



Universiteit
Leiden
The Netherlands

Epigenome-wide association study reveals CpG sites associated with thyroid function and regulatory effects on KLF9

Weihs, A.; Chaker, L.; Martin, T.C.; Braun, K.V.E.; Campbell, P.J.; Cox, S.R.; ... ; Teumer, A.

Citation

Weihs, A., Chaker, L., Martin, T. C., Braun, K. V. E., Campbell, P. J., Cox, S. R., ... Teumer, A. (2023). Epigenome-wide association study reveals CpG sites associated with thyroid function and regulatory effects on KLF9. *Thyroid*, 33(3), 301-311. doi:10.1089/thy.2022.0373

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](#)

Downloaded from: <https://hdl.handle.net/1887/3754058>

Note: To cite this publication please use the final published version (if applicable).

Open camera or QR reader and
scan code to access this article
and other resources online.



Epigenome-Wide Association Study Reveals CpG Sites Associated with Thyroid Function and Regulatory Effects on *KLF9*

Antoine Weihs,^{1,*} Loyal Chaker,^{2,3,*} Tiphaine C. Martin,^{4,5,*} Kim V.E. Braun,⁶ Purdey J. Campbell,⁷
Simon R. Cox,⁸ Myriam Fornage,^{9,10} Christian Gieger,^{11–13} Hans J. Grabe,^{1,14} Harald Grallert,^{11,12}
Sarah E. Harris,⁸ Brigitte Kühnel,^{11,12} Riccardo E. Marioni,¹⁵ Nicholas G. Martin,¹⁶ Daniel L. McCartney,¹⁵
Allan F. McRae,¹⁷ Christa Meisinger,¹⁸ Joyce B.J. van Meurs,^{19,20} Jana Nano,^{12,21} Matthias Nauck,^{22,23}
Annette Peters,^{11–13,21} Holger Prokisch,^{24,25} Michael Roden,^{26–28} Elizabeth Selvin,^{29,30} Marian Beekman,³¹
Diana van Heemst,³² Eline P. Slagboom,³¹ Brenton R. Swenson,³³ Adrienne Tin,^{29,34} Pei-Chien Tsai,^{5,35,36}
Andre Uitterlinden,¹⁹ W. Edward Visser,² Henry Völzke,^{23,37} Melanie Waldenberger,^{11–13} John P. Walsh,^{7,38}
Anna Köttgen,^{29,39} Scott G. Wilson,^{5,7,40} Robin P. Peeters,² Jordana T. Bell,^{5,†}
Marco Medici,^{2,†} and Alexander Teumer^{1,23,37,41,†}

¹Department of Psychiatry and Psychotherapy; ²²Institute of Clinical Chemistry and Laboratory Medicine; ³⁷Institute for Community Medicine; University Medicine Greifswald, Greifswald, Germany.

²Erasmus MC Academic Center for Thyroid Diseases, Department of Internal Medicine and ³Erasmus MC Academic Center for Thyroid Diseases, Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands.

⁴Department of Oncological Sciences, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA.

⁵Department of Twin Research and Genetic Epidemiology, St Thomas' Hospital Campus, King's College London, London, United Kingdom.

Departments of ⁶Epidemiology, ¹⁹Internal Medicine, and ²⁰Orthopedics and Sports Medicine, Erasmus Medical Center, Rotterdam, The Netherlands.

⁷Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, Nedlands, Australia.

⁸Lothian Birth Cohorts, Department of Psychology; ¹⁵Centre for Genomic and Experimental Medicine, Institute of Genetics and Cancer; University of Edinburgh, Edinburgh, United Kingdom.

⁹Brown Foundation Institute of Molecular Medicine, McGovern Medical School, Houston, Texas, USA.

¹⁰Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas, USA.

¹¹Research Unit Molecular Epidemiology; ¹²Institute of Epidemiology; ²⁴Institute of Neurogenetics, Computational Health Center; Helmholtz Munich, Neuherberg, Germany.

¹³German Center for Cardiovascular Research (DZHK), Partner Site Munich Heart Alliance, Munich, Germany.

¹⁴German Centre for Neurodegenerative Diseases (DZNE), Site Rostock, Greifswald, Germany.

¹⁶QIMR Berghofer Medical Research Institute, Brisbane, Australia.

¹⁷Institute for Molecular Bioscience, The University of Queensland, St Lucia, Australia.

¹⁸Epidemiology, Medical Faculty, University of Augsburg, Augsburg, Germany.

²¹Institute for Medical Informatics, Biometrics and Epidemiology, Ludwig-Maximilians-Universität (LMU) Munich, Munich, Germany.

²³DZHK (German Centre for Cardiovascular Research), Partner Site Greifswald, Greifswald, Germany.

²⁵Institute of Human Genetics, School of Medicine, Technical University Munich, Munich, Germany.

²⁶Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research; ²⁷Division of Endocrinology and Diabetology, Medical Faculty; Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

²⁸German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany.

²⁹Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.

³⁰Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins School of Medicine, Baltimore, Maryland, USA.

³¹Section of Molecular Epidemiology, Department of Biomedical Data Sciences; ³²Section of Gerontology and Geriatrics, Department of Internal Medicine; Leiden University Medical Center, Leiden, Netherlands.

³³Cardiovascular Health Research Unit, University of Washington, Seattle, Washington, USA.

³⁴Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi, USA.

³⁵Department of Biomedical Sciences, Chang Gung University, Taoyuan, Taiwan.

³⁶Division of Pediatric Infectious Diseases, Department of Pediatrics, Chang Gung Memorial Hospital, Taoyuan City, Taiwan.

³⁸Medical School, University of Western Australia, Crawley, Australia.

³⁹Institute of Genetic Epidemiology, Faculty of Medicine and Medical Center—University of Freiburg, Freiburg, Germany.

⁴⁰School of Biomedical Sciences, University of Western Australia, Perth, Australia.

⁴¹Department of Population Medicine and Lifestyle Diseases Prevention, Medical University of Białystok, Białystok, Poland.

*These authors contributed equally to this work.

†These authors jointly supervised this work.

Background: Thyroid hormones play a key role in differentiation and metabolism and are known regulators of gene expression through both genomic and epigenetic processes including DNA methylation. The aim of this study was to examine associations between thyroid hormones and DNA methylation.

Methods: We carried out a fixed-effect meta-analysis of epigenome-wide association study (EWAS) of blood DNA methylation sites from 8 cohorts from the ThyroidOmics Consortium, incorporating up to 7073 participants of both European and African ancestry, implementing a discovery and replication stage. Statistical analyses were conducted using normalized beta CpG values as dependent and log-transformed thyrotropin (TSH), free thyroxine, and free triiodothyronine levels, respectively, as independent variable in a linear model. The replicated findings were correlated with gene expression levels in whole blood and tested for causal influence of TSH and free thyroxine by two-sample Mendelian randomization (MR).

Results: Epigenome-wide significant associations (p -value $< 1.1 \times 10^{-7}$) of three CpGs for free thyroxine, five for free triiodothyronine, and two for TSH concentrations were discovered and replicated (combined p -values = 1.5×10^{-9} to 4.3×10^{-28}). The associations included CpG sites annotated to *KLF9* (cg00049440) and *DOTIL* (cg04173586) that overlap with all three traits, consistent with hypothalamic–pituitary–thyroid axis physiology. Significant associations were also found for CpGs in *FKBP5* for free thyroxine, and at *CSNK1D/LINC01970* and *LRRC8D* for free triiodothyronine. MR analyses supported a causal effect of thyroid status on DNA methylation of *KLF9*. DNA methylation of cg00049440 in *KLF9* was inversely correlated with *KLF9* gene expression in blood. The CpG at *CSNK1D/LINC01970* overlapped with thyroid hormone receptor alpha binding peaks in liver cells. The total additive heritability of the methylation levels of the six significant CpG sites was between 25% and 57%. Significant methylation QTLs were identified for CpGs at *KLF9*, *FKBP5*, *LRRC8D*, and *CSNK1D/LINC01970*.

