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^{-3}), SULT2A1 (rs2637125; p = 2.61 × 10^{-3}), ARPC3 (rs740160; p = 1.56 × 10^{-16}), TRIM4 (rs17277546; p = 6.50 × 10^{-11}), BMF (rs7181230; p = 5.44 × 10^{-11}), HHEX (rs2497306; p = 4.64 × 10^{-3}), BCL2L11 (rs6738028; p = 1.72 × 10^{-3}), and CYP2C9 (rs2155570; p = 2.29 × 10^{-3}). 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.

* E-mail: tim.spector@kcl.ac.uk

† These authors contributed equally to this work.

‡ These senior authors were joint senior authors on this work.
Introduction

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 by the adrenal gland, is the most abundant circulating steroid in humans. It shows a significant physiological decline after the age of 25 and diminishes about 95% by the age of 85 years, which 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 concentrations 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, Trim4, 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.

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 by the adrenal gland, is the most abundant circulating steroid in humans. It shows a significant physiological decline after the age of 25 and diminishes about 95% by the age of 85 years, which 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 concentrations 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 14,946 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, λG, which is defined as the median Z (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 λG 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 = 3.15×10⁻⁸) on chromosome 7q22.1, rs2637125 (p = 2.61×10⁻⁸) on chromosome 19q13.3, and rs2497306 (p = 4.6×10⁻⁸) 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 \times 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, 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' UTR of the TRIM4 and CYP3A43 genes (rs17277546, p = 4.50 \times 10^{-11}). Furthermore, we identified two additional significant loci associated with DHEAS, one on chromosome 2q13 (rs6738028, p = 1.72 \times 10^{-8}), and another on chromosome 7 within the ARPC1A gene (rs740160 located 161 kb downstream of rs11761528, p = 1.56 \times 10^{-16}) (Table 2; Figure 1D). In total, we found eight independent SNPs associated with serum DHEAS concentrations at conventional genome-wide significant level (p<5 \times 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 1. Descriptive statistics of serum levels of DHEAS (μmol/L) for each cohort.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Assay</th>
<th>Mean Age (Range)</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>R51</td>
<td>Immunoassay</td>
<td>69 (55–98)</td>
<td>4.34</td>
<td>2.88</td>
<td>3.70</td>
<td>0.10</td>
<td>23.08</td>
<td>22.98</td>
<td>740</td>
</tr>
<tr>
<td>SHIP</td>
<td>Immunoassay</td>
<td>51 (20–79)</td>
<td>1.90</td>
<td>1.21</td>
<td>1.64</td>
<td>0.31</td>
<td>8.90</td>
<td>8.59</td>
<td>1832</td>
</tr>
<tr>
<td>FHS</td>
<td>Immunoassay</td>
<td>51 (25–80)</td>
<td>7.05</td>
<td>5.07</td>
<td>5.35</td>
<td>0.27</td>
<td>29.86</td>
<td>29.59</td>
<td>1571</td>
</tr>
<tr>
<td>GOOD</td>
<td>MassSpec</td>
<td>19 (18–20)</td>
<td>6.31</td>
<td>2.33</td>
<td>6.04</td>
<td>1.27</td>
<td>15.10</td>
<td>13.83</td>
<td>924</td>
</tr>
<tr>
<td>InCHIANTI</td>
<td>Immunoassay</td>
<td>67 (23–94)</td>
<td>3.16</td>
<td>2.98</td>
<td>2.25</td>
<td>0.02</td>
<td>33.06</td>
<td>33.04</td>
<td>518</td>
</tr>
<tr>
<td>HABC</td>
<td>Immunoassay</td>
<td>74 (69–80)</td>
<td>1.58</td>
<td>1.12</td>
<td>1.40</td>
<td>0.00</td>
<td>9.93</td>
<td>9.93</td>
<td>696</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.
*Pre-conditional p values were 0.612, 1.90 \times 10^{-10}, and 1.94 \times 10^{-7} for rs740160, rs17277546, and rs6738028, respectively.

