

## Carriers of a VEGFA enhancer polymorphism selectively binding chop/ddit3 are predisposed to increased circulating levels of thyroid stimulating hormone

Ahluwalia, Tarunveer Singh; Troelsen, Jesper; Balslev-Harder, Marie; Bork-Jensen, Jette; Thuesen, Betina Heinsbæk; Cerqueira, Charlotte; Linneberg, Allan; Grarup, Niels; Pedersen, Oluf; Hansen, Torben; Dalgaard, Louise Torp

*Published in:*  
Journal of Medical Genetics

*DOI:*  
[10.1136/jmedgenet-2016-104084](https://doi.org/10.1136/jmedgenet-2016-104084)

*Publication date:*  
2017

*Document Version*  
Peer reviewed version

### *Citation for published version (APA):*

Ahluwalia, T. S., Troelsen, J., Balslev-Harder, M., Bork-Jensen, J., Thuesen, B. H., Cerqueira, C., Linneberg, A., Grarup, N., Pedersen, O., Hansen, T., & Dalgaard, L. T. (2017). Carriers of a VEGFA enhancer polymorphism selectively binding chop/ddit3 are predisposed to increased circulating levels of thyroid stimulating hormone. *Journal of Medical Genetics*, 54(3), 166-175. Advance online publication. <https://doi.org/10.1136/jmedgenet-2016-104084>

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

### **Take down policy**

If you believe that this document breaches copyright please contact [rucforsk@kb.dk](mailto:rucforsk@kb.dk) providing details, and we will remove access to the work immediately and investigate your claim.

**CARRIERS OF A VEGFA ENHANCER POLYMORPHISM  
 SELECTIVELY BINDING CHOP/DDIT3 ARE PREDISPOSED TO  
 INCREASED CIRCULATING LEVELS OF THYROID  
 STIMULATING HORMONE**

Journal:	<i>Journal of Medical Genetics</i>
Manuscript ID	jmedgenet-2016-104084.R1
Article Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	<p>Ahluwalia, Tarunveer; University of Copenhagen, The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences; Steno Diabetes Center AS</p> <p>Troelsen, Jesper; Roskilde Universitet, Department of Science and Environment</p> <p>Balslev-Harder, Marie; University of Copenhagen, 1The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences</p> <p>Bork-Jensen, Jette; University of Copenhagen, The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences</p> <p>Thuesen, Betina; Research Centre for Prevention and Health</p> <p>Cerqueira, Charlotte; Research Centre for Prevention and Health</p> <p>Linneberg, Allan; Research Centre for Prevention and Health</p> <p>Grarup, Niels; Kobenhavns Universitet Sundhedsvidenskabelige Fakultet, The Novo Nordisk Foundation Center for Basic Metabolic Research</p> <p>Pedersen, Oluf; Kobenhavns Universitet Sundhedsvidenskabelige Fakultet, The Novo Nordisk Foundation Center for Basic Metabolic Research</p> <p>Hansen, Torben; Kobenhavns Universitet Sundhedsvidenskabelige Fakultet, The Novo Nordisk Foundation Center for Basic Metabolic Research</p> <p>Dalgaard, Louise; Roskilde Universitet, Department of Science and Environment</p>
Keywords:	thyroid, Molecular genetics, Metabolic disorders, insulin resistance, transcription factor

1  
2  
3 **CARRIERS OF A *VEGFA* ENHANCER POLYMORPHISM SELECTIVELY**  
4 **BINDING CHOP/DDIT3 ARE PREDISPOSED TO INCREASED**  
5 **CIRCULATING LEVELS OF THYROID STIMULATING HORMONE**  
6  
7

8 Running title: Circulating TSH association with a *VEGFA* functional polymorphism  
9

10  
11  
12 Tarunveer Singh Ahluwalia<sup>1,2</sup>, Jesper Thorvald Troelsen<sup>3</sup>, Marie Balslev-Harder<sup>1</sup>, Jette Bork-  
13 Jensen<sup>1</sup>, Betina Heinsbæk Thuesen<sup>4</sup>, Charlotte Cerqueira<sup>4</sup>, Allan Linneberg<sup>4,5,6</sup>, Niels Grarup<sup>1</sup>, Oluf  
14 Pedersen<sup>1</sup>, Torben Hansen<sup>1,7</sup>, Louise Torp Dalgaard<sup>3</sup>  
15  
16

17 <sup>1</sup>The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics,  
18 Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

19 <sup>2</sup>Steno Diabetes Center, Gentofte, Denmark

20 <sup>3</sup>Department of Science and Environment, Roskilde University, Roskilde, Denmark

21 <sup>4</sup>Research Centre for Prevention and Health, The Capital Region, Glostrup, Denmark

22 <sup>5</sup>Department of Clinical Experimental Research, Rigshospitalet, Copenhagen, Denmark

23 <sup>6</sup>Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of  
24 Copenhagen, Copenhagen, Denmark

25 <sup>7</sup>Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark  
26  
27  
28  
29

30 **Address correspondence to:**  
31

32 Associate professor Louise Torp Dalgaard, MSc, PhD, Department of Science and Environment,  
33 Roskilde University, Universitetsvej 1, DK-4000 Roskilde, Denmark. E-mail: ltd@ruc.dk, phone:  
34 +4546742713  
35  
36  
37

38 Word counts: Text: 4136, Abstract: 250, References: 49, Tables: 2, Supplementary tables: 7  
39

40 Figures: 4, Supplementary figures: 4  
41  
42  
43

44 **KEYWORDS**  
45  
46

47 thyroid, genetics, metabolic disorders, insulin resistance, transcription factor  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**ABSTRACT**

**Background:** Levels of serum thyroid stimulating hormone (TSH) indicate thyroid function, because thyroid hormone negatively controls TSH release. Genetic variants in the vascular endothelial growth factor A (*VEGFA*) gene are associated with TSH levels. The aim was to characterize the association of *VEGFA* variants with TSH in a Danish cohort and to identify and characterize functional variants.

**Methods:** We performed an association study of the *VEGFA* locus for circulating TSH levels in 8445 Danish individuals. Lead variants were tested for allele-specific effects *in vitro* using luciferase reporter and gel-shift assays.

**Results:** Four SNPs in *VEGFA* were associated with circulating TSH (rs9472138, rs881858, rs943080 and rs4711751). For rs881858, the presence of each G allele was associated with a corresponding decrease in TSH levels of 2.3% ( $P=8.4 \times 10^{-9}$ ) and an increase in circulating free T4 levels ( $P=0.0014$ ). Rs881858 is located in a binding site for CHOP (C/EBP homology protein) and C/EBP $\beta$  (ccaat enhancer binding protein  $\beta$ ). Reporter-gene analysis showed increased basal enhancer activity of the rs881858 A-allele versus the G-allele ( $34.5 \pm 9.9\%$  (average  $\pm$  SEM),  $P=0.0012$ ), while co-expression of CHOP effectively suppressed the rs881858 A-allele activity. The A-allele showed stronger binding to CHOP in gel-shift assays.

**Conclusions:** VEGF is an important angiogenic signal required for tissue expansion. We show that *VEGFA* variation giving allele-specific response to transcription factors with overlapping binding sites associate closely with circulating TSH levels. Because CHOP is induced by several types of intracellular stress, this indicates that cellular stress could be involved in the normal or pathophysiological response of the thyroid to TSH.

**ABBREVIATIONS**

TSH thyroid stimulating hormone, VEGFA vascular endothelial growth factor A, CHOP c/EBP  
homology protein, c/EBP $\beta$  ccaat enhancer binding protein  $\beta$ , CEBPB *C/EBP $\beta$*  gene symbol, SNP  
single nucleotide polymorphism, Chr chromosome, EAF effect allele frequency, BS binding site, ds  
double-stranded, DTT dithiothreitol, GWAS genome wide association study, BMI body mass index,  
eQTL expressed quantitative trait locus

## INTRODUCTION

The thyroid gland is an essential regulator whole body energy expenditure and metabolic rate. Circulating levels and activities of thyroid hormones, their activating enzymes (deiodinases) and the regulating hormones TSH (thyrotropin/thyroid stimulating hormone) and TRH (thyrotropin releasing hormone) are precisely balanced to ensure the euthyroid state. Circulating levels of TSH comprise a clinically valuable indicator of thyroid function, and in the absence of pituitary or hypothalamic failure, an increased level of TSH is a very sensitive marker of decreased thyroid function. Clinical reference levels for TSH define elevated TSH levels above 4mU/L as being associated with clinically decreased thyroid function<sup>1</sup>. However, there is a marked inter-individual, while low intra-individual variability in circulating TSH levels, as well as in the hypothalamic-pituitary-thyroid axis<sup>2</sup>, which appears to be highly heritable as evidenced by heritability estimates of 65% derived from twin studies<sup>3-5</sup>.

Subclinical hypothyroidism (TSH above 4mU/L with T4 levels within the reference range) is associated with an impaired metabolic phenotype, cardiovascular risk factors, elevated blood total cholesterol and blood pressure increase, decreased glomerular filtration rate and bone fractures<sup>6</sup>. However, association between TSH within the reference range and obesity is not well established. A meta-analysis encompassing 29 studies, found 18 of these to report a positive association<sup>7</sup>. There have been reports of positive correlations between TSH levels and BMI in obese or over-weight individuals, suggesting decreased thyroid function in these subjects<sup>8-10</sup>, but also cross sectional population-based reports have been made for the association between increased TSH levels and BMI<sup>11-13</sup>.

Recent genome wide association studies (GWAS) for circulating serum TSH levels have focused on common variants (minor allele frequency (MAF) > 5%) and have identified at least 26 genomic loci to date<sup>14-17</sup> of which one was the Vascular Endothelial Growth Factor A gene (*VEGFA*) (Fig. 1A).