Conclusions: We report novel associations between TSH, thyroid hormones, and blood-based DNA methylation. This study advances our understanding of thyroid hormone action particularly related to *KLF9* and serves as a proof-of-concept that integrations of EWAS with other -omics data can provide a valuable tool for unraveling thyroid hormone signaling in humans by complementing and feeding classical *in vitro* and animal studies.

Keywords: thyroid function, DNA methylation, *KLF9*, Mendelian randomization, gene expression

Introduction

THYROID HORMONES PLAY a key role in differentiation and metabolism. Thyroid dysfunction, a condition affecting 5–10% of the adult population, is associated with an increased risk of weight changes, cardiovascular diseases, osteoporosis, psychiatric disorders, and mortality.^{1,2} Prohormone thyroxine and the biologically active triiodothyronine are the main circulating thyroid hormones and are regulated by thyrotropin (TSH). Thyroid hormones are known regulators of gene expression through both genomic and epigenetic processes;³ however, the underlying exact molecular mechanisms remain unknown. One of these processes includes DNA methylation (DNAm), which predominantly occurs at CpG sites, and is a key regulator of gene expression.

While there has been little research so far in humans, various animal models provided evidence of thyroid hormones influencing DNAm. This was supported, for example, in a study by Kyono et al,⁴ showing in developing frog tissue that triiodothyronine directly controls DNA methyltransferase 3a (*Dnmt3a*) expression, responsible for *de novo* DNAm. A recent study of two cohorts of Australian adolescents further revealed two and six CpG sites that were associated with TSH and free triiodothyronine (fT3) concentrations, respectively, indicating that although DNAm is highly tissue specific, the analysis of DNAm in blood may provide deeper insights into thyroid hormone action and/or regulation.⁵

Motivated by the conclusion that larger sample sizes are necessary to replicate the findings and to reveal additional associations, we conducted a large-scale epigenome-wide association study (EWAS) assessing the effects of free thy-

roxine (fT4), fT3, and TSH on changes of DNAm in whole blood encompassing up to 7073 individuals of the ThyroidOmics Consortium (www.thyroidomics.com). We furthermore replicated findings in independent samples, improved the generalizability of the results by including individuals of both European and African ancestry, and provided insights into the underlying molecular mechanisms by analyzing gene expression and assessing causality using Mendelian randomization (MR).

Materials and Methods

Study population

We conducted the EWAS using population-based studies from the ThyroidOmics Consortium. The discovery stage included three cohorts: SHIP-Trend,⁶ ARIC,⁷ and KORA.⁸ Replication of the findings was sought in the Rotterdam Study (RS),⁹ TwinsUK,¹⁰ the Lothian Birth Cohorts of 1921 and 1936 (LBC1921 and LBC1936),¹¹ and the Brisbane Systems Genetics Study (BSGS).¹² All study protocols were conducted in accordance with the Declaration of Helsinki and approved by the respective local ethics committees. All subjects provided written informed consent. Detailed information is given in Tables 1 and 2 and in the Supplementary Data.

Biomarker measurements

DNAm was measured from whole blood using Illumina Infinium BeadChip arrays. Details regarding the cohorts, as well as the fT4, fT3, and TSH assays applied, are provided in Table 2.

TABLE 1. BASELINE STATISTICS OF THE COHORTS INCLUDED IN THE EPIGENOME-WIDE ASSOCIATION STUDY ANALYSIS

	SHIP-Trend	ARIC	KORA	RS	TWINSUK	LBC1936	LBC1921	BSGS
Ethnicity	Discovery European ancestry	Discovery African ancestry	Discovery European ancestry	Replication European ancestry	Replication European ancestry	Replication European ancestry	Replication European ancestry	Replication European ancestry
Age, mean (SD), years	50.9 (13.52)	56.2 (5.77)	61.01 (8.86)	59.85 (8.16)	57.20 (10.36)	69.54 (0.83)	79.05 (0.58)	21.0 (18.88)
Body mass index, mean (SD), kg/m ²	27.09 (4.01)	29.99 (6.24)	28.81 (4.78)	27.47 (4.79)	27.06 (5.23)	27.70 (4.41)	26.13 (4.07)	20.41 (3.63)
Female, n (%)	117 (50)	2061 (64)	882 (51.07)	388 (54.11)	573 (95.98)	390 (47.10)	215 (57.95)	(33% missing) 284 (48.14)
Smoking status, n (%)								
Never smoker	93 (39.74)	1459 (45.14)	755 (43.77)	209 (29.15)	354 (59.3)	391 (47.22)	165 (44.59)	NA
Former smoker	91 (38.89)	942 (29.15)	720 (41.74)	315 (43.93)	178 (29.9)	341 (41.18)	178 (48.11)	NA
Current smoker	50 (21.37)	831 (25.71)	250 (14.5)	193 (26.92)	65 (10.9)	96 (11.60)	27 (7.30)	NA
ft3, mean (SD), pmol/L	4.80 (0.58)	NA	4.86 (0.61)	NA	NA	NA	NA	4.89 (0.66)
ft4, mean (SD), pmol/L	13.28 (1.60)	14.29 (2.96)	13.77 (1.91)	15.64 (2.10)	NA	15.16 (2.20)	13.96 (2.29)	12.67 (1.42)
TSH, mean (SD), mIU/L	1.31 (0.63)	1.99 (3.62)	1.45 (0.90)	2.26 (1.72)	1.65 (2.23)	2.05 (1.47)	1.77 (1.20)	1.59 (0.92)

Sample sizes per cohort and trait can be found in Figure 1 and in the Supplementary Data.

BSGS, Brisbane Systems Genetics Study; ft3, free triiodothyronine; ft4, free thyroxine; NA, not available; RS, Rotterdam study; TSH, thyrotropin.

Statistical analysis

The EWAS was undertaken in each cohort separately and subsequently meta-analyzed using quantile–quantile normalized beta CpG values as the dependent and log-transformed hormone levels as the independent variable. Each cohort followed the same analysis plan (Supplementary Data), and the results were processed by our in-house quality control pipeline (Supplementary Fig. S1). The EWAS results were corrected for inflation and bias where applicable using the *Bacon* R-package (Supplementary Fig. S2).¹³ A detailed overview of the analyses is provided in Figure 1 and in the Supplementary Data.

Findings of the discovery stage passing significance after the Bonferroni correction (p -value $<0.05/450,000 = 1.1E-07$) were replicated in additional independent samples. A successfully replicated site was defined by p -value <0.05 in the replication stage with consistent effect direction, and a p -value $<1.1E-07$ in the discovery and replication combined meta-analysis.

Correlation with gene expression levels

The association of replicated CpG sites with gene expression levels of genes within ± 500 kb vicinity was assessed in up to 713 blood samples of the KORA study (Supplementary Data).¹⁴ Results passing the false discovery rate (FDR) <0.05 were declared significant.

