Genetic Determinants of Circulating Estrogen Levels, and Evidence of a Causal Effect of Estradiol on Bone Density in Men

Anna L. Eriksson[1*], John R.B. Perry[2,3*], Andrea D. Coviello[4], Graciela E. Delgado[5], Luigi Ferrucci[6], Andrew R. Hoffman[7], Ilpo T. Huhtaniemi[8,9], M. Arfan Ikram[10], Magnus K. Karlsson[11], Marcus E. Kleber[5], Gail A. Laughlin[12], Yongmei Liu[13], Mattias Lorentzon[1,14], Kathryn L. Lunetta[15,16], Dan Mellström[1,14], Joanne M. Murabito[17], Anna Murray[3], Maria Nethander[1], Carrie M. Nielson[18], Inga Prokopenko[19,20], Stephen R. Pye[21], Leslie J. Raffel[22], Fernando Rivadeneira[10,23], Priya Srikanth[18], Lisette Stolk[23], Alexander Teumer[24,25], Thomas G. Travison[26], André G. Uitterlinden[10,23], Dhananjay Vaidya[27], Dirk Vanderschueren[28], Joseph M. Zmuda[29], Winfried März[30,31], Eric S. Orwoll[32], Pamela Ouyang[27], Liesbeth Vandenput[1], Frederick CW. Wu[33], Frank H. de Jong[23‖], Shalender Bhasin[34‖], Douglas P. Kiel[16,26‖], Claes Ohlsson[1‖],

[1]Centre for Bone and Arthritis Research, Institute of Medicine, Sahlgrenska University Hospital, 413 45 Gothenburg, Sweden

[2]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Box 285, Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge CB20QQ, UK

[3]University of Exeter Medical School, University of Exeter, Exeter EX1 2LU, UK

[4]Duke University School of Medicine, Durham, NC 27710, USA

[5]V[th] Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim 68167, Germany

[6]Longitudinal Studies Section, Clinical Research Branch, Gerontology Research Center, National Institute on Aging, Baltimore, MD 21224, USA

[7]Division of Endocrinology, Stanford University School of Medicine, Stanford, CA 94305, USA

[8]Department of Surgery and Cancer, Imperial College London, Hammersmith Campus, London, W12 0NN, UK

[9]Department of Physiology, Institute of Biomedicine, University of Turku, Turku 20100, Finland

[10]Department of Epidemiology, Erasmus MC, Rotterdam 3000CA, Netherlands

[11]Department of Orthopaedics and Clinical Sciences, Skåne University Hospital, Lund University, 217 74 Malmö, Sweden

[12]Family Medicine and Public Health, University of California-San Diego, San Diego, CA 92093, USA

[13]Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA

[14]Geriatric Medicine, Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, University of Gothenburg and Geriatric Medicine, Sahlgrenska University Hospital, 43180 Mölndal, Sweden

[15]Boston University School of Public Health, Boston, MA 02118, USA

[16]Framingham Heart Study, Framingham, MA 07012, USA

[17]Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA 02118, USA

[18]School of Public Health, Oregon Health & Science University, Portland, OR 97239, USA

[19]Department of Genomics of Common Disease, School of Public Health, Imperial College London, London, W12 0NN, UK

[20]Hammersmith Hospital, Burlington Danes Building, Du Cane Road, London, W12 0NN, UK

[21]Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, The University of Manchester, Manchester Academic Health Science Centre, Oxford Road, Manchester, M13 9PT, UK

[22]Division of Genetic and Genomic Medicine, Department of Pediatrics, University of California, Irvine, Orange, CA 92868, USA

[23]Department of Internal Medicine, Erasmus MC, Rotterdam 3000CA, Netherlands

[24]Institute for Community Medicine, University Medicine Greifswald, 17475, Greifswald, Germany

[25]Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17475, Greifswald, Germany

[26]Institute for Aging Research, Hebrew Senior Life and Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA 02131, USA

[27]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287 USA

[28]Department of Clinical and Experimental Medicine, Katholieke Universiteit Leuven, Laboratory of Clinical and Experimental Endocrinology, Leuven, B-3000, Belgium

[29]Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA 15261, USA

[30]Synlab Academy, Synlab Holding Deutschland GmbH, Mannheim 68161, Germany

[31]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, 8036 Graz, Austria

[32]Bone & Mineral Unit, Oregon Health & Science University, Portland, OR 97239, USA

[33]Andrology Research Unit, Centre for Endocrinology and Diabetes, Institute of Human Development, Faculty of Medical and Human Sciences, The University of Manchester, Central Manchester University Hospitals NHS Foundation Trust, Old St Mary's Building, Hathersage RoadManchester, M13 9WL, UK

[34]Research Program in Men's Health: Aging and Metabolism, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115, USA

* These authors contributed equally to this work

⁋ These authors were joint senior authors on this work.

Précis: *CYP19A1* was the main regulator of estrogen levels in this GWAS, with additional loci on chromosome X and in *TRIM4* and *CYP11B1/B2*. Findings in the study strengthen the importance of E2 for bone health.

Short title: Genetic Determinants of Estrogen Levels in Men

***Corresponding author and person to whom reprint requests should be made*:**

Claes Ohlsson

Centre for Bone and Arthritis Research

Klin Farm Lab, Vita Stråket 11
Dept. of Internal Medicine and Clinical Nutrition,
Sahlgrenska University Hospital,
SE-41345 Gothenburg, Sweden
E-mail: claes.ohlsson@medic.gu.se

Funding information

This work was supported by the Swedish Research Council, the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, Knut and Alice Wallenberg Foundation, the Torsten and Ragnar Söderberg's Foundation and the Novo Nordisk Foundation.

Disclosure summary

Vaidya Dhananjay is a consultant for Consumable Science Inc. Ilpo T Huhtaniemi received grants from Ferring Pharmaceutical, and did consultation for Novartis and Takeda. Winfried März is employed with Synlab Holding Deutschland GmbH. The other authors have no disclosures.

Abstract

Context

Serum estradiol (E2) and estrone (E1) levels exhibit substantial heritability. No genome-wide association study (GWAS) of estrogen levels has been performed in men of European origin.

Objective

To investigate the genetic regulation of serum E2 and E1 in men.

Design, setting and participants

GWAS in 11,097 men of European origin from nine epidemiological cohorts.

Main Outcome Measures

Genetic determinants of serum E2 and E1 levels.

Results

Variants in/near *CYP19A1* demonstrated the strongest evidence for association with E2, resolving to three independent signals. Two additional independent signals were found on the X chromosome; *FAM9B*, rs5934505 (p-value 3.4×10^{-8}) and *Xq27.3*, rs5951794 (p-value 3.1×10^{-10}). E1 signals were found in *CYP19A1* (rs2899472, p-value 5.5×10^{-23}), in *TRIM4* (rs17277546, $p = 5.8 \times 10^{-14}$) and in *CYP11B1/B2* (rs10093796, p-value 1.2×10^{-8}).

E2 signals in *CYP19A1* and *FAM9B* were associated with bone mineral density (BMD). Mendelian Randomization analysis suggested a causal effect of serum E2 on BMD in men. 1 pg/ml genetically increased E2 was associated with a 0.048 SD increase in lumbar spine BMD (p-value 2.8×10^{-12}).

In men and women combined, *CYP19A1* alleles associated with higher E2 levels were associated with lower degrees of insulin resistance.

Conclusions

Our findings confirm that *CYP19A1* is an important genetic regulator of E2 and E1 levels, and strengthen the causal importance of E2 for bone health in men. We also report 2 new independent loci on the X-chromosome for E2, one new locus each in *TRIM4* and *CYP11B1/B2*, for E1.

***Keywords*:** estradiol, estrone, GWAS, men, BMD, insulin sensitivity

Introduction

17 ß-estradiol (E2) and estrone (E1) are the major biologically active estrogens in men. E2 is more potent than E1. Aromatase, encoded by the *CYP19A1* gene, is the key enzyme responsible for the final step in the synthesis of both E2 and E1. E2 is formed from aromatization of testosterone, and E1 is formed from aromatization of androstenedione. E2 can also be formed from conversion of E1 by 17ß-hydroxysteroid dehydrogenase (1).

In men, the circulating levels of E2 and E1 are determined by both genetic and environmental factors. The heritability for E2 in men has been estimated to be ~30-45% and for E1 ~40% (2, 3). Early studies of the genetic regulation of circulating E2 and E1 levels were hampered by their small size and the use of immunoassays with poor specificity, precision and accuracy at lower concentrations. However, in 2010 Orwoll and colleagues performed a large study of 5,000 elderly men of European, Asian and African origin in Sweden, the United States, Hong Kong and Tobago (4). Serum sex steroid levels were measured using gas chromatography – mass spectrometry (GC-MS), thereby avoiding the previously mentioned problems with immunoassays. In addition to geographical differences in E2 and E1 levels, suggestive of environmental influences, they also found racial differences. Both E2 and E1 levels, as well as the estradiol to testosterone and estrone to androstenedione ratios, were higher in Black than in Asian and Caucasian men (4). These data suggested that genetically determined differences in aromatase activity among Black, Asian and Caucasian men might be responsible for the observed racial differences in E2 and E1 levels.

We made a first attempt to find genetic loci involved in the determination of estrogen levels in men by analyzing 604 SNPs in 50 candidate sex steroid-related genes (5). In a screening cohort, the *CYP19A1* SNP rs2470152 showed the most significant association with E2 levels

measured by GC-MS. This was confirmed in two replication cohorts. Rs2470152 was also significantly associated with E1 levels in all three cohorts (n=5531) (5).

