The following full text is a publisher’s version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/197679

Please be advised that this information was generated on 2019-05-13 and may be subject to change.
Genome-wide analyses identify a role for SLC17A4 and AADAT in thyroid hormone regulation

Alexander Teumer, Layal Chaker, Stefan Groeneweg, Yong Li, Celia Di Munno, Caterina Barbieri et al.

Thyroid dysfunction is an important public health problem, which affects 10% of the general population and increases the risk of cardiovascular morbidity and mortality. Many aspects of thyroid hormone regulation have only partly been elucidated, including its transport, metabolism, and genetic determinants. Here we report a large meta-analysis of genome-wide association studies for thyroid function and dysfunction, testing 8 million genetic variants in up to 72,167 individuals. One-hundred-and-nine independent genetic variants are associated with these traits. A genetic risk score, calculated to assess their combined effects on clinical end points, shows significant associations with increased risk of both overt (Graves’ disease) and subclinical thyroid disease, as well as clinical complications. By functional follow-up on selected signals, we identify a novel thyroid hormone transporter (SLC17A4) and a metabolizing enzyme (AADAT). Together, these results provide new knowledge about thyroid hormone physiology and disease, opening new possibilities for therapeutic targets.
Thyroid dysfunction is a common clinical condition, affecting ~10% of the general adult population. Adequate thyroid hormone levels are essential for normal growth and differentiation, regulation of energy metabolism, and physiological function of virtually all human tissues. Thyroxine (T4) is the prohormone produced by the thyroid, which is largely converted into the active hormone 3,3',5'-triiodothyronine (T3) in peripheral tissues. Circulating T4 levels are regulated by the hypothalamus–pituitary–thyroid (HPT) axis, in which pituitary thyroid-stimulating hormone (TSH) stimulates T4 production. In turn, T4 and T3 negatively regulate TSH synthesis via a negative feedback loop.

To exert their actions, T4 and T3 cross the membranes of target cells via specific transporters. Once intracellular, they are metabolized, including the conversion of T4 to T3, followed by binding of T3 to its nuclear receptor to regulate transcription of target genes. Both T4 and T3 transport and metabolism are therefore key determinants of thyroid hormone action.

In daily clinical practice, thyroid function is assessed by measuring circulating TSH and free T4 (FT4) levels, with increased TSH indicating hypothyroidism and decreased TSH indicating hyperthyroidism. FT4 levels are decreased in overt hypothyroidism and in the reference range in subclinical hypo and hyperthyroidism. In the last decade, it has become clear that not only overt but also subclinical hypothyroidism and hyperthyroidism are associated with several pathological conditions, such as atrial fibrillation, coronary heart disease, stroke, depression, as well as cardiovascular and overall mortality. More recently, studies have shown that even variation in thyroid function within the normal range is associated with many of these complications. Despite the physiological significance of thyroid hormones, as well as the prevalence and clinical importance of thyroid dysfunction, many key players in the regulation of thyroid hormone bioavailability and action, including its transport and metabolism, still need to be elucidated.

Genome-wide association studies (GWAS) performed so far have revealed genetic variants in about 30 loci robustly associated with thyroid function. However, these variants explain only <9% of the heritability of TSH and FT4 variation, while in total, it has been estimated at 65 and 39–80% for TSH and FT4, respectively, suggesting that many loci still await discovery. Here, we report the results of a large meta-analysis of GWAS for circulating TSH and FT4 levels, as well as for hypo and hyperthyroidism, followed by independent replication and functional studies. Results are complemented with genetic risk score (GRS) analyses, gene expression, co-localization analyses, and associations with various clinical phenotypes to discover new pathways underlying thyroid function and disease. We identify 109 significant independent genetic associations with these traits. The GRS shows a significant association with increased risk of both Graves’ disease and subclinical thyroid disease, as well as clinical complications. Finally, we identify a novel thyroid hormone transporter and a metabolizing enzyme. Together, these results enhance our knowledge about thyroid hormone physiology and disease.

Results

New loci affecting thyroid hormone levels. Our GWAS meta-analyses and replication in up to 72,167 subjects of European ancestry with TSH levels within the reference range (Supplementary Data 1) discovered 19 novel loci for circulating TSH levels and 16 novel loci for circulating FT4 levels (Tables 1 and 2, Supplementary Figures 2–5), leading to a total of 42 and 21 known and novel associated loci for these two traits. As illustrated in Fig. 1, TSH and FT4 capture distinct and complementary genetic underpinnings of thyroid function. Some of the novel loci include genes that have been previously implicated in thyroid development (GLIS3), thyroid hormone action and transport (NCOR1, TTR, SLC1B1), thyroid hormone metabolism (DIO2, DIO3OS), and thyroid cancer (e.g., HES1, SPATA13, DIRC3, ID4) by various candidate gene studies of monogenic diseases and animal models. Multiple independent variants were found for PDE8B, DIO1, DIO2, TSHR, and CAPZB.

Across the 42 TSH and 21 FT4 loci, allelic heterogeneity (i.e., independently associated single-nucleotide polymorphisms (SNPs) at the same locus) was detected at 11 and 7 loci, respectively, by using linkage structure information and summary statistics-based conditional analyses (Supplementary Tables 1 and 2). All significant associations together accounted for 33% and 21% of the genetic variance of TSH and FT4, respectively, explained by all common and low-frequency variants with a minor allele frequency (MAF) >1%.

Since TSH and FT4 regulation are inversely correlated through the HPT axis, we investigated the association of the TSH-associated loci with FT4 levels, and vice versa. As shown in Fig. 1 and in Supplementary Tables 1 and 2, we observed overlapping associations (Bonferroni-corrected threshold \( p < 8.2 \times 10^{-5} \)) at various loci (TSH: FGF7, PDE8B, DET1, ITPK1, VEGFA, GLIS3, NFIA, and MBIP, and FT4: FOXE1 and GLIS3), although only GLIS3 showed genome-wide significance for both traits. All alleles associated with higher TSH were associated with lower FT4, with the exception of MBIP and FOXE1.

Hypo and hyperthyroidism are more prevalent in women than in men. However, sex-stratified GWAS meta-analyses for TSH and FT4 did not show any significant gene-by-sex interaction in our samples (Supplementary Figure 6, Supplementary Tables 1 and 2).

