

1 Genetic analyses on the health impacts of testosterone highlight effects on female-specific diseases and
2 sex differences

3 Jaakko T. Leinonen¹, Nina Mars^{1*}, Leevi E. Lehtonen^{1*}, Ari Ahola-Olli^{1*}, Sanni Ruotsalainen¹, Terho
4 Lehtimäki^{2,3}, Mika Kähönen^{3,4}, Olli Raitakari^{5,6,7}, FinnGen, Mark Daly^{1,8}, Tiinamaija Tuomi^{1,9,10,11,12}, Samuli
5 Ripatti^{1,8,13}, Matti Pirinen^{1,13,14}, Taru Tukiainen¹

6 1 = Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; 2 = Department of Clinical Chemistry, Finnish
7 Cardiovascular Research Center – Tampere; 3 = Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland; 4 = Department
8 of Clinical Physiology, Finnish Cardiovascular Research Center – Tampere; 5 = Centre for Population Health Research, University of Turku and Turku
9 University Hospital; 6 = Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku; 7 = Department of Clinical
10 Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland; 8 = Broad Institute of MIT and Harvard, Cambridge, MA, USA, 9 =
11 Department of Endocrinology, Abdominal Centre, Helsinki University Hospital, Helsinki, Finland, 10 = Folkhalsan Research Center, Helsinki, Finland,
12 11 = Research Programs Unit, Clinical and Molecular Metabolism, University of Helsinki, Helsinki, Finland, 12 = Lund University Diabetes Centre,
13 Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden, 13 = Department of Public Health, Faculty of Medicine, University of
14 Helsinki, Helsinki, Finland, 14 = Helsinki Institute for Information Technology HIIT and Department of Mathematics and Statistics, University of
15 Helsinki, Helsinki, Finland. * Authors contributed equally to this work. Correspondence: taru.tukiainen@helsinki.fi

16 Abstract

17 Testosterone (T) is linked with diverse characteristics of human health, yet, whether these associations
18 reflect correlation or causation remains debated. Here, we provide a broad perspective on the role of T on
19 complex diseases in both sexes leveraging genetic and health registry data from the UK Biobank and FinnGen
20 (total N=625,650).

21 We find genetically predicted T affects sex-biased and sex-specific traits, with a particularly pronounced
22 impact on female reproductive health. We show T levels are intricately involved in metabolism, sharing many
23 associations with sex hormone binding globulin (SHBG), but report lack of direct causality behind most of
24 these associations. Across other disease domains, including behavior, we find little evidence for a significant
25 contribution from normal variation in T levels. Highlighting T's unique biology, we show T associates with
26 antagonistic effects on stroke risk and reproduction in males and females.

27 Overall, we underscore the involvement of T in both male and female health, and the complex mechanisms
28 linking T levels to disease risk and sex differences.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

29 Introduction

30 Testosterone (T) is the male sex hormone responsible for regulation of development of primary and
31 secondary male sexual characteristics. Individual variation in T levels has been suggested to shape human
32 physiology broadly, including effects on disease risk in both males and females (1-3). Epidemiological studies
33 and randomized clinical trials for T replacement therapy have observed associations between serum T levels
34 and various traits ranging from type 2 diabetes (T2D) and cardiovascular disease to body composition and
35 behavior (1-11). Yet, these studies have yielded partly mixed results, and, in many instances, the proposed
36 relationships between T, complex traits and disease remain elusive (2, 7-11).

37 Besides the disease links, T is a known driver for sex differences. After puberty, males and females differ
38 extensively with respect to their average T levels, with males showing roughly 7-15-fold higher serum total T
39 concentrations (1, 12). This difference largely results from the testicular T production in males that far
40 exceeds the amount of T produced in the ovaries and the adrenal gland in females, and is known to directly
41 contribute to variation in, for instance, body composition between the sexes (1, 12).

42 In the human body, the majority of T is bound to a carrier molecule, whereas only a small fraction (1-3%) of
43 this total T exists as free T, considered to represent the most potent form of T in terms of biological activity
44 (13-15). Most of the remaining T in circulation is tightly bound by sex hormone-binding globulin (SHBG), and
45 the bulk of the rest remains attached to carrier proteins like serum albumin (13, 14). Non-SHBG bound T is
46 often approximated with free androgen index (FAI) (Figure 1A) (4, 13, 14).

47 Complex physiological processes regulate circulating T levels, allowing T levels to fluctuate based on internal
48 and external stimuli on a daily basis (16, 17). However, twin studies indicate that heritability of serum T is
49 relatively high in both sexes, up to 65% in males (18, 19). Until recently, few genetic variants were shown to
50 associate with circulating T levels and related traits in the general population (4, 18, 20, 21). Yet, through the
51 emergence of genetic and biomarker data from large biobanks, currently more than a hundred loci for serum
52 total T, SHBG and bioavailable T (i.e. free T and FAI) have been identified, with evidence for sex-specific
53 genetic effects (22-24).

54 Recent efforts have used genetics to address the potential causal contribution of adult T to selected complex
55 traits in both sexes, including T2D, body composition and hormonal cancers, or have examined the effects of
56 free T more broadly in males based on the UK Biobank data (22, 23, 25-27). Here, we extend these
57 investigations combining data from the UK Biobank (N=408,186) and FinnGen (N=217,464). Leveraging the
58 extensive healthcare registry data of FinnGen, sex-specific polygenic scores (PGS), and public data from
59 genome-wide association studies (GWAS), we studied the role of T in both sexes across traits ranging from
60 metabolic conditions and sex-specific reproductive disorders to neurological and behavioral endpoints.

61 Combining multiple analysis strategies, we illustrate causal links between T and sex-specific diseases and sex-
62 biased traits, including many reproductive conditions in females, while finding that the association between
63 T and many metabolic phenotypes may instead largely stem from shared etiology. Overall, our data provides
64 cues into the biology of T and its contribution to disease risk and complex traits, stressing the unique genetic
65 architecture of serum T levels, and emphasizing the crucial role of the hormone also for women's health.

66

67

68

69

70

71

72

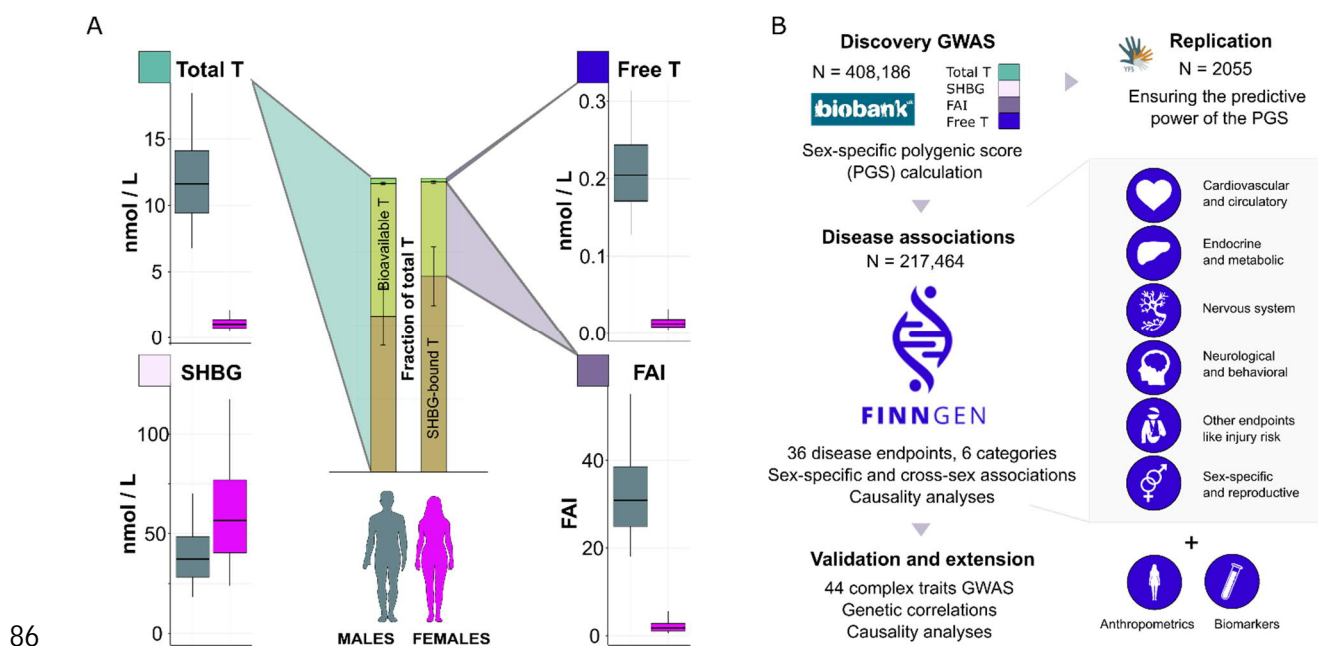
73

74

75

76 Results

77 We utilized the rich biochemical and health information available in two population-scale genetic datasets
 78 and analysis methods building on GWAS discovery (Figure 1). In brief, we conducted sex-stratified GWAS for
 79 T, SHBG, FAI and free T, using data available in the UK Biobank (Figure 1A), from which we built sex-specific
 80 PGS for these four traits. The PGSs capture the combined genetic effects on T and SHBG levels, and therefore
 81 serve as a proxy for cumulative post-pubertal T exposure. Using an external dataset (Young Finns Study; YFS),
 82 we validated the performance of the PGSs. We then investigated the effects of the PGSs on a wide range of
 83 diseases across diverse clinical entities using the FinnGen study (Figure 1B). Lastly, we evaluated causal
 84 relationships and genetic correlations between the studied T traits and complex traits, leveraging publicly
 85 available GWAS summary statistics.



87 Figure 1. Illustration of the studied traits and study overview. A) Sex-specific distributions for serum total T and SHBG
 88 and calculated FAI and free T levels for the UK Biobank participants included in the genome-wide association study
 89 (GWAS). The box plots show median (black line), lower and upper quartiles (colored area of the box) and the error bars
 90 indicate 5% and 95% quantiles. B) Overview of the study design to assess the contribution of T to health and disease
 91 using genetic approaches and biobank data. We conducted the discovery GWAS in the UK Biobank, built sex-specific
 92 PGSs for the four T-related traits, validated the PGSs in the Young Finns Study (YFS), and performed complex disease
 93 and trait associations in FinnGen (release 5) and using publicly available GWAS data.

94 GWAS and polygenic scores for testosterone traits

95 We identified more than a hundred genome-wide significant ($p < 5 \times 10^{-8}$) loci for all testosterone traits (up to
96 263 loci for SHBG in males) in the UK Biobank GWAS (Methods; Supplementary Tables 1-9). In both sexes,
97 common variants (allele frequency $> 1\%$) covered a large proportion of the trait variability, with the SNP
98 heritability (h^2) estimates ranging from 10% for total T in females to 28% for SHBG in males (Supplementary
99 Table 9). The associated loci were enriched for genes affecting steroid hormone biosynthesis, metabolism
100 and excretion, with preferential expression in the liver for all the studied traits, in line with recent findings
101 (Supplementary Figures 1&2, Supplementary Table 10) (22-24).

102 The loci affecting SHBG were largely shared between the sexes (genetic correlation (r_g) = 0.88, $p = 9.7 \times 10^{-197}$),
103 for FAI sharing was intermediate ($r_g = 0.54$, $p = 5.8 \times 10^{-26}$), but we observed a near-zero genetic correlation
104 estimates for both serum T and free T between males and females ($r_g = 0.08$ and 0.05 , respectively, $p > 0.05$),
105 indicating sex-specific genetic determinants, as previously reported (22, 23). Co-localization analyses
106 between the male and female GWAS further confirmed the widespread sex-specificity of the genetic loci
107 (Methods, Supplementary Figure 3 and Supplementary Tables 11-19) (23).

108 Reflecting the sex-specific genetic architecture, we observed strong genetic correlation between total T and
109 SHBG only in males ($r_g = 0.78$ in males vs. 0.05 in females, Supplementary Figure 3 and Supplementary Table
110 9). In females, instead, the genetic determinants for FAI and free T were shared with SHBG (e.g. $r_g = -0.80$ with
111 FAI). Highlighting these connections, SHBG was found causal for total T levels in males (genetic causality
112 proportion (GCP) = 0.80, $p = 5.8 \times 10^{-5}$, Methods), whereas in females SHBG appeared to control especially FAI
113 and free T fractions (GCP = 0.83, $p = 3.8 \times 10^{-7}$ for free T, Supplementary Table 9).