Meta-analyses of genome-wide association studies (GWAS) enable a comprehensive analysis of the whole genome in a large number of subjects. Chen and colleagues performed a GWAS in 3,495 Chinese men, where E2 concentrations were determined using an immunoassay. They found two independent SNPs in the *CYP19A1* gene to be associated with E2 levels (rs2414095 and rs2445762) (6). These findings further strengthened the evidence for a major role of *CYP19A1* in the regulation of serum E2 levels in men, but due to the relatively small sample size and low power, genetic loci in other regions of the genome could have been missed. To date no GWAS has been performed in men of European origin. In women, a smaller GWAS meta-analysis of 1,583 postmenopausal women found no genome-wide significant SNPs. Among variants that were suggestively associated with E2, several were located at the *CYP19A1* locus (7).

Both E2 and testosterone regulate bone mass (8). Studies of men with non-functional estrogen receptor alpha (ERα) (9), and inactivating mutations of the *CYP19A1* gene (10), have demonstrated that estrogens are important for peak bone mass acquisition in men. Population based studies have shown that in men, low serum levels of E2 are associated with a lower bone mineral density (BMD), higher rates of bone loss and an increased risk of fractures (8, 11-14). Some studies also show a smaller contribution of testosterone to BMD in men (8, 11). The relative contribution of androgens versus estrogens in the regulation of bone mass in men remains incompletely understood, and studies showing evidence of a causal effect of serum E2 on BMD in men are still sparse (15).

Mendelian randomization is a method used to strengthen or refute the causality of a biomarker, such as E2, and an outcome measure of interest, such as BMD, when a

randomized controlled trial is not possible. Mendelian randomization uses genetic data and relies on the principle that due to the random assortment of genetic variants at conception, these genetic variants are independent of many factors that bias observational studies, such as confounding and reverse causation. Therefore, if a biomarker is etiologically involved in an outcome measure, the genetic factors that influence the biomarker will also influence the outcome measure (16). To date, no Mendelian randomization has been performed to investigate causality between E2 levels and BMD in men.

Case reports of men with aromatase deficiency due to an inactivating mutation of the *CYP19A1* gene, mechanistic animal studies and clinical studies also suggest that estrogen signaling through ERα is important for insulin sensitivity in men (17-23). Thus, genetic factors regulating estrogen levels may also be of relevance for the regulation of insulin sensitivity in men.

Here we present the results of the first GWAS of estrogen levels combining several population-based cohorts of men of European origin. We also present results of our analyses of the association of resultant genome wide significant associations with two major estrogen related traits –bone mineral density and insulin sensitivity.

Methods

Study samples

The discovery stage of the E2 GWAS included 11,097 men of European origin drawn from nine epidemiological cohorts: the Framingham Heart Study (FHS), the Gothenburg Osteoporosis and Obesity Determinants (GOOD) study, the Invecchiare in Chianti (InCHIANTI) study, the LUdwigshafen RIsk and Cardiovascular Health (LURIC) study, the Multi-Ethnic Study of Atherosclerosis (MESA) study, the Osteoporotic Fractures in Men (MrOS) Sweden Gothenburg study, the MrOS Sweden Malmö study, the MrOS US Study, and the Rotterdam 1 (RS1) study. Replication of one SNP displaying considerable heterogeneity in genome wide significant fixed effect models, but nominal significance only in random effects models, was performed in the European Male Ageing Study (EMAS, n=1,641). EMAS is a cohort of men predominantly of European origin with only 0.62 % (n=21) of the sample used here being of non-European descent.

The discovery stage of the E1 GWAS included 7,570 men of European origin drawn from six of the above-mentioned cohorts: FHS, GOOD, MrOS Sweden Gothenburg, MrOS Sweden Malmö, MrOS US and RS1.

Exclusion criteria included chemical or surgical castration and/or medications affecting sex hormones such as steroid 5-alpha reductase inhibitors, and sex hormone antagonists. All studies were approved by local ethics committees and all participants provided written informed consent. Characteristics of the study samples and detailed descriptions of the participating cohorts, genotyping, quality control and imputation procedures are provided in the Supplementary Appendix and in Supplemental Tables 1, 2 and 3.

Sex hormone measurements

In six discovery cohorts (FHS, GOOD, MrOS Sweden Gothenburg, MrOS Sweden Malmö, and MrOs US), measurements of E1 and E2 were performed using either the GC-MS or the liquid chromatography tandem mass spectrometry (LC-MS/MS) technique. In the remaining discovery cohorts (LURIC, InCHIANTI, MESA and RS-1) measurements were performed using immunoassays. In the replication cohort (EMAS), E2 was measured using the GC-MS technique. Methods for all measurements are given in the Supplementary Appendix.

Genotyping and statistical analyses

Nine discovery and one replication study populations were genotyped using a variety of genotyping platforms including Illumina (HumanHap 550k, 610k, 1M-Duo, Omni1-Quad, Omni express) and Affymetrix (500K Dual GeneChip + 50K gene-centered MIP set, Array 6.0) (Supplemental Table 2). To increase genomic coverage and allow the evaluation of the same SNPs across as many study populations as possible, each study imputed genotype data based on the HapMap CEU Build 36. Algorithms were used to infer unobserved genotypes in a probabilistic manner using either MACH (http://www.sph.umich.edu/csg/abecasis/MACH), or IMPUTE2 (24). We analyzed only those SNPs (genotyped or imputed) which had a minor allele frequency of >0.01 and an imputation quality of ≥0.3. The X chromosome was available for analysis in 6 cohorts (FHS, GOOD, LURIC, MrOS Sweden Gothenburg, MrOS Sweden Malmö and MrOS US), in this study. Imputations of the X-chromosome were performed in all of these cohorts except MrOS US.

Altogether, ~2.5 million SNPs were tested for association with serum E2 and E1 in the discovery stage. Genome-wide association analyses were performed using an additive genetic linear regression model adjusted for: 1) age and BMI (E2 and E1) or: 2) age, BMI, testosterone and SHBG (E2 only), in each of the discovery cohorts. In FHS, a linear mixed

effect model with a random effect to account for relationships was used. Imputed genotypes were analyzed in all cohorts taking the genotype uncertainties into account. The meta-analyses were performed in the METAL software (https://www.sph.umich.edu/csg/abecasis/MACH), using an inverse-variance weighted fixed effect model. Random effects models were used when fixed effect models displayed heterogeneity defined as an I^2-value > 50% (25). These models were calculated using the R-package (http://www.r-project.org). A threshold of $p < 5\times10^{-8}$ was established *a priori* as the level for genome-wide significance in the discovery analyses (26).

Approximate conditional analyses for E2 and E1 were performed using the Genome-wide Complex Trait Analysis (GCTA) software (27), and the genotypes of the EPIC Norfolk study cohort used as a reference panel to estimate patterns of Linkage Disequilibrium (28). The GC-corrected and quality control filtered meta-analysis results and a condition list containing the lead SNPs of the final loci were used as input for the conditional analysis. An additional association was declared when the conditional P-value was below the genome-wide significance threshold. Subsequently, this SNP was added to the list of conditional analysis SNPs and the conditional analysis was performed again in a stepwise fashion until no additional significant independent associations were found.

Gene expression analyses

We analyzed associations between identified SNPs associated with serum estrogen levels and gene expression in the eQTL dataset generated by the GTEx Consortium (version 6p), which was obtained from http://www.gtexportal.org/ (29)

Associations with testosterone

Associations with serum testosterone concentrations were retrieved from the discovery dataset of our previously published GWAS of testosterone levels (30).

Associations with other traits

We hypothesized, based on data in the literature, that our genome wide significant SNPs and secondary signals from conditional analyses could be associated with BMD and/or insulin sensitivity. To test these hypotheses we searched publicly available databases for associations with lumbar spine (LS) and femoral neck (FN) BMD in men (www.gefos.org) (31). Data on glycemic traits in men and women combined were downloaded from http://www.magicinvestigators.org/downloads/ (32, 33). Data on glycemic traits in men and women separately were contributed by MAGIC investigators (32, 33). HOMA-IR was calculated as (fasting insulin x fasting glucose)/22.5.

Mendelian Randomization of serum E2 on BMD

To investigate if E2 has a causal effect on BMD we performed a summary statistic two sample inverse variance weighted Mendelian Randomization (34). We selected the 5 top loci from our E2 meta-analysis and extracted summary statistics (β and SE) from the corresponding SNPs in both our E2 study and the GEFOS study on LS and FN BMD. The variant specific associations were used to create an inverse variance weighted estimate of the causal effect size and its standard error.

Results

We performed a GWAS of serum E2 and E1 concentrations, investigating ~2.5 million SNPs in up to 11,097 men. In analyses of autosomal chromosomes, all 9 discovery cohorts (n=11,097) were included in the discovery analyses of E2, and six cohorts (n=7,570) were included in the discovery analyses of E1.

In analyses of the X-chromosome, six cohorts (n=8,953) were included in the discovery analyses of E2, and five cohorts (n=6,917) were included in the discovery analyses of E1.