1  
2  
3 Both the VEGF protein and its receptor, KDR (Kinase Insert Domain Receptor), are highly  
4 expressed in the thyroid gland<sup>18</sup>. *In vivo*, TSH administration increases VEGF release from the  
5 thyroid gland<sup>19</sup> and treatment of isolated, cultured thyrocytes with TSH also stimulates VEGF  
6 release<sup>20</sup>. Variants regulating the activity of the *VEGFA* locus are likely contributors to the observed  
7 TSH association<sup>14</sup>, because TSH levels indicate thyroid function.  
8  
9

10  
11  
12 In order to investigate the molecular and genetic mechanisms in the *VEGFA* locus controlling  
13 circulating levels of TSH, we performed an association study using densely spaced SNPs of the  
14 *VEGFA* genomic region in three population-based cohorts from Denmark comprising 8,445  
15 individuals and meta-analysed them. The lead variants were further investigated to identify allele  
16 specific effects using *in vitro* cell based assays to elucidate the molecular mechanism supporting the  
17 observed clinical findings.  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27

## 28 29 **METHODS**

### 30 31 **Genetic association analysis**

#### 32 33 *Study participants*

34  
35 The study was conducted in accordance with the Helsinki declaration and approved by the Danish  
36 Data Protection Board and by the Ethical Committee of Copenhagen County. Informed written  
37 consent was obtained from all subjects before participation. The genetic association analyses were  
38 performed in three Danish study cohorts (Inter99, Health2006 and Health2008) that have been  
39 described previously elsewhere<sup>21</sup>: 1) The study (ClinicalTrials.gov ID-no: NCT00289237) is a  
40 population-based study for ischemic heart disease<sup>22</sup>, 2) The Health 2006 Study (Ethical committee  
41 approval number: KA20060011) is a population based study comprising individuals aged between  
42 18-69 years from the South Western part of greater Copenhagen area<sup>23</sup>. This study was designed to  
43 address chronic health issues, 3) The Health 2008 Study (Ethical committee approval number:  
44 KA20060011) is a cross sectional population based study<sup>24</sup>.  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### *Biochemical and anthropometric measurements*

The biochemical and anthropometric information on and phenotypical characterization of study participants is described in Table S1 and Table S2 and has been presented previously<sup>22-24</sup>.

### *Genotyping, variant calling and quality checks*

DNA extraction, genotyping and genotype call processing has been described previously<sup>21 25</sup>. A total of 8,445 individuals ( $n_{\text{Inter99}}:5,420$ ,  $n_{\text{Health2006}}:2,442$ ,  $n_{\text{Health2008}}:583$ ) with complete phenotype and genotype data participated for the serum TSH association analyses. The genotyping platform was Illumina Human Exome 12v1.0 containing 263,894 SNPs (including 16,024 custom SNPs identified from a recent exome sequencing study in Danes<sup>21 25</sup>) post quality control.

### *SNP Selection*

SNPs from the *VEGFA* gene region (6:43737946-6:43754224 GRCh37) and within the adjoining flanking region ( $\pm 75$  Kb) were selected covering a total of 166.2 kb (6:43662946-6:43829224 GRCh37) region. After removing SNPs with a  $\text{MAF} < 0.005$  a total of 15 non-coding SNPs were available from the *VEGFA* gene region for this study. Based on this SNP selection, a Bonferroni corrected  $p$  value corresponding 0.0033 was set as the significance threshold for SNP-TSH association testing.

### *SNP-TSH association testing and meta-analyses*

The association testing between the SNPs and the serum TSH levels was performed individually for each cohort using the additive linear regression model adjusting for gender, age and first five principal components as covariates. The fasting circulating measures of TSH were transformed to natural log scale before the association testing to control for non-normalised data. Prior to the association testing, individuals with known thyroid pathologies and those with out of range TSH values ( $< 0.4$  mIU/L and  $> 4.0$  mIU/L) were removed. The method for meta-analyses was as previously described<sup>21</sup>.

### *SNP metabolic traits association*



1  
2  
3 Associations between a SNP and metabolic traits were tested among normal glucose tolerant (NGT)  
4 individuals from the Inter99 cohort using the general linear model assuming an additive genetic  
5 effect for the SNP. Association was tested with baseline measures and changes during follow-up ( $\Delta$   
6 values: *Follow up – baseline measurements*) and was adjusted for gender and age. A  $p$ -value of  
7  $p < 0.05$  was considered significant for the follow-up and single SNP-metabolic trait testing. All  
8 association analyses were performed using PLINK v1.07  
9 (<http://pngu.mgh.harvard.edu/purcell/plink/>), and R version 3.1.1 (<http://www.r-project.org/>).  
10  
11  
12

### 13 *Linkage Disequilibria (LD) estimations*

14 LD estimation and proxy search was performed using 1000 genomes project data and an LD heat-  
15 map depicting pairwise  $r^2$  values is available in Fig. S1. Non-genotyped SNPs in LD with TSH-  
16 associated SNPs were retrieved from the ENSEMBL genome browser and SNPs with  $LD \geq 0.8$  were  
17 inspected for evidence of differential allele effects using Genome Browser  
18 (<https://genome.ucsc.edu/>) (Table S3).  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30

### 31 ***In vitro* molecular biology studies**

#### 32 *Reporter gene analysis*

33 The *VEGFA* proximal promoter (hg19, chr6:43737097-43738057), corresponding to 849bp of  
34 promoter and 252bp of the first intron, was cloned into pGL4.10. Genomic regions containing  
35 rs881858 and rs9472138 were amplified by PCR from homozygous carriers and cloned into *VEGFA*  
36 pro/GL4.10 down-stream of the *luc* gene to generate rs88 A/GL4.10, rs88 G/GL4.10, rs94 C/GL4.10  
37 and rs94 T/GL4.10. All constructs were confirmed by sequencing. Plasmid DNA was prepared using  
38 Qiagen Maxi Prep kit (Qiagen, Copenhagen Ø, Denmark) and ethanol precipitated. Transfections  
39 were made in human embryonic kidney cells (HEK)-293 cells (American Tissue Type Culture  
40 Collection, Rockville, MD, U.S.A.) using polyethylene imine (PEI25). Cells were harvested after  
41 24hrs for luciferase assays (Dual Light, Thermofisher Scientific, Copenhagen Ø, Denmark).  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Expression vectors for C/EBP $\beta$  and CHOP10 were a gift from Peter Johnson (Addgene plasmid  
4 #12557) and David Ron (Addgene plasmid # 21899), respectively.  
5  
6

#### 7 *Electrophoretic mobility shift assay (EMSA)*

8  
9 Nuclear extracts from HEK293 cells were prepared as described previously<sup>26</sup>. Some extracts were  
10 prepared following incubation of cells with 1mM dithiothreitol (DTT) for 16hrs to induce CHOP10  
11 via the unfolded protein response<sup>27</sup>. Complementary oligos representing the SNPs rs881858 were  
12 annealed and labelled with  $\alpha$ -<sup>32</sup>P-dATP (3000 Ci/mmol) by Klenow fill-in and purified using NICK  
13 columns (GE Healthcare, Brøndby, Denmark). Binding reactions were made as described  
14 previously<sup>26</sup>, separated with non-denaturing polyacrylamide gel-electrophoresis and visualized on  
15 phosphor-imager screen. Screens were scanned using a Molecular Dynamics Storm Scanner and the  
16 protein/DNA complexes analyzed using Image-Quant Software version 3.5. Oligonucleotides used  
17 for cloning and EMSA are listed in Table S3.  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

#### 29 *Statistics for molecular biology experiments*

30  
31 Results are expressed as mean value  $\pm$  SEM. Statistical analysis was performed in GraphPad Prism  
32 software. Effects of SNP constructs were tested using ANOVA with post-hoc t-test and Bonferroni  
33 correction. Differences between treatments were considered significant at a P-value<0.05 (two-  
34 tailed).  
35  
36  
37  
38  
39  
40  
41  
42

## 43 **RESULTS**

### 44 *VEGFA and TSH association analyses*

45  
46 We investigated three population-based Danish cohort studies for association analyses of the  
47 VEGFA locus with circulating TSH: The Inter99, the Health2006 and Health2008 cohorts<sup>23 24</sup> (Table  
48 S1). We searched for genotyped SNPs located up to 75kb up- and downstream of the *VEGFA*  
49 transcription start site (TSS) (*VEGFA*: Chr6:43737946-43754224). This region was defined based on  
50 the localization of chromatin marks within this region (Chromatin Interaction Analysis by paired-end  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

sequencing, ChIA-PET indicating interaction with the *VEGFA* gene and localization of nearby transcripts MRPS18A (upstream of VEGFA) and LOC100132354 and C6orf223 downstream of VEGFA) and we identified 15 SNPs in this region (Table 1, Table S4).

Confidential: For Review Only

**Table 1: *VEGFA* SNPs significantly associated with circulating levels of thyroid stimulating hormone (TSH)**

SNP name	Position (build 37/hg19)	Location wrt <i>VEGFA</i>	Alleles (effect/ other)	EAF	Inter99 <i>n</i> =5,420		Health2006 <i>n</i> =2,442		Health2008 <i>n</i> =583		Combined		
					Effect	<i>P</i>	Effect	<i>P</i>	Effect	<i>P</i>	<i>N</i>	<i>P</i>	<i>I</i> <sup>2</sup> ( <i>P</i> <sub>HET</sub> )
#rs9472138*	43,811,762	Downstream	T/C	0.27	-0.046	$4.4 \times 10^{-6}$	-0.051	$1.0 \times 10^{-3}$	-0.041	0.15	8,443	$5.9 \times 10^{-9}$	0 (0.94)
#rs881858*	43,806,609	Downstream	G/A	0.28	-0.046	$2.9 \times 10^{-6}$	-0.047	$1.9 \times 10^{-3}$	-0.038	0.18	8,445	$8.4 \times 10^{-9}$	0 (0.96)
#rs943080**	43,826,672	Downstream	T/C	0.51	-0.024	0.0053	-0.012	0.35	-0.040	0.12	8,440	0.0016	0 (0.59)
#rs4711751**	43,828,582	Downstream	T/C	0.51	-0.023	0.0089	-0.012	0.32	-0.043	0.09	8,402	0.002	0 (0.58)

#SNPs in LD ( $r^2 > 0.4$ ). \*SNPs in strong LD ( $r^2: 0.73, D': 0.94$ ). \*\*SNPs in strong LD ( $r^2: 1.0, D': 1.0$ ). EAF: Effect allele frequency. *VEGFA*: Vascular endothelial growth factor A. *I*<sup>2</sup>: heterozygosity at meta-analyses level. *P*<sub>HET</sub>: *P* value for heterozygosity. List of SNPs not reaching study wide association with circulating TSH-levels is given in Suppl. Table S4.

