

# Gene Set of Nuclear-Encoded Mitochondrial Regulators Is Enriched for Common Inherited Variation in Obesity

Nadja Knoll<sup>1</sup>, Ivonne Jarick<sup>2</sup>, Anna-Lena Volckmar<sup>1</sup>, Martin Klingenspor<sup>3</sup>, Thomas Illig<sup>4,5</sup>, Harald Grallert<sup>4</sup>, Christian Gieger<sup>6</sup>, Heinz-Erich Wichmann<sup>7,8,9</sup>, Annette Peters<sup>10</sup>, Johannes Hebebrand<sup>1</sup>, André Scherag<sup>11</sup>, Anke Hinney<sup>1\*</sup>

**1** Department of Child and Adolescent Psychiatry, University of Duisburg-Essen, Essen, Germany, **2** Institute of Medical Biometry and Epidemiology, Philipps-University of Marburg, Marburg, Germany, **3** Molecular Nutritional Medicine, Technical University of Munich, Else Kröner-Fresenius Center, Freising-Weihenstephan, Germany, **4** Research Unit of Molecular Epidemiology, Helmholtz Center Munich – German Research Center for Environmental Health, Neuherberg, Germany, **5** Hannover Unified Biobank, Hannover Medical School, Hannover, Germany, **6** Institute of Genetic Epidemiology, Helmholtz Center Munich – German Research Center for Environmental Health, Neuherberg, Germany, **7** Institute of Epidemiology I, Helmholtz Center Munich – German Research Center for Environmental Health, Neuherberg, Germany, **8** Institute of Medical Informatics, Biometry, and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany, **9** Munich University Hospital, Campus Grosshadern, Munich, Germany, **10** Institute of Epidemiology II, Helmholtz Center Munich – German Research Center for Environmental Health, Neuherberg, Germany, **11** Institute for Medical Informatics, Biometry and Epidemiology, University of Duisburg-Essen, Essen, Germany

## Abstract

There are hints of an altered mitochondrial function in obesity. Nuclear-encoded genes are relevant for mitochondrial function (3 gene sets of known relevant pathways: (1) 16 nuclear regulators of mitochondrial genes, (2) 91 genes for oxidative phosphorylation and (3) 966 nuclear-encoded mitochondrial genes). Gene set enrichment analysis (GSEA) showed no association with type 2 diabetes mellitus in these gene sets. Here we performed a GSEA for the same gene sets for obesity. Genome wide association study (GWAS) data from a case-control approach on 453 extremely obese children and adolescents and 435 lean adult controls were used for GSEA. For independent confirmation, we analyzed 705 obesity GWAS trios (extremely obese child and both biological parents) and a population-based GWAS sample (KORA F4,  $n = 1,743$ ). A meta-analysis was performed on all three samples. In each sample, the distribution of significance levels between the respective gene set and those of all genes was compared using the leading-edge-fraction-comparison test (cut-offs between the 50<sup>th</sup> and 95<sup>th</sup> percentile of the set of all gene-wise corrected p-values) as implemented in the MAGENTA software. In the case-control sample, significant enrichment of associations with obesity was observed above the 50<sup>th</sup> percentile for the set of the 16 nuclear regulators of mitochondrial genes ( $p^{\text{GSEA},50} = 0.0103$ ). This finding was not confirmed in the trios ( $p^{\text{GSEA},50} = 0.5991$ ), but in KORA ( $p^{\text{GSEA},50} = 0.0398$ ). The meta-analysis again indicated a trend for enrichment ( $p^{\text{MAGENTA},50} = 0.1052$ ,  $p^{\text{MAGENTA},75} = 0.0251$ ). The GSEA revealed that weak association signals for obesity might be enriched in the gene set of 16 nuclear regulators of mitochondrial genes.

**Citation:** Knoll N, Jarick I, Volckmar A-L, Klingenspor M, Illig T, et al. (2013) Gene Set of Nuclear-Encoded Mitochondrial Regulators Is Enriched for Common Inherited Variation in Obesity. PLoS ONE 8(2): e55884. doi:10.1371/journal.pone.0055884

**Editor:** Katriina Aalto-Setälä, University of Tampere, Finland

**Received:** September 24, 2012; **Accepted:** January 3, 2013; **Published:** February 8, 2013

**Copyright:** © 2013 Knoll et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by grants of the German Federal Ministry of Education and Research (BMBF and NGFNplus, grant numbers 01KU0903, 01GS0830, and 01GS0820), the European Community's Seventh Framework Program (FP7/2007-2013, grant number 245009) and the IFORES program of the University of Duisburg-Essen. The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München - German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. 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: Anke.Hinney@uni-due.de

## Introduction

Heritability estimates for variance of the body mass index (BMI) range between 40 to 70% [1]. The latest analysis of nearly 250,000 individuals confirmed or newly identified 32 polygenic loci that are associated with BMI [2]. These variants, however, only explain about 1.5% of the BMI variance. Because of the polygenic nature and the small effect sizes of these loci [1,2], an estimation revealed that even an increase to a sample size of 730,000 individuals would not lead to a substantial increase in the explained BMI variance (<5%).

To discover further genetic variation associated with a certain trait, single-locus-oriented genome-wide association studies (GWAS) should be extended to pathway-based approaches or

gene set enrichment analyses (GSEA). As these analyses concentrate on the combined effects across several loci, a gain in statistical power is expected and new genetic insight in the trait of interest might be revealed [3–5]. For example, Lui et al. [6] performed a pathway-based GWA analysis for BMI and body fat mass in US whites and showed the vasoactive intestinal peptide (VIP) pathway to be significantly associated with the investigated traits. This finding was recently confirmed by Evangelou et al. [7] in a subsample of the EPIC-Norfolk study [8].

It is well known that mitochondria are the cellular power plants whose main function is ATP production via oxidative phosphorylation (OXPHOS). Impairment of mitochondrial function might influence body weight. Indeed, there are hints that mitochondria (size, number) and mitochondrial function are altered in obesity.

Adult obese individuals showed smaller mitochondria in skeletal muscle tissue [9] and a reduced complex I activity in both skeletal muscle and cardiomyocytes [9,10] compared to lean subjects. Obesity was also associated with a reduced mRNA expression of the mitochondrial transcription factor A (Tfam) and the nuclear respiratory factor-1 (NRF1), which are two key regulators for mitochondrial biogenesis, and with strongly reduced protein and mRNA expressions of NADH dehydrogenase 6 (ND6), a subunit of complex I in the respiratory chain [10].

The mitochondrial DNA codes only for 37 genes, of which 22 encode tRNAs, two rRNAs and 13 OXPHOS subunits of the complexes I, III, IV and V. Additionally, more than 1,000 nuclear-encoded mitochondrial genes are necessary to ensure complete mitochondrial function and biogenesis [11–14].

Segrè et al. [15] performed a GSEA on nuclear-encoded mitochondrial genes for association with type 2 diabetes mellitus (T2DM) or related glycemic traits, and grouped these genes into three sets: (1) 16 autosomal nuclear regulators of mitochondrial genes based on the literature [12,16–20]; (2) 91 autosomal OXPHOS genes [15,21]; and (3) 966 autosomal nuclear-encoded human mitochondrial genes taken from the MitoCarta compendium [13]. Segrè et al. [15] did not find enrichment in any of the three gene sets in association with T2DM or related glycemic traits.

Based on the central role of mitochondria in metabolism and findings of an altered mitochondrial function we performed a GSEA focusing on obesity to address the question if gene sets of nuclear encoded mitochondrial genes are enriched for modest association signals that collectively influence obesity risk.

## Materials and Methods

### Study Subjects

**Ethics Statement.** Written informed consent was given by all participants and in case of minors by their parents. The study was approved by the Ethics Committees of the Universities of Marburg and Essen and the Bavarian Medical Association. It was conducted in accordance with the Declaration of Helsinki.