Estradiol

In the model adjusted for age and BMI (Model 1), two loci were associated with E2 concentrations at the genome-wide significance threshold of $p < 5 \times 10^{-8}$ in the discovery analyses (Figure S1A). The strongest association was found within the *CYP19A1* locus on chromosome 15q21.1 (rs727479, effect size 1.39 pg/ml per effect allele, (SE 0.12), $p = 8.2 \times 10^{-30}$) (Table 1, Figures 1A, S2A, S3A). This SNP, which is located in the second intron of the gene, showed heterogeneity of effect size across studies as indicated by an I^2 value of 57% (25). To take this heterogeneity into account, we additionally calculated a random effects model, which was also genome-wide significant (effect size = 1.35 pg/ml SE 0.19, $p = 2.0 \times 10^{-12}$).

The second locus was found on the X-chromosome where one SNP, rs5934505, reached genome-wide significance ($p = 3.4 \times 10^{-8}$). Rs5934505 is located 79 kb downstream of the *FAMily with sequence similarity 9, member B* (*FAM9B*) gene (Xp22.31) (Table 1, Figures 1A, S2B, S3B). There was heterogeneity of effect size across studies for this SNP ($I^2 = 72\%$). A random effects model displayed nominal, but not genome-wide, significance in the same direction as the result from the fixed effect meta-analysis (C-allele associated with higher E2

302 levels, effect size 0.74 pg/ml per effect allele (SE 0.24), p = 0.002). Therefore, we attempted
303 replication for rs5934505 in the EMAS cohort (n = 1,641). In this cohort, the C-allele was
304 also associated with higher E2 levels; effect size of 1.59 pg/ml per effect allele (SE 0.39), p-
305 value 5.2 x 10^{-5}.

306 In the model that was adjusted for testosterone and SHBG levels, in addition to age and BMI
307 (Model 2 (Figure S1B)), the associations between E2 and the *CYP19A1* locus remained
308 significant (rs727479: p =3.1 x 10^{-43} (Table 1, Figures 1B, S2C, S3C)). In this analysis, the I^2
309 value was 69%, but the random effects model was genome wide significant (effect size 1.42
310 pg/ml per effect allele (SE 0.20), p-value 3.5 x 10^{-13}). A novel genome wide significant locus
311 on the X-chromosome also appeared in this analysis. Rs5951794 (p = 3.1 x 10^{-10}, I^2 = 6%) is
312 located in the distal part of the long arm on chromosome X (Xq27.3), approximately 137 Mb
313 from the *FAM9B* SNP rs5934505 (Table 1, Figures 1B, S2D, S3D).

314 To identify multiple statistically independent SNPs within the same genomic region, we
315 performed stepwise approximate conditional analyses (GCTA) for each of the genome-wide
316 significant loci. In the model adjusted for testosterone and SHBG, the analysis revealed two
317 additional genome-wide significant SNPs in *CYP19A1* locus; rs2899472 in intron 4
318 (conditional p-value 1.1 x 10^{-8}) and rs16964258 in intron 1 (conditional p-value 8.2 x 10^{-15})
319 (Table 1, Figures 1B, S2C, S3E-F). In the model adjusted for age and BMI only, no additional
320 independent associations were found.

321 In Model 1, rs727479 explained 0.9% of the overall variance of E2 levels. When rs5934505
322 (*FAM9B*) was added, 1.1% of the variance was explained. In Model 2, independent *CYP19A1*
323 SNPs explained 1.3% of the overall variance in E2 levels. When rs5951794 (*Chr X*) was
324 added, 1.4% of the variance was explained.

In Model 1, rs727479 explained 0.9% of the overall variance of E2 levels. When the other identified SNP from Model 1, rs5934505 (*FAM9B*) was added, 1.1% of the overall variance in E2 levels was explained. In Model 2, independent *CYP19A1* SNPs explained 1.3% of the overall variance in E2 levels. When the other genome wide significant SNP from Model 2, rs5951794 (*Chr X*), was added, 1.4% of the overall variance in E2 levels was explained.

Estrone

Three genome-wide significant loci, located on chromosomes 7, 8 and 15, respectively, were associated with E1 levels (Figure S1C). The strongest association was found for the *CYP19A1* locus on chromosome 15. The lead SNP was rs2899472 ($p = 5.5 \times 10^{-23}$) (Table 1, Figures 1C, S2E, S3G). Because of heterogeneity in effect size at this variant ($I^2 = 59\%$), a random effects model was run, which was genome wide significant (effect size 2.55 pg/ml per effect allele, SE 0.41, $p = 4.6 \times 10^{-10}$). In conditional analyses of this locus, the SNP with the most significant association with E2, rs727479, was also genome wide significantly associated with E1 (conditional p-value 3.5×10^{-10}) (Table 1, Figures 1C, S2E, S3H).

On chromosome 7, the SNP most significantly associated with E1 levels was rs17277546 ($p = 5.8 \times 10^{-14}$), located in the 3' UTR of the *Tripartite motif containing 4* (*TRIM4*) gene (Table 1, Figures 1C, S2F, S3I). On chromosome 8, the SNP most significantly associated with E1 levels was rs10093796 ($p = 1.2 \times 10^{-8}$). This SNP is located between the *CYP11B1* and the *CYP11B2* genes (Table 1, Figures 1C, S2G, S3J).

Estrone is not derived from testosterone and not bound to SHBG in the circulation. Therefore no analyses of E1 adjusted for these parameters were performed.

Independent *CYP19A1* SNPs explained 1.5% of the overall variance in E1 levels. Rs17277546 (*TRIM4)* and rs10093796 (*CYP11B1*/B2) explained 0.5% and 0.1% respectively

of the variance. In total, 2.1% of the overall variance in E1 levels was explained by these genome wide significant SNPs.

Gene expression analyses

In the GTEx database, two of the *CYP19A1* SNPs were robustly associated with the expression level of *CYP19A1*. The alleles associated with higher E2 levels were associated with higher gene expression levels (rs727479: β 0.23, $p = 1.9 \times 10^{-5}$ (skin), and rs2899472: β 0.20, $p = 9 \times 10^{-8}$ (whole blood)). Rs727479 was also associated with the expression level of *signal peptide peptidase like 2A* (*SPPL2A*) (β 0.18, $p = 1.3 \times 10^{-4}$(transformed fibroblasts). *SPPL2A* is located 442 kB upstream of *CYP19A1*. The E1 associated SNP on chromosome 8, rs10093796, was associated with the expression levels of two adjacent genes in several tissues (*Lys6/Neurotoxin1* (*LYNX1*)) pancreas β 0.68, $p = 5.6 \times 10^{-9}$ and *Lymphocyte Antigen 6 Complex, Locus K* (*LY6K*) skin β 0.32, $p=2.3 \times 10^{-7}$). *LYNX1* and *LY6K* are located 95 kB and 168 kB respectively upstream of *CYP11B1*. The other SNPs in our study were not associated with expression levels in the GTEx database.

Associations with estrogen related traits

To further investigate the physiological relevance of our E2 GWAS findings, we performed look up analyses of other GWAS which had data on phenotypes known or suspected to be related to E2 levels.

Testosterone

To better understand the mechanism underlying the association between our E2-related SNPs and E2 levels, we studied the association between these SNPs and serum testosterone levels. If the effect of the SNPs on E2 levels was exerted upstream of the aromatase enzyme, one would expect that SNPs to be associated with higher testosterone as well as higher E2 levels. On the other hand, if the effect of the SNPs on E2 levels were exerted through alteration in

either the amount or the activity of the aromatase enzyme, only E2 levels would be expected to be increased, with no increase in testosterone levels. The C-allele of the E2 X chromosome SNP rs5934505 (*FAM9B*) was positively associated with levels of both testosterone and E2, suggesting that the effect of rs5934505 is exerted upstream of aromatase (Table 2, Figure 2). Indeed, we have previously reported that the X chromosome SNP rs5934505 (*FAM9B*) is associated with circulating testosterone levels in men ($p = 1.6 \times 10^{-8}$) (30). None of the other E2 SNPs were associated with increased levels of testosterone, suggesting that these SNPs are affecting either the amount or the activity of aromatase or estradiol clearance. In fact, the G-allele of the other E2 X-chromosome SNP, rs5951794, was associated with increased E2 levels and slightly decreased testosterone levels (effect size -7.68 ng/dl per effect allele (SE 3.05), p = 0.01) (Table 2, Figure 2). Additionally, for *CYP19A1* SNPs, there were indications of associations with testosterone in the opposite direction compared to E2, but these associations did not reach statistical significance (rs727479 p = 0.05 and rs16964258 p = 0.26) (Table 2, Figure 2).

BMD

The primary SNP in *CYP19A1*, rs727479, and the secondary signals rs2899472 and rs16964258, were all significantly associated with LS BMD in men ($p \leq 0.01$; Table 3). Rs727479 and rs2899472 were also associated with FN BMD in men ($p < 0.01$). The direction of the effect was the same for all markers, *i.e.* alleles associated with higher levels of E2 were associated with a higher BMD. Moreover, rs5934505 (*FAM9B*) was associated with both FN (p = 0.01) and LS ($p = 7 \times 10^{-6}$) BMD. As in the case of *CYP19A1* SNPs, the allele associated with higher E2 levels was associated with a higher BMD (Table 3).

Mendelian Randomization E2 and BMD

The data from the GEFOS database show associations between individual SNPs and BMD, but do not provide information on possible causality between the E2 levels resulting from these SNPs and BMD. To overcome this we performed a summary statistic Mendelian Randomization analysis which suggested that there is a causal effect of serum E2 on BMD. A 1 pg/ml genetically increased E2 was associated with a 0.048 SD (SE 0.008), $p = 2.8 \times 10^{-12}$ increase in LS BMD. For the femoral neck the increase was 0.037 SD (SE 0.007, $p = 4.4 \times 10^{-8}$ (Figure 3).