1  
2  
3  
4 Four *VEGFA* region SNPs (rs9472138, rs881858, rs943080, and rs4711751) associated with fasting  
5 serum TSH levels at a study-wide significance level ( $P_{combined} < 0.0033$ ), (Table 1, Fig. 1A and Fig.  
6 S2) in up to 8445 individuals following combined-meta analysis of the three cohorts. All four SNPs  
7 were in LD with each other ( $r^2 > 0.4$ ), and one SNP was a known signal (*VEGFA* rs9472138)<sup>14</sup> for  
8 circulating TSH. All the significantly associating SNPs were common (MAF > 0.05).  
9

#### 15 ***Genomic marks qualifying VEGFA SNPs for further investigation***

16  
17 Common SNPs in high LD ( $r^2 > 0.8$ ) (proxy SNPs) with the four significant SNPs in *VEGFA*:  
18 rs9472138, rs881858, rs943080 and rs4711751 was obtained from the ENSEMBL browser yielding  
19 12 common SNPs linked with rs9472138 and rs881858 and 3 SNPs with rs943080 and rs4711751  
20 (Table S5, Fig. S1), previously genotyped in the 1000genomes project. These SNPs were considered  
21 functional candidates to explain the observed genetic association, because they are common and in  
22 high LD with the lead SNPs. Using ENCODE data-tracks on Genome Browser (GRCh37/hg19)<sup>28</sup>  
23 these SNPs were evaluated based on presence of open chromatin structure (DNase seq),  
24 conservation, marks of H3K27Ac (Histone 3, Lysine 27 acetylation) and H3K4Me1 (Histone 3,  
25 Lysine 4 mono-methylation) and indication of protein binding to the SNP region by ChIP-seq.  
26 Furthermore, it was also assessed if the SNP altered the binding site for factors shown to bind the  
27 region by ChIP (Table S5, Fig.1B). From this, it was evident that rs881858 was highly conserved  
28 with G being the ancestral allele. Furthermore, rs881858 was located in a region showing marks  
29 characteristic of active regulatory elements (*H3K27Ac* and *H3K4Me1* and displaying evidence of  
30 close three dimensional proximity to the *VEGFA* promoter (Chromosome Interaction Analysis-  
31 Paired End Tags (ChIA-PET). The region was also DNase hypersensitive suggesting an open  
32 chromatin structure (Fig. 1B).  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

52 Moreover, rs881858 was directly located in a site bound by the transcription factor *c/EBPβ* (encoded  
53 by *CEBPB*) (by ChIP-seq) in multiple cell lines, and predicted to be bound by CHOP. Thus, for  
54 rs881858 there is strong evidence of an allele-specific regulatory role. Performing the same analysis  
55  
56  
57  
58  
59  
60

--

1  
2  
3 for the other SNPs did not reveal equal evidence of regulatory activity, conservation or protein  
4 binding to their vicinity (Table S5, Fig. S3). Rs881858 is located at the 3' end of the c/EBP $\beta$  binding  
5 site BS in a position, which does not confer specificity to the c/EBP $\beta$  binding according to the  
6 position weight matrix (PWM) for c/EBP $\beta$  (Fig. 1C)<sup>29</sup>, whereas the BS for CHOP is predicted to  
7 prefer the A-allele of the rs881858 compared with the G-allele. Thus, based on ENSEMBL and  
8 ENCODE data and differential predicted binding affinities to c/EBP $\beta$  and CHOP, the A and G alleles  
9 of rs881858 could confer differential responses.  
10  
11  
12  
13  
14  
15  
16  
17  
18

19  
20 *Association of G-allele of VEGFA rs881858 with decreased circulating TSH, increased thyroid*  
21 *hormone levels and metabolic traits*  
22  
23

24  
25 The effect size of rs881858 on circulating TSH levels was -0.092 ( $P=2.2\times 10^{-11}$ ) for the G-allele  
26 corresponding to an additive 2.3% decrease in TSH level per allele. Furthermore, free T4 levels were  
27 correspondingly increased in GG subjects compared with AA or AG subjects (GG: 15.4 (14.2-16.6)  
28 vs AA: 15.0 (13.8-16.3) pmol/L,  $P=0.0014$ ), indicating a slightly increased thyroid function in GG  
29 individuals and an altered set point for the TSH/T4 axis (Table 2). There was no available  
30 information on circulating levels of total T3 or thyroid hormone binding globulin for these subjects.  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Table 2: Associations between *VEGFA* rs881858 G-allele and metabolic traits in 4,373 normal glucose tolerant (NGT) Danish subjects (Inter99) at baseline**

<i>VEGFA</i> rs881858	GG	GA	AA	*Effect	SE	<i>P</i>
<i>n</i>	347	1,754	2,272	-	-	-
Age (years)	45 (40-50)	45 (40-50)	45 (40-50)	-	-	-
TSH (mIU/L)	1.11 (0.77-1.56)	1.23 (0.88-1.75)	1.29 (0.90-1.85)	-0.092	0.013	<b>2.2 × 10<sup>-11</sup></b>
Free T4 (pmol/L)	15.4 (14.2-16.6)	15.0 (13.9-16.3)	15.0 (13.8-16.3)	0.009	0.002	<b>1.4 × 10<sup>-3</sup></b>
BMI (kg/m <sup>2</sup> )	25.1 (22.8-28.2)	24.9 (22.8-27.6)	24.9 (22.7-27.5)	0.14	0.095	0.11
Leptin (ng/ml)	5.5 (2.5-10.9)	5.3 (2.4-11.0)	5.3 (2.4-11.3)	0.01	0.022	0.62
HbA1c (%)	5.8 (5.5-6.1)	5.8 (5.5-6.0)	5.8 (5.5-6.0)	0.016	0.008	0.056
Fasting plasma glucose (mmol/L)	5.4 (5.1-5.6)	5.3 (5.0-5.6)	5.3 (5.0-5.6)	0.021	0.008	<b>0.016</b>
2-hour glucose during OGTT (mmol/L)	5.6 (4.8-6.4)	5.6 (4.7-6.4)	5.5 (4.7-6.3)	0.072	0.026	<b>6.1 × 10<sup>-3</sup></b>
Fasting serum insulin (pmol/L)	32 (23-47.5)	31 (22-45)	31 (21-44)	0.023	0.013	0.073
2-hour insulin during OGTT (pmol/L)	133 (86-221)	131 (83-202)	134 (83-204)	0.013	0.017	0.43
ISI <sub>Matsuda</sub>	3.03 (2.2-4.2)	3.15 (2.24-4.44)	3.19 (2.25-4.47)	-0.02	0.012	0.12
HOMA-IR	1.25 (0.92-1.92)	1.20 (0.83-1.81)	1.20 (0.82-1.76)	0.027	0.013	<b>0.041</b>
Insulinogenic index	85.2 (53.3-137.2)	76.0 (47.6-125.1)	77.0 (49.7-129.2)	0.013	0.017	0.45
Disposition index	227.0 (169.3-339.8)	226.0 (159.6-341.0)	231.6 (162.7-352.7)	-0.006	0.015	0.68

Values correspond to median (interquartile range) in non-transformed traits. \* G allele as the effect allele assuming an additive genetic model for log transformed traits. ISI Matsuda, HOMA-IR, Insulinogenic index and Disposition index were calculated as described in Supplementary Table S3. *P* values in bold indicate significant results ( $P < 0.05$ ).

Since increased circulating TSH levels are associated with an impaired metabolic phenotype, it was tested if rs881858 associated with measures of glucose tolerance. The GG genotype of rs881858 was associated with slightly increased fasting and 2-hour post OGTT plasma glucose levels ( $P = 0.016$  and  $6.1 \times 10^{-3}$ ) (Table 2) and increased HOMA-IR values ( $P = 0.041$ ) among glucose-tolerant subjects. Body mass index, HbA1c and circulating leptin levels were not associated with carrier-status of the rs881858. Moreover, we investigated measures of insulin release derived from OGTT data in relation to rs881858, but the insulinogenic index and the disposition index were not different between genotypes (Table 2).

1  
2  
3 *Association of G-allele of VEGFA rs881858 with fasting circulating TSH, thyroid hormone levels*  
4  
5 *and metabolic traits after 5-yr follow-up*  
6  
7

8  
9 We studied changes in metabolic traits over the 5 yr follow-up period were studied among glucose  
10 tolerant individuals from the Inter99 cohort (Table S2, Table S6). Glucose (HbA1c: 0.66% and  
11 fasting plasma glucose: 0.86%) and insulin measures (fasting serum insulin: 3.2%; HOMA-IR:  
12 4.1%) improved over a mean follow-up time of 5.4 years (Table S2). However, none of these  
13 measures were associated with the *VEGFA* rs881858 polymorphism. Moreover, changes in BMI,  
14 circulating TSH and T4 levels were also not associated with *VEGFA* rs881858 (Table S6). This  
15 indicates that the variant could act by modulating a given set point for TSH, since the phenotype of  
16 carriers appears to be stable during follow-up.  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26

27 *Reporter-gene analysis of VEGFA rs881858 and rs9472138 alleles*  
28  
29

30  
31 To determine if the most significant TSH associated *VEGFA* SNPs rs881858 or rs9472138 could  
32 confer functional changes to the *VEGFA* promoter activity, luciferase reporter vectors representing  
33 the SNPs were tested by transfection in HEK293 cells followed by luciferase assays (Fig. 2). Basal  
34 activity of the *VEGFA* minimal promoter was very high in HEK293 cells (not shown). The activity  
35 of the rs881858 A allele was significantly higher than the G-allele (P=0.0012) (Fig. 2A), while there  
36 was no difference between the C and the T-allele of *VEGFA* rs9472138.  
37  
38  
39  
40  
41  
42  
43  
44

45 The A-allele of *VEGFA* rs881858 is predicted to create a novel binding site for the transcription  
46 factor CHOP. To test the response of this site to CHOP, HEK293 cells were transfected with  
47 increasing amounts of CHOP expression vector in the presence of either rs881858 A- or G-allele  
48 reporter vector (Fig. 2B). With no CHOP over-expression the *VEGFA* A-allele had increased activity  
49 compared with the G-allele (as in Fig. 2A). Low amounts of CHOP expression vector increased the  
50 G-allele reporter activity, while there was no difference in the activity of the A-allele. Furthermore,  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 increasing amounts of CHOP activity resulted in significantly decreased activity of the A-allele,  
4 indicating that the activity of this site is repressed by CHOP. Thus, the main action of CHOP on the  
5 predicted binding site created by rs881858 A was to decrease reporter gene activity, which is in line  
6 with CHOP being a transcriptional repressor<sup>30</sup>. When over-expressing c/EBP $\beta$  both the *VEGFA* A-  
7 and the G-alleles of rs881858 responded by increasing luciferase activities 25% (P<0.05) at low  
8 levels of c/EBP $\beta$  while decreasing at higher amounts of c/EBP $\beta$  (Fig. 2C). Since the rs881858 is not  
9 located in the core binding site of c/EBP $\beta$  (Fig. 1B), this is compatible with an equal response to  
10 c/EBP $\beta$  by either SNP allele.  
11  
12  
13  
14  
15  
16  
17  
18  
19

