**Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits**

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

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1111/acel.12490</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:29408427">http://nrs.harvard.edu/urn-3:HUL.InstRepos:29408427</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>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 <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits


1Institute for Community Medicine, University Medicine Greifswald, 17475, Greifswald, Germany
2Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17475, Greifswald, Germany
3Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, USA
4Bioinformatics Core Facility, The Sahlgrenska Academy, University of Gothenburg, Gothenburg 40530, Sweden
5Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
6Geriatric Unit, Azienda Sanitaria di Firenze (ASF), Florence, Italy
7Molecular Epidemiology, Leiden University Medical Center, Leiden 2300 RC, The Netherlands
8Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892, USA
9Medizinische Klinik und Poliklinik IV, Klinikum der Universitaet Muenchen, 80336, Munich, Germany
10Department of Epidemiology, Subdivision Genetic Epidemiology, Erasmus Medical Center, Postbus 2040, 3000 CA, Rotterdam, The Netherlands
11Department of Internal Medicine, Erasmus Medical Center, Postbus 2040,3000 CA, Rotterdam, The Netherlands
12Division of Endocrinology, Diabetes, and Metabolism, Department of Medicine, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA
13Section of Geriatrics, Department of Clinical and Experimental Medicine, University of Parma, Parma, Italy
14Geriatric Rehabilitation Department, University-Hospital of Parma, Parma, Italy
15Framingham Heart Study, Framingham, MA 01702, USA
16Department of Neurology, Boston University School of Medicine, Boston, MA 02118, USA
17Section of Endocrinology, Diabetes & Nutrition, Boston University School of Medicine, Boston, MA 02118, USA
18Institute for Translational Genomics and Population Sciences, Los Angeles BioMedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA 90502, USA
19Department of Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, USA
20Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano 39100, Italy
21Affiliated Institute of the University of Lübeck, 23562, Lübeck, Germany
22Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, 41345, Gothenburg, Sweden
23Translational Gerontology Branch, National Institute on Aging, Baltimore, MD 21225, USA
24Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, 17475, Greifswald, Germany
25Institute for Clinical Chemistry and Laboratory Medicine, Regensburg University Medical Center, D-93053, Regensburg, Germany
26Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
27The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, 2100, Copenhagen, Denmark
28Division of Epidemiology, Department of Population Health, New York University School of Medicine, New York, NY 10016, USA
29Department of Epidemiology, Erasmus University Medical Center, Rotterdam 3000 CA, The Netherlands
30Leiden University Medical Center, Medical Statistics and Bioinformatics, Leiden 2300 RC, The Netherlands
31Department of Nutrition, Harvard School of Public Health, Boston, MA 02115, USA
32Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA 02115, USA
33Department of Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston, MA 02115, USA
34Research Centre for Prevention and Health, Capital Region of Denmark, Copenhagen DK-2600, Denmark
35Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam 3000 CA, The Netherlands
36Centre for Medical Systems Biology, Leiden 2300 RC, The Netherlands
37Department of Internal Medicine, Erasmus University Medical Center, Rotterdam 3000 CA, The Netherlands
38Department of Physiology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, 41345, Gothenburg, Sweden

Correspondence
Robert C. Kaplan, 1300 Morris Park Avenue, Albert Einstein College of Medicine, Bronx, NY 10461, USA. Tel.: 718 430 4076; fax: x718-430-3588; e-mail: robert.kaplan@einstein.yu.edu

*These authors contributed equally to this work.

Accepted for publication 1 April 2016
Interactions, which were characterized by different genotype–phenotype associations between men and women, were found only for associations of IGFBP-3 concentrations with SNPs at the loci IGFBP3 and SORCS2. Analyses of SNPs, gene expression, and protein levels suggested that interplay between IGFBP3 and genes within the NUBP2 locus (IGFALS and HAGH) may affect circulating IGF-I and IGFBP-3 concentrations. The IGF-I-decreasing allele of SNP rs934073, which is an eQTL of ASXL2, was associated with lower adiposity and higher likelihood of survival beyond 90 years. The known longevity-associated variant rs2153960 (FOXO3) was observed to be a genomewide significant SNP for IGF-I concentrations. Bioinformatics analysis suggested enrichment of putative regulatory elements among these IGF-I- and IGFBP-3-associated loci, particularly of rs646776 at CELSR2.

In conclusion, this study identified several loci associated with circulating IGF-I and IGFBP-3 concentrations and provides clues to the potential role of the IGF axis in mediating effects of known (FOXO3) and novel (ASXL2) longevity-associated loci.

Key words: aging; genomewide association study; growth hormone axis; IGF-I; IGFBP-3; longevity.

Introduction

The insulin-like growth factor (IGF) axis is an evolutionarily conserved system that plays important biologic roles in embryonic development, growth, and adulthood (Le Roith, 1997). IGF-I mediates most of the activity of growth hormone (GH). The GH/IGF system consists of two ligands (IGF-I and IGF-II), six IGF-binding proteins (IGFBP-1-6), and three IGF receptor subtypes (IGF-I receptor, IGF-II receptor, and insulin receptor) (Jones & Cleemans, 1995). IGF-I promotes mitosis and cell cycle progression and is involved in human postnatal growth and development. Circulating IGF-I is mainly bound to IGF-binding proteins (IGFBPs) (Fowkes, 1997), which affect activity (Lee et al., 1997) and half-life of IGF-I (Guler et al., 1989). From the clinical point of view, the measurement of IGF-I and IGFBP-3 blood concentrations is an important tool in establishing the diagnosis as well as monitoring treatment of GH-related diseases (Ho & Participants, 2007; Cohen et al., 2008; Melmed et al., 2009).

Circulating concentrations of IGF-I and IGFBP-3 have been associated with risk of type 2 diabetes, cardiovascular diseases, cancer, and mortality in epidemiological studies (Juu et al., 2002; Vasan et al., 2003; Renihan et al., 2004; Kaplan et al., 2007; Friedrich et al., 2009; Burgers et al., 2011; Rajpathak et al., 2012). In animal models, diminished IGF-I/insulin signaling has been associated with extended lifespan (Ziv & Hu, 2011), although the role of the IGF axis in human longevity remains inconclusive. Human genetic studies have suggested an association between polymorphisms in IGF-I signaling pathway genes and longevity (Willcox et al., 2008; Ziv & Hu, 2011; Di Bona et al., 2014). Heritability studies have provided evidence for a substantial genetic contribution to circulating concentrations of IGF-I (with heritability estimates ~40–60%) and IGFBP-3 (80%) (Harrela et al., 1996; Hong et al., 1996; Souren et al., 2007). Identifying genetic determinants of circulating IGF-I and IGFBP-3 could lead to a better understanding of the biological basis of these factors in relation to human health and might identify pathways or susceptibility biomarkers that may assist with the development of interventions that target IGF-I, its receptors, and binding proteins.