114 To study the impacts of T in datasets where T measurements are not directly available, we next constructed
115 sex-specific genetic predictors for T levels, PGS, for each trait applying the LDpred algorithm (28) to the sex-
116 specific GWASs (Methods). We tested the predictive ability of the PGS in the YFS where the phenotypic
117 variance explained by the PGS (R^2) ranged between 1.1% (male free T) to 9.2% (male SHBG), indicating the
118 PGS predict T and SHBG levels in an independent cohort (Supplementary Table 20). Notably, the sex-specific

119 PGSs for T and free T had no predictive value in the opposite sex (Supplementary Table 20 and Supplementary
120 Figure 4).

121 Studying the links between cumulative T exposure and disease

122 We continued by using the PGS to study how post-pubertal T exposure associates with disease risk, the
123 associations potentially implying causal relationships (28). To this end we used the FinnGen data, consisting
124 of 217,464 (94,478 males, 122,986 females) Finnish participants, representing roughly 5% of the Finnish adult
125 population, with genotypes linked to up to 46 years of follow-up within nationwide healthcare registries (29).
126 We studied 36 diseases with potential links to hormones from the following categories: 1) endocrine and
127 metabolic problems, 2) sex-specific endpoints (many specific to females, e.g. postmenopausal bleeding,
128 (PMB)), 3) cardiovascular and circulatory system, 4) nervous system disease, 5) behavioral and neurological
129 diagnoses and 6) other endpoints like injury risk (Supplementary Table 21). The number of cases ranged from
130 229 individuals diagnosed with hirsutism to 68,774 statin users.

131 We first tested whether the PGS are associated with disease risk in a sex-specific manner, observing 32
132 associations ($p < 0.0014$, after Bonferroni correction for 36 independent tests; Supplementary Table 21 and
133 Supplementary Figure 5). The associations involved mainly endocrine, metabolic and sex-specific disorders,
134 highlighting in particular female-specific endpoints (Figure 2A&B, and Supplementary Table 21). In males,
135 both total T and SHBG PGSs often associated with reduced disease risk, whereas we saw few associations to
136 free T. In females, higher total T, FAI and free T PGSs generally increased risk for multiple diseases, often
137 showing inverse associations to SHBG. Given the shared association profile with SHBG, we additionally
138 included the SHBG PGS as a covariate in the analyses to distinguish true T-driven effects.

139 Underscoring T's and SHBG's involvement in metabolism, larger PGS values for total T and SHBG were
140 associated with reduced T2D risk and statin use in males (for total T in males, $HR=0.94$, $p=1.4e-17$ and 0.96 ,
141 $p=1.1e-15$, respectively). SHBG appeared protective of these endpoints also in females ($HR=0.87$, $p=4.8e-58$
142 and 0.95 , $p=1.1e-21$), whereas higher female-specific free T PGS increased risk for both ($HR=1.09$, $p=1.9e-22$
143 and $HR=1.02$, $p=0.0011$). The SHBG adjusted analyses nevertheless suggested the T and free T associations

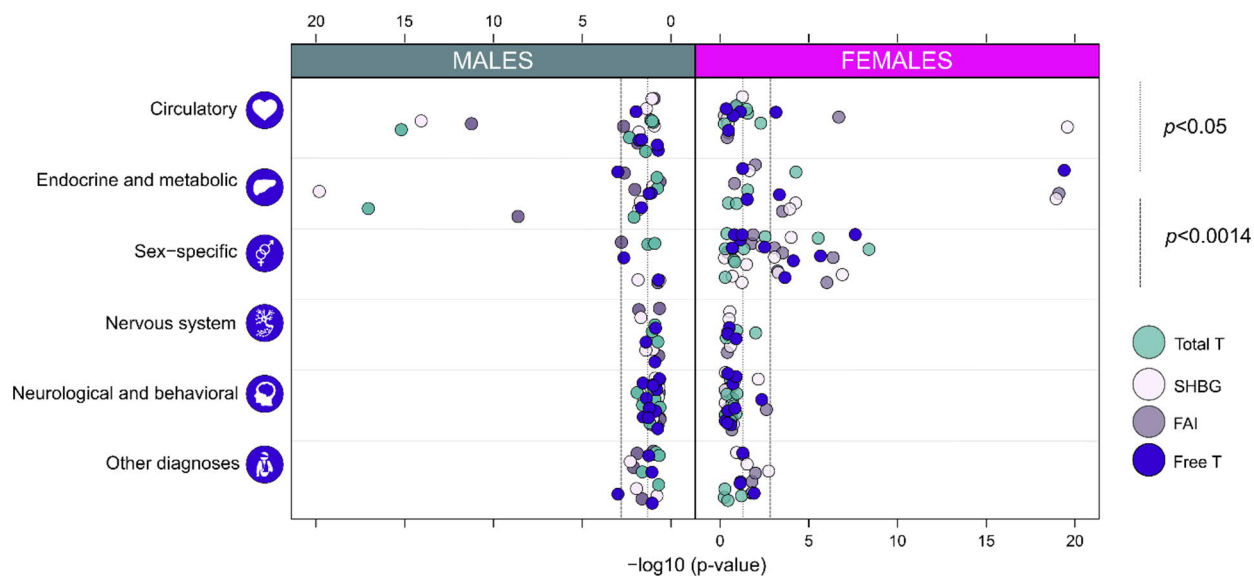
144 to these metabolic traits in either of the sexes were not primarily attributable to androgen action
145 (Supplementary Table 21 and Supplementary Figure 6). In females, higher T PGS was additionally associated
146 with lower hypothyroidism risk (HR=0.97, $p=6.5e-05$), persisting SHBG adjustment (HR=0.97, $p=8.6e-05$). We
147 also detected suggestive associations to bone strength and injury risk in both sexes, but with the exception
148 of SHBG and osteoporosis in females (HR=1.08, $p=0.00015$), none of these findings survived correction for
149 multiple testing.

150 In the sex-specific category, we replicated the known associations of T to PCOS and breast cancer risk in
151 females (22) (HR=1.02, $p=2.8e-06$ and HR=1.04, $p=0.0001$ for free T PGS) (Figure 2B and Supplementary Table
152 20). Of the novel endpoints, we robustly linked T with hirsutism and post-menopausal bleeding (HR=1.45,
153 $p=2.7e-08$ and HR=1.05, $p=0.00032$ for free T PGS). The association to hirsutism (excessive hair growth in a
154 male-type fashion) appeared particularly pronounced, with the risk of the condition almost doubling with a
155 2SD change in free T. With the exception of PCOS, total T and free T displayed similar associations to these
156 traits. The SHBG adjusted analyses suggested these associations were androgen dependent (Supplementary
157 Table 21 and Supplementary Figure 6). Whereas SHBG adjustment generally showed negligible effects on
158 total T associations, for free T and FAI, linked also with infertility risk in females (HR=1.04, $p=0.00076$ for FAI),
159 some confounding by SHBG (showing generally positive effects on reproductive traits, e.g., HR=0.98,
160 $p=0.00116$ for irregular menstruation) was evident.

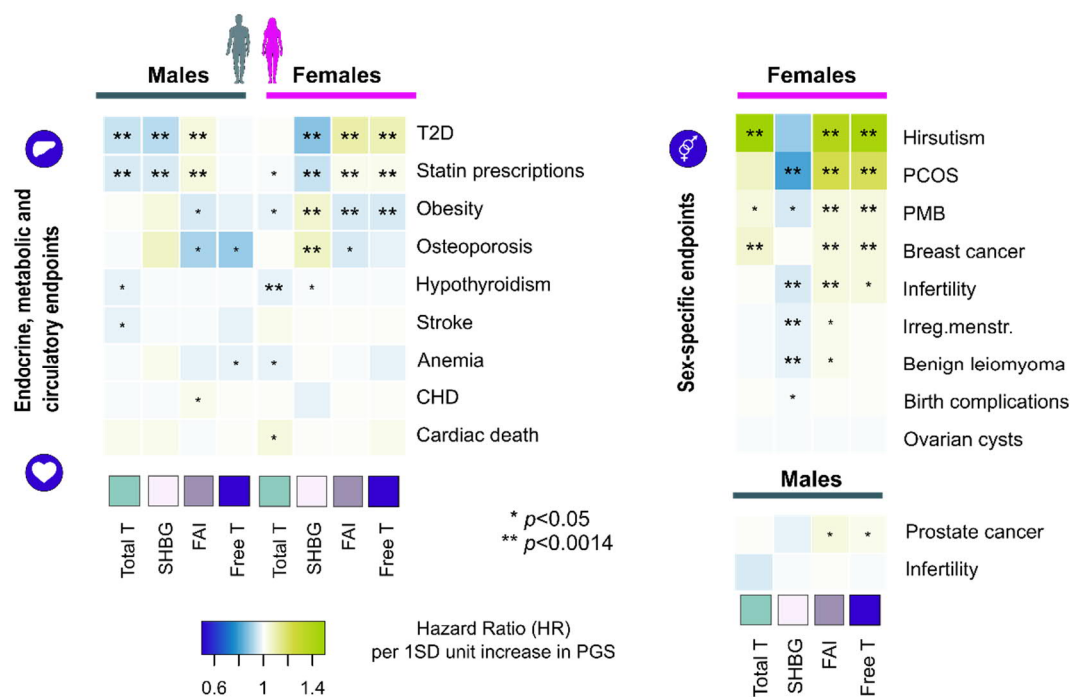
161 We observed no statistically significant associations to other diseases (Figure 2A). We for example detected
162 no associations to any of the 13 neurological/behavioral endpoints studied, including Alzheimer's disease,
163 alcohol use, and anxiety disorders (all $p>0.0014$) in both sexes. However, the PGSs did show some nominal
164 clinically-relevant associations, e.g., higher PGSs for male free T was linked with prostate cancer and lower
165 osteoporosis incidence (HR=1.03, $p=0.0083$ and HR=0.89, $p=0.0023$, respectively). Male total T associated
166 with stroke (HR=0.97, $p=0.017$), and female total T with cardiac death (HR=1.04, $p=0.049$) and Parkinson's
167 (HR=1.08, $p=0.017$). Both male free T and female total T showed associations to lower anemia risk (HR=0.97,
168 $p=0.042$ HR=0.98, $p=0.041$, respectively, Figure 2, and Supplementary Table 21).

169

A



B



170

171 Figure 2. Results from the PGS associations with disease endpoints in the FinnGen. Panel A illustrates the distribution of association
 172 p -values by disease category for each biomarker PGS separately for both sexes. For clarity p -values capped at $1e-20$. Panel B shows
 173 hazard ratios per one SD increase in PGS for 20 traits from endocrine, metabolic, circulatory and sex-specific categories. * $p < 0.05$,
 174 ** $p < 0.0014$, corresponding to Bonferroni correction for 36 independent traits. PCOS = polycystic ovary syndrome, PMB = post-
 175 menopausal bleeding, CHD = coronary heart disease.

176

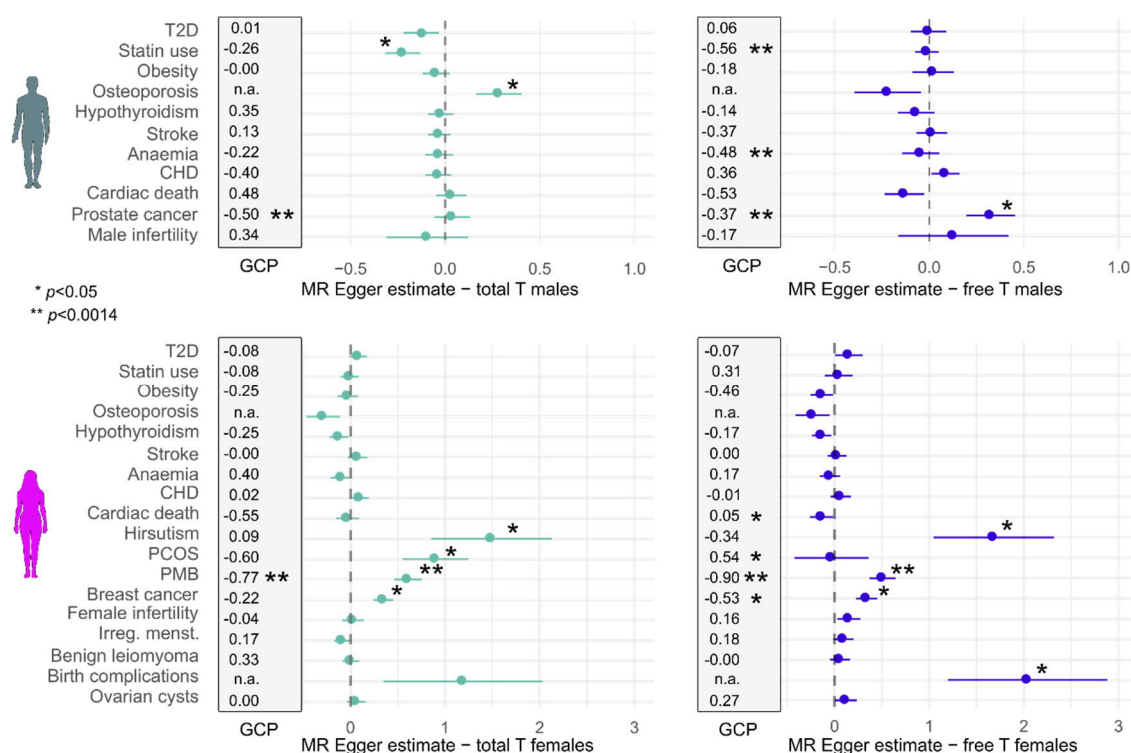
177 Evaluation of causal relationships between T and disease

178 Although the PGS associations may imply direct causality of the first trait to another, these are also prone for
179 confounders including genetic pleiotropy (30). We estimated causal relationships between T and the studied
180 endpoints using two complementary MR methods that correct for potential pleiotropy: LCV (31),
181 representing a genome-wide approach, and MR-Egger (32) that uses significantly associated SNPs. For the
182 sex-specific PGS data, in 23/252 instances one or both methods suggested evidence for a causal relationship
183 ($p < 0.0014$) between a PGS and a disease (Figure 3, Supplementary Figure 7, and Supplementary Table 22).
184 Finally, we distinguished between the effects of T and SHBG by including the latter as a covariate in
185 multivariable MR Egger analyses (33).

