Full Title: Patterns of recent natural selection on genetic loci associated with sexually differentiated human body size and shape phenotypes

Short Title: Recent natural selection on human sexually differentiated phenotypes

Audrey M. Arner[1], Kathleen E. Grogan[1,2], Mark Grabowski[3,4], Hugo Reyes-Centeno[3,5], George H. Perry[1,3,6,*]

[1]Departments of Anthropology and Biology, Pennsylvania State University, University Park, Pennsylvania, United States of America [2]Departments of Anthropology and Biology, University of Cincinnati, Cincinnati, Ohio, United States of America [3]DFG Center for Advanced Studies, University of Tübingen, Tübingen, Germany [4]Research Centre in Evolutionary Anthropology and Palaeoecology, Liverpool John Moores University, Liverpool, United Kingdom [5]Department of Anthropology & William S. Webb Museum of Anthropology, University of Kentucky, Lexington, Kentucky, United States of America [6]Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, Pennsylvania, United States of America

*Corresponding author: G.H.P (ghp3@psu.edu)

Keywords: GWAS, sexual dimorphism, UK Biobank, natural selection, height

Abstract

Levels of sex differences for human body size and shape phenotypes are hypothesized to have adaptively reduced following the agricultural transition as part of an evolutionary response to relatively more equal divisions of labor and new technology adoption. In this study, we tested this hypothesis by studying genetic variants associated with five sexually differentiated human phenotypes: height, body mass, hip circumference, body fat percentage, and waist circumference. We first analyzed genome-wide association (GWAS) results for UK Biobank individuals (~197,000 females and ~167,000 males) to identify a total of 119,023 single nucleotide polymorphisms (SNPs) significantly associated with at least one of the studied phenotypes in females, males, or both sexes ($P<5x10^{-8}$). From these loci we then identified 3,016 SNPs (2.5%) with significant differences in the strength of association between the female- and male-specific GWAS results at a low false-discovery rate (FDR<0.001). Genes with known roles in sexual differentiation are significantly enriched for co-localization with one or more of these SNPs versus SNPs associated with the phenotypes generally but not with sex differences (2.93-fold enrichment; permutation test; P=0.0041). We also confirmed that the identified variants are disproportionately associated with greater phenotype effect sizes in the sex with the stronger association value. We then used the singleton density score statistic, which quantifies recent (within the last ~3,000 years; post-agriculture adoption in Britain) changes in the frequencies of alleles underlying polygenic traits, to identify a signature of recent positive selection on alleles associated with greater body fat percentage in females (permutation test; P=0.0038; FDR=0.0380), directionally opposite to that predicted by the sex differentiation reduction hypothesis. Otherwise, we found no evidence of positive selection for sex difference-associated alleles for any other trait. Overall, our results

challenge the longstanding hypothesis that sex differences adaptively decreased following subsistence transitions from hunting and gathering to agriculture.

Author Summary

There is uncertainty regarding the evolutionary history of human sex differences for quantitative body size and shape phenotypes. In this study we identified thousands of genetic loci that differentially impact body size and shape trait variation between females and males using a large sample of UK Biobank individuals. After confirming the biological plausibility of these loci, we used a population genomics approach to study the recent (over the past 3,000 years) evolutionary histories of these loci in this population. We observed significant increases in the frequencies of alleles associated with greater body fat percentage in females. This result is contradictory to longstanding hypotheses that sex differences have adaptively decreased following subsistence transitions from hunting and gathering to agriculture.

Main Text

Introduction

In many vertebrate species, it is not uncommon for morphological phenotypes to have average size and shape differences between females and males [1]. Traits with average phenotype values that differ by sex but with overlapping trait distributions – such as human height and body fat percentage – are described as 'sexually differentiated' traits. Technically, the more commonly-used term 'sexually dimorphic' is specific to non-overlapping traits; for example, exaggerated ornamentation in male guppies, peacocks, and mandrills [2–5].

Some sexually differentiated traits are believed by many researchers to be the result of sexual selection. In species with high inter-male competition for mates, larger males may have a competitive advantage that results in increased fitness [6]. Perhaps as a result, the magnitude of sexually differentiated phenotypes are often greater in polygynous species with high competition (e.g. gorillas) and lower in monogamous species (e.g. gibbons) [6–8]. Finally, there are major differences in the degree of body size and shape sex differences between closely related species, suggesting the potential for the relatively rapid evolution of sexually differentiated traits [6].

In humans, females and males exhibit significant but relatively subtle differences in many anthropometric phenotypes [9]. For example, European and African males are an average of approximately 9% taller in height and 15% heavier in body mass than females from the same populations [10,11]. Humans also exhibit sexually differentiated biometric and disease phenotypes, especially those related to immune function [12,13].

83 There is little consensus regarding the evolutionary history of sexually differentiated traits in our
84 species [11,14,15]. Current levels of human sex differences may partially reflect an evolutionary
85 history of complex interactions between our biology and culture, but the timing, direction, and
86 forces responsible for any adaptive changes in these patterns are debated. For example, it has been
87 hypothesized that the degree of sexual differentiation for body size phenotypes in Europe likely
88 decreased following recent (within the past ~10,000 years) shifts from a foraging-based
89 subsistence strategy to one relying primarily on food production, in response to less pronounced
90 divisions of labor and mobility [16]. However, other scholars suggest that sexually differentiated
91 phenotypes would respond to selection at too slow an evolutionary rate to respond to recent
92 environmental and cultural changes; thus, any such recent changes are more likely to reflect
93 genetic drift and/or non-genetic responses to environmental changes rather than natural selection
94 [17].
95
96 In this study, we combined genomic and evolutionary analyses to quantify how recent changes in
97 lifestyle and culture might be affecting the underlying genetic basis of human sexually
98 differentiated traits. First, we applied a genome wide association study (GWAS)-based approach
99 to identify genetic variants associated with sexually differentiated phenotypes based on data from
100 the UK Biobank study [18]. With a very large number of UK Biobank participants (~361,000
101 individuals in the dataset we analyzed), this analysis represents a powerful extension of several
102 previous GWAS-based studies of the genetic architecture of human sexually differentiated
103 anthropometric traits [19]. We then used the Singleton Density Score (SDS), a statistic that
104 identifies signatures of polygenic adaptation that acted within the last ~3,000 years [20], a period
105 following the transition to agriculture in present-day United Kingdom [21].

106

107 **Results**

108

109 We analyzed genome-wide association study (GWAS) summary statistics for the following five
110 sexually differentiated anthropometric phenotypes that were produced by the Neale Lab [22] using
111 data from the UK Biobank [18]: height, body mass, hip circumference, body fat percentage, and
112 waist circumference. We chose these traits given their relevance to the motivating evolutionary
113 hypothesis and because they have been extensively studied from anthropological and/or genomics
114 perspectives [9,23–25]. We did not include body mass index (BMI) in our set of body size and
115 shape phenotypes following concerns that have been voiced about this metric related to its failure
116 to quantify body shape when indicating obesity and obesity-related health risks [26,27] and other,
117 ethical issues [28]. We instead included the two constituents of BMI (height and body mass) as
118 independent variables in our analysis.

119

120 ***Identification of phenotype-associated SNPs***

121 We analyzed summary statistics from GWAS that were performed separately in ~194,000 females
122 and ~167,000 males of white British genetic ancestry on ~13.8 million autosomal SNPs [22]. SNPs
123 with minor allele frequencies < 0.05 and low imputation quality were filtered out. We further
124 restricted our analysis to SNPs that 1) passed these filters in both females and males and 2) had
125 SDS values available [20], given our motivation to conduct downstream evolutionary analyses.
126 These filtering steps resulted in ~4.4 million genome-wide SNPs for each sex-stratified GWAS.
127 For each phenotype, we identified significantly phenotype-associated SNPs present in females,
128 males, or both using the genome-wide significance threshold of $P = 5x10^{-8}$ commonly applied in

129 UK Biobank studies [29–31]. We identified the following total (not yet pruned for linkage
130 disequilibrium) numbers of phenotype-associated SNPs significant in either females, males, or
131 both: 67,738 for height, 15,669 for body mass, 12,580 for hip circumference, 10,538 for body fat
132 percent, and 7,765 for waist circumference (Fig 1A-B; S1 Table).
133
134 ***SNPs disproportionately associated with female or male trait variation***
135 From the sets of phenotype-associated SNPs that were significant in females, males, or both for
136 each phenotype (SexDiff-associated SNPs), we identified those SNPs with significant differences
137 in the statistical strengths of association for the female vs. male-specific GWAS results using the
138 t-SexDiff statistic [19]. This statistic estimates the probability of difference between female-
139 specific and male-specific effect sizes given their standard errors and considering the genome-
140 wide correlation between male and female effect sizes for each trait (see Methods). To account for
141 multiple testing, we performed this analysis with four different false-discovery rate (FDR) cutoffs
142 for each phenotype: 0.05, 0.01, 0.005, and 0.001.
143
144 At the most stringent FDR cutoff of 0.001, we identified the following number of SexDiff-
145 associated SNPs: 677 for height, 541 for body mass, 808 for hip circumference, 439 for body fat
146 percentage, and 551 for waist circumference (Fig 1C-D; S1 Table). In addition to identifying
147 SexDiff-associated SNPs, for each phenotype we calculated the proportion of SexDiff-associated
148 SNPs at the FDR threshold of 0.001 to the number of phenotype-associated SNPs, which ranged
149 from 0.0010 for height to 0.0718 for waist circumference.