**Discovery.** The initial GSEA was performed in a case-control GWAS sample comprising 453 (extremely) obese children and adolescents and 435 normal weight or lean adult controls [22]. Using lean adults who were never overweight or obese during childhood (as assessed by interview) as control group reduces the chances of misclassification compared to the use of lean children as controls who might become overweight or obese in adulthood [23]. The measured body mass index (BMI; in  $\text{kg}/\text{m}^2$ ) was assessed for extremeness using age- and sex-specific percentile criteria for the German population from the National Nutrition Survey I [24]. According to this reference population, all cases were at least overweight (BMI  $\geq 90^{\text{th}}$  percentile), and 84.4% were extremely obese (BMI  $\geq 99^{\text{th}}$  percentile), the lean controls had a BMI of  $18.31 \pm 1.11 \text{ kg}/\text{m}^2$  (Table S1).

**Confirmation.** For independent confirmation a family-based GWAS sample which consisted of 705 obesity trios, each comprising one (extremely) obese child or adolescent (index case) and both biological parents, was used [22]. All index cases were at least overweight (BMI  $\geq 90^{\text{th}}$  percentile), and 83.8% were extremely obese (BMI  $\geq 99^{\text{th}}$  percentile) with regard to reference data from the German National Nutrition Survey I ([24]; Table S1).

In addition, we analyzed a population-based sample comprising 1,743 adult participants, which is a sub-sample of the total KORA F4 sample (Cooperative Health Research in the Region of Augsburg, Table S1, [25]). This sample was analyzed as case-

control sample (KORA-CC): all individuals with a BMI  $\geq 30 \text{ kg}/\text{m}^2$  were categorized as obese cases ( $n = 463$ ) and those with a BMI  $< 25 \text{ kg}/\text{m}^2$  were coded as normal weight controls ( $n = 483$ ). This population-based sample was converted into a case-control sample, as it was recently shown that genetic markers with an effect in the extremes of a trait are detected more solidly within a case-control design compared to a linear regression design, even for smaller sample sizes [26]. This is due to the fact that linear regression results are mostly influenced by the majority of individuals with a moderate trait and only little by the few individuals of the extremes [26].

### Meta-analysis

We performed a meta-analysis of all three samples using the METAL software package ([www.sph.umich.edu/csg/abecasis/metal](http://www.sph.umich.edu/csg/abecasis/metal)). In more detail, meta-analysis was carried out using the inverse variance method by assuming a fixed effect model. We adopted the method of Kazeem and Farrell [27] to meta-analyze single marker information from the two case-control analyses and from the family-based TDTs.

### Genotyping

All three samples were genotyped by the Affymetrix Genome-Wide Human SNP Array 6.0. For quality control, only individuals with a genome-wide SNP call rate (CR)  $\geq 95\%$  were included. Furthermore, in each of the 705 trios from the family-based GWAS sample the percentage of markers with observed Mendelian inconsistent genotype calls (genome-wide 906,703 SNPs) was below 5%. Focusing on the approximate 870,000 autosomal SNPs, the following genotyping quality control filters were applied to each sample separately (Table S2): (1) sample CR  $\geq 95\%$ ; (2) MAF  $\geq 1\%$  in the whole sample (case-control and population-based sample) and MAF  $\geq 5\%$  in the set of all parents of the family-based sample, respectively; (3) two-sided exact p-value  $\geq 0.001$  of the test for Hardy-Weinberg-Equilibrium (HWE) [28] in the whole KORA sample, in the parents of the family-based sample and in the controls of the case-controls GWAS sample, respectively. Additionally, after setting all Mendelian inconsistent calls to “missing” in the family-based sample, we claimed (4) at least one major allele and one minor allele transmission at each SNP. 703,015/641,991/659,502 autosomal SNPs passed this QC in the case-control/family-based/population-based sample, respectively and were used for the analysis (Table S2).

### Statistical Analyses

**Gene set enrichment analysis (GSEA) for sets of mitochondrial genes.** Each gene set enrichment analysis (GSEA) is based on the idea of comparing gene association signals of gene sets with biological plausibility for the given phenotype to those of the genome-wide set of genes [3,4,15]. First, the analyzed SNPs are linked to their corresponding gene. Secondly, a test statistic (e.g. p-value) aggregating the SNP information is derived for each gene. Finally, the distribution of test statistics is compared between gene sets (details see below).

We tested those three mitochondrial gene sets for enrichment of obesity association signals, which were considered in the study of Segrè et al. [15] who tested for an enrichment of association signals with T2DM and related glycemic traits: (1) a set of 16 autosomal nuclear regulators of mitochondrial genes based on the literature [12,16–20], (2) a set of 91 autosomal oxidative phosphorylation (OXPHOS) genes [15,21], and (3) a list of 966 autosomal nuclear-encoded human mitochondrial genes taken from the MitoCarta compendium which are over 80% of all assumed mitochondrial genes [13].

**GSEA – Discovery.** In the case-control GWAS sample, the Cochran-Armitage trend test for an additive mode of inheritance was applied to each autosomal SNP. Afterwards, SNPs were mapped onto genes. For this purpose, a list of human gene transcripts (n = 26,914 for the hg18 March 2006 version) was downloaded from the UCSC Genome Browser (<http://genome.ucsc.edu/>). After exclusion of genes with two or more transcripts on separate chromosomes or with more than 1 Mb distance on the same chromosome, a total of 17,680 autosomal genes were followed-up. In consistency with Segrè et al. [15], SNPs that are located within 110 kb upstream and 40 kb downstream to the most extreme transcript start and end site of a gene were assigned to this gene. These boundaries were chosen as they represent the 99<sup>th</sup> percentiles of the distances of cis-eQTLs from transcript start and end sites of adjacent genes [29]. Genes without SNPs in their extended gene boundaries (n = 55) were discarded from GSEA testing.

Subsequently, each gene was assigned a gene-wise empirically corrected p-value  $P_g$ . To determine  $P_g$ , the lowest observed p-value  $P_{g,\min}$  of each gene was first determined and stored. Secondly, 10,000 permutations of the genotype data were performed using PLINK. In each permutation, affection status was flipped for all SNPs to generate the null distribution. Finally,  $P_g$  was calculated as the fraction of permutations whose minimal p-value per gene was equal to or smaller than  $P_{g,\min}$ . To achieve maximal accuracy, for those genes with  $P_g \leq 0.01$  (0.001), the procedure was repeated with 100,000 (1,000,000) permutations.

Prior to GSEA testing, we addressed physical clustering of genes by excluding all genes with the identical selected SNP as compared to the gene with the lowest  $P_g$  for the SNP already in the gene set. This exclusion was done in order to avoid significant gene set enrichment based on identical association signals [15].

Our alternative hypothesis was that gene p-value ranks in one of the three gene subsets of interest were skewed towards high ranks compared to the full autosomal set of genes. To test this hypothesis, we applied: (1) the leading-edge-fraction-comparison test as proposed by Segrè et al. [15] with  $P^{\text{cut-off}} = 95^{\text{th}}$ ,  $75^{\text{th}}$  and  $50^{\text{th}}$  percentile of the set of autosomal gene-wise p-values and with 10,000 samplings from the full autosomal gene p-value distribution (with corresponding GSEA p-values:  $P^{\text{GSEA},95}$ ,  $P^{\text{GSEA},75}$  and  $P^{\text{GSEA},50}$ ). This test is based on the idea of comparing the fraction of genes with gene p-values below a certain cut-off (i.e. above a certain percentile; leading edge fraction) in the full set of genes and in the gene subset of interest. Here, the null distribution of such fractions in the gene subset is derived by randomly sampling the same number of gene p-values from the full set of gene p-values. The GSEA p-value is then determined by dividing the number of samplings with equal or larger leading edge fraction as the observed one by the number of samplings generated. To test the robustness of this test, we additionally ran three alternative one-sided GSEA tests as proposed by Segrè et al. [15]: (2) the Wilcoxon-Mann-Whitney test ( $P^{\text{GSEA},\text{WMW}}$ ); (3) the Kolmogorov-Smirnov test ( $P^{\text{GSEA},\text{KS}}$ ); and (4) the t-test ( $P^{\text{GSEA},t}$ ).