186 Despite the many associations (Figure 2), we observed no clear evidence for T being directly causal for
187 metabolic traits. For example, the link between total T, osteoporosis and statins in males appeared
188 confounded by SHBG (the effect being close to zero in the adjusted model, (Figure 3 and Supplementary
189 Table 22)). We also consistently saw no significant causality between T and T2D, including both sex-combined
190 and sex-specific T2D GWASs from FinnGen (Figure 3, Supplementary Figure 7 and Supplementary Table 22).

191 In contrast, the causality analyses supported the role of T in the regulation of female reproductive health.
192 Adjusting for the effects of SHBG in the multivariable MR Egger analyses seemed to strengthen these
193 causality estimates (Supplementary Table 22). Under both MR Egger models, we observed causality between
194 total and free T and postmenopausal bleeding and hirsutism ($\beta = 0.61$, $p = 4.5e-05$ and $\beta = 2.11$, $p = 0.002$,
195 respectively, for free T in multivariable MR Egger). For PCOS MR Egger supported causality for total T ($\beta = 0.90$,
196 $p = 0.012$) and LCV for free T (GCP = 0.54, $p = 0.0017$) (Figure 3 and Supplementary Table 22). Additionally, both
197 approaches indicated causality between T and hormonal cancers (e.g. GCP = -0.55, $p = 0.041$ and $\beta = 0.34$,
198 $p = 0.0031$ for female free T and breast cancer; GCP = -0.37, $p = 2.8e-16$; $\beta = 0.32$, $p = 0.014$ for male free T and
199 prostate cancer). Notably, genetically predicted free T was linked with increased cancer risk also in the
200 opposite sex (GCP = -0.37, $p = 6.6e-19$ for male free T and breast cancer, and GCP = -0.74, $p = 0.0041$ for female
201 free T and prostate cancer).

202 Lastly, some causal relationships were detected despite no robust associations in the original PGS analyses.
 203 For example, T seemed to protect females from seropositive rheumatoid arthritis (RA) (Multivariable MR
 204 Egger $\beta=-0.48$, $p=0.0025$ for free T), and T and SHBG were linked with injury risk in both sexes (e.g. GCP=0.53,
 205 $p=0.0018$ and $\beta=0.17$, $p=1.3e-05$ for SHBG increasing forearm/elbow injuries in males). Finally, LCV suggested
 206 a causal relationship between higher SHBG and ADHD in both sexes, and higher free T and increased risk for
 207 conduct disorder but decreased risk for emotionally unstable personality in males ($p<0.0014$, Supplementary
 208 Table 22), pointing to potential hormonal involvement in the regulation of neuronal processes.



209 Figure 3. Results from the causality analyses in FinnGen, showing MR estimates for total and free T in males and females. The
 210 figure includes traits from the endocrine, metabolic, circulatory and sex-specific categories shown in Figure 2. Shown are both genetic
 211 causality proportion (GCP) estimates from LCV (within the grey boxes, 0=no causality, 1= fully causal, minus sign indicates suggested
 212 reverse causality i.e. the endpoint affecting T), and MR Egger betas and standard error for each trait (horizontal lines). The MR Egger
 213 value in the x-axis corresponds to SD increase in risk/per SD increase in T/free T. * $p<0.05$, ** $p<0.0014$, n.a.=not applicable due to low
 214 heritability estimate in FinnGen under the LCV model.
 215
 216
 217

218 Results from the cross-sex PGS associations in FinnGen

219 Owing to the unique genetic architecture, the sex-specific PGS for T and free T do not predict the
220 corresponding hormone levels in the opposite sex (Supplementary Table 20). Given this, we reasoned that
221 cross-sex analyses, i.e., analysis of the effect of a sex-specific PGS in the opposite sex, would provide us with
222 an additional means to assess if the original associations stem from T action, and to detect potential
223 antagonistic effects for the PGSs between the sexes.

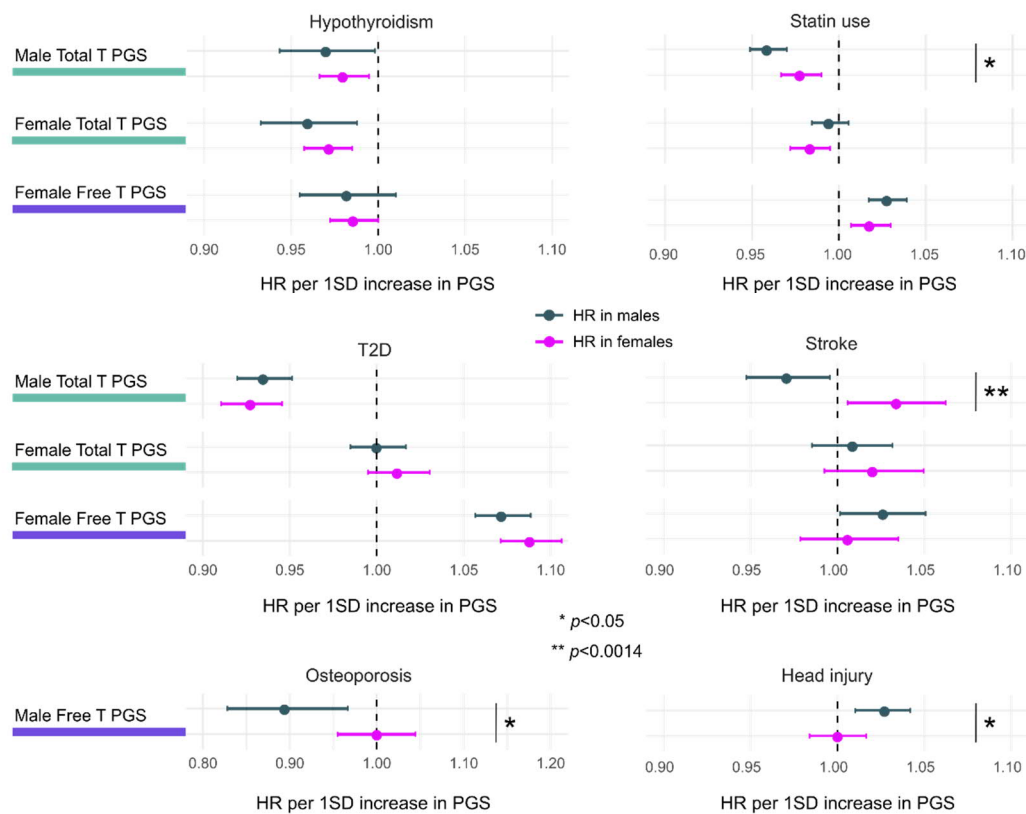
224 Aligning with the results from the MR analyses, 15/20 of nominally significant ($p < 0.05$) sex-specific
225 associations remained similar (Z-test $p > 0.05$ for difference in PGS effects) also in the cross-sex analyses,
226 pointing to shared genetic etiology rather than T action as the basis for many metabolic associations. For
227 example, both female and male total T PGSs, with and without SHBG adjustment, associated with
228 hypothyroidism risk with a similar effect size in the opposite sex, and all PGS associations to T2D were
229 replicated in the other sex (Figure 4A, and Supplementary Tables 23&24).

230 The associations differed (Z-test $p < 0.05$) in the opposite sex for male total T PGS and stroke, female total T
231 PGS and anemia, and male free T PGS, head injury risk and osteoporosis. (Figure 4A, and Supplementary
232 Tables 23&24). In addition, although an effect of male total T PGS on statin use was observed also for females,
233 it was attenuated. Intriguingly, while higher male total T PGSs reduced stroke risk in males (HR=0.97,
234 $p=0.017$), this associated with increased risk for stroke in females (HR=1.03, $p=0.017$). This was the only
235 endpoint for which we detected evidence for sexual antagonism. The antagonistic effect enhanced with SHBG
236 adjustment (HR=0.96, $p=0.011$ in males and HR=1.07, $p=0.0010$ in females). The results indicate the genetic
237 effects on stroke risk may be partly sex-specific (34), with potential interplay from sex hormones.

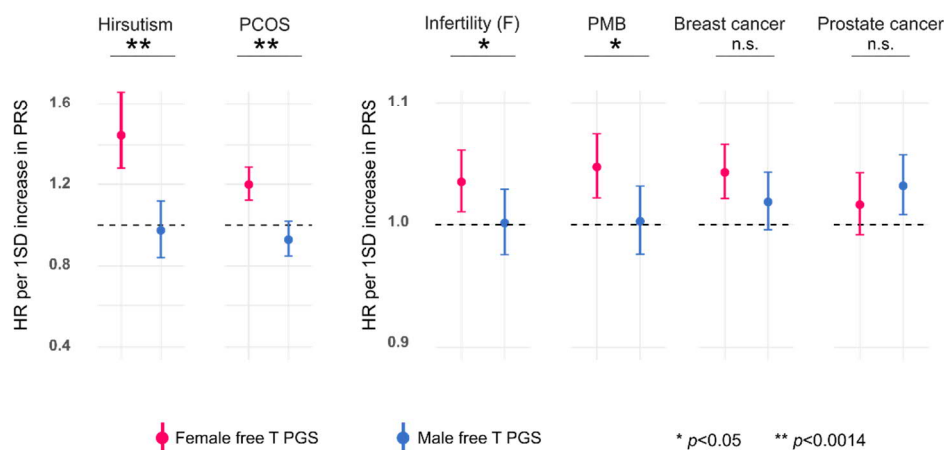
238 Finally, the cross-sex analyses implied increased androgen load as a direct contributor to poorer reproductive
239 health in females, agreeing with the MR-based causality assessments. For the reproductive endpoints that
240 were associated with female free T PGS, i.e., infertility, PMB, PCOS and hirsutism, the male free T PGS had
241 no predictive power (all $p < 0.05$, Figure 4B, Supplementary Tables 23&24). Nonetheless, the effect sizes were
242 not attenuated in a statistically significant manner for breast and prostate cancers, ($p > 0.05$, Figure 4B),

243 echoing the causality analyses and pointing to some degree of shared genetic risk, irrespective of free T levels,
 244 between male and female hormonal cancers.

A



B



245

246

247

248

249

250

Figure 4. Results from the cross-sex analyses. A) Cross-sex PGS associations with hypothyroidism, statin use, T2D and stroke. $p < 0.05$ suggests that the effects of a given PGS vary depending on sex B) Illustration of cases with statistical evidence for male and female PGSs having different effects (hirsutism, polycystic ovary syndrome (PCOS), infertility and post-menopausal bleeding (PMB)). No evidence for such difference for breast and prostate cancers. *Chi-Squared $p < 0.05$, ** = Chi-Squared $p < 0.0014$, n.s. = not significant. Panels A and B show HR point estimates with 95% confidence intervals.

251 Extending the FinnGen discoveries

252 We next sought to validate and refine the FinnGen discoveries in additional datasets, extending our analysis
253 to include quantitative traits not available in large numbers in FinnGen. To this end, we used genetic
254 correlation analysis, allowing for estimation of the extent to which two traits are affected by the same genetic
255 factors (35, 36), followed by causality estimations.