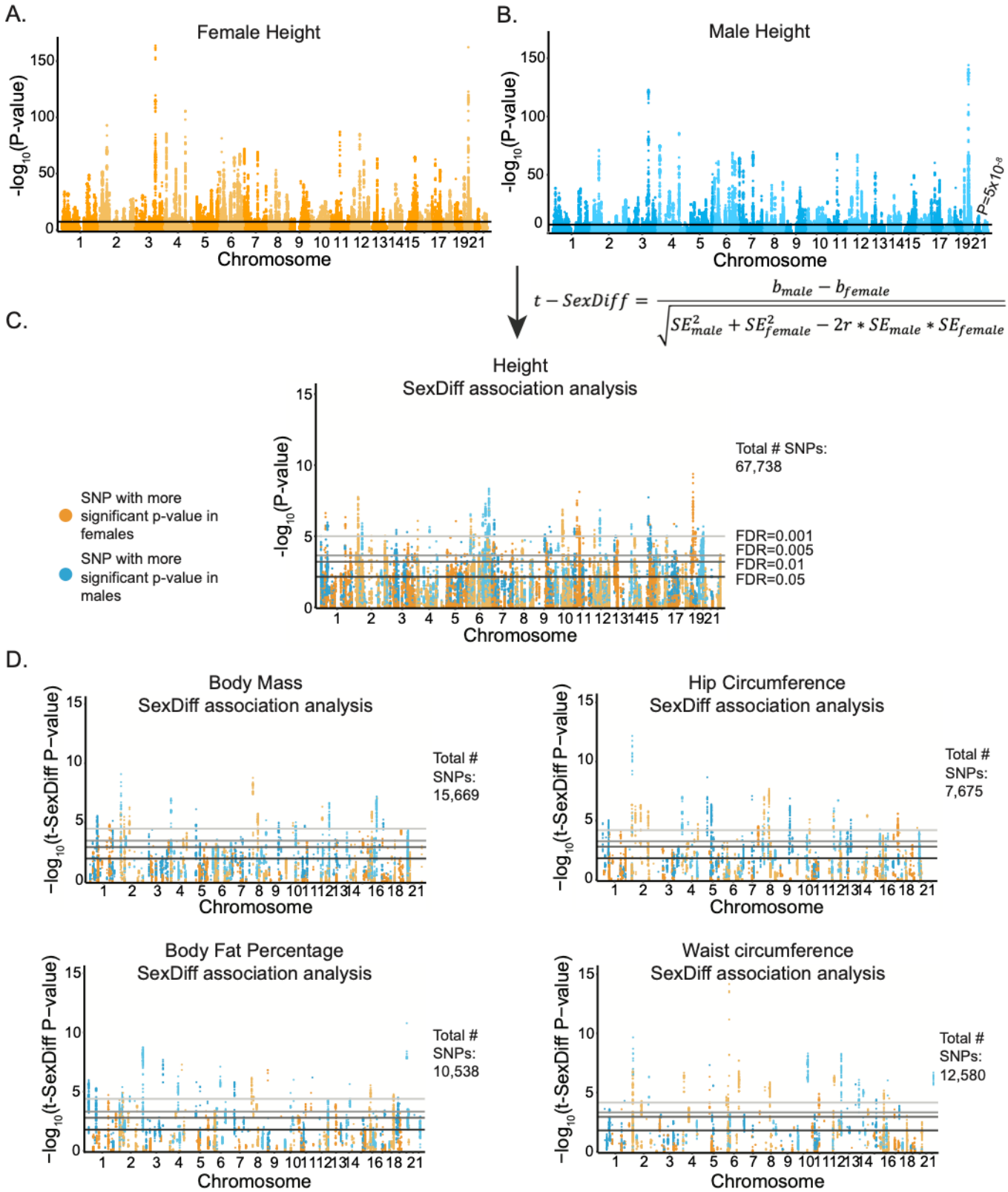

Fig 1. SexDiff-associated SNPs for five anthropometric phenotypes. (A) Manhattan plot depicting -$\log_{10}$ P-values for the association of each genome-wide SNP with female height. The black line corresponds to the genome-wide significance threshold of $P=5x10^{-8}$. (B) Manhattan plot for SNP associations with male height. (C) For each of the 67,738 SNPs significantly associated female and/or male height, we used the equation shown to test whether the SNP was disproportionately associated with height between the sexes (height SexDiff-associated SNPs). The plot depicts -$\log_{10}$ P-values for the t-SexDiff statistic. Gray bars correspond to four different FDR cutoffs. (D) SexDiff association analyses for significant phenotype-associated SNPs (number of SNPs included in each analysis is shown to the right of each plot) for four additional anthropometric traits: body mass, hip circumference, body fat percentage, and waist circumference.

159 ***Co-localization of SexDiff-associated SNPs and loci with known roles in sexual differentiation***

160 Prior to conducting evolutionary analyses on the anthropometric trait SexDiff-associated SNPs,
161 we assessed their biological plausibility in two ways. First, we tested whether these SNPs are
162 significantly more likely to be located within or nearby (+/- 10,000 base pairs) genes previously
163 known to be involved in sexual differentiation (Gene Ontology term GO:0007548) compared to
164 SNPs that are significantly associated with the phenotype but not significantly associated with sex
165 differences [32,33]. Although not all SexDiff-associated SNPs are expected to be located nearby
166 genes already known to be involved in sexual differentiation, we would expect at least some
167 enrichment if we are identifying true SexDiff associations with our approach.

168

169 For each t-SexDiff FDR threshold, we determined the number of unique sexual differentiation
170 (GO:0007548) genes with one or more co-localized SexDiff-associated SNPs (combined across
171 the five phenotypes in our study). We similarly identified the number of all genes (of those in the
172 GO database) that were co-localized with at least one SexDiff-associated SNP and calculated the
173 proportion of the number of GO:0007548 genes to the total number of genes. We conducted the
174 same analysis for the set of SNPs associated with the phenotypes in general but *not* associated with
175 sexual differentiation. Finally, we estimated the GO:0007548 enrichment ratio for SexDiff-
176 associated to non SexDiff-associated SNPs by dividing the two proportions. By focusing this
177 analysis on counts of genes rather than counts of SNPs, we limit potential linkage disequilibrium
178 (LD)-based enrichment inflation. That is, a gene would only be counted once regardless of how
179 many SexDiff-associated SNPs are located within or nearby that gene.

180

181 For example, at our most stringent FDR threshold (0.001), the 3,016 total SexDiff-associated SNPs
182 were located within or nearby 9 unique GO:0007548 genes and 162 total genes (to be included in
183 Dryad Digital Repository deposition). Thus, the proportion of GO:0007548 genes = 0.0555
184 (9/162). In contrast, the 100,646 non-SexDiff-associated but still phenotype-associated SNPs for
185 this same FDR threshold were located within or nearby 52 unique GO:0007548 genes and 2,544
186 total genes, for a proportion of 0.0204, resulting in an enrichment ratio = 2.72 (0.0555/0.0204). In
187 other words, GO:0007548 genes with known roles in sexual differentiation are >2.7 times more
188 likely to be co-localized with one or more SexDiff-associated SNPs than with one or more non-
189 SexDiff-associated but still phenotype-associated SNPs at the FDR<0.001 analysis level.
190
191 We repeated this analysis for each FDR threshold (Fig 2A; S2 Table). The enrichment ratio steadily
192 increased with increasingly stringent FDR significance thresholds. This pattern is consistent with
193 expectations if our analyses are identifying true SexDiff-associated loci.
194
195 To test whether the proportion of GO:0007548 genes co-localized with at least one SexDiff-
196 associated SNP was significantly greater than expected by chance, we used a simple permutation
197 scheme. From the set of 2,191 total GO classified genes that were co-localized with at least one
198 phenotype-associated SNP, we randomly selected the same number of total genes that were co-
199 localized with one or more SexDiff-associated SNPs at the FDR cutoff being considered (e.g. 162
200 genes at FDR<0.001) and counted the number of GO:0007548 genes represented. We repeated
201 this process 10,000 times for each FDR cutoff and compared the resulting distributions to the
202 actual number of GO:0007548 genes co-localized with at least one SexDiff-associated SNP to
203 calculate empirical P-values. We used this permutation-based approach rather than a Fisher's exact

