Eight Common Genetic Variants Associated with Serum DHEAS Levels Suggest a Key Role in Ageing Mechanisms

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

Published Version
doi:10.1371/journal.pgen.1002025

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:5146966

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Eight Common Genetic Variants Associated with Serum DHEAS Levels Suggest a Key Role in Ageing Mechanisms

Guangju Zhai1, Alexander Teumer2,*, Lisette Stolk3,*, John R. B. Perry5,*, Liesbeth Vandenput4,*, Andrea D. Covilleo8,*, Annemarie Koster9,*, Jordana T. Bell1,6, Shalender Bhasin10, Joel Eriksson7, Anna Eriksson7, Florian Ernst2, Luigi Ferrucci11, Timothy M. Frayling5, Daniel Glass1, Elin Grundberg1,12, Robin Haring13, Asa K. Hedman9, Albert Hofman4,14, Douglas P. Kiel15, Heyo K. Kroemer16, Yongmei Liu17, Kathryn L. Lunetta18, Marcello Maggio19, Mattias Lorzenton7, Massimo Mangino1, David Melzer5, Iva Miljkovic20, MuTHER Consortium, Alexandra Nica12,21, Brenda W. J. Penninx22, Ramachandran S. Vasan8,23, Fernando Ravindere13,4, Kerrin S. Small1,12, Nicole Soranzo1,12, André G. Uitterlinden3,4,14, Henry Völzke24, Scott G. Wilson1,25,26, Li Xi27, Wei Vivian Zhuang18, Tamara B. Harris9,*, Joanne M. Murabito8,23, Claes Ohlsson7*, Anna Murray51, Frank H. de Jong3*, Tim D. Spector1,*, Henri Wallaschofski13

1 Department of Twin Research and Genetic Epidemiology, King’s College London, London, United Kingdom, 2 Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany, 3 Department of Internal Medicine, Erasmus MC Rotterdam, Rotterdam, The Netherlands, 4 Netherlands Consortium of Healthy Ageing, Rotterdam, The Netherlands, 5 Genetics of Complex Traits, Peninsula Medical School, University of Exeter, Exeter, United Kingdom, 6 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, 7 Department of Internal Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, 8 Sections of General Internal Medicine, Preventive Medicine, Cardiology and Endocrinology, Diabetes and Nutrition, Department of Medicine, Erasmus School of Medicine, Boston, Massachusetts, United States of America, 9 Laboratory for Epidemiology, Demography, and Biometry, National Institute on Aging, Bethesda, Maryland, United States of America, 10 Section of Endocrinology, Diabetes, and Nutrition, Claude D. Pepper Older Americans Independence Center, Boston University School of Medicine, Boston, Massachusetts, United States of America, 11 Clinical Research Branch, National Institute on Aging, Baltimore, Maryland, United States of America, 12 Wellcome Trust Sanger Institute, Hinxton, United Kingdom, 13 Institute for Chemical and Laboratory Medicine, University of Greifswald, Greifswald, Germany, 14 Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands, 15 Hebrew Senior Life Institute for Aging Research and Harvard Medical School, Boston, Massachusetts, United States of America, 16 Center of Pharmacology and Experimental Therapeutics, Department of Pharmacology, University of Greifswald, Greifswald, Germany, 17 Department of Epidemiology and Prevention, Wake Forest University Health Sciences, Winston-Salem, North Carolina, United States of America, 18 Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, United States of America, 19 Department of Internal Medicine and Biomedical Sciences, Section of Geriatrics, University of Parma, Parma, Italy, 20 Department of Epidemiology, University of Pittsburgh, Pittsburg, Pennsylvania, United States of America, 21 Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland, 22 Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands, 23 The National Heart Lung and Blood Institute’s Framingham Heart Study, Framingham, Massachusetts, United States of America, 24 Institute for Community Medicine, University of Greifswald, Greifswald, Germany, 25 Department of Endocrinology and Diabetes, Sir Charles Gardiner Hospital, Nedlands, Australia, 26 School of Medicine and Pharmacology, University of Western Australia, Nedlands, Australia, 27 Molecular Medicine – Computational Biology, Pfizer Worldwide R&D, Groton, Connecticut, United States of America

