless whether they are monomorphic or dimorphic, harbor cryptic differences in genetic architecture between the sexes, resulting in sexually discordant phenotypic effects from sexually concordant regulatory changes. This provides an alternative route to dimorphism through sex-specific genetic architecture, rather than differential regulation of shared architecture.

Keywords: sexual dimorphism, genetic architecture, between-sex genetic correlation, rFM, knock-out

20 Introduction

21 In organisms with separate sexes, different
22 evolutionary interests of males and females can lead to
23 divergent trait optima, which can be realized through the
24 evolution of sexual dimorphism. For traits to change from
25 monomorphic to dimorphic, the underlying genetic
26 mechanisms need to be decoupled between males and
27 females. However, even in species with sex chromosomes,
28 males and females share the vast majority of their genome
29 (Bachtrog *et al.*, 2014), leading to the expectation that
30 traits are controlled by the same loci in both sexes (Lande,
31 1980). This shared genomic architecture is typically
32 considered a source of significant constraint on the
33 evolution of dimorphism (Stewart & Rice, 2018), as traits
34 would need to first become genetically decoupled
35 between females and males before divergence can occur
36 (Lande, 1980; Poissant *et al.*, 2010; Hermansen *et al.*,
37 2018). Shared trait architecture can lead to intra-locus
38 sexual conflict (Rice & Chippindale, 2001), where alleles at
39 a locus have different fitness effects in males and females,
40 and is this assumed to limit the degree to which the sexes
41 can achieve their respective fitness optima (Hansen, 2006).
42 Indeed, the constraints on the evolution of sexual
43 dimorphism (SD) are often considered both pervasive and
44 persistent, resulting in enduring sexually antagonistic
45 selection on many traits (Rice & Chippindale, 2001;
46 Chenoweth *et al.*, 2008; Poissant *et al.*, 2010; Ruzicka *et*
47 *al.*, 2019). This persistent constraint is however difficult to
48 reconcile with the fact that sexual dimorphism evolves
49 rapidly (Stewart & Rice, 2018), is seen in a broad array of
50 traits, and differs markedly among related species (Owens
51 & Hartley, 1998).

52 It has been suggested that sexual dimorphism arises
53 from regulatory differences between males and females
54 (Ellegren & Parsch, 2007; Mank, 2017), and there are good

55 examples of this (e.g. Galouzis & Prud'homme, 2021).
56 Indeed, recent genome-wide scans in fruit flies have
57 shown that protein coding sequence differences are
58 overrepresented among evolutionarily persistent variants
59 thought to be maintained by sexual antagonism (Ruzicka
60 *et al.*, 2019). This might suggest that conflict over coding
61 sequence variation is much harder to resolve compared to
62 conflict over gene expression. However, functional
63 studies have revealed that the genes underlying some
64 dimorphisms are not expressed differently between the
65 sexes (Khila *et al.*, 2012). This indicates that sex-biased
66 expression alone cannot explain all dimorphism, and other
67 mechanisms may exist.

68 Another perspective on the genetics of sexually
69 dimorphic traits stems from investigations grounded in
70 quantitative genetic theory (Lande, 1980). By comparing
71 the phenotypes of individuals of known relatedness,
72 usually through breeding designs or pedigrees, one can
73 estimate the between-sex genetic correlation (r_{fm}) for a
74 trait of interest. This correlation describes the extent to
75 which a particular genotype affects both male and female
76 phenotypes in the same way. If $r_{fm} \approx 1$, genotypes affect
77 males and females similarly (i.e. brothers and sisters look
78 alike), while if $r_{fm} \approx 0$, male and female phenotypes vary
79 independently (Lande, 1980). This estimate of r_{fm} is based
80 on autosomal additive standing genetic variation and
81 measures the additive effects of the many genetic variants
82 that exist in that population at that time. It can therefore
83 be used to predict the extent to which a population can
84 respond to sexually divergent selection. Since this r_{fm}
85 estimate is based on the additive genetic variance, we will
86 denote it here as r_{fm}^A for clarity.

87 Average estimates of r_{fm}^A are often close to one
88 (Poissant *et al.*, 2010), suggesting that there is little
89 standing sex-specific genetic variation. However, these

90 estimates are also interpreted by many to reflect the
 91 extent to which the autosomal genetic architecture
 92 underlying the trait is shared between the sexes
 93 (Chenoweth *et al.*, 2008; Poissant *et al.*, 2010; Griffin *et al.*,
 94 2013; e.g. Stewart & Rice, 2018). In other words, a strongly
 95 positive r_{fm}^A is interpreted to mean that the gene network
 96 that produces the phenotypic trait value is largely identical
 97 between the sexes, suggesting that genetic architecture
 98 needs to be decoupled before SD can evolve. Furthermore,
 99 if r_{fm}^A is an evolutionary important constraint, one would
 100 expect those traits with weak r_{fm}^A to be more likely to
 101 evolve SD, resulting in a negative relationship
 102 (Bonduriansky & Rowe, 2005; Fairbairn & Roff, 2006;
 103 Poissant *et al.*, 2010). Alternatively, selection in favor of SD
 104 may drive reductions in r_{fm}^A , leading to the same
 105 prediction. This negative association is supported by the
 106 prevailing evidence (Poissant *et al.*, 2010), however the
 107 correlation varies widely between studies, and r_{fm}^A is

108 generally a poor predictor of SD. Furthermore, r_{fm}^A has
 109 been shown to be quickly eroded under artificial selection
 110 (Delph *et al.*, 2011).

111 r_{fm}^A estimates provide a high-level statistical
 112 description of genotype to phenotype mapping across the
 113 sexes and are an aggregate across standing genetic
 114 variation in the population. However, we know very little
 115 about the loci that underlie this statistic. In particular, we
 116 do not know whether variation in protein coding sequence
 117 is more or less likely to cause sexually discordant
 118 phenotypic effects than expression variation. Here, we use
 119 high-throughput phenotype data from a genome-wide
 120 panel of gene knock-outs in mice to reveal unexpected
 121 differences in the gene expression architecture between
 122 the sexes (The International Mouse Phenotyping
 123 Consortium *et al.*, 2016; International Mouse Phenotyping
 124 Consortium *et al.*, 2017). We find that although most
 125 phenotypic traits are dimorphic, even many monomorphic