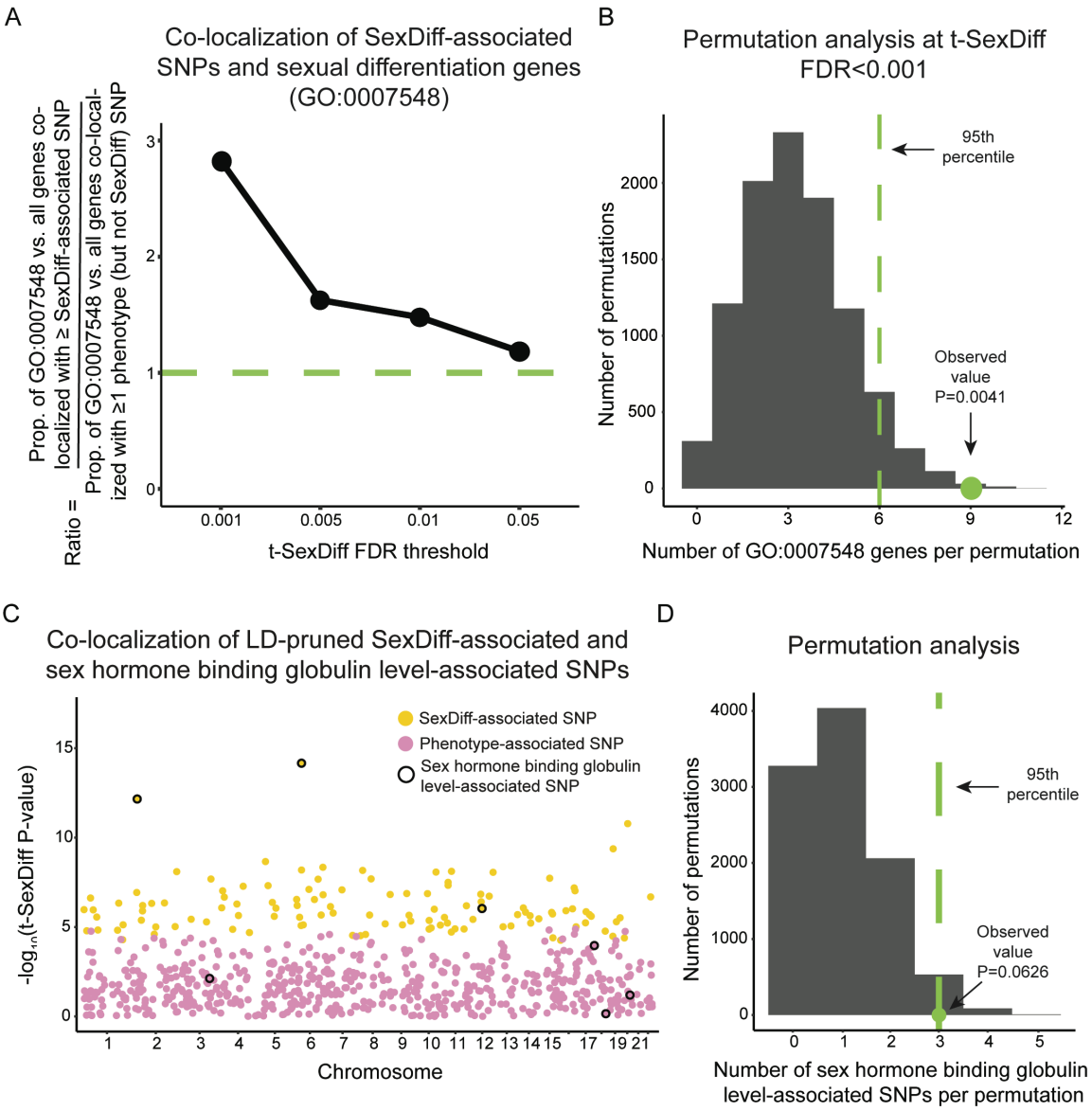

Fig 2. Co-localization of SexDiff-associated SNPs and loci with known roles in sexual differentiation. (A) For each t-SexDiff FDR threshold, we computed the proportion of the number of genes in the "Sexual Differentiation" Gene Ontology category (GO:0007548) to the number of all Gene Ontology genes with at least one co-localized SexDiff-associated SNP (+/- 10,000 base pairs). We also computed the same proportion for the set of SNPs significant associated with our studied phenotypes in general but *not* with sexual differentiation. Values shown are ratios of these two proportions at each t-SexDiff FDR threshold. The green line indicates the 1:1 ratio that would be expected in the absence of any disproportionate co-localization between SexDiff-associated SNPs and GO:00007548 genes. (B) Permutation analysis of the number of genes involved in sexual differentiation co-localized with at least one SexDiff-associated SNPs at our most stringent FDR threshold (q<0.001). From the set of 2,191 total GO-classified genes that were co-localized with at least one phenotype-associated SNP, we randomly selected 162 genes, the number of total genes that were co-localized with one or more SexDiff-associated SNPs at the FDR<0.001 cutoff. Of these 162 genes, we counted and recorded the number of GO:0007548 genes represented. We repeated this process 10,000 times and computed an empirical P-value (P=0.0041) as the proportion of permutations with a greater than or equal number of GO:0007548 genes as the observed value for FDR<0.001 SexDiff-associated SNPs (9 genes). Results for similar analyses based on SexDiff FDR thresholds 0.005, 0.01, and 0.05 are shown in S1 Fig. (C) Manhattan plot depicting the -log10 values for the t-SexDiff statistic for a curated set of LD-pruned SexDiff and phenotype-associated SNPs. Specifically, we selected a maximum of one SNP with the lowest t-SexDiff P-value across any of the five studied phenotypes from each of 1,703 approximately LD-independent blocks of the human genome. Yellow dots indicate SexDiff-associated SNPs at an FDR of 0.001 (for at least one of the five traits), and pink dots indicate phenotype-associated SNPs that did not cross the FDR=0.001 threshold for any trait. SNPs outlined in black are also significantly associated with the sex hormone binding globulin levels as annotated in the GWAS catalog. (D) From the curated set of 693 total LD-pruned SexDiff and phenotype-associated SNPs, we randomly selected 117 SNPs (the number of pruned SexDiff-associated SNPs) and counted how many were also associated with sex hormone globulin blinding levels. We repeated this process 10,000 times and computed an empirical P-value of P=0.0626.

227 test because some genes were co-localized with both SexDiff-associated SNPs and non-SexDiff
228 phenotype-associated SNPs. The observed proportion of GO:0007548 genes is very unlikely to be
229 explained by chance at the most stringent FDR<0.001 threshold (P=0.0041; Fig 2B), with
230 increasing probability for the FDR thresholds of 0.005, 0.01, and 0.05 (P=0.0513, P=0.0829, and
231 P=0.2111, respectively; S1 Fig). Accordingly, all subsequent analyses were limited to SNPs
232 classified based on the FDR<0.001 analysis.
233
234 In a second assessment of the biological plausibility of our results, we queried the GWAS Catalog
235 [34] to identify any pleiotropic co-occurrence between our SexDiff-associated (FDR<0.001) or
236 phenotype-associated SNPs and significant associations for all other GWAS Catalog traits. To
237 avoid linkage disequilibrium (LD)-based result inflation of co-occurrence, we selected a maximum
238 of one SNP (zero if no SNPs in the region were significantly associated with a trait) with the lowest
239 t-SexDiff P-value across any of the five studied traits for each of 1,703 approximately LD-
240 independent blocks of the human genome [35]. Of the 693 total SNPs in this curated dataset, 117
241 SNPs (17%) were significantly associated with sex differences for at least one of the five traits,
242 with the remaining 576 SNPs (83%) associated with one or more of the five phenotypes but not
243 sex differences.
244
245 Upon reviewing the overlap with the curated set of SexDiff-associated SNPs and the GWAS
246 Catalog trait associations, we immediately noticed "Sex hormone binding globulin level"
247 annotations [36–38] for 3 of the 117 SexDiff-associated SNPs (2.6%), including the two SNPs
248 with the most extreme t-SexDiff P-values (Fig 2C). In contrast, sex hormone binding globulin
249 level associations were annotated for 4 of the 576 phenotypes but not SexDiff-associated SNPs

250 (0.7%; Fisher's exact test; P=0.098). We also used a permutation analysis to estimate the
251 probability of observing 3 SNPs with both SexDiff- and sex hormone binding globulin level-
252 associations, given the occurrence rate for all loci in this curated SNP subset (P=0.063; Fig 2D).
253
254 Altogether, these results suggest that the anthropometric trait SexDiff-associated SNPs we
255 identified – especially those discovered with the most stringent FDR<0.001 threshold – are at least
256 enriched for SNPs that do underlie (or are linked to those that do) sexual differentiation variation.
257
258 ***SexDiff-associated SNPs are associated with greater trait effect sizes in the expected sex***
259 Our statistical approach technically identifies genetic variants with significantly different strengths
260 of association for a given trait between males and females. Before proceeding, we sought to
261 confirm that these identified variants are disproportionately associated with greater phenotype
262 effect sizes in the sex with the stronger association value, and not merely with lower trait value
263 variance.
264
265 To do so, for each phenotype we first divided our SexDiff-associated SNPs into those with
266 disproportionately stronger associations in females vs. males. To avoid linkage disequilibrium
267 (LD)-based result inflation, we pruned each of these sets of SNPs to include only the one SNP (if
268 any) with the lowest t-SexDiff FDR value located within each of 1,703 approximately LD-
269 independent blocks of the human genome [35]. For comparison to the disproportionate female and
270 male SexDiff-associated SNPs, we prepared similar LD-pruned sets of all SNPs significantly
271 associated with the phenotype, regardless of their associations with sex differences, for each of the
272 five traits (to be included in Dryad Digital Repository deposition). Then, for every SNP in each