20  
21 The interaction between different amounts of c/EBP $\beta$  and CHOP was tested by co-transfection  
22 experiments in which varying ratios of c/EBP $\beta$  and CHOP were used with reporter vectors. For all  
23 combinations of CHOP in the presence of c/EBP $\beta$ , CHOP repressed the A-allele luciferase activity,  
24 while having no effect on the G-allele (Fig. 2D). Thus, based on reporter-gene assays, the A-allele of  
25 *VEGFA* rs881858 creates a novel response element of CHOP effectively repressing the minimal  
26 promoter activity of *VEGFA*, while this has no effect on the response to c/EBP $\beta$ . Furthermore, the A-  
27 allele confers higher reporter-gene activity in the basal state compared to the G-allele.  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38

### 39 ***Binding affinities of VEGFA rs881858 (A/G) alleles***

40  
41 Electrophoretic mobility shift assays (EMSA) were made assess *in vitro* transcription factor binding  
42 of *VEGFA* rs881858. Double-stranded (ds)-oligos representing the A-allele and the G-allele and  
43 encompassing both the c/EBP $\beta$  and the CHOP binding sites were compared with known c/EBP $\beta$  and  
44 the CHOP binding sites. Labeled 'A' and 'G' oligos formed two strong binding complexes (lane 1  
45 and 6) of which the double band could be removed by competition (COMP) with unlabelled A and  
46 G-oligos, as well as with un-labeled CHOP (lane 2 and 9) and c/EBP $\beta$  (lane 5) oligos (Fig. 3A). This  
47 complex contains CHOP as well as c/EBP $\beta$  protein, because of the efficient competition by  
48 corresponding unlabelled oligos, and the reduction in band intensity, when adding C/EBP $\beta$  antibody  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 to the binding reaction (lane 10, 'supershift'). Moreover, the intensity of the CHOP complex was  
4  
5 increased for the A oligo compared with G, indicating stronger binding of the A probe to CHOP  
6  
7 protein (Fig. 3A, lane 1 vs. lane 6, Fig. 3B, lane 5 vs. 6 and 9 vs. 10). One complex was specific for  
8  
9 A and G oligos as these efficiently out-compete the radio-labeled probes (Fig. 3A and 3B, arrow),  
10  
11 whereas competition using oligos containing known CHOP or c/EBP $\beta$  sites did not remove the  
12  
13 complex (Fig. 3A, lane 1 vs. 2, 5 vs. 6, 5 vs. 9 and Fig. 3B, lane 1 vs. 2). Thus, both A and G  
14  
15 versions of rs881858 can bind CHOP and c/EBP $\beta$  *in vitro*.  
16  
17

18  
19 We also compared A and G oligos with oligos representing known CHOP or C/EBP $\beta$  binding sites<sup>30</sup>  
20  
21 <sup>31</sup> (Fig. 3B). CHOP and C/EBP $\beta$  oligos form complexes with the same mobility shift as A and G  
22  
23 oligos. The same lower complex formed with either the CHOP or C/EBP $\beta$  probe, consistent with  
24  
25 CHOP and C/EBP $\beta$  forming heterodimers. UPR induction increased C/EBP $\beta$ , A and G complex  
26  
27 quantity (Fig. 3B). Furthermore, excess unlabelled CHOP oligo efficiently removed the VEGFA A  
28  
29 and G probe binding at the lower complex, but not the top complex (arrow, Fig. 3B). The A oligo  
30  
31 consistently formed more lower complex than the G oligo, indicating increased binding strength of  
32  
33 this probe. The top complex, specific for the rs881858 site (Fig. 3A and 3B, arrows) suggests that  
34  
35 additional proteins may bind the *VEGFA* rs881858 site.  
36  
37

38  
39 The c/EBP $\beta$  probe formed a faint complex (lane 3 and 7 Fig. 3B, lanes 1-3, Fig. 3C), most likely  
40  
41 c/EBP $\beta$  homo-dimer, because it was super-shifted with c/EBP $\beta$  antibody and competed with un-  
42  
43 labeled c/EBP $\beta$  probe. The known CHOP binding site of the *TRIB3* promoter was used as probe for  
44  
45 EMSA (Fig. 3C, lane 4-10). As expected the formed complex formed contained CHOP, shown by  
46  
47 efficient competition by unlabelled CHOP ds-oligo (lane 4, Fig. 3C) and decreased complex  
48  
49 formation with addition of CHOP antibody (lane 5, Fig. 3C). However, since excess of A or G did  
50  
51 not remove *TRIB3* probe binding, this indicates that the investigated *VEGFA* binding site has lower  
52  
53 affinity for CHOP than the *TRIB3* site with no difference between the A-allele and the G-allele of  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 the rs881858. In summary, the *VEGFA* rs881858 site forms a binding site for both CHOP and  
4  
5 c/EBP $\beta$ , where the A-allele forms more CHOP complex compared with the G-allele, but where this  
6  
7 site measured against a well-characterized CHOP binding site has lower affinity for CHOP.  
8  
9

10 Public eQTL databases were search in order to determine, if base-line thyroid *VEGFA* mRNA levels  
11  
12 correlated with the SNPs significantly associated with TSH. Data from the GTEX database<sup>32</sup> is  
13  
14 shown in Figure S4 indicating that base-line thyroid *VEGFA* transcript levels do not depend on  
15  
16 genotypes of the SNPs rs881858, rs9472138, rs943080 or rs4711751.  
17  
18  
19  
20  
21  
22

## 23 DISCUSSION AND CONCLUSIONS

24  
25  
26 We focused on the genetic and molecular characterization of the *VEGFA* locus with the aim to  
27  
28 identify the functional variant(s) explaining the association with circulating TSH levels<sup>14-17</sup>. We  
29  
30 performed a dense association-mapping for *VEGFA* SNPs using data from 8445 Danish individuals,  
31  
32 and identified 4 SNPs significantly associated with circulating TSH levels, all of which were located  
33  
34 more than 50kb 3' of the coding region. Among the top hits we identified *VEGFA* rs881858, in high  
35  
36 LD to the *VEGFA* GWAS SNP rs9472138<sup>14</sup>. The carriers of *VEGFA* rs881858 G allele had a 2.3%  
37  
38 decrease in circulating TSH levels in an additive manner. Moreover, the identified SNP rs881858  
39  
40 provided functional evidence of allele specific effects at a *VEGFA* regulatory region binding CHOP  
41  
42 and C/EBP $\beta$  proteins, thus connecting cellular stress activated pathways with *VEGFA* gene  
43  
44 regulatory activity and thyroid function, because CHOP production is activated by several types of  
45  
46 cellular stress, such as endoplasmic reticulum stress, nutrient deprivation or oxidative stress<sup>33</sup>.  
47  
48  
49

50  
51 Thus, data here and elsewhere<sup>14</sup> support that common variation in the *VEGFA* locus is an important  
52  
53 determinant of circulating TSH levels. For the *VEGFA* locus, the variants most highly associated  
54  
55 with TSH, rs881858 and rs9472138 are located >50kb 3' of the *VEGFA* coding region, and equally  
56  
57  
58  
59  
60

1  
2  
3 close to a long non-coding RNA (Loc100132354), the function of which is not characterized.  
4  
5 However, chromosomal interaction analysis identified contact points between the rs881858 SNP  
6  
7 region and the promoter of *VEGFA* indicating control of *VEGFA* gene activity (Fig. 1A).  
8  
9

10  
11 *VEGFA* is important for angiogenesis, homeostatic responses and organ growth in multiple tissues or  
12 cell types (white adipose tissue<sup>34 35</sup>, islets of Langerhans<sup>36 37</sup>), and has also been shown to control T4  
13 to T3 conversion in hypothalamic tanycytes and thereby the feedback control of thyroid hormones to  
14 TRH and TSH release<sup>38 39</sup>. *VEGFA* rs881858 is also a known GWAS locus for chronic kidney  
15 disease (CKD) and kidney function<sup>40</sup> suggesting a possibility that this SNP controls responses in  
16 multiple organs, also as *VEGFA* is expressed ubiquitously. Since T4 levels were correspondingly  
17 increased in G-allele carriers, whose TSH levels were decreased, rs881858 seem to act primarily on  
18 the thyroid gland.  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

29 The G-allele is associated with increased *VEGFA* response to TSH, an increased T4 release, and  
30 presumably increased T3 conversion in the hypothalamus and therefore resulting in decreased TSH  
31 release as a hypothalamic-pituitary response. However, it is also a possibility that the action of  
32 rs881858 can be on both *VEGFA* expression in the thyroid gland as well as on the hypothalamic  
33 tanycytes. Consistent with CHOP being a repressor<sup>41</sup> A-allele carriers have lower T4 and increased  
34 TSH: For A-carriers the induction of ER-stress may result in an impaired response to stimulate  
35 thyroid growth via *VEGFA* due to increased CHOP binding (Fig.4). Since thyroid VEGF is increased  
36 by TSH<sup>19</sup>, this is consistent with rs881858 being more important in thyroid than in hypothalamus.  
37 Moreover, CHOP deficient mice have increased angiogenesis, showing that CHOP normally acts to  
38 limit angiogenesis<sup>42</sup>.  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

51  
52 Interestingly, the genomic region containing rs881858 and rs9472138 is only modestly associated  
53 with levels of circulating VEGF, while regions situated both 5' of the *VEGFA* coding regions and  
54 further 3' in the C6orf223 locus are highly associated with circulating VEGF levels<sup>43 44</sup>. There is  
55  
56  
57  
58  
59  
60

1  
2  
3 little or no LD between these regions and the LD block harbouring the SNPs associated with  
4 circulating TSH in the current study (rs881858, rs94772138, rs943080 and rs4711751) and these  
5 signals appear to be independent. Thus, it is conceivable that several regions in or near the *VEGFA*  
6 locus control different aspects of VEGF regulation and release. It seems reasonable that regulatory  
7 mechanisms operating in thyroid tissue to govern for example stress induced VEGF production could  
8 differ from the molecular mechanisms contributing to the regulation of circulating VEGF levels.  
9 Although we do not know the contribution of the thyroid gland to the sVEGF pool it is likely to be  
10 only a minor contributor compared with endothelial-released VEGF.  
11  
12  
13  
14  
15  
16  
17  
18  
19