Insulin sensitivity

The publicly available GWAS results for measures of insulin sensitivity included only autosomal chromosomes, and did not include results for men and women separately. Thus the following results apply for men and women combined. Insulin resistance expressed as HOMA-IR was negatively associated with the E2 increasing A-alleles of rs727479 (p = 0.004) and rs2899472 (p = 0.003) in *CYP19A1*. This was due to a negative association of these alleles with fasting insulin (p = 0.003 for rs727479 and p = 0.017 for rs2899472) (Supplemental Table 4). Adjustments for BMI had no effect on the results (BMI-adjusted fasting insulin p = 0.002 for rs727479 and p = 0.031 for rs2899472). There were no associations with fasting glucose for these SNPs. The MAGIC investigators also provided us with data not publicly available on fasting insulin and fasting glucose for men and women separately (fasting insulin: men n ≈ 26,000, women n ≈ 32,000, fasting glucose: men n ≈ 36,000, women n ≈ 43,000). In this dataset, the association between rs727479 and fasting insulin was significant in women (β -0.014 (SE 0.004), p = 0.002). In men the direction of the association was the same as in women, but was not statistically significant (rs727479: β -0.006 (SE 0.005), p = 0.19).

Discussion

In this GWAS, SNPs in the *CYP19A1* gene showed the strongest associations with both E1 and E2 levels. This confirms data from previous studies (5, 6, 35) and establishes *CYP19A1* as an important genetic regulator of estrogen levels in men. We found three independent signals in *CYP19A1*, which extends the results from previous studies. We also identified two additional signals for E2 on chromosome X and two additional signals for E1, on chromosomes 7 and 8, respectively. Moreover, SNPs found to be associated with E2 levels in this study were also associated with known or suspected estrogen-related traits including BMD and insulin sensitivity. Mendelian randomization analysis using the independent E2 SNPs suggests a causal effect of E2 on BMD in men.

The finding of several independent signals for both E1 and E2 in *CYP19A1* is consistent with the findings in the previously reported GWAS in Chinese men, where two independent SNPs were found. This strengthens the conception that the regulation of estrogen levels is governed by more than one signal in the gene. The organization of *CYP19A1* is rather complex. The gene consists of a 30-kb coding region and a 93-kb regulatory region including 10 tissue-specific promoters (36). There are four blocks of linkage disequilibrium (LD) in the gene. Rs727479, which displayed the most significant association with E2 levels in our study, is located in intron 2 in LD block 4, which covers 50 kB including the entire coding region, exons/promoters I.6, I.3 and PII, through 5.8 kb downstream of exon 10 (37). Rs727479 has been associated with E2 levels in previous candidate gene studies investigating haplotype-tagging SNPs in *CYP19A1*, as well as in more comprehensive studies investigating larger numbers of SNPs in many genes, in both men (35, 38, 39) and postmenopausal women (40). Moreover, rs727479 was the second most significant SNP in the GWAS of E2 levels in postmenopausal women performed by Prescott and colleagues, although it did not reach

genome-wide significance ($p = 5 \times 10^{-7}$), perhaps due to the relatively low number of study participants (7). In all of these studies, the direction of the effect was the same as in our study: the A-allele was associated with higher E2 levels. The most significant SNP in the male Chinese GWAS performed by Chen and colleagues, rs2414095, is in very strong linkage ($r^2 = 0.96$) with rs727479 (6), and it is also located in intron 3 in LD block 4. The findings from our gene expression analyses that rs727479 is associated with the expression of *CYP19A1* in two tissues further support the relevance of this SNP in the regulation of E2 levels.

To our knowledge, the *CYP19A1* loci rs2899472 and rs16964258 have not been linked to E2 levels in previous studies. Rs2899472 is located in intron 4, in LD-block 4. Rs16964258 is located in a different region of the gene; intron 1, between LD blocks 1 and 2. Interestingly, the SNP most significantly associated with estrogen levels in our previous extended candidate gene study, rs2470152 (5), is also located in this region, 10 kb downstream of rs16964258. The D` for rs2470152 and rs16964258 is 1.0 but the r^2 is 0.062, indicating that the SNPs are probably linked but, due to different allele frequencies, they are not proxy SNPs of one another.

The signal in the *FAM9B* region on the X-chromosome, rs5934505, has not been associated with E2 levels before, but associations of this locus with testosterone levels are known from our earlier testosterone GWAS (30), a finding which was later replicated by Jin and colleagues in a smaller GWAS in men (n = 3,225) (41). Because testosterone is the precursor of estradiol, it is likely that the association of rs5934505 in the *FAM9B* region with E2 levels is mediated through the regulation of testosterone production and not through the conversion of testosterone to E2 per se. Rs5934505 is located in a CNV-insertion area (Xp22), 145 kb upstream of the *family with sequence similarity 9, member A* (*FAM9A*), and 79 kb downstream of *family with sequence similarity 9, member B* (*FAM9B*) genes. Both genes are expressed exclusively in the testes, and share 46% amino acid identity. Very little is known

about their functions (42). The *Kallman syndrome 1* (*KAL1*) gene is located 214 kb downstream of rs5934505. *KAL1* encodes the extracellular matrix glycoprotein anosmin-1 implicated in the embryonic migration of gonadotropin releasing hormone and olfactory neurons. Deleterious mutations in *KAL1* cause X-linked Kallmann syndrome, characterized by hypogonadotropic hypogonadism and anosmia (43), but there are no previous data supporting that minor alterations in the function of KAL1 are associated with sex steroid levels. Moreover, rs5934505 is correlated ($r^2 = 0.35$) with another SNP, rs5978985, in this region, which was associated with male puberty in a recent GWAS (44).

The other signal on the X chromosome, rs5951794, has not previously been associated with sex steroid levels, and the mechanism underlying the association in our study is not known. In contrast to rs5934505 *(FAM9B)*, rs5951794 was not associated with higher testosterone levels. Therefore, the effect of this SNP would be expected to be exerted through alteration in the amount or activity of the aromatase enzyme or through regulation of estradiol clearance. In fact, rs5951794 was associated with slightly lower levels of testosterone. This might be the result of E2 mediated suppression of LH, which in turn would result in decreased testosterone levels. Rs5951794 is located approximately 65 kb downstream of a region rich in micro-RNAs (*MIRs 506-510, 513-514*), expressed mainly in the testes (45). Aside from the micro-RNA cluster, *Fragile X mental retardation 1* (*FMR1*) is the closest gene located approximately 700 kb downstream of rs5951794. Keeping the distance in mind, one could speculate that rs5951794 could affect the regulation of *FMR1*, a gene which in addition to its crucial role in the pathogenesis of Fragile X Syndrome associated mental retardation, is also the leading molecular cause of premature ovarian failure (46).

The E1 signal rs17277546 in the *TRIM4* gene has also been shown to be associated in our previous GWAS of dehydroepiandrosterone sulphate (DHEAS) concentrations (47). Serum levels of DHEAS and DHEA are highly collinear (48). Serum levels of DHEAS could

494 therefore be a marker of serum levels of DHEA. In our earlier GWAS, the G allele was
495 associated with higher levels of DHEAS, and in the present study, the G allele was associated
496 with higher levels of estrone. Thus, an increased amount of adrenal derived precursors for
497 estrogen synthesis is a possible explanation for the present findings. *TRIM4* is a member of
498 the *tripartite motif (TRIM) family*. Members of this family have been implicated in many
499 biological processes including cell differentiation, apoptosis and transcriptional regulation
500 (49). The mechanism relating rs17277546 to DHEAS levels is not known, but in our previous
501 GWAS, we found that rs17277546 is strongly associated with expression levels of *TRIM4* in
502 cell lines from liver and adipose tissue in publically available databases. This indicates that
503 rs17277546 is a functional SNP, or linked to such a SNP (47).

504 The chromosome 8 signal, rs10093618 is located 1.5 kb upstream of the *CYP11B1* gene. The
505 product of *CYP11B1*, the steroid 11β-hydroxylase enzyme, catalyzes the conversion of 11-
506 deoxycortisol to cortisol, representing the final step in cortisol biosynthesis, and 11-
507 deoxycorticosterone to corticosterone. Deficiency of this enzyme leads to congenital adrenal
508 hyperplasia. Hyperandrogenism is a hallmark of this condition since accumulated precursors
509 are shunted into the androgen synthesis pathway (50). One could thus speculate that
510 rs10093618, or an unknown variant linked with it, affects the production or efficiency of the
511 steroid 11β-hydroxylase enzyme and thereby regulates the level of adrenal precursors for the
512 sex steroid synthesis pathway, notably androstenedione, which is a direct precursor in estrone
513 biosynthesis.

514 Because serum E2 levels in men are positively associated with BMD, the SNPs associated
515 with higher E2 levels would be expected to be associated with higher BMD. In fact, in our
516 previous extended candidate gene study there was such an association between the lead
517 *CYP19A1* SNP, rs2470152, and BMD (5). Thus, the association in the present study between
518 E2 associated SNPs in *CYP19A1* as well as *FAM9B,* and BMD is a plausible finding. In fact,

519 rs5934505 is in complete linkage, r^2 1.0, with rs5934507, which was identified as the only
520 male specific signal in our previous GWAS of BMD (31). Because of the known association
521 of rs5934505 with testosterone, the BMD signal was thought to be mediated via testosterone
522 levels in the BMD GWAS. Given the findings in the present study of an association between
523 rs5934505 and E2, it seems more likely that the association with BMD is mediated at least in
524 part via E2 levels rather than solely via a direct effect of testosterone (Figure 3).