273 set of pruned SNPs, we calculated the $\log_2$ ratio of the female trait effect size to the male trait effect
274 size.
275
276 Each of the five sets of female SexDiff-associated SNPs had strongly positive average ratios,
277 meaning that the female effect sizes for these SNPs were larger than the male effect sizes, and *vice*
278 *versa* for the male SexDiff-associated SNPs (Fig 3A). Mean $\log_2$ ratios were significantly different
279 from 0 for each of the 10 SexDiff-associated SNP subset (one-sided t-tests; $P=3.3x10^{-5}$ and lower;
280 $FDR=3.3x10^{-5}$ and lower; S3 and S4 Tables). Conversely, the mean $\log_2$ ratios for the five sets of
281 general trait-associated SNPs were near 0 (Fig 3A). Mean $\log_2$ ratios for each of the 10 SexDiff-
282 associated SNPs were also significantly different from the corresponding mean $\log_2$ ratios of the
283 phenotype-associated SNPs (Permutation analyses [see methods]; $P<0.001$, $FDR=0.001$). These
284 results confirm that our statistical process successfully and reliably identifies SNPs with sex
285 disproportionate effects on phenotypic trait variation. We thus proceeded to study the evolutionary
286 histories of these SNPs.
287
288 ***Signatures of positive selection on SexDiff-associated alleles***
289 We next tested whether the SexDiff-associated SNPs identified in our analyses are significantly
290 enriched for signatures of recent (~3,000 years) positive selection, using the Singleton Density
291 Score statistic (SDS) [20]. Briefly, alleles affected by recent positive selection are predicted to be
292 found on haplotypes with relatively fewer singleton mutations; the SDS quantifies this pattern. In
293 turn, the trait-SDS statistic reflects directionality with respect to an associated phenotype. A
294 positive trait-SDS value reflects a recent increase in the frequency of a phenotype-increasing allele,

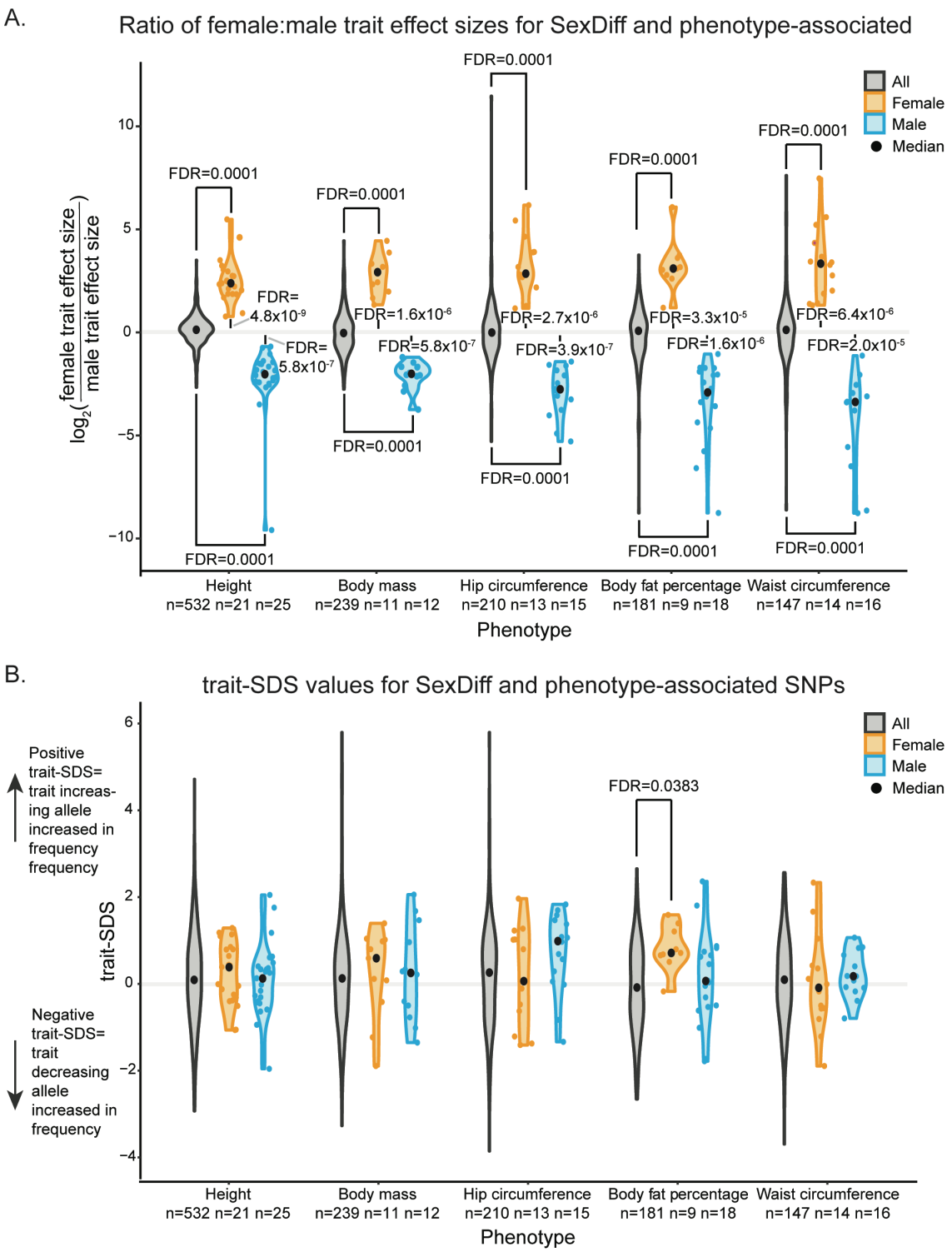

Fig 3. Sex-specific effect size ratios and trait-SDS scores for anthropometric trait-associated SNPs. For each anthropometric trait, a maximum of one SNP per each of 1,703 approximately LD-independent blocks of the human genome was included (within each block, the SNP with the strongest statistical significance) from the sets of i) all SNPs associated with trait variation, and the subsets of those SNPs disproportionately associated with variation in ii) females and iii) males (SexDiff-associated SNPs at FDR<0.001). (A) Violin plots of $\log_2$ ratios of female trait effect size to male effect size (calculated separately for each SNP). Mean $\log_2$ ratios for each set of female and male SexDiff pruned SNPs were compared to those for the corresponding phenotype-association set using a permutation scheme (FDR values indicated at the top and bottom of the plot), and to 0 using a one-sided t-test (FDR values indicated in the middle of the plot). (B) Violin plots of trait-SDS values. Positive trait-SDS values reflects recent increases in the frequencies of phenotype-increasing alleles, while negative trait-SDS values reflect increases in the frequencies of phenotype-decreasing alleles. The trait-SDS distributions for each set of female and male SexDiff pruned SNPs were compared to those for the corresponding phenotype-association set using a permutation analysis

307 while a negative trait-SDS value reflects an increase in the frequency of a phenotype-decreasing
308 allele [20]. We performed our analyses with previously-published SDS data computed from the
309 whole genome sequences of 3,195 British individuals from the UK10K project [20,39]. This
310 dataset is ideal for integration with the UK BioBank genotype-phenotype association data, given
311 study population similarity.
312
313 To examine the evolutionary histories of alleles associated with anthropometric trait sex
314 differences, we aimed to determine whether any signatures of positive selection for SexDiff-
315 associated SNPs differed in either magnitude or direction from any such signatures for alleles
316 associated with the phenotypes (but not necessarily sex differences) themselves. For this analysis,
317 we again considered the sets of LD-pruned SexDiff-associated SNPs divided into those that are
318 disproportionately associated with trait variation in females vs. those disproportionately associated
319 with trait variation in males for each phenotype.
320
321 Using a permutation analysis, we tested whether the average trait-SDS values for the pruned
322 female or male SexDiff-associated SNPs were significantly different than the value for the
323 corresponding pruned set of all phenotype-associated SNPs (Fig 3B; S5 Table). For example, there
324 were n=21 pruned SexDiff-associated SNPs disproportionately associated with female height and
325 n=532 pruned SNPs associated with height generally (but not necessarily with sex differences).
326 We randomly selected 21 of the 532 pruned height-associated SNPs and calculated the average
327 trait-SDS score. We repeated this process 10,000 times and calculated an empirical P-value based
328 on the proportion (multiplied by two; see Methods) of permuted observations with equal or more
329 extreme average trait-SDS values than the observed average trait-SDS for the actual female height

330 SexDiff-associated SNPs (average trait-SDS=0.240; P=0.66). This test was then repeated for the
331 male SexDiff-associated SNPs for height, and the female and male SexDiff-associated SNPs for
332 each of the remaining four anthropometric traits. We also performed a similar permutation but
333 with the additional step of matching minor allele frequencies (+/- 0.05) between the original set of
334 SexDiff-associated SNPs and the set of phenotype-associated SNPs selected in each permutation.
335 The pattern of results was unchanged relative to the above (S6 Table).
336
337 For nine of the comparisons, trait-SDS distributions for the sets of pruned SexDiff-associated
338 SNPs were not significantly different (following FDR adjustment for ten tests) from those for
339 SNPs associated with the corresponding general phenotype (Fig 3B; S5 Table). However,
340 compared to all body fat percentage-associated SNPs (n=181 pruned SNPs; average trait-SDS=-
341 0.124), the average trait-SDS value for the set of female SexDiff-associated SNPs for this
342 phenotype was significantly elevated (n=9 pruned SNPs; average trait-SDS=0.827; P=0.0038;
343 FDR=0.038). In other words, the average frequencies of alleles associated with greater body fat
344 percentage in females have increased in frequency over the past ~3,000 years at a faster rate than
345 expected based on the pattern for SNPs associated with body fat percentage generally, a potential
346 signature of polygenic selection on this sub-trait.
347
348 While our hypothesis testing framework is focused primarily on analyses of recent signatures of
349 positive natural selection, we also investigated whether the sets of LD-pruned SexDiff-associated
350 SNPs were enriched for signatures of relatively more ancient positive selection. To do so we
351 conducted an identical permutation analysis as above with the trait-SDS values, but instead using
352 absolute values of the integrated haplotype score (iHS) statistic, a within-population haplotype-