Additional characterization of the findings

To analyze whether DNAm sites altered by thyroid hormones levels might be influenced by binding of thyroid hormones on thyroid hormone receptors elements, we extracted the overlap and distance between replicated CpGs and binding sites of the thyroid hormone receptor alpha from ChIPAtlas (https://chip-atlas.org/peak_browser).¹⁵ The data were based on human liver¹⁶ and brain cell lines (GEO-ID: GSE129039). Lookup of transcription factor footprint and binding site associations in thyroid tissue was conducted using eFORGE-TF (<https://eforge-tf.altiusinstitute.org>).¹⁷ Additive heritability separating the overall genetic variation from environmental influences on selected DNAm sites was estimated using summary statistics from a previous published data set.¹⁸ To evaluate whether the DNAm sites are altered by *cis* and *trans* genetic variants, we explored GoDMC (www.godmc.org.uk).¹⁹

MR analysis

To test for a possible causal effect of the thyroid hormones on DNAm levels of significantly associated CpG sites, we conducted a two-sample MR using the R-package *TwoSampleMR*.²⁰ As genetic instruments for ft3 were unavailable, we only conducted the MR on ft4- and TSH-associated sites using independent genome-wide significantly associated index single nucleotide polymorphisms (SNPs) from a large genome-wide association study (GWAS) as instruments for the thyroid hormone levels.²¹ The association of these SNPs with DNAm levels (methylation quantitative trait locus [mQTLs]) as outcomes was assessed in 1662 blood samples of the KORA study. The mQTLs were estimated

TABLE 2. TECHNICAL DETAILS OF THE COHORTS USED IN THE EPIGENOME-WIDE ASSOCIATION STUDY ANALYSIS

	<i>SHIP-Trend</i>	<i>ARIC</i>	<i>KORA</i>	<i>RS</i>	<i>TWINSUK</i>	<i>LBC1936</i>	<i>LBC1921</i>	<i>BSGS</i>
Used for Cohort reference (PMID)	Discovery 35348705	Discovery 2646917	Discovery 16032513	Replication 26386597	Replication 22253318	Replication 22253310, 29546429	Replication 22253310, 29546429	Replication 22563384
Medium	Whole blood	Whole blood	Whole blood	Whole blood	Whole blood	Whole blood	Whole blood	Whole blood
DNAm platform	EPIC Illumina	450K Illumina	450K Illumina	450K Illumina	450K Illumina	450K Illumina	450K Illumina	450K Illumina
DNAm array and QC pipeline	CPACOR ^a	BMIQ ^b	CPACOR ^a	CPACOR ^a	GenomeStudio ^c	BMIQ ^b	BMIQ ^b	Inhouse pipeline ^d
Prediction or real proportion of WBC, method if prediction	Prediction, Houseman's method	Prediction, Houseman's method	Prediction, Houseman's method	Houseman's (CD8T and CD4T) and real proportion WBC	Prediction, Houseman's method	Prediction, Houseman's method	Prediction, Houseman's method	Prediction, Houseman's method
TSH assay	Dimension Vista System, Siemens	Elecsys 2010, Roche	Dimension Vista, Siemens	Elecsys 2010, Roche	ARCHITECT, Abbott	ARCHITECT, Abbott	ARCHITECT, Abbott	ARCHITECT, Abbott
ft3 assay	Dimension Vista, Siemens	Not measured	Dimension Vista, Siemens	Not measured	Not measured	Not measured	Not measured	ARCHITECT Free T3, Abbott
ft4 assay	Dimension Vista, Siemens	Elecsys 2010, Roche	Dimension Vista, Siemens	Elecsys 2010, Roche	Not measured	ARCHITECT, Abbott	ARCHITECT, Abbott	ARCHITECT, Abbott

Sample size per cohort and trait can be found in Figure 1 and in the Supplementary Data.

^aSee Lehne et al.³⁸

^bSee Teschendorff et al.³⁹

^cSee Tsaprouni et al.⁴⁰

^dSee McRae et al.⁴¹

BMIQ, Beta Mixture Quantile dilation; CPACOR, incorporating Control Probe Adjustment and reduction of global CORrelation; DNAm, DNA methylation; WBC, white blood cell.

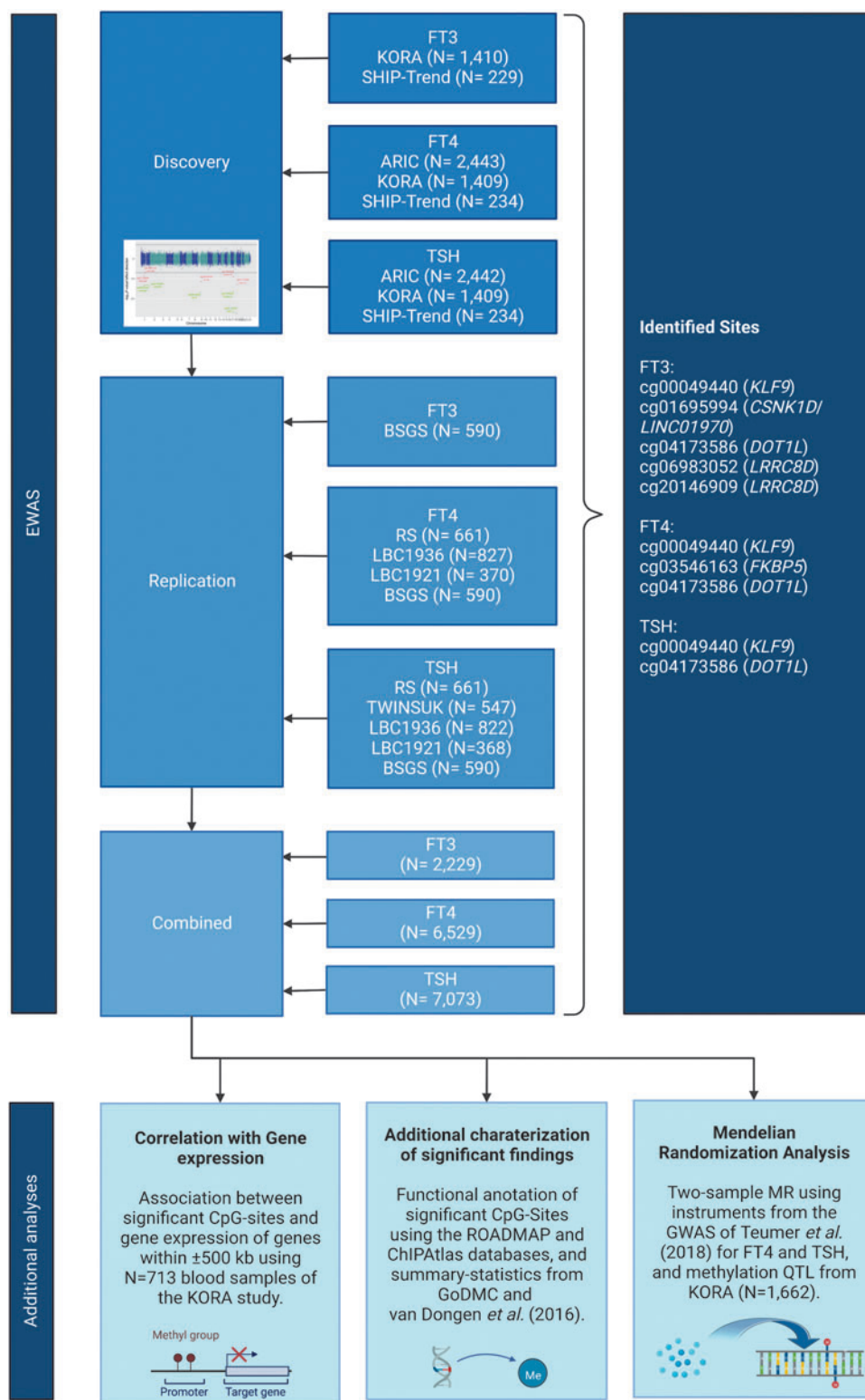


FIG. 1. Flowchart of the analyses performed in the current study. Image created with BioRender.com

by regressing the residuals of the methylation beta values adjusted for sex, age, technical factors, and estimated white blood cell type composition²² on the SNP allele dosage.