20  
21 Although we present evidence for differential effects of alleles of *VEGFA* rs881858 by both reporter  
22 assays and direct binding, and have retrieved and inspected SNPs with high LD to this allele; we  
23 have not performed an extensive investigation of all linked SNPs in the *VEGFA* region for evidence  
24 of regulatory activity, and thus cannot exclude that additional functional SNPs may exist. Another  
25 limitation of our study is that we have no available data to show association between rs881858  
26 alleles and levels of *VEGFA* mRNA transcript, protein levels or TSH stimulated VEGF release. Data  
27 from public eQTL databases shows no association between the investigated SNPs and thyroid  
28 *VEGFA* mRNA levels (Fig. S4).  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39

40  
41 When examining metabolic traits association with *VEGFA* rs881858 GG homozygous subjects were  
42 slightly more insulin resistant, while having similar measures of obesity and similar insulin secretion  
43 capacity (Table 2). Adipose tissue-specific knock-out of *Vegfa* results in inability to expand the  
44 adipose tissue, when demands for fat storage increases, exemplified by high-fat feeding<sup>34</sup>. Our results  
45 indicate that in the human setting the A-allele of *VEGFA* rs881858 via CHOP binding generates  
46 lower reporter-gene activity suggesting a decreased *VEGFA* response to cellular stress. This would  
47 result in an impaired angiogenic response of the A-allele, which is consistent with increased TSH  
48 levels. However, this is seemingly at odds with the GG-homozygous subjects being more insulin  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 resistant, because in mouse models insulin resistance is observed when the angiogenic response and  
4 therefore adipogenesis is impaired<sup>34 35</sup>. Thus, further studies are necessary to determine the tissue-  
5 specific effects of VEGFA and genetic variation on different human tissues. The *VEGFA* SNP  
6 rs9472138 has previously been associated with visceral obesity and insulin resistance in women<sup>45</sup>,  
7 further underlining the importance of genetic variation in *VEGFA* also for human adipose tissue  
8 expansion. Of note, increased circulating TSH is associated with obesity and impaired cardio-  
9 metabolic health, which could indirectly affect the observed associations between insulin resistance  
10 and *VEGFA* rs881858. Mendelian randomization studies in large population based cohort could  
11 potentially resolve this.  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26

## 27 ACKNOWLEDGMENTS

28  
29  
30 This study was supported by the Danish Medical Research Council (grant DFF – 1331-00033), The Novo  
31 Nordisk Foundation, the Danish Diabetes Association, the Gerda and Aage Haensch Foundation, the A.P.  
32 Møller Foundation for the Advancement of Medical Science, Novo Nordisk A/S, Roskilde University and  
33 University of Copenhagen. The Novo Nordisk Foundation Center for Basic Metabolic Research is an  
34 independent Research Center at the University of Copenhagen partially funded by an unrestricted donation  
35 from the Novo Nordisk Foundation ([www.metabol.ku.dk](http://www.metabol.ku.dk)). Further, this work is supported by the European  
36 Commission as an integrated project under the 6<sup>th</sup> Framework Programme (LSHM-CT-2005-018734,  
37 HepAdip). The Danish Obesity Research Centre (DanORC; [www.danorc.dk](http://www.danorc.dk)) is supported by The Danish  
38 Council for Strategic Research (grant 2101-06-0005). TSA was funded by the Gene Diet Interactions in  
39 Obesity (GENDINOBS, [www.gendinob.dk](http://www.gendinob.dk)) project and acknowledges the same. The authors wish to thank  
40 Christa Persson, Annemette Forman, Tina Lorentzen, and Marianne Stendal for technical assistance, Arne  
41 Nielsen for data management, and Grete Lademann and Trine Toldsted for secretarial support.  
42  
43

## 44 References

- 45  
46 1. Laurberg P, Andersen S, Carle A, Karmisholt J, Knudsen N, Pedersen IB. The TSH upper reference limit:  
47 where are we at? *Nat Rev Endocrinol* 2011;**7**(4):232-9.
- 48 2. Andersen S, Pedersen KM, Bruun NH, Laurberg P. Narrow individual variations in serum T(4) and T(3) in  
49 normal subjects: a clue to the understanding of subclinical thyroid disease. *J Clin Endocrinol Metab*  
50 2002;**87**(3):1068-72.
- 51 3. Hansen PS, Brix TH, Bennedbaek FN, Bonnema SJ, Kyvik KO, Hegedus L. Genetic and environmental  
52 causes of individual differences in thyroid size: a study of healthy Danish twins. *J Clin Endocrinol*  
53 *Metab* 2004;**89**(5):2071-7.
- 54 4. Hansen PS, Brix TH, Sorensen TI, Kyvik KO, Hegedus L. Major genetic influence on the regulation of the  
55 pituitary-thyroid axis: a study of healthy Danish twins. *J Clin Endocrinol Metab* 2004;**89**(3):1181-7.  
56  
57  
58  
59  
60

5. Panicker V, Wilson SG, Spector TD, Brown SJ, Falchi M, Richards JB, Surdulescu GL, Lim EM, Fletcher SJ, Walsh JP. Heritability of serum TSH, free T4 and free T3 concentrations: a study of a large UK twin cohort. *Clin Endocrinol (Oxf)* 2008;**68**(4):652-9.
6. Taylor PN, Razvi S, Pearce SH, Dayan CM. Clinical review: A review of the clinical consequences of variation in thyroid function within the reference range. *J Clin Endocrinol Metab* 2013;**98**(9):3562-71.
7. de Moura Souza A, Sichieri R. Association between serum TSH concentration within the normal range and adiposity. *Eur J Endocrinol* 2011;**165**(1):11-5.
8. Chikunguwo S, Brethauer S, Nirujogi V, Pitt T, Udomsawaengsup S, Chand B, Schauer P. Influence of obesity and surgical weight loss on thyroid hormone levels. *Surg Obes Relat Dis* 2007;**3**(6):631-5; discussion 35-6.
9. Iacobellis G, Ribaldo MC, Zappaterreno A, Iannucci CV, Leonetti F. Relationship of thyroid function with body mass index, leptin, insulin sensitivity and adiponectin in euthyroid obese women. *Clin Endocrinol (Oxf)* 2005;**62**(4):487-91.
10. Sakurai M, Nakamura K, Miura K, Yoshita K, Takamura T, Nagasawa SY, Morikawa Y, Ishizaki M, Kido T, Naruse Y, Nakashima M, Nogawa K, Suwazono Y, Nakagawa H. Association between a serum thyroid-stimulating hormone concentration within the normal range and indices of obesity in Japanese men and women. *Intern Med* 2014;**53**(7):669-74.
11. Knudsen N, Laurberg P, Rasmussen LB, Bulow I, Perrild H, Ovesen L, Jorgensen T. Small differences in thyroid function may be important for body mass index and the occurrence of obesity in the population. *J Clin Endocrinol Metab* 2005;**90**(7):4019-24.
12. Asvold BO, Bjoro T, Vatten LJ. Association of serum TSH with high body mass differs between smokers and never-smokers. *J Clin Endocrinol Metab* 2009;**94**(12):5023-7.
13. Fox CS, Pencina MJ, D'Agostino RB, Murabito JM, Seely EW, Pearce EN, Vasan RS. Relations of thyroid function to body weight: cross-sectional and longitudinal observations in a community-based sample. *Arch Intern Med* 2008;**168**(6):587-92.
14. Porcu E, Medici M, Pistis G, Volpato CB, Wilson SG, Cappola AR, Bos SD, Deelen J, den Heijer M, Freathy RM, Lahti J, Liu C, Lopez LM, Nolte IM, O'Connell JR, Tanaka T, Trompet S, Arnold A, Bandinelli S, Beekman M, Bohringer S, Brown SJ, Buckley BM, Camaschella C, de Craen AJ, Davies G, de Visser MC, Ford I, Forsen T, Frayling TM, Fugazzola L, Gogele M, Hattersley AT, Hermus AR, Hofman A, Houwing-Duistermaat JJ, Jensen RA, Kajantie E, Kloppenburg M, Lim EM, Masciullo C, Mariotti S, Minelli C, Mitchell BD, Nagaraja R, Netea-Maier RT, Palotie A, Persani L, Piras MG, Psaty BM, Raikonen K, Richards JB, Rivadeneira F, Sala C, Sabra MM, Sattar N, Shields BM, Soranzo N, Starr JM, Stott DJ, Sweep FC, Usala G, van der Klauw MM, van Heemst D, van Mullem A, Vermeulen SH, Visser WE, Walsh JP, Westendorp RG, Widen E, Zhai G, Cucca F, Deary IJ, Eriksson JG, Ferrucci L, Fox CS, Jukema JW, Kiemeneij LA, Pramstaller PP, Glessinger D, Shuldiner AR, Slagboom EP, Uitterlinden AG, Vaidya B, Visser TJ, Wolffenbuttel BH, Meulenbelt I, Rotter JI, Spector TD, Hicks AA, Toniolo D, Sanna S, Peeters RP, Naitza S. A meta-analysis of thyroid-related traits reveals novel loci and gender-specific differences in the regulation of thyroid function. *PLoS Genet* 2013;**9**(2):e1003266.
15. Malinowski JR, Denny JC, Bielinski SJ, Basford MA, Bradford Y, Peissig PL, Carrell D, Crosslin DR, Pathak J, Rasmussen L, Pacheco J, Kho A, Newton KM, Li R, Kullo IJ, Chute CG, Chisholm RL, Jarvik GP, Larson EB, McCarty CA, Masys DR, Roden DM, de Andrade M, Ritchie MD, Crawford DC. Genetic variants associated with serum thyroid stimulating hormone (TSH) levels in European Americans and African Americans from the eMERGE Network. *PLoS One* 2014;**9**(12):e111301.
16. Zhan M, Chen G, Pan CM, Gu ZH, Zhao SX, Liu W, Wang HN, Ye XP, Xie HJ, Yu SS, Liang J, Gao GQ, Yuan GY, Zhang XM, Zuo CL, Su B, Huang W, Ning G, Chen SJ, Chen JL, Song HD, China Consortium for Genetics of Autoimmune Thyroid D. Genome-wide association study identifies a novel susceptibility gene for serum TSH levels in Chinese populations. *Hum Mol Genet* 2014;**23**(20):5505-17.
17. Kwak SH, Park YJ, Go MJ, Lee KE, Kim SJ, Choi HS, Kim TH, Choi SH, Lim S, Kim KW, Park do J, Kim SS, Lee JY, Park KS, Jang HC, Cho NH. A genome-wide association study on thyroid function and anti-thyroid peroxidase antibodies in Koreans. *Hum Mol Genet* 2014;**23**(16):4433-42.