Our prior genomewide association study (GWAS) (N = 10 280) identified four loci with genomewide significant (P < 5 × 10⁻⁸)
associations with circulating IGF-I and IGFBP-3 concentrations, including SNPs in or near IGFBP3, TNS3, SORCS2, and NUBP2/IGFALS, as well as three additional loci with suggestive associations (P < 1 × 10^{-5}) in or near RPA3, SP0CK2, and FOXO3 (Kaplan et al., 2011). Some of these genes are involved in well-described IGF regulatory or signaling pathways (such as IGFBP3 genes) and are believed to influence traits that are also associated with concentrations or bioactivity of IGFs (e.g., FOXO3 locus associated with longevity (Willcox et al., 2008) and IGFBP3 locus associated with hip osteoarthritis (Evans et al., 2014)). To identify additional genetic variants with smaller effect sizes and enable sex-specific analyses, we expanded our GWAS meta-analysis to include up to a total of 30,884 individuals of European ancestry from 21 studies with measured circulating concentrations of IGF-I and IGFBP-3. In addition, using published GWAS data, we also performed lookups of associations of identified IGF-I and IGFBP-3 loci with survival beyond 90 years and other age-related clinical traits.

Results

Characteristics of study samples

An overview of the study samples and data collection methods can be found in Tables S1 and S2 (Supporting information). Analyses of IGF-I included up to 30,884 individuals (14,424 men and 16,460 women) from 21 studies and analyses of IGFBP-3 included up to 18,995 individuals (8,053 men and 10,942 women) from 13 studies.

Loci associated with circulating IGF-I and IGFBP-3 concentrations

An overview of the GWAS meta-analysis results is given by the Manhattan plots in Fig. 1 and in Fig. S1 (Supporting information). There was no indication of inflated test statistics (i.e., due to unaccounted population stratification) as seen by the quantile–quantile (QQ) plots, and genomic control lambda ranged from 1.02 to 1.08 for the meta-analysis results (Fig. S2, Supporting information) and from 0.98 to 1.08 (median 1.01) for the individual GWAS results. All lead SNPs (independent SNPs with the smallest P-value within a locus, see Methods) had a good imputation quality across the studies (median imputation quality >0.9).

After the final stage, which combines results of stages 1 and 2 plus de novo genotyping in stage 3, we found seven genomewide significant loci (P < 5.0 × 10^{-8}) associated with circulating IGF-I concentration (Table 1). In addition to the known locus near TNS3, we identified new loci in or near GCKR, IGF1, FOXO3, ASXL2, NUBP2, and GHSR associated with IGF-I concentrations.

We found genomewide significant associations with IGFBP-3 concentrations for SNPs in or near IGFBP3, TNS3, NUBP2, and SORCS2, thus confirming all four previously known loci (Table 1). The SNPs at TNS3 and NUBP2 were genomewide significantly associated with both IGF-I and IGFBP-3 concentrations and had the same direction of effect for each circulating protein.

For six of ten genomewide significant SNPs, effects were in the same direction of association for IGF-I concentrations and IGFBP-3 concentrations (Table 1).

Detailed results of the significant associations after the final stage can be found in Table 1. Results of the individual analysis stages appear in Table S3 (Supporting information). Regional association plots are shown in Figs S3 and S4 (Supporting information).

Bivariate analysis of IGF-I and IGFBP-3

We performed a bivariate analysis of IGF-I and IGFBP-3. By leveraging shared variance between the two outcomes, this analysis can have improved power to identify SNPs associated with both IGF-I and IGFBP-3 concentrations, especially in the case of SNPs that have opposite effects on positively correlated traits (Aschard et al., 2014). The bivariate analysis identified a new locus at CELSR2 (Table 1), which had nominal association with IGFBP-3 (P = 2.08 × 10^{-9}) and IGF-1 (P = 0.0096) in the univariate analysis. SNP rs646776 at CELSR2 had opposite effects on the two traits, being negatively associated with IGF-1 and positively associated with IGFBP-3 (Table 1, Table S4 and Fig. S5, Supporting information). In addition, SNPs at IGFBP3, TNS3, NUBP2, SORCS2, GCKR, IGF1, and FOXO3, identified in the univariate analysis, also showed genomewide significant associations in the bivariate analysis.

Interaction by sex

The sex-stratified analyses revealed no additional discoveries that were not detected in the overall population. Although the direction of effect was similar for IGFBP3 and SORCS2 SNPs within sex subgroups, these two SNPs were found to have significantly different association effect sizes between men and women for IGFBP-3, consistent with stronger associations in women. These findings of sex interaction maintained statistical significance after Bonferroni correction for the 12 tested genomewide significant lead SNPs (P < 0.004) (Table S5, Supporting information).

Gene-based analysis (VEGAS)

Gene-based analyses showed several significant IGF-I-associated genes within or close to the GCKR GWAS locus: EIF2B4, FNDC4, GCKR, IFT172, PPM1G, SNX17, ZNF513, GTF3C2, KRTP3, MPV17, and NRBP1 (associated with IGF-I). New gene-based associations that were not covered by a single SNP GWAS association were found for C6orf173 (chromosome 6) on IGF-I concentration. The following genes of the NUBP2 GWAS locus were associated with circulating IGFBP-3 concentration: EME2, IGFL5, MAKP81P3, MRPS34, NME3, NUBP2, HS3ST6, RPL3L, SEPX1, and SPSB. Two genes, IGBP1 and IGFBP3, at the IGFBP3 locus were associated IGFBP-3 concentration.

Lookup for expression quantitative trait loci associations

A lookup of the lead SNPs for cis expression quantitative trait loci (eQTL) was performed in the publicly available database of whole blood eQTL associations (Westra et al., 2013). For the following SNPs, one or more cis eQTL associations were found: rs1065656 (NUBP2), rs11977526 (IGFBP3), rs2153960 (FOXO3), rs509035 (GHSR), rs780093 (GCKR), rs934073 (ASXL2), and rs978458 (IGF1) (Table 2).