Given the high degree of functional homology between the mouse and human genome, we selected from The International Mouse Phenotyping Consortium database that when manipulated in mice cause abnormal thyroid physiology (i.e., hormone levels, \( n = 26 \)) or morphology (\( n = 51 \)), and assessed whether their human homologs contained SNPs significantly \(( p < 2.5 \times 10^{-5} \) for physiology and \( p < 1.9 \times 10^{-5} \) for morphology, see Methods) associated in our FT4 and TSH GWAS (Supplementary Data 2). Of these candidate genes, SNPs in CGA (rs6924373) and TPO (rs9678281) contained significant associations that did not reach genome-wide significance in our GWAS for TSH. These associations were tested for replication in 9011 independent samples and achieved genome-wide significance for TSH (\( p < 5 \times 10^{-8} \)) in the combined dataset (Supplementary Figure 7A). Overall, these results highlight the potential of nested candidate gene approaches in GWAS summary results and emphasize the functional conservation of genes regulating thyroid function between mice and humans.

Relation to hypo and hyperthyroidism. Genetic variants that determine variation in circulating TSH and FT4 levels within the reference range (i.e., the individual HPT-axis setpoint) are expected to differ from variants that underlie thyroid dysfunction (hypo or hyperthyroidism). To clarify this, we also conducted a case–control GWAS meta-analysis of increased TSH levels (i.e., hyperthyroidism), including cases with TSH levels above the cohort-specific reference range (\( n = 3340 \)) and controls with TSH levels within the reference range (\( n = 49,983 \)). The decreased TSH level (i.e., hyperthyroidism) GWAS meta-analysis included cases with TSH below the reference range (\( n = 1840 \)) and the same controls as in the increased TSH GWAS. The distribution of sex and age groups of these subjects is provided in Supplementary Table 3. Since in both GWAS analyses, cases were defined on the basis of...
The table contains the list of independent SNPs and additional independent associations of replicated TSH susceptibility loci. The values are provided for the combined discovery and replication sample, for additional independent hits (*) for the discovery stage only.

Bold values of the hyper and hypothyroidism p-values indicate significance after Bonferroni correction for the 61 independent TSH-associated SNPs tested (p < 8.2E-4).

A1 effect allele, A1 frequency of A1, SE standard error of the effect, P association p-value, N percentage of total variation across studies that is due to heterogeneity, N sample size.
marker of autoimmune hypothyroidism, as well as in patients with Graves’ disease (i.e., autoimmune hyperthyroidism) from the BioBank Japan Project. For TPOAb positivity, significant associations were found for MAF, SPATA13, and VAV3 (Supplementary Table 5). SPATA13 and VAV3 have previously been linked to self-reported diagnosed hypothyroidism\(^22\), while no studies have investigated their potential autoimmune origin. The observed associations of these gene variants with variation in TSH levels within the normal range could therefore be due to a mild early stage of thyroid autoimmunity, instead of reflecting physiological differences in the HPT-axis setpoint.

For Graves’ disease, only the psoriasis\(^22\) PSORS1C1 locus showed a significant association, consistent with shared genetic determinants between these two autoimmune diseases\(^23\).

Detailed results of the mouse candidate analysis, results of pathway analyses, and look-ups for pleiotropy of the TSH, FT4, hypo and hyperthyroidism loci are described in Supplementary Note 1–3.

**Gene expression analyses.** To obtain insights into gene expression patterns and potential effector transcripts at the identified loci, we assessed whether the 94 independent index variants from the TSH, FT4, hypo and hyperthyroidism GWAS were correlated with transcript levels of nearby (cis-) or distant (trans-) genes. The results of 22 published expression quantitative trait loci (eQTL) studies were interrogated (Methods), assessing the relation between the genetic variants and gene expression patterns in a total of 127 different tissues and cell types. First, we evaluated the presence of eQTLs in at least one tissue or cell type: 38 variants showed eQTL effects (Fig. 2a and Supplementary Data 3). While many variants were associated with transcript expression in only one or few (<8) tissues, two variants located on chromosome 17 showed ubiquitous associations with gene expression: the FT4-associated variant rs11078333 at NCOR1 locus and the TSH-associated variant rs199461 at the NSF locus. The FT4-increasing allele at rs11078333 was associated with higher expression levels of NCOR1 in blood and brain, but also affected the expression of ADORA2B (increased) and ZSWIM7 and TTC19 (decreased) in many other tissues (including thyroid for TTC19). NCOR1 is an essential nuclear co-repressor that is recruited by thyroid hormone receptors in the absence of thyroid hormone to mediate transcriptional repression. At the NSF locus, the TSH increasing allele at rs199461 increased expression of KANSL1 and LRRC37A2 and decreased expression of WNT3 in several tissues, including thyroid. Consistent with known thyroid physiology, the majority of TSH-associated variants acted as eQTLs in thyroid tissue (Fig. 2a), with 45% of the variants being thyroid-specific eQTLs. In contrast, none of the nine FT4 eQTLs acted exclusively on the thyroid but were also associated with transcript expression changes in multiple known thyroid hormone effector organs, including liver, muscle, and adipose tissue.

Second, we used a summary-based Mendelian randomization (SMR) method coupled with testing for heterogeneity of effects (HEIDI) to assess co-localization, i.e., to investigate whether the overlap between eQTLs and GWAS hits could be attributable to the same underlying causative variant\(^24\). In thyroid tissue, we found evidence for co-localization with differential gene expression at 13 different GWAS loci: PD8EB, PRDM11, MBIP (with RP11-116N8.1 expression), NKX2-3, NSP (with WNT3 expression), IGFB2BP2, FOXA2, SLC25A37, and C9orf92 for TSH; AADAT, NEK6 (with both NEK6 and PSMB7 expressions) for FT4; and TPO, PDE8B, and PDE10A for hyperthyroidism (Supplementary Data 4). At these loci, our findings implicate the causal gene among the many genes present in the locus. For example, the FT4-associated variant rs6854291-influenced transcript levels of AADAT in thyroid while there were no effects on
transcript levels of the neighboring MFAP3L gene, implicating AADAT as the causal gene underlying the FT4 association at this locus (Fig. 2b, Supplementary Data 4). We also observed that independent variants at the same locus were associated with gene expression in different tissues. For example, while the index variant rs6885099 at PDE8B co-localized with changes in PDE8B expression in thyroid, the independent variant rs1119208 was associated with PDE8B expression in pancreas (Supplementary Data 4). Notably, also the variants at AADAT and SLC17A4 co-localized with gene expression in pancreas (Fig. 2c), which is of interest given the close interrelations between thyroid hormone signaling, insulin regulation, and glucose metabolism.25,26.