doi:10.1371/journal.pgen.1002025.t001

doi:10.1371/journal.pgen.1002025.t002

### 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>I^2 index*</th>
<th>Effect direction in each study</th>
<th>Gene</th>
<th>Distance to the gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discovery meta-analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11761528</td>
<td>7</td>
<td>98956737</td>
<td>0.08</td>
<td>T</td>
<td>0.16 (0.01)</td>
<td>3.15 \times 10^{-6}</td>
<td>0.12</td>
<td>--</td>
<td>--</td>
<td>ZKSCAN5</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 \times 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 \times 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 \times 10^{-9}</td>
<td>0.00</td>
<td>--</td>
<td>HHEX</td>
<td>25 kb</td>
</tr>
<tr>
<td>rs2185570</td>
<td>19</td>
<td>96741260</td>
<td>0.13</td>
<td>C</td>
<td>0.06 (0.01)</td>
<td>2.29 \times 10^{-8}</td>
<td>0.00</td>
<td>--</td>
<td>CYP2C9</td>
<td></td>
</tr>
</tbody>
</table>

| **Conditional analysis** |     |                       |      |               |            |         |            |                               |                       |                     |
| rs740160       | 7   | 98795816              | 0.05 | T             | 0.15 (0.02) | 1.56 \times 10^{-6} | 0.02 | ++-------- | ARPC1A                     | intron                |                     |
| rs17277546    | 7   | 99327507              | 0.05 | A             | 0.11 (0.02) | 4.50 \times 10^{-11} | 0.00 | --         | TRIM4/CYP3A43                | 3'UTR                 |                     |
| rs6738028     | 2   | 111665798             | 0.40 | G             | 0.04 (0.01) | 1.72 \times 10^{-8} | 0.00 | --         | BCL2L11                     |                     | --62 kb             |

*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.
*Pre-conditional p values were 0.612, 1.90 \times 10^{-10}, and 1.94 \times 10^{-7} for rs740160, rs17277546, and rs6738028, respectively.
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<5×10^{-8}. doi:10.1371/journal.pgen.1002025.g001
Figure 2. Regional linkage disequilibrium plots. For rs6738028 (A), rs740160 (B), rs11761528 (C), rs17277546 (D), rs2497306 (E), rs2185570 (F), rs7181230 (G), and rs2637125 (H). Note: p values from the conditional analysis were used for (a) and (b), both of them became genome-wide significant in the conditional analysis. Annotation key: ▲ - framestop or splice; ▼ - Non-Synonymous; ■ - Synonymous or UTR; ● - nothing; * - TFBScons; ** - MCS44 Placental.

doi:10.1371/journal.pgen.1002025.g002

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 with p<5.0×10^{-8} include CYP3A4, CYP3A43, CYP3A5, and CYP3A7 on chromosome 7, and CYP2C8 and CYP2C9 on chromosome 10 for all three pathways, and SULT2A1 for RXR pathway. The best index SNPs are rs17277546 for CYP3A4 and CYP3A43, rs4646450 for CYP3A5 and CYP3A7, rs2185570 for CYP2C9, rs11572169 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 rs11572169 were associated with DHEAS p values of 8.8×10^{-17} and 4.8×10^{-5}, 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 rs11572169 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].

Table 3. Association between DHEAS-associated SNPs and related gene expression levels in different human tissues.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>SNP (effect allele)</th>
<th>Position</th>
<th>LCL* (n = 777)</th>
<th>Adipose tissue* (n = 776)</th>
<th>Skin tissue* (n = 667)</th>
<th>Liver tissue† (n = 427)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Beta (SE)</td>
<td>P value</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
</tr>
<tr>
<td>BCL2L11</td>
<td>2</td>
<td>rs6738028 (G)</td>
<td>111665798</td>
<td>0.07 (0.02)</td>
<td>0.0003</td>
<td>0.02 (0.005)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRIM4</td>
<td>7</td>
<td>rs17277546 (A)</td>
<td>99327507</td>
<td>0.15 (0.04)</td>
<td>0.0001</td>
<td>0.13 (0.04)</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SULT2A1</td>
<td>19</td>
<td>rs2637125 (A)/rs2457231**</td>
<td>53093705</td>
<td>0.0006 (0.007)</td>
<td>0.93</td>
<td>-0.009 (0.007)</td>
<td>0.19</td>
</tr>
</tbody>
</table>

*from MuTHER consortium and beta (SE) were from linear regression modelling; LCL – lymphoblastoid cell lines.

* from reference 27 and effect size was not reported.

**P value in liver expression is for rs2547231, data is not available for rs2637125, but two SNPs are in strong LD (r^2 = 0.658).
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 – CYP1A1 and CYP1A2, 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 CYP1A1 and rs4919686 with $p = 0.04$ for CYP1A2. The decline in serum DHEAS concentrations with increasing age has been proposed as a putative biomarker of ageing [21]. We found that two putative ageing genes – BCL2L11 and BMF [42] are associated with serum DHEAS concentrations. Both of these 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.