Results of the mQTL analyses were available for 31 and 56 instruments for FT4 and TSH, respectively. The inverse-variance weighted fixed-effect MR was conducted as primary analysis, and the weighted median²³ and MR-Egger²⁴ as

sensitivity analyses representing less powerful but more robust methods regarding violations of the validity of the instruments. Significance was assessed by the primary analysis p -value < 0.05 divided by the number of tested CpGs per thyroid function trait. The effects can be interpreted as the proportion increase in DNAm per change in one standard deviation of the thyroid hormone level.

Results

Study sample characteristics

In the discovery stage EWAS based on 3 cohorts, we included 4085 individuals for TSH, 1639 for fT3, and 4081 for fT4. In the replication stage using 5 independent cohorts, 2988 additional individuals were included in the TSH analysis, 590 for fT3, and 2448 for fT4 (Table 1 and Supplementary Data).

DNAm sites associated with thyroid hormone levels

We discovered and replicated two novel CpG sites associated with TSH (cg00049440 in *KLF9* and cg04173586 in *DOTIL*) and three novel sites associated with fT4 (cg00049440 in *KLF9*, cg04173586 in *DOTIL*, and cg03546163 in *FKBP5*). Of the five fT3-associated CpGs (in *KLF9*, *DOTIL*, *LRRC8D*, and near *CSNK1D/LINC01970*), the two CpGs in *LRRC8D* (cg06983052 and cg20146909) are novel. The replicated results are provided in Table 3, the significant results of the discovery stage are shown in Supplementary Table S1, and the cohort-specific results are listed in Supplementary Table S2. In total, all but one site identified in the discovery analysis were successfully replicated. Detailed annotation of the replicated CpGs is shown in Supplementary Table S3.

As indicated by the Chicago plots, all replicated sites were associated with higher methylation for higher TSH (Fig. 2) and with lower methylation for both higher fT4 (Fig. 3) and fT3 (Fig. 4).

In detail, the TSH meta-analysis revealed two associations cg00049440 ($\beta_{\text{combined}}=0.13$; $p\text{-value}=9.57\text{E-}16$) and cg04173586 ($\beta_{\text{combined}}=0.07$; $p\text{-value}=1.53\text{E-}09$). The CpG site cg00049440 is located inside a promotor region in intron 1 of the *KLF9* gene (Supplementary Fig. S3), and cg04173586 is located in intron 1 of *DOTIL* (Supplementary Fig. S4).

The fT4 analysis identified three sites (cg00049440, cg03546163, and cg04173586), two of which, namely cg00049440 in *KLF9* ($\beta_{\text{combined}}=-0.75$; $p\text{-value}=4.44\text{E-}20$; Supplementary Fig. S5) and cg04173586 in *DOTIL* ($\beta_{\text{combined}}=-0.43$; $p\text{-value}=4.03\text{E-}11$; Supplementary Fig. S6), were also identified in the TSH analysis. The additional association is cg03546163 ($\beta_{\text{combined}}=-0.84$; $p\text{-value}=1.45\text{E-}16$), which lies within the second intron of *FKBP5* (Supplementary Fig. S7).

The EWAS on fT3 identified five sites cg00049440, cg01695994, cg04173586, cg06983052, and cg20146909, two of which, namely cg00049440 ($\beta_{\text{combined}}=-1.46$; $p\text{-value}=1.69\text{E-}17$; Supplementary Fig. S8) and cg04173586 ($\beta_{\text{combined}}=-1.93$; $p\text{-value}=4.33\text{E-}28$; Supplementary Fig. S9), were also identified in the fT4 and TSH, and one CpG, namely cg06983052 ($\beta_{\text{combined}}=-1.04$; $p\text{-value}=5.45\text{E-}14$; Supplementary Fig. S10), was identified in the fT4 analysis (Supplementary Table S1). The remaining two sites are cg01695994 ($\beta_{\text{combined}}=-1.37$; $p\text{-value}=5.84\text{E-}17$), which is located 1518 bp downstream of *LINC01970* and 14,809 bp upstream of *CSNK1D* (Supplementary Fig. S11), and cg20146909 ($\beta_{\text{combined}}=-1.04$; $p\text{-value}=6.51\text{E-}14$), which is located inside intron 1 of the gene *LRRC8D* (Supplementary Fig. S12).

The DNAm sites we identified for TSH, fT4, and fT3 were not colocated with SNPs or genes previously highlighted through GWAS of thyroid function.

TABLE 3. CpG SITES SIGNIFICANTLY ASSOCIATED WITH THYROID FUNCTION AFTER THE REPLICATION STAGE

CpG probeID	Chr: position (build 37)	Discovery estimate	p	Sample size	Replication estimate	p	Sample size	Combined estimate	p	Sample size	Nearest gene
fT4	cg00049440	-0.790	4.53E-15	4078	-0.674	1.45E-06	2448	-0.750	4.44E-20	6526	<i>KLF9</i>
	cg03546163	-0.849	3.11E-13	3878	-0.825	1.00E-04	1251	-0.843	1.45E-16	5129	<i>FKBP5</i>
	cg04173586	-0.445	5.14E-09	4038	-0.380	1.88E-03	2444	-0.427	4.03E-11	6482	<i>DOTIL</i>
fT3	cg00049440 ^a	-1.349	2.61E-12	1637	-1.876	1.81E-07	590	-1.463	1.69E-17	2227	<i>KLF9</i>
	cg01695994 ^a	-1.112	8.43E-10	1638	-2.497	8.63E-12	590	-1.371	5.84E-17	2228	<i>CSNK1D/LINC01970</i>
	cg04173586 ^a	-1.842	7.98E-14	1607	-2.013	1.51E-16	586	-1.935	4.33E-28	2193	<i>DOTIL</i>
	cg06983052	-0.875	5.90E-08	1639	-1.474	8.49E-09	590	-1.040	5.45E-14	2229	<i>LRRC8D</i>
	cg20146909	-0.889	3.38E-09	1639	-1.811	5.67E-08	590	-1.036	6.51E-14	2229	<i>LRRC8D</i>
TSH	cg00049440	0.143	5.19E-14	4082	0.088	1.11E-03	2988	0.125	9.57E-16	7070	<i>KLF9</i>
	cg04173586	0.077	4.52E-08	4041	0.061	8.52E-03	2984	0.072	1.53E-09	7025	<i>DOTIL</i>

The effect estimates and p -values of the discovery and replication combined meta-analysis are provided.

^aKnown association.