Additionally, using a similar strategy we performed lookup in the MuTHER consortium (Grundberg et al., 2012) for cis eQTL associations found in fat cell, skin cell, and lymphoblastic cell lines (LCL). After Bonferroni correction for 297 lookups of three different traits (P < 5.6 × 10^{-8}), rs1065656 (NUBP2) showed significant associations, specifically with FAHD1 (all tissues) and HAGH (fat cells and LCL) (Table S6, Supporting information). Furthermore, both genes were also significant in whole blood cis eQTL for the SNP rs1065656 (NUBP2) (Table 2).
GWAS reveals IGF-I- and IGFBP-3-associated loci, A. Teumer et al.

© 2016 The Authors. Aging Cell published by the Anatomical Society and John Wiley & Sons Ltd.
**Fig. 1** Manhattan plots of the combined stage 1 and 2 meta-analysis results of IGF-I and IGFBP-3 traits in the men and women combined sample. SNPs are plotted on the x-axis according to their position on each chromosome with the -log10 association P-value on the y-axis. The upper solid horizontal line indicates the threshold for genomewide significance. Known hits are colored in orange and new findings in blue. Plots are truncated on the y-axis to 20.

<table>
<thead>
<tr>
<th>Trait</th>
<th>SNP</th>
<th>A1</th>
<th>A2</th>
<th>F1</th>
<th>P</th>
<th>P</th>
<th>Chr</th>
<th>Position</th>
<th>Nearest gene</th>
<th>Gene distance</th>
<th>Direction effect</th>
<th>IGF-I</th>
<th>IGFBP-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGF-I*</td>
<td>rs700753</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>1.60E-23</td>
<td>4.2</td>
<td>7</td>
<td>46,720,209</td>
<td>TNS3</td>
<td>561067</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGF-I</td>
<td>rs780093</td>
<td>T</td>
<td>C</td>
<td>2.19E-13</td>
<td>24.5</td>
<td>2</td>
<td>27,596,107</td>
<td>GCKR</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGF-I</td>
<td>rs978458</td>
<td>T</td>
<td>C</td>
<td>1.56E-10</td>
<td>0.0</td>
<td>12</td>
<td>101,326,369</td>
<td>IGFI</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGF-I</td>
<td>rs2153960</td>
<td>A</td>
<td>G</td>
<td>5.16E-09</td>
<td>22.5</td>
<td>6</td>
<td>109,082,339</td>
<td>FOXL2</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGF-I</td>
<td>rs934073</td>
<td>C</td>
<td>G</td>
<td>6.48E-09</td>
<td>21.8</td>
<td>2</td>
<td>25,790,669</td>
<td>NUBP2</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGF-I</td>
<td>rs1065656</td>
<td>C</td>
<td>G</td>
<td>1.17E-08</td>
<td>47.9</td>
<td>16</td>
<td>1,778,837</td>
<td>NUBP2</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGF-I</td>
<td>rs590935</td>
<td>A</td>
<td>G</td>
<td>2.09E-08</td>
<td>0.0</td>
<td>3</td>
<td>173,664,143</td>
<td>GHSR</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGFBP-3*</td>
<td>rs11977526</td>
<td>A</td>
<td>G</td>
<td>4.16E-161</td>
<td>51.5</td>
<td>7</td>
<td>45,918,692</td>
<td>G/AF1408</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGFBP-3*</td>
<td>rs700753</td>
<td>C</td>
<td>G</td>
<td>1.11E-46</td>
<td>26.7</td>
<td>7</td>
<td>46,720,209</td>
<td>TNS3</td>
<td>561067</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGFBP-3*</td>
<td>rs1065656</td>
<td>C</td>
<td>G</td>
<td>8.55E-23</td>
<td>24.1</td>
<td>16</td>
<td>1,778,837</td>
<td>NUBP2</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGFBP-3*</td>
<td>rs4234798</td>
<td>T</td>
<td>G</td>
<td>8.86E-19</td>
<td>0.0</td>
<td>4</td>
<td>7,270,834</td>
<td>SOCS5</td>
<td>0</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bivariate analysis</td>
<td>rs646776</td>
<td>T</td>
<td>C</td>
<td>6.87E-09</td>
<td>26.1/43.1</td>
<td>1</td>
<td>109,620,053</td>
<td>CELSR2</td>
<td>152</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Known association.