**GSEA – Confirmation.** In the family-based GWAS sample, for each SNP a transmission disequilibrium test (TDT; [30]) was calculated using the PLINK v1.07 software [31] (<http://pngu.mgh.harvard.edu/purcell/plink/>) and assuming an additive allelic model of inheritance. Families with missing genotypes were excluded from TDT analysis of the respective SNP. In total, for only 0.37% of all SNPs more than 5% of trios were excluded from the TDT (due to genotyping failures and/or Mendelian inconsistencies). Gene-wise empirically corrected p-values for the family-based GWAS sample were based on randomly flipping the parentally transmitted allele for each family and each permutation.

In KORA-CC each SNP was tested by the Cochran-Armitage trend test for an association with obesity, and gene-wise p-values were determined as for the discovery CC sample. For both the family-based GWAS sample and KORA-CC the remaining GSEA procedure was performed as described for the discovery sample.

**Meta-Analysis Gene set Enrichment of variant Associations (MAGENTA).** The Meta-Analysis Gene set Enrichment of variant Associations (MAGENTA) software provided by Segrè et al. [15] was specifically designed for the application to large-genome-wide association study meta-analyses in which individual genotypes are not available. In this context, it is not possible to evaluate statistical gene-wise significance via standard phenotype permutation procedures as described above. Instead, a linear regression-based approach accounting for physical gene size, the number of SNPs and their genetic properties (LD between SNPs, number of recombination hotspots and genetic distance of the gene), was proposed to determine gene-wise corrected p-values.

After bringing together the single marker information from the two case-control samples and the family-based TDTs (in application of the METAL software; for details see ‘Meta-analysis’), we applied MAGENTA to the single marker p-values of the meta-analysis. Leading edge fraction tests for the 95<sup>th</sup>, the 75<sup>th</sup> and the 50<sup>th</sup> percentile cut-off were performed as well as the alternatively included exact Wilcoxon-Mann-Whitney test. These cut-offs were chosen because simulations showed that for modest effects the 95<sup>th</sup> percentile and for weak effects the 75<sup>th</sup> percentile yielded the optimal power to detect gene set enrichment [15] and with additional regard to our discovery findings.

For reasons of comparability, in addition to our permutation-based GSEA testing procedure, we applied MAGENTA to all three samples individually. Regression-corrected gene p-values ( $p^{\text{MAGENTA}}$ ) and permutation-based gene p-values ( $p^{\text{GSEA}}$ ) were shown to be highly correlated ( $r = 0.95$ ; see Table 1 & 2).

## Results

### Discovery

In our case-control sample, the effective gene set size for the GSEA analyses of all human autosomal genes was 10,180, since 55 genes did not have any genotyped SNPs within their extended gene boundaries (110 kb upstream and 40 kb downstream to the most extreme transcript boundaries) and 7,445 genes were removed due to physical clustering (see Methods). In total, all human autosomal genes were covered by 521,469 unique SNPs (73.03% of all autosomal SNPs which can be found on the Affymetrix SNP array 6.0). Among the lists of the 16 nuclear regulators of mitochondrial genes and the 91 OXPHOS genes, all genes had SNPs in their extended boundaries. Two genes of the 91 OXPHOS genes were removed due to physical clustering. There were 1,014 unique SNPs (0.14% of all SNPs) that fell within the gene regions of the nuclear regulators of mitochondrial genes and 2,781 unique SNPs (0.39% of all SNPs) that were located within the gene regions of the OXPHOS genes. Furthermore, those 965 autosomal nuclear-encoded human mitochondrial genes that contained SNPs in their gene regions were covered by 35,223 unique SNPs (4.93% of all SNPs), whereas due to physical clustering the corresponding effective gene set size was 880.

The first gene set of 16 nuclear regulators of mitochondrial genes was enriched for obesity association signals ( $P^{\text{GSEA},\text{WMW}} = 0.0075$ ,  $P^{\text{GSEA},\text{KS}} = 0.0195$ ,  $P^{\text{GSEA},t} = 0.0053$ ; Fig. 1 & Table 1). This enrichment was found above the 50<sup>th</sup> percentile ( $P^{\text{GSEA},50} = 0.0103$ ). The enrichment remained significant after

**Table 1.** Discovery: GSEA and MAGENTA for obesity in the case-control GWAS sample of 453 (extremely) obese cases and 435 lean controls.

Gene set	total number of genes	Effective number of genes involved	number of SNPs (703,015)	% of all autosomal SNPs involved	$p^{GSEA,WMW}$ , Wilcoxon-Mann-Whitney test	$p^{GSEA,KS}$ , Kolmogorov-Smirnov Test	$p^{GSEA,t}$ , t-Test	$p^{GSEA,95}$ , 95 <sup>th</sup> percentile cut-off test <sup>a</sup>	$p^{GSEA,75}$ , 75 <sup>th</sup> percentile cut-off test <sup>b</sup>	$p^{GSEA,50}$ , 50 <sup>th</sup> percentile cut-off test <sup>c</sup>	$p^{MAGENTA,95}$ , 95 <sup>th</sup> percentile cut-off test <sup>a</sup>	$p^{MAGENTA,75}$ , 75 <sup>th</sup> percentile cut-off test <sup>b</sup>	$p^{MAGENTA,50}$ , 50 <sup>th</sup> percentile cut-off test <sup>c</sup>
1) Nuclear regulators of mitochondrial genes	16	1,014	0.14	<b>0.0075</b>	<b>0.0195</b>	<b>0.0053</b>	0.5644	0.0796	<b>0.0103</b>	<b>0.0043</b>	0.575	<b>0.0074</b>	<b>0.0099</b>
2) Oxidative phosphorylation genes	91	2,781	0.39	0.6225	0.8586	0.6374	0.2873	0.5643	0.8447	0.6565	0.7495	0.7369	
3) Nuclear-encoded mitochondrial genes	966	35,223	4.93	0.3841	0.2502	0.4104	0.6437	0.1905	0.8969	0.5287	0.7372	0.7577	
all autosomal genes	17,680	10,180	521,469	73.03	reference	reference	reference	reference	reference	reference	reference	reference	reference

<sup>a</sup>cut-off = 0.0216,<sup>b</sup>cut-off = 0.1631,<sup>c</sup>cut-off = 0.3951,<sup>d</sup>exact GSEA Wilcoxon-Mann-Whitney test; GSEA and MAGENTA p-values below 0.05 are highlighted in bold.  
doi:10.1371/journal.pone.0055884.t001

**Table 2.** Confirmation & Meta-analysis: GSEA and MAGENTA for the gene set of 16 nuclear regulators of mitochondrial genes in 705 trios, KORA-CC and for meta-analysis.

Sample	total number of genes involved	effective number of autosomal SNPs	% of all SNPs involved	$p^{GSEA,WMW}$ , Wilcoxon-Mann-Whitney test	$p^{GSEA,KS}$ , Kolmogorov-Smirnov-Test	$p^{GSEA,t}$ , t-Test	$p^{GSEA,95}$ , 95 <sup>th</sup> percentile cut-off test <sup>a</sup>	$p^{GSEA,75}$ , 75 <sup>th</sup> percentile cut-off test <sup>b</sup>	$p^{GSEA,50}$ , 50 <sup>th</sup> percentile cut-off test <sup>c</sup>	$p^{MAGENTA,WMW}$ , Wilcoxon-Mann-Whitney test <sup>d</sup>	$p^{MAGENTA,95}$ , 95 <sup>th</sup> percentile cut-off test <sup>a</sup>	$p^{MAGENTA,75}$ , 75 <sup>th</sup> percentile cut-off test <sup>b</sup>	$p^{MAGENTA,50}$ , 50 <sup>th</sup> percentile cut-off test <sup>c</sup>
705 obesity trios	16	919	0.14	0.7879	0.7930	0.7588	1.0000	0.3711	0.5991	0.6817	1	0.6024	0.7683
463 cases and 483 controls (KORA-CC) <sup>e</sup>	16	933	0.14	<b>0.0260</b>	<b>0.0431</b>	<b>0.0211</b>	<b>0.0432</b>	0.1939	<b>0.0398</b>	<b>0.0083</b>	0.1918	0.1888	<b>0.0405</b>
meta-analysis16	16	1,036	0.14	-	-	-	-	-	-	<b>0.0357</b>	0.5587	<b>0.0251</b>	0.1052

<sup>a</sup>Trios: cut-off = 0.0382, KORA-CC: cut-off = 0.0486, Meta-analysis: cut-off = 0.0443.