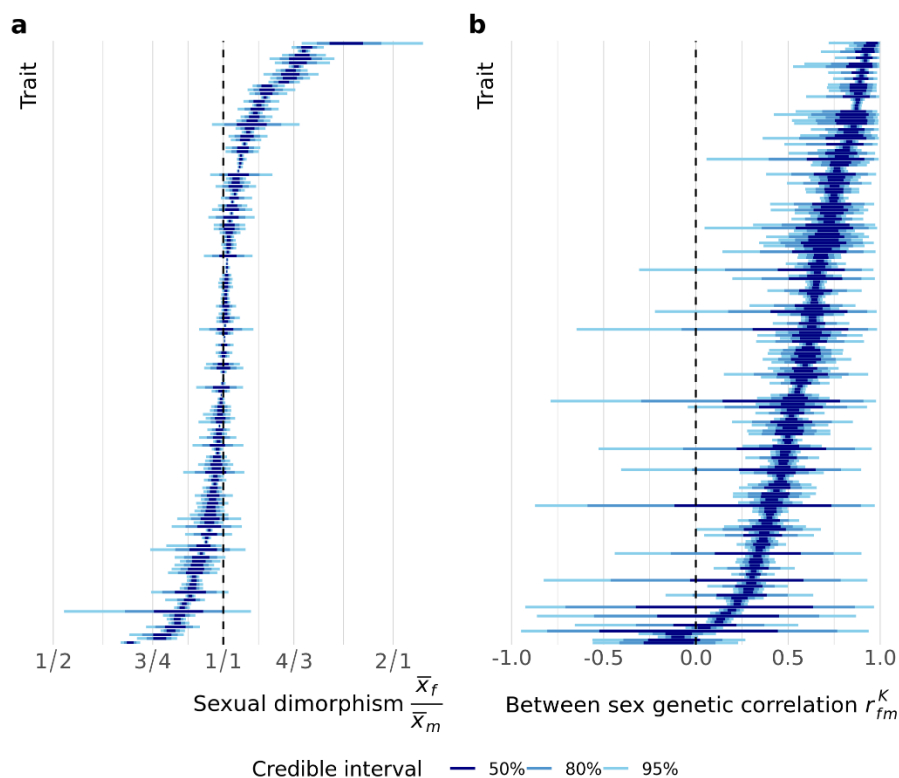


Figure 1: (a) Estimates and associated uncertainty for sexual dimorphism for each trait analyzed. Each horizontal line displays the credible intervals for one trait, where traits have been arranged by the posterior median. Shaded regions indicated the credible intervals of 50%, 80% and 95% of the posterior densities from a multilevel model. Sexual dimorphism is averaged across the wild-type genotypes, and defined as the ratio of female and male means. (b) As in (a), but depicting the between-sex genetic correlation r_{fm}^K . Note that the traits have been arranged independently in each panel.

126 traits harbor sex-dependent architectures, suggesting that
127 many traits may harbor cryptic sex-specific variation.
128 Changes in both sexes to these loci through expression
129 may provide a way for SD to rapidly evolve, as traits are
130 already partially decoupled and the phenotypic effect
131 differs between males and females. These findings imply
132 that the evolutionary constraint in SD may be more easily
133 overcome than previously thought and explain the broad
134 diversity of sexual dimorphism observed in nature, as well
135 as the apparent rapid evolution of many sexually
136 dimorphic traits.

137 Results

138 We evaluated the sex-specific effects of alterations to
139 gene expression, by leveraging data from large-scale high-
140 throughput phenotyping of gene knock-out lines from the
141 International Mouse Phenotyping Consortium (IMPC) (The
142 International Mouse Phenotyping Consortium *et al.*,
143 2016). We obtained data for all continuous traits from the
144 main IMPC pipeline for which at least 100 genotypes were
145 available. The IMPC uses highly standardized phenotyping
146 assays on C57BL/6 inbred mice. Both control mice and
147 phenotype knock-out lines are tested continuously, with
148 the eventual goal of knocking out each gene in the mouse
149 genome. This immense scientific effort provides an
150 unprecedented opportunity to quantify the between-sex
151 genetic correlation across many traits and many genotypes
152 in highly standardized conditions.

153 *Sexual dimorphism and r_{fm}^K of mouse traits*

154 If males and females share the genetic architecture of
155 traits, knock-outs should affect the phenotype of both
156 sexes similarly, and as architectures diverge the knock-out
157 effects should diverge as well. We estimated the genetic
158 correlation between males and females analogous to the
159 conventional approach outlined above (r_{fm}^A). However, to

160 delineate the knock-out lines from the traditional
161 approach, we denote these estimates as r_{fm}^K , where K
162 denotes the genetic variance-covariance matrix between
163 knock-out genotypes (Figure S1). Note that r_{fm}^K measures
164 the correlation between the phenotypic effects of genetic
165 knock-outs, while r_{fm}^A measures the correlation for
166 genome-wide additive genetic variance.

167 For each of 260 traits, we obtained all available
168 observations. On average, traits were measured in 8,069
169 control mice, as well as in 21,513 mice across 1,713
170 different knock-out genotypes. Per knock-out line, seven
171 females and seven males were typically phenotyped.

172 For each trait we obtained posterior distributions for
173 SD and the between-sex genetic correlation (r_{fm}^K) by fitting
174 a Bayesian multilevel model. SD was expressed as the ratio
175 of means (for Figure 1) and as the “sexual dimorphism
176 index”: $\frac{\bar{x}_{larger\ sex}}{\bar{x}_{smaller\ sex}} - 1$ (for downstream analyses). Since
177 mice are sexually dimorphic for body size and many traits
178 scale with body size, we included a standardized
179 population level effect of body weight in the model.
180 Models without body size adjustment produced
181 qualitatively similar results (see supplementary material).
182 Additionally, we added group level intercepts for known
183 sources of variance, this included the phenotyping center,
184 the date of testing, as well as variation in testing conditions
185 indicated by the IMPC. Using a Bayesian approach allowed
186 us to evaluate and propagate the uncertainty in the
187 estimate of r_{fm}^K in downstream analyses. This can be
188 important since this correlation can be biased towards 0 if
189 it is difficult to estimate (Griffin *et al.*, 2013). Out of 260
190 traits tested, 202 traits passed our model evaluation
191 procedure and were used for further inference.

192 Many of the measured traits showed substantial SD
193 (Figure 1a), confirming a previous report on the IMPC data
194 (International Mouse Phenotyping Consortium *et al.*,
195 2017), with an average SD index of 0.09 [0.08, 0.10]
196 (posterior median [95% Credible Interval]). As the large
197 sample size in this study makes it possible to distinguish
198 small effects that have little biological relevance, we
199 evaluated SD using equivalence testing (Wellek, 2010). We
200 compared the 95% credible intervals (CI) of the SD index
201 for each trait with a region of practical equivalence (ROPE)
202 between 0 and 0.05 (Kruschke, 2018) (i.e. between 0 and
203 5% difference in absolute magnitude). When the entire CI
204 falls outside the ROPE, we can be confident the sexes differ
205 by more than 5% and the trait is considered dimorphic. We
206 consider a trait monomorphic if we are confident there is
207 less than a 5% difference, so when the entire CI falls within
208 the ROPE. Under this decision rule (Kruschke, 2018),
209 dimorphic traits roughly equal monomorphic traits. 49 out
210 of the 156 traits (31.4%) were found to be clearly

211 dimorphic, while 47 traits (30.1%) to be monomorphic. and
212 60 traits (38.5%) were not classified, as their credible
213 interval overlapped the 5% threshold. Some of the most
214 monomorphic traits include calcium levels in the blood and
215 the time spent on the periphery of an open field. Strongly
216 dimorphic traits include a variety of immune function
217 related traits, such as spleen weight and counts of
218 different T-cell types, as well as glucose tolerance (Table
219 S1).