256 We selected 44 traits with publicly available GWAS summary statistics, identical to (e.g., T2D, breast and
257 prostate cancers) or closely reflecting the studied disease phenotypes (heel bone mineral density (HBMD),
258 mood swings) from FinnGen, adding anthropometric traits to the analyses (Supplementary Table 25). For
259 most female-specific phenotypes studied in FinnGen, including hirsutism and PMB, we had no comparable
260 phenotypes, as there are no published GWAS available.

261 We found evidence of significant genetic correlation in 72/352 instances ($p < 0.0011$, corresponding to
262 Bonferroni correction for 44 independent tests) (Figure 5, Supplementary Table 25) with the results reflecting
263 the FinnGen PGS associations. We observed significant genetic correlations to traits related to metabolism,
264 including many biomarkers and anthropometrics, but detected only few correlations to behavioral traits, and
265 no significant correlations to neurological or temperamental traits (Figure 5, Supplementary Table 25).
266 Notably, in all cases where we observed genetic correlation to behavioral traits, these had clear links to
267 metabolism (smoking, sleep duration and exercise).

268 The genetic factors increasing serum total T and SHBG appeared to promote a favorable metabolic profile in
269 males, supporting the PGS findings. Despite correlating with increased BMI, total T and SHBG were positively
270 correlated to adiponectin, high-density lipoprotein (HDL) and lower waist-to-hip ratio (WHR) ($r_g > 0.20$,
271 $p < 0.0011$) whilst lowering triglycerides and T2D incidence in males ($r_g < -0.25$, $p < 0.0011$). In contrast, higher
272 FAI and free T fractions in females associated with negative metabolic effects, including higher WHR ($r_g = 0.25$,
273 $p = 1.6e-22$) and lower HDL cholesterol ($r_g = -0.18$, $p = 4.2e-05$), whilst free T showed no significant correlations
274 to these metabolism-related traits in males. Strong correlations were observed also between SHBG and

275 metabolic traits in females, including negative associations to markers of liver damage (alanine transaminase
276 (ALAT), $r_g=-0.21$, $p=1.7e-08$) and gamma-glutamyl transferase (GGT), $r_g=-0.16$, $p=0.0001$).

277 We observed significant genetic correlations to hormonal cancers in both sexes, in line with recent findings
278 (22, 25). In males, the genetic factors increasing FAI and free T promoted prostate cancer ($r_g=0.12$, $p=0.0004$),
279 whereas in females these increased especially the risk of estrogen receptor (ER) positive breast cancer
280 ($r_g=0.15$, $p=9.1e-09$). Finally, consistent with existing genetic, epidemiological and experimental data (1, 25),
281 the genetic correlation analyses pointed towards shared genetic background for T, hemoglobin levels and
282 body fat in both sexes (e.g. $r_g=0.15$, $p=1.0e-07$ and $r_g=-0.14$, $p=0.0002$ for hemoglobin and body fat,
283 respectively, with male free T).

284 In the LCV and MR-Egger analyses we found statistically significant evidence ($p<0.0011$) of a causal
285 relationship in 7% (26/354) of instances across the 44 traits (Figure 5 and Supplementary Table 26).
286 Reflecting the results from FinnGen, despite the genetic correlations to many traits there was no evidence
287 for significant causality, but the few suggested causal relationships involved traits with clear biological links
288 to T function.

289 As examples of expected causal relationships (1, 25, 37), our analyses supported the contribution of free T
290 levels to male-pattern baldness (MPB, $GCP=0.44$, $p=7.7e-19$) and hemoglobin levels ($GCP=0.64$, $p=0.00085$)
291 (Figure 5 and Supplementary Table 26). In addition, despite the lack of significant genetic correlations
292 between these traits, higher T in males was linked to increased number of children fathered (NCF, $GCP=0.62$,
293 $p=1.2e-15$ for total T, $GCP=0.42$, $p=0.011$ for free T), raising speculation about the potential evolutionary
294 benefits of males maintaining adequate T levels.

295 The MR Egger analyses supported the causality of female total T and free T to ER+ breast cancer ($\beta=0.282$,
296 $p=0.00014$ and $\beta=0.250$, $p=0.002$, respectively). Instead of T, in these analyses prostate cancer risk was linked
297 to SHBG levels in males ($GCP=-0.64$, $p=0.00019$) (38). Additionally, SHBG increased lymphocyte count in both
298 sexes ($GCP=0.39$, $p=1.6e-07$ in males, $GCP=0.42$, $p=0.00014$ in females) (39), and in males, we linked SHBG
299 also with reduced risk of erectile dysfunction ($GCP=0.42$, $p=0.00032$). In females, we observed causality

300 between SHBG and age at menopause ($\beta=-0.756$, $p=0.0010$), suggesting that inherited differences in SHBG
301 levels likely modify reproductive phenotypes across female lifespan.

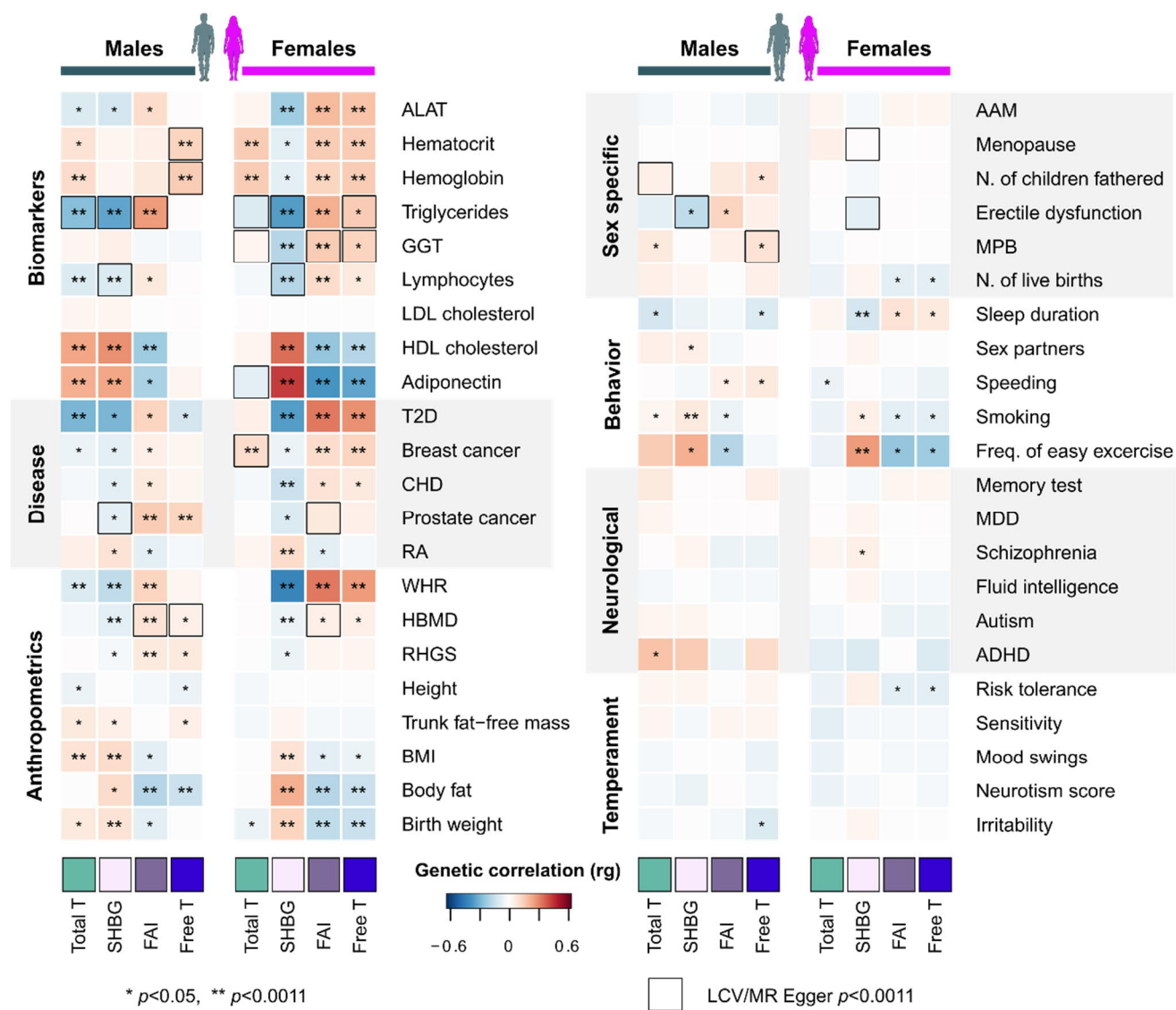
302 In most instances, we however found no evidence of a significant causal relationship between T levels and
303 the studied traits (Supplementary Table 26). Yet, emphasizing the intricate relationship between hormone
304 levels and metabolism (40), some connections to metabolism-related biomarkers emerged: for example, it
305 appeared that triglycerides may causally influence T (GCP=-0.48, $p=1.3e-30$) and free T (GCP=-0.80, $p=1.4e-$
306 06) levels in females. In combination with the PGS analyses, these results thus further suggest that whilst
307 some sex-biased phenotypes may be directly related to T levels, in most instances T's relationship to complex
308 traits and diseases is not straightforward.

309

310

311

312



323 Discussion

324 Since its discovery in the early 20th century, testosterone (T) has been proposed to modify phenotypes and
325 diseases that differ between the sexes, due to the extensive male-female differences in circulating T levels.
326 Besides the biological impacts of normal variation in T levels, the use of T replacement therapy in medical
327 practice has been globally on rise, and the potential risks and benefits associated with T supplementation
328 remain debated (7). To provide a broad and systematic perspective into the function of T as a regulator of
329 health and disease in males and females we leveraged the UK Biobank resource to construct PGSs - predictors
330 for genetically determined T levels - for both males and females, which we then associated with a uniquely
331 rich collection of disease endpoints from 217,464 participants in FinnGen. This combination allowed us to
332 extend and refine recent efforts that have utilized genetic data to understand the disease impacts of T,
333 concentrating on a limited set of phenotypes or only to males (22, 23, 25-27). In addition, by taking advantage
334 of the sex-specific genetic determinants of T in cross-sex analyses, and through careful Mendelian
335 Randomization strategies, we could pinpoint T's causal effects on adult health.

336 Based on our analyses, three major themes emerged regarding T's contribution to disease. First, we report
337 that the studied PGSs associated with disease risk especially in females. Secondly, we highlight distinct
338 association profiles for total T and free T in males, consistent with proposed divergent biological effects for
339 the bound and unbound T fractions (13-15). Underscoring the potential role of SHBG as a confounder, the
340 former closely correlates with SHBG levels in males, and the latter in females (13-15). Thirdly, taking into
341 account such confounding, we generally observed causal relationships between a genetic predisposition to
342 higher T levels and several sex-specific and sex-biased phenotypes with clear biological links to T, but less
343 contribution to most other phenotypes, echoing experimental data and findings from recent MR studies (22,
344 25).

345 Besides the causal links to several female-specific reproductive endpoints, sex-biased traits and hormonal
346 cancers, significant associations involved diseases and traits from the metabolic and endocrine categories.
347 We stress the apparent complexity of these relationships. For many metabolic traits SHBG – either directly

348 or potentially through its action on other hormones - appeared to confound T's associations and causality
349 estimates. Beyond this T/SHBG relationship, we furthermore estimate that some of the observed
350 associations to metabolic health, including T2D and hypothyroidism, may generally reflect widespread
351 genetic pleiotropy and thus overall complex shared genetic etiology rather than T action.

352 Based on our data, we further speculate that normal variation in T levels - contrary to popular beliefs - has
353 only modest effects on most phenotypes. Particularly, the grounds to explain some temperamental and
354 neurological phenotypes like anxiety and emotional instability with heritable differences in adult T levels (41,
355 42) appears unsubstantiated in the light of our study. Although we cannot fully exclude a causal relationship
356 between T and some of these phenotypes, our data suggests that especially without larger sample sizes or
357 refined phenotyping, efforts to link T levels to behavior will likely be unproductive. Taken together,
358 supporting recent recommendations, our data thus suggests that the risks and benefits of using T as a medical
359 treatment should be carefully weighted, given T's complex and indirect relationship to most phenotypes and
360 potential adverse and beneficial outcomes in both sexes (7, 25).