525 Although an association between serum E2 levels and BMD in men has been shown in earlier
526 association studies, a causal relation has not been demonstrated. In this study, using
527 Mendelian Randomization analysis, we provide evidence that there is a causal effect of E2 on
528 BMD. For instance, in the RS-1 cohort, where the E2 levels were 12.7 pg/ml (SD 6.6), 1 SD
529 of genetically instrumented decrease in E2 would result in a 6.6 x 0.048 = 0.32 SD decrease in
530 LS BMD and 6.6 x 0.037 = 0.24 SD decrease in FN BMD.

531 According to Johnell and coworkers, the relative risk for hip fracture in men aged 65 was 2.94
532 (95 % CI 2.02-4.27) for each SD decrease in FN BMD (51). Using this information of the
533 association between FN-BMD and hip fracture risk together with the causal effect of serum
534 estradiol on FN-BMD as estimated in the present MR analysis, 1 SD (using the SD of serum
535 estradiol from the RS-1 cohort) decrease in genetically instrumented E2 level could increase
536 the relative risk for hip fracture by 47 %.

537 In this study, SNPs in *CYP19A1* that were associated with higher E2 levels, were also
538 associated with improved insulin sensitivity and lower fasting insulin in men and women
539 combined. In men, the role of estrogens in the regulation of insulin sensitivity is not fully
540 understood. However, mechanistic studies and clinical trials suggest that estrogen signaling is
541 important in the regulation of insulin sensitivity in men (18, 20, 22, 23). Furthermore, men
542 with aromatase deficiency due to an inactivating mutation of the *CYP19A1* gene are

overweight or obese, and display and insulin resistance, which often improves with estrogen replacement therapy (17).

The strengths of our study include the large sample size, with 11,097 men in the discovery analysis of E2 levels, and the large proportion of serum samples analyzed using the MS-technique. This enabled us to find multiple signals in the *CYP19A1* locus, and new signals on other chromosomes, for both E1 and E2. A potential weakness of our study is that not all samples were analyzed by MS. As a result of the lower specificity of the immunoassays, weaker genetic signals might have been missed. It is likely that future studies with even larger numbers of samples analyzed by MS could uncover signals not found in this study. Nevertheless we believe that due to the large proportion of samples analyzed by MS our findings are robust and the risk for false positive signals is low. We also found SNP associations with BMD and measures of insulin sensitivity. Additionally, the Mendelian Randomization analysis provides evidence of a causal effect of E2 on BMD in men. The mechanisms underlying some of the associations in our study should be further investigated to expand our understanding of the regulation of sex steroid levels.

Acknowledgements

The authors thank the MAGIC investigators for providing us with data on fasting insulin and fasting glucose for men and women separately.

562 ## References

563 1. Payne AH, Hales DB. Overview of steroidogenic enzymes in the pathway from cholesterol
564 to active steroid hormones. *Endocr Rev*. 2004;25(6):947-70.
565 2. Travison TG, Zhuang WV, Lunetta KL, Karasik D, Bhasin S, Kiel DP, Coviello AD,
566 Murabito JM. The heritability of circulating testosterone, oestradiol, oestrone and sex
567 hormone binding globulin concentrations in men: the Framingham Heart Study. *Clin*
568 *Endocrinol (Oxf)*. 2014;80(2):277-82.
569 3. Bogaert V, Taes Y, Konings P, Van Steen K, De Bacquer D, Goemaere S, Zmierczak H,
570 Crabbe P, Kaufman JM. Heritability of blood concentrations of sex-steroids in relation to
571 body composition in young adult male siblings. *Clin Endocrinol (Oxf)*. 2008;69(1):129-35.
572 4. Orwoll ES, Nielson CM, Labrie F, Barrett-Connor E, Cauley JA, Cummings SR, Ensrud K,
573 Karlsson M, Lau E, Leung PC, Lunggren O, Mellstrom D, Patrick AL, Stefanick ML,
574 Nakamura K, Yoshimura N, Zmuda J, Vandenput L, Ohlsson C. Evidence for geographical
575 and racial variation in serum sex steroid levels in older men. *J Clin Endocrinol Metab*.
576 2010;95(10):E151-60.
577 5. Eriksson AL, Lorentzon M, Vandenput L, Labrie F, Lindersson M, Syvanen AC, Orwoll
578 ES, Cummings SR, Zmuda JM, Ljunggren O, Karlsson MK, Mellstrom D, Ohlsson C.
579 Genetic variations in sex steroid-related genes as predictors of serum estrogen levels in men. *J*
580 *Clin Endocrinol Metab*. 2009;94(3):1033-41.
581 6. Chen Z, Tao S, Gao Y, Zhang J, Hu Y, Mo L, Kim ST, Yang X, Tan A, Zhang H, Qin X,
582 Li L, Wu Y, Zhang S, Zheng SL, Xu J, Mo Z, Sun J. Genome-wide association study of sex
583 hormones, gonadotropins and sex hormone-binding protein in Chinese men. *J Med Genet*.
584 2013;50(12):794-801.
585 7. Prescott J, Thompson DJ, Kraft P, Chanock SJ, Audley T, Brown J, Leyland J, Folkerd E,
586 Doody D, Hankinson SE, Hunter DJ, Jacobs KB, Dowsett M, Cox DG, Easton DF, De Vivo I.

587 Genome-wide association study of circulating estradiol, testosterone, and sex hormone-
588 binding globulin in postmenopausal women. *PLoS One*. 2012;7(6):e37815.
589 8. Ohlsson C, Borjesson AE, Vandenput L. Sex steroids and bone health in men. *Bonekey*
590 *Rep*. 2012;1:2.
591 9. Smith EP, Boyd J, Frank GR, Takahashi H, Cohen RM, Specker B, Williams TC, Lubahn
592 DB, Korach KS. Estrogen resistance caused by a mutation in the estrogen-receptor gene in a
593 man. *N Engl J Med*. 1994;331(16):1056-61.
594 10. Morishima A, Grumbach MM, Simpson ER, Fisher C, Qin K. Aromatase deficiency in
595 male and female siblings caused by a novel mutation and the physiological role of estrogens.
596 *J Clin Endocrinol Metab*. 1995;80(12):3689-98.
597 11. Khosla S, Melton LJ, 3rd, Robb RA, Camp JJ, Atkinson EJ, Oberg AL, Rouleau PA,
598 Riggs BL. Relationship of volumetric BMD and structural parameters at different skeletal
599 sites to sex steroid levels in men. *J Bone Miner Res*. 2005;20(5):730-40.
600 12. Mellstrom D, Vandenput L, Mallmin H, Holmberg AH, Lorentzon M, Oden A,
601 Johansson H, Orwoll ES, Labrie F, Karlsson MK, Ljunggren O, Ohlsson C. Older men with
602 low serum estradiol and high serum SHBG have an increased risk of fractures. *J Bone Miner*
603 *Res*. 2008;23(10):1552-60.
604 13. Amin S, Zhang Y, Felson DT, Sawin CT, Hannan MT, Wilson PW, Kiel DP. Estradiol,
605 testosterone, and the risk for hip fractures in elderly men from the Framingham Study. *Am J*
606 *Med*. 2006;119(5):426-33.
607 14. Amin S, Zhang Y, Sawin CT, Evans SR, Hannan MT, Kiel DP, Wilson PW, Felson DT.
608 Association of hypogonadism and estradiol levels with bone mineral density in elderly men
609 from the Framingham study. *Ann Intern Med*. 2000;133(12):951-63.
610 15. Finkelstein JS, Lee H, Leder BZ, Burnett-Bowie SA, Goldstein DW, Hahn CW, Hirsch
611 SC, Linker A, Perros N, Servais AB, Taylor AP, Webb ML, Youngner JM, Yu EW. Gonadal

612 steroid-dependent effects on bone turnover and bone mineral density in men. *J Clin Invest*.
613 2016;126(3):1114-25.
614 16. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian
615 randomization: using genes as instruments for making causal inferences in epidemiology. *Stat*
616 *Med*. 2008;27(8):1133-63.
617 17. Zirilli L, Rochira V, Diazzi C, Caffagni G, Carani C. Human models of aromatase
618 deficiency. *J Steroid Biochem Mol Biol*. 2008;109(3-5):212-8.
619 18. Van Sinderen ML, Steinberg GR, Jorgensen SB, To SQ, Knower KC, Clyne CD,
620 Honeyman J, Chow JD, Herridge KA, Jones ME, Simpson ER, Boon WC. Hepatic glucose
621 intolerance precedes hepatic steatosis in the male aromatase knockout (ArKO) mouse. *PLoS*
622 *One*. 2014;9(2):e87230.
623 19. Zhu L, Martinez MN, Emfinger CH, Palmisano BT, Stafford JM. Estrogen signaling
624 prevents diet-induced hepatic insulin resistance in male mice with obesity. *Am J Physiol*
625 *Endocrinol Metab*. 2014;306(10):E1188-97.
626 20. Davis KE, M DN, Sun K, W MS, J DB, J AZ, Zeve D, L DH, D WC, L MG, Xu Y, Z
627 VW, S AK, Clegg DJ. The sexually dimorphic role of adipose and adipocyte estrogen
628 receptors in modulating adipose tissue expansion, inflammation, and fibrosis. *Mol Metab*.
629 2013;2(3):227-42.
630 21. Cooke PS, Heine PA, Taylor JA, Lubahn DB. The role of estrogen and estrogen receptor-
631 alpha in male adipose tissue. *Mol Cell Endocrinol*. 2001;178(1-2):147-54.
632 22. Juang PS, Peng S, Allehmazedeh K, Shah A, Coviello AD, Herbst KL. Testosterone with
633 dutasteride, but not anastrazole, improves insulin sensitivity in young obese men: a
634 randomized controlled trial. *J Sex Med*. 2014;11(2):563-73.