Fig. 2 Impact on gene expression of index SNP. a shows tissues in which an expression QTL (eQTL) was found in LD ($r^2 > 0.8$) with FT4, hypothyroidism, and TSH index SNPs. SNPs are ordered according to trait they are associated with and then by genomic position; squares are colored according to the LD between the eQTL and the index variant, as depicted in the legend. When multiple eQTLs were detected in the same tissue, the eQTL with the highest LD is shown. b and c illustrate results of the summary-based Mendelian randomization (SMR) test for FT4 levels and expression QTLs at AADAT and SLC17A4 loci, respectively. The upper box shows the regional association curve with FT4 levels, with level of significance of the SMR test (y axis) for each transcript in the locus indicated by a diamond positioned at the center of the transcript. A significant SMR test indicates an association of the transcript level of the respective genes with the trait. The lower box shows the regional association distribution with changes in expression of the highlighted transcript in pancreas. In both boxes, x axis refers to GRCh37/hg19 genomic coordinates.
In vitro studies. Thyroid hormone action in target tissues is importantly determined by the amount of T3 available for receptor binding inside the cell. Therefore, the transport of thyroid hormone across the cell membrane and its metabolism inside the cell represent crucial regulatory layers in thyroid hormone signaling. Although several key players in thyroid hormone signaling have been described over the last decades, including deiodinases and several thyroid hormone transporters, many others remain to be identified. Based on their associations with circulating FT4 levels and the co-localization studies, SLC17A4 and AADAT were further studied in vitro to explore a direct role in thyroid hormone signaling.

SLC17A4 is an organic anion transporter that is particularly expressed in the liver, kidney, and gastrointestinal tract. We transiently over-expressed human SLC17A4 (hSLC17A4) in COS-1 cells and observed increased cellular T3 (Fig. 3a) and T4 (Fig. 3b) accumulation compared to empty-vector transfected control cells. These effects were even stronger upon co-transfection with the intracellular thyroid hormone-binding protein CRYM, after indicated incubation times at 37 °C. All uptake levels are expressed relative to the amount of radio-labeled T3 or T4 added to the cells at the start of the incubation (1 nM (5 × 10^{-4} c.p.m.)) [^{125}I]-T3 or [^{125}I]-T4). All results are presented as means ± SEM (n = 4). In the presence of CRYM, T3 and T4 accumulation in hSLC17A4 transfected cells was significantly higher compared to empty-vector control cells at all time points (one-way ANOVA with a Bonferroni-corrected post hoc test, p < 0.001). T3 (c) and T4 (d) saturation curves in COS-1 cells transiently transfected with hSLC17A4 in the absence of CRYM. All data points are corrected for background thyroid hormone uptake in control cells and presented relatively to the amount of internalized thyroid hormone in the presence of the lowest substrate concentration (0.003 μM for T3 and 0.001 μM for T4, respectively). Apparent IC_{50} values were determined by standard second order polynomial regression analyses implemented in GraphPad Prism (La Jolla, USA).

Together, these findings strongly indicate that SLC17A4 encodes a high-affinity T3 and T4 transporter.

AADAT encodes a mitochondrial amino-transferase with broad substrate specificity, which acts on kynurenine and α-aminoadipate, important intermediates in tryptophan, and lysine metabolism. The association of circulating FT4 with the AADAT locus suggested that AADAT may also be involved in thyroid hormone metabolism. In that case, it could facilitate the oxidative deamination of the alanine side-chain of thyroid hormone, yielding a pyruvic acid moiety. Therefore, lysates of AADAT over-expressing COS-1 cells were incubated with T4 and T3 in the presence of the co-factor pyridoxal phosphate and the co-substrate α-ketoglutaric acid, and the reaction mixtures were analyzed by ultra-performance liquid chromatography (UPLC). The results demonstrated effective time-and AADAT concentration-dependent conversion of T4 and T3 to their pyruvic acid metabolites TK3 and TK4 (Fig. 4), with saturation occurring at substrate concentrations between 10 and 100 μM. Importantly, this is well below the reported Km values of AADAT for α-aminoadipate (0.9 mM) and kynurenine (4.7 mM).

Given the observed effects of SLC17A4 and AADAT on T3 and T4 transport and metabolism, we additionally tested the associations of the identified genetic variants in SLC17A4 and AADAT with T3 levels and the T3/T4 ratio (Supplementary Table 6). SLC17A4-rs9356988 was associated with both the T3/T4 ratio and T3 levels.

Genetic TSH and FT4 risk score associations. To assess the cumulative clinical impact of our GWAS findings, we calculated a weighted GRS for TSH and FT4 levels, which included all
conversion of T3 to TK3 depends on the incubation time (b) concentrations between 10 and 100 µM on the UPLC column and is corrected for the TK3 production in lysates derived from empty-vector-transfected control cells (which was nearly absent). All association with Graves' indicates a polygenic contribution of TSH-associated variants with the limited overlap observed between TSH and FT4 loci.

hyperthyroidism (Supplementary Table 7), which is consistent with a GRS in the lowest quartile had an odds ratio of 2.53 (odds ratio (OR) = 2.53, 95% CI: 1.52, 4.11), whereas a higher FT4 GRS was associated with a higher risk of goiter (OR = 1.52, p = 7.9 × 10^{-3}) and higher thyroid volume (Δvol = 9%, p = 3.8 × 10^{-3}). In addition, a higher TSH GRS was associated with a lower risk of schizophrenia (OR = 0.94, p = 0.01), shorter height (Δheight:beta = -0.05, p = 2.0 × 10^{-11}), and reduced kidney function (ΔeGFR = -1%, p = 1.4 × 10^{-5}), as well as higher LDL (ΔLDL:beta = 0.04, p = 4.9 × 10^{-3}) and total cholesterol levels (Δcholesterol:beta = 0.05, p = 1.1 × 10^{-5}). A higher FT4 GRS was additionally associated with taller height (Δheight:beta = 0.04, p = 2.9 × 10^{-4}), lower BMI (ΔBMI:beta = -0.04, p = 2.7 × 10^{-3}), and lower LDL (ΔLDL:beta = -0.06, p = 3.1 × 10^{-4}) and total cholesterol levels (Δcholesterol:beta = -0.05, p = 6.1 × 10^{-3}) (Supplementary Table 9).

These associations match clinical and epidemiological observations.

Discussion

With 8 million genetic variants tested in up to 72,167 individuals, we present results of the largest GWAS on thyroid function and dysfunction performed so far. We identified 109 significantly and independently associated genetic variants, doubling the number of loci known to regulate thyroid function, which explain a substantial part of the variation in these traits. Importantly, we detected associations between these variants and thyroid diseases as well as various clinical end points, and functionally

Figure 4 AADAT converts T3 and T4 to their respective pyruvic acid metabolites. a shows the conversion of T3 and T4 to their pyruvic acid metabolites TK3 and TK4 in cell lysates of hAADAT over-expressing COS-1 cells. Cell lysates were incubated with [125I]-T3 or [125I]-T4 (2 × 10^5 c.p.m.) in the presence of a range of concentrations between 10 and 100 µM (d). The percentage conversion reflects the amount of TK3 as a percentage of the total radioactivity eluted from the UPLC column and is corrected for the TK3 production in lysates derived from empty-vector-transfected control cells (which was nearly absent). All data are presented as means ± SEM (n = 3).
characterized a new thyroid hormone transporter as well as a new thyroid hormone metabolizing enzyme.