<sup>b</sup>Trios: cut-off = 0.2216, KORA-CC: cut-off = 0.2611, Meta-analysis: cut-off = 0.2969.

<sup>c</sup>Trios: cut-off = 0.4687, KORA-CC: cut-off = 0.5085, Meta-analysis: cut-off = 0.5619.

<sup>d</sup>Exact GSEA Wilcoxon-Mann-Whitney test.

<sup>e</sup>BMI  $\geq 30$  (cases) vs. BMI  $< 25$  (controls); GSEA p-values below 5% are highlighted in bold. doi:10.1371/journal.pone.0055884.t002

Bonferroni correction for the three gene sets tested except for the Kolmogorov-Smirnov-Test.

## Confirmation

In the family-based sample (705 trios), we did not observe enrichment of association signals to (early onset extreme) obesity in the first gene set ( $p^{GSEA,50} = 0.5991$ ,  $p^{GSEA,WMW} = 0.7879$ ,  $p^{GSEA,KS} = 0.7930$ ,  $p^{GSEA,t} = 0.7588$ ; Fig. 2A & Table 2).

For the second independent case-control sample (KORA-CC), we found significant enrichment for association signals for obesity for the first gene set ( $p^{GSEA,WMW} = 0.0260$ ,  $p^{GSEA,KS} = 0.0431$ ,  $p^{GSEA,t} = 0.0211$ ). The enrichment was found above the 50<sup>th</sup> percentile ( $p^{GSEA,50} = 0.0398$ ), confirming the initial finding of the discovery case-control sample (Fig. 2B & Table 2).

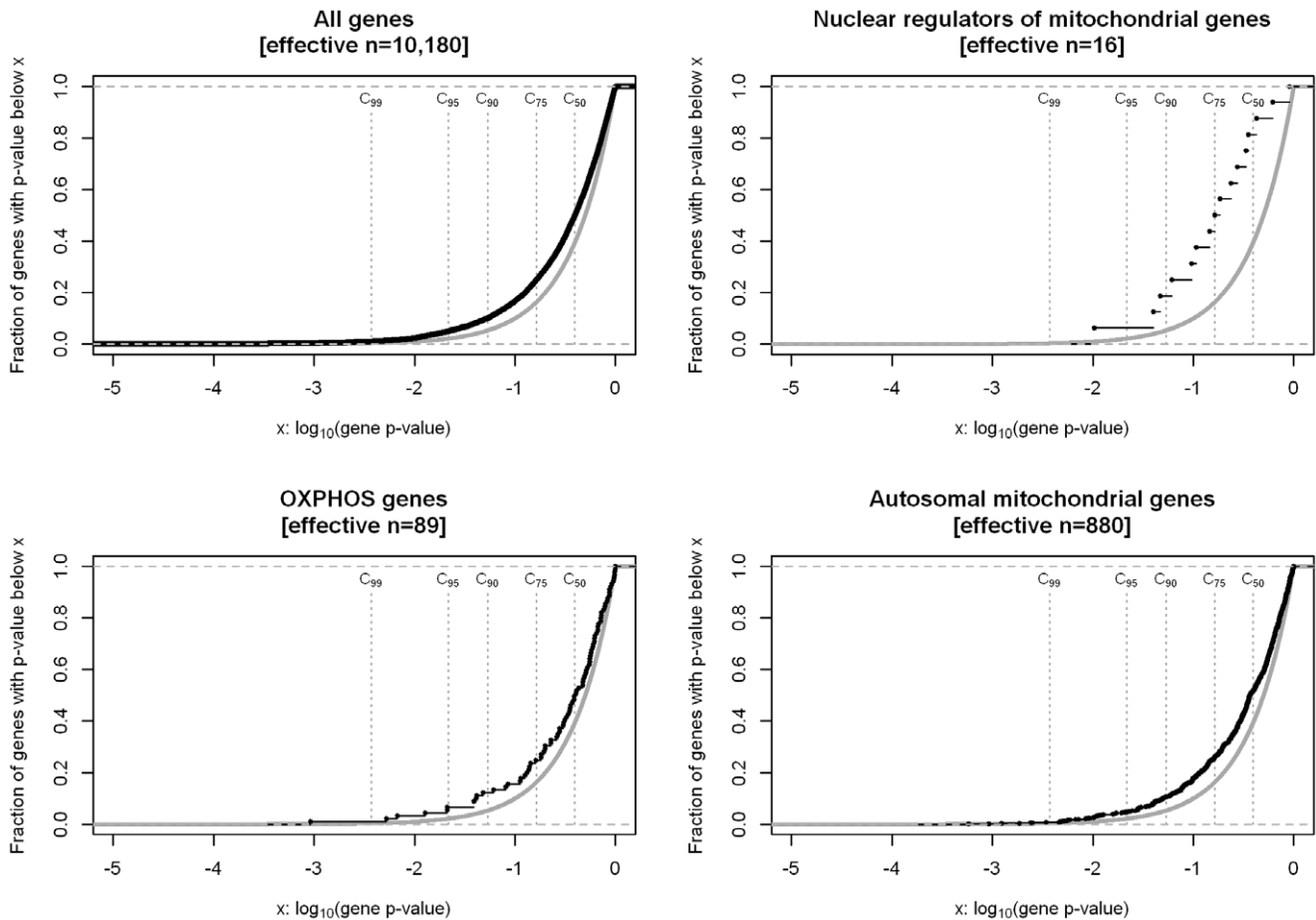
## Meta-analysis

Due to the different design of the study samples (case-control and family-based), the permutation-based GSEA testing procedure and explicitly the determination of gene-wise corrected p-values could not directly be applied in a meta-analysis of all three samples. Alternatively, we applied the two software toolkits METAL and MAGENTA which were both specifically designed for the application to meta-analyze data. Significant enrichment for the first gene set was found ( $p^{MAGENTA,WMW} = 0.0357$ ). This enrichment was not found above the 50<sup>th</sup> percentile as in the discovery, but rather above the 75<sup>th</sup> percentile ( $p^{MAGENTA,50} = 0.1052$ ,  $p^{MAGENTA,75} = 0.0251$ ; Fig. 2C & Table 2). Thus, the effect of enrichment remained stable in the meta-analysis.

For most genes, the SNP with minimal single-marker p-value differed in the three analyzed samples (Table 3). For example, rs12033461 was the best SNP for *ESRRG* in the family-based sample, while it was rs11577585 in KORA-CC. The LD between these two SNPs was  $r^2 = 0$  in the parents of the family-based sample. Generally, these LDs were quite low (Table 3), so association signals seem to be independent, and our gene-based approach which considers large parts of common variation within a gene might be a useful approach to handle multiple ancestral mutations [32,33].