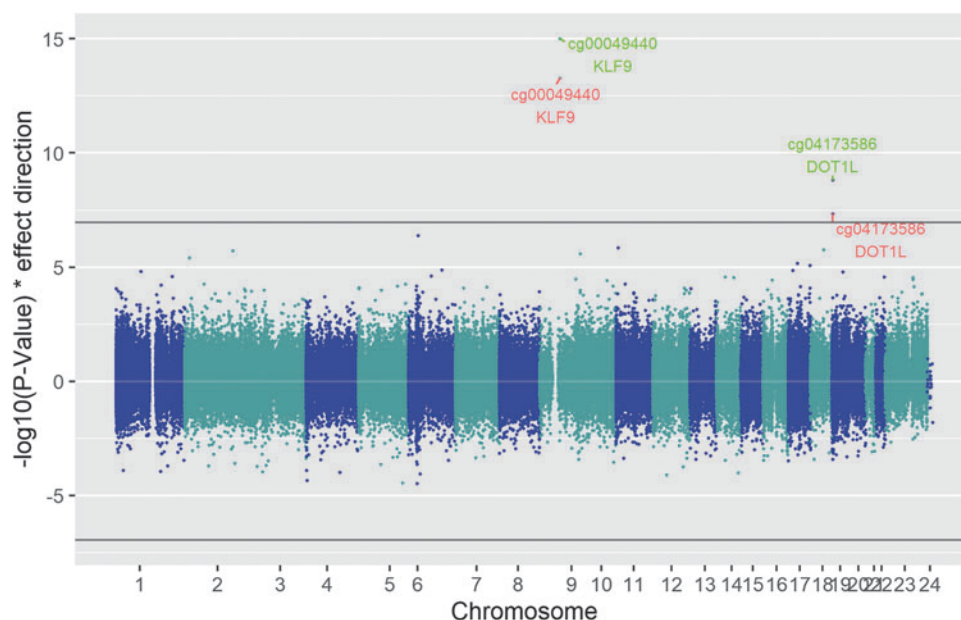


FIG. 2. Chicago plot showing the results of the TSH discovery analysis. Red labels show the significant sites (p -value $< 1.1E-07$) of the discovery analysis, while green labels show results of the combined analysis. The black line shows the epigenome-wide significance cutoff ($1.1E-07$). Points plotted in the upper panel show a positive correlation with TSH, while those plotted below have a negative association. TSH, thyrotropin.

Effects on gene expression in blood

As DNAm represents an important regulator of gene expression, we tested the association of the methylation levels of our replicated findings with mRNA levels in blood of nearby genes. Of all 146 CpG-mRNA associations tested, only cg00049440 passed the level of significance (FDR < 0.05) and was negatively correlated with *KLF9* gene expression ($\beta = -0.109$; p -value = $3.00E-7$; Supplementary Table S4). The CpG cg00049440 is located in an island shore and in the promoter region of *KLF9*, and was associated with circulating TSH, fT3, and fT4 levels.

Overlap of DNAm sites with thyroid hormone receptor binding sites

Except for cg20146909, all CpGs that were associated with thyroid function fall in promoters and enhancer

regions shared between several tissues including blood, liver, and brain (Supplementary Table S3). As thyroid hormone receptor footprints are dynamic, capable of chromatin remodeling and dependent on thyroid hormones levels,^{25,26} we furthermore studied whether DNAm sites associated with fT4, fT3, and TSH fall in thyroid hormone receptor binding sites in liver and neural cells where the data were available (see the Materials and Methods section). The cg01695994 at *CSNK1D/LINC01970* overlapped with thyroid hormone receptor alpha binding peaks in liver cells. We did not identify any other overlap between thyroid hormone receptor binding site peaks and our six significant DNAm sites (distance between 517 and 2219 bp; Supplementary Table S3). This could be either because of tissue specificity or because there are more complex relationships between thyroid hormone levels and DNAm.

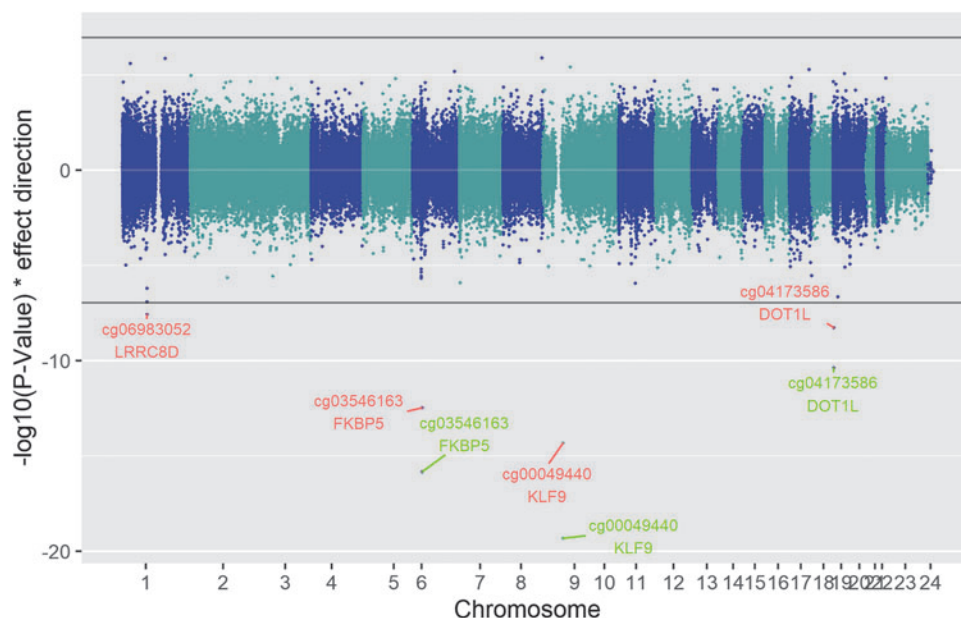
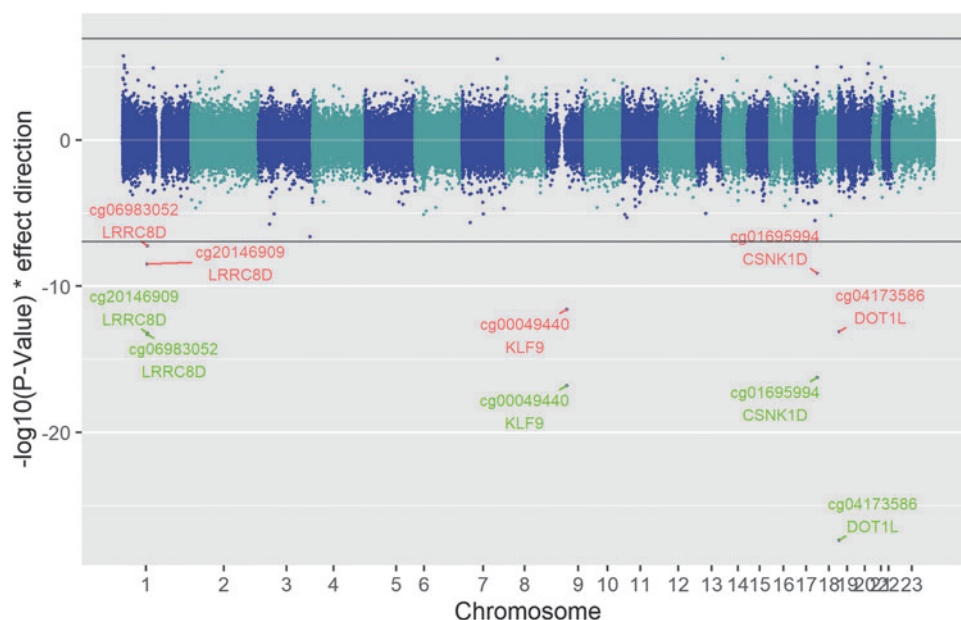


FIG. 3. Chicago plot showing the results of the fT4 discovery analysis. Red labels show the significant sites (p -value $< 1.1E-07$) of the discovery analysis, while green labels show results of the combined analysis. The black line shows the epigenome-wide significance cutoff ($1.1E-07$). Points plotted in the upper panel show a positive correlation with fT4, while those plotted below have a negative association. fT4, free thyroxine.

FIG. 4. Chicago plot showing the results of the fT3 discovery analysis. Red labels show the significant sites (p -value $< 1.1 \times 10^{-7}$) of the discovery analysis, while green labels show results of the combined analysis. The black line shows the epigenome-wide significance cutoff (1.1×10^{-7}). Points plotted in the upper panel show a positive correlation with fT3, while those plotted below have a negative association. fT3, free triiodothyronine.