**Table 2** Results of significant whole blood eQTL associations of the genomewide significant lead SNPs

<table>
<thead>
<tr>
<th>SNP</th>
<th>GWAS locus</th>
<th>eQTL p-value</th>
<th>Chr</th>
<th>Probe center position</th>
<th>Probe name</th>
<th>SNP alleles</th>
<th>Effect allele</th>
<th>Effect direction</th>
<th>EQLT gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1065656</td>
<td>NUBP2</td>
<td>3.66E-04</td>
<td>16</td>
<td>1,829,861</td>
<td>1710332</td>
<td>C/G</td>
<td>C</td>
<td>+</td>
<td>FAHD1</td>
</tr>
<tr>
<td>rs1065656</td>
<td>NUBP2</td>
<td>4.28E-12</td>
<td>16</td>
<td>1,799,259</td>
<td>4900333</td>
<td>C/G</td>
<td>C</td>
<td>–</td>
<td>HAGH</td>
</tr>
<tr>
<td>rs1065656</td>
<td>NUBP2</td>
<td>3.78E-04</td>
<td>16</td>
<td>1,809,162</td>
<td>1780356</td>
<td>C/G</td>
<td>C</td>
<td>–</td>
<td>HAGH</td>
</tr>
<tr>
<td>rs1065656</td>
<td>NUBP2</td>
<td>9.35E-04</td>
<td>16</td>
<td>1,760,172</td>
<td>5270575</td>
<td>C/G</td>
<td>C</td>
<td>+</td>
<td>MAPK8B3</td>
</tr>
<tr>
<td>rs1065656</td>
<td>NUBP2</td>
<td>3.45E-05</td>
<td>16</td>
<td>1,762,997</td>
<td>6110307</td>
<td>C/G</td>
<td>C</td>
<td>+</td>
<td>MRPS34</td>
</tr>
<tr>
<td>rs1065656</td>
<td>NUBP2</td>
<td>7.75E-34</td>
<td>16</td>
<td>1,760,510</td>
<td>6450424</td>
<td>C/G</td>
<td>C</td>
<td>+</td>
<td>NME3</td>
</tr>
<tr>
<td>rs1065656</td>
<td>NUBP2</td>
<td>1.09E-14</td>
<td>16</td>
<td>1,778,738</td>
<td>6960730</td>
<td>C/G</td>
<td>C</td>
<td>–</td>
<td>NUBP2</td>
</tr>
<tr>
<td>rs1065656</td>
<td>NUBP2</td>
<td>9.87E-05</td>
<td>16</td>
<td>1,766,816</td>
<td>2850433</td>
<td>C/G</td>
<td>C</td>
<td>–</td>
<td>SPB3</td>
</tr>
<tr>
<td>rs2135960</td>
<td>FOXO3</td>
<td>2.95E-06</td>
<td>6</td>
<td>109,129,272</td>
<td>7200189</td>
<td>G/A</td>
<td>G</td>
<td>–</td>
<td>HS.133419</td>
</tr>
<tr>
<td>rs509035</td>
<td>GH5R</td>
<td>5.97E-05</td>
<td>3</td>
<td>173,706,575</td>
<td>870202</td>
<td>G/A</td>
<td>A</td>
<td>+</td>
<td>TNFSF10</td>
</tr>
<tr>
<td>rs780093</td>
<td>GCKR</td>
<td>2.46E-04</td>
<td>2</td>
<td>27,440,911</td>
<td>5960546</td>
<td>T/C</td>
<td>T</td>
<td>–</td>
<td>EF2B4</td>
</tr>
<tr>
<td>rs780093</td>
<td>GCKR</td>
<td>5.70E-04</td>
<td>2</td>
<td>27,440,904</td>
<td>6370494</td>
<td>T/C</td>
<td>T</td>
<td>–</td>
<td>EF2B4</td>
</tr>
<tr>
<td>rs780093</td>
<td>GCKR</td>
<td>2.69E-05</td>
<td>2</td>
<td>27,518,384</td>
<td>430239</td>
<td>T/C</td>
<td>T</td>
<td>–</td>
<td>NRPB1</td>
</tr>
<tr>
<td>rs780093</td>
<td>GCKR</td>
<td>1.00E-10</td>
<td>2</td>
<td>27,453,289</td>
<td>3360468</td>
<td>T/C</td>
<td>T</td>
<td>+</td>
<td>SNX17</td>
</tr>
<tr>
<td>rs934073</td>
<td>ASXL2</td>
<td>5.96E-04</td>
<td>2</td>
<td>25,816,038</td>
<td>650075</td>
<td>G/C</td>
<td>G</td>
<td>+</td>
<td>ASXL2</td>
</tr>
<tr>
<td>rs978458</td>
<td>IGFI</td>
<td>1.71E-03</td>
<td>12</td>
<td>101,115,257</td>
<td>990136</td>
<td>T/C</td>
<td>T</td>
<td>+</td>
<td>C12ORF48</td>
</tr>
<tr>
<td>rs11977526</td>
<td>IGFBP3</td>
<td>1.84E-05</td>
<td>7</td>
<td>45,918,692</td>
<td>6840372</td>
<td>G/A</td>
<td>A</td>
<td>–</td>
<td>IGFBP3</td>
</tr>
</tbody>
</table>

Chr, chromosome; eQTL, expression quantitative trait loci; GWAS, genomewide association study.

mRNA of probe and gene names marked in bold showed also significant association with circulating IGFBP-3 levels (P < 5 × 10⁻⁴).

**Associations of gene expression with circulating IGF-I and IGFBP-3 concentrations**

We next sought to link associations between SNPs and circulating IGF-I and IGFBP-3 concentrations, with cis eQTL associations of the same SNPs. In the 986 samples of the SHIP-TREND cohort, we examined associations between whole blood mRNA expression levels of the genes located in a 500-kb vicinity of our significant lead SNPs and circulating IGF-I and IGFBP-3 concentrations. Significance of the 323 array probe trait associations was defined by a false discovery rate (FDR) < 0.05. Only mRNA levels of genes in vicinity of the NUBP2 GWAS locus were significantly associated with IGF-I concentration (gene SEPX1) or IGFBP-3 concentration (genes HAGH and RPS2). Of note, HAGH was the gene on which the corresponding lead SNP (rs1065656) had also a significant cis eQTL. The complete gene expression association results are listed in Table S7 (Supporting information).

© 2016 The Authors. Aging Cell published by the Anatomical Society and John Wiley & Sons Ltd.
Association with plasma protein levels

In 197 samples of the SHIP-TREND cohort, peptides of the following proteins that were encoded by genes in a 500-kb vicinity of the lead SNPs were examined for protein quantitative trait analyses (pQTL): insulin-like growth factor-binding protein complex acid labile subunit (ALS encoded by IGFALS at the NUBP2 locus), 28S ribosomal protein S34, mitochondrial (RT34 encoded by MRPS34 at the NUBP2 locus), insulin-like growth factor-binding protein 3 (IGFBP3) at the IGFALS locus, and coiled-coil domain-containing protein 121 (CC121 encoded by CCDC121 at the GCKR locus). Of the 32 tested SNP peptide pairs, peptides of the ALS protein had significant pQTL at FDR <0.05 (Table S8, Supporting information). Furthermore, strong associations were found in the same samples for the pQTL-associated peptides of ALS and IBP3 with circulating levels of IGF-I and IGFBP-3 (P < 1.0 × 10⁻⁶).

Allelic heterogeneity of the NUBP2 locus

To further examine the NUBP2 locus, we performed a conditional analysis of this locus based on the meta-analysis results adjusting for the lead SNP rs1065656. The analyses revealed an independent genomewide significant association of a second SNP rs11644716 with IGFALS-3 (P = 6.3 × 10⁻¹⁵) and an opposite effect direction of the minor allele C (r² = 0.03) compared with the lead SNP rs1065656 (MAF = 31%) (P = 1.8 × 10⁻⁵) and has an eQTL for HAGH (probe 4900333: P = 3.2 × 10⁻⁵; probe 1780356: P = 0.003) and a pQTL with ALS (eight peptides with P-value <0.01). In all cases, the effect directions based on the minor allele of rs11644716 were opposite of that for the minor allele of rs1065656.