220 Traits showed a wide variety of estimates for r_{fm}^K , from
221 a correlation close to 1 between the phenotypes of the
222 sexes down to correlations indistinguishable from 0 (Figure
223 1b). The average correlation was clearly positive, but not
224 as strong as we expected (0.650 [0.622, 0.689]).
225 Surprisingly, very few traits showed a strong concordance
226 between male and female effects, with fewer than 5% of
227 traits having an estimate above 0.9. Some of the traits with
228 the highest correlation are body temperature and eye

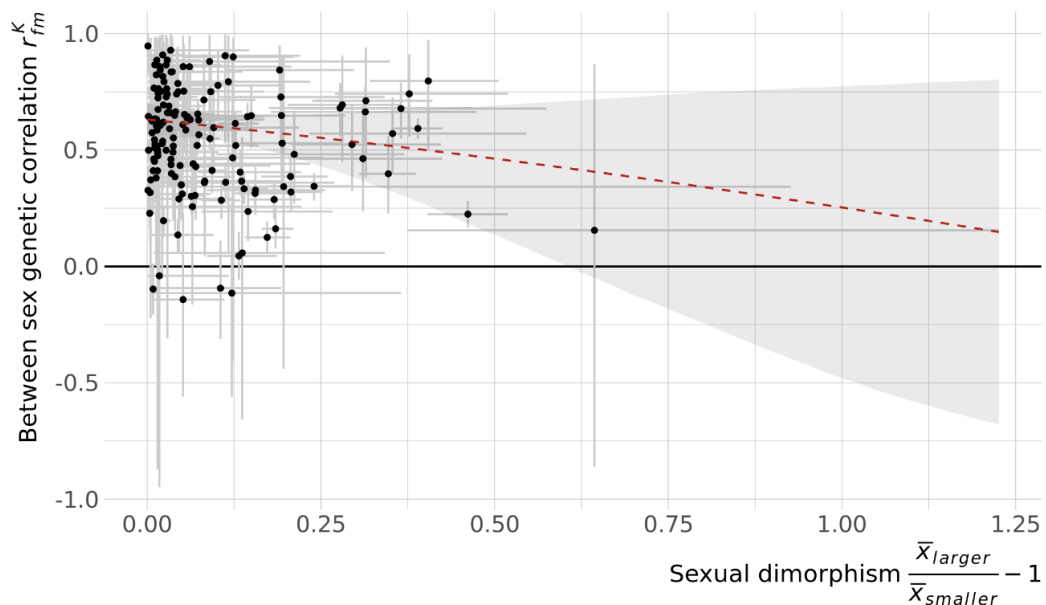


Figure 2: The between-sex genetic correlation does not depend on sexual dimorphism in the trait. Each point is a trait, with error bars indicating the 95% credible interval (CI) in the estimates. The red line represents the model fit of a linear model on the Fisher-transformed r_{fm}^K , with the shaded region indicating the 95% credible interval, including propagation of trait level uncertainty. Sexual dimorphism is expressed as the SD ratio.

229 morphology, while several immune phenotypes have a
230 correlation close to 0 (Table S1).

231 To test the constraint that high r_{fm}^K places on the
232 evolution of dimorphism, we assessed whether r_{fm}^K is
233 lower for more dimorphic traits, which we would expect if
234 dimorphism is more often associated with a reduced inter-
235 sexual correlation. We fitted a linear model with Fisher-
236 transformed r_{mf}^K values as the dependent variable and
237 sexual dimorphism (expressed as the SD index),
238 propagating the uncertainty in both variables from the
239 trait-level models. Contrary to expectation, the between-
240 sex genetic correlation is not associated with sexual
241 dimorphism (Figure 2, slope: -0.49 [-1.34, 0.35]). Although
242 there is a trend in the expected direction, the relationship
243 is non-significant, and r_{fm}^K at monomorphism (i.e. the
244 intercept) is only slightly higher than the overall average:
245 0.630 [0.557, 0.698].

246 To investigate whether there were differences in the
247 genetic architecture of dimorphism between trait types
248 (Poissant *et al.*, 2010), we assigned each of the traits one
249 of four categories: behavior, morphology, physiology or
250 immunity (Table S1). We repeated the linear model
251 regressing r_{mf}^K on SD, now including trait category and the
252 SD:trait category interaction as additional parameters.
253 There is no evidence that the relationship between r_{mf}^K
254 and SD is different for different trait categories (Figure S2).
255 The average r_{mf}^K of trait categories, estimated at
256 monomorphism, can also not clearly be distinguished
257 (Figure S3).

258 Male and female genetic variances were often
259 unbalanced, and there was a clear tendency for male
260 genetic variance to be larger ($\frac{V_{G(m)}}{V_{G(f)}} = 1.14 [1.04, 1.23]$).
261 Thus, knock-out mutations have, on average, substantially
262 larger phenotypic effects in males. It has been noted
263 previously that mutations have larger fitness effects in

264 male *Drosophila* (Sharp & Agrawal, 2013), and differences
265 in genetic variance between the sexes may contribute
266 toward the evolution of dimorphism, even under a strong
267 between-sex genetic correlation (Wyman & Rowe, 2014).
268 However, we found no relation between the imbalance of
269 sex-specific variances and the level of SD (slope: 0.03 [-
270 0.26, 0.30]).

271 *Development of size dimorphism and r_{fm}^K*

272 Body size is dimorphic in many species, including the
273 mouse, yet it has been found numerous times that r_{fm}^G for
274 this trait is close to 1 (Roff, 2012). Nonetheless, sexual size
275 dimorphism can often be rapidly altered in response to the
276 environment (Badyaev, 2002), making this an important
277 trait to study in order to better understand the link
278 between the evolution of SD and sex-specific
279 architectures. As sexual size dimorphism (SSD) is
280 established through variable development rates and
281 times, it is especially useful to understand when in
282 development the effect of body size loci diverges between
283 the sexes. Unfortunately, there is very little data available
284 for the development of r_{fm}^G , with studies usually including
285 only 2 or 3 time points (Poissant & Coltman, 2009). In
286 contrast, the IMPC measures body weight weekly from
287 week 3 through 16, providing the opportunity to estimate
288 when during development the effects of expression
289 changes become sex-biased.

290 Using the same modelling approach described above,
291 we obtained estimates for SSD and r_{fm}^K at each week
292 (Figure 3). SSD increases strongly at the start of this period,
293 more than doubling between weeks 3 and 7 (Figure 3a).
294 r_{fm}^K decreases during that same time (Figure 3b), and both
295 parameters stabilize around 8 weeks. The two variables
296 follow a roughly linear negative relationship during
297 development (Figure 3c). A developmental link between
298 SSD and r_{fm}^K may be the result of sexually antagonistic