*HHHEX* 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, *HHHEX* 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 rs11761528 which is located in the middle of the region, are located in four genes - ZKSCAN3, ARPC1A, and TRIM4/CYP3A43. ZKSCAN3 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]. rs11761528 is located in the intron of the ZKSCAN3 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 CYP5A7 which has been reported to increase the clearance of DHEA and DHEAS [50] 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 rs11761520 and become non-significant in the conditional analysis.

In the pathway analysis, two DHEAS-associated SNPs (rs2185570 and rs17277546) were contained in all three pathways we found and one SNP (rs2637125) was contained in the XR2 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 = 1,892), 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
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, SNPTTEST, QUICKTEST, or MERLIN for the association testing. The summary results from each cohort were meta-analyzed by Z-score pooling method implemented in Meta (http://www.sph.umich.edu/csg/abecasis/mach/). 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 included index to assess between-study heterogeneity and the inverse variance weighted method to estimate the effect size. Genome-wide significance was defined as \( p < 5 \times 10^{-8} \). 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 batch effects 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 (95th 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. Retinoic 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/metacore.php).

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

Table S1 87 SNPs associated with DHEAS in men, women, and combined meta-analysis with p<5\times10^{-8}.

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

Text S1 Supplementary Note.
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 ‘Centre of Knowledge Interchange’ program of the Siemens AG. The DHEAS reagents used were sponsored by Siemens Healthcare Diagnostics, Eschborn, formerly DPC Bierrmann GmbH, Bad Nauheim, Germany. Novo Nordisc 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-6-2101, N01-AG-6-2105, and N01-AG6-2106. The genome-wide association study was funded by NIA grant 1RO1AG032990-01A1 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 to The Johns Hopkins University, contract number HHHS268200720096C. This research was supported (in part) by the Intramural Research Program of the NIH, National Institute on Aging.

R51: 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 (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI) - Netherlands Consortium of Healthy Aging (NCHA) project nr. 050-060-0182 and funding the European Commission (HEALTH-F2-2008-201865, GEFOs: HEALTH-F2-2008-35627, TREAT-0014). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), 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 Science and Technology of Rotterdam. We thank Pascal Arp, Mila Jhamai, Dr. Michael Moorthoor, Marijn Verkert, and Sander Bervots for their help in creating the GWAS database. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We would like to thank Dr. Tobias A. Knoch, Karol Estrada, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MedGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technologie under grants #01 AK 803-AH and #01 IG 07015 G for access to their grid resources.

FHS: DHEAS-genotype genome-wide association analyses were supported by the National Institute of Aging (Genetics of Reproductive Life Period and Health Outcomes, R21AG023598, JMM, KI, DPK, WWZ). “The Framingham Heart Study of the National Heart and Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine was supported by the National Heart, Lung and Blood Institute’s Framingham Heart Study Contract no. N01-HC-25195 and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). Analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHArE) project. A portion of this research was conducted using the Linux Cluster for Genetic Analysis (LinGa-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.” DPK received support from NIH (Grant R01 AR/AG 41398).

**GOOD:** Financial support was received from the Swedish Research Council (K2010-5X-05094-19-3, 2006-3032), the Swedish Foundation for Strategic Research, the ALE/LUA research grant, Lundbeckfonden, the Lundberg Foundation, the Torsten and Ragnar Söderberg’s Foundation, Petrus and Augusta Hedlunds Foundation, the Västra Götaland Foundation, the Göteborg Medical Society, the Novo Nordisk foundation and the European Commission grant HEALTH-F2-2008-201865-GEFOS. We would like to thank Dr. Tobias A. Knoch, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MedGRID 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-H 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/RP97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD R21336); the InCHIANTI Follow-up 1 (2001–2003) was funded by the U.S. National Institute on Aging (Contracts: N1-AG-1-I and N1-AG-1-I-2111); the InCHIANTI Follow-ups 2 and 3 studies (2004–2010) were financed by the U.S. National Institute on Aging (Contract: N01-AG-5-0002); supported in part by the Intramural research program of the National Institute on Aging, National Institutes of Health, Baltimore, Maryland. JRPB is funded by a Sir Henry Wellcome Postdoctoral Fellowship (092417/Z/10/Z).

**Author Contributions:**