Almost all previously identified TSH and FT4-associated SNPs were also genome-wide significantly associated with the respective trait in our analyses: the 20 TSH and 4 FT4 associations of the sex-combined GWAS of Porcu et al.\(^\text{11}\), the two additional TSH SNPs as well as the FT4 association revealed in Taylor et al.\(^\text{12}\), and 17 of the 21 genome-wide significantly TSH-associated SNPs identified by Gudmundsson et al.\(^\text{13}\). The remaining loci of the latter study were discovered in our study via the mouse candidate analysis and were also associated with hypothryroidism (TPO), associated with FT4, hypo and hyperthyroidism (FOXE1), or had a \(p < 1 \times 10^{-6}\) (FOX1, ELK3, SIVA1). Only two FT4-associated loci, LPCAT2/CAPNS2 and NETO1/FBXO15, identified in the sex-stratified analysis of Porcu et al. did not replicate in our sex-specific GWAS \((p \geq 0.01)\). While all but 2 of the 18 cohorts (the Old Order Amish and the Baltimore longitudinal study on Aging) of Porcu et al. as well as four of the seven cohorts (TwinUK GWAS, SardiNIA, Val Bovera and BHS) of Taylor et al. were also included in our GWAS, we more than doubled the sample size for both traits, thereby significantly increasing our power to detect new loci.

Consistent with HPT-axis physiology, most TSH-associated variants acted on gene expression levels in thyroid tissue, while the FT4-associated variants had more widespread effects on multiple known thyroid hormone target tissues. In addition, four of the newly identified variants were either associated with the risk of TPO Ab positivity or Graves’ disease, suggesting an underlying autoimmune-related pathophysiology. All of these findings confirm that GWAS in the general population provides a valuable method to identify genes implicated in thyroid physiology and/or thyroid disease. Moreover, these insights can be successfully translated into experimental evidence, as illustrated by our in vitro studies on SLC17A4 and AADAT.

To investigate the combined effect of the thyroid hormone associated risk variants, we calculated a GRS. The GRS of TSH-associated variants was significantly associated with the risk of hypothyroidism and hyperthyroidism. For some FT4-associated hits, the lack of association with TSH levels can be explained on physiological grounds. For example, the identified SNP in DIO1 decreases the enzymatic activity of the protein, leading to less T4 to T3 conversion, resulting in higher T4 levels, but lower T3 levels, resulting in no net effect on feedback to the pituitary and therefore no effect on TSH levels. Similar hypotheses could be postulated for other loci involved in thyroid hormone metabolism, such as DIO3OS and AADAT. To assess whether these genetically estimated TSH levels are clinically relevant or merely reflect physiological inter-individual differences in TSH levels (i.e., HPT-axis setpoint), we tested the GRS against various clinical end points. These analyses showed significant associations with thyroid diseases, and also with altered lipid levels (total and LDL cholesterol) and height, which are both known to be affected by hypo and hyperthyroidism. For example, short stature is one of the key characteristics of patients suffering from congenital hypothyroidism. Interestingly, associations were also found with kidney function and schizophrenia, for which the causal relationships are less apparent. Thyroid hormone has been shown to influence kidney development and filtration function\(^\text{37,38}\). Likewise, rodent and human studies have shown that both hypo and hyperthyroidism lead to disrupted prenatal glial cell development, which is an important step in the development of schizophrenia\(^\text{39}\), while various psychiatric diseases including schizophrenia are also thought to influence thyroid function via central effects on the HPT axis\(^\text{40}\). Future studies are needed to clarify the mechanisms underlying these associations, as it is not possible to solve these potential inverse causal relationships solely with GWAS results. Irrespective of the direction of the effects, our results suggest that the presence of kidney dysfunction and psychiatric symptoms in patients with thyroid disease deserve attention. Given the substantial increase in number of TSH and FT4-associated variants, which explain a substantial part of the variation in these traits, future studies should start exploring the use of these markers to predict the individual HPT-axis setpoint. This predicted setpoint could be used to guide treatment of thyroid diseases, which is important as despite normalized TSH and FT4 levels, a substantial part of treated patients still have persistent hypo or hyperthyroid complaints, leading to a lower quality of life\(^\text{41}\). This could be due to the fact that the TSH and FT4 levels are normalized to within the population-based reference ranges, but still deviate from the patient’s individual

---

**Fig. 5** Associations of genetic risk scores with hypothyroidism and hyperthyroidism. The y axis shows the probability of hypothyroidism (red) or hyperthyroidism (blue) with the \(p\)-value of the association test of the trait on the risk score. The x axis shows the percentage of risk alleles carried based on a weighted genetic risk score (GRS) built using the 61 TSH-associated (a) and 31 FT4-associated GWAS SNPs (b). The gray histogram shows the distribution of the GRS in the study sample.
setpoint. For this purpose, a GWAS on the TSH/FT4 ratio could prove to be a more sensitive method to identify more variants, which specifically affect the HPT-axis setpoint.

When interpreting the results of our GWAS studies on increased and decreased TSH levels, it is important to realize that these studies were performed in population-based cohorts, and not in dedicated thyroid disease patient cohorts. Individuals on thyroid medication or a history of thyroid surgery were excluded, resulting in a relative overrepresentation of individuals with subclinical forms of thyroid dysfunction. The identified variants are therefore expected to be a mix of variants, which have been previously associated with hypo or hyperthyroidism (e.g., TPO, FOXE1, and ATFXN2) and variants that lead to a TSH level, which is slightly above or below the population-based reference ranges. These latter effects can either reflect true mild thyroid dysfunction with increased risk of clinical consequences or merely reflect a deviation from the individual HPT-axis setpoint with no clinical consequences. While our GRS analyses suggest that carrying multiple risk alleles leads to an increased risk of overt thyroid dysfunction and related clinical consequences, the exact sequencing of each individual variant needs to be clarified in future studies.