- 1  
2  
3 18. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, Hodge CL, Haase J, Janes J, Huss JW, III, Su  
4 AI. BioGPS: an extensible and customizable portal for querying and organizing gene annotation  
5 resources. *Genome Biol* 2009;**10**(11):R130.
- 6 19. Sato K, Yamazaki K, Shizume K, Kanaji Y, Obara T, Ohsumi K, Demura H, Yamaguchi S, Shibuya M.  
7 Stimulation by thyroid-stimulating hormone and Grave's immunoglobulin G of vascular endothelial  
8 growth factor mRNA expression in human thyroid follicles in vitro and flt mRNA expression in the  
9 rat thyroid in vivo. *J Clin Invest* 1995;**96**(3):1295-302.
- 10 20. Viglietto G, Romano A, Manzo G, Chiappetta G, Paoletti I, Califano D, Galati MG, Mauriello V, Bruni P,  
11 Lago CT, Fusco A, Persico MG. Upregulation of the angiogenic factors PIGF, VEGF and their  
12 receptors (Flt-1, Flk-1/KDR) by TSH in cultured thyrocytes and in the thyroid gland of thiouracil-fed  
13 rats suggest a TSH-dependent paracrine mechanism for goiter hypervascularization. *Oncogene*  
14 1997;**15**(22):2687-98.
- 15 21. Ahluwalia TS, Allin KH, Sandholt CH, Sparso TH, Jorgensen ME, Rowe M, Christensen C, Brandslund I,  
16 Lauritzen T, Linneberg A, Husemoen LL, Jorgensen T, Hansen T, Grarup N, Pedersen O. Discovery  
17 of coding genetic variants influencing diabetes-related serum biomarkers and their impact on risk of  
18 type 2 diabetes. *J Clin Endocrinol Metab* 2015;**100**(4):E664-71.
- 19 22. Jorgensen T, Borch-Johnsen K, Thomsen TF, Ibsen H, Glumer C, Pisinger C. A randomized non-  
20 pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99.  
21 *Eur J Cardiovasc Prev Rehabil* 2003;**10**(5):377-86.
- 22 23. Thuesen BH, Cerqueira C, Aadahl M, Ebstrup JF, Toft U, Thyssen JP, Fenger RV, Hersoug LG, Elberling  
23 J, Pedersen O, Hansen T, Johansen JD, Jorgensen T, Linneberg A. Cohort Profile: the Health2006  
24 cohort, research centre for prevention and health. *Int J Epidemiol* 2014;**43**(2):568-75.
- 25 24. Byberg S, Hansen AL, Christensen DL, Vistisen D, Aadahl M, Linneberg A, Witte DR. Sleep duration  
26 and sleep quality are associated differently with alterations of glucose homeostasis. *Diabet Med*  
27 2012;**29**(9):e354-60.
- 28 25. Albrechtsen A, Grarup N, Li Y, Sparso T, Tian G, Cao H, Jiang T, Kim SY, Korneliussen T, Li Q, Nie C,  
29 Wu R, Skotte L, Morris AP, Ladenvall C, Cauchi S, Stancakova A, Andersen G, Astrup A, Banasik  
30 K, Bennett AJ, Bolund L, Charpentier G, Chen Y, Dekker JM, Doney AS, Dorkhan M, Forsen T,  
31 Frayling TM, Groves CJ, Gui Y, Hallmans G, Hattersley AT, He K, Hitman GA, Holmkvist J, Huang  
32 S, Jiang H, Jin X, Justesen JM, Kristiansen K, Kuusisto J, Lajer M, Lantieri O, Li W, Liang H, Liao  
33 Q, Liu X, Ma T, Ma X, Manijak MP, Marre M, Mokrosinski J, Morris AD, Mu B, Nielsen AA,  
34 Nijpels G, Nilsson P, Palmer CN, Rayner NW, Renstrom F, Ribel-Madsen R, Robertson N,  
35 Rolandsson O, Rossing P, Schwartz TW, Group DESIRS, Slagboom PE, Sterner M, Consortium D,  
36 Tang M, Tarnow L, Tuomi T, van't Riet E, van Leeuwen N, Varga TV, Vestmar MA, Walker M,  
37 Wang B, Wang Y, Wu H, Xi F, Yengo L, Yu C, Zhang X, Zhang J, Zhang Q, Zhang W, Zheng H,  
38 Zhou Y, Altshuler D, t Hart LM, Franks PW, Balkau B, Froguel P, McCarthy MI, Laakso M, Groop  
39 L, Christensen C, Brandslund I, Lauritzen T, Witte DR, Linneberg A, Jorgensen T, Hansen T, Wang  
40 J, Nielsen R, Pedersen O. Exome sequencing-driven discovery of coding polymorphisms associated  
41 with common metabolic phenotypes. *Diabetologia* 2013;**56**(2):298-310.
- 42 26. Troelsen JT, Mitchelmore C, Spodsberg N, Jensen AM, Noren O, Sjostrom H. Regulation of lactase-  
43 phlorizin hydrolase gene expression by the caudal-related homeodomain protein Cdx-2. *Biochem J*  
44 1997;**322** ( Pt 3):833-38.
- 45 27. Roybal CN, Yang S, Sun CW, Hurtado D, Vander Jagt DL, Townes TM, Abcouwer SF. Homocysteine  
46 increases the expression of vascular endothelial growth factor by a mechanism involving endoplasmic  
47 reticulum stress and transcription factor ATF4. *J Biol Chem* 2004;**279**(15):14844-52.
- 48 28. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome  
49 browser at UCSC. *Genome Res* 2002;**12**(6):996-1006.
- 50 29. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull  
51 M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E.  
52 TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic*  
53 *Acids Res* 2006;**34**(Database issue):D108-10.
- 54 30. Tang QQ, Lane MD. Role of C/EBP homologous protein (CHOP-10) in the programmed activation of  
55 CCAAT/enhancer-binding protein-beta during adipogenesis. *Proc Natl Acad Sci U S A*  
56 2000;**97**(23):12446-50.
- 57  
58  
59  
60



- 1  
2  
3 31. Ohoka N, Yoshii S, Hattori T, Onozaki K, Hayashi H. TRB3, a novel ER stress-inducible gene, is induced  
4 via ATF4-CHOP pathway and is involved in cell death. *EMBO J* 2005;**24**(6):1243-55.
- 5 32. Mele M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM,  
6 Pervouchine DD, Sullivan TJ, Johnson R, Segre AV, Djebali S, Niarchou A, Consortium GT, Wright  
7 FA, Lappalainen T, Calvo M, Getz G, Dermitzakis ET, Ardlie KG, Guigo R. Human genomics. The  
8 human transcriptome across tissues and individuals. *Science* 2015;**348**(6235):660-5.
- 9 33. Brewer JW. Regulatory crosstalk within the mammalian unfolded protein response. *Cell Mol Life Sci*  
10 2014;**71**(6):1067-79.
- 11 34. Sung HK, Doh KO, Son JE, Park JG, Bae Y, Choi S, Nelson SM, Cowling R, Nagy K, Michael IP, Koh  
12 GY, Adamson SL, Pawson T, Nagy A. Adipose vascular endothelial growth factor regulates  
13 metabolic homeostasis through angiogenesis. *Cell Metab* 2013;**17**(1):61-72.
- 14 35. Elias I, Franckhauser S, Ferre T, Vila L, Tafuro S, Munoz S, Roca C, Ramos D, Pujol A, Riu E, Ruberte J,  
15 Bosch F. Adipose tissue overexpression of vascular endothelial growth factor protects against diet-  
16 induced obesity and insulin resistance. *Diabetes* 2012;**61**(7):1801-13.
- 17 36. Brissova M, Aamodt K, Brahmachary P, Prasad N, Hong JY, Dai C, Mellati M, Shostak A, Poffenberger  
18 G, Aramandla R, Levy SE, Powers AC. Islet microenvironment, modulated by vascular endothelial  
19 growth factor-A signaling, promotes beta cell regeneration. *Cell Metab* 2014;**19**(3):498-511.
- 20 37. Reinert RB, Brissova M, Shostak A, Pan FC, Poffenberger G, Cai Q, Hundemer GL, Kantz J, Thompson  
21 CS, Dai C, McGuinness OP, Powers AC. Vascular endothelial growth factor-a and islet  
22 vascularization are necessary in developing, but not adult, pancreatic islets. *Diabetes*  
23 2013;**62**(12):4154-64.
- 24 38. Langlet F, Levin BE, Luquet S, Mazzone M, Messina A, Dunn-Meynell AA, Balland E, Lacombe A,  
25 Mazur D, Carmeliet P, Bouret SG, Prevoit V, Dehouck B. Tanycytic VEGF-A boosts blood-  
26 hypothalamus barrier plasticity and access of metabolic signals to the arcuate nucleus in response to  
27 fasting. *Cell Metab* 2013;**17**(4):607-17.
- 28 39. Fonseca TL, Correa-Medina M, Campos MP, Wittmann G, Werneck-de-Castro JP, Arrojo e Drigo R,  
29 Mora-Garzon M, Ueta CB, Caicedo A, Fekete C, Gereben B, Lechan RM, Bianco AC. Coordination  
30 of hypothalamic and pituitary T3 production regulates TSH expression. *J Clin Invest*  
31 2013;**123**(4):1492-500.
- 32 40. Kottgen A, Pattaro C, Boger CA, Fuchsberger C, Olden M, Glazer NL, Parsa A, Gao X, Yang Q, Smith  
33 AV, O'Connell JR, Li M, Schmidt H, Tanaka T, Isaacs A, Ketkar S, Hwang SJ, Johnson AD,  
34 Dehghan A, Teumer A, Pare G, Atkinson EJ, Zeller T, Lohman K, Cornelis MC, Probst-Hensch NM,  
35 Kronenberg F, Tonjes A, Hayward C, Aspelund T, Eiriksdottir G, Launer LJ, Harris TB, Rumpfer  
36 E, Mitchell BD, Arking DE, Boerwinkle E, Struchalin M, Cavalieri M, Singleton A, Giallauria F,  
37 Metter J, de Boer IH, Haritunians T, Lumley T, Siscovick D, Psaty BM, Zillikens MC, Oostra BA,  
38 Feitosa M, Province M, de Andrade M, Turner ST, Schillert A, Ziegler A, Wild PS, Schnabel RB,  
39 Wilde S, Munzel TF, Leak TS, Illig T, Klopp N, Meisinger C, Wichmann HE, Koenig W, Zgaga L,  
40 Zemunik T, Kolcic I, Minelli C, Hu FB, Johansson A, Igl W, Zaboli G, Wild SH, Wright AF,  
41 Campbell H, Ellinghaus D, Schreiber S, Aulchenko YS, Felix JF, Rivadeneira F, Uitterlinden AG,  
42 Hofman A, Imboden M, Nitsch D, Brandstatter A, Kollerits B, Kedenko L, Magi R, Stumvoll M,  
43 Kovacs P, Boban M, Campbell S, Endlich K, Volzke H, Kroemer HK, Nauck M, Volker U, Polasek  
44 O, Vitart V, Badola S, Parker AN, Ridker PM, Kardina SL, Blankenberg S, Liu Y, Curhan GC, Franke  
45 A, Rochat T, Paulweber B, Prokopenko I, Wang W, Gudnason V, Shuldiner AR, Coresh J, Schmidt  
46 R, Ferrucci L, Shlipak MG, van Duijn CM, Borecki I, Kramer BK, Rudan I, Gyllenstein U, Wilson JF,  
47 Wittman JC, Pramstaller PP, Rettig R, Hastie N, Chasman DI, Kao WH, Heid IM, Fox CS. New loci  
48 associated with kidney function and chronic kidney disease. *Nat Genet* 2010;**42**(5):376-84.
- 49 41. Ron D, Habener JF. CHOP, a novel developmentally regulated nuclear protein that dimerizes with  
50 transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene  
51 transcription. *Genes Dev* 1992;**6**(3):439-53.
- 52 42. Loinard C, Zouggar Y, Rueda P, Ramkhalawon B, Cochain C, Vilar J, Recalde A, Richart A, Charue D,  
53 Duriez M, Mori M, Arenzana-Seisdedos F, Levy BI, Heymes C, Silvestre JS. C/EBP homologous  
54 protein-10 (CHOP-10) limits postnatal neovascularization through control of endothelial nitric oxide  
55 synthase gene expression. *Circulation* 2012;**125**(8):1014-26.
- 56  
57  
58  
59  
60