635 23. Finkelstein JS, Lee H, Burnett-Bowie SA, Pallais JC, Yu EW, Borges LF, Jones BF,
636 Barry CV, Wulczyn KE, Thomas BJ, Leder BZ. Gonadal steroids and body composition,
637 strength, and sexual function in men. *N Engl J Med*. 2013;369(11):1011-22.
638 24. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method
639 for the next generation of genome-wide association studies. *PLoS Genet*. 2009;5(6):e1000529.
640 25. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-
641 analyses. *BMJ*. 2003;327(7414):557-60.
642 26. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for
643 genomewide association studies of nearly all common variants. *Genet Epidemiol*.
644 2008;32(4):381-5.
645 27. Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V, Thorleifsson G,
646 Zillikens MC, Speliotes EK, Magi R, Workalemahu T, White CC, Bouatia-Naji N, Harris TB,
647 Berndt SI, Ingelsson E, Willer CJ, Weedon MN, Luan J, Vedantam S, Esko T, Kilpelainen
648 TO, Kutalik Z, Li S, Monda KL, Dixon AL, Holmes CC, Kaplan LM, Liang L, Min JL,
649 Moffatt MF, Molony C, Nicholson G, Schadt EE, Zondervan KT, Feitosa MF, Ferreira T,
650 Lango Allen H, Weyant RJ, Wheeler E, Wood AR, Estrada K, Goddard ME, Lettre G,
651 Mangino M, Nyholt DR, Purcell S, Smith AV, Visscher PM, Yang J, McCarroll SA, Nemesh
652 J, Voight BF, Absher D, Amin N, Aspelund T, Coin L, Glazer NL, Hayward C, Heard-Costa
653 NL, Hottenga JJ, Johansson A, Johnson T, Kaakinen M, Kapur K, Ketkar S, Knowles JW,
654 Kraft P, Kraja AT, Lamina C, Leitzmann MF, McKnight B, Morris AP, Ong KK, Perry JR,
655 Peters MJ, Polasek O, Prokopenko I, Rayner NW, Ripatti S, Rivadeneira F, Robertson NR,
656 Sanna S, Sovio U, Surakka I, Teumer A, van Wingerden S, Vitart V, Zhao JH, Cavalcanti-
657 Proenca C, Chines PS, Fisher E, Kulzer JR, Lecoeur C, Narisu N, Sandholt C, Scott LJ,
658 Silander K, Stark K, Tammesoo ML, Teslovich TM, Timpson NJ, Watanabe RM, Welch R,
659 Chasman DI, Cooper MN, Jansson JO, Kettunen J, Lawrence RW, Pellikka N, Perola M,

660 Vandenput L, Alavere H, Almgren P, Atwood LD, Bennett AJ, Biffar R, Bonnycastle LL,
661 Bornstein SR, Buchanan TA, Campbell H, Day IN, Dei M, Dorr M, Elliott P, Erdos MR,
662 Eriksson JG, Freimer NB, Fu M, Gaget S, Geus EJ, Gjesing AP, Grallert H, Grassler J,
663 Groves CJ, Guiducci C, Hartikainen AL, Hassanali N, Havulinna AS, Herzig KH, Hicks AA,
664 Hui J, Igl W, Jousilahti P, Jula A, Kajantie E, Kinnunen L, Kolcic I, Koskinen S, Kovacs P,
665 Kroemer HK, Krzelj V, Kuusisto J, Kvaloy K, Laitinen J, Lantieri O, Lathrop GM, Lokki
666 ML, Luben RN, Ludwig B, McArdle WL, McCarthy A, Morken MA, Nelis M, Neville MJ,
667 Pare G, Parker AN, Peden JF, Pichler I, Pietilainen KH, Platou CG, Pouta A, Ridderstrale M,
668 Samani NJ, Saramies J, Sinisalo J, Smit JH, Strawbridge RJ, Stringham HM, Swift AJ, Teder-
669 Laving M, Thomson B, Usala G, van Meurs JB, van Ommen GJ, Vatin V, Volpato CB,
670 Wallaschofski H, Walters GB, Widen E, Wild SH, Willemsen G, Witte DR, Zgaga L, Zitting
671 P, Beilby JP, James AL, Kahonen M, Lehtimaki T, Nieminen MS, Ohlsson C, Palmer LJ,
672 Raitakari O, Ridker PM, Stumvoll M, Tonjes A, Viikari J, Balkau B, Ben-Shlomo Y,
673 Bergman RN, Boeing H, Smith GD, Ebrahim S, Froguel P, Hansen T, Hengstenberg C,
674 Hveem K, Isomaa B, Jorgensen T, Karpe F, Khaw KT, Laakso M, Lawlor DA, Marre M,
675 Meitinger T, Metspalu A, Midthjell K, Pedersen O, Salomaa V, Schwarz PE, Tuomi T,
676 Tuomilehto J, Valle TT, Wareham NJ, Arnold AM, Beckmann JS, Bergmann S, Boerwinkle
677 E, Boomsma DI, Caulfield MJ, Collins FS, Eiriksdottir G, Gudnason V, Gyllensten U,
678 Hamsten A, Hattersley AT, Hofman A, Hu FB, Illig T, Iribarren C, Jarvelin MR, Kao WH,
679 Kaprio J, Launer LJ, Munroe PB, Oostra B, Penninx BW, Pramstaller PP, Psaty BM,
680 Quertermous T, Rissanen A, Rudan I, Shuldiner AR, Soranzo N, Spector TD, Syvanen AC,
681 Uda M, Uitterlinden A, Volzke H, Vollenweider P, Wilson JF, Witteman JC, Wright AF,
682 Abecasis GR, Boehnke M, Borecki IB, Deloukas P, Frayling TM, Groop LC, Haritunians T,
683 Hunter DJ, Kaplan RC, North KE, O'Connell JR, Peltonen L, Schlessinger D, Strachan DP,
684 Hirschhorn JN, Assimes TL, Wichmann HE, Thorsteinsdottir U, van Duijn CM, Stefansson

685 K, Cupples LA, Loos RJ, Barroso I, McCarthy MI, Fox CS, Mohlke KL, Lindgren CM. Meta-
686 analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism
687 in the genetic basis of fat distribution. *Nat Genet*. 2010;42(11):949-60.
688 28. Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A, Wareham N. EPIC-Norfolk:
689 study design and characteristics of the cohort. European Prospective Investigation of Cancer.
690 *Br J Cancer*. 1999;80 Suppl 1:95-103.
691 29. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue
692 gene regulation in humans. *Science*. 2015;348(6235):648-60.
693 30. Ohlsson C, Wallaschofski H, Lunetta KL, Stolk L, Perry JR, Koster A, Petersen AK,
694 Eriksson J, Lehtimaki T, Huhtaniemi IT, Hammond GL, Maggio M, Coviello AD, Ferrucci L,
695 Heier M, Hofman A, Holliday KL, Jansson JO, Kahonen M, Karasik D, Karlsson MK, Kiel
696 DP, Liu Y, Ljunggren O, Lorentzon M, Lyytikainen LP, Meitinger T, Mellstrom D, Melzer D,
697 Miljkovic I, Nauck M, Nilsson M, Penninx B, Pye SR, Vasan RS, Reincke M, Rivadeneira F,
698 Tajar A, Teumer A, Uitterlinden AG, Ulloor J, Viikari J, Volker U, Volzke H, Wichmann HE,
699 Wu TS, Zhuang WV, Ziv E, Wu FC, Raitakari O, Eriksson A, Bidlingmaier M, Harris TB,
700 Murray A, de Jong FH, Murabito JM, Bhasin S, Vandenput L, Haring R. Genetic
701 determinants of serum testosterone concentrations in men. *PLoS Genet*. 2011;7(10):e1002313.
702 31. Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, Oei L,
703 Albagha OM, Amin N, Kemp JP, Koller DL, Li G, Liu CT, Minster RL, Moayyeri A,
704 Vandenput L, Willner D, Xiao SM, Yerges-Armstrong LM, Zheng HF, Alonso N, Eriksson J,
705 Kammerer CM, Kaptoge SK, Leo PJ, Thorleifsson G, Wilson SG, Wilson JF, Aalto V, Alen
706 M, Aragaki AK, Aspelund T, Center JR, Dailiana Z, Duggan DJ, Garcia M, Garcia-Giralt N,
707 Giroux S, Hallmans G, Hocking LJ, Husted LB, Jameson KA, Khusainova R, Kim GS,
708 Kooperberg C, Koromila T, Kruk M, Laaksonen M, Lacroix AZ, Lee SH, Leung PC, Lewis
709 JR, Masi L, Mencej-Bedrac S, Nguyen TV, Nogues X, Patel MS, Prezelj J, Rose LM, Scollen