Thyroid hormone is importantly metabolized through enzymatic deiodination by DIO1–3, but also undergoes alternative metabolic reactions, including conjugation with sulfate or glucuronic acid and modification of the alanine side-chain. The latter includes the conversion of T3 and T4 to their respective pyruvic acid metabolites TK3 and TK4, which requires the oxidative deamination of their alanine side-chain49. TK3 and TK4 have been detected in urine and bile of rat injected with radio-labeled T3 and T442. Although these and other studies suggested an important role for the liver and kidney in the formation of these pyruvic acid metabolites, the involved enzyme(s) had not been identified. Our functional analyses demonstrated that AADAT effectively catalyzes the transamination of T4 and in particular T3 to TK4 and TK3, respectively. Moreover, AADAT is highly expressed in the liver, gastrointestinal tract, and kidney in human43. Taken together, AADAT activity may thus be critical for the rate of thyroid hormone metabolism, which likely underlies the association of AADAT with circulating FT4. Although the specific impact of the associated variant on AADAT expression has not been assessed yet, our eQTL co-localization studies indicate that the index SNP decreases AADAT transcript levels in the thyroid, and this in turn leads to increased circulating FT4 levels.

The functional analyses further demonstrate that human SLC17A4 is able to transport both T4 and T3. The protein belongs to the solute carrier 17 family, whose members transport various organic anions, such as p-aminomimidic acid. Genetic variation in the SLC17A4 locus has been associated with the progression of elevated serum urate levels to gout27,44. According to the GTEx data resource and previously reported studies27, SLC17A4 is predominantly expressed in human small intestinal and colonic epithelial cells, pancreas, liver, and kidney cortex, which could imply a role for this transporter in the metabolic clearance and entero-hepatic cycle of thyroid hormone. Future studies should investigate the pharmacokinetic properties of SLC17A4, its relative contributions to thyroid hormone transport in various individual tissues, as well as the effects of the identified SLC17A4 variants on thyroid hormone transport.

The findings from our functional studies do not only provide new insights into thyroid hormone physiology, but may also have important clinical implications. Hypothyroidism is treated with levothyroxine (LT4), which is inexpensive and administered orally. In recent decades, various factors have been identified which help to determine LT4 dose, such as weight, gastrointestinal diseases, and interfering drugs45,46. Despite this knowledge, ineffective LT4 supplementation is still a major clinical problem, as 30–50% of patients are either under- or over-treated and therefore remain at risk for the symptoms and complications associated with thyroid dysfunction45,46. Therefore, the identification of SLC17A4 as a thyroid hormone transporter and AADAT as a thyroid hormone metabolizing enzyme provides new insights into thyroid hormone physiology and opens up a potential avenue for novel therapeutic targets or optimization of existing ones to improve the care of patients suffering from thyroid dysfunction. All genetic findings in our study were limited to common or low-frequency SNPs, whereas rare SNPs or structural variants may also contribute to the yet unexplained variance of thyroid function. Large whole-exome or -genome sequencing studies are required to reveal these rare variant associations47. Furthermore, additional GWAS with increased sample size will help to reveal the yet undiscovered associations of common and low-frequency SNPs. Our ThyroidOmnics Consortium (http://www.thyroidomics.com) provides a well-established infrastructure to address these knowledge gaps in future projects.

**Methods**

**Inclusion studies.** Discovery meta-analyses included data from 22 independent cohorts with 54,288 subjects for the TSH analyses, and from 19 cohorts with 49,269 subjects for FT4, 53,423 subjects (3440 cases) for hyperthyroidism, and 51,823 subjects (1840 cases) for hypothyroidism (Supplementary Data 1). Selected SNPs from the TSH or FT4 analyses were carried forward for replication with in-silico GWAS data from 5 cohorts (9053 subjects) and de novo genotyping in additional 5 cohorts (13,330 subjects). All subjects gave informed consent and studies were approved by the cohort-specific ethics committees.

We used the results of the GWAS of TPOab positivity that included 18,297 subjects48 for a look-up of all the TSH-associated loci or their HapMapII proxies (r² > 0.8 in a 1 Mb window) that were available in that dataset to assess their relation to autoimmune hypothyroidism. A complementary look-up was performed for the 52 SNPs that were available in a GWAS on Graves’ disease diagnosed by clinical examinations, circulating thyroid hormone and TSH concentrations, serum levels of antibodies against thyroglobulin, thyroid microsomes, and TSH receptors, ultrasonography, [99mTc] TCO4 (technetium-99m pertechnetate) (or [ 123I] (radioactive iodine)) uptake and thyroid scintigraphy using the data of the BioBank Japan Project (BBI) including 1747 patients and 6420 controls (Supplementary Data 1).

**Trait definition.** In each study, only subjects with TSH levels within the cohort-specific reference range were included for the TSH and FT4 analyses. TSH and FT4 were analyzed as continuous variables after inverse normal transformation.

Increased TSH was defined by a level above the upper limit of the cohort-specific TSH reference range, while decreased TSH was defined by a level below the lower limit of the reference range. For both increased and decreased TSH analyses, the comparison group consisted of subject with a TSH level within the cohort-specific reference range. Exclusion criteria for all analyses were non-European ancestry, use of thyroid medication (defined as ATC (Anatomical Therapeutic Chemical) code H03), or previous thyroid surgery.

**GWAS in individual studies.** In each study of the discovery GWAS, genotyping was performed on genome-wide arrays. Genome-wide data were imputed to the 1000 Genomes, phase 1 version 3 (March 2012) ALL populations reference panel, including the X chromosome. Quality control before imputation was applied in each study separately. Details on study-specific genotyping and imputation information are provided in the Supplementary Data 1. In the individual study GWAS, the association of the SNPs was analyzed using linear regression for TSH and FT4, and logistic regression for decreased and increased TSH. The genotype–phenotype association was conducted using an additive genetic model on SNP dosages, thus taking genotype uncertainties of imputed SNPs into account. The analyses for TSH and FT4 were initially sex-stratified and meta-analyzed as a second step. The analyses were adjusted for age, age-squared (to account for non-linear effects), and relevant study-specific covariates such as principal components for population stratification, study center, and other drug exposure (e.g., by inclusion of the kinship matrix as a random effect), if applicable. The family-based cohorts GARP, Sardinia, ValBorbera, MICROS, TwinsUK, LL, and FHS conducted additional analyses on the men and women–combined sample, with additional adjustment for sex, to properly account for their family relatedness.

**Statistical methods for meta-analysis.** Result files from individual studies included in this analysis underwent extensive quality control before meta-analysis: file format checks as well as plausibility and distributions of association results including effects, standard errors, allele frequencies, and imputation quality of the
SNPs were obtained by using the gwasqc() function of the GWAtoolbox package v2.2.4.22. Additionally, the known associations of rs6885099 in PDE6B with TSH and rs2235544 in DIO1 were checked for consistent direction and size in each study. All cohort-specific genomic control values (λg) ranged from 0.94 to 1.14 (median 1.00) for the continuous trait and from 0.68 to 1.04 (median 0.91) for the dichotomous trait GWAS.