- 1  
2  
3 43. Choi SH, Ruggiero D, Sorice R, Song C, Nutile T, Vernon Smith A, Concas MP, Traglia M, Barbieri C,  
4 Ndiaye NC, Stathopoulou MG, Lagou V, Maestrale GB, Sala C, Debette S, Kovacs P, Lind L,  
5 Lamont J, Fitzgerald P, Tonjes A, Gudnason V, Toniolo D, Pirastu M, Bellenguez C, Vasani RS,  
6 Ingelsson E, Leutenegger AL, Johnson AD, DeStefano AL, Visvikis-Siest S, Seshadri S, Ciullo M.  
7 Six Novel Loci Associated with Circulating VEGF Levels Identified by a Meta-analysis of Genome-  
8 Wide Association Studies. *PLoS Genet* 2016;**12**(2):e1005874.
- 9 44. Debette S, Visvikis-Siest S, Chen MH, Ndiaye NC, Song C, Destefano A, Safa R, Azimi NM, Sawyer D,  
10 Marteau JB, Xanthakis V, Siest G, Sullivan L, Pfister M, Smith H, Choi SH, Lamont J, Lind L, Yang  
11 Q, Fitzgerald P, Ingelsson E, Vasani RS, Seshadri S. Identification of cis- and trans-acting genetic  
12 variants explaining up to half the variation in circulating vascular endothelial growth factor levels.  
13 *Circ Res* 2011;**109**(5):554-63.
- 14 45. Burgdorf KS, Gjesing AP, Grarup N, Justesen JM, Sandholt CH, Witte DR, Jorgensen T, Madsbad S,  
15 Hansen T, Pedersen O. Association studies of novel obesity-related gene variants with quantitative  
16 metabolic phenotypes in a population-based sample of 6,039 Danish individuals. *Diabetologia*  
17 2012;**55**(1):105-13.
- 18 46. Bonnefond A, Saulnier PJ, Stathopoulou MG, Grarup N, Ndiaye NC, Roussel R, Nezhad MA, Dechaume  
19 A, Lantieri O, Hercberg S, Lauritzen T, Balkau B, El-Sayed Moustafa JS, Hansen T, Pedersen O,  
20 Froguel P, Charpentier G, Marre M, Hadjadj S, Visvikis-Siest S. What is the contribution of two  
21 genetic variants regulating VEGF levels to type 2 diabetes risk and to microvascular complications?  
22 *PLoS One* 2013;**8**(2):e55921.
- 23 47. Fullwood MJ, Han Y, Wei CL, Ruan X, Ruan Y. Chromatin interaction analysis using paired-end tag  
24 sequencing. *Curr Protoc Mol Biol* 2010;**Chapter 21**:Unit 21 15 1-25.
- 25 48. Li G, Fullwood MJ, Xu H, Mulawadi FH, Velkov S, Vega V, Ariyaratne PN, Mohamed YB, Ooi HS,  
26 Tennakoon C, Wei CL, Ruan Y, Sung WK. ChIA-PET tool for comprehensive chromatin interaction  
27 analysis with paired-end tag sequencing. *Genome Biol* 2010;**11**(2):R22.
- 28 49. Daily K, Patel VR, Rigor P, Xie X, Baldi P. MotifMap: integrative genome-wide maps of regulatory motif  
29 sites for model species. *BMC Bioinformatics* 2011;**12**:495.  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Legends to figures

**Figure 1:** A) Schematic representation of SNPs in the *VEGFA* gene associated with circulating TSH, serum VEGF, Type 2 diabetes or insulin resistance. Chromosomal base pair annotations is given for the hg19 assembly. The citations for the articles is: Porcu et al.<sup>14</sup>, Debbette et al.<sup>44</sup>, Choi et al.<sup>43</sup>, Burgdorf et al.<sup>45</sup> and Bonnefond et al.<sup>46</sup>. B) Genomic region surrounding the *VEGFA* gene. Simplified representations of ENCODE sub-tracks from Genome Browser are shown. TSS: transcription start site, PROM: promoter, H3K27AC: Histone 3, Lysine 27 Acetylation (a mark of active and regulatory genomic DNA), H3K4Me1: Histone 3, Lysine 4 Mono-methylation (a mark of active enhancers), Conservation: 24 placental mammals, ChIA-PET: Chromatin Interaction Analysis Paired-End Tags (ChIA-PET) from ENCODE/Genome Institute of Singapore-Ruan<sup>47 48</sup>, c/EBP $\beta$  binding; Chromatin-immunoprecipitation using antibody directed against C/EBP $\beta$  followed by sequencing, DNase HS: DNase hyper sensitivity. Genome browser screen-shots of the area are shown in Suppl. Fig. S3. B) The binding sites for c/EBP $\beta$  and CHOP aligned showing the binding preferences of c/EBP $\beta$  and CHOP at the position of rs881858. TSS: Transcription start site. UIPAC nucleotide abbreviations N: Any, M: A or C, R: A or G, K: G or T, D: A or G or T. Binding site logos were from <http://motifmap.ics.uci.edu/><sup>49</sup>.

**Figure 2:** Reporter gene activities of examined *VEGFA* gene variants: rs881858 A or G, and rs9472138 C or T, respectively, refer to enhancer plasmids containing the *VEGFA* minimal promoter as well as the DNA regions surrounding rs881858 and rs9472138 and representing the different version of the SNPs. A) Basal activity of *VEGFA* minimal promoter and SNP-containing regions. Shown are relative luciferase activities of plasmids transfected into HEK203 cells. pGL4.10: promoter-less plasmid. Min pro: *VEGFA* minimal promoter plasmid. Luciferase activities were normalized to beta-galactosidase activity and are presented relative to the activity of the *VEGFA* minimal promoter, B) and C) Response of the rs881858 A and G alleles to separate CHOP and c/EBP $\beta$  over-expression, respectively. Shown are relative luciferase activities of rs881858 enhancer plasmids harbouring A or G co-transfected with CHOP or c/EBP $\beta$  expression plasmids in HEK293 cells. Luciferase activities were normalized to beta-galactosidase activity and are presented relative to the activity of the *VEGFA* minimal promoter (not shown on graph). D) Response of the rs881858 A and G alleles to varying levels of concurrent CHOP and c/EBP $\beta$  over-expression (5ng per well). Relative luciferase activities of rs881858 enhancer plasmids harbouring A or G co-transfected with CHOP and c/EBP $\beta$  expression plasmids in HEK293 cells. Experiments were performed 4 times in triplicate. \* P<0.05, \*\*P<0.01 by t-test for the indicated comparison or compared against the basal activity of the allele.

**Figure 3:** Electrophoretic Mobility Shift Assay (EMSA) experiments probing the *VEGFA* rs881858 region for binding using HEK293 nuclear extract. A) Lane 1-4: <sup>32</sup>P-labeled A-allele as probe ('A'), 5-10: <sup>32</sup>P-labeled G-allele as probe ('G'). Nuclear extracts from DTT treated HEK293 in all lanes. B) Comparison of complex formation using probes for the A-allele, G-allele, CHOP and c/EBP $\beta$  binding sites. C) Identification of complexes binding to CHOP and C/EBP $\beta$  using known binding sites for these proteins. Abbreviations: Probe: <sup>32</sup>P-labeled ds-oligo as indicated. COMP: Competition using unlabelled ds-oligo, Ab.: Antibody used for super-shift of complexes. N.E.: Nuclear Extract. CHOP: ds-oligo having the CHOP binding site from the TRIB3 (tribbles pseudokinase 3) gene<sup>31</sup>.