## Discussion

The gene variants discovered by single-locus-oriented GWAS have explained only about 1.5% of the total BMI variance so far [2]. As GSEA approaches concentrate on the combined effects of several loci to potentially reveal new insight into the genetic impact, we performed a GSEA to analyze if autosomal nuclear-encoded mitochondrial genes are enriched for association signals for obesity. The three mitochondrial gene sets as well as the GSEA and MAGENTA procedure were adopted from Segrè et al. [15]. While Segrè et al. [15] did not find enrichment in association with T2DM and related glycemic traits, we observed enrichment for obesity association signals in the gene set of the 16 regulators of nuclear-encoded mitochondrial genes (gene set 1) in two independent case-control GWAS data sets (total  $n = 1,834$ ). However, the enrichment was not detectable in a family-based GWAS sample of 705 obesity trios. The enrichment for obesity association signals was found for the 50<sup>th</sup> percentile, i.e. gene adjusted p-values between  $\sim 0.2$  and  $0.5$ . By applying MAGENTA to the discovery sample, we also found enrichment above the 75<sup>th</sup> percentile (Table 1). This observation was similarly present only in the meta-analysis, but not in any of the confirmation samples individually. Our results support the hypothesis that a GSEA may detect combined association effects of several loci [3–7]. None of the



**Figure 1. Empirical cumulative distribution functions (ECDF) of  $P_g$  in four different gene sets in the Discovery.** A case-control GWAS sample of 453 (extremely) obese cases and 435 lean controls was analyzed. In each panel the grey line represents the ECDF of the uniform distribution (null hypotheses of no association) and the black line represents the ECDF of the respective gene set.  $P_{gr}$ , gene-wise corrected p-value. doi:10.1371/journal.pone.0055884.g001

above described 16 genes revealed significant association to obesity in a single locus-oriented approach, as none of these genes has been found in the list of 32 BMI loci reported in the latest and largest meta-analysis so far [2].

It was not possible to identify one specific weight associated candidate gene, as in both case-control approaches the SNPs/genes with the lowest p-values differed (Table 3). Our finding underscores that the combined effect of several loci leads to an association with the investigated trait, rather than a single gene of a set.

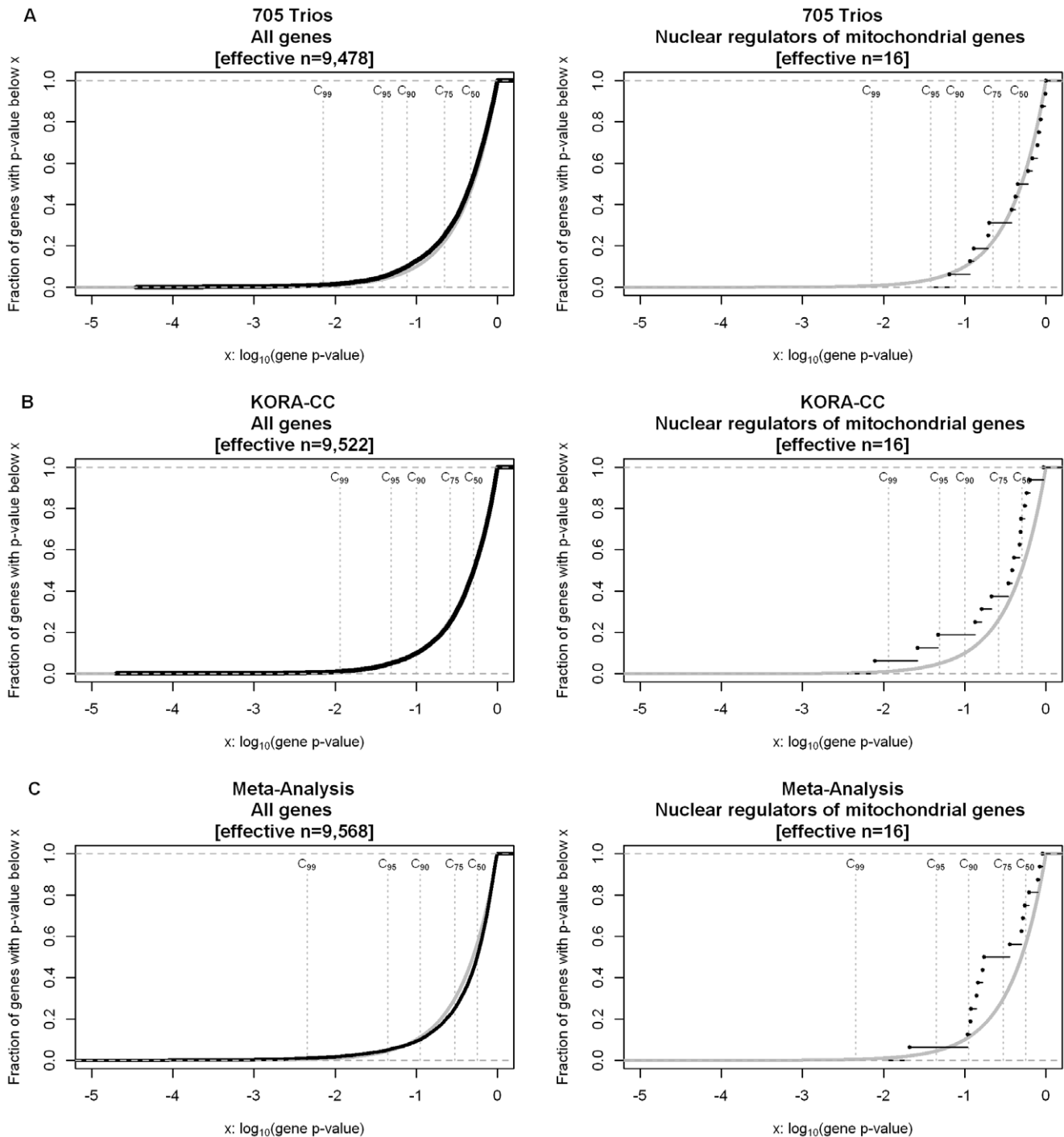
One limitation of our analysis is that it is based exclusively on autosomal mitochondrial genes. According to the MitoCarta compendium there are 1012 unique mitochondrial genes [13] of which 13 are protein coding genes of mtDNA (1.3% of all mitochondrial genes) and 31 are X-chromosomal (one is X/Y-chromosomal; 3.1% of all mitochondrial genes). As most GWAS primarily focus on autosomal SNPs, genes of mtDNA and sex chromosomes were not included in the analysis of Segrè et al. [15]. For reasons of comparability we also only focused on autosomal mitochondrial genes. However, due to the fact that both mtDNA and sex-chromosomal genes represent less than 5% of all mitochondrial genes, the impact of these genes on the enrichment analysis might be small. Anyhow, variation in the mtDNA has probably more relevant effects on the mitochondrial function than variation in autosomal genes, as for instance mtDNA does not

comprise UTRs or introns. However, mtDNA SNPs are not the focus of our GSEA.

Another limitation of our GSEA is that it is based on GWAS data, i.e. common variants. Rare variants, which could have a stronger impact on the investigated trait, are thus hardly addressed in our analyses.

We evaluated the robustness of our results pertaining to the 16 nuclear regulators of mitochondrial function. Besides the leading-edge fraction test several other statistical tests recommended were performed (Wilcoxon-Mann-Whitney-test, Kolmogorov Smirnov test, t-test) to demonstrate independence from the method choice. Although the Kolmogorov Smirnov test of the discovery stage revealed only nominal significance, results of all tests were similar regarding significance within a fixed tested sample and gene set. The MAGENTA software was additionally applied to each single sample in order to maximally guarantee robustness (within each sample) and comparability (between different samples and the meta-analysis) of the results. Apart from a few exceptions, we found high levels of agreement of the p-values determined by both methods (leading-edge fraction test vs. MAGENTA) within a sample (Table 1 & 2).

In addition to the robustness regarding method choice another strength of our study was that we used both case-control and family-based samples. We observed consistent evidence for enrichment in two case-control samples but failed



**Figure 2. Empirical cumulative distribution functions (ECDF) of  $P_g$  in all autosomal genes and gene set 1.** For independent confirmation of the initial finding, GSEA was performed in 705 obesity trios (A) and in 463 obese cases and 483 normal weight or lean controls of the KORA-CC sample (B). In addition, a meta-analysis of all three study samples (from Discovery and Confirmation) was performed (C). In each panel the grey line represents the ECDF of the uniform distribution (null hypotheses of no association) and the black line represents the ECDF of the respective gene set.  $P_g$ , gene-wise corrected p-value. doi:10.1371/journal.pone.0055884.g002

to detect it in the family-based GWAS sample of 705 obesity trios. Possibly the family-based sample was too small for a confirmation. Moreover, if the effect was mainly driven by lean and normal weight subjects, the frequencies of the variants would be very low in the predominantly obese trio parents [34] again resulting in power issues. A third explanation might be

genetic heterogeneity including both locus and allelic heterogeneity.