Impacts of genetic variation on thyroid-DNA_m sites

The total additive heritability of individual DNA_m levels (h^2) at our six significant CpG sites is above 25% (cg06983052) and up to 57% (cg06983052) in blood (Supplementary Table S3). DNA_m heritability estimates at cg00049440 and cg03546163 were explained by 72% and 100% of common genetic variants in the genome (h^2_{SNPs}/h^2), respectively.¹⁸ However, the proportion of heritability explained by common genetic variants for the five remaining signals was less than 30% and near 0% for cg04173586.

Using mQTL results identified in GoDMC,¹⁹ several independent *cis* and *trans* genetic variants impact the blood DNA_m levels of all significant CpG sites except *DOT1L* (Supplementary Table S5). Interestingly, DNA_m levels at cg00049440 (*KLF9*) and cg03546163 (*FKBP5*) were significantly associated with several genetic variants in *trans* located in an intronic region of the *THRB* gene (rs9310736 and rs869785), which encodes the nuclear hormone receptor for triiodothyronine. However, none of these variants were identified in thyroid hormone GWAS^{21,27} or as eQTLs in the GTEx database (<https://gtexportal.org>).

Causal effects of thyroid status on DNA_m

We conducted an MR analysis to test if variation in TSH and fT4 causally affects the changes in DNA_m of the replicated EWAS findings. MR is a framework using SNPs as instrumental variables to assess an unconfounded exposure–outcome association and thus allowing a causal inference.^{28,29} To maximize statistical power, we applied a two-sample MR using published large-scale GWAS results for selecting instruments for TSH and fT4 as exposure. Because no instruments for fT3 from former GWAS were available, no causal effects of triiodothyronine on DNA_m as outcome could be assessed. The MR results revealed a significant effect of genetically determined TSH levels on cg00049440 at *KLF9* ($\beta = 0.005$; p -value = 0.01). The MR-Egger and the weighted median MR analysis were nominally

significant with the same effect directions and similar effect sizes underpinning the robustness of the findings. No heterogeneity (Q-statistics p -value = 0.35) or directional pleiotropy (MR-Egger intercept p -value = 0.33) being both indicators of wrong instruments was detected.

All MR results of the second TSH-associated CpG site as well as the fT4-associated sites had p -values > 0.05 (Supplementary Table S6). As an additional sensitivity analysis addressing potential horizontal pleiotropy through associations with thyroperoxidase (TPO), we removed instruments that were associated (p -value < 0.05) with TPO antibody positivity in a recent GWAS.³⁰ After excluding five TPO-associated SNPs for TSH, significance changed only for the weighted median MR result for TSH on cg00049440 to p -value = 0.06 with a similar effect estimate. No instruments needed to be excluded for the MR of fT4 (Supplementary Table S6).

Discussion

In the current analyses, we revealed three significantly associated CpGs for fT4, five for fT3, and two for TSH. An overview of our findings in relation to known mechanisms is illustrated in Figure 5. Additionally to the known associations between fT3 and Krueppel-like factor 9 (*KLF9*) and DOT1-like histone lysine methyltransferase (*DOT1L*)-related CpG sites,⁵ we found *de novo* associations regarding TSH and fT4 with effect directions consistent with hypothalamic–pituitary–thyroid axis physiology (i.e., opposite effects for TSH and fT4/fT3).

The ubiquitously expressed *KLF9* gene is part of the Sp1 C2H2-type zinc finger transcription factor family that binds to GC box elements located in the promoter (NCBI RefSeq database). *KLF9* was found to be a triiodothyronine response gene in mice²⁶ and frogs,³¹ where triiodothyronine regulated gene expression by binding to and activating thyroid hormone receptors, which in turn regulate *KLF9* transcription via the thyroid response element.²⁶ Furthermore, there is evidence that this mechanism is also present in human cells

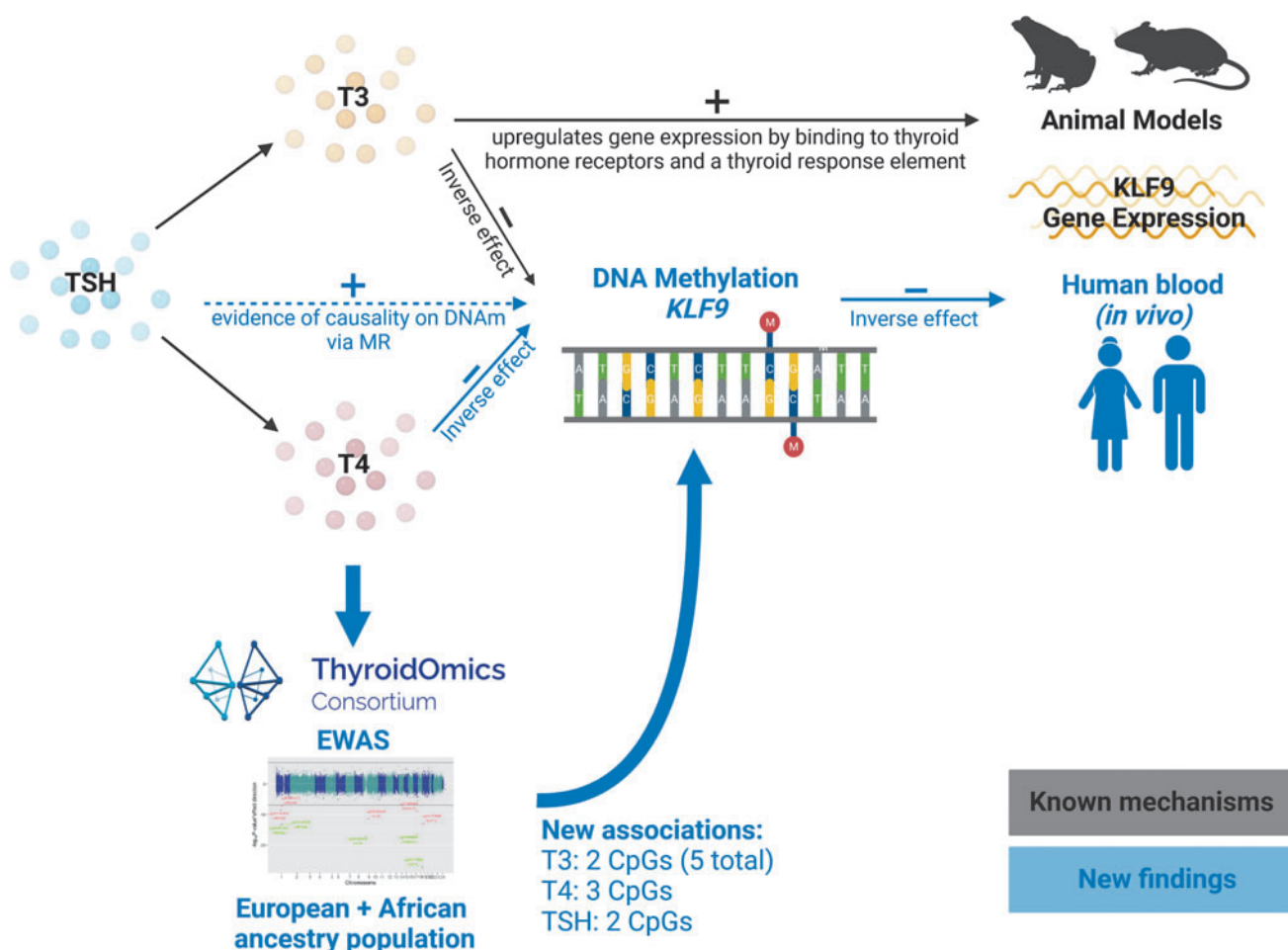


FIG. 5. Overview of the study analyses and findings. Known mechanisms and associations are colored in black, and new findings revealed by this study are marked in blue color. Image created with BioRender.com

including hepatocytes³² and neural cells.³³ Our *KLF9* findings exemplify how thyroid function affects gene expression via DNAm in humans. In particular, we found that higher thyroid hormone levels (with physiologically fitting lower TSH levels) were associated with lower DNAm levels at cg00049440, located in the promotor region of *KLF9*, which in turn was associated with increased *KLF9* expression in blood cells (Fig. 5).