353 based statistic that can be used to identify putative signatures of positive selection from across the
354 past ~25,000 years [40]. We used previously-published iHS data [41] computed for the Great
355 Britain population (GBR) from the 1000 Genomes Project [42]. Absolute iHS values were not
356 significantly different for any of the sets of SexDiff-associated SNPs relative to the corresponding
357 phenotype-associated SNPs (S2 Fig, S7 Table).
358
359 Finally, because positive selection may occur more often within or nearby genic regions of the
360 genome [40], we also tested whether our sets of pruned SexDiff-associated SNPs and
361 corresponding phenotype-associated SNPs were unevenly distributed among genic (annotated
362 genes +/- 10,000 bp) and intergenic regions. None of the ten sets of SexDiff-associated SNPs were
363 enriched for genic regions. However, we found that SNPs significantly associated with male hip
364 circumference variation were more likely to be intergenic relative to general hip circumference-
365 associated loci (S8 Table).
366

367 **Discussion**

368 Using a sex-stratified GWAS framework for five sexually differentiated anthropometric
369 phenotypes, we identified 3,016 SNPs that were disproportionately associated with either female
370 or male trait variation at a low false discovery rate (FDR<0.001). We confirmed the biological
371 plausibility of these results by showing that genes with known roles in sexual differentiation are
372 significantly enriched for SexDiff-associated SNPs. Together, these results confirm the importance
373 of considering sex differences when investigating the genetic structure of human polygenic traits
374 [43]. We then used a statistic that quantifies changes in the frequencies of alleles underlying

375 polygenic traits over the past ~3,000 years to identify a signature of recent positive selection on
376 SNPs associated with increased female body fat percentage in the British study population.
377
378 We must emphasize that inferring selection signals from GWAS data should be approached with
379 great care, as even subtle uncorrected population structure can impact GWAS and downstream
380 results [44]. For example, data from the GIANT consortium were previously used to identify
381 strong signatures of polygenic selection for height across the genome [20]. However, subtle
382 population structure in the GIANT sample led to effect-size estimate biases, in turn resulting in
383 false signals of polygenic selection for SNPs not crossing the genome-wide significance threshold
384 and impacting results for significant SNPs as well [44]. In contrast, these issues were much less
385 prevalent using GWAS summary statistics from the UK Biobank, in which population structure is
386 minimized [44–46]. In light of these considerations, in our study we have i) used UK Biobank
387 GWAS summary statistics only, ii) focused solely on phenotype-associated SNPs below the
388 genome-wide significance threshold, and iii) restricted our evolutionary analyses to direct
389 comparisons between SNPs significantly associated with individual phenotypes and a sub-
390 phenotype (i.e., sex differences).
391
392 Our study further demonstrates the value of GWAS-based approaches for testing anthropological
393 hypotheses [47]. Concerning the evolution of human body size and shape phenotypes, our results
394 fail to provide support for the prevailing notion of recent (i.e., subsequent to agriculture) adaptive
395 reductions in levels of sex differences for such traits. Specifically, using large samples of genomes
396 from British individuals we did not observe significant differences in the recent evolutionary
397 trajectories of SNPs disproportionately associated with female or male variation in height, body

398 mass, hip circumference, and waist circumference relative to the trajectories of SNPs associated
399 with these traits generally.
400
401 We note that we made a number of conservative choices (for example, with aggressive pruning to
402 account for linkage disequilibrium) in our analytical approach, meaning that our failure to reject
403 the null hypothesis for each of these four traits should not be interpreted as evidence that no
404 selection on them occurred. Still, even with our conservative analytical approach we did find
405 evidence that the average frequencies of alleles disproportionately associated with greater female
406 body fat percentage significantly increased over the past 3,000 years, a pattern consistent with
407 polygenic adaptation. Given that females have higher average body fat percentages than men in
408 historic and contemporary populations, the direction of polygenic adaptation in the population we
409 studied would actually be opposite to expectations under hypotheses of recent adaptive reductions
410 in anthropometric trait sex differences in agricultural societies. However, since SNPs can be
411 pleiotropically associated with multiple phenotypes (32), we cannot definitively conclude that
412 positive selection acted directly on female body fat percentage. Regardless, at the very least we
413 did not find positive support for the prevailing hypothesis concerning the evolution of sex
414 differences in recent human evolution.
415
416 **Methods**
417
418 ***Subjects and dataset generation***
419 We used genome-wide association study (GWAS) summary statistics generated from analyses of
420 UK Biobank data [18,22]. The original GWAS analyses were restricted to 361,194 unrelated

421 individuals (194,174 females and 167,020 males) of white British ancestry (based on the
422 combination of self-report and genetic ancestry analysis) who did not have sex chromosome
423 aneuploidies. GWAS summary statistics for each phenotype were computed separately for females
424 and males. Because every phenotype was not defined for all individuals, some of the analyses
425 contributing to our study included fewer than 361,194 individuals (S9 Table). Approximately 40
426 million SNPs were originally available for analysis, of which only those with minor allele
427 frequencies > 0.001, an INFO score (imputation quality) > 0.8, and a Hardy-Weinberg Equilibrium
428 P-value > $1x10^{-10}$ were retained, resulting in datasets of 13,791,468 SNPs.
429
430 We then subsequently filtered out SNPs with minor allele frequencies < 0.05 and those that were
431 not associated with a singleton density score (SDS; for description, see below) [20]. These filters
432 resulted in a total genome-wide dataset of ~4.4 million SNPs for each phenotype.
433
434 We chose five sexually differentiated anthropometric phenotypes from the GWAS summary
435 statistics for analysis (S9 Table). To identify phenotype-associated SNPs in female and/or male
436 individuals, we applied the commonly used genome wide significance threshold (i.e., to account
437 for the large number of tested SNPs) of $P=5x10^{-8}$ to the male-specific and female-specific GWAS
438 summary statistics (to be included in Dryad Digital Repository deposition).
439
440 ***Scan for SNPs that are disproportionately associated with male or female trait variation***
441 For each genome-wide SNP significantly associated with a phenotype in females, males, or both
442 sexes, we evaluated whether there was a significant difference in the statistical strengths of

443 association for the female vs. male-specific GWAS results using the following t-statistic (t-
444 SexDiff) [19].
445

446 $$t - SexDiff = \frac{b_{male} - b_{female}}{\sqrt{SE_{male}^2 + SE_{female}^2 - 2r * SE_{male} * SE_{female}}}$$

447

448 b_{male} refers to the male-specific beta value and b_{female} refers to the female-specific beta value for
449 each genome-wide SNP. SE_{male} and SE_{female} refer to the male-specific and female-specific standard
450 error for each genome-wide SNP, respectively. The correlation *r* (rho) for each phenotype was
451 calculated as the Spearman rank correlation coefficient between b_{male} and b_{female} using the cor.test()
452 function in R (version 3.5.1). t-SexDiff was converted to a two-sided P-value using the R function
453 pt(). The effects of multiple testing were considered by computing the False Discovery Rate (FDR)
454 for each t-SexDiff P-value using the R function p.adjust [48].
455

456 Of the phenotype-associated SNPs that were significantly associated with a given phenotype for
457 females, males, or both sexes, we identified SexDiff-associated SNPs at four different FDR
458 thresholds: <0.05, <0.01, <0.005, and <0.001 (to be included in Dryad Digital Repository
459 deposition).
460

461 ***Assessing the biological plausibility of the SexDiff-associated SNPs***

462 In order to confirm the biological plausibility of our SexDiff-associated SNPs, we tested whether
463 regions of the genome functionally linked to sexual differentiation are more likely than expected
464 by chance to contain one or more of our SexDiff-associated SNPs. Separately for each t-SexDiff

465 FDR threshold, we counted the number of unique genes in the Gene Ontology database [32,33]
466 that contained at least one SexDiff-associated SNP, including within a +/-10,000 base-pair (bp)
467 window around the gene to encompass potential regulatory regions. We then counted the number
468 of these genes with known links to processes of sexual differentiation corresponding to the Gene
469 Ontology (GO) term GO:0007548 and computed the proportion of the number of GO:0007548
470 genes to the number of all genes co-localized with >= 1 SexDiff-associated SNP. We repeated this
471 analysis for significant phenotype- but not SexDiff-associated SNPs. In the absence of enrichment
472 for SexDiff-associated SNPs, the ratio of these two proportions is expected to equal one.
473
474 We then used the following permutation scheme for each t-SexDiff FDR cutoff. There was a total
475 of 2,191 GO-classified genes overlapping one or more phenotype-associated SNP (whether
476 SexDiff-associated or not; each t-SexDiff FDR cutoff started with the same genome-wide
477 significant set of phenotype-associated SNPs so the total number of 2,191 co-localized genes
478 applies to each FDR cutoff). Given the number of these genes overlapping >=1 SexDiff-associated
479 SNP for a given FDR cutoff, we randomly selected that number of unique genes from the pool of
480 2,191 genes and counted the number of GO:0007548 genes. We then repeated this procedure
481 10,000 times and computed an empirical P-value as the proportion of permuted data sets with an
482 equal or greater to number of sexual differentiation genes when compared to our observation for
483 the associated FDR threshold.
484
485 We studied pleiotropic relationships between our anthropometric SexDiff and phenotype traits and
486 other traits with the aid of the GWAS Catalog (48; accessed 18 Feb 2021). After concatenating all
487 SexDiff (FDR<0.001) and phenotype-associated SNPs from across the five studied traits, we