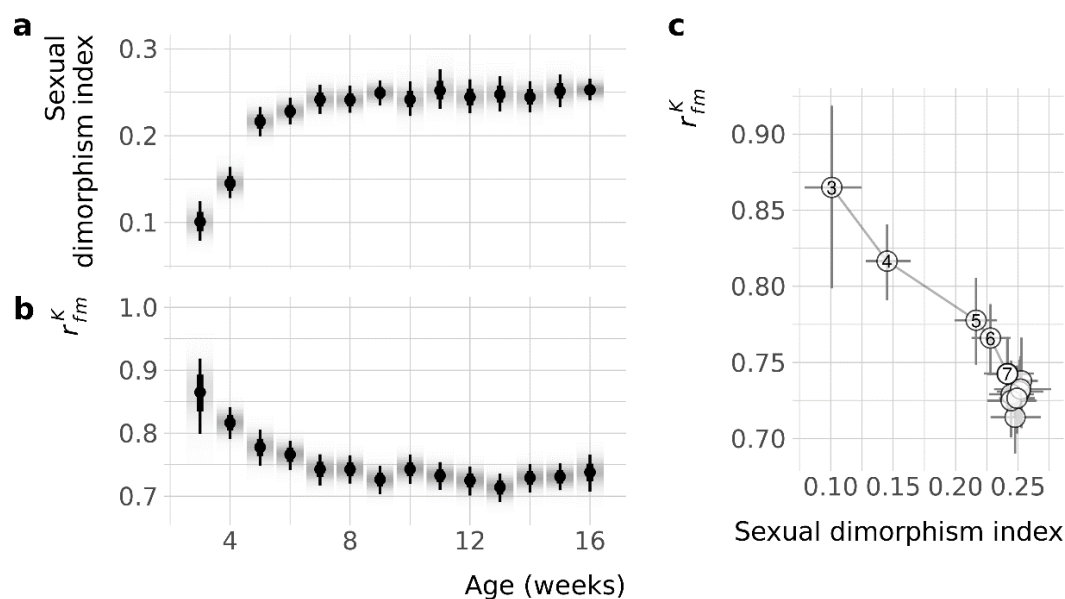


Figure 3: The between sex genetic correlation decreases as size dimorphism increases over development. (a) Estimates for sexual dimorphism in body mass for wildtype mice. Points indicate the posterior median with wide and narrow line segments denoting the 66% and 95% credible intervals respectively, and the density gradient represents the posterior density. (b) As in (a), but depicting the between sex genetic correlation. (c) Association of sexual size dimorphism and the r_{fm}^K during development. Points are posterior medians with 95% credible intervals, as in (a) and (b), with lines connecting subsequent week. Weeks 3 through 7 are numbered.

299 selection mainly acting in adulthood. This would bias sex-
 300 specific loci to be expressed only later in development,
 301 driving an increasing SSD and decreasing r_{fm} .
 302 Alternatively, strong trait integration during early
 303 development may pose significant constraints on the
 304 divergence of the sexes before 6 weeks.

305 *Identification of knock-out genotypes with sexually* 306 *discordant effects*

307 To gain insight into the extent to which sex-specific
 308 architectures are shared between different traits, we
 309 quantified to what extent knock-out genotypes have
 310 consistent sexually concordant or discordant effects. We
 311 separated the sexually concordant and discordant effect of
 312 each genotype on a trait by projecting the estimated effect
 313 (Best Linear Unbiased Predictor) along two independent
 314 axes (Ruzicka *et al.*, 2019), the positive and negative

315 diagonal of a female vs male plot (as in Figure S1). Then, in
 316 order to differentiate knockouts with strong versus weak
 317 discordant effects, we looked for genotypes with a
 318 consistently low or high ranking along the discordant axis.

319 We identified five knock-out genotypes that
 320 consistently had smaller sexually discordant effects,
 321 compared to other genotypes (Figure 4). Those five
 322 genotypes also had much smaller concordant effects,
 323 indicating that their phenotypes are consistently average.
 324 Unsurprisingly, these were five wildtype genotypes.
 325 Additionally, 24 genotypes had larger than average
 326 discordant effects (Figure 4, Table S2). These genotypes
 327 tended to affect the sexes differently, across many traits.
 328 An analysis of Gene Ontologies for the genes that were
 329 knocked out in these genotypes, revealed no significantly
 330 overrepresented categories. In contrast to the 29

331 discordant genotypes, 292 genotypes (out of 2543) had
332 consistently small or large concordant effects. This
333 difference suggests that traits are more likely to genetically
334 co-vary in their average value, rather than in their
335 dimorphism.

336 *Sex-biased gene expression and fertility*

337 Many investigations into the evolutionary significance
338 of gene expression to SD have focused on sex-biased gene
339 expression (Grath & Parsch, 2016). Of specific interest are
340 expression differences in the gonads, where most sex-
341 biased expression occurs. In these studies, it is often
342 assumed that gonadal expression bias reflects important
343 sex-specific fertility functions, however, it is usually not
344 possible to verify this. Combining previously published
345 gonadal expression data (Rinn *et al.*, 2004) with fertility

346 data from the IMPC database, however, allowed us to test
347 whether the expression knock-out of sex-biased genes
348 causes sex-specific infertility.

349 As predicted, fertility status was significantly
350 associated with expression bias category (i.e. male-biased,
351 female-biased or unbiased; $\chi^2_6 = 76.6$, $p < 0.001$, Figure S4).
352 Gene knockouts of female-biased or unbiased genes led to
353 male-limited infertility in 1.5% of cases, but this increased
354 to 11% of cases when knocking out male-biased genes.
355 Female-limited fertility on the other hand was less
356 common in general and showed no increase with knock-
357 outs of female-biased genes (Figure S4), possibly because
358 female gametogenesis is largely encoded during fetal
359 development and then arrested.

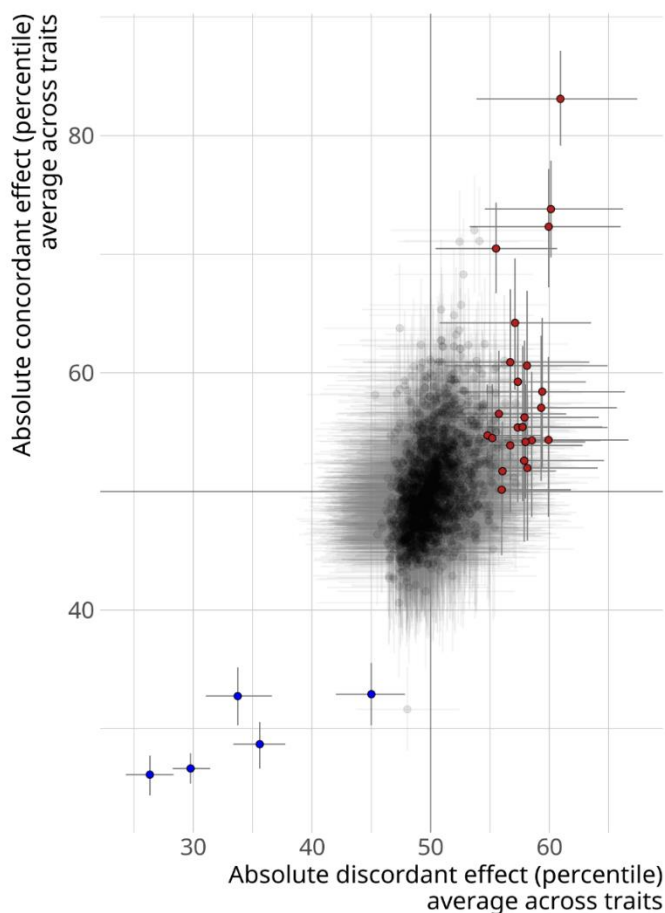


Figure 4: Identifying genotypes with consistent sexually discordant effects. Each point is a genotype, having been tested for at least 50 traits, with error bars denoting 95% credible intervals (CIs). The average percentile rank for the absolute sexually discordant effect of a genotype is plotted along the x-axis. The y-axis shows the average percentile rank for the absolute concordant effect. Red points indicate genotypes that tend to have more sexually discordant effects than other genotypes, while blue points are genotypes that have less discordant effects (CI does not overlap 50th percentile).

360 Discussion

361 Using the extensive phenotyping effort of gene knock-
362 out mouse lines by the IMPC, we have tested for the extent
363 of overlap in trait genetic architecture between males and
364 females. Even in the mouse, which is relatively
365 monomorphic when compared to many other vertebrates,
366 it is surprisingly common for traits to show clear
367 differences between the sexes after controlling for body
368 size. This therefore suggests that sexual dimorphism is not
369 the exception but the norm across many crucial somatic
370 traits.