Abstract
Dehydroepiandrosterone sulphate (DHEAS) is the most abundant circulating steroid secreted by adrenal glands—yet its function is unknown. Its serum concentration declines significantly with increasing age, which has led to speculation that a relative DHEAS deficiency may contribute to the development of common age-related diseases or diminished longevity. We conducted a meta-analysis of genome-wide association data with 14,846 individuals and identified eight independent common SNPs associated with serum DHEAS concentrations. Genes at or near the identified loci include ZKSCAN5 (rs11761528; p = 3.15 × 10−6), SULT2A1 (rs2637125; p = 2.61 × 10−6), ARPC1A (rs740160; p = 1.56 × 10−6), BCL2L11 (rs3153738; p = 4.50 × 10−6), ZFCH10 (rs1811230; p = 5.44 × 10−6), HHEX (rs26947306; p = 4.64 × 10−5), BCL2L11 (rs6738028; p = 1.72 × 10−5), and CYP2C9 (rs2185570; p = 2.29 × 10−5). These genes are associated with type 2 diabetes, lymphoma, actin filament assembly, drug and xenobiotic metabolism, and zinc finger proteins. Several SNPs were associated with changes in gene expression levels, and the related genes are connected to biological pathways linking DHEAS with ageing. This study provides much needed insight into the function of DHEAS.


Editor: Greg Gibson, Georgia Institute of Technology, United States of America

Received November 18, 2010; Accepted January 27, 2011; Published April 14, 2011

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Funding: This is a collaborative work and supported by many sources of funding. Due to the limit of words here, the details and list of the funding sources are provided in the Acknowledgments. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* These authors contributed equally to this work.
† These senior authors were joint senior authors on this work.
**Author Summary**

Dehydroepiandrosterone sulphate (DHEAS), mainly secreted by the adrenal gland, is the most abundant circulating steroid in humans. It acts as an inactive precursor which is converted initially into DHEA and thereafter into active androgens and estrogens in peripheral target tissues [1]. In humans the serum concentration of circulating DHEAS is 100- to 500-fold or 1000 to 10,000 higher than that of testosterone and estradiol respectively. Unlike DHEA, which is swiftly cleared from the circulation and shows diurnal variation, serum DHEAS concentrations are stable and facilitate accurate measurement and diagnosis of pathology [2].

DHEAS is distinct from the other major adrenal steroids (cortisol and aldosterone) in showing a significant physiological decline after the age of 25 and diminishes about 95% by the age of 85 years [3]. This age-related decline has led to speculation that a relative DHEAS deficiency may contribute to the development of common age-related diseases or diminished longevity. Twin- and family-based studies have shown that there is a substantial genetic effect with heritability estimate of 60%, but no specific genes regulating serum DHEAS concentration have been identified to date. Here we take advantage of recent technical and methodological advances to examine the effects of common genetic variants on serum DHEAS concentrations. By examining 14,846 Caucasian individuals, we show that eight common genetic variants are associated with serum DHEAS concentrations. Genes at or near these genetic variants include BCL2L11, ARPC1A, ZKSCAN5, TRIM8, HHEX, CYP2C9, BMF, and SULT2A1. These genes have various associations with steroid hormone metabolism—co-morbidities of ageing including type 2 diabetes, lymphoma, actin filament assembly, drug and xenobiotic metabolism, and zinc finger proteins—suggesting a wider functional role for DHEAS than previously thought.

**Introduction**

Dehydroepiandrosterone sulphate (DHEAS), mainly secreted by the adrenal gland, is the most abundant circulating steroid in humans. Despite these observations, the physiological function of DHEAS and its importance in maintaining health are poorly understood. Although previous twin [22,23] and family-based studies [24,25] have shown that there is a substantial genetic effect with a heritability estimate of 60% [22], no specific genes regulating serum DHEAS concentration in healthy individuals have been identified to date. Therefore, the current study meta-analyzed the results of genome-wide association studies (GWAS) performed in a total of 14,940 individuals from seven cohorts to identify common genetic variants associated with serum DHEAS concentrations. The findings not only advance understanding of how serum DHEAS concentration is regulated by genes but also provide clues as to its mechanism of action as well as Mendelian randomisation principles [26].

**Results**

We carried out a meta-analysis of 8,563 women and 6,281 men of European origin from collaborating studies: TwinsUK (n = 4,906), Framingham Heart Study (FHS) (n = 3,183), SHIP (n = 1,832), Rotterdam Study (RS1) (n = 1,597), InCHIANTI (n = 1,182), Health ABC (n = 1,222), and GOOD (n = 924). Serum samples were collected either after overnight fasting or non-fasting in each cohort and DHEAS was measured by either immunoassay or liquid chromatography tandem mass spectrometry (LC-MS/MS) methods (Table 1). Mean age differed across the cohorts from 19 to 74 years in men and 50 to 74 years in women and corresponding mean DHEAS concentrations varied from 1.20 to 7.05 μmol/L (Table 1).

Each cohort performed GWA tests for log transformed DHEAS on ~2.5 million imputed single nucleotide polymorphisms (SNPs) in men and women separately with adjustment for age, and additionally for age and sex for those cohorts who had data in both men and women. Then Z-scores from each cohort were pooled for the meta-analysis at each SNP.