Importantly, the causality in this cascade was supported by MR analyses showing directionally consistent associations with genetically determined TSH levels. Previous studies showed no overlap between variants detected in TSH and fT4 GWAS,²¹ while MR analyses showed that well-known thyroid status-related endpoints such as cholesterol levels and atrial fibrillation were associated with TSH but not with fT4 variants.^{34,35} This strongly suggests that TSH-associated genetic variants are a more valid representation of thyroid function, which likely explains the absence of associations in our fT4 MR analyses.

We hereby show in humans that next to the known pathway via binding to thyroid hormone receptor beta on a response element in the *KLF9* promoter, thyroid hormones can also affect *KLF9* gene expression via *KLF9* DNAm (Fig. 5). In this context, it is noteworthy to mention that in

a study in developing *Xenopus* brain, triiodothyronine has been shown to promote DNA demethylation. This is in part by promoting TET3 recruitment to discrete genomic regions, and in part by inducing genes that encode DNA demethylation enzymes.³⁶

The second multi-trait association, *DOTIL*, represents a ubiquitously expressed gene coding for a methyltransferase that methylates lysine-79 of histone H3. A link to thyroid hormones was found in *Xenopus laevis* where triiodothyronine induces *DotIL* expression during metamorphosis.³⁷ Our study partially translates these results to humans, where the thyroid function-associated CpG cg04173586 is located in an enhancer region. However, the correlation of DNAm with *DOTIL* mRNA levels in 699 blood tissue samples was not significant.

This points to a limitation of our study as blood represents only one of the possible target tissues for both DNAm and gene expression analyses. Additionally, the sample size might have been too small to achieve the required statistical power, although the effect direction suggests an inverse association (Supplementary Table S4). The advantage of correlating both measured DNAm and mRNA is that this approach does not depend on the availability of known genetic factors needed to predict these measurements in alternative approaches such as colocalization.

In addition to two new and three known associations with fT3 revealed in our study, the remaining three associations discovered in the study of Lafontaine et al.⁵ were statistically significant ($p < 0.05/6$) in our EWAS. Considering that our replication cohort for fT3 was also included in Lafontaine et al, we restricted this lookup to the discovery stage results (Supplementary Table S7) providing an independent sample set.

Of note, the CpG cg20146909 in *LRRC8D* that was associated with fT3 in our study was suggestive of an association with the study by Lafontaine et al.⁵ Thus, we confirmed this candidate and revealed a second site in this gene, namely cg06983052, that was associated with both fT3 and fT4.

The two previously reported TSH associations at *FOXK2* and the lncRNA *AC012506.1* were not replicated in our study (Supplementary Table S7). This could be because the previous findings were false positive results, or because these associations are specific to adolescents or young adults who accounted for most participants in the former study.

In addition to data sets from European ancestry, our study also included a substantial proportion of DNAm data of African ancestry individuals via the ARIC cohort. However, no ancestry-specific heterogeneity of the effect estimates was observed for our replicated EWAS results (Supplementary Table S2 and Supplementary Fig. S13).

In conclusion, we conducted an EWAS by developing and applying a standardized analyses and quality control workflow in a large sample of both European and African ancestry individuals, replicated the statistically significant associations in independent cohorts, and characterized the findings by integrating multiple types of -omics data. Our study revealed DNAm associated with thyroid function at five distinct loci. In particular, we showed that DNAm is another previously unknown route underlying thyroid hormone-induced *KLF9* expression, which is a well-known thyroid hormone household gene. This study serves as a proof-of-concept that integration of EWAS with other -omics data can provide a valuable tool to unravel thyroid hormone signaling in humans by complementing and feeding classical *in vitro* and animal studies.

Authors' Contributions

Project design: A.Te., J.T.B., L.C., M.M., R.P.P., and T.C.M. Data collection: A.F.M., C.G., C.M., D.v.H., E.P.S., H.G., H.J.G., H.P., J.P.W., J.T.B., M.F., M.N., M.R., M.W., N.G.M., P.J.C., S.E.H., S.G.W., and S.R.C. Cohort study management: A.P., A.U., C.M., E.P.S., H.V., J.B.J.v.M., J.T.B., M.B., M.F., M.N., N.G.M., R.P.P., S.R.C., and T.C.M. Subject recruitment: C.M., H.V., N.G.M., and S.R.C. Drafting of the article: A.Te., A.W., and T.C.M. Interpretation of results: A.K., A.Te., A.W., E.S., J.N., J.P.W., J.T.B., M.M., S.G.W., and T.C.M. Statistical methods and analysis: A.F.M., A.Te., A.Ti., A.W., B.K., B.R.S., D.L.M., J.N., K.V.E.B., L.C., P.J.C., P.-C.T., R.E.M., S.E.H., and T.C.M. Critical review of the article: A.F.M., A.K., A.P., A.Te., A.Ti., A.U., A.W., B.K., B.R.S., C.G., C.M., D.L.M., D.v.H., E.P.S., E.S., W.E.V., H.G., H.J.G., H.P., H.V., J.B.J.v.M., J.N., J.P.W., J.T.B., K.V.E.B., L.C., M.B., M.F., M.M., M.N., M.R., M.W., N.G.M., P.J.C., P.-C.T., R.E.M., R.P.P., S.E.H., S.G.W., S.R.C., and T.C.M.

Author Disclosure Statement

H.J.G. has received travel grants and speakers honoraria from Fresenius Medical Care, Neuraxpharm, Servier, and Janssen Cilag as well as research funding from Fresenius Medical Care not related to this work. R.E.M. has received speaker fees from Illumina and is an advisor to the Epigenetic Clock Development Foundation. All other authors declare no conflicts of interest.

Funding Information

Acknowledgements and funding sources are listed in the Supplementary Note section in the Supplementary Data.