All meta-analyses were carried out in duplicate by three independent analysts. We conducted a genome-wide meta-analysis applying inverse-variance weighting as implemented in Metaf36. SNPs with MAF ≤ 0.05 or an imputation quality score ≤ 0.4 were excluded prior to the meta-analyses resulting in a median of 9,653,808 SNPs per cohort (IQR: 9,302,604–10,705,092). During the meta-analysis, the study-specific results were corrected by their specific λg if > 1. Results were checked for possible errors like use of incorrect unit, trait transformation, or association model by plotting the association p-values of the analyses against those from a z-score-based meta-analysis for verifying overall concordance. SNPs that were present in < 75% of the total sample size contributing to the respective meta-analysis (separately for autosomal and X-chromosomal SNPs) or with a MAF ≤ 0.01 (hypo- and hyperthyroidism) were excluded (because of the low number of cases in the analysis) were excluded from subsequent analyses. Finally, data for up to 8,048,941 genotyped or imputed autosomal and X-chromosomal SNPs were available after the discovery stage meta-analysis of TSH, FT4, and up to 5,965,951 SNPs after hypo- and hyperthyroidism.

Quantile–quantile plots of the meta-analysis results are provided in Supplementary Figures 10 and 11. To assess whether there was an inflation of p-values in the meta-analysis results attributed to reasons other than polygenicity, we performed LD score regression40. The LD score corrected λgc values of the meta-analysis results ranged from 1.00 to 1.04, supporting the absence of unaccounted population stratification. Genomic-wide significance was defined as a p-value of < 5 × 10−8, corresponding to a Bonferroni correction of one million independent tests. Unless stated otherwise, all reported p-values are two-sided. The I2 statistic was used to evaluate between-study heterogeneity41.

Gene-by-sex interaction on the circulating TSH and FT4 levels were obtained for each SNP by the discovery meta-analysis results from men (TSH: n = 24,618; FT4: n = 22,315) and women using a t-test. Test statistics were calculated using the formula t = (βmax − βmean)/sqrt(SEβmax + SEβmean), assuming independent effect sizes between men and women. To evaluate the presence of independent SNPs within each locus, SNPs were clustered based on their correlation with the SNP showing the lowest p-value at that locus (index SNP) using the software PLINK32 and the genotypes of the combined individuals of the 1000Genomes phase3 all ethinities reference panel (linkage disequilibrium pruning using r2 ≤ 0.01 within windows of ±1 Mb). The loci were named according to the nearest gene of the index SNP. Genomic positions correspond to build 37 (GRCh37).

Replication analysis. The genome-wide significant index SNPs of newly identified loci from the sex-combined TSH (n = 22) and FT4 meta-analyses (n = 19) were taken forward to the replication stage (Supplementary Table 10). When SNPs were not available in the in silico replication datasets, a proxy SNP in LD with a significant index SNP was taken forward to the replication stage (Supplementary Table 10). When SNPs were obtained by using the gwasqc() function of the GWAtoolbox package v2.2.4.22. Additionally, the known associations of rs6885099 in PDE6B with TSH and rs2235544 in DIO1 were checked for consistent direction and size in each study. All cohort-specific genomic control values (λg) ranged from 0.94 to 1.14 (median 1.00) for the continuous trait and from 0.68 to 1.04 (median 0.91) for the dichotomous trait GWAS.

All meta-analyses were carried out in duplicate by three independent analysts. We conducted a genome-wide meta-analysis applying inverse-variance weighting as implemented in Metaf36. SNPs with MAF ≤ 0.05 or an imputation quality score ≤ 0.4 were excluded prior to the meta-analyses resulting in a median of 9,653,808 SNPs per cohort (IQR: 9,302,604–10,705,092). During the meta-analysis, the study-specific results were corrected by their specific λg if > 1. Results were checked for possible errors like use of incorrect unit, trait transformation, or association model by plotting the association p-values of the analyses against those from a z-score-based meta-analysis for verifying overall concordance. SNPs that were present in < 75% of the total sample size contributing to the respective meta-analysis (separately for autosomal and X-chromosomal SNPs) or with a MAF ≤ 0.01 (hypo- and hyperthyroidism) were excluded (because of the low number of cases in the analysis) were excluded from subsequent analyses. Finally, data for up to 8,048,941 genotyped or imputed autosomal and X-chromosomal SNPs were available after the discovery stage meta-analysis of TSH, FT4, and up to 5,965,951 SNPs after hypo- and hyperthyroidism.

Quantile–quantile plots of the meta-analysis results are provided in Supplementary Figures 10 and 11. To assess whether there was an inflation of p-values in the meta-analysis results attributed to reasons other than polygenicity, we performed LD score regression40. The LD score corrected λgc values of the meta-analysis results ranged from 1.00 to 1.04, supporting the absence of unaccounted population stratification. Genomic-wide significance was defined as a p-value of < 5 × 10−8, corresponding to a Bonferroni correction of one million independent tests. Unless stated otherwise, all reported p-values are two-sided. The I2 statistic was used to evaluate between-study heterogeneity41.

Gene-by-sex interaction on the circulating TSH and FT4 levels were obtained for each SNP by the discovery meta-analysis results from men (TSH: n = 24,618; FT4: n = 22,315) and women using a t-test. Test statistics were calculated using the formula t = (βmax − βmean)/sqrt(SEβmax + SEβmean), assuming independent effect sizes between men and women. To evaluate the presence of independent SNPs within each locus, SNPs were clustered based on their correlation with the SNP showing the lowest p-value at that locus (index SNP) using the software PLINK32 and the genotypes of the combined individuals of the 1000Genomes phase3 all ethinities reference panel (linkage disequilibrium pruning using r2 ≤ 0.01 within windows of ±1 Mb). The loci were named according to the nearest gene of the index SNP. Genomic positions correspond to build 37 (GRCh37).

Replication analysis. The genome-wide significant index SNPs of newly identified loci from the sex-combined TSH (n = 22) and FT4 meta-analyses (n = 19) were taken forward to the replication stage (Supplementary Table 10). When SNPs were not available in the in silico replication datasets, a proxy SNP in LD with a significant index SNP was taken forward to the replication stage (Supplementary Table 10). When SNPs were obtained by using the gwasqc() function of the GWAtoolbox package v2.2.4.22. Additionally, the known associations of rs6885099 in PDE6B with TSH and rs2235544 in DIO1 were checked for consistent direction and size in each study. All cohort-specific genomic control values (λg) ranged from 0.94 to 1.14 (median 1.00) for the continuous trait and from 0.68 to 1.04 (median 0.91) for the dichotomous trait GWAS.