In all our individual GWAS, λGC, which is defined as the median χ2 (1 degree of freedom) association statistic across SNPs divided by its theoretical median under the null distribution [27], ranged from 0.984 to 1.023, indicating that there was no population stratification or it was very minor. Further, we corrected for population stratification by applying the genomic control method [27]; the λGC in the meta-analysis is 1.017. In addition, the effect direction was consistent across all the cohorts and there is no between-study heterogeneity as indicated by I2 ranging between 0 and 0.12 (Table 2).

We found 44 SNPs were associated with serum DHEAS concentrations in men at conventional genome-wide significance (p<5×10−8), which are all located on chromosome 7q22.1 (Figure 1B; Table S1). All these SNPs except for three were significant in women (Figure 1A; Table S1). In addition, 19 SNPs located on chromosome 19q13.3 were found in women to be associated with serum DHEAS concentrations with p<5×10−8. In the sex-combined meta-analysis, the significance became stronger for all these SNPs (Figure 1C; Table S1). Further, we found 8 SNPs located on chromosome 10q23.33 which represents two regions more than 2 Mb apart, 12 SNPs on chromosome 15q15.1, and in addition, 4 SNPs on chromosome 19q13.3 were associated with serum DHEAS concentrations with p<5×10−8. Together we found a total of 87 SNPs associated with serum DHEAS concentrations with p<5×10−8, representing five chromosomal regions of less than 1 Mb each (Table S1).

The most significantly associated SNPs in each of these five regions are presented in Table 2. The minor allele of rs11761528 (p = 3.15×10−30) on chromosome 7q22.1, rs2637125 (p = 2.61×10−15) on chromosome 19q13.3, and rs2497306 (p = 4.6×10−10) and rs2185570 (p = 2.29×10−8) on chromosome 10q22.33 (more than 2 Mb apart), were negatively associated with...
DHEAS concentrations. In comparison, the minor allele of rs7181230 (p = 5.44 × 10^{-11}) on chromosome 15q15.1 was positively associated with serum DHEAS concentrations. Based on the HapMap3 release2 CEU data, the significant 87 SNPs from within the five regions have low pair-wise r^2, indicating potentially multiple independent signals. To verify this, we performed a conditional meta-analysis with adjustment for the five most significant SNPs plus age and sex in each cohort.

After this adjustment, all other SNPs on chromosome 10, 15, 26, and 19 became non-significant (Figure 1D). However, on chromosome 7, we found two independent signals; one defined by rs11761528 and a second located 370 kb upstream in the 3' intron of the TRIM4 gene (rs740160 located 161 kb downstream of rs11761528, p = 1.72 × 10^{-8}), and another on chromosome 7 within the ARPC1A gene (rs740160 located 161 kb downstream of rs11761528, p = 1.56 × 10^{-6}) (Table 2; Figure 1D). In total, we found eight independent SNPs associated with serum DHEAS concentrations at conventional genome-wide significant level (p<5 × 10^{-8}) (Table 2). The effect was consistently in the same direction across all cohorts (Table 2). No heterogeneity among cohorts was observed (Table 2). These SNPs together explained ~4% of the total and ~7% of genetic variance of serum DHEAS concentrations based on TwinsUK data. To further look at whether the magnitude of these genetic association varies with age,

**Table 2. SNPs associated with serum DHEAS concentrations: genome-wide results of meta-analysis of men and women combined.**