361 Having comprehensively mapped the impacts of T across diverse complex disease and traits, we can start
362 drawing inferences on the role of T as a contributor to the male-female differences. Indeed, in such instances
363 where causality of T was implicated for a sex-shared trait, the effect estimates often aligned with the
364 direction of the phenotypic sex bias in the given trait, i.e., higher T levels associated with typical male
365 characteristics. For instance, our work highlighted causal connections between increasing free T, higher
366 hemoglobin and higher bone strength, backed up by previous experimental observations (1, 43, 44).
367 Extrapolating from the MR results, we estimate that ~10-20% of the mean difference in hemoglobin levels
368 between males and females may result from average differences in free T levels, consistent with the notion
369 that T directly affects male-female differences in, e.g., athletic capacity (1). Moreover, higher free T levels
370 were found causal to masculine external features like hirsutism and baldness, and in females higher
371 bioavailable T levels correlated genetically with a shift of metabolism into a male-like direction (e.g.,
372 increased BMI, WHR, and poorer blood lipids, yet reduced risk for obesity and reduced body fat levels).

373 Overall, the limited evidence for the causal involvement of T for most of the studied traits however suggests
374 that most phenotypic sex differences are not attributable to a linear relationship between T levels and a
375 phenotype. Instead the impact of T can be mediated through a threshold effect, potentially at a given
376 developmental time window, acting as a switch that results in more global rewiring of biological processes
377 and thereby in systematic male-female differences. Under this model, the within-sex variability in T may have
378 non-existent or very subtle effects on many phenotypes, complicating the detection of potential causality of
379 T.

380 Finally, our study provides some unique insight into the potential causes behind the sex differences in T
381 levels. In the cross-sex PGS analyses, we found only one case where there was clear evidence of antagonistic
382 effects for the T PGS between the sexes. The male-specific total T PGS protected from stroke in males,
383 whereas the same PGS had opposite effects on stroke risk in females. Although we cannot conclude that the
384 action of T truly drives these associations, the result agrees with a degree of sex-specificity in the genetic
385 disease mechanisms for stroke (34).

386 The observation that the genetic variants responsible for regulating T levels are largely distinct between
387 males and females moreover raises speculation about the evolutionary forces maintaining sex differences
388 and shaping the biology of T. The cross-sex genetic correlations for traits related to fitness (e.g., reproductive
389 success) are generally expected to be low, due to potentially conflicting evolutionary pressures (45). We
390 indeed associated T positively with reproductive success in males (T increasing the number of children
391 fathered), but negatively with both pre- and postmenopausal reproductive health in females with evidence
392 for a causal role of T behind these associations. We may thus speculate that there exists a selective advantage
393 in relation to reproductive success to maintain higher T levels in males, whereas the opposite may be true
394 for females, potentially promoting the widespread sex differences across multiple traits.

395 Although based on extensive data sets, our study still has some limitations. We stress that studying T levels
396 differs drastically from studying T action, serum T levels serving only as a proxy for the latter. Furthermore,
397 although we extrapolate that the PGS for total T and SHBG used in the study explain a substantial fraction of

398 the variance in these traits in the studied cohorts, for example the male free T PGS seemed to yield less
399 accurate predictions. Genetic studies are also prone to confounding by pleiotropy, whereby a gene influences
400 multiple traits via independent biological pathways (30). Although we opted for adjusting for the effect of
401 both SHBG and BMI in our analyses – both known confounders for testosterone levels (Supplementary Figure
402 8) (46, 47) - it remains possible that these connections still affect our findings. Genetic pleiotropy may
403 confound also MR analyses, and despite the vast potential of MR in establishing causal relationships (48), we
404 generally propose caution in interpreting these findings. In our case the MR models did not always agree on
405 causality (Supplementary Tables 22 and 27), common to recent MR studies assessing the function of T (22,
406 23, 25). When based on large number of variants, the MR analyses often include variants that do not fulfill
407 the strict definition of an instrumental variable. Here we show via MR and cross-sex PGS analyses that such
408 pleiotropic variants may underlie for example T's associations with T2D and hypothyroidism.

409 Importantly, our setting does not allow for assessing the effects of fetal T exposure, which may be critical,
410 e.g., for neurological traits (49). We also emphasize that our results are based on normal variation in T levels,
411 not on supraphysiological T injections. Additionally, many of T's effects depend on its conversion to estradiol
412 also in males, and we cannot rule this out as a potential confounder in our study. Finally, the data used in our
413 study does not allow for assessing the effects of acute changes in hormone secretion, and personal
414 differences in the response to such fluctuations may be crucial for some phenotypes.

415 Despite these challenges, we were able to highlight several novel albeit often expected relationships with
416 genetically determined T levels, human health, and sex differences. Besides the gained medical insight,
417 underscoring some critical factors that should be considered when assessing these relationships, with this
418 study we thus provide a reference point for future genetic and epidemiological studies studying the action
419 of T.

420

421

422

423 Methods

424 Genotype and phenotype data from the UK Biobank

425 The genetic association analysis was based on data from the UK Biobank, a population-based biobank
426 consisting of 502 637 subjects (aged 37-73 years)(50). At recruitment, participants provided electronic signed
427 consent. Ethics approval for the UK Biobank study was obtained from the North West Centre for Research
428 Ethics Committee (11/NW/0382). All experiments were performed in accordance to relevant guidelines and
429 regulations including the Declaration of Helsinki ethical principles for medical research. This study was run
430 under UK Biobank application number 22627.

431 Multiple biochemical assays have been performed on the entire UK Biobank cohort, including measurements
432 of serum testosterone, SHBG and serum albumin. These measurements were performed once for each
433 participant for using Beckman Coulter DXI 800 Chemiluminescent Immunoassay, with competitive binding
434 for testosterone and two-step sandwich for SHBG, referenced in
435 https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_biochemistry.pdf. The document contains
436 also details about the extensive QC procedures for all biochemical measurements in the UK Biobank. For
437 example, assay consistency was monitored by using internal QC samples between batches, and external
438 quality assurance schemes against the ISO 17025:2005 standard.

439 We restricted our study to encompass 408,186 individuals from the white British subset. We removed
440 outliers for genotype heterozygosity and missingness, as well as samples with sex chromosome aneuploidies,
441 mismatches between reported and inferred sex, and samples that UK Biobank did not use in relatedness
442 calculations (50). We did not exclude related samples since our analysis method (BOLT-LMM) allows for their
443 inclusion.

444 For the GWASs, we used biochemically measured testosterone and SHBG, and calculated FAI and free T. For
445 calculation of FAI, we used the formula $100 * \text{Testosterone} / \text{SHBG}$ (nmol/ml). Calculated free T was derived
446 using the Vermeulen equation using directly measured albumin values for each participant in the equation,
447 as described in (51, 52).

448 All four traits were separated by sex and log transformed. A linear regression model was fit for each trait with
449 BMI and age as covariates, as well as menopause status for females. Subjects with residuals values of ± 5 SD
450 from the mean were excluded from the analyses, serving as a further QC step to exclude outliers potentially
451 reflecting medical conditions or drug use affecting androgen levels. Inverse normalized values of the
452 remaining residuals were used as phenotype values for the GWAS analyses. After the QC-steps, our study
453 included altogether 177,499 males and 205,141 females.

454 Genetic association analysis and definition of the lead SNPs

455 The GWAS analyses were performed using BOLT-LMM (v2.3.2)(51, 52). Imputed SNPs were restricted to
456 variants with $MAF \geq 0.1\%$ and imputation quality ≥ 0.7 (50). 1000 Genomes European data was used as
457 reference LD scores for calibrating the BOLT-LMM statistic. First 10 Principal Components were used as
458 quantitative covariates in the runs. A linear regression model was fit for each trait with BMI and age as
459 covariates, as well as menopause status for females. Genetic correlation analyses, heritability estimates and
460 number of loci found implied the results remained consistent with different covariate configurations
461 (Supplementary Figure 8), but we chose to include body mass index (BMI), known to associate with T levels
462 (28, 29), as a covariate in our GWAS. Including up to 127 covariates (based on (53)), e.g., assay center, dilution
463 factors, blood draw time, and socioeconomic status indicators, or excluding related individuals from the
464 analysis all showed negligible effects on the genetic findings we report here (Supplementary Figure 8).

465 Independent lead SNPs were selected for each chromosome by recursively taking the SNP with the lowest p-
466 value (until none below the p -value threshold $5e-08$ were left) from the GWAS summary statistics and
467 removing all SNPs 500kb on each side of it from the next round. The chromosomal positions of these 1mb
468 windows were stored, and overlapping windows were merged into the final list of loci. The SNP with the
469 lowest p-value in each of these windows was selected as the lead SNP.

470 Pathway, tissue enrichment and co-localisation analyses

471 Tissue and gene set enrichment analyses were carried out with SNP2GENE and GENE2FUNC implemented in
472 FUMA using default settings (54). For testing in which tissues the genes residing in the GWAS loci were

473 preferably expressed, we used the full distribution of SNP p-values and the GTEx v6 30 general tissue types.
474 For pathway analysis, to assess whether the genes in the GWAS loci are overrepresented in pre-defined gene
475 sets via hypergeometric tests, we selected manually curated KEGG-pathways (55). For co-localisation
476 analyses to assess whether the genetic loci showed evidence for shared genetic effects between males and
477 females, and to estimate the maximum posterior probability (MAP) for the loci being shared, we used gwas-
478 pw (56).

479 Replication in the Young Finns Cohort and calculation of PGS

480 The Cardiovascular Risk in Young Finns Study (YFS) is a longitudinal follow-up of 3,596 subjects at baseline.
481 The baseline survey was conducted in 1980 and subsequent follow-ups involving the whole sample were held
482 in 1983, 1986, 2001, 2007, 2011, and 2017. For testosterone and SHBG, we used data on 2001 follow-up
483 (Subjects aged 24-39 yrs). A venous blood sample was drawn from antecubital vein after 12-hour overnight
484 fast. Serum was aliquoted and stored in -70 Celsius degrees until analysis. In males, total testosterone
485 quantification was performed in 2009 with competitive radioimmunoassay (Spectria Testosterone kit, Orion
486 Diagnostica, Espoo, Finland) and Bio-Rad Lyphocheck control serums 1, 2, and 3 were used in quality control.
487 Before quantification, serum aliquots had been melted three times. Total testosterone was quantified first
488 and aliquots were re-frozen before SHBG quantification. In females, total testosterone quantification was
489 performed in 2011. SHBG quantification for males was done in 2009 and for females in 2011 with Spectria
490 SHBG IRMA kit (Orion Diagnostica, Espoo, Finland). Free testosterone was estimated using Vermeulen's
491 formula. As albumin concentration was not available, we used fixed albumin concentration of 43 g/l.

492 In YFS, genotyping was performed at the Wellcome Trust Sanger Institute (UK) using customized Illumina
493 Human Map 670k bead array. The custom content on 670k array replaced some poor performing probes on
494 Human610 and added more CNV content. As quality control, we excluded individuals and probes with over
495 5% of missingness (--geno and --mind filters in Plink). Variants deviating from Hardy-Weinberg equilibrium (p
496 $< 1 \times 10^{-6}$) and minor allele frequency below 1% were excluded. Related samples were excluded ($n=51$) with
497 π -hat cut-off of 0.2. Total of 2,442 individuals and 546,674 variants passed the quality control measures. The

498 mean call rate across all included markers after the quality control was 0.9984. Next, imputation was
499 performed using population-specific Sequencing Initiative Suomi (SISu) as reference panel. We examined the
500 association of the UK Biobank GWAS lead SNPs with the corresponding T trait in YFS. If the annotated lead
501 SNP was not available in YFS, we used LDstore2 (v2.0b) (57) to calculate LD in a 100kb window around the
502 lead SNP in the UK Biobank imputed data and selected the closest SNP with $R^2 > 0.8$ with the lead SNP as a
503 proxy.

504 To construct PGS we applied the LDpred (28) method to the sex-specific GWAS results from the UK Biobank
505 for total T, SHBG, FAI and free T using 1000 Genomes Europeans as LD reference and the default LD radius to
506 account for LD. We then used the weights from the LDpred infinitesimal model to construct genome-wide
507 PGSs for each individual in the YFS with Plink 2.0 (11 Feb 2018). Only variants imputed with high confidence
508 (imputation $INFO \geq 0.8$) were included in PGS calculation. Variants in chromosomes 1-22 and chrX were
509 included. In males, allele dosage of 2 was used for X-chromosomal haploid variants. To evaluate the PGS
510 prediction accuracy in the YFS, we calculated the R^2 for each trait using using linear regression with z-score
511 normalized PGS as predictor, age and 10 PCs as covariates and z-score normalized T trait as outcome.