The 16 regulators of nuclear-encoded mitochondrial genes (gene set 1) are transcription factors and/or co-activators (Table 3). Although there are hints of disturbed mitochondria or mitochondrial function among obese individuals [9,10], from the findings of

**Table 3.** Best SNPs of nuclear regulators of mitochondrial genes (gene set 1) in each sample and linkage disequilibrium between best SNPs of the three different study samples.

Sample	Discovery				Confirmation				Confirmation				Meta-analysis											
	453 cases and 435 controls				705 obesity trios				463 cases and 483 controls (KORA-CC) <sup>a</sup>				453 cases and 435 controls, 705 obesity trios & 463 cases and 483 controls (KORA-CC) <sup>a</sup>											
gene id	Gene	Best SNP	Best SNP p-value	LD: best SNP in Trios – best SNP in CC [ $r^2$ ] <sup>c</sup>	number of SNPs in gene region	Gene p-value	Best SNP in gene <sup>b</sup>	Best SNP p-value	number of SNPs in gene region	Gene p-value	Best SNP in gene <sup>b</sup>	Best SNP p-value	LD: best SNP in Trios – best SNP in CC [ $r^2$ ] <sup>c</sup>	number of SNPs in gene region	Gene p-value	Best SNP in gene <sup>b</sup>	Best SNP p-value	LD: best SNP in Trios – best SNP in CC [ $r^2$ ] <sup>c</sup>	number of SNPs in gene region	Gene p-value	Best SNP in gene <sup>b</sup>	Best SNP p-value		
ESRRA	0.0616	rs2429455 <sup>§</sup>	0.0065	0.403	12	0.3778	rs1059440 <sup>§</sup>	0.0719	10	0.3853	rs11231740 <sup>§</sup>	0.0681	0.129	11	0.1662	rs4930702 <sup>§</sup>	0.0161	0.004	12	0.1662	rs4930702 <sup>§</sup>	0.0161	0.004	12
ESRRG	0.4240	rs2185226 <sup>#</sup>	0.0035	0.001	335	0.8585	rs12033461 <sup>#</sup>	0.0155	316	0.9301	rs11577585 <sup>#</sup>	0.0185	0	320	0.8458	rs7531250 <sup>#</sup>	0.0090	0.003	349	0.8458	rs7531250 <sup>#</sup>	0.0090	0.003	349
GABPA	0.0400	rs2051180 <sup>§</sup>	0.0022	0.945	36	0.7873	rs11087972 <sup>§</sup>	0.1216	35	0.0261	rs7284014 <sup>§</sup>	0.0012	0.024	32	0.1191	rs2829866 <sup>§</sup>	0.0048	0.206	37	0.1191	rs2829866 <sup>§</sup>	0.0048	0.206	37
GABPB1	0.3542	rs4775886 <sup>§</sup>	0.0205	0	37	0.9953	rs12910368 <sup>#</sup>	0.3374	32	0.1336	rs16963477 <sup>§</sup>	0.0071	0	35	0.1091	rs16963477 <sup>§</sup>	0.0046	0	38	0.1091	rs16963477 <sup>§</sup>	0.0046	0	38
GABPB2	0.9071	rs3754210 <sup>§</sup>	0.2138	0.072	19	<b>0.0644</b>	<b>rs4970989<sup>§</sup></b>	<b>0.0052</b>	<b>19</b>	0.4848	rs267738 <sup>§</sup>	0.0539	NA	19	0.6166	rs7526955 <sup>§</sup>	0.0666	0.243	22	0.6166	rs7526955 <sup>§</sup>	0.0666	0.243	22
MEF2A	0.1643	rs7173943 <sup>§</sup>	0.0057	0.001	59	0.4216	rs4313794 <sup>§</sup>	0.0185	60	0.4036	rs7175248 <sup>§</sup>	0.0156	0.065	59	0.1462	rs7173943 <sup>§</sup>	0.0039	0.001	65	0.1462	rs7173943 <sup>§</sup>	0.0039	0.001	65
MYC	0.2750	rs11990827 <sup>§</sup>	0.0094	0.001	60	0.6785	rs4395860 <sup>§</sup>	0.0422	55	0.4720	rs13252644 <sup>§</sup>	0.0194	0	59	0.9125	rs12155669 <sup>§</sup>	0.0923	0.384	62	0.9125	rs12155669 <sup>§</sup>	0.0923	0.384	62
NRF1	0.2379	rs2693737 <sup>§</sup>	0.0104	0.013	51	0.9805	rs9792084 <sup>§</sup>	0.1758	52	0.6346	rs11771549 <sup>§</sup>	0.0418	0.012	47	0.8021	rs11771549 <sup>§</sup>	0.0756	0.012	56	0.8021	rs11771549 <sup>§</sup>	0.0756	0.012	56
NR1P1	0.1075	rs2776043 <sup>#</sup>	0.0052	0.005	47	0.6005	rs17274722 <sup>§</sup>	0.0506	46	<b>0.0078</b>	<b>rs10482862<sup>§</sup></b>	<b>0.0003</b>	<b>0.004</b>	<b>43</b>	0.1730	rs10482862 <sup>§</sup>	0.0047	0.004	49	0.1730	rs10482862 <sup>§</sup>	0.0047	0.004	49
PPARA	0.6169	rs3744749 <sup>§</sup>	0.0358	0.009	55	0.1999	rs12170325 <sup>§</sup>	0.0084	46	0.2128	rs4253754 <sup>#</sup>	0.0084	0.008	48	0.4980	rs4253655 <sup>#</sup>	0.0219	0.021	55	0.4980	rs4253655 <sup>#</sup>	0.0219	0.021	55
PPARD	0.1854	rs9658085 <sup>#</sup>	0.0097	0.016	40	0.8156	rs2894401 <sup>§</sup>	0.1644	22	0.3453	rs2267666 <sup>#</sup>	0.0201	0.134	35	0.3595	rs9658085 <sup>#</sup>	0.0181	0.016	40	0.3595	rs9658085 <sup>#</sup>	0.0181	0.016	40
PPARGCTA	0.1456	rs17574213 <sup>*</sup>	0.0034	0.001	75	0.8988	rs10517032 <sup>§</sup>	0.0641	71	0.4964	rs17576576 <sup>§</sup>	0.0163	0.009	67	0.5493	rs7682906 <sup>§</sup>	0.0182	0.089	78	0.5493	rs7682906 <sup>§</sup>	0.0182	0.089	78
PPARGCTB	0.3352	rs10069462 <sup>§</sup>	0.0081	0.017	110	0.4494	rs7713955 <sup>§</sup>	0.0141	102	0.5462	rs10065816 <sup>§</sup>	0.0166	0.145	104	0.5180	rs11746690 <sup>#</sup>	0.0096	0.016	114	0.5180	rs11746690 <sup>#</sup>	0.0096	0.016	114
SIRT1	<b>0.0104</b>	<b>rs7895833<sup>§</sup></b>	<b>0.0011</b>	<b>0.304</b>	<b>16</b>	0.1282	rs10509291 <sup>§</sup>	0.0190	16	0.0471	rs16924888 <sup>§</sup>	0.0053	0.01	15	<b>0.0209</b>	<b>rs17712705<sup>§</sup></b>	<b>0.0019</b>	<b>0.146</b>	<b>17</b>	<b>0.0209</b>	<b>rs17712705<sup>§</sup></b>	<b>0.0019</b>	<b>0.146</b>	<b>17</b>
SP1	0.0964	rs4759082 <sup>§</sup>	0.0126	same SNP	16	0.1949	rs4759082 <sup>§</sup>	0.0436	11	0.1624	rs2016266 <sup>§</sup>	0.0207	0.209	14	0.1173	rs4759082 <sup>§</sup>	0.0101	same SNP	16	0.1173	rs4759082 <sup>§</sup>	0.0101	same SNP	16
YY1	0.0470	rs8007801 <sup>§</sup>	0.0034	0.011	25	0.1156	rs9291 <sup>§</sup>	0.0079	26	0.5759	rs2766692 <sup>§</sup>	0.0732	0.015	25	0.1391	rs9291 <sup>§</sup>	0.0085	same SNP	26	0.1391	rs9291 <sup>§</sup>	0.0085	same SNP	26

- Table 1 will be continued -.