Associations with serum metabolites

Lead SNPs associated with IGF-I and IGFBP-3 concentrations were examined in a published metabolite-SNP association database (Suhre et al., 2011; Shin et al., 2014). The IGF-I-associated SNP rs780093 at GCKR locus was associated with glucose/mannose ratio (P = 9.4 × 10⁻¹⁴), and the IGFBP-3-associated SNP rs4234798 at SORCS2 locus was associated with caprylate (8:0)/phenylalanine ratio (P = 7.3 × 10⁻⁸).

Associations of top loci with age-related traits

We also examined the associations of the IGF-I- and IGFBP-3-associated SNPs with anthropometric traits (height, BMI, waist-to-hip ratio, and fat percentage) (Heid et al., 2010; Lango Allen et al., 2010; Speliotes et al., 2010), bone mineral density (Estrada et al., 2012), risk of type 2 diabetes (Voight et al., 2010; Morris et al., 2012) and related traits (fasting glucose, 2-h glucose, HbA1c, fasting insulin, proinsulin, HOMA-IR, and HOMA-B) (Dupuis et al., 2010; Saxena et al., 2010; Soranzo et al., 2010), and coronary artery disease (Coronary Artery Disease Genetics C, 2011; Schunkert et al., 2011; Consortium CAD and Deloukas, 2013) (Table S9, Supporting information). Many nominal associations were expected because of the known influence of the IGF system on these traits. Of note is the finding that for rs780093 in the GCKR locus, the allele associated with higher IGF-I concentration was already known to be associated with elevated risk of type 2 diabetes (P = 3.7 × 10⁻¹⁶), as well as higher levels of fasting glucose, fasting insulin, and HOMA-IR (all P < 2.0 × 10⁻¹⁶), lower 2-h glucose levels (P = 1.7 × 10⁻¹⁶), increased height (P = 2.0 × 10⁻¹⁶), lower waist-to-
hip ratio ($P = 0.0003$), and higher lumbar spine bone mineral density ($P = 0.002$).

Three additional loci (GHSR, CELSR2, and FOXO3) showed strong associations with height (all $P < 1.0 \times 10^{-4}$). The IGFBP-3-increasing allele of SNP rs646776 (CELSR2 locus) was associated with increased risk of coronary artery disease ($P = 9.4 \times 10^{-15}$). The IGF-I-decreasing allele of SNP rs934073 at ASXL2 showed a nominal association with survival beyond 90 years ($P = 0.018$) as well as higher levels of BMI ($P = 0.008$) and fat percentage ($P = 9.4 \times 10^{-4}$) and lower lumbar spine bone mineral density ($P = 0.004$).

We further performed lookups of top IGF-I- and IGFBP-3-associated SNPs ($P < 10^{-6}$ in the meta-analysis of stage 1 and stage 2) for associations with survival beyond 90 years using published GWAS data (Broer et al., 2014) (Tables S10 and S11, Supporting information). Among 15 independent circulating IGF-I-associated SNPs defined based on linkage disequilibrium (LD) (settings $r^2 > 0.01$, 1 Mb distance), the SNP rs10457180 ($r^2 = 0.96$ with the lead SNP rs2153960) at FOXO3 ($P = 8.6 \times 10^{-5}$) and SNP rs11892454 ($r^2 = 0.71$ with the lead SNP rs934073) at ASXL2 ($P = 0.003$) reached statistical significance after Bonferroni correction for 15 independent tests. Among 13 independent circulating IGFBP-3-associated SNPs, the SNP rs9398172 ($r^2 = 1$ with the lead SNP rs2153960) at FOXO3 ($P = 2.5 \times 10^{-4}$) remained significantly associated with survival beyond 90 years after Bonferroni correction.

**Enrichment of putative regulatory elements among loci associated with circulating IGF-1 and IGFBP-3 concentrations**

We examined whether the identified SNPs fall within regulatory elements in the epigenetic ENCODE and ROADMAP data for associated SNPs using Haploreg (http://www.broadinstitute.org/mammals/haploreg/haploreg.php) (Ward & Kellis, 2012) and RegulomeDB (http://regulomedb.org/) (Boyle et al., 2012) (Table S11, Supporting information). Lower scores indicate stronger evidence for the presence of a regulatory element. To determine whether these IGF-I- and IGFBP-3-associated loci are enriched for regions likely to affect gene expression, we further examined the distribution of scores among these SNPs compared with all RegulomeDB SNPs. We found that these identified SNPs are highly enriched for low Regulome scores ($P < 2.2 \times 10^{-16}$ by multinomial method, Fig. 3A). The genomic and representative epigenetic context surrounding rs646776 (CELSR2 locus), a SNP with a Regulome score of 1f, is shown as an example (Fig. 3). SNP rs646776 localizes to a genomic region of high LD (Fig. 3B) and lies within peaks of histone marks associated with regulatory elements, and a DNase hypersensitivity region (Fig. 3C) (Kent et al., 2002). In addition, SNP rs646776 falls in ChIP-identified binding regions for CTCF, POLR2A, REST, and TAF7 (data not shown).

**Discussion**

Our second GWAS report from the CHARGE IGF Working Group, here expanded to include more than ~30 000 individuals, revealed several SNPs influencing circulating levels of IGF-I and IGFBP-3 which also have been associated with other metabolic and age-related traits. These included several loci already implicated in the biology of GH/IGF-I (IGF1, IGFBP3, IGFLS, GHSR, FOXO3) as well other novel findings including rs934073 SNP on chr 2, which is an eQTL for polycomb group gene ASXL2 associated with reduced circulating IGF-I. Cross-reference of IGF-I- and IGFBP-3-associated SNPs against published GWAS of age-related traits identified the ASXL2 SNP as a (to our knowledge) novel locus for longevity as defined as survival beyond 90 years. All genomewide significant associations with IGFBP-3 and IGF-I levels that were reported in our preceding study (Kaplan et al., 2011) could be confirmed here using a larger sample. Additionally, our preliminary finding of an association of circulating IGF-I level with rs2153960 in the FOXO3 gene reached genomewide significance using this larger sample size. Moreover, bioinformatics analysis also suggests enrichment of putative regulatory elements among these IGF-I- and IGFBP-3-associated loci, particularly of rs646776 at CELSR2.