371 Furthermore, traits are affected differently by knock-
372 out mutations depending on the sex of the individual. This
373 clearly illustrates that studies of gene function must
374 account for sex, as knock-out effects may only be easily
375 detectable in one of the sexes (International Mouse
376 Phenotyping Consortium *et al.*, 2017). Alterations in gene
377 expression are often thought to be a common mechanism
378 to resolve intra-locus sexual conflict by making gene
379 expression sex-biased or sex-specific (Grath & Parsch,
380 2016). This assumes a shared genetic architecture, which
381 is differentially regulated between the sexes. Our work
382 suggests that the underlying architecture may differ
383 between the sexes in many cases, and the low estimates
384 of r_{fm}^K that we recover highlight a different potential role
385 of gene expression in the evolution of SD.

386 Mutations of large regulatory effect can often be
387 expected to alter SD, providing one way to resolve intra-
388 locus sexual conflict. However, these regulatory changes
389 need not result in sex-biased gene expression, as our work
390 suggests that regulatory changes in both sexes, in this case
391 elimination of expression in both sexes through knockouts,
392 often predominantly only affect the phenotype of one. In
393 other words, sexually concordant regulatory changes can
394 result in sexually discordant phenotypic effects, and our

395 results suggest that this commonly occurs. This provides
396 an alternative route to dimorphism through sex-specific
397 genetic architecture, rather than differential regulation of
398 shared architecture. This could, for example, be the result
399 of interactions with sex-biased genes in the same
400 regulatory network, or of a sex-bias in the size of the cell
401 populations expressing the gene. It appears likely that the
402 modulation of gene expression, either through sex-bias in
403 the downstream phenotypic effects or in the expression
404 itself, is a major contributor to the evolution of SD.

405 Although mutations of large effect, especially gene
406 deletions, can have deleterious effects on other traits
407 through pleiotropy, many genes are non-essential
408 (Amsterdam *et al.*, 2004; Liao & Zhang, 2007; Georgi *et al.*,
409 2013). This suggests significant regulatory potential in the
410 evolution of SD. Additionally, the knockout mutations
411 assessed here likely represent an extreme form of
412 regulatory variation, which we would expect to have
413 similar, if less drastic, sex-specific effects, and more often
414 contribute to SD.

415 As others have previously indicated (Cowley & Atchley,
416 1988; Reeve & Fairbairn, 2001; Bonduriansky & Rowe,
417 2005), r_{fm}^A may not be as strong an indicator of constraint
418 as was originally suggested (Lande, 1980). While r_{fm}^A is
419 very useful in describing the potential for the standing
420 genetic variation to alter SD in a single or a few
421 generations, it cannot detect decoupling in trait
422 architectures which are currently lacking variation. Our
423 results indicate that even high r_{fm}^A traits may be
424 susceptible to changes in SD, as most traits have cryptic
425 parts of the genetic architecture in which new mutations
426 will likely have sex discordant effects. Importantly,
427 changes in the architecture itself, such as changes in gene
428 pathways or the recruitment of new transcription factors,

429 are not necessary to have occurred, contrasting with a
430 common interpretation of a strong r_{fm}^A .

431 A potential limitation of this study is that the mice are
432 inbred, resulting in genome-wide homozygosity. This
433 means that the phenotypic variation is expected to be
434 relatively small, making the effects of knockouts appear
435 stronger. Additionally, the effects of dominance and
436 epistasis are artificially limited. As it has been suggested
437 that sex-specific dominance is pervasive (Grieshop &
438 Arnqvist, 2018), and epistatic interactions could be
439 affected by sex as well, our estimates of r_{fm}^K could
440 potentially be biased upwards. It is also important to note
441 that sex-linked genetic architecture can allow for the
442 evolution of dimorphism. However, given the relatively
443 small size and limited gene content of the mouse Y
444 chromosome (Soh *et al.*, 2014), the role of the Y in sex-
445 specific genetic architecture for a broad array of somatic
446 traits is unclear.

447 The vast majority of genotypes were neither strongly
448 nor weakly discordant across traits, suggesting there are
449 very few or no “sex-specific genes” or “SD genes” but
450 rather many different genes have sex-specific effects on
451 different traits. The few genotypes that did show some
452 consistently discordant effects had no functional
453 categories in common, also suggesting that SD is regulated
454 differently in different traits. As we identified more
455 genotypes that had consistently large concordant effects,
456 the genetic covariance between trait means is likely
457 stronger than between SD of different traits. Large-scale
458 analyses in a multivariate framework are needed to fully
459 clarify the covariance of expression variance across traits
460 and sex, in order to come to a complete understanding of
461 the evolutionary constraints on SD.

462 In conclusion, using a dataset of unprecedented size,
463 we have demonstrated that traits harbor a surprising
464 amount of sex-specific genetic architecture, as sexes

465 respond variably to knock-out mutations. These results
466 may help explain why SD is common, evolvable and
467 variable, even under supposed strong genetic constraints.
468 While these differences clearly indicate that the genotype-
469 to-phenotype mapping is sex-dependent for most traits, it
470 remains unclear what underlying mechanisms are the
471 cause for this. We hope future work will help elucidate
472 proximate causes and evolutionary consequences of this
473 work.

474 Methods

475 We obtained data from the online IMPC genotype-
476 phenotype database. We selected phenotypes for analysis
477 by requesting all uni-dimensional continuous traits,
478 excluding legacy pipelines. We also excluded traits that
479 were not measured in both sexes, fitness-related traits
480 (such as reproductive screening), body size (we analyzed
481 body size separately), traits with fewer than 100
482 genotypes, and traits that were clearly not actually
483 continuous (such as a count of the number of ribs). After
484 triage, we had 260 traits for which we downloaded all
485 available phenotype data, including both knock-out
486 phenotypes and control data. On average we obtained
487 data for 8,069 control mice and 21,513 mice from 1,713
488 knock-out lines, per trait.

489 Sexual dimorphism and r_{fm}^K of mouse traits

490 As we were interested in estimating a single value for
491 r_{fm}^K per trait, we collapsed different sources of genetic
492 variance into genotypes. As some gene knock-outs were
493 performed in different genetic backgrounds, some genes
494 had multiple allelic knock-outs, and some genes were
495 tested in different zygosity, we defined each unique
496 gene:allele:background:zygosity combination as a
497 separate genotype. Note that the genetic backgrounds are
498 all C57BL/6 mice, but a different sub-strain.

499 To each of the trait datasets, we fitted a Bayesian linear
 500 mixed model with the goal of estimating both the
 501 between-sex genetic correlation (r_{fm}^K) and sexual
 502 dimorphism (SD). We opted for the analysis of single traits
 503 as opposed to multivariate models, since phenotypes have
 504 been measured across differing sets of individuals and
 505 knock-outs. Additionally, the univariate models were
 506 computationally expensive, with each model taking
 507 several days to a week to fit, and multivariate models
 508 would be logistically unfeasible. Each model had one of the
 509 phenotypes as the dependent variable, which was
 510 standardized (centered and scaled to unit variance) and
 511 transformed (see below). We included sex as a population
 512 level effect (also called fixed effect), allowing an average
 513 level of dimorphism across genotypes, although we did not
 514 directly use this parameter as our measurement of SD (see
 515 below). We also included body mass as a population level
 516 parameter, since mice are size dimorphic. Body mass was
 517 standardized (centered and scaled to unit variance) prior
 518 to analysis. All analyses were repeated without body mass,
 519 and the qualitatively similar results can be found in the
 520 supplementary material, although we only recommend
 521 interpretation of the results accounting for body size.