- Table 1 continued -.

<sup>a</sup>BMI  $\geq 30$  (cases) vs. BMI  $< 25$  (controls);<sup>b</sup>location of SNP: \*, exon;<sup>#</sup>intronic;<sup>§</sup>upstream of gene and.<sup>¶</sup>downstream of gene;<sup>‡</sup>Linkage Disequilibrium (LD)

was calculated in the parents of the family-based GWAS sample by use of HaploView 4.2;

<sup>§</sup>SNP-wise p-values of the meta-analysis were derived by application of the METAL software (for details see 'Materials and Methods, Meta-analysis');<sup>¶</sup>LD between rs7895833 and rs16924888;  $r^2 = 0.581$ ; best gene of each sample and the meta-analysis is indicated in bold letter.

ESRRA, Estrogen related receptor alpha; ESRRG, Estrogen related receptor gamma; GABPA, GA-binding protein alpha subunit; GABPB1, GA-binding protein beta subunit 1; GABPB2, GA-binding protein beta subunit 2; MEF2A,

Myocyte-specific enhancer factor 2A; MYC, Myelocytomatosis viral oncogene homolog (avian); NRF1, Nuclear respiratory factor 1; NR1P1, Nuclear receptor-interacting protein 1; PPARGA, Peroxisome proliferator-activated receptor

alpha; PPARGB, Peroxisome proliferator-activated receptor beta; PPARGCTA, Peroxisome proliferator-activated receptor gamma coactivator 1 alpha; PPARGCTB, Peroxisome proliferator-activated receptor gamma coactivator 1 beta;

SIRT1, Sirtuin 1; SP1, Specificity protein 1; YY1, Transcriptional repressor protein YY1.

doi:10.1371/journal.pone.0055884.t003



**Table 4.** Animal models (knockout, alterations in the expression and mutations) of the nuclear regulators of mitochondrial genes (gene set 1) in relation to obesity or related traits.

Gene	Phenotype	Reference
<i>ESRRA</i>	<i>ERRα</i> <sup>-/-</sup> mice with reduced body weight and fat mass, and resistance to a high-fat diet-induced obesity	[38]
<i>ESRRG</i>	No body weight/body fat associated phenotype	
<i>GABPA</i>	= <i>NRF2</i> ; targeted knock-out (k.o.) of <i>Nrf2</i> in mice leads to 20% lower body weight after <i>ad libitum</i> diet compared to wild type littermates, lower adipose tissue mass, smaller adipocytes and protects against weight gain and obesity otherwise induced by a high fat diet	[39]
<i>GABPB1</i>	No body weight/body fat associated phenotype	
<i>GABPB2</i>	No body weight/body fat associated phenotype	
<i>MEF2A</i>	No body weight/body fat associated phenotype	
<i>MYC</i>	Transgenic mice overexpressing <i>c-myc</i> in the liver show lower body weight increase and lower fat accumulation in adipose tissue compared to control mice on a high fat diet of 3 months	[40]
<i>NRF1</i>	No body weight/body fat associated phenotype	
<i>NRIP1</i>	Formerly known as <i>RIP140</i> ; k.o. mice viable and morphologically normal, but 15–20% less heavier than wild-type or heterozygous littermates	[35]
<i>PPARA</i>	<i>PPARα</i> -null mice on two different backgrounds ( <i>Sv/129</i> or <i>C57BL/6N</i> ) were not obese, but had hepatic accumulation of fat and larger gonadal adipose stores compared to wild type controls	[41]
<i>PPARD</i>	<i>PPARδ</i> -null mice are smaller than controls and have smaller gonadal fat stores	[42]
<i>PPARGC1A</i>	Female <i>PGC-1α</i> <sup>-/-</sup> mice show increased body fat and hepatic steatosis after short term starvation	[43]
<i>PPARGC1B</i>	<i>PGC-1b</i> k.o. mice with reduced body weight and fat mass	[44]
<i>SIRT1</i>	<i>Sirt1</i> transgenic (knockin) mice are lighter and have less white adipose tissue per body weight than wild type littermates	[36]
<i>SP1</i>	No body weight/body fat associated phenotype	
<i>YY1</i>	No body weight/body fat associated phenotype	

*ESRRA*, Estrogen related receptor alpha; *ESRRG*, Estrogen related receptor gamma; *GABPA*, GA-binding protein alpha subunit; *GABPB1*, GA-binding protein beta subunit 1; *GABPB2*, GA-binding protein beta subunit 2; *MEF2A*, Myocyte-specific enhancer factor 2A; *MYC*, Myelocytomatosis viral oncogene homolog (avian); *NRF1*, Nuclear respiratory factor 1; *NRIP1*, Nuclear receptor-interacting protein 1; *PPARA*, Peroxisome proliferator-activated receptor alpha; *PPARD*, Peroxisome proliferator-activated receptor delta; *PPARGC1A*, Peroxisome proliferator-activated receptor gamma coactivator 1 alpha; *PPARGC1B*, Peroxisome proliferator-activated receptor gamma coactivator 1 beta; *SIRT1*, Sirtuin 1; *SP1*, Specificity protein 1; *YY1*, Transcriptional repressor protein YY1.  
doi:10.1371/journal.pone.0055884.t004

this GSEA, we cannot conclude if or to what extent these nuclear-encoded regulators of gene set 1 are involved in mitochondrial disturbance. There are mouse models for nine of the 16 genes which showed that knockout (k.o.) or alterations in the expression of these genes are related to leanness or related traits (Table 4). For example, *Nrip1* k.o. mice are viable and morphologically normal, but 15–20% lighter than the wild-type or heterozygous littermates [35]. *Nrf2* (= *Gabpa*) k.o. mice are characterized by decreased adipose tissue mass and protected against a high fat diet induced obesity. In addition, *Sirt1* transgenic (knockin) mice were lighter than wild type littermates and had less white adipose tissue per body weight [36]. The enrichment of association signals in gene set 1 supports the findings from the above mentioned animal studies that these genes could be potential candidate genes for obesity/leanness and related traits. Similarly, in a recent GSEA, Vimalaewaren et al. [37] found enrichment of association signals for a gene set of 547 obesity-susceptibility candidate genes in a large meta-analysis of 123,564 individuals [2].

In summary, a GSEA on autosomal nuclear-encoded genes relevant for mitochondrial function revealed that a gene set of 16 nuclear encoded regulators of mitochondrial genes was enriched for weak obesity association signals. Initially, this enrichment was

found in a case-control approach and independently confirmed in another case-control sample.

## Supporting Information

**Table S1 Basic phenotypical characteristics of the family-based, the case-control and the population-based GWAS sample.**

(DOC)

**Table S2 Quality control of SNPs.**

(DOC)

## Acknowledgments

We thank all participants for their voluntary contribution.

## Author Contributions

Conceived and designed the experiments: AH JH IJ AS NK MK. Performed the experiments: NK IJ AS. Analyzed the data: IJ AS. Contributed reagents/materials/analysis tools: ALV TI HG CG HEW AP AH JH. Wrote the paper: NK IJ AS AH.