710 S, Siggeirsdottir K, Smith AV, Svensson O, Trompet S, Trummer O, van Schoor NM, Woo J,
711 Zhu K, Balcells S, Brandi ML, Buckley BM, Cheng S, Christiansen C, Cooper C, Dedoussis
712 G, Ford I, Frost M, Goltzman D, Gonzalez-Macias J, Kahonen M, Karlsson M,
713 Khusnutdinova E, Koh JM, Kollia P, Langdahl BL, Leslie WD, Lips P, Ljunggren O, Lorenc
714 RS, Marc J, Mellstrom D, Obermayer-Pietsch B, Olmos JM, Pettersson-Kymmer U, Reid
715 DM, Riancho JA, Ridker PM, Rousseau F, Slagboom PE, Tang NL, Urreizti R, Van Hul W,
716 Viikari J, Zarrabeitia MT, Aulchenko YS, Castano-Betancourt M, Grundberg E, Herrera L,
717 Ingvarsson T, Johannsdottir H, Kwan T, Li R, Luben R, Medina-Gomez C, Palsson ST,
718 Reppe S, Rotter JI, Sigurdsson G, van Meurs JB, Verlaan D, Williams FM, Wood AR, Zhou
719 Y, Gautvik KM, Pastinen T, Raychaudhuri S, Cauley JA, Chasman DI, Clark GR, Cummings
720 SR, Danoy P, Dennison EM, Eastell R, Eisman JA, Gudnason V, Hofman A, Jackson RD,
721 Jones G, Jukema JW, Khaw KT, Lehtimaki T, Liu Y, Lorentzon M, McCloskey E, Mitchell
722 BD, Nandakumar K, Nicholson GC, Oostra BA, Peacock M, Pols HA, Prince RL, Raitakari
723 O, Reid IR, Robbins J, Sambrook PN, Sham PC, Shuldiner AR, Tylavsky FA, van Duijn CM,
724 Wareham NJ, Cupples LA, Econs MJ, Evans DM, Harris TB, Kung AW, Psaty BM, Reeve J,
725 Spector TD, Streeten EA, Zillikens MC, Thorsteinsdottir U, Ohlsson C, Karasik D, Richards
726 JB, Brown MA, Stefansson K, Uitterlinden AG, Ralston SH, Ioannidis JP, Kiel DP,
727 Rivadeneira F. Genome-wide meta-analysis identifies 56 bone mineral density loci and
728 reveals 14 loci associated with risk of fracture. *Nat Genet*. 2012;44(5):491-501.
729 32. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, Rybin D, Liu
730 CT, Bielak LF, Prokopenko I, Amin N, Barnes D, Cadby G, Hottenga JJ, Ingelsson E,
731 Jackson AU, Johnson T, Kanoni S, Ladenvall C, Lagou V, Lahti J, Lecoeur C, Liu Y,
732 Martinez-Larrad MT, Montasser ME, Navarro P, Perry JR, Rasmussen-Torvik LJ, Salo P,
733 Sattar N, Shungin D, Strawbridge RJ, Tanaka T, van Duijn CM, An P, de Andrade M,
734 Andrews JS, Aspelund T, Atalay M, Aulchenko Y, Balkau B, Bandinelli S, Beckmann JS,

735 Beilby JP, Bellis C, Bergman RN, Blangero J, Boban M, Boehnke M, Boerwinkle E,
736 Bonnycastle LL, Boomsma DI, Borecki IB, Bottcher Y, Bouchard C, Brunner E, Budimir D,
737 Campbell H, Carlson O, Chines PS, Clarke R, Collins FS, Corbaton-Anchuelo A, Couper D,
738 de Faire U, Dedoussis GV, Deloukas P, Dimitriou M, Egan JM, Eiriksdottir G, Erdos MR,
739 Eriksson JG, Eury E, Ferrucci L, Ford I, Forouhi NG, Fox CS, Franzosi MG, Franks PW,
740 Frayling TM, Froguel P, Galan P, de Geus E, Gigante B, Glazer NL, Goel A, Groop L,
741 Gudnason V, Hallmans G, Hamsten A, Hansson O, Harris TB, Hayward C, Heath S, Hercberg
742 S, Hicks AA, Hingorani A, Hofman A, Hui J, Hung J, Jarvelin MR, Jhun MA, Johnson PC,
743 Jukema JW, Jula A, Kao WH, Kaprio J, Kardia SL, Keinanen-Kiukaanniemi S, Kivimaki M,
744 Kolcic I, Kovacs P, Kumari M, Kuusisto J, Kyvik KO, Laakso M, Lakka T, Lannfelt L,
745 Lathrop GM, Launer LJ, Leander K, Li G, Lind L, Lindstrom J, Lobbens S, Loos RJ, Luan J,
746 Lyssenko V, Magi R, Magnusson PK, Marmot M, Meneton P, Mohlke KL, Mooser V,
747 Morken MA, Miljkovic I, Narisu N, O'Connell J, Ong KK, Oostra BA, Palmer LJ, Palotie A,
748 Pankow JS, Peden JF, Pedersen NL, Pehlic M, Peltonen L, Penninx B, Pericic M, Perola M,
749 Perusse L, Peyser PA, Polasek O, Pramstaller PP, Province MA, Raikkonen K, Rauramaa R,
750 Rehnberg E, Rice K, Rotter JI, Rudan I, Ruokonen A, Saaristo T, Sabater-Lleal M, Salomaa
751 V, Savage DB, Saxena R, Schwarz P, Seedorf U, Sennblad B, Serrano-Rios M, Shuldiner AR,
752 Sijbrands EJ, Siscovick DS, Smit JH, Small KS, Smith NL, Smith AV, Stancakova A,
753 Stirrups K, Stumvoll M, Sun YV, Swift AJ, Tonjes A, Tuomilehto J, Trompet S, Uitterlinden
754 AG, Uusitupa M, Vikstrom M, Vitart V, Vohl MC, Voight BF, Vollenweider P, Waeber G,
755 Waterworth DM, Watkins H, Wheeler E, Widen E, Wild SH, Willems SM, Willemsen G,
756 Wilson JF, Witteman JC, Wright AF, Yaghootkar H, Zelenika D, Zemunik T, Zgaga L,
757 Wareham NJ, McCarthy MI, Barroso I, Watanabe RM, Florez JC, Dupuis J, Meigs JB,
758 Langenberg C. A genome-wide approach accounting for body mass index identifies genetic

759 variants influencing fasting glycemic traits and insulin resistance. *Nat Genet*. 2012;44(6):659-
760 69.
761 33. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E,
762 Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Magi R, Morris AP, Randall J, Johnson
763 T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan
764 A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JR, Egan JM, Lajunen T,
765 Grarup N, Sparso T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P,
766 Cavalcanti-Proenca C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G,
767 Ingelsson E, An P, O'Connell J, Luan J, Elliott A, McCarroll SA, Payne F, Roccasecca RM,
768 Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y,
769 Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A,
770 Bonnycastle LL, Borch-Johnsen K, Bottcher Y, Brunner E, Bumpstead SJ, Charpentier G,
771 Chen YD, Chines P, Clarke R, Coin LJ, Cooper MN, Cornelis M, Crawford G, Crisponi L,
772 Day IN, de Geus EJ, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A,
773 Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ,
774 Grundy S, Gwilliam R, Gyllensten U, Hadjadj S, Hallmans G, Hammond N, Han X,
775 Hartikainen AL, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA,
776 Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PR, Jorgensen T,
777 Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P,
778 Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley
779 R, Mangino M, Manning AK, Martinez-Larrad MT, McAteer JB, McCulloch LJ, McPherson
780 R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S,
781 Narisu N, Neville MJ, Oostra BA, Orru M, Pakyz R, Palmer CN, Paolisso G, Pattaro C,
782 Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AF, Pichler I, Polasek O, Posthuma D,
783 Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S,

784 Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet
785 P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurethsson G, Sijbrands EJ, Silveira A,
786 Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvanen AC, Tanaka T,
787 Thorand B, Tichet J, Tonjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M,
788 Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A,
789 Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemsen G, Witteman JC,
790 Yarnell JW, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, Borecki IB,
791 Loos RJ, Meneton P, Magnusson PK, Nathan DM, Williams GH, Hattersley AT, Silander K,
792 Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper
793 C, Dedoussis GV, Serrano-Rios M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW,
794 Ebrahim S, Marmot M, Kao WH, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T,
795 Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M,
796 Campbell H, Wilson JF, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J,
797 Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BW, Boomsma DI, Deloukas P,
798 Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn
799 CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruokonen A, Jarvelin MR,
800 Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis GR, Wareham NJ, Sladek
801 R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC,
802 Barroso I. New genetic loci implicated in fasting glucose homeostasis and their impact on
803 type 2 diabetes risk. *Nat Genet*. 2010;42(2):105-16.
804 34. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with
805 multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37(7):658-65.
806 35. Ahn J, Schumacher FR, Berndt SI, Pfeiffer R, Albanes D, Andriole GL, Ardanaz E,
807 Boeing H, Bueno-de-Mesquita B, Chanock SJ, Clavel-Chapelon F, Diver WR, Feigelson HS,
808 Gaziano JM, Giovannucci E, Haiman CA, Henderson BE, Hoover RN, Kolonel LN, Kraft P,