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Position in base pair</th>
<th>Freq</th>
<th>Effect Allele</th>
<th>Beta (SE)*</th>
<th>P value</th>
<th>( \ldots ) index</th>
<th>Effect direction in each study</th>
<th>Gene</th>
<th>Distance to the gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery meta-analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1761526</td>
<td>7</td>
<td>98956737</td>
<td>0.08</td>
<td>T</td>
<td>-0.16 (0.01)</td>
<td>3.15 × 10^{-6}</td>
<td>0.12</td>
<td>-----------------</td>
<td>ZKSCAN5</td>
<td>intron</td>
</tr>
<tr>
<td>rs2637125</td>
<td>19</td>
<td>53093705</td>
<td>0.15</td>
<td>A</td>
<td>0.09 (0.01)</td>
<td>2.61 × 10^{-9}</td>
<td>0.00</td>
<td>++ + + + + + + +</td>
<td>SULT2A1</td>
<td>12 kb</td>
</tr>
<tr>
<td>rs7181230</td>
<td>15</td>
<td>38148033</td>
<td>0.33</td>
<td>G</td>
<td>0.05 (0.01)</td>
<td>5.44 × 10^{-11}</td>
<td>0.00</td>
<td>+ + + + + + + +</td>
<td>BMF</td>
<td>23 kb</td>
</tr>
<tr>
<td>rs2497306</td>
<td>10</td>
<td>94475191</td>
<td>0.49</td>
<td>C</td>
<td>0.04 (0.01)</td>
<td>4.64 × 10^{-9}</td>
<td>0.00</td>
<td>+ + + + + + + +</td>
<td>HHEX</td>
<td>25 kb</td>
</tr>
<tr>
<td>rs2185570</td>
<td>10</td>
<td>96741260</td>
<td>0.13</td>
<td>C</td>
<td>0.06 (0.01)</td>
<td>2.29 × 10^{-8}</td>
<td>0.00</td>
<td>+ + + + + + + +</td>
<td>CYP2C9</td>
<td>-2 kb</td>
</tr>
<tr>
<td>Conditional analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs740160</td>
<td>7</td>
<td>98795816</td>
<td>0.05</td>
<td>T</td>
<td>0.15 (0.02)</td>
<td>1.56 × 10^{-6}</td>
<td>0.02</td>
<td>+ + + + + + + +</td>
<td>ARPC1A</td>
<td>intron</td>
</tr>
<tr>
<td>rs17277546</td>
<td>7</td>
<td>99327507</td>
<td>0.05</td>
<td>A</td>
<td>-0.11 (0.02)</td>
<td>4.50 × 10^{-11}</td>
<td>0.00</td>
<td>+ + + + + + + +</td>
<td>TRIM4/CYP3A43 UTR</td>
<td></td>
</tr>
<tr>
<td>rs6738028</td>
<td>2</td>
<td>111665798</td>
<td>0.40</td>
<td>G</td>
<td>-0.04 (0.01)</td>
<td>1.72 × 10^{-8}</td>
<td>0.00</td>
<td>+ + + + + + + +</td>
<td>BCL2L11</td>
<td>-62 kb</td>
</tr>
</tbody>
</table>

*Beta was expressed as natural log changes in serum DHEAS concentration in µmol/L per copy of the risk allele.
\*Index for between-study heterogeneity: 0.25 – low, 0.50 – moderate and 0.75 – high heterogeneity.

1pre-conditional p values were 0.612, 1.90 × 10^{-9}, and 1.94 × 10^{-7} for rs740160, rs17277546, and rs6738028, respectively.

doi:10.1371/journal.pgen.1002025.t002
we carried out an interaction analysis between age and each of these 8 SNPs on serum DHEAS concentrations by including an interaction term of age×SNP in the linear regression model in each cohort and then meta-analyzed the results. We found that there was no significant interaction between age and each of these SNPs (all p values ≥0.05).

The genes at, or near the identified SNPs, include BCL2L11 on chromosome 2, ZKSCAN5, ARPC1A, TRIM4 and CYP3A43 on chromosome 7, HHEX and CYP2C9 on chromosome 10, BMF on chromosome 15, and SULT2A1 on chromosome 19 (Figure 2). To explore the potentially functional impacts and likely genetic mechanisms, we used two resources: Genome-wide expression data from the Multiple Tissue Human Expression Resource (MuTHER) [28] (http://www.muther.ac.uk/) based on ~777 unselected UK twins sampled for skin, adipose tissue, and lymphoblastoid cell lines (LCLs) (more details in Text S1); and published gene expression data in human liver [29]. We found that 3 DHEAS-associated SNPs were clearly associated with the related

Figure 1. Manhattan plots for the genome-wide meta-analysis results. Green dots indicate the SNPs with $p<5\times10^{-8}$. doi:10.1371/journal.pgen.1002025.g001
gene expression levels in at least one tissue after accounting for multiple testing (Table 3). These specific transcript associations provide further evidence for the likely functional gene at each locus.