522 To estimate r_{fm}^K we added group level parameters (also
 523 called random effects) of genotype for each sex, and their
 524 correlation. Finally, we added group level intercepts for
 525 known sources of variation when they were present, which
 526 were 1) the phenotyping center in which testing was
 527 performed, a parameter encoding several methodological
 528 differences (“meta group”), and 2) the date of testing. This
 529 leads to the final model definition (in *lme4/brms* syntax):
 530 *phenotype* ~ *weight* + *sex* + (*0* + *sex* | *genotype*) + (*1* |
 531 *center*) + (*1* | *meta_group*) + (*1* | *date*). In mathematical
 532 notation, following Gelman & Hill (2006):

$$533 \text{ trait}_i \sim N(\alpha_{j[l],k[l],l[i],m[i]} + \beta_{\text{sex},j[l]}(\text{sex})$$

$$534 + \beta_3(\text{body mass}), \sigma^2)$$

$$535 \begin{pmatrix} \beta_{\text{female},j} \\ \beta_{\text{male},j} \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_{\beta_{\text{female},j}} \\ \mu_{\beta_{\text{male},j}} \end{pmatrix}, \begin{pmatrix} \sigma_{\beta_{\text{female},j}}^2 & \rho_{\beta_{\text{female},j}\beta_{\text{male},j}} \\ \rho_{\beta_{\text{male},j}\beta_{\text{female},j}} & \sigma_{\beta_{\text{male},j}}^2 \end{pmatrix} \right)$$

536 for genotype $j = 1, \dots, J$

$$537 \alpha_k \sim N(\mu_{\alpha_k}, \sigma_{\alpha_k}^2), \text{ for center } k = 1, \dots, K$$

$$538 \alpha_l \sim N(\mu_{\alpha_l}, \sigma_{\alpha_l}^2), \text{ for meta group } l = 1, \dots, L$$

$$539 \alpha_m \sim N(\mu_{\alpha_m}, \sigma_{\alpha_m}^2), \text{ for date } m = 1, \dots, M$$

540 Parameter values were estimated using the *brms*
 541 (Bürkner, 2017, 2018) interface to the probabilistic
 542 programming language Stan (Carpenter et al., 2017). We
 543 used weakly informative prior distributions, with priors of
 544 $N(0, 1)$ for the intercept and $N(0, 2)$ for the effect of body
 545 mass. For the group level standard deviations and residual
 546 standard deviation we used the positive range of unit
 547 student-t distributions with 5 degrees of freedom. Finally,
 548 we used an LKJ prior with $\eta = 1$ for r_{fm}^K , which is uniform
 549 over the range -1 to 1. Posterior distributions were
 550 obtained using Stan’s no-U-turn HMC sampler, with 2
 551 chains of 8000 iterations, with the first 4000 used as warm-
 552 up and discarded. We additionally set the max tree-depth
 553 to 20 and the adapt delta parameter to 0.9. To evaluate
 554 the ability of our models to accurately estimate the
 555 between-sex genetic correlation, even though the sample
 556 size for each genotype was limited, we performed a
 557 simulation study (figure S7), confirming that our approach
 558 recovers the true value for r_{fm}^K .

559 In order to satisfy the assumption of approximately
 560 normal residuals, we preceded each analysis by estimation
 561 of a Box-Cox transformation, following the established
 562 methods by the IMPC (Kurbatova et al., 2019), using the
 563 simplified model definition: *phenotype* ~ *weight* + *sex* + (*0*

564 + sex | genotype). We estimated the transform using the
565 *bcnPower* method in the *car* package (Fox *et al.*, 2019),
566 with model fitting performed by *lme4* (Bates *et al.*, 2015).

567 After fitting all 260 trait models, we performed model
568 criticism. For each model we obtained the maximum \hat{R}
569 parameter, the number of divergences and the minimum
570 effective sample size. We removed all models that had a
571 maximum \hat{R} of more than 1.05, more than 2.5% divergent
572 draws, or a minimum effective sample size of less than
573 400. Finally, we performed visual posterior predictive
574 checks (Gabry *et al.*, 2019), and removed models that did
575 not reproduce the observed data distribution. Considering
576 the computational effort required for each of these
577 models, as well as that the number of successful models
578 was more than large enough for the analyses we wished to
579 perform, we did not attempt to remedy the failing models.
580 We performed visual checks to confirm that the excluded
581 traits did not have a bias in SD or r_{fm}^K . After model
582 criticism, 202 out of 260 models remained.

583 For each of these models we derived posterior
584 distributions of r_{fm}^K . Note that *brms* estimates standard
585 deviations and correlations directly, so no parameter
586 transformation was necessary. We then derived posterior
587 distributions of SD by predicting average male and female
588 phenotypes for wildtype (i.e. control group) mice. When
589 there were multiple genetic background variations in
590 which a trait was tested, we used the marginal means
591 across backgrounds. To make SD estimates comparable
592 across traits, we used a mean standardized effect size for
593 SD, the SD index: $\frac{\bar{x}_{larger\ sex}}{\bar{x}_{smaller\ sex}} - 1$, i.e. the ratio between
594 larger and smaller, divided by the residual standard
595 deviation. Note that the SD index requires that
596 comparisons to zero are biologically meaningful (i.e. traits
597 are measured on a ratio scale), which was not true for all
598 the traits in our data set, such as body temperature,

599 indices and fractional measures. We therefore performed
600 back transformations of the marginal means to the original
601 scale, and we only calculated SD for 156 out of 202 traits.

602 After obtaining the posteriors for each trait, we used a
603 linear model to test for a relationship between r_{fm}^K and SD.
604 In order to account for uncertainty in those estimates we
605 performed random draws from the posterior distributions
606 of those estimates to create 500 datasets. For each of
607 those samples we ran one MCMC chain of a $Zr_{fm}^K \sim SD$
608 *index* model using the *brm_multiple* function, and
609 performed inference on the combined set of 500 chains.
610 Note that we performed a Z-transformation on r_{fm}^K , also
611 called the Fisher transformation, to stabilize the variance.
612 Additionally, we performed the same procedure for the
613 ratio of the genetic variances: $\frac{V_G(larger)}{V_G(smaller)}$, which was log
614 transformed before analysis.

615 *Development of size dimorphism and r_{fm}^K*

616 To quantify sexual size dimorphism during
617 development, and associated changes in r_{fm}^K , we split the
618 body mass data into different ages. Mice were weighed
619 once a week, with most mice being measured between 4
620 and 16 weeks of age. For each week, we ran the same
621 analysis as for the separate traits outlined above.

622 *Identification of knock-out genotypes with sexually 623 discordant effects*

624 The concordant and discordant nature of knock-out
625 genotypes was determined by evaluating whether the
626 genotypes were consistently ranked low or high along the
627 concordant and discordant axes across traits. For each
628 trait, we used the multilevel model that was used to
629 estimate SD and r_{fm}^K , described above, to obtain estimates
630 of the male and female trait values for the measured
631 genotypes. We extracted the posteriors for the male and
632 female parameter for the genotype group term (BLUP).
633 Note that these estimates are adjusted for body weight

634 and environmental effects, have already undergone
635 parameter shrinkage, and are centered around zero. We
636 then translated the male and female phenotypes into
637 concordant and discordant effects, by rotating the axes so
638 that the concordant axis is the positive diagonal (female =
639 male) and the discordant axis is the negative diagonal
640 (female = -male). The absolute value along the two
641 diagonal axes was taken, so that the effect of a genotype
642 is larger when it is further from the population average.
643 Since the size of the discordant effects of a genotype is
644 strongly affected by the trait architecture (i.e. r_{mf}^K), we
645 assigned genotypes percentile ranks to aid comparison
646 across traits.