512 Estimation of heritability and calculation of the genetic correlations by LDSC

513 SNP-based heritability for the studied T traits and genetic correlations between these and with 44 additional
514 phenotypes were estimated using linkage disequilibrium score regression (LDSC)(35). The summary statistics
515 for the 44 traits were downloaded directly from the source repositories and analysed locally, for the original
516 sources please see references in Supplementary Table 25. For the genetic correlation analyses, pre-computed
517 LD Scores from 1000 Genomes Europeans excluding the HLA region were used. For 23 traits we performed
518 the analyses using sex-specific GWAS results and compared these to data from sex-combined GWAS
519 (Supplementary Table 27). Generally, genetic correlation results using either sex-specific or sex-combined
520 GWAS data were highly similar.

521 Disease associations in FinnGen

522 To assess if the PGS for studied traits associate with disease risk we utilized the FinnGen study (data freeze
523 5), consisting of 217,464 (94,478 males, 122,986 females) (29). FinnGen is comprised of Finnish prospective
524 epidemiological and disease-based cohorts and voluntary biobank samples collected by hospital biobanks.
525 The genotypes have been linked to national hospital discharge (available from 1968), death (1969–), cancer
526 (1953–) and medication reimbursement (1964–) registries as well as the registry on medication purchases
527 (1995–). The samples were genotyped with Illumina and Affymetrix arrays (Illumina Inc., San Diego, and
528 Thermo Fisher Scientific, Santa Clara, CA, USA). The genotypes have been imputed with using the SISu v3
529 population-specific reference panel developed from high-quality data for 3,775 high-coverage (25-30x)
530 whole-genome sequencing in Finns. The detailed genotype imputation workflow can be found at
531 <https://dx.doi.org/10.17504/protocols.io.xbgfijw>. The dataset uses genome build 38 (hg38).

532 For PGS analyses, we used same variant weights (LDpred infinitesimal model) as for YFS, and calculated
533 genome-wide PGSs for each individual with PLINK2 (v2.00a2.3LM). Variants in chromosomes 1-22 and
534 chromosome X (imputed with high confidence, imputation INFO ≥ 0.7) were included (total number of
535 variants ranging from 6,535,263 for female total T to 6,536,405 for female SHBG) and we used genotype
536 dosages to incorporate imputation uncertainty. In males, allele dosage of 2 was used for X-chromosomal
537 haploid variants. We studied the PGS associations to 36 disease endpoints with potential links to androgens,
538 representing six loosely defined disease categories. For details of the studied phenotypes see Supplementary
539 Table 21, www.finnngen.fi and risteys.finnngen.fi. Cox proportional hazards models were used for estimating
540 hazard ratios (HRs) and 95% CIs, with age as the time scale and 10 first principal components of ancestry and
541 genotyping batch as covariates. The proportionality assumption for Cox models was assessed with
542 Schoenfeld residuals and log-log plots. For the cross-sex analyses, we took the sex-specific PGSs, and checked
543 whether these would associate with the studied endpoints in the other sex, using the z-test to compare
544 equality between the original and cross-sex associations. We additionally performed SHBG-adjusted PGS
545 associations to all endpoints to control for potential confounding of SHBG to total and free T.

546 Patients and control participants in FinnGen provided informed consent for biobank research, based on the
547 Finnish Biobank Act. Alternatively, older research cohorts, collected prior the start of FinnGen (in August
548 2017), were collected based on study-specific consents and later transferred to the Finnish biobanks after
549 approval by Valvira, the National Supervisory Authority for Welfare and Health. Recruitment protocols
550 followed the biobank protocols approved by Valvira. The Coordinating Ethics Committee of the Hospital
551 District of Helsinki and Uusimaa (HUS) approved the FinnGen study protocol Nr HUS/990/2017.

552 The FinnGen study is approved by Finnish Institute for Health and Welfare (THL), approval number
553 THL/2031/6.02.00/2017, amendments THL/1101/5.05.00/2017, THL/341/6.02.00/2018,
554 THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019, Digital and population data
555 service agency VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3 the Social Insurance Institution
556 (KELA) KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019, and Statistics Finland
557 TK-53-1041-17. The Biobank Access Decisions for FinnGen samples and data utilized in FinnGen Data Freeze
558 6 include: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71,
559 BB2019_7, BB2019_8, BB2019_26, Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank
560 HUS/359/2017, Auria Biobank AB17-5154, Biobank Borealis of Northern Finland_2017_1013, Biobank of
561 Eastern Finland 1186/2018, Finnish Clinical Biobank Tampere MH0004, Central Finland Biobank 1-2017, and
562 Terveystalo Biobank STB 2018001.

563 Causality analyses

564 MR analyses treat genetic variants as instrumental variables and their reliability depends on two key
565 assumptions: 1) alleles are randomly assigned, and 2) that alleles that influence exposure do not influence
566 the outcome via any other means. The first assumption is controlled by using BOLT-LMM as our model in the
567 primary GWAS analysis, but the second is harder to control when using a large number of SNPs as
568 instrumental variables. In line with the observed wide-spread genetic pleiotropy affecting most complex
569 traits, we noted that the GWAS loci contained many genes associated with pleiotropic effects on human
570 phenotypes (for example, *LIN28B* (58), *GCKR* (56) and *TYK2*(59)). Therefore, given the vast polygenicity of the

571 studied T traits, we chose latent causal variable (LCV) (31) and MR-Egger (32) as our primary MR methods,
572 designed to take into account pleiotropy-induced confounding when assessing causal relationships. LCV has
573 been proposed to provide more unbiased causality estimates than conventional MR approaches, whereas
574 MR Egger should provide accurate causality estimates under the InSIDE assumption (the genetic variants
575 have pleiotropic effects that are independent in magnitude and are thus not mediated by a single confounder
576 exposure), besides its recommended use as a sensitivity analysis for conventional MR (31, 32). For
577 comparison we also ran conventional MR analyses (Inverse-Variance Weighed (IVW)), that remains more
578 sensitive for confounding by genetic correlation and pleiotropy (31). To extend the basic MR Egger analysis
579 and to tease out the potential effects of SHBG on causality estimates of total and free T, we used multivariable
580 MR Egger (33). LCV reports genetic causality proportion (GCP) as an estimate of causality, under a model
581 where genetic correlation between two traits is mediated by a latent variable having a causal effect on each
582 trait. $GCP=1$ means trait 1 is fully correlated with the latent variable, and hence fully causal to trait 2. A high
583 GCP value and a statistically significant effect support partial genetic causality between the traits, and suggest
584 that interventions targeting trait 1 are likely to affect trait 2. The p -value obtained in the analysis refers to
585 the null hypothesis that the $GCP=0$. A highly significant p -value does not require a high GCP. Positive GCP
586 value indicates causality of trait1 to trait2, whereas a negative value indicates support for causality of trait 2
587 to trait 1. LCV also estimates genetic correlation between the traits. To estimate the how much T could
588 explain sex differences in hemoglobin we calculated $((FT_m-FT_f/SD_{FT_m}) * \beta * SD_{H_m}) / ((H_m-$
589 $H_f)/SD_{H_m})$, where FT = mean free Testosterone, m=males, f=females, SD = standard deviation, β =MR
590 estimate, H=Hemoglobin in UK Biobank based on (60). We applied the LCV, MR-Egger, multivariate MR-Egger
591 and IWV models locally using R 4.0.2. The MR analyses were run using TwoSampleMR (v0.5.2) (61) and
592 MendelianRandomization (v0.5.0) R packages (62). For the traits from public GWAS included in genetic
593 correlation analyses, in 16 out 44 instances we could perform two sample MR (phenotype data not based on
594 UK Biobank samples, Supplementary Table 24). FinnGen represents an independent research cohort from
595 the UK Biobank and thus all FinnGen causality analyses were two-sample MR analyses.

596

597 Acknowledgements / Funding statement

598 The research has been conducted using the UK Biobank Resource under application number 22627. The
599 FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016)
600 and eleven industry partners (AbbVie Inc, AstraZeneca UK Ltd, Biogen MA Inc, Celgene Corporation, Celgene
601 International II Sarl, Genentech Inc, Merck Sharp & Dohme Corp, Pfizer Inc., GlaxoSmithKline, Sanofi, Maze
602 Therapeutics Inc., Janssen Biotech Inc). The funders had no role in study design, data collection and analysis,
603 decision to publish or preparation of the manuscript. The following biobanks are acknowledged for collecting
604 the FinnGen project samples: Auria Biobank (<https://www.auria.fi/biopankki>), THL Biobank
605 (<https://thl.fi/fi/web/thl-biopankki>), Helsinki Biobank (<https://www.terveyskyla.fi/helsinginbiopankki>),
606 Biobank Borealis of Northern Finland (<https://www oulu.fi/university/node/38474>), Finnish Clinical Biobank
607 Tampere (https://www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere),
608 Biobank of Eastern Finland (<https://ita-suomenbiopankki.fi>), Central Finland Biobank
609 (<https://www.ksshp.fi/fi-FI/Potilaalle/Biopankki>), Finnish Red Cross Blood Service Biobank
610 (<https://www.veripalvelu.fi/verenluovutus/biopankkitoiminta>) and Terveystalo Biobank
611 (<https://www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/>). All Finnish Biobanks are
612 members of BBMRI.fi infrastructure (www.bbmri.fi). The Young Finns Study has been financially supported
613 by the Academy of Finland: grants 322098 (T.L), 286284, 134309 (Eye), 126925, 121584, 124282, 129378
614 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State
615 Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals
616 (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular
617 Research; Finnish Cultural Foundation; The Sigrid Juselius Foundation; Tampere Tuberculosis Foundation;
618 Emil Aaltonen Foundation; Yrjo Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; Diabetes
619 Research Foundation of Finnish Diabetes Association; This project has received funding from the European
620 Union's Horizon 2020 research and innovation programme under grant agreements No 848146 for To Aition
621 and grant agreement 755320 for TAXINOMISIS; European Research Council (grant 742927 for MULTIEPIGEN
622 project); Tampere University Hospital Supporting Foundation and Finnish Society of Clinical Chemistry. We

623 greatly thank all UK Biobank, the Young Finns study, and FinnGen participants, as well as the principal
624 investigators, laboratory personnel and data management teams behind these efforts. This work was
625 supported by the Academy of Finland (<https://www.aka.fi/en/>) grants 331671 to N.M., 312072 to Tm.T.,
626 288509 and 319181 to M.P., and 315589 and 320129 to T.T., and Academy of Finland Center of Excellence in
627 Complex Disease Genetics grants 312075 to M.D., 312062 and 336820 to S.R. and 312076 to M.P.), by the
628 Finnish Foundation for Cardiovascular Research (<https://www.sydantutkimussaatio.fi/en/>) (S.R.), the Sigrid
629 Juselius Foundation (<https://sigridjuselius.fi/en/>) (S.R., M.P., T.T.) and University of Helsinki
630 (<https://www.helsinki.fi/en>) HiLIFE Fellow and Grand Challenge grants (S.R. and M.P.), and three-year
631 research project grant (T.T.).

632

633 References

- 634 1. Handelsman DJ, Hirschberg AL, Bermon S. Circulating Testosterone as the Hormonal Basis of Sex
635 Differences in Athletic Performance. *Endocr Rev.* 2018;39(5):803-29.
- 636 2. Tyagi V, Scordo M, Yoon RS, Liporace FA, Greene LW. Revisiting the role of testosterone: Are we
637 missing something? *Rev Urol.* 2017;19(1):16-24.
- 638 3. Davis SR, Wahlin-Jacobsen S. Testosterone in women—the clinical significance. *The Lancet*
639 *Diabetes & Endocrinology.* 2015;3(12):980-92.
- 640 4. Vandeput L, Ohlsson C. Genome-wide association studies on serum sex steroid levels. *Mol Cell*
641 *Endocrinol.* 2014;382(1):758-66.
- 642 5. Kloner RA, Carson C, Dobs A, Kopecky S, Mohler ER. Testosterone and Cardiovascular Disease. *J Am*
643 *Coll Cardiol.* 2016;67(5):545-57.
- 644 6. Gagliano-Jucá T, Basaria S. Testosterone replacement therapy and cardiovascular risk. *Nature*
645 *Reviews Cardiology.* 2019;16(9):555-74.
- 646 7. Bhasin S, Brito JP, Cunningham GR, Hayes FJ, Hodis HN, Matsumoto AM, Snyder PJ, Swerdloff RS, Wu
647 FC, Yialamas MA. Testosterone Therapy in Men With Hypogonadism: An Endocrine Society* Clinical Practice
648 Guideline. *The Journal of Clinical Endocrinology & Metabolism.* 2018;103(5):1715-44.
- 649 8. Celec P, Ostatníková D, Hodosy J. On the effects of testosterone on brain behavioral functions. *Front*
650 *Neurosci.* 2015;9:12-.
- 651 9. Venturelli E, Orenti A, Fabricio ASC, Garrone G, Agresti R, Paolini B, Bonini C, Gion M, Berrino F,
652 Desmedt C, Coradini D, Biganzoli E. Observational study on the prognostic value of testosterone and adiposity
653 in postmenopausal estrogen receptor positive breast cancer patients. *BMC Cancer.* 2018;18(1):651-.
- 654 10. Stattin P, Lumme S, Tenkanen L, Alfthan H, Jellum E, Hallmans G, Thoresen S, Hakulinen T,
655 Luostarinen T, Lehtinen M, Dillner J, Stenman U-H, Hakama M. High levels of circulating testosterone are not
656 associated with increased prostate cancer risk: A pooled prospective study. 2004;108(3):418-24.
- 657 11. Eisenegger C, Naef M, Snozzi R, Heinrichs M, Fehr E. Prejudice and truth about the effect of
658 testosterone on human bargaining behaviour. *Nature.* 2010;463(7279):356-9.
- 659 12. Clark RV, Wald JA, Swerdloff RS, Wang C, Wu FCW, Bowers LD, Matsumoto AM. Large divergence in
660 testosterone concentrations between men and women: Frame of reference for elite athletes in sex-specific
661 competition in sports, a narrative review. 2019;90(1):15-22.