Further, we carried out gene ontology and pathway analyses using a gene set enrichment analysis (GSEA) approach in MAGENTA [30] which consists of four main steps: First, DNA variants, e.g. SNP, are mapped onto genes. Second, each gene is assigned a gene association score that is a function of its regional SNP association p-values. Third, confounding effects on gene association scores are identified and corrected for, without requiring genotype data. Fourth, a GSEA-like statistical test is applied to predefined biologically relevant gene sets to determine whether any of the gene sets are enriched for highly ranked gene association scores compared to randomly sampled gene sets of identical size from the genome. More details of these four steps are described in the method section. In this analysis, we identified three pathways which passed our significance threshold (false discovery rate (FDR) < 0.05): xenobiotic metabolism with FDR = 0.001 (pathway database: KEGG and Ingenuity), retinoid X receptor (RXR) function with FDR = 0.003 (pathway database: Ingenuity), and linoleic acid metabolism with FDR = 0.02 (pathway database: KEGG) (Figure S1). The top significant genes include CYP2A6, CYP3A4, CYP3A5, and CYP5A1 on chromosome 10, and CYP2C8 and CYP2C9 on chromosome 10 for all three pathways, and SULT2A1 for RXR pathway. The best index SNPs are rs17277546 for CYP2A6 and CYP3A4, rs4646450 for CYP3A5 and CYP5A1, rs11761528 for CYP2C8, and rs2637125 for SULT2A1. The full list of the genes in each of the three pathways and the best index SNPs for each gene are listed in Table S2. Three SNPs – rs17277546, rs2185570, and rs2637125 are the DHEAS-associated SNPs found in our meta-analysis. Both rs4646450 and rs11761528 were associated with DHEAS with p values of $8.8 \times 10^{-17}$ and $4.8 \times 10^{-16}$, respectively, but become non-significant in the conditional meta-analysis because rs4646450 is in linkage disequilibrium (LD, $r^2 = 0.429$) with rs11761528 which is the most significant DHEAS-associated SNP while rs11761528 is in high LD ($r^2 = 0.778$) with rs2185570. Intriguingly, two pathways - xenobiotic metabolism and linoleic acid metabolism, have been linked to ageing in model organisms [31–36].

**Discussion**

This is the first meta-analysis of GWA studies on serum DHEAS in 14,846 Caucasian subjects. We found 8 common SNPs that implicate nearby genes that are independently associated with serum DHEAS concentrations and provide clues to its role in ageing.

Among the genes identified, SULT2A1, a specialized sulphotransferase which converts DHEA to DHEAS in the adrenal cortex, is an obvious candidate gene [3]. SULT2A1 has a broad substrate specificity, which includes conversion of pregnenolone, 17α-hydroxypregnenolone, and DHEA to their respective sulphated products [37]. Once sulphated by SULT2A1, pregnenolone and 17α-hydroxypregnenolone are no longer available as substrates for HSD3B2. Therefore, SULT2A1 sulphation of pregnenolone and 17α-hydroxypregnenolone removes these substrates from the mineralocorticoid and glucocorticoid biosynthetic pathways. This suggests that high levels of SULT2A1 would ensure the formation of DHEAS [3].

Variation in SULT2A1 expression has previously been associated with variation of DHEAS concentration [38]. The SULT2A1 gene is predominantly expressed in the adrenal cortex and to a lesser extent in the liver. We found that rs2547231 (p = 1.76 x 10^{-17}), located 12 kb downstream of SULT2A1, was strongly associated with expression levels of SULT2A1 in human liver tissues. Although this SNP is not the most strongly associated with serum DHEAS, it is itself in strong LD with the most significant SNP rs2637125 ($r^2 = 0.658$). However, we did not find a significant association with SULT2A1 expression levels in LCL, skin, and adipose tissues, suggesting a tissue specific effect. The SULT2B1b is also reported to play a role in sulphation of DHEA, but in comparison, the strongest signal from that genomic region was rs10417472 with a p = 0.06. In contrast, enzymatic removal of the sulphate group from DHEAS to form DHEA is performed by steroid sulphatase gene (STS), but that gene is on the X chromosome and so was not assessed in this meta-analysis.

CYP2C9 is an important cytochrome P450 enzyme, accounts for approximately 17–20% of the total P450 content in human liver, and catalyzes many reactions involved in drug metabolism as well as synthesis of cholesterol, steroids and other lipids [39]. We found that rs2185570 located in the CYP2C9 gene region is associated with variables.
with serum DHEAS concentrations. This SNP is in strong LD with rs4066116 and rs4917639 ($r^2 = 0.67$ for both) which were found to be associated with acenocoumarol [40] and warfarin maintenance dosage [41] respectively in recent GWAS.

Two other cytochrome P450 enzymes – CYP11A1 and CYP17A1, are two important enzymes which are required in the synthesis of DHEAS in the adrenal gland [3], however, the strongest signals in the genomic region were rs2993036 with $p = 0.29$ for CYP11A1 and rs4919686 with $p = 0.04$ for CYP17A1.

The decline in serum DHEAS concentrations with increasing age has been proposed as a putative biomarker for ageing [21]. We found that two putative ageing genes – BCL2L11 and BMF [42] are associated with serum DHEAS concentrations. Both of them encode proteins which belong to the BCL2 family and act as anti- or pro-apoptotic regulators that are involved in a wide variety of cellular activities. BCL2L11 has been implicated in chronic lymphocytic leukaemia (rs17483466, $P = 2.36 \times 10^{-10}$) [43] and follicular lymphoma (rs3789068, $P$ for trend = 0.0004) [44]. The DHEAS-associated SNP rs6738028 is not however one of the same SNPs associated with lymphocytic leukaemia and follicular lymphoma nor is it in LD with them. Nevertheless, rs6738028 is strongly associated with BCL2L11 gene expression levels in both LCL and adipose tissues, suggesting its putative regulatory role. There is relatively little data on the human BMF gene or the protein product, but Bmf$^{-/-}$ mice show altered immune and hematopoietic phenotypes as well as defects in uterovaginal development. However, we were not able to detect any association between rs1781230 and the expression levels of BMF in the tissues we studied.