1  
2  
3 CEBP: ds-oligo having the c/EBP $\beta$  binding site from the c/EBP $\alpha$  gene promoter<sup>30</sup>. Labels to the left  
4 of images indicate formed complexes: c/EBP $\beta$  homodimer, CHOP/c/EBP $\beta$  heterodimer and arrows  
5 indicate rs881858 specific complexes not binding CHOP or c/EBP $\beta$ . N.s.: non-specific. Shown are  
6 representative blots (n=2-4 of each).  
7  
8  
9  
10

11 **Figure 4:** A model illustrating the possible involvement of *VEGFA* rs881858 in regulating thyroid  
12 function and the set point of TSH. The rs881858 A-allele is repressed by CHOP, which is activated  
13 by the cellular stress, resulting in less *VEGFA* enhancer activity, with predicted less thyrocyte VEGF  
14 produced. The result is predicted to be reduced angiogenesis and less thyroid compensatory  
15 expansion. This result will be less T4 production, less feedback inhibition of TRH and TSH release  
16 and therefore increased circulating TSH levels. A higher TSH level will result in increased  
17 stimulation of the TSH receptor and a higher degree of activation of the ER-stress response. Dashed  
18 lines illustrate rs881858 A-allele related effects.  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Supplementary Figure legends

### Figure S1: Meta-Analysis Forest plot of the association between TSH and the G-allele of *VEGFA* rs881858

The association between TSH and the G-allele of rs881858 of the *VEGFA* locus was meta-analysed in three cohorts: Inter99, Health2006 and Health2008. Shown are the  $\beta$ -values and SE for each cohort with corresponding 95% confidence intervals and overall P-value evaluated using a fixed effect model. A total of 8,445 individuals from three Danish cohorts (Inter99, Health2006, and Health2008) were combined for inverse variance meta-analyses, where weights are proportional to the squared standard errors of the effect estimates. Genomic inflation factor ( $\lambda$ ) was at acceptable levels ( $\lambda_{TSH} = 1.0$ ) after the meta-analyses. A chi-square test for heterogeneity (I) was implemented, to estimate the heterogeneity in effect sizes across different participating cohorts using METAL software (<http://csg.sph.umich.edu/abecasis/metal/>).

**Figure S2.** LD Heat Map of pairwise  $r^2$  values of SNPs studied from the *VEGFA* gene region. LD estimations and proxy search were performed using 1000 genomes project data implemented in SNP Annotation Proxy search tool (<http://www.broadinstitute.org/mpg/snap/>). An LD heat-map depicting pairwise  $r^2$  values as colors (least correlated  $r^2=0$  as dark blue and most correlated  $r^2=1.0$  as light shade of blue). Non-genotyped SNPs in LD with TSH associated SNPs were retrieved from the ENSEMBL genome browser ([http://grch37.ensembl.org/Homo\\_sapiens/Info/Index](http://grch37.ensembl.org/Homo_sapiens/Info/Index)) using CEU data, only SNPs with  $LD \geq 0.8$  were investigated further for evidence of differential allele effects.

**Figure S3:** Genome browser ([www.genome.ucsc.edu](http://www.genome.ucsc.edu)) screen shots of genomic regions containing the SNPs summarized in Table S5.

**Figure S4:** Expressed quantitative trait locus (eQTL) data for rs9472138, rs881858, rs943080 and rs4711751 for *VEGFA* mRNA in thyroid tissue. The mRNA levels of *VEGFA* are presented according to genotype of each of the SNPs. Data were retrieved from the GTEX portal ([www.gtexportal.org](http://www.gtexportal.org))<sup>32</sup>.

Figure 1

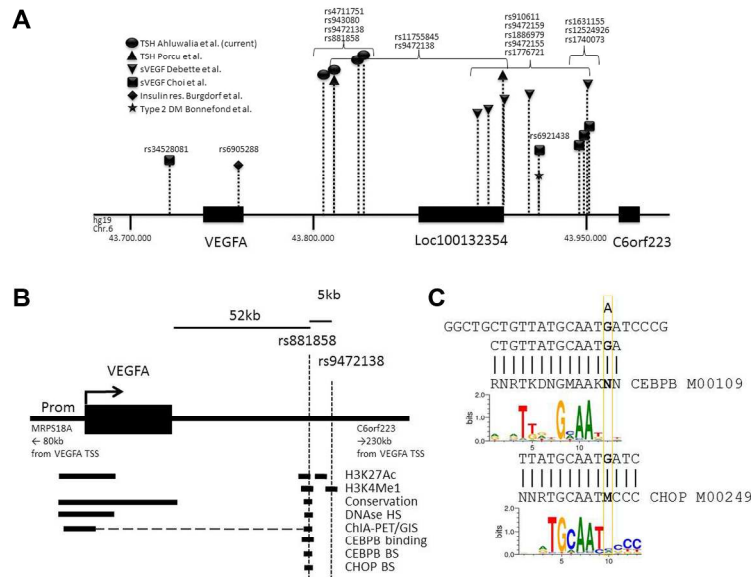


Figure 1: A) Schematic representation of SNPs in the VEGFA gene associated with circulating TSH, serum VEGF, Type 2 diabetes or insulin resistance. Chromosomal base pair annotations is given for the hg19 assembly. The citations for the articles is: Porcu et al.14, Debbette et al. 44, Choi et al. 43, Burgdorf et al. 45 and Bonnefond et al. 46. B) Genomic region surrounding the VEGFA gene. Simplified representations of ENCODE sub-tracks from Genome Browser are shown. TSS: transcription start site, PROM: promoter, H3K27Ac: Histone 3, Lysine 27 Acetylation (a mark of active and regulatory genomic DNA), H3K4Me1: Histone 3, Lysine 4 Mono-methylation (a mark of active enhancers), Conservation: 24 placental mammals, ChIA-PET: Chromatin Interaction Analysis Paired-End Tags (ChIA-PET) from ENCODE/Genome Institute of Singapore-Ruan47 48, c/EBP-β binding; Chromatin-immunoprecipitation using antibody directed against C/EBPβ followed by sequencing, DNase HS: DNase hyper sensitivity. Genome browser screen-shots of the area are shown in Suppl. Fig. S3. C) The binding sites for c/EBPβ and CHOP aligned showing the binding preferences of c/EBPβ and CHOP at the position of rs881858. TSS: Transcription start site. UIPAC nucleotide abbreviations N: Any, M: A or C, R: A or G, K: G or T, D: A or G or T. Binding site logos were from

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

<http://motifmap.ics.uci.edu/49>.  
Fig. 1  
209x297mm (300 x 300 DPI)

Confidential: For Review Only

Figure 2

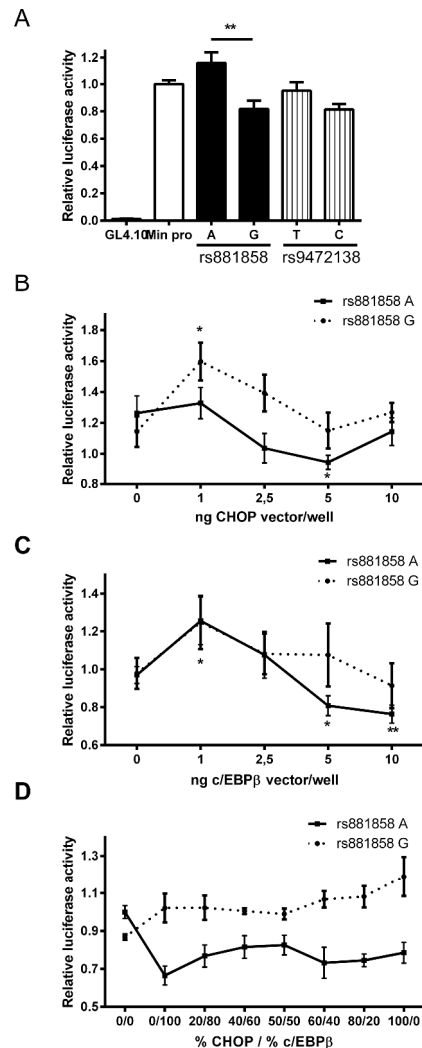


Figure 2: Reporter gene activities of examined VEGFA gene variants: rs881858 A or G, and rs9472138 C or T, respectively, refer to enhancer plasmids containing the VEGFA minimal promoter as well as the DNA regions surrounding rs881858 and rs9472138 and representing the different version of the SNPs. A) Basal activity of VEGFA minimal promoter and SNP-containing regions. Shown are relative luciferase activities of plasmids transfected into HEK203 cells. pGL4.10: promoter-less plasmid. Min pro: VEGFA minimal promoter plasmid. Luciferase activities were normalized to beta-galactosidase activity and are presented relative to the activity of the VEGFA minimal promoter, B) and C) Response of the rs881858 A and G alleles to separate CHOP and c/EBPβ over-expression, respectively. Shown are relative luciferase activities of rs881858 enhancer plasmids harbouring A or G co-transfected with CHOP or c/EBPβ expression plasmids in HEK293 cells. Luciferase activities were normalized to beta-galactosidase activity and are presented relative to the activity of the VEGFA minimal promoter (not shown on graph). D) Response of the rs881858 A and G alleles to varying levels of concurrent CHOP and c/EBPβ over-expression (5ng per well). Relative luciferase activities of rs881858 enhancer plasmids harbouring A or G co-transfected with CHOP and c/EBPβ



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

expression plasmids in HEK293 cells. Experiments were performed 4 times in triplicate. \* P<0.05, \*\*P<0.01 by t-test for the indicated comparison or compared against the basal activity of the allele.

Fig. 2  
197x489mm (600 x 600 DPI)

Confidential: For Review Only

Figure 3

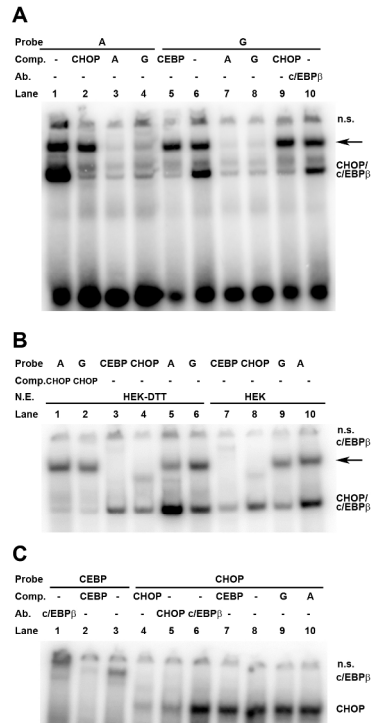


Figure 3: Electrophoretic Mobility Shift Assay (EMSA) experiments probing the VEGFA rs881858 region for binding using HEK293 nuclear extract. A) Lane 1-4: 32P-labeled A-allele as probe ('A'), 5-10: 32P-labeled G-allele as probe ('G'). Nuclear extracts from DTT treated HEK293 in all lanes. B) Comparison of complex formation using probes for the A-