- 662 13. Goldman AL, Bhasin S, Wu FCW, Krishna M, Matsumoto AM, Jasuja R. A Reappraisal of Testosterone's
663 Binding in Circulation: Physiological and Clinical Implications. *Endocr Rev.* 2017;38(4):302-24.
- 664 14. Cunningham SK, Loughlin T, Culliton M, McKenna TJ. The Relationship between Sex Steroids and Sex-
665 Hormone-Binding Globulin in Plasma in Physiological and Pathological Conditions. 1985;22(5):489-97.
- 666 15. Mendel CM. The Free Hormone Hypothesis: A Physiologically Based Mathematical Model*. *Endocr*
667 *Rev.* 1989;10(3):232-74.
- 668 16. Pasquali R. The hypothalamic-pituitary-adrenal axis and sex hormones in chronic stress and obesity:
669 pathophysiological and clinical aspects. *Ann N Y Acad Sci.* 2012;1264(1):20-35.
- 670 17. Trost LW, Mulhall JP. Challenges in Testosterone Measurement, Data Interpretation,
671 and Methodological Appraisal of Interventional Trials. *The journal of sexual medicine.* 2016;13(7):1029-46.
- 672 18. Ohlsson C, Wallaschofski H, Lunetta KL, Stolk L, Perry JRB, Koster A, Petersen A-K, Eriksson J,
673 Lehtimäki T, Huhtaniemi IT, Hammond GL, Maggio M, Coviello AD, Group ES, Ferrucci L, Heier M, Hofman A,
674 Holliday KL, Jansson J-O, Kähönen M, Karasik D, Karlsson MK, Kiel DP, Liu Y, Ljunggren Ö, Lorentzon M,
675 Lyytikäinen L-P, Meitinger T, Mellström D, Melzer D, Miljkovic I, Nauck M, Nilsson M, Penninx B, Pye SR, Vasan
676 RS, Reincke M, Rivadeneira F, Tajar A, Teumer A, Uitterlinden AG, Ulloor J, Viikari J, Völker U, Völzke H,
677 Wichmann HE, Wu T-S, Zhuang WV, Ziv E, Wu FCW, Raitakari O, Eriksson A, Bidlingmaier M, Harris TB, Murray
678 A, de Jong FH, Murabito JM, Bhasin S, Vandenput L, Haring R. Genetic Determinants of Serum Testosterone
679 Concentrations in Men. *PLOS Genetics.* 2011;7(10):e1002313.
- 680 19. Harris JA, Vernon PA, Boomsma DIJBG. The Heritability of Testosterone: A Study of Dutch Adolescent
681 Twins and Their Parents. 1998;28(3):165-71.
- 682 20. Ruth KS, Campbell PJ, Chew S, Lim EM, Hadlow N, Stuckey BGA, Brown SJ, Feenstra B, Joseph J,
683 Surdulescu GL, Zheng HF, Richards JB, Murray A, Spector TD, Wilson SG, Perry JRB. Genome-wide association
684 study with 1000 genomes imputation identifies signals for nine sex hormone-related phenotypes. *Eur J Hum*
685 *Genet.* 2016;24(2):284-90.
- 686 21. Koenis MMG, Brouwer RM, van Baal GCM, van Soelen ILC, Peper JS, van Leeuwen M, Delemarre-van
687 de Waal HA, Boomsma DI, Hulshoff Pol HE. Longitudinal Study of Hormonal and Physical Development in
688 Young Twins. *The Journal of Clinical Endocrinology & Metabolism.* 2013;98(3):E518-E27.
- 689 22. Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, Beaumont RN, Wittemans L, Martin
690 S, Busch AS, Erzurumluoglu AM, Hollis B, O'Mara TA, McCarthy MI, Langenberg C, Easton DF, Wareham NJ,
691 Burgess S, Murray A, Ong KK, Frayling TM, Perry JRB, The Endometrial Cancer Association C. Using human
692 genetics to understand the disease impacts of testosterone in men and women. *Nat Med.* 2020;26(2):252-8.
- 693 23. Flynn E, Tanigawa Y, Rodriguez F, Altman RB, Sinnott-Armstrong N, Rivas MA. Sex-specific genetic
694 effects across biomarkers. *Eur J Hum Genet.* 2020.
- 695 24. Sinnott-Armstrong N, Naqvi S, Rivas M, Pritchard JK. GWAS of three molecular traits highlights core
696 genes and pathways alongside a highly polygenic background. *eLife.* 2021;10:e58615.
- 697 25. Mohammadi-Shemirani P, Chong M, Pigeyre M, Morton RW, Gerstein HC, Paré G. Effects of lifelong
698 testosterone exposure on health and disease using Mendelian randomization. *eLife.* 2020;9:e58914.
- 699 26. Hayes BL, Robinson T, Kar SP, Ruth KS, Tsilidis KK, Frayling TM, Murray A, Martin RM, Lawlor DA,
700 Richmond RC. Do sex hormones confound or mediate the effect of chronotype on breast and prostate
701 cancer? A Mendelian randomization study. 2021:2021.04.20.21255783.
- 702 27. Harrison S, Davies NM, Howe LD, Hughes A. Testosterone and socioeconomic position: Mendelian
703 Randomization in 306,248 men and women participants of UK Biobank. 2020:2020.11.06.20226944.
- 704 28. Vilhjálmsdóttir Bjarni J, Yang J, Finucane Hilary K, Gusev A, Lindström S, Ripke S, Genovese G, Loh P-R,
705 Bhatia G, Do R, Hayeck T, Won H-H, Ripke S, Neale Benjamin M, Corvin A, Walters James TR, Farh K-H,
706 Holmans Peter A, Lee P, Bulik-Sullivan B, Collier David A, Huang H, Pers Tune H, Agartz I, Agerbo E, Albus M,
707 Alexander M, Amin F, Bacanu Silviu A, Begemann M, Belliveau Richard A, Jr., Bene J, Bergen Sarah E,
708 Bevilacqua E, Bigdeli Tim B, Black Donald W, Bruggeman R, Buccola Nancy G, Buckner Randy L, Byerley W,
709 Cahn W, Cai G, Champion D, Cantor Rita M, Carr Vaughan J, Carrera N, Catts Stanley V, Chambert Kimberly D,
710 Chan Raymond CK, Chen Ronald YL, Chen Eric YH, Cheng W, Cheung Eric FC, Chong Siow A, Cloninger CR,
711 Cohen D, Cohen N, Cormican P, Craddock N, Crowley James J, Curtis D, Davidson M, Davis Kenneth L,
712 Degenhardt F, Del Favero J, DeLisi Lynn E, Demontis D, Dikeos D, Dinan T, Djurovic S, Donohoe G, Drapeau E,
713 Duan J, Dudbridge F, Durmishi N, Eichhammer P, Eriksson J, Escott-Price V, Essioux L, Fanous Ayman H, Farrell

714 Martillas S, Frank J, Franke L, Freedman R, Freimer Nelson B, Friedl M, Friedman Joseph I, Fromer M,
715 Genovese G, Georgieva L, Gershon Elliot S, Giegling I, Giusti-Rodriguez P, Godard S, Goldstein Jacqueline I,
716 Golimbet V, Gopal S, Gratten J, Grove J, de Haan L, Hammer C, Hamshere Marian L, Hansen M, Hansen T,
717 Haroutunian V, Hartmann Annette M, Henskens Frans A, Herms S, Hirschhorn Joel N, Hoffmann P, Hofman
718 A, Hollegaard Mads V, Hougaard David M, Ikeda M, Joa I, Julia A, Kahn Rene S, Kalaydjieva L, Karachanak-
719 Yankova S, Karjalainen J, Kavanagh D, Keller Matthew C, Kelly Brian J, Kennedy James L, Khrunin A, Kim Y,
720 Klovins J, Knowles James A, Konte B, Kucinskas V, Kucinskiene Zita A, Kuzelova-Ptackova H, Kahler Anna K,
721 Laurent C, Keong Jimmy Lee C, Lee SH, Legge Sophie E, Lerer B, Li M, Li T, Liang K-Y, Lieberman J, Limborska
722 S, Loughland Carmel M, Lubinski J, Linnqvist J, Macek M, Jr., Magnusson Patrik KE, Maher Brion S, Maier W,
723 Mallet J, Marsal S, Mattheisen M, Mattingdal M, McCarley Robert W, McDonald C, McIntosh Andrew M,
724 Meier S, Meijer Carin J, Melegh B, Melle I, Meshulam-Gately Raquelle I, Metspalu A, Michie Patricia T, Milani
725 L, Milanova V, Mokrab Y, Morris Derek W, Mors O, Mortensen Preben B, Murphy Kieran C, Murray Robin M,
726 Myin-Germeys I, Müller-Myhsok B, Nelis M, Nenadic I, Nertney Deborah A, Nestadt G, Nicodemus Kristin K,
727 Nikitina-Zake L, Nisenbaum L, Nordin A, O'Callaghan E, O'Dushlaine C, O'Neill FA, Oh S-Y, Olincy A, Olsen L,
728 Van Os J, Pantelis C, Papadimitriou George N, Papiol S, Parkhomenko E, Pato Michele T, Paunio T, Pejovic-
729 Milovancevic M, Perkins Diana O, Pietilinen O, Pimm J, Pocklington Andrew J, Powell J, Price A, Pulver Ann E,
730 Purcell Shaun M, Quested D, Rasmussen Henrik B, Reichenberg A, Reimers Mark A, Richards Alexander L,
731 Roffman Joshua L, Roussos P, Ruderfer Douglas M, Salomaa V, Sanders Alan R, Schall U, Schubert Christian R,
732 Schulze Thomas G, Schwab Sibylle G, Scolnick Edward M, Scott Rodney J, Seidman Larry J, Shi J, Sigurdsson E,
733 Silagadze T, Silverman Jeremy M, Sim K, Slominsky P, Smoller Jordan W, So H-C, Spencer Chris CA, Stahl Eli A,
734 Stefansson H, Steinberg S, Stogmann E, Straub Richard E, Strengman E, Strohmaier J, Stroup TS, Subramaniam
735 M, Suvisaari J, Svrakic Dragan M, Szatkiewicz Jin P, Sderman E, Thirumalai S, Toncheva D, Tooney Paul A,
736 Tosato S, Veijola J, Waddington J, Walsh D, Wang D, Wang Q, Webb Bradley T, Weiser M, Wildenauer
737 Dieter B, Williams Nigel M, Williams S, Witt Stephanie H, Wolen Aaron R, Wong Emily HM, Wormley
738 Brandon K, Wu Jing Q, Xi Hualin S, Zai Clement C, Zheng X, Zimprich F, Wray Naomi R, Stefansson K, Visscher
739 Peter M, Adolfsson R, Andreassen Ole A, Blackwood Douglas HR, Bramon E, Buxbaum Joseph D, Børglum
740 Anders D, Cichon S, Darvasi A, Domenici E, Ehrenreich H, Esko T, Gejman Pablo V, Gill M, Gurling H, Hultman
741 Christina M, Iwata N, Jablensky Assen V, Jonsson Erik G, Kendler Kenneth S, Kirov G, Knight J, Lencz T,
742 Levinson Douglas F, Li Qingqin S, Liu J, Malhotra Anil K, McCarroll Steven A, McQuillin A, Moran Jennifer L,
743 Mortensen Preben B, Mowry Bryan J, Nthen Markus M, Ophoff Roel A, Owen Michael J, Palotie A, Pato
744 Carlos N, Petryshen Tracey L, Posthuma D, Rietschel M, Riley Brien P, Rujescu D, Sham Pak C, Sklar P, St. Clair
745 D, Weinberger Daniel R, Wendland Jens R, Werge T, Daly Mark J, Sullivan Patrick F, O'Donovan Michael C,
746 Kraft P, Hunter DJ, Adank M, Ahsan H, Aittomäki K, Baglietto L, Berndt S, Blomquist C, Canzian F, Chang-
747 Claude J, Chanock SJ, Crisponi L, Czene K, Dahmen N, Silva IdS, Easton D, Eliassen AH, Figueroa J, Fletcher O,
748 Garcia-Closas M, Gaudet MM, Gibson L, Haiman CA, Hall P, Hazra A, Hein R, Henderson BE, Hofman A, Hopper
749 JL, Irwanto A, Johansson M, Kaaks R, Kibriya MG, Lichtner P, Lindström S, Liu J, Lund E, Makalic E, Meindl A,
750 Meijers-Heijboer H, Müller-Myhsok B, Muranen TA, Nevanlinna H, Peeters PH, Peto J, Prentice RL, Rahman
751 N, Sánchez MJ, Schmidt DF, Schmutzler RK, Southey MC, Tamimi R, Travis R, Turnbull C, Uitterlinden AG, van
752 der Luijt RB, Waisfisz Q, Wang Z, Whittemore AS, Yang R, Zheng W, Kathiresan S, Pato M, Pato C, Tamimi R,
753 Stahl E, Zaitlen N, Pasaniuc B, Belbin G, Kenny EE, Schierup MH, De Jager P, Patsopoulos NA, McCarroll S, Daly
754 M, Purcell S, Chasman D, Neale B, Goddard M, Visscher PM, Kraft P, Patterson N, Price AL. Modeling Linkage
755 Disequilibrium Increases Accuracy of Polygenic Risk Scores. *The American Journal of Human Genetics*.
756 2015;97(4):576-92.
757 29. Mars NJ, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, Ahola-Olli A, Kurki M,
758 Karjalainen J, Palta P, Neale BM, Daly M, Salomaa V, Palotie A, Widén E, Ripatti S. Polygenic and clinical risk
759 scores and their impact on age at onset of cardiometabolic diseases and common cancers. 2019:727057.
760 30. Richardson TG, Harrison S, Hemani G, Davey Smith G. An atlas of polygenic risk score associations to
761 highlight putative causal relationships across the human phenome. *eLife*. 2019;8:e43657.
762 31. O'Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and
763 complex traits. *Nat Genet*. 2018;50(12):1728-34.
764 32. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger
765 method. *Eur J Epidemiol*. 2017;32(5):377-89.