488 pruned this total set of SNPs to a maximum of one SNP per each of approximately 1,703
489 approximately LD-independent blocks of the human genome [35]. For each block, we specifically
490 chose the one SNP with the lowest t-SexDiff P-value (if a block did not have any significant
491 phenotype-associated SNPs for any trait, zero SNPs were included from that block). There were
492 693 total SNPs in this curated dataset, 117 of which (17%) were significantly associated with sex
493 differences for at least one of the 5 traits, with 576 (83%) significantly associated with one or more
494 of the five phenotypes but not sex differences. After reviewing the full results (data to be included
495 in the Dryad Digital Repository Deposition), we used a permutation analysis to estimate the
496 probability that SexDiff-associated SNPs were more likely than expected by chance to also be
497 associated with the "sex hormone binding globulin level" trait. Specifically, from our curated set
498 of 693 SNPs, we randomly selected the number of SNPs equal to the number of observed SexDiff-
499 associated SNPs (117) and counted how many were also associated with sex hormone binding
500 globulin level as per the GWAS catalog. We repeated this procedure 10,000 times and computed
501 the proportion of permuted data sets with equal or more extreme numbers of sex hormone binding
502 globulin level-associated SNPs compared to the result for the SexDiff-associated SNPs.
503
504 ***Assessing the sex-specific effects of SexDiff-associated SNPs***
505 For each phenotype we split our SexDiff-associated SNPs into those that had lower P-values (as
506 identified in the original sex specific GWAS for phenotype-associated SNPs) in females than
507 males (female SexDiff-associated SNPs) and those that had lower P-values in males than females
508 (male SexDiff-associated SNPs). We separately pruned each set of female SexDiff-associated
509 SNPs and male SexDiff-associated SNPs to account for linkage disequilibrium. Specifically, if
510 there was more than one female SexDiff-associated SNP in one of 1,703 approximately LD-

511 independent segments of the genome [35], we only kept the female SexDiff-associated SNP for
512 that segment with the lowest P-value for association with the phenotype in females. We did the
513 same for male SexDiff-associated SNPs in males. For each set of pruned SexDiff-associated SNPs,
514 we then calculated the $\log_2$ ratio of the female effect size to male effect size.
515
516 The set of phenotype-associated SNPs was pruned in a similar fashion to the above, with a
517 maximum of one SNP per each of the 1,703 approximately LD-independent genome segments,
518 chosen as the SNP with the most significant P-value in the segment regardless of whether it was
519 most significant in females or males (to be included in Dryad Digital Repository deposition) from
520 among those below the genome-wide significance threshold.
521
522 We then used a permutation to estimate the probability that the $\log_2$ ratio of female effect size to
523 male effect size was significantly for each of the ten sets of SexDiff-associated SNPs. From the
524 pruned set of phenotype-associated SNP, we randomly selected a number of SNPs equal to the
525 number of observed SexDiff-associated SNPs for that phenotype. We repeated this procedure
526 10,000 times and computed the proportion of permuted data sets with equal or more extreme
527 average $\log_2$ ratio. We computed FDR values to account for multiple testing.
528
529 Finally, we used a one-sided t-test to determine whether our $\log_2$ ratios for each set of SexDiff-
530 associated SNPs was significantly different from zero. We again computed FDR values to account
531 for multiple testing
532
533 ***Identification of signatures of positive selection***

534 To test the hypothesis that SexDiff-associated SNPs have been affected by recent (past ~3,000
535 years) positive selection in recent human evolution, we used the Singleton Density Score (SDS)
536 statistic [20]. We used a database of genome-wide SNP SDS scores that were computed [20] using
537 3,195 whole genome sequences from British individuals the UK10K project [39]. For the pruned
538 sets of female and male SexDiff-associated SNPs, we fixed the sign of SDS scores so that positive
539 values indicate an increased frequency of the trait-increasing allele.
540
541 We used permutations to estimate the probability that the average trait-SDS value for each trait
542 and sex could be observed by chance given the distribution trait-SDS scores for SNPs significantly
543 associated with the corresponding phenotype (regardless of SexDiff-association; S10 Table). From
544 the pruned set of phenotype-associated SNPs, we then randomly selected a number of SNPs equal
545 to the number of observed SexDiff-associated SNPs for that phenotype and the sex and FDR
546 threshold being considered, and we calculated the average trait-SDS score for that set of SNPs.
547 We repeated this procedure 10,000 times and computed the proportion of permuted data sets with
548 equal or more extreme average trait-SDS scores compared to actual result for the observed
549 SexDiff-associated SNPs. This proportion was then multiplied by two to account for the two-tailed
550 nature of this test (i.e., the average trait-SDS values for SexDiff-associated SNPs could have been
551 significantly greater than or less than that for the phenotype-associated SNPs). We computed FDR
552 values to account for multiple testing.
553
554 We also performed trait-SDS permutation analysis with minor allele frequency (MAF) matching
555 between SexDiff and phenotype-associated SNPs. Specifically, for each SexDiff-associated SNP
556 we identified all phenotype-associated SNPs with +/- 0.05 MAF. We then randomly selected one

557 of these SNPs for inclusion in the permuted dataset and removed it from our list of phenotype-
558 associated SNPs to be drawn from in the MAF matching to other SexDiff-associated SNPs so that
559 it wouldn't be represented twice in the same permutation. We repeated this procedure 10,000
560 times, randomizing the input order of SexDiff-associated SNPs each time.
561
562 For our iHS-based analysis, we used a database of genome-wide SNP iHS scores that were
563 previously computed [41] using data from the 1000 Genomes project [42]. We specifically used
564 iHS data from the Great Britain (GBR) population due to population similarity with our GWAS
565 data. We considered the absolute value of the standardized iHS for each SNP, and we removed
566 SNPs from our dataset that did not have a corresponding iHS score. To estimate the probability
567 that the average iHS score for each trait and sex could be observed by chance given the distribution
568 of iHS scores for phenotype-associated SNPs, we used a permutation scheme identical to that used
569 for the primary trait-SDS analysis (the version without consideration of minor allele frequencies).
570
571 To test whether there were significant differences in the distributions of pruned SexDiff-associated
572 SNPs and corresponding pruned phenotype-associated SNPs, we determined whether or not each
573 SNP was located within +/- 10,000 bp of a gene annotated in the Gene Ontology database. We
574 compared the ratio of genic:intergenic SNPs for each of our ten sets of SexDiff-associated and
575 corresponding phenotype -associated SNPs for each trait using a permutation analysis (S8 Table).
576 For each set of SexDiff-associated SNPs, we randomly selected the same number of SNPs from
577 the corresponding set of phenotype-associated SNPs and counted the number of SNPs in intergenic
578 regions. We then repeated this procedure 10,000 times and computed an empirical P-value as the
579 proportion of permuted data sets with an equal or greater to number of SNPs in intergenic regions

when compared to our observation for the associated set of SexDiff-associated SNPs. This proportion was then multiplied by two to account for the two-tailed test (i.e., the number of intergenic SexDiff-associated SNPs could have been significantly greater than or less than that of phenotype-associated SNPs). We computed FDR values to account for multiple testing.

Computational resources

All analyses were conducted in R (version 3.5.1) with Advanced CyberInfrastructure computational resources provided by The Institute for CyberScience at Pennsylvania State University. All scripts are available at https://github.com/audreyarner/dimorphism-evolution.

Data availability

Data from all steps of our analyses will be made available on Dryad prior to publication of this manuscript.

Acknowledgments

We thank the DFG Center for Advanced Studies "Words, Bones, Genes, Tools" for their support and S. Marciniak, C. Bergey, J. Tung, and D. Puts and the Sex Differences Interest Group for their helpful discussions. This research was conducted with Advanced CyberInfrastructure computational resources provided by The Institute for CyberScience at Penn State.