All meta-analyses were carried out in duplicate by three independent analysts. We conducted a genome-wide meta-analysis applying inverse-variance weighting as implemented in Metaf36. SNPs with MAF ≤ 0.05 or an imputation quality score ≤ 0.4 were excluded prior to the meta-analyses resulting in a median of 9,653,808 SNPs per cohort (IQR: 9,302,604–10,705,092). During the meta-analysis, the study-specific results were corrected by their specific λg if > 1. Results were checked for possible errors like use of incorrect unit, trait transformation, or association model by plotting the association p-values of the analyses against those from a z-score-based meta-analysis for verifying overall concordance. SNPs that were present in < 75% of the total sample size contributing to the respective meta-analysis (separately for autosomal and X-chromosomal SNPs) or with a MAF ≤ 0.01 (hypo- and hyperthyroidism) were excluded (because of the low number of cases in the analysis) were excluded from subsequent analyses. Finally, data for up to 8,048,941 genotyped or imputed autosomal and X-chromosomal SNPs were available after the discovery stage meta-analysis of TSH, FT4, and up to 5,965,951 SNPs after hypo- and hyperthyroidism.

Quantile–quantile plots of the meta-analysis results are provided in Supplementary Figures 10 and 11. To assess whether there was an inflation of p-values in the meta-analysis results attributed to reasons other than polygenicity, we performed LD score regression40. The LD score corrected λgc values of the meta-analysis results ranged from 1.00 to 1.04, supporting the absence of unaccounted population stratification. Genomic-wide significance was defined as a p-value of < 5 × 10−8, corresponding to a Bonferroni correction of one million independent tests. Unless stated otherwise, all reported p-values are two-sided. The I2 statistic was used to evaluate between-study heterogeneity41.

Gene-by-sex interaction on the circulating TSH and FT4 levels were obtained for each SNP by the discovery meta-analysis results from men (TSH: n = 24,618; FT4: n = 22,315) and women using a t-test. Test statistics were calculated using the formula t = (βmax − βmean)/sqrt(SEβmax + SEβmean), assuming independent effect sizes between men and women. To evaluate the presence of independent SNPs within each locus, SNPs were clustered based on their correlation with the SNP showing the lowest p-value at that locus (index SNP) using the software PLINK32 and the genotypes of the combined individuals of the 1000Genomes phase3 all ethinities reference panel (linkage disequilibrium pruning using r2 ≤ 0.01 within windows of ±1 Mb). The loci were named according to the nearest gene of the index SNP. Genomic positions correspond to build 37 (GRCh37).

Replication analysis. The genome-wide significant index SNPs of newly identified loci from the sex-combined TSH (n = 22) and FT4 meta-analyses (n = 19) were taken forward to the replication stage (Supplementary Table 10). When SNPs were not available in the in silico replication datasets, a proxy SNP in LD with a significant index SNP was taken forward to the replication stage (Supplementary Table 10). When SNPs were obtained by using the gwasqc() function of the GWAtoolbox package v2.2.4.22. Additionally, the known associations of rs6885099 in PDE6B with TSH and rs2235544 in DIO1 were checked for consistent direction and size in each study. All cohort-specific genomic control values (λg) ranged from 0.94 to 1.14 (median 1.00) for the continuous trait and from 0.68 to 1.04 (median 0.91) for the dichotomous trait GWAS.
incubating 0.1-100 μM of T3 or T4 in the presence of 2 × 10^5 c.p.m. 125I-labeled TSH and FT4, respectively. In case the trait was logarithm-transformed in the association effects correspond to a change in the look-up GWAS trait (or natural log). The probability of disease was calculated using the formula 1/(1 + exp(−(β0 + β1x))). Where β0 and β1 correspond to the intercept and GWAS-related effect in the regression model, respectively.

Pathway analyses. We performed Data-Driven Expression Prioritized Integration for Complex Traits (DEPICT) analyses to prioritize the most likely causal genes at the associated loci and identify enriched pathways and tissues. We used DEPICT version 1 r6194 and included variants with GC-corrected p-value < 1 × 10^-5 from discovery GWAS as input. The following parameters were applied in the DEPICT analyses: 50 repetitions used to compute the false discovery rate, 500 permutations used to adjust for biases such as gene length, and 500 null permutations used to run repetitions and permutations.

Genes for network analysis were selected using the associated genes from DEPICT using the lead SNPs from the analyses of the discovery GWAS. The Ingenuity Pathway Analysis Software Tool (IPA; Ingenuity® Systems, CA, USA) Network was used in order to perform pathway analysis (core analysis). Molecules and/or relationships considered were the ones available in the IPA Knowledge Base for mammals. Confidence filters considered only relationships, direct and indirect, where the confidence is experimentally observed or high (predicted). Networks were generated with a maximum size of 35 genes and allowing up to 25 networks per analysis. We did not restrict for tissue and cell lines or mutations. The networks are constructed using the IPA algorithm with the Ingenuity Knowledge Base as a reference set generating a score as well as a p-value. IPA computes a score for each network according to the fit of that network to the user-defined set of focus genes. The score indicates the likelihood of the focus genes in a network being found together due to random chance. The significance of p-value is calculated using the right-tailed Fisher exact test.

Look-up of pleiotropy. The look-up of additional traits associated with the replicated GWAS findings were performed using the PhenoScanner[33] database with the default search options, whereas the independent SNPs of the replicated TSH, FT4, hypo, and hyperthyroidism-associated loci or their proxies (r^2 > 0.8) obtained by SNPs[34] were used as input. Look-up results with genome-wide significant p-value (5 × 10^-8) were reported.