Supplementary Material

Supplementary Data

References

- Cooper DS, Biondi B. Subclinical thyroid disease. *Lancet* 2012;379(9821):1142–1154; doi: 10.1016/S0140-6736(11)60276-6
- Wondisford, Fredric E; Radovick S, (eds). *Clinical Management of Thyroid Disease*, 1st ed. Elsevier: Philadelphia, USA; 2009.
- Wu Y, Koenig RJ. Gene regulation by thyroid hormone. *Trends Endocrinol Metab* 2000;11(6):207–211; doi: 10.1016/S1043-2760(00)00263-0
- Kyono Y, Sachs LM, Bilesimo P, et al. Developmental and thyroid hormone regulation of the DNA methyltransferase 3a gene in *Xenopus* tadpoles. *Endocrinology* 2016;157(12):4961–4972; doi: 10.1210/en.2016-1465
- Lafontaine N, Campbell PJ, Castillo-Fernandez JE, et al. Epigenome-wide association study of thyroid function traits identifies novel associations of fT3 with KLF9 and DOT1L. *J Clin Endocrinol Metab* 2021;106(5):e2191–e2202; doi: 10.1210/clinem/dgaa975
- Völzke H, Schössow J, Schmidt CO, et al. Cohort profile update: The study of health in Pomerania (SHIP). *Int J Epidemiol* 2022; doi: 10.1093/ije/dyac034
- Wright JD, Folsom AR, Coresh J, et al. The ARIC (atherosclerosis risk in communities) study: JACC focus seminar 3/8. *J Am Coll Cardiol* 2021;77(23):2939–2959; doi: 10.1016/j.jacc.2021.04.035
- Holle R, Happich M, Löwel H, et al. KORA—A research platform for population based health research. *Das Gesundheitswes* 2005;67(S 01):19–25; doi: 10.1055/s-2005-858235
- Hofman A, Brusselle GGO, Murad SD, et al. The Rotterdam study: 2016 objectives and design update. *Eur J Epidemiol* 2015;30(8):661–708; doi: 10.1007/s10654-015-0082-x
- Moayyeri A, Hammond CJ, Valdes AM, et al. Cohort profile: TwinsUK and healthy ageing twin study. *Int J Epidemiol* 2013;42(1):76–85; doi: 10.1093/ije/dyr207
- Taylor AM, Pattie A, Deary IJ. Cohort profile update: The Lothian birth cohorts of 1921 and 1936. *Int J Epidemiol* 2018;47(4):1042–1042r; doi: 10.1093/ije/dyy022
- Powell JE, Henders AK, McRae AF, et al. The Brisbane systems genetics study: Genetical genomics meets complex trait genetics. *PLoS One* 2012;7(4):e35430; doi: 10.1371/journal.pone.0035430
- van Iterson M, van Zwet EW, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biol* 2017;18(1):19; doi: 10.1186/s13059-016-1131-9

14. Schurmann C, Heim K, Schillert A, et al. Analyzing illumina gene expression microarray data from different tissues: Methodological aspects of data analysis in the metaxpress consortium. *PLoS One* 2012;7(12):e50938; doi: 10.1371/journal.pone.0050938
15. Oki S, Ohta T, Shioi G, et al. ChIP-Atlas: A data-mining suite powered by full integration of public ChIP-seq data. *EMBO Rep* 2018;19(12); doi: 10.15252/EMBR.201846255
16. Partridge EC, Chhetri SB, Prokop JW, et al. Occupancy maps of 208 chromatin-associated proteins in one human cell type. *Nature* 2020;583(7818):720–728; doi: 10.1038/s41586-020-2023-4.
17. Breeze CE. Cell type-specific signal analysis in epigenome-wide association studies. *Methods Mol Biol* 2022;2432: 57–71; doi: 10.1007/978-1-0716-1994-0_5
18. Van Dongen J, Nivard MG, Willemsen G, et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat Commun* 2016;7(1): 1–13; doi: 10.1038/ncomms11115
19. Min JL, Hemani G, Hannon E, et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat Genet* 2021;53(9):1311–1321; doi: 10.1038/s41588-021-00923-X
20. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018;7; doi: 10.7554/eLife.34408
21. Teumer A, Chaker L, Groeneweg S, et al. Genome-wide analyses identify a role for SLC17A4 and AADAT in thyroid hormone regulation. *Nat Commun* 2018;9(1):4455; doi: 10.1038/s41467-018-06356-1
22. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012;13(1):86; doi: 10.1186/1471-2105-13-86
23. Bowden J, Davey Smith G, Haycock PC, et al. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016;40(4):304–314; doi: 10.1002/gepi.21965
24. Bowden J, Smith GD, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015; 44(2); doi: 10.1093/ije/dyv080
25. Grøntved L, Waterfall JJ, Kim DW, et al. Transcriptional activation by the thyroid hormone receptor through ligand-dependent receptor recruitment and chromatin remodelling. *Nat Commun* 2015 61 2015;6(1):1–11; doi: 10.1038/ncomms8048
26. Denver RJ, Williamson KE. Identification of a thyroid hormone response element in the mouse Krüppel-like factor 9 gene to explain its postnatal expression in the brain. *Endocrinology* 2009;150(8):3935–3943; doi: 10.1210/en.2009-0050
27. Zhou W, Brumpton B, Kabil O, et al. GWAS of thyroid stimulating hormone highlights pleiotropic effects and inverse association with thyroid cancer. *Nat Commun* 2020;11(1):3981; doi: 10.1038/s41467-020-17718-z
28. Ebrahim S, Davey Smith G. Mendelian randomization: Can genetic epidemiology help redress the failures of observational epidemiology? *Hum Genet* 2008;123(1):15–33; doi: 10.1007/s00439-007-0448-6
29. Medici M, Peeters RP, Teumer A, et al. The importance of high-quality mendelian randomisation studies for clinical thyroidology. *Lancet Diabetes Endocrinol* 2019;50(5):668–681; doi: 10.1016/S2213-8587(19)30145-7
30. Medici M, Porcu E, Pistis G, et al. Identification of novel genetic Loci associated with thyroid peroxidase antibodies and clinical thyroid disease. *PLoS Genet* 2014;10(2): e1004123; doi: 10.1371/journal.pgen.1004123
31. Furlow JD, Kanamori A. The transcription factor basic transcription element-binding protein 1 is a direct thyroid hormone response gene in the frog *Xenopus laevis*. *Endocrinology* 2002;143(9):3295–3305; doi: 10.1210/EN.2002-220126
32. Cvaro A, Devito L, Milton FA, et al. A thyroid hormone receptor/KLF9 axis in human hepatocytes and pluripotent stem cells. *Stem Cells* 2015;33(2):416–428; doi: 10.1002/stem.1875
33. Morte B, Gil-Ibáñez P, Bernal J. Regulation of gene expression by thyroid hormone in primary astrocytes: Factors influencing the genomic response. *Endocrinology* 2018;159(5):2083–2092; doi: 10.1210/EN.2017-03084
34. Marouli E, Kus A, Del Greco M F, et al. Thyroid function affects the risk of stroke via atrial fibrillation: A Mendelian randomization study. *J Clin Endocrinol Metab* 2020; doi: 10.1210/clinem/dgaa239
35. Kuś A, Marouli E, Del Greco M F, et al. Variation in normal range thyroid function affects serum cholesterol levels, blood pressure and type 2 diabetes risk: A Mendelian randomization study. *Thyroid* 2020; doi: 10.1089/thy.2020.0393
36. Raj S, Kyono Y, Sifuentes CJ, et al. Thyroid hormone induces DNA demethylation in *Xenopus* tadpole brain. *Endocrinology* 2020;161(11); doi: 10.1210/endocr/bqaa155
37. Wen L, Fu L, Shi Y. Histone methyltransferase Dot1L is a coactivator for thyroid hormone receptor during *Xenopus* development. *FASEB J* 2017;31(11):4821–4831; doi: 10.1096/fj.201700131R
38. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol* 2015;16(1):37; doi: 10.1186/s13059-015-0600-x
39. Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 2013;29(2):189–196; doi: 10.1093/bioinformatics/bts680
40. Tsaprouni LG, Yang T, Bell J, et al. Epigenetics cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics* 2014;9(10):1382–1396; doi: 10.4161/15592294.2014.969637
41. McRae AF, Powell JE, Henders AK, et al. Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol* 2014;15(5):R73; doi: 10.1186/gb-2014-15-5-r73

Address correspondence to:

Alexander Teumer, PhD

Department of Psychiatry and Psychotherapy

University Medicine Greifswald

Ellernholzstr. 1-2

Greifswald 17475

Germany

E-mail: ateumer@uni-greifswald.de