Our study reveals clues about mechanisms of IGF system regulation through the interplay of IGFBP-3 and genes within the NUBP2 locus, including IGFLS and HAGH. The IGFLS encodes the insulin-like growth factor-binding protein complex, acid labile subunit protein (ALS) which forms a ternary complex with IGF-I and IGFBP-3 (Firth et al., 1998; Twigg & Baxter, 1998). Like in most GWAS, our analyses cannot establish which is the causative SNP or gene of a locus. However, IGFLS seems to be a strong candidate supported by the associations of SNPs in the vicinity of IGFLS with both circulating IGFBP-3 and plasma levels of the protein coded by IGFLS. Although the sample size for the plasma proteome analyses was restricted to 197 probands of SHIP-TREND, the observed associations of the second signal rs11644716 achieved a moderately high level of statistical significance ($P < 0.0015$). IGFLS was not abundantly expressed in whole blood cells, and its transcript levels were neither associated with IGF-I nor IGFBP-3 concentration. Certainly, the protein ALS encoded by IGFLS is abundant in serum, but the gene is translated in liver. In contrast to IGFLS, HAGH (hydroxycycligluthionate hydrolase) was sufficiently expressed in whole blood cells and the amount of its mRNA was strongly correlated with circulating IGFBP-3 serum concentration. Taking into account the eQTL association of the lead SNP rs1065656 with HAGH and the correlation with the mRNA and IGFBP-3, this SNP might influence the IGFBP-3 levels by modulating the amount of HAGH mRNA, whereas both the levels of mRNA and IGFBP-3 are reduced per copy of the minor allele. Although this chain of associations was revealed in whole blood, it might be present in other tissues as well because significant eQTL for rs1065656 with HAGH were observed in other tissues studied in the MuTHER dataset (Grundberg et al., 2012). Given the more pronounced association with ALS and the less significant eQTL with HAGH of the second signal compared with the lead SNP, rs11644716 might reduce the level of circulating IGFBP-3 indirectly by reducing the amount of ALS per minor allele. Of note, there was no nonsynonymous SNP in LD in the 1000 Genomes v3 dataset ($R^2 > 0.8$, SHIP cohort) for both SNPs which could have helped narrow down the functional mechanism.

Taken together with genotype–phenotype association data assembled by others, our study revealed that IGF-I- and IGFBP-3-associated SNPs had expected associations with anthropometric and age-related chronic disease traits (e.g., bone mineral density, disordered carbohydrate metabolism). We also found that SNPs associated with reduced IGF-I levels tended to be associated with longer survival defined as death after 90 years (Broer et al., 2014). This is consistent with an observation from prior analysis of candidate genes associated with the insulin and IGF signaling axis (van Heemst et al., 2005). In addition, rs934073, an eQTL for additional sex comb-like 2 (ASXL2), was a genomewide significant SNP associated with lower IGF-I level which also has enriched frequency in adults older than 90 years of age (Broer et al., 2014). Other traits associated with the IGF-I-decreasing allele of rs934073 included greater adiposity and reduced lumbar spine (but not femoral neck) bone mineral density. ASXL2 is a polycomb group protein with known functions in development that has also been associated with pediatric cancer, but to our knowledge has not before been suggested as a
longevity gene (Huether et al., 2014). The SNP in the known longevity gene FOXO3 has not only been previously associated with reduced fasting insulin and HOMA-IR (Willcox et al., 2008), but here it was found to produce lower circulating IGF-I levels. Multiple genetic determinants for circulating IGF-I in normal and IGF1R resistance states might partially explain the U-shaped association of circulating IGF-I concentration with mortality (Suh et al., 2008; Burgers et al., 2011).

While our meta-analysis encompasses a large number of samples from multiple cohorts, this may lead to limitations. Given the different origin of the cohorts (Table S1, Supporting information), heterogeneity in the association results might occur due to the different genetic background and the patterns of intake of nutrients across the individual studies.

In summary, this project extends our prior work (Kaplan et al., 2011) through the identification of several new loci related to circulating IGF-I levels. Multiple genetic determinants for circulating IGF-I in normal and IGF1R resistance states might partially explain the U-shaped association of circulating IGF-I concentration with mortality (Suh et al., 2008; Burgers et al., 2011).

Experimental procedures

Participating studies

In total, the CHARGE IGF Working Group included 21 and 13 studies that participated in the association analysis for IGF-I (N = 27 520, 53% women) and IGFBP-3 (N = 18 995, 58% women), respectively. Four of the cohorts (N = 10 280) were previously included in a GWAS meta-analysis of IGF-I and IGFBP-3 levels (Kaplan et al., 2011). Imputed SNPs for chromosome X were available for 16 670 and 11 959 individuals with IGF-I and IGFBP-3 measurements, respectively. Additionally, up to 3364 individuals (55% women) with IGF-I from one study were available for de novo genotyping of selected SNPs. Detailed information on participant characteristics, IGF-I and IGFBP-3 measurements, and genotyping of all studies participated in the different analyses and stages is given in Table S2 (Supporting information). All participants provided informed consent, and human subjects’ research review was obtained from each participating cohort.

Statistical analyses

GWAS in individual studies

Each study of the GWAS stages performed genotyping on genomewide arrays and imputed SNPs using the HapMap2 reference panel. Detailed
information on genotyping and imputation is provided in Table S2 (Supporting information). Association analyses in individual studies were performed on IGF-I and IGFBP-3 levels measured in ng ml\(^{-1}\) using a multiple linear regression with an additive genetic model based on allele dosages adjusted for age and stratified by sex. All cohorts accounted for relatedness, population substructure using genetic principal components, study center, and laboratory batch of IGF measurement where applicable. Individuals of non-European ancestry, with missing phenotypic data, diagnosed growth hormone deficiency, or known use of human growth hormones were excluded prior to the analyses.

**Meta-analysis**

From each cohort’s result file, monomorphic SNPs as well as SNPs with an imputation quality below 0.3 were excluded prior to the meta-analysis. All study-specific GWAS results were corrected by the genomic inflation factor \(\hat{\lambda}_{GC}\) if \(>1\). Due to the IGF-I and IGFBP-3 assay-based differences in both effect sizes and variances of measurements across cohorts, a sample size-weighted \(z\)-score-based meta-analysis implemented in METAL (Willer et al., 2010) was conducted, and the meta-analysis \(P\)-values were corrected for genomic inflation. After meta-analysis, SNPs with a MAF \(<1\%\) were removed from subsequent analyses.