HHEX encodes a member of the homeobox family of transcription factors, many of which are involved in developmental processes. This gene has been found to be associated with type 2 diabetes by several recent GWAS [45–51]. The risk alleles of the diabetes-associated SNPs rs1111875 and rs5015480 are associated with low serum DHEAS concentrations although the $p$ values ($p = 0.0009$ for both SNPs) didn’t reach to the GWAS significance level. This is consistent with the observation in which the low serum DHEAS concentrations were associated with insulin resistance [6,7]. The identified DHEAS-associated SNP rs2497306 is in moderate LD with rs1111875 and rs5015480 ($r^2 = 0.38$). And the major allele of rs2497306 is associated with increasing serum DHEAS concentrations. The reason for the observed association is unknown. Studies showed that insulin infusion increases the metabolic clearance of DHEA and DHEAS [52,53], resulting in decreased DHEA and DHEAS concentrations, and DHEA administration significantly enhances insulin sensitivity attenuating the age-related decline in glucose tolerance [54], partly explaining why the diabetes-associated gene is also associated with DHEAS. Interestingly, HHEX null mice show cardiovascular, endocrine, liver, muscle, nervous system, and metabolic phenotypes, suggesting extensive multisystem roles for the protein product of this gene. The findings could help dissect causal pathways for the observed associations between DHEAS, insulin resistance, age-related decline in glucose tolerance [54], and other age related phenotypes [55].

Three identified DHEAS-associated SNPs on chromosome 7 (Figure S2), which were independent, and 161 kb downstream (rs740160) and 370 kb upstream (rs17277546) apart from rs1176528 which is located in the middle of the region, are located in four genes - ZKSCAN1, ARPC1A, and TRIM4/CYP3A43. ZKSCAN5 encodes a zinc finger protein of the Kruppel family and is expressed ubiquitously in adult and fetal tissues with the strongest expression in testis [56]. rs1176528 is located in the intron of the ZKSCAN5 gene. It is the strongest signal we found and explains 1% of the total variance of serum DHEAS concentration alone. ARPC1A encodes one of seven subunits of the human Arp2/3 protein complex which has been implicated in actin polymerization and filament assembly in cells [57]. TRIM4 encodes a member of the tripartite motif (TRIM) family whereas CYP3A43 is another cytochrome P450 enzyme. The potential mechanisms for the association are unknown, but we found that rs17277546 is strongly associated with expression levels of TRIM4 not CYP3A43, suggesting TRIM4 is the possible candidate for DHEAS. However, rs17277546 is the best index SNP for both CYP3A45 and CYP3A44 genes in the pathway analysis, indicating a fine mapping in this region is needed to reveal the potential mechanism for the association. Further, the region harbours many other genes including CIP5A7 which has been reported to increase the clearance of DHEA and DHEAS [59,60] and a common haplotype polymorphism in the gene has been associated with DHEAS [59,60]. However, none of the DHEAS-associated SNPs are associated with its expression levels in the tissues we studied, and the best index SNP rs4646450 for CYP5A7 found in our pathway analysis is in LD with rs1176528 and become non-significant in the conditional analysis.

In the pathway analysis, two DHEAS-associated SNPs (rs2105570 and rs17277546) were contained in all three pathways we found and one SNP (rs2637125) was contained in the XR4 function pathway. Intriguingly, components of the xenobiotic metabolism pathway have been linked to ageing in model organisms, for example, age-associated changes in expression of genes involved in xenobiotic metabolism have been identified in rats [31,32], up-regulation of xenobiotic detoxification genes has been observed in long-lived mutant mice [33], and adrenal xenobiotic-metabolizing activities increase with ageing in guinea pigs [34]. Furthermore, linoleic acid metabolism has also been linked to changes with ageing in rat cardiac muscle [35] and in human skin fibroblasts [36]. Taken together, these findings suggest that molecular pathways involved in ageing and longevity may also underlie DHEAS regulation, suggesting shared genetic components in both processes and corroborating a role for DHEAS as a marker of biological ageing.