Data availability


Received: 20 February 2018 Accepted: 31 August 2018
Published online: 26 October 2018

References


1Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. 2DZHK (German Center for Cardiovascular Research), Partner Site Greifswald, Greifswald, Germany. 3Erasmus MC Academic Center for Thyroid Diseases, Rotterdam, The Netherlands. 4Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. 5Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. 6Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Milan, Italy. 7 amer Ducato Diabetes Center Copenhagen, Gentofte, Denmark. 8Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK. 9Department of Epidemiology, University of Edinburgh, Edinburgh, UK. 10Department of Psychology, University of Edinburgh, Edinburgh, UK. 11Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. 12McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA. 13Department of Biostatistics, University of Washington, Seattle, WA, USA. 14Department of Nutrition, Exercise, and Sports, Faculty of Science, University of Copenhagen, Copenhagen, Denmark. 15Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands. 16Pathwest Laboratory Medicine WA, Nedlands, WA 6009, Australia. 17School of Biomedical Sciences, University of Western Australia, Crawley, WA 6009, Australia. 18Bimeta, Clinical Research Center Ghent, Ghent University Hospital, Ghent, Belgium. 19Human Genetics Center, University of Texas Health Science Center, 1200 Pressler Street, Houston, TX 77030, USA. 20Department of Endocrinology and Diabetes, Sir Charles Gardiner Hospital, Nedlands, WA 6005, Australia. 21Department of Cardiology, Ghent University Hospital, Ghent, Belgium. 22Department of Medicine and Surgery, University of Parma, University Hospital of Parma, Parma, Italy. 23Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, The Capital Region, Copenhagen, Denmark. 24Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche Monserrato, Monserrato, Italy. 25Dipartimento di Scienze Biomediche, Università degli Studi di Sassari, Sassari, Italy. 26Centre for Cognitive Ageing and Cogntive Epidemiology, University of Edinburgh, Edinburgh, UK. 27Department of Psychology, University of Edinburgh, Edinburgh, UK. 28Max Planck Institute for Biology of Ageing, Cologne, Germany. 29Department of Nephrology and Hypertension, University of Erlangen-Nürnberg, Erlangen, Germany. 30Department of Nephrology and Medical Intensive Care, Charité – Universitätsmedizin Berlin, Berlin, Germany. 31 Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany. 32Department of General Practice and Primary Health Care, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. 33Folkhälsoan Research Center, Helsinki, Finland. 34National Institute for Health and Welfare, Helsinki, Finland. 35Translational Gerontology Branch, National Institute on Aging, Baltimore, MD, USA. 36Department of Clinical Chemistry, Ghent University Hospital, Ghent, Belgium. 37Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK. 38National...
Heart, Lung, and Blood Institute’s Framingham Heart Study and the Center for Population Studies, Framingham, MA, USA. 39Division of Endocrinology, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA. 40Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lubeck, Boziano, Italy. 41Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands. 42Helmholtz Zentrum München, Institute of Epidemiology, German Research Center for Environmental Health, Neuherberg, Germany. 43Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany. 44German Research Centre Division of Cancer Epidemiology, Heidelberg, Germany. 45Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland. 46Medical Genetics Section, University of Edinburgh Centre for Genomic and Experimental Medicine and MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK. 47Leiden University Medical Center, Geriatrics and Gerontology, Leiden, The Netherlands. 48Vrije Universiteit Medisch Centrum, Amsterdam, The Netherlands. 49Department of Biomedical Data Sciences, Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. 50Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany. 51Medical School, University of Western Australia, Crawley, WA 6009, Australia. 52Department of Medicine, Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA. 53Faculty of Biology, University of Freiburg, Freiburg, Germany. 54Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands. 55Einhoven Laboratory for Experimental Vascular Medicine, LUMC, Leiden, The Netherlands. 56Children’s Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, Finland. 57PEDEGO Research Unit, MRC Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland. 58Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan. 59Institute of Epidemiology, Friedrich-Löffler-Institut (FLI), Federal Research Institute for Animal Health, Greifswald, Germany. 60Department of Endocrinology, Ghent University Hospital, Ghent, Belgium. 61Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands. 62Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. 63Department of Medical Genetics, Division of Genetic Epidemiology, Molecular and Clinical Pharmacology, Medical University of Innsbruck, Innsbruck, Austria. 64Laboratory for Genotyping Development, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. 65Department of Biostatistics, Boston University, Boston, MA, USA. 66Department of Clinical Experimental Research, Rigshospitalet, Copenhagen, Denmark. 67Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan. 68Institute for Medical Epidemiology, Biostatistics and Informatics, Martin-Luther-University Halle-Wittenberg, Halle, Germany. 69Chair of Epidemiology, Ludwig-Maximilians-Universität München, at UNIKA-T Augsburg, Augsburg, Germany. 70Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium. 71Department of Pharmaceutical Sciences, University of Basel, Basel, Switzerland. 72Department of Internal Medicine, Division of Endocrinology, Radboud University Medical Center, Nijmegen, The Netherlands. 73Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 74Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka, Japan. 75Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita, Japan. 76Institute of Clinical Chemistry and Laboratory Medicine, Medicine Greifswald, Greifswald, Germany. 77Center for Integrative Genomics, University of Lausanne, and Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. 78Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, Netherlands. 79Cardiovascular Health Research Unit, Departments of Medicine and Health Services, University of Washington, Seattle, WA, USA. 80Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA. 81Lady Davis Institute, Jewish General Hospital, Montreal, QC H3T 1E2, Canada. 82The Department of Twin Research and Genetic Epidemiology, King’s College London, St. Thomas’ Campus, Lambeth Palace Road, London SE1 7EH, UK. 83Biobanking and Cardiovascular Epidemiology, Ghent University Hospital, Ghent, Belgium. 84Department of Internal Medicine (Cardiology), Ghent University, Ghent, Belgium. 85The Institute for Translational Genomics and Population Sciences, Departments of Pediatrics and Medicine, LAMBioMed at Harbor-UCLA Medical Center, Torrance, CA, USA. 86Laboratory of Genetics and Genomics, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA. 87Department of Epidemiology, Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. 88The Welcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1HH, UK. 89Department of Public Health (Section on Epidemiology), Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. 90Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, UK. 91Institute of Cardiovascular and Medical Sciences, Faculty of Medicine, University of Glasgow, Glasgow, UK. 92AZ Sint Jan Brugge, Brugge, Belgium. 93Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 94Diseases & Population (DaP) Geninfo Lab, School of Life Sciences, Westlake University and Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China. 95Institute of Aging Research and the Affiliated Hospital, School of Medicine, Hangzhou Normal University, Hangzhou, Zhejiang, China. 96Division of Endocrinology, Diabetes, and Metabolism, University of Pennsylvania School of Medicine, Philadelphia, PA, USA. 97Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 98Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 625 N Wolfe Street, Baltimore, MD 21205, USA. These authors contributed equally: Alexander Teumer, Layal Chaker, Stefan Groeneweg, Yong Li, Celia Di Munno, Caterina Barbieri. These authors jointly supervised this work: Henry Völzke, Karen Raikow, Anna Köttgen, Theo J. Visser, Marco Medici. Deceased: Theo J. Visser. A full list of consortium members appears below.

Lifelines Cohort Study

Behrooz Z. Alizadeh99, H. Marike Boezen99, Lude Franke100, Tim van der Harst101, Gerjan Navis102, Marianne Rots103, Harold Snieder99, Morris A. Swertz100 & Ciska Wijmenga100

99Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 100Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 101Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 102Department of Internal Medicine, Division of Nephrology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 103Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.