Our multistage design had two GWAS stages (stages 1 and 2) and an additional stage (stage 3) with \(de\ novo\) genotyping data \((N = 3364\) individuals) to confirm novel loci. After stage 1 GWAS, all 19 lead SNPs from all traits with a \(P < 10^{-6}\) were taken forward to stage 2. All IGF-I lead SNPs of novel loci that had a combined stage 1 and stage 2 \(P < 10^{-8}\) (except GCKR) were selected for \(de\ novo\) replication in an additional cohort. An overview of the design and the significantly associated loci at each stage is provided in Fig. S6 (Supporting information). Details on SNP selection and quality control are given in the Appendix S1 (Supporting information). Regional association plots were generated using LocusZoom (Pruim et al., 2010).

**Assessment of independent signals**

To define a lead SNP of each locus, the association results of a GWAS stage with \(P\)-values \(<1 \times 10^{-5}\) were grouped based on the LD structure of the HapMap release 28 CEU dataset using PLINK (settings \(r^2 >0.01, 1\) Mb distance) (Purcell et al., 2007). Due to the strong association of the IGFBP3 locus with IGFBP-3, only one lead SNP was selected regardless of several grouped results.

The analysis of secondary signals in the NUBP2 locus was performed using the software gCTA (Yang et al., 2011) and the genotypes of the SHIP cohort as a reference, and was verified by an analysis using the genotypes of the NHS/HPFS cohorts as a reference.

**Sex interaction analysis**

Sex interactions on IGF-I and IGFBP-3 levels were obtained by comparing, for each SNP, the stage 2 meta-analysis \(z\)-scores from men \((z_{\text{men}})\) (IGF-I: \(N = 12\,917,\) IGFBP-3: \(N = 8052\)) and women \((z_{\text{women}})\) (IGF-I: \(N = 14\,602,\) IGFBP-3: \(N = 10\,942\)) using the formula \(z_{\text{interaction}} = (z_{\text{men}} - z_{\text{women}})/\sqrt{2},\) assuming independent effect sizes between men and women, and matched to a common effect allele.

**Bivariate meta-analysis of IGF-I and IGFBP-3**

The stage 2 meta-analysis \(z\)-scores of the combined samples IGF-I and IGFBP-3 were used to calculate a bivariate meta-analysis implemented in the function multipheno.T2 of the \texttt{x-package gtx} (version 0.0.8. http://CRAN.R-project.org/package=gtx). The function corresponds closely with Hotelling’s \(T^2\) test and calculates a multiphenotype association test for each marker based on the meta-analysis result \(z\)-statistics that is equivalent to using the subject-specific data to perform a multivariate analysis of variance.

**Gene-based analysis**

Genomewide gene-based tests which account for both gene length and LD between SNPs were performed by \texttt{vegas} 0.8.27 (Versatile Gene-Based Association Study) (Liu et al., 2010) using SNP \(P\)-value results from the overall meta-analyses. SNPs were allocated to one or more autosomal genes using gene boundaries \(\pm 50\) kb. We performed \(1 \times 10^{7}\) permutations and defined a gene-based \(P\)-value \(<1 \times 10^{-6}\) as gene-based genomewide significant.

**Gene expression and eQTL analysis**

For each of the lead SNPs of the significant loci after final stage, significant cis eQTL associations in whole blood, lymphocytes, subcutaneous fat, muscle, and skin were looked up in the publically available association result databases (Grundberg et al., 2012; Westra et al., 2013). Association analysis of whole blood gene expression data with serum IGF-I and IGFBP-3 levels was conducted in 986 samples of the SHIP-TREND cohort (Schurmann et al., 2012).

**Association with plasma protein levels**

Plasma proteome data were obtained as described in Appendix S1 (Supporting information) using liquid chromatography–mass spectrometry (LC-MS). \texttt{mascot} (in-house \texttt{mascot} server v2.3.2; Matrix Science, London, GB) search algorithm was used to match the generated peak lists with a human \texttt{fasta}-formatted database containing 20 268 unique sequence entries (reviewed human database, release of October 2011). Prior to data analyses, all peptide intensity values were log10-transformed and median–median-normalized. Association analyses between peptides and serum IGF-I and IGFBP-3 levels were performed by linear regression, adjusted for age, sex, and the MS processing batch. Associations of a SNP with the peptides were conducted by linear regression, adjusted for age, sex, and the first four principal components of a peptide-level-based principal component analysis. Protein intensities used for analyses were obtained by averaging the corresponding peptide intensities that passed the QC filter, and were put instead of the peptide intensities into the association model. All measured peptides that passed QC and that belonged to proteins which were encoded by genes located in a 500-kb vicinity of our lead SNPs were selected for association analyses. The assignment of protein names (uniprot identifiers) to the corresponding genes was performed using the DAVID gene conversion tool (http://david.abcc.ncifcrf.gov). Finally, after QC the following proteins measured in 197 SHIP-TREND samples were available: ALS, CC121, IBP3, and RT34.

**Lookups of top loci in association with IGF correlated traits**

Top SNPs associated with levels of IGF-I and IGFBP-3 were examined in relationship to other phenotypes using published data on serum metabolites (Suhre et al., 2011; Shin et al., 2014), anthropometric traits (Heid et al., 2010; Lango Allen et al., 2010; Speliotes et al., 2010), bone mineral density (Estradá et al., 2012), diabetes (Voight et al., 2010; Morris et al., 2012) and glycemic traits (Dupuis et al., 2010; Saxena et al., 2010; Soranzo et al., 2010), coronary artery disease (Coronary Artery Disease Genetics C, 2011; Schunkert et al., 2011; Consortium CAD and Deloukas, 2013), and survival beyond 90 years (Broer et al., 2014). Detailed information of the published datasets used including its references is given in Table S9 (Supporting information).
GWAS reveals IGF-I- and IGFBP-3-associated loci, A. Teumer et al.

Assessment of regulatory elements associated with identified loci
ENCODE and ROADMAP data were assessed using HAPLOGREG (http://www.broadinstitute.org/mammals/haploreg/haplogreg.php) and REGULOME (http://regulomedb.org). Statistical analysis of individual Regu-
logene scores was performed using Monte Carlo sampling of 10 SNPs (the size of our “observed data” pool). RegulomeDB assigns scores to SNP loci based on the presence of histone marks, predicted and experimentally validated transcription factor binding, DNase hypersensitivity, and other evidence for regulatory function. Scores range from 1 to 7, with lower scores indicating stronger evidence for the presence of a regulatory element. For the purpose of this analysis, score subcategories (1a, 1b, etc.) were merged. A multinomial test was performed for a statistical comparison between the observed distribution and the background distribution. The LD plot in Fig. 3B was generated using HAPLOGREG 4.2 with genetic data downloaded from version 3, Release 2, using the genomic region Chr1:10959000-10963000, and analysis panel CEU + TSI. Histone mark and DNase tracks in Fig. 3C were downloaded from UCSC Genome Browser.