In summary, this first GWAS identified eight independent SNPs associated with serum DHEAS concentrations. The related genes have various associations with steroid hormone metabolism, comorbidities of ageing including type 2 diabetes, lymphoma, actin filament assembly, drug and xenobiotic metabolism, and zinc fingers - suggesting a wider functional role for DHEAS than previously thought.

Methods

Study population

Seven study samples contributed to this meta-analysis of GWA studies on serum DHEAS concentrations, comprising a total of 14,046 men and women of Caucasian origin. The consortium was made up of populations from TwinsUK (n = 4,906), Framingham Heart Study (FHS) (n = 3,183), SHIP (n = 1,832), Rotterdam Study (RS1) (n = 1,597), InCHIANTI (n = 1,182), Health ABC (n = 1,222), and GOOD (n = 924). Full details can be found in Text S1.

DHEAS methods

Blood samples were collected from each of the study participants either after overnight fasting or non-fasting and the serum levels of DHEAS were measured by either immunoassay or liquid chromatography tandem mass spectrometry (LC-MS/MS) methods (Text S1). Because the distribution of the serum DHEAS

GWAS of DHEAS
levels was skewed, we log transformed the concentrations and the transformed data used in the subsequent analysis.

**Genotyping and imputation**

Seven study populations were genotyped using a variety of genotyping platforms including Illumina (HumanHap 317k, 550k, 610k, 1M-Duo BeadChip) and Affymetrix (array 500K, 6.0). Each cohort followed a strict quality control on the genotyping data. More details on the quality control and filtering criteria can be found in Text S1. In order 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 in either MACH [http://www.sph.umich.edu/csg/abecasis/MACH], or IMPUTE [61]. We exclude non-genotyped SNPs with an imputation quality score <0.2 and SNPs with allele frequency <0.01 from meta-analysis.

**Statistical method**

Each study performed genome-wide association testing for serum concentrations of DHEAS across approximately 2.5 million SNPs under an additive genetic model separately in men and women (Text S1). The analyses were adjusted for age. In addition, the association testing was performed in the combined men and women data with adjustment for age and sex. Studies used PLINK, GenABEL, SNPTEST, QUICKTEST, or MERLIN for the association testing. The summary results from each cohort were meta-analyzed by Z-score pooling method implemented in Metal [http://www.sph.umich.edu/csg/abecasis/Metal/]. We chose this method to minimize the impact of the different assays used for serum DHEAS measurements. Specifically, for each study, we converted the two-sided P value after adjustment for population stratification by the genomic control method to a Z statistic that was signed to reflect the direction of the association given the reference allele. Each Z score was then weighted; the squared weights were chosen to sum to 1, and each sample-specific weight was proportional to the square root of the effective number of individuals in the sample. We summed the weighted Z statistics across studies and converted the summary Z score to a two-sided P value. We also used I^2 index to assess between-study heterogeneity and the inverse variance weighted method to estimate the effect size. Genome-wide significance was defined as p<5×10^{-8}. The association between the DHEAS-associated SNPs and the related gene expression levels in MuTHER data were examined by mixed linear regression modelling which takes both family structure and gene expression levels in MuTHER data were examined by mixed linear regression modelling which takes both family structure and gene expression levels into account. The significance was determined as p<0.006 after accounting for multiple testing (Bonferroni method, correcting 9 independent tests).

**Pathway analysis.** Meta-Analysis Gene-set Enrichment of variaNT Associations (MAGENTA) was used to explore pathway-based associations in the full GWAS dataset. MAGENTA implements a gene set enrichment analysis (GSEA) based approach, the methodology of which is described in Segre et al [30]. Briefly, each gene in the genome is mapped to a single index SNP with the lowest P-value within a 110 kb upstream, 40 kb downstream window. This P-value, representing a gene score, is then corrected for confounding factors such as gene size, SNP density and LD-related properties in a regression model. Genes within the HLA-region were excluded from analysis due to difficulties in accounting for gene density and LD patterns. Each mapped gene in the genome is then ranked by its adjusted gene score. At a given significance threshold (5th and 75th percentiles of all gene scores), the observed number of gene scores in a given pathway, with a ranked score above the specified threshold percentile, is calculated. This observed statistic is then compared to 1,000,000 randomly permuted pathways of identical size. This generates an empirical GSEA P-value for each pathway. Significance was determined when an individual pathway reached a false discovery rate (FDR)<0.05 in either analysis. In total, 2529 pathways from Gene Ontology, PANTHER, KEGG and Ingenuity were tested for enrichment of multiple modest associations with serum DHEAS levels.

**Ethics statement**

All studies were approved by local ethics committees and all participants provided written informed consent as stated in Text S1.