- 766 33. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian
767 randomization to correct for both measured and unmeasured pleiotropy. *Stat Med*. 2017;36(29):4705-18.
- 768 34. Traylor M, Rutten-Jacobs LCA, Holliday EG, Malik R, Sudlow C, Rothwell PM, Maguire JM, Koblar SA,
769 Bevan S, Boncoraglio G, Dichgans M, Levi C, Lewis CM, Markus HS. Differences in Common Genetic
770 Predisposition to Ischemic Stroke by Age and Sex. 2015;46(11):3042-7.
- 771 35. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, Duncan L, Perry JRB, Patterson N,
772 Robinson EB, Daly MJ, Price AL, Neale BM, ReproGen C, Psychiatric Genomics C, Genetic Consortium for
773 Anorexia Nervosa of the Wellcome Trust Case Control C. An atlas of genetic correlations across human
774 diseases and traits. *Nat Genet*. 2015;47(11):1236-41.
- 775 36. van Rheenen W, Peyrot WJ, Schork AJ, Lee SH, Wray NR. Genetic correlations of polygenic disease
776 traits: from theory to practice. *Nature reviews Genetics*. 2019;20(10):567-81.
- 777 37. Choi MH, Yoo YS, Chung BC. Biochemical Roles of Testosterone and Epitestosterone to 5 α -Reductase
778 as Indicators of Male-Pattern Baldness. *J Invest Dermatol*. 2001;116(1):57-61.
- 779 38. Gann PH, Hennekens CH, Ma J, Longcope C, Stampfer MJ. Prospective Study of Sex Hormone Levels
780 and Risk of Prostate Cancer. *JNCI: Journal of the National Cancer Institute*. 1996;88(16):1118-26.
- 781 39. Brand JS, van der Schouw YT, Dowsett M, Folkert E, Luben RN, Wareham NJ, Khaw KT. Testosterone,
782 SHBG and differential white blood cell count in middle-aged and older men. *Maturitas*. 2012;71(3):274-8.
- 783 40. Palmisano BT, Zhu L, Eckel RH, Stafford JM. Sex differences in lipid and lipoprotein metabolism.
784 *Molecular Metabolism*. 2018;15:45-55.
- 785 41. McHenry J, Carrier N, Hull E, Kabbaj M. Sex differences in anxiety and depression: Role of
786 testosterone. *Front Neuroendocrinol*. 2014;35(1):42-57.
- 787 42. O'Connor DB, Archer J, Wu FCW. Effects of Testosterone on Mood, Aggression, and Sexual Behavior
788 in Young Men: A Double-Blind, Placebo-Controlled, Cross-Over Study. *The Journal of Clinical Endocrinology
789 & Metabolism*. 2004;89(6):2837-45.
- 790 43. Snyder PJ, Kopperdahl DL, Stephens-Shields AJ, Ellenberg SS, Cauley JA, Ensrud KE, Lewis CE, Barrett-
791 Connor E, Schwartz AV, Lee DC, Bhasin S, Cunningham GR, Gill TM, Matsumoto AM, Swerdloff RS, Basaria S,
792 Diem SJ, Wang C, Hou X, Cifelli D, Dougar D, Zeldow B, Bauer DC, Keaveny TM. Effect of Testosterone
793 Treatment on Volumetric Bone Density and Strength in Older Men With Low Testosterone: A Controlled
794 Clinical Trial. *JAMA Intern Med*. 2017;177(4):471-9.
- 795 44. Roy CN, Snyder PJ, Stephens-Shields AJ, Artz AS, Bhasin S, Cohen HJ, Farrar JT, Gill TM, Zeldow B, Cella
796 D, Barrett-Connor E, Cauley JA, Crandall JP, Cunningham GR, Ensrud KE, Lewis CE, Matsumoto AM, Molitch
797 ME, Pahor M, Swerdloff RS, Cifelli D, Hou X, Resnick SM, Walston JD, Anton S, Basaria S, Diem SJ, Wang C,
798 Schrier SL, Ellenberg SS. Association of Testosterone Levels With Anemia in Older Men: A Controlled Clinical
799 Trial. *JAMA Intern Med*. 2017;177(4):480-90.
- 800 45. Connallon T, Matthews G. Cross-sex genetic correlations for fitness and fitness components:
801 Connecting theoretical predictions to empirical patterns. *Evol Lett*. 2019;3(3):254-62.
- 802 46. Vihma V, Naukkarinen J, Turpeinen U, Hämäläinen E, Kaprio J, Rissanen A, Heinonen S, Hakkarainen
803 A, Lundbom J, Lundbom N, Mikkola TS, Tikkanen MJ, Pietiläinen KH. Metabolism of sex steroids is influenced
804 by acquired adiposity—A study of young adult male monozygotic twin pairs. *The Journal of Steroid
805 Biochemistry and Molecular Biology*. 2017;172:98-105.
- 806 47. Eriksson J, Haring R, Grarup N, Vandenput L, Wallaschofski H, Lorentzen E, Hansen T, Mellström D,
807 Pedersen O, Nauck M, Lorentzon M, Nystrup Husemoen LL, Völzke H, Karlsson M, Baumeister SE, Linneberg
808 A, Ohlsson C. Causal relationship between obesity and serum testosterone status in men: A bi-directional
809 mendelian randomization analysis. *PLoS One*. 2017;12(4):e0176277.
- 810 48. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in
811 epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89-R98.
- 812 49. Suderman M, Simpkin A, Sharp G, Gaunt T, Lyttleton O, McArdle W, Ring S, Davey Smith G, Relton C.
813 Sex-associated autosomal DNA methylation differences are wide-spread and stable throughout childhood.
814 2017:118265.
- 815 50. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O,
816 O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J.
817 The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203-9.

- 818 51. Vermeulen A, Verdonck L, Kaufman JM. A Critical Evaluation of Simple Methods for the Estimation of
819 Free Testosterone in Serum. *The Journal of Clinical Endocrinology & Metabolism*. 1999;84(10):3666-72.
- 820 52. Emadi-Konjin P, Bain J, Bromberg IL. Evaluation of an algorithm for calculation of serum
821 "Bioavailable" testosterone (BAT). *Clin Biochem*. 2003;36(8):591-6.
- 822 53. Sinnott-Armstrong N, Tanigawa Y, Amar D, Mars NJ, Aguirre M, Venkataraman GR, Wainberg M, Ollila
823 HM, Pirruccello JP, Qian J, Shcherbina A, Rodriguez F, Assimes TL, Agarwala V, Tibshirani R, Hastie T, Ripatti
824 S, Pritchard JK, Daly MJ, Rivas MA. Genetics of 38 blood and urine biomarkers in the UK Biobank.
825 2019:660506.
- 826 54. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of
827 genetic associations with FUMA. *Nature Communications*. 2017;8(1):1826.
- 828 55. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene
829 and protein annotation. *Nucleic Acids Res*. 2015;44(D1):D457-D62.
- 830 56. Pickrell JK, Berisa T, Liu JZ, Séguérel L, Tung JY, Hinds DA. Detection and interpretation of shared
831 genetic influences on 42 human traits. *Nat Genet*. 2016;48(7):709-17.
- 832 57. Benner C, Havulinna AS, Järvelin MR, Salomaa V, Ripatti S, Pirinen M. Prospects of Fine-Mapping
833 Trait-Associated Genomic Regions by Using Summary Statistics from Genome-wide Association Studies. *Am*
834 *J Hum Genet*. 2017;101(4):539-51.
- 835 58. Leinonen JT, Chen Y-C, Pennonen J, Lehtonen L, Junna N, Tukiainen T, Panula P, Widén E. LIN28B
836 affects gene expression at the hypothalamic-pituitary axis and serum testosterone levels. *Sci Rep*.
837 2019;9(1):18060.
- 838 59. Sivakumaran S, Agakov F, Theodoratou E, Prendergast James G, Zgaga L, Manolio T, Rudan I,
839 McKeigue P, Wilson James F, Campbell H. Abundant Pleiotropy in Human Complex Diseases and Traits. *The*
840 *American Journal of Human Genetics*. 2011;89(5):607-18.
- 841 60. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F,
842 Kostadima MA, Lambourne JJ, Sivapalaratnam S, Downes K, Kundu K, Bomba L, Berentsen K, Bradley JR,
843 Daugherty LC, Delaneau O, Freson K, Garner SF, Grassi L, Guerrero J, Haimel M, Janssen-Megens EM, Kaan A,
844 Kamat M, Kim B, Mandoli A, Marchini J, Martens JHA, Meacham S, Megy K, O'Connell J, Petersen R, Sharifi N,
845 Sheard SM, Staley JR, Tuna S, van der Ent M, Walter K, Wang S-Y, Wheeler E, Wilder SP, Iotchkova V, Moore
846 C, Sambrook J, Stunnenberg HG, Di Angelantonio E, Kaptoge S, Kuijpers TW, Carrillo-de-Santa-Pau E, Juan D,
847 Rico D, Valencia A, Chen L, Ge B, Vasquez L, Kwan T, Garrido-Martín D, Watt S, Yang Y, Guigo R, Beck S, Paul
848 DS, Pastinen T, Bujold D, Bourque G, Frontini M, Danesh J, Roberts DJ, Ouwehand WH, Butterworth AS,
849 Soranzo N. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease.
850 *Cell*. 2016;167(5):1415-29.e19.
- 851 61. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J,
852 Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM, Relton C, Martin RM, Davey Smith G,
853 Gaunt TR, Haycock PC. The MR-Base platform supports systematic causal inference across the human
854 phenome. *eLife*. 2018;7:e34408.
- 855 62. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian
856 randomization analyses using summarized data. *Int J Epidemiol*. 2017;46(6):1734-9.

857