647 For all genotypes that were tested for at least 100
648 phenotypes, we then calculated the average concordant
649 and discordant rank across traits. Credible intervals (CIs)
650 for this average were calculated by computing that
651 average for 100 random draws of the posteriors. We then
652 categorized genotypes as less or more discordant than
653 average by checking whether the CI overlapped a median
654 rank (50th percentile in Figure 4).

655 For the genotypes that were more discordant than
656 average, we extracted which gene had been knocked out
657 and analyzed the associated gene ontology (GO) terms.
658 Using *goseq* (Young *et al.*, 2010) we tested for
659 overrepresented GO terms, using the hypergeometric
660 method for obtaining p-values. Finally, we adjusted the p-
661 values to control the false discovery rate (Benjamini &
662 Hochberg, 1995).

663 *Sex-biased gene expression and fertility*

664 We obtained published gene expression profiles of
665 male and female gonadal tissue from the ArrayExpress
666 database under accession number E-GEOD-1148 (Rinn *et al.*,
667 2004). Using *limma* (Ritchie *et al.*, 2015), we calculated
668 the difference in expression between the sexes (log₂ fold

669 change), and empirical Bayes moderated t-statistics with
670 adjusted p-values. We then classified genes as sex-biased
671 if the fold-change was at least 2, and the adjusted p-values
672 was significant ($\alpha = 0.05$). Genes that did not satisfy both
673 those criteria were categorized as unbiased.

674 We then obtained female and male specific fertility
675 data from the IMPC (phenotypes IMPC_FER_019_001 and
676 IMPC_FER_001_001), which are binary traits (fertile vs.
677 infertile) where each sex has been allowed to breed with a
678 wildtype mate. Combining these we defined four fertility
679 categories: fertile, female-limited infertile, male-limited
680 infertile and infertile. To test for an association between
681 gene expression category and fertility outcome after
682 knock-out, we performed a 3x4 chi-squared test for
683 independence.

684 Software

685 All analyses were performed in R v3.6.1 (R Core Team,
686 2019). Specific R packages used in the analyses are listed
687 above, and the *tidyverse* (Wickham *et al.*, 2019) was used
688 for general data handling and visualization.

689 References

- 690 Amsterdam, A., Nissen, R.M., Sun, Z., Swindell, E.C.,
691 Farrington, S. & Hopkins, N. 2004. Identification
692 of 315 Genes Essential for Early Zebrafish
693 Development. *Proc. Natl. Acad. Sci. U. S. A.* **101**:
694 12792–12797.
- 695 Bachtrog, D., Mank, J.E., Peichel, C.L., Kirkpatrick, M., Otto,
696 S.P., Ashman, T.-L., *et al.* 2014. Sex
697 Determination: Why So Many Ways of Doing It?
698 *PLoS Biol.* **12**: e1001899.
- 699 Badyaev, A.V. 2002. Growing apart: an ontogenetic
700 perspective on the evolution of sexual size
701 dimorphism. *Trends Ecol. Evol.* **17**: 369–378.
- 702 Bates, D., Mächler, M., Bolker, B. & Walker, S. 2015. Fitting
703 Linear Mixed-Effects Models Using lme4. *J. Stat.*
704 *Softw.* **67**: 1–48.

- 705 Benjamini, Y. & Hochberg, Y. 1995. Controlling the False
706 Discovery Rate: A Practical and Powerful
707 Approach to Multiple Testing. *J. R. Stat. Soc. Ser.
708 B Methodol.* **57**: 289–300.
- 709 Bonduriansky, R. & Rowe, L. 2005. Intralocus Sexual
710 Conflict and the Genetic Architecture of Sexually
711 Dimorphic Traits in *Prochyliza Xanthostoma*
712 (diptera: Piophilidae). *Evolution* **59**: 1965–1975.
- 713 Bürkner, P.-C. 2018. Advanced Bayesian Multilevel
714 Modeling with the R Package brms. *R J.* **10**: 395–
715 411.
- 716 Bürkner, P.-C. 2017. brms: An R Package for Bayesian
717 Multilevel Models Using Stan. *J. Stat. Softw.* **80**:
718 1–28.
- 719 Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D.,
720 Goodrich, B., Betancourt, M., *et al.* 2017. Stan: A
721 Probabilistic Programming Language. *J. Stat.
722 Softw.* **76**.
- 723 Chenoweth, S.F., Rundle, H.D. & Blows, M.W. 2008.
724 Genetic Constraints and the Evolution of Display
725 Trait Sexual Dimorphism by Natural and Sexual
726 Selection. *Am. Nat.* **171**: 22–34.
- 727 Cheverud, J.M., Vaughn, T.T., Pletscher, L.S., Peripato, A.C.,
728 Adams, E.S., Erikson, C.F., *et al.* 2001. Genetic
729 architecture of adiposity in the cross of LG/J and
730 SM/J inbred mice. *Mamm. Genome* **12**: 3–12.
- 731 Cowley, D.E. & Atchley, W.R. 1988. Quantitative Genetics
732 of *Drosophila Melanogaster*. II. Heritabilities and
733 Genetic Correlations between Sexes for Head and
734 Thorax Traits. *Genetics* **119**: 421–433.
- 735 Cowley, D.E., Atchley, W.R. & Rutledge, J.J. 1986.
736 Quantitative Genetics of *Drosophila*
737 *Melanogaster*. I. Sexual Dimorphism in Genetic
738 Parameters for Wing Traits. *Genetics* **114**: 549–
739 566.
- 740 Delph, L.F., Steven, J.C., Anderson, I.A., Herlihy, C.R. & Iii,
741 E.D.B. 2011. Elimination of a Genetic Correlation
742 Between the Sexes Via Artificial Correlational
743 Selection. *Evolution* **65**: 2872–2880.
- 744 Ellegren, H. & Parsch, J. 2007. The evolution of sex-biased
745 genes and sex-biased gene expression. *Nat. Rev.
746 Genet.* **8**: 689–698.
- 747 Fairbairn, D.J. & Roff, D.A. 2006. The quantitative genetics
748 of sexual dimorphism: assessing the importance
749 of sex-linkage. *Heredity* **97**: 319–328.
- 750 Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-
751 Bovy, G., *et al.* 2019. *car: Companion to Applied*
752 *Regression*.
- 753 Gabry, J., Simpson, D., Vehtari, A., Betancourt, M. &
754 Gelman, A. 2019. Visualization in Bayesian
755 workflow. *J. R. Stat. Soc. Ser. A Stat. Soc.* **182**:
756 389–402.
- 757 Galouzis, C.C. & Prud'homme, B. 2021. Transvection
758 regulates the sex-biased expression of a fly X-
759 linked gene. *Science* **371**: 396–400. American
760 Association for the Advancement of Science.
- 761 Gelman, A. & Hill, J. 2006. Data Analysis Using Regression
762 and Multilevel/Hierarchical Models. Cambridge
763 University Press.
- 764 Georgi, B., Voight, B.F. & Bucan, M. 2013. From mouse to
765 human: evolutionary genomics analysis of human
766 orthologs of essential genes. *PLoS Genet.* **9**.
- 767 Grath, S. & Parsch, J. 2016. Sex-Biased Gene Expression.
768 *Annu. Rev. Genet.* **50**: 29–44.
- 769 Grieshop, K. & Arnqvist, G. 2018. Sex-specific dominance
770 reversal of genetic variation for fitness. *PLOS Biol.*
771 **16**: e2006810.
- 772 Griffin, R.M., Dean, R., Grace, J.L., Ryden, P. & Friberg, U.
773 2013. The Shared Genome Is a Pervasive
774 Constraint on the Evolution of Sex-Biased Gene
775 Expression. *Mol. Biol. Evol.* **30**: 2168–2176.
- 776 Hanrahan, J.P. & Eisen, E.J. 1973. Sexual dimorphism and
777 direct and maternal genetic effects on body
778 weight in mice. *Theor. Appl. Genet.* **43**: 39–45.
- 779 Hansen, T.F. 2006. The Evolution of Genetic Architecture.
780 *Annu. Rev. Ecol. Evol. Syst.* **37**: 123–157.
- 781 Hermansen, J.S., Starrfelt, J., Voje, K.L. & Stenseth, N.C.
782 2018. Macroevolutionary consequences of sexual
783 conflict. *Biol. Lett.* **14**: 20180186.
- 784 International Mouse Phenotyping Consortium, Karp, N.A.,
785 Mason, J., Beaudet, A.L., Benjamini, Y., Bower, L.,
786 *et al.* 2017. Prevalence of sexual dimorphism in
787 mammalian phenotypic traits. *Nat. Commun.* **8**.