601 **Bibliography**

602 1. Fairbairn DJ. Allometry for sexual size dimorphism: pattern and process in the
603 coevolution of body size in males and females. Annu Rev Ecol Syst. 1997 Nov;28(1):659–
604 87.
605 2. Khramtsova EA, Davis LK, Stranger BE. The role of sex in the genomics of human complex
606 traits. Nat Rev Genet. 2019 Mar 23;20(3):173–90.
607 3. Sharma E, Künstner A, Fraser BA, Zipprich G, Kottler VA, Henz SR, et al. Transcriptome
608 assemblies for studying sex-biased gene expression in the guppy, *Poecilia reticulata*. BMC
609 Genomics. 2014 May 26;15(1):400.
610 4. Setchell JM, Jean Wickings E. Dominance, status signals and coloration in male mandrills
611 *(Mandrillus sphinx)*. Ethology. 2005 Jan 1;111(1):25–50.
612 5. Loyau A, Jalme M Saint, Sorci G. Intra- and intersexual selection for multiple traits in the
613 peacock *(Pavo cristatus)*. Ethology. 2005 Sep 1;111(9):810–20.
614 6. Plavcan JM. Sexual dimorphism in primate evolution. Am J Phys Anthropol.
615 2001;116(S33):25–53.
616 7. Lindenfors P, Tullberg BS. Phylogenetic analyses of primate size evolution: the
617 consequences of sexual selection. Biol J Linn Soc. 1998 Aug;64(4):413–47.
618 8. Morris JS, Cunningham CB, Carrier DR. Sexual dimorphism in postcranial skeletal shape
619 suggests male-biased specialization for physical competition in anthropoid primates. J
620 Morphol. 2019 May 20;280(5):731–8.
621 9. Sidorenko J, Kassam I, Kemper KE, Zeng J, Lloyd-Jones LR, Montgomery GW, et al. The
622 effect of X-linked dosage compensation on complex trait variation. Nat Commun. 2019

623 Dec 8;10(1):3009.

624 10. Ruff C. Variation in human body size and shape. Annu Rev Anthropol. 2002
625 Oct;31(1):211–32.

626 11. Stulp G, Barrett L. Evolutionary perspectives on human height variation. Biol Rev. 2016
627 Feb;91(1):206–34.

628 12. McCombe PA, Greer JM, Mackay IR. Sexual dimorphism in autoimmune disease. Curr Mol
629 Med. 2009;9(9):1058–79.

630 13. Natri H, Garcia AR, Buetow KH, Trumble BC, Wilson MA. The pregnancy pickle: Evolved
631 immune compensation due to pregnancy underlies sex differences in human diseases.
632 Trends Genet. 2019;35(7):478–88.

633 14. Dunsworth HM. Expanding the evolutionary explanations for sex differences in the
634 human skeleton. Evol Anthropol Issues, News, Rev. 2020 May 2;29(3):108–16.

635 15. Touraille P. Human Sex Differences in Height: Evolution due to Gender Hierarchy? In: M.
636 Ah-King, editor. Challenging Popular Myths of Sex, Gender and Biology. Springer, Cham;
637 2013. p. 65–75.

638 16. Frayer DW. Sexual dimorphism and cultural evolution in the Late Pleistocene and
639 Holocene of Europe. J Hum Evol. 1980 Jul;9(5):399–415.

640 17. Rogers AR, Mukherjee A. Quantitative genetics of sexual dimorphism in human body size.
641 Evolution. 1992 Feb;46(1):226–34.

642 18. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An open
643 access resource for identifying the causes of a wide range of complex diseases of middle
644 and old age. PLoS Med. 2015;12(3):e1001779.

645 19. Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, et al. Sex-stratified
646 genome-wide association studies including 270,000 individuals show sexual dimorphism
647 in genetic loci for anthropometric traits. Gibson G, editor. PLoS Genet. 2013 Jun
648 6;9(6):e1003500.
649 20. Field Y, Boyle EA, Telis N, Gao Z, Gaulton KJ, Golan D, et al. Detection of human
650 adaptation during the past 2000 years. Science. 2016 Nov 11;354(6313):760–4.
651 21. Woodbridge J, Fyfe RM, Roberts N, Downey S, Edinborough K, Shennan S. The impact of
652 the Neolithic agricultural transition in Britain: a comparison of pollen-based land-cover
653 and archaeological 14C date-inferred population change. J Archaeol Sci. 2014 Nov
654 1;51:216–24.
655 22. http://www.nealelab.is/uk-biobank/ [Internet]. [cited 2020 Apr 5]. Available from:
656 http://www.nealelab.is/uk-biobank/
657 23. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of
658 common variation in the genomic and biological architecture of adult human height. Nat
659 Genet. 2014 Nov 5;46(11):1173–86.
660 24. Gray JP, Wolfe LD. Height and sexual dimorphism of stature among human societies. Am
661 J Phys Anthropol. 1980;53(3):441–56.
662 25. Rawlik K, Canela-Xandri O, Tenesa A. Evidence for sex-specific genetic architectures
663 across a spectrum of human complex traits. Genome Biol. 2016 Dec 29;17(1):166.
664 26. Gómez-Ambrosi J, Silva C, Galofré JC, Escalada J, Santos S, Millán D, et al. Body mass
665 index classification misses subjects with increased cardiometabolic risk factors related to
666 elevated adiposity. Int J Obes. 2012 Feb 17;36(2):286–94.

667 27. Nevill AM, Stewart AD, Olds T, Holder R. Relationship between adiposity and body size
668 reveals limitations of BMI. Am J Phys Anthropol. 2006 Jan;129(1):151–6.
669 28. Humphreys S. The unethical use of BMI in contemporary general practice. Br J Gen Pract.
670 2010 Sep;60(578):696–7.
671 29. de Kovel CGF, Francks C. The molecular genetics of hand preference revisited. Sci Rep.
672 2019 Dec 12;9(1):5986.
673 30. Loh PR, Genovese G, Handsaker RE, Finucane HK, Reshef YA, Palamara PF, et al. Insights
674 into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. Nature.
675 2018;559(7714):350–5.
676 31. Rask-Andersen M, Karlsson T, Ek WE, Johansson Å. Genome-wide association study of
677 body fat distribution identifies adiposity loci and sex-specific genetic effects. Nat
678 Commun. 2019;10(1):339.
679 32. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology:
680 tool for the unification of biology. Nat Genet. 2000 May;25(1):25–9.
681 33. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing
682 strong. Nucleic Acids Res. 2019 Jan 8;47(D1):D330–8.
683 34. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The
684 NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays
685 and summary statistics 2019. Nucleic Acids Res. 2019;47(D1):D1005–12.
686 35. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of
687 shared genetic influences on 42 human traits. Nat Genet. 2016 Jul 16;48(7):709–17.
688 36. Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, et al. Using human

689 genetics to understand the disease impacts of testosterone in men and women. Nat
690 Med. 2020 Feb 10;26(2):252–8.
691 37. Prescott J, Thompson DJ, Kraft P, Chanock SJ, Audley T, Brown J, et al. Genome-Wide
692 Association Study of Circulating Estradiol, Testosterone, and Sex Hormone-Binding
693 Globulin in Postmenopausal Women. Gorlova OY, editor. PLoS One. 2012 Jun
694 4;7(6):e37815.
695 38. Coviello AD, Haring R, Wellons M, Vaidya D, Lehtimäki T, Keildson S, et al. A Genome-
696 Wide Association Meta-Analysis of Circulating Sex Hormone–Binding Globulin Reveals
697 Multiple Loci Implicated in Sex Steroid Hormone Regulation. Gibson G, editor. PLoS
698 Genet. 2012 Jul 19;8(7):e1002805.
699 39. Walter K, Min JL, Huang J, Crooks L, Memari Y, McCarthy S, et al. The UK10K project
700 identifies rare variants in health and disease. Nature. 2015 Oct 14;526(7571):82–90.
701 40. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A Map of Recent Positive Selection in the
702 Human Genome. Hurst L, editor. PLoS Biol. 2006 Mar 7;4(3):e72.
703 41. Johnson KE, Voight BF. Patterns of shared signatures of recent positive selection across
704 human populations. Nat Ecol Evol. 2018 Apr 19;2(4):713–20.
705 42. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, et al. A global
706 reference for human genetic variation. Vol. 526, Nature. Nature Publishing Group; 2015.
707 p. 68–74.
708 43. Clayton JA. Applying the new SABV (sex as a biological variable) policy to research and
709 clinical care. Physiol Behav. 2018 Apr 1;187:2–5.
710 44. Berg JJ, Harpak A, Sinnott-Armstrong N, Joergensen AM, Mostafavi H, Field Y, et al.

711 Reduced signal for polygenic adaptation of height in UK Biobank. Elife. 2019 Mar
712 21;8:e39725.
713 45. Barton N, Hermisson J, Nordborg M. Why structure matters. Elife. 2019 Mar
714 21;8:e45380.
715 46. Sohail M, Maier RM, Ganna A, Bloemendal A, Martin AR, Turchin MC, et al. Polygenic
716 adaptation on height is overestimated due to uncorrected stratification in genome-wide
717 association studies. Elife. 2019 Mar 21;8:e39702.
718 47. Vukelic A, Cohen JA, Sullivan AP, Perry GH. Extending Genome-wide Association Study
719 Results to Test Classic Anthropological Hypotheses: Human Third Molar Agenesis and the
720 "Probable Mutation Effect." Hum Biol. 2017;89(2):157.
721 48. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful
722 approach to multiple testing. J R Stat Soc Ser B. 1995 Jan;57(1):289–300.
723

724 **Figure Captions**

725 **Fig 1. SexDiff-associated SNPs for five anthropometric phenotypes.** (A) Manhattan plot
726 depicting -$\log_{10}$ P-values for the association of each genome-wide SNP with female height. The
727 black line corresponds to the genome-wide significance threshold of $P=5x10^{-8}$. (B) Manhattan plot
728 for SNP associations with male height. (C) For each of the 67,738 SNPs significantly associated
729 female and/or male height, we used the equation shown to test whether the SNP was
730 disproportionately associated with height between the sexes (height SexDiff-associated SNPs).
731 The plot depicts -$\log_{10}$ P-values for the t-SexDiff statistic. Gray bars correspond to four different
732 FDR cutoffs. (D) SexDiff association analyses for significant phenotype-associated SNPs (number
733 of SNPs included in each analysis is shown to the right of each plot) for four additional
734 anthropometric traits: body mass, hip circumference, body fat percentage, and waist
735 circumference.