## References

- Hebebrand J, Volckmar AL, Knoll N, Hinney A (2010) Chipping away the 'missing heritability': GIANT steps forward in the molecular elucidation of obesity - but still lots to go. *Obes Facts* 3: 294–303. Review.
- Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, et al. (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 42: 937–948.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545–15550.
- Wang K, Li M, Bucan M (2007) Pathway-based approaches for analysis of genomewide association studies. *Am J Hum Genet* 81: 1278–1283.
- Torkamani A, Topol EJ, Schork NJ (2008) Pathway analysis of seven common diseases assessed by genome-wide association. *Genomics* 92: 265–272.
- Liu YJ, Guo YF, Zhang LS, Pei YF, Yu N, et al. (2010) Biological pathway-based genome-wide association analysis identified the vasoactive intestinal peptide (VIP) pathway important for obesity. *Obesity* 18: 2339–2346.
- Evangelou M, Rendon A, Ouwehand WH, Wernisch L, Dudbridge F (2012) Comparison of methods for competitive tests of pathway analysis. *PLOS ONE* 7: e41018.
- Day N, Oakes S, Luben R, Khaw K, Bingham S, et al. (1999) EPIC-Norfolk: study design and characteristics of the cohort. *European Prospective Investigation of Cancer. Brit J Canc* 80: 95–103.
- Kelley DE, He J, Menshikova EV, Ritov VB (2002) Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes. *Diabetes* 51: 2944–2950.
- Niemann B, Chen Y, Teschner M, Li L, Silber RE, et al. (2011) Obesity induces signs of premature cardiac aging in younger patients: the role of mitochondria. *J Am Coll Cardiol* 57: 577–585.
- Garesse R and Vallejo CG (2001) Animal mitochondrial biogenesis and function: a regulatory cross-talk between two genomes. *Gene* 263: 1–16. Review.
- Goffart S and Wiesner RJ (2003) Regulation and co-ordination of nuclear gene expression during mitochondrial biogenesis. *Exp Physiol* 88: 33–40. Review.
- Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, et al. (2008) A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134: 112–123.
- Bar-Yaacov D, Blumberg A, Mishmar D (2012) Mitochondrial-nuclear co-evolution and its effects on OXPHOS activity and regulation. *Biochim Biophys Acta* 1819: 1107–1111.
- Segrè AV, DIAGRAM Consortium, MAGIC investigators, Groop L, Mootha VK, et al. (2010) Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet* 6, pii: e1001058.
- Kelly DP, Scarpulla RC (2004) Transcriptional regulatory circuits controlling mitochondrial biogenesis and function. *Genes Dev* 18: 357–368. Review.
- Wan J, Jiang L, Lü Q, Ke L, Li X, et al. (2010) Activation of PPARdelta up-regulates fatty acid oxidation and energy uncoupling genes of mitochondria and reduces palmitate-induced apoptosis in pancreatic beta-cells. *Biochem. Biophys. Res Commun* 391: 1567–1572.
- Finck BN and Kelly DP (2006) PGC-1 coactivators: inducible regulators of energy metabolism in health and disease. *J Clin Invest* 116: 615–622. Review.
- Giguère V (2008) Transcriptional control of energy homeostasis by the estrogen-related receptors. *Endocr Rev* 29, 677–696. Review.
- Yu J, Auwerx J (2009) The role of sirtuins in the control of metabolic homeostasis. *Ann N Y Acad Sci* 1173: E10–19. Review.
- Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, et al. (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34: 267–273.
- Scherag A, Dina C, Hinney A, Vatin V, Scherag S, et al. (2010) Two new Loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and German study groups. *PLoS Genet* 6(4): e1000916.
- Hinney A, Nguyen TT, Scherag A, Friedel S, Brönnner G, et al. (2007) Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (FTO) variants. *PLoS One* 2(12): e1361.
- Hebebrand J, Hesecker H, Himmelmann GW, Schäfer H, Remschmidt H (1994) Altersperzentilen für den Body Mass Index aus Daten der Nationalen Verzehrstudie einschließlich einer Übersicht zu relevanten Einflussfaktoren. *Aktuelle Ernährungsmedizin* 19: 259–265.
- Rückert IM, Heier M, Rathmann W, Baumeister SE, Döring A, et al. (2011) Association between markers of fatty liver disease and impaired glucose regulation in men and women from the general population: the KORA-F4-study. *PLoS One* 6(8): e22932.
- Pütter C, Pechlivanis S, Nöthen MM, Jöckel KH, Wichmann HE, et al. (2011) Missing heritability in the tails of quantitative traits? A simulation study on the impact of slightly altered true genetic models. *Hum Hered* 72: 173–181.
- Kazeem GR, Farrall M (2005) Integrating case-control and TDT studies. *Ann Hum Genet* 69: 329–335.
- Wigginton JE, Cutler DJ, Abecasis GR (2005) A note on exact tests of Hardy-Weinberg equilibrium. *Am. J Hum Genet* 76: 887–893.
- Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, et al. (2008) High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet* 4(10): e1000214.
- Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52: 506–516.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
- Pennisi E (1998) A closer look at SNPs suggests difficulties. *Science* 281: 1787–1789.
- Neale BM, Sham PC (2004) The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 75: 353–62.
- Hebebrand J, Wulfstange H, Goerg T, Ziegler A, Hinney A, et al. (2000) Epidemic obesity: are genetic factors involved via increased rates of assortative mating? *Int J Obes Relat Metab Disord* 24: 345–353.
- White R, Leonardsson G, Rosewell I, Ann Jacobs M, Milligan S, et al. (2000) The nuclear receptor co-repressor nrip1 (RIP140) is essential for female fertility. *Nat Med* 6: 1368–1374.
- Bordone L, Cohen D, Robinson A, Motta MC, van Veen E, et al. (2007) SIRT1 transgenic mice show phenotypes resembling calorie restriction. *Aging Cell* 6: 759–767.
- Vimaleswaran KS, Tachmazidou I, Zhao JH, Hirschhorn JN, Dudbridge F, et al. (2012) Candidate genes for obesity-susceptibility show enriched association within a large genome-wide association study for BMI. *Hum Mol Genet* 21: 4537–4542.
- Luo J, Sladek R, Carrier J, Bader JA, Richard D, et al. (2003) Reduced fat mass in mice lacking orphan nuclear receptor estrogen-related receptor alpha. *Mol Cell Biol* 23: 7947–7956.
- Pi J, Leung L, Xue P, Wang W, Hou Y, et al. (2010) Deficiency in the nuclear factor E2-related factor-2 transcription factor results in impaired adipogenesis and protects against diet-induced obesity. *J Biol Chem* 285: 9292–9300.
- Riu E, Ferre T, Hidalgo A, Mas A, Franckhauser S, et al. (2003) Overexpression of c-myc in the liver prevents obesity and insulin resistance. *FASEB J* 17: 1715–1717.
- Akiyama TE, Nicol CJ, Fievet C, Staels B, Ward JM, et al. (2001) Peroxisome proliferator-activated receptor-alpha regulates lipid homeostasis, but is not associated with obesity: studies with congenic mouse lines. *J Biol Chem* 276: 39088–39093.
- Peters JM, Lee SS, Li W, Ward JM, Gavrilova O, et al. (2000) Growth, adipose, brain, and skin alterations resulting from targeted disruption of the mouse peroxisome proliferator-activated receptor beta(delta). *Mol Cell Biol* 20: 5119–5128.
- Leone TC, Lehman JJ, Finck BN, Schaeffer PJ, Wende AR, et al. (2005) PGC-1alpha deficiency causes multi-system energy metabolic derangements: muscle dysfunction, abnormal weight control and hepatic steatosis. *PLoS Biol* 3(4): e101.
- Lelliott CJ, Medina-Gomez G, Petrovic N, Kis A, Feldmann HM, et al. (2006) Ablation of PGC-1beta results in defective mitochondrial activity, thermogenesis, hepatic function, and cardiac performance. *PLoS Biol* N; 4(11): e369.