809 Ma J, Le Marchand L, Overvad K, Palli D, Stattin P, Stampfer M, Stram DO, Thomas G,
810 Thun MJ, Travis RC, Trichopoulos D, Virtamo J, Weinstein SJ, Yeager M, Kaaks R, Hunter
811 DJ, Hayes RB. Quantitative trait loci predicting circulating sex steroid hormones in men from
812 the NCI-Breast and Prostate Cancer Cohort Consortium (BPC3). *Hum Mol Genet*.
813 2009;18(19):3749-57.
814 36. Bulun SE, Sebastian S, Takayama K, Suzuki T, Sasano H, Shozu M. The human CYP19
815 (aromatase P450) gene: update on physiologic roles and genomic organization of promoters. *J*
816 *Steroid Biochem Mol Biol*. 2003;86(3-5):219-24.
817 37. Haiman CA, Stram DO, Pike MC, Kolonel LN, Burtt NP, Altshuler D, Hirschhorn J,
818 Henderson BE. A comprehensive haplotype analysis of CYP19 and breast cancer risk: the
819 Multiethnic Cohort. *Hum Mol Genet*. 2003;12(20):2679-92.
820 38. Travis RC, Schumacher F, Hirschhorn JN, Kraft P, Allen NE, Albanes D, Berglund G,
821 Berndt SI, Boeing H, Bueno-de-Mesquita HB, Calle EE, Chanock S, Dunning AM, Hayes R,
822 Feigelson HS, Gaziano JM, Giovannucci E, Haiman CA, Henderson BE, Kaaks R, Kolonel
823 LN, Ma J, Rodriguez L, Riboli E, Stampfer M, Stram DO, Thun MJ, Tjonneland A,
824 Trichopoulos D, Vineis P, Virtamo J, Le Marchand L, Hunter DJ. CYP19A1 genetic variation
825 in relation to prostate cancer risk and circulating sex hormone concentrations in men from the
826 Breast and Prostate Cancer Cohort Consortium. *Cancer Epidemiol Biomarkers Prev*.
827 2009;18(10):2734-44.
828 39. Huhtaniemi IT, Pye SR, Holliday KL, Thomson W, O'Neill TW, Platt H, Payne D, John
829 SL, Jiang M, Bartfai G, Boonen S, Casanueva FF, Finn JD, Forti G, Giwercman A, Han TS,
830 Kula K, Lean ME, Pendleton N, Punab M, Silman AJ, Vanderschueren D, Labrie F, Wu FC.
831 Effect of polymorphisms in selected genes involved in pituitary-testicular function on
832 reproductive hormones and phenotype in aging men. *J Clin Endocrinol Metab*.
833 2010;95(4):1898-908.

834 40. Haiman CA, Dossus L, Setiawan VW, Stram DO, Dunning AM, Thomas G, Thun MJ,
835 Albanes D, Altshuler D, Ardanaz E, Boeing H, Buring J, Burtt N, Calle EE, Chanock S,
836 Clavel-Chapelon F, Colditz GA, Cox DG, Feigelson HS, Hankinson SE, Hayes RB,
837 Henderson BE, Hirschhorn JN, Hoover R, Hunter DJ, Kaaks R, Kolonel LN, Le Marchand L,
838 Lenner P, Lund E, Panico S, Peeters PH, Pike MC, Riboli E, Tjonneland A, Travis R,
839 Trichopoulos D, Wacholder S, Ziegler RG. Genetic variation at the CYP19A1 locus predicts
840 circulating estrogen levels but not breast cancer risk in postmenopausal women. *Cancer Res*.
841 2007;67(5):1893-7.
842 41. Jin G, Sun J, Kim ST, Feng J, Wang Z, Tao S, Chen Z, Purcell L, Smith S, Isaacs WB,
843 Rittmaster RS, Zheng SL, Condreay LD, Xu J. Genome-wide association study identifies a
844 new locus JMJD1C at 10q21 that may influence serum androgen levels in men. *Hum Mol*
845 *Genet*. 2012;21(23):5222-8.
846 42. Martinez-Garay I, Jablonka S, Sutajova M, Steuernagel P, Gal A, Kutsche K. A new gene
847 family (FAM9) of low-copy repeats in Xp22.3 expressed exclusively in testis: implications
848 for recombinations in this region. *Genomics*. 2002;80(3):259-67.
849 43. Oliveira LM, Seminara SB, Beranova M, Hayes FJ, Valkenburgh SB, Schipani E, Costa
850 EM, Latronico AC, Crowley WF, Jr., Vallejo M. The importance of autosomal genes in
851 Kallmann syndrome: genotype-phenotype correlations and neuroendocrine characteristics. *J*
852 *Clin Endocrinol Metab*. 2001;86(4):1532-8.
853 44. Day FR, Bulik-Sullivan B, Hinds DA, Finucane HK, Murabito JM, Tung JY, Ong KK,
854 Perry JR. Shared genetic aetiology of puberty timing between sexes and with health-related
855 outcomes. *Nat Commun*. 2015;6:8842.
856 45. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A,
857 Kamphorst AO, Landthaler M, Lin C, Socci ND, Hermida L, Fulci V, Chiaretti S, Foa R,
858 Schliwka J, Fuchs U, Novosel A, Muller RU, Schermer B, Bissels U, Inman J, Phan Q, Chien

M, Weir DB, Choksi R, De Vita G, Frezzetti D, Trompeter HI, Hornung V, Teng G, Hartmann G, Palkovits M, Di Lauro R, Wernet P, Macino G, Rogler CE, Nagle JW, Ju J, Papavasiliou FN, Benzing T, Lichter P, Tam W, Brownstein MJ, Bosio A, Borkhardt A, Russo JJ, Sander C, Zavolan M, Tuschl T. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*. 2007;129(7):1401-14.

46. Pu D, Xing Y, Gao Y, Gu L, Wu J. Gene variation and premature ovarian failure: a meta-analysis. *Eur J Obstet Gynecol Reprod Biol*. 2014;182C:226-37.

47. Zhai G, Teumer A, Stolk L, Perry JR, Vandenput L, Coviello AD, Koster A, Bell JT, Bhasin S, Eriksson J, Eriksson A, Ernst F, Ferrucci L, Frayling TM, Glass D, Grundberg E, Haring R, Hedman AK, Hofman A, Kiel DP, Kroemer HK, Liu Y, Lunetta KL, Maggio M, Lorentzon M, Mangino M, Melzer D, Miljkovic I, Nica A, Penninx BW, Vasan RS, Rivadeneira F, Small KS, Soranzo N, Uitterlinden AG, Volzke H, Wilson SG, Xi L, Zhuang WV, Harris TB, Murabito JM, Ohlsson C, Murray A, de Jong FH, Spector TD, Wallaschofski H. Eight common genetic variants associated with serum DHEAS levels suggest a key role in ageing mechanisms. *PLoS Genet*. 2011;7(4):e1002025.

48. Tivesten A, Vandenput L, Carlzon D, Nilsson M, Karlsson MK, Ljunggren O, Barrett-Connor E, Mellstrom D, Ohlsson C. Dehydroepiandrosterone and its sulfate predict the 5-year risk of coronary heart disease events in elderly men. *J Am Coll Cardiol*. 2014;64(17):1801-10.

49. Yan J, Li Q, Mao AP, Hu MM, Shu HB. TRIM4 modulates type I interferon induction and cellular antiviral response by targeting RIG-I for K63-linked ubiquitination. *J Mol Cell Biol*. 2014;6(2):154-63.

50. Krone N, Arlt W. Genetics of congenital adrenal hyperplasia. *Best Pract Res Clin Endocrinol Metab*. 2009;23(2):181-92.

882 51. Johnell O, Kanis JA, Oden A, Johansson H, De Laet C, Delmas P, Eisman JA, Fujiwara S,
883 Kroger H, Mellstrom D, Meunier PJ, Melton LJ, 3rd, O'Neill T, Pols H, Reeve J, Silman A,
884 Tenenhouse A. Predictive value of BMD for hip and other fractures. *J Bone Miner Res*.

885

886 **Legends to figures**

887 **Figure 1A-C. Manhattan plots for the genome-wide meta-analysis results.**

888 (A) E2 adjusted for age and BMI, (B) E2 adjusted for, age, BMI, testosterone and SHBG and
889 (C) E1 adjusted for age and BMI. Red line indicates $p = 5 \times 10^{-8}$. Genome-wide significant
890 loci are indicated by green color. In the analysis of E1, one SNP on chromosome 1 reached
891 the threshold for genome-wide significance ($p < 5 \times 10^{-8}$), but had a minor allele frequency of
892 < 0.01 in all but two cohorts. Therefore this SNP was discarded from further analyses.

893 **Figure 2. Proposed mechanisms underlying the associations between genome wide**
894 **significant SNPs and serum levels of E2 and T**

895 SNPs associated with elevated levels of both E2 and T are expected to be located upstream of
896 T. SNPs associated with elevated levels of E2 but no increase in T levels are expected to be
897 affecting aromatase activity or estradiol clearance. The allele associated with increased serum
898 E2 is given for each SNP. The proposed effect of E2 on BMD is also indicated; upwards
899 arrow represents increase, downwards arrow represents decrease and arrow in parentheses
900 represents non-significant decrease.

901 **Figure 3. Forest plot of Mendelian Randomization analyses showing the effect of E2 on**
902 **BMD.**

903 Effect size of E2 on BMD expressed as SD increase in BMD per pg/ml E2. The horizontal
904 lines represent confidence interval; the central vertical line represents precision. The values
905 are based on a meta-analysis of all five E2 associated SNPs (rs727479, rs2899472,
906 rs16964258, rs5934505, rs5951794). The horizontal axis shows the scale of the effects.
907 LSBMD = lumbar spine BMD, FNBMD = femoral neck BMD.

908

909