**Supporting Information**

**Figure S1** Three pathways which were associated with DHEAS. The genes which are near the DHEAS-associated SNPs are highlighted by red circles. a. Xenobiotic metabolism pathway; b. Retinoid X receptor (RXR) function pathway; c. Linoleic acid metabolism pathway; d. Legends for the pathway figures. The pathway figures were made using MetaCore from GeneGo (http://www.genego.com/metaknowledge.php). (TIF)

**Figure S2** Regional linkage disequilibrium plots for three SNPs on chromosome 7 in one plot. (TIF)

**Table S1** 87 SNPs associated with DHEAS in men, women, and combined meta-analysis with p<5×10^{-8}. (XLS)

**Table S2** Pathway analysis results – list of all pathways, significant pathways, and significant genes with the best index SNPs. (XLS)

**Text S1** Supplementary Note. (DOC)

**Acknowledgments**

**TwinsUK:** The study was funded by the Wellcome Trust, Arthritis Research Campaign, European Community’s Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F2-2008-201865-GEFOS and Seventh Framework Programme grant 200000 Treat OA/ FP7/2007-2013), ENGAGE project grant agreement HEALTH-F1-2007-201413, and the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from the UK Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy’s & St Thomas’ NHS Foundation Trust in partnership with King’s College London. TDS is an NIHR senior investigator. The project also received support from a Biotechnology and Biological Sciences Research Council project grant (G200254). The authors acknowledge the funding and support of the National Eye Institute (NEI) via an NIH/Center for Inherited Disease Research (CIDR) genotyping project. We thank the staff from the Genotyping Facilities at the Wellcome Trust Sanger Institute for sample preparation, quality control, and genotyping; Le Centre National de Génotypage, France, for genotyping; Duke University, NC, USA, for genotyping; and the Finnish Institute of Molecular Medicine, Finnish Genome Center, University of Helsinki. Genotyping was also done by CIDR as part of an NEI/NIH project grant. **SHIP:** Computing resources have been made available by the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities (HLRB project h1251). SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9505, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural
Affairs as well as the Social Ministry of the Federal State of Mecklenburg - West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg West Pomerania. This work is also part of the research project Greifswald Approach to Individualized Medicine (GANI_MED). The GANI_MED consortium is funded by the Federal Ministry of Education and Research and the Ministry of Cultural Affairs of the Federal State of Mecklenburg – West Pomerania (03IS2061A). The University of Greifswald is a member of the ‘Center of Knowledge Interchange’ program of the Siemens AG. The DHEAS reagents used were sponsored by Siemens Healthcare Diagnostics, Eschborn, formerly DPC Biemann GmbH, Bad Nauheim, Germany. Novo Nordisk provided partial grant support for the determination of serum samples and data analysis.

Health ABC: This study was supported by National Institute on Aging contracts N01-AG-2-2101, N01-AG-6-2103, and N01-AG-6-2106. The genome-wide association study was funded by NIA grant 1R01AG032990-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health. The University of Greifswald is a member of the ‘Center of Knowledge Interchange’ program of the Siemens AG. The DHEAS reagents used were sponsored by Siemens Healthcare Diagnostics, Eschborn, formerly DPC Biemann GmbH, Bad Nauheim, Germany. Novo Nordisk provided partial grant support for the determination of serum samples and data analysis.

RS1: The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-013; RIDE2), the Netherlands Genomics Initiative (NGI) - Netherlands Consortium of Healthy Ageing (NCHA) project nr. 050-060-018 and funded by the European Commission (HEALTH-F2-2006-201865, GEFOS; HEALTH-F2-2008-35627, TREAT-OA). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMWs), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Ministry of Industry and Trade of the Province of Rotterdam. The Dutch Genomics Consortium was supported by the Dutch Ministry of Economic Affairs as part of the Next Generation Sequencing program (NGIS). The database was supported by the National Institutes of Health (http://www.ncbi.nlm.nih.gov). The authors wish to thank Dr. Tobias A. Knoch, Karol Estrada, Luc V. de Zeeuw, Anis Abseirs, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technologie under grants #01 AK 803 A-A and # 01 IG 07015 G for access to their grid resources. We would also like to thank Karol Estrada, Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands for advice regarding the grid resources.

InCHIANTI: The InCHIANTI study baseline (1998–2000) was supported as a “targeted project” (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9166 and 263 MD R21336); the InCHIANTI Follow-up 1 (2001–2003) was supported by the U.S. National Institute on Aging (Contracts: N1-AG-1 and N1-AG-12111); the InCHIANTI Follow-up 2 and 3 studies (2004–2010) were financed by the U.S. National Institute on Aging (Contract: N01-AJ-5-0020). Support was provided by the Institutional Research Program of the National Institute on Aging, National Institutes of Health, Baltimore, Maryland, with additional support from the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.


References