Funding
Funding sources for the individual studies included in this work are listed in Supporting information.

Conflict of interest
The authors do not have conflict of interest.

References


Broer L, Buchman AS, Deelen J, Evans DS, Faul JD, Lunetta KL, Sebastiani P, Smith JA, Smith AV, Tanaka T, Yu L, Arnold AM, Aschard H, Vilhjalmsson BJ, Greiliche N, Morange PE, Tregouet DA, Kraft P (2014) Assessment of regulatory elements associated with identified loci based on the presence of histone marks, predicted and experimentally validated transcription factor binding, DNase hypersensitivity, and other evidence for regulatory function. Scores range from 1 to 7, with lower scores indicating stronger evidence for the presence of a regulatory element. For the purpose of this analysis, score subcategories (1a, 1b, etc.) were merged. A multinomial test was performed for a statistical comparison between the observed distribution and the background distribution. The LD plot in Fig. 3B was generated using HAPLOGREG 4.2 with genetic data downloaded from version 3, Release 2, using the genomic region Chr1:10959000-10963000, and analysis panel CEU + TSI. Histone mark and DNase tracks in Fig. 3C were downloaded from UCSC Genome Browser.

Funding
Funding sources for the individual studies included in this work are listed in Supporting information.

Conflict of interest
The authors do not have conflict of interest.

References


Broer L, Buchman AS, Deelen J, Evans DS, Faul JD, Lunetta KL, Sebastiani P, Smith JA, Smith AV, Tanaka T, Yu L, Arnold AM, Aschard H, Vilhjalmsson BJ, Greiliche N, Morange PE, Tregouet DA, Kraft P (2014) Assessment of regulatory elements associated with identified loci based on the presence of histone marks, predicted and experimentally validated transcription factor binding, DNase hypersensitivity, and other evidence for regulatory function. Scores range from 1 to 7, with lower scores indicating stronger evidence for the presence of a regulatory element. For the purpose of this analysis, score subcategories (1a, 1b, etc.) were merged. A multinomial test was performed for a statistical comparison between the observed distribution and the background distribution. The LD plot in Fig. 3B was generated using HAPLOGREG 4.2 with genetic data downloaded from version 3, Release 2, using the genomic region Chr1:10959000-10963000, and analysis panel CEU + TSI. Histone mark and DNase tracks in Fig. 3C were downloaded from UCSC Genome Browser.

Funding
Funding sources for the individual studies included in this work are listed in Supporting information.

Conflict of interest
The authors do not have conflict of interest.

References


Broer L, Buchman AS, Deelen J, Evans DS, Faul JD, Lunetta KL, Sebastiani P, Smith JA, Smith AV, Tanaka T, Yu L, Arnold AM, Aschard H, Vilhjalmsson BJ, Greiliche N, Morange PE, Tregouet DA, Kraft P (2014) Assessment of regulatory elements associated with identified loci based on the presence of histone marks, predicted and experimentally validated transcription factor binding, DNase hypersensitivity, and other evidence for regulatory function. Scores range from 1 to 7, with lower scores indicating stronger evidence for the presence of a regulatory element. For the purpose of this analysis, score subcategories (1a, 1b, etc.) were merged. A multinomial test was performed for a statistical comparison between the observed distribution and the background distribution. The LD plot in Fig. 3B was generated using HAPLOGREG 4.2 with genetic data downloaded from version 3, Release 2, using the genomic region Chr1:10959000-10963000, and analysis panel CEU + TSI. Histone mark and DNase tracks in Fig. 3C were downloaded from UCSC Genome Browser.

Funding
Funding sources for the individual studies included in this work are listed in Supporting information.

Conflict of interest
The authors do not have conflict of interest.

References


Broer L, Buchman AS, Deelen J, Evans DS, Faul JD, Lunetta KL, Sebastiani P, Smith JA, Smith AV, Tanaka T, Yu L, Arnold AM, Aschard H, Vilhjalmsson BJ, Greiliche N, Morange PE, Tregouet DA, Kraft P (2014) Assessment of regulatory elements associated with identified loci based on the presence of histone marks, predicted and experimentally validated transcription factor binding, DNase hypersensitivity, and other evidence for regulatory function. Scores range from 1 to 7, with lower scores indicating stronger evidence for the presence of a regulatory element. For the purpose of this analysis, score subcategories (1a, 1b, etc.) were merged. A multinomial test was performed for a statistical comparison between the observed distribution and the background distribution. The LD plot in Fig. 3B was generated using HAPLOGREG 4.2 with genetic data downloaded from version 3, Release 2, using the genomic region Chr1:10959000-10963000, and analysis panel CEU + TSI. Histone mark and DNase tracks in Fig. 3C were downloaded from UCSC Genome Browser.

Funding
Funding sources for the individual studies included in this work are listed in Supporting information.

Conflict of interest
The authors do not have conflict of interest.

References


Broer L, Buchman AS, Deelen J, Evans DS, Faul JD, Lunetta KL, Sebastiani P, Smith JA, Smith AV, Tanaka T, Yu L, Arnold AM, Aschard H, Vilhjalmsson BJ, Greiliche N, Morange PE, Tregouet DA, Kraft P (2014) Assessment of regulatory elements associated with identified loci based on the presence of histone marks, predicted and experimentally validated transcription factor binding, DNase hypersensitivity, and other evidence for regulatory function. Scores range from 1 to 7, with lower scores indicating stronger evidence for the presence of a regulatory element. For the purpose of this analysis, score subcategories (1a, 1b, etc.) were merged. A multinomial test was performed for a statistical comparison between the observed distribution and the background distribution. The LD plot in Fig. 3B was generated using HAPLOGREG 4.2 with genetic data downloaded from version 3, Release 2, using the genomic region Chr1:10959000-10963000, and analysis panel CEU + TSI. Histone mark and DNase tracks in Fig. 3C were downloaded from UCSC Genome Browser.