736

737 **Fig 2. Co-localization of SexDiff-associated SNPs and loci with known roles in sexual**
738 **differentiation.** (A) For each t-SexDiff FDR threshold, we computed the proportion of the number
739 of genes in the "Sexual Differentiation" Gene Ontology category (GO:0007548) to the number of
740 all Gene Ontology genes with at least one co-localized SexDiff-associated SNP (+/- 10,000 base
741 pairs). We also computed the same proportion for the set of SNPs significant associated with our
742 studied phenotypes in general but *not* with sexual differentiation. Values shown are ratios of these
743 two proportions at each t-SexDiff FDR threshold. The green line indicates the 1:1 ratio that would
744 be expected in the absence of any disproportionate co-localization between SexDiff-associated
745 SNPs and GO:00007548 genes. (B) Permutation analysis of the number of genes involved in
746 sexual differentiation co-localized with at least one SexDiff-associated SNPs at our most stringent

747 FDR threshold (q<0.001). From the set of 2,191 total GO-classified genes that were co-localized
748 with at least one phenotype-associated SNP, we randomly selected 162 genes, the number of total
749 genes that were co-localized with one or more SexDiff-associated SNPs at the FDR<0.001 cutoff.
750 Of these 162 genes, we counted and recorded the number of GO:0007548 genes represented. We
751 repeated this process 10,000 times and computed an empirical P-value (P=0.0041) as the
752 proportion of permutations with a greater than or equal number of GO:0007548 genes as the
753 observed value for FDR<0.001 SexDiff-associated SNPs (9 genes). Results for similar analyses
754 based on SexDiff FDR thresholds 0.005, 0.01, and 0.05 are shown in S1 Fig. (C) Manhattan plot
755 depicting the -log10 values for the t-SexDiff statistic for a curated set of LD-pruned SexDiff and
756 phenotype-associated SNPs. Specifically, we selected a maximum of one SNP with the lowest t-
757 SexDiff P-value across any of the five studied phenotypes from each of 1,703 approximately LD-
758 independent blocks of the human genome. Yellow dots indicate SexDiff-associated SNPs at an
759 FDR of 0.001 (for at least one of the five traits), and pink dots indicate phenotype-associated SNPs
760 that did not cross the FDR=0.001 threshold for any trait. SNPs outlined in black are also
761 significantly associated with the sex hormone binding globulin levels as annotated in the GWAS
762 catalog. (D) From the curated set of 693 total LD-pruned SexDiff and phenotype-associated SNPs,
763 we randomly selected 117 SNPs (the number of pruned SexDiff-associated SNPs) and counted
764 how many were also associated with sex hormone globulin blinding levels. We repeated this
765 process 10,000 times and computed an empirical P-value of P=0.0626.
766
767 **Fig 3. Sex-specific effect size ratios and trait-SDS scores for anthropometric trait-associated**
768 **SNPs.** For each anthropometric trait, a maximum of one SNP per each of 1,703 approximately
769 LD-independent blocks of the human genome was included (within each block, the SNP with the

770 strongest statistical significance) from the sets of i) all SNPs associated with trait variation, and
771 the subsets of those SNPs disproportionately associated with variation in ii) females and iii) males
772 (SexDiff-associated SNPs at FDR<0.001). (A) Violin plots of $\log_2$ ratios of female trait effect size
773 to male effect size (calculated separately for each SNP). Mean $\log_2$ ratios for each set of female
774 and male SexDiff pruned SNPs were compared to those for the corresponding phenotype-
775 association set using a permutation scheme (FDR values indicated at the top and bottom of the
776 plot), and to 0 using a one-sided t-test (FDR values indicated in the middle of the plot). (B) Violin
777 plots of trait-SDS values. Positive trait-SDS values reflects recent increases in the frequencies of
778 phenotype-increasing alleles, while negative trait-SDS values reflect increases in the frequencies
779 of phenotype-decreasing alleles. The trait-SDS distributions for each set of female and male
780 SexDiff pruned SNPs were compared to those for the corresponding phenotype-association set
781 using a permutation analysis
782
783 **S1 Fig. Permutation enrichment distribution at each FDR threshold.** Permutation analysis of
784 the number of genes involved in sexual differentiation for all anthropometric SNPs at every FDR
785 threshold. Data are the frequency of distribution of our results for 10,000 permuted data sets. The
786 empirical P-value represents the probability that the observed value of sexual differentiation genes
787 from our SexDiff-associated SNP pool is equal to or greater than those from a randomly selected
788 set.
789
790 **S2 Fig. Sex-specific iHS scores for anthropometric SexDiff and phenotype-associated SNPs.**
791 The iHS distributions for each set of female and male SexDiff pruned SNPs were compared to

792 those for the corresponding phenotype-association set using a permutation analysis. None of the
793 distributions were significantly different.
794
795 **S1 Table** Observed number of SexDiff-associated SNPs at each FDR threshold for every
796 phenotype.
797 [a]Ratio of SexDiff-associated SNPs at the FDR threshold of 0.001 to the number of phenotype-
798 associated SNPs
799
800 **S2 Table** Observed unique sexual differentiation genes (SDG) and total number of genes for
801 SexDiff-associated SNPs and Non SexDiff-associated SNPs
802 [a]Proportion of SDG genes observed in each SNP group [b]Ratio of proportion of SDG genes
803 observed in the group of SexDiff-associated SNPs to the proportion of SDG genes observed in the
804 group of non SexDiff-associated SNPs
805
806 **S3 Table** Observed $\log_2$ ratio of female to male beta values and p-values for each set of Female
807 SexDiff-associated SNPs
808 [a]Number of pruned SexDiff-associated SNPs at an FDR threshold of 0.001 [b]Mean $\log_2$ ratio of
809 female trait effect size to the male trait effect size [c]One-sided t-test P-value comparing distribution
810 of the $\log_2$(ratio) [d]Permutation P-value of the probability that the mean $\log_2$(ratio) could be
811 observed by chance when compared to phenotype-associated SNPs
812
813 **S4 Table** Observed $\log_2$ ratio of female to male beta values and p-values for each set of Male
814 SexDiff-associated SNPs

815 [a]Number of pruned SexDiff-associated SNPs at an FDR threshold of 0.001 [b]Mean $\log_2$ ratio of
816 female trait effect size to the male trait effect size [c]One-sided t-test P-value comparing distribution
817 of the $\log_2$(ratio) [d]Permutation P-value of the probability that the mean $\log_2$(ratio) could be
818 observed by chance when compared to phenotype-associated SNPs
819
820 **S5 Table** Observed trait-SDS and permutation P-values for each set of Female SexDiff-associated
821 SNPs and Male SexDiff-associated SNPs permuted against phenotype-associated SNPs.
822 [a]Number of pruned SexDiff-associated SNPs at an FDR threshold of 0.001 [b]Average trait-SDS
823 score of pruned set of SexDiff-associated SNPs [c]Permutation P-value of the probability that the
824 trait-SDS score for each sex could be observed by chance when compared to phenotype-associated
825 SNPs
826
827 **S6 Table** Observed trait-SDS and permutation P-values for each set of Female SexDiff-associated
828 SNPs and Male SexDiff-associated SNPs permuted against phenotype-associated SNPs matched
829 for minor allele frequency.
830 [a]Number of pruned SexDiff-associated SNPs at an FDR threshold of 0.001 [b]Average trait-SDS
831 score of pruned set of SexDiff-associated SNPs [c]Permutation P-value of the probability that the
832 trait-SDS score for each sex could be observed by chance when compared to phenotype-associated
833 SNPs matched for minor allele frequency
834
835 **S7 Table** Observed average |iHS| scores and permutation P-values for each set of Female SexDiff-
836 associated SNPs and Male SexDiff-associated SNPs permuted against phenotype-associated
837 SNPs.

838 [a]Number of pruned SexDiff-associated SNPs at an FDR threshold of 0.001 [b]Average |iHS| score
839 of pruned set of SexDiff-associated SNPs [c]Permutation P-value of the probability that the |iHS|
840 score for each sex could be observed by chance when compared to phenotype-associated SNPs
841
842 **S8 Table** Observed number of intergenic SNPs and permutation P-values for each set of Female
843 SexDiff-associated SNPs and Male SexDiff-associated SNPs permuted against phenotype-
844 associated SNPs.
845 [a]Number of pruned SexDiff-associated SNPs at an FDR threshold of 0.001 [b]Permutation P-value
846 of the probability that the number of intergenic SNPs in each set of SexDiff-associated SNPs could
847 be observed by chance when compared to phenotype-associated SNPs
848
849 **S9 Table** Phenotype information
850 [a]As referred to in the Neale lab manifest released on July 31, 2018 [b] Correlation for each phenotype
851 calculated as the Spearman rank correlation coefficient between beta values of men and women
852
853 **S10 Table** Observed trait-SDS for each set of pruned phenotype-associated SNP groups.