- 788 Khila, A., Abouheif, E. & Rowe, L. 2012. Function,
789 Developmental Genetics, and Fitness
790 Consequences of a Sexually Antagonistic Trait.
791 *Science* **336**: 585–589. American Association for
792 the Advancement of Science.
- 793 Kruschke, J.K. 2018. Rejecting or Accepting Parameter
794 Values in Bayesian Estimation. *Adv. Methods*
795 *Pract. Psychol. Sci.* **1**: 270–280.
- 796 Kurbatova, N., Karp, N., Mason, J. & Haselimashhadi, H.
797 2019. *PhenStat: Statistical analysis of phenotypic*
798 *data*. Bioconductor version: Release (3.10).
- 799 Lande, R. 1980. Sexual dimorphism, sexual selection and
800 adaptation in polygenic characters. *Evolution* **34**:
801 292–305.
- 802 Liao, B.-Y. & Zhang, J. 2007. Mouse duplicate genes are as
803 essential as singletons. *Trends Genet.* **23**: 378–
804 381.
- 805 Mank, J.E. 2017. The transcriptional architecture of
806 phenotypic dimorphism. *Nat. Ecol. Evol.* **1**: 1–7.
807 Nature Publishing Group.
- 808 Owens, I.P.F. & Hartley, I.R. 1998. Sexual dimorphism in
809 birds: why are there so many different forms of
810 dimorphism? *Proc. R. Soc. Lond. B Biol. Sci.* **265**:
811 397–407.
- 812 Poissant, J. & Coltman, D.W. 2009. The ontogeny of cross-
813 sex genetic correlations: an analysis of patterns. *J.*
814 *Evol. Biol.* **22**: 2558–2562.
- 815 Poissant, J., Wilson, A.J. & Coltman, D.W. 2010. Sex-
816 specific genetic variance and the evolution of
817 sexual dimorphism: a systematic review of cross-
818 sex genetic correlations. *Evolution* **64**: 97–107.
- 819 R Core Team. 2019. *R: A Language and Environment for*
820 *Statistical Computing*. R Foundation for Statistical
821 Computing, Vienna, Austria.
- 822 Reeve, J.P. & Fairbairn, D.J. 2001. Predicting the evolution
823 of sexual size dimorphism. *J. Evol. Biol.* **14**: 244–
824 254.
- 825 Rice, W.R. & Chippindale, A.K. 2001. Intersexual
826 ontogenetic conflict. *J. Evol. Biol.* **14**: 685–693.
- 827 Rinn, J.L., Rozowsky, J.S., Laurenzi, I.J., Petersen, P.H., Zou,
828 K., Zhong, W., *et al.* 2004. Major Molecular
829 Differences between Mammalian Sexes Are
830 Involved in Drug Metabolism and Renal Function.
831 *Dev. Cell* **6**: 791–800.
- 832 Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi,
833 W., *et al.* 2015. limma powers differential
834 expression analyses for RNA-sequencing and
835 microarray studies. *Nucleic Acids Res.* **43**: e47–
836 e47.
- 837 Roff, D.A. 2012. *Evolutionary Quantitative Genetics*.
838 Springer Science & Business Media.
- 839 Ruzicka, F., Hill, M.S., Pennell, T.M., Flis, I., Ingleby, F.C.,
840 Mott, R., *et al.* 2019. Genome-wide sexually
841 antagonistic variants reveal long-standing
842 constraints on sexual dimorphism in fruit flies.
843 *PLOS Biol.* **17**: e3000244.
- 844 Sharp, N.P. & Agrawal, A.F. 2013. Male-biased fitness
845 effects of spontaneous mutation in drosophila
846 melanogaster. *Evolution* **67**: 1189–1195.
- 847 Soh, Y.Q.S., Alföldi, J., Pyntikova, T., Brown, L.G., Graves,
848 T., Minx, P.J., *et al.* 2014. Sequencing the Mouse
849 Y Chromosome Reveals Convergent Gene
850 Acquisition and Amplification on Both Sex
851 Chromosomes. *Cell* **159**: 800–813.
- 852 Stewart, A.D. & Rice, W.R. 2018. Arrest of sex-specific
853 adaptation during the evolution of sexual
854 dimorphism in *Drosophila*. *Nat. Ecol. Evol.* **2**:
855 1507–1513.
- 856 The International Mouse Phenotyping Consortium,
857 Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L.,
858 Wong, M.D., *et al.* 2016. High-throughput
859 discovery of novel developmental phenotypes.
860 *Nature* **537**: 508–514.
- 861 Wellek, S. 2010. Testing statistical hypotheses of
862 equivalence and noninferiority. Chapman and
863 Hall/CRC.
- 864 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan,
865 L., François, R., *et al.* 2019. Welcome to the
866 Tidyverse. *J. Open Source Softw.* **4**: 1686.
- 867 Wyman, M.J. & Rowe, L. 2014. Male Bias in Distributions
868 of Additive Genetic, Residual, and Phenotypic
869 Variances of Shared Traits. *Am. Nat.* **184**: 326–
870 337.

871 Young, M.D., Wakefield, M.J., Smyth, G.K. & Oshlack, A.
872 2010. Gene ontology analysis for RNA-seq:
873 accounting for selection bias. *Genome Biol.* **11**:
874 R14.

875

876 Acknowledgements

877 We gratefully acknowledge support from a
878 Canada 150 Research Chair and the European
879 Research Council (grant agreement 680951). We
880 thank Locke Rowe, Michael Whitlock and members of
881 the Mank Lab for helpful comments and suggestions.
882 Finally, we are indebted to all contributors to the
883 IMPC project and their commitment to open science.

884 Author Contributions

885 Both authors conceived of the study, WvdB
886 performed the data analysis, and both authors wrote
887 the manuscript.

901

888 Competing Interests statement

889 The authors declare they have no competing
890 interests.

891 Data Availability Statement

892 No new data was collected for this study. All raw
893 phenotype data is available from the International
894 Mouse Phenotyping Consortium
895 (<https://www.mousephenotype.org/>). The gene
896 expression profiles of male and female gonadal tissue
897 is available from the ArrayExpress database under
898 accession number E-GEOD-1148. All estimates used in
899 down-stream analyses are available in the
900 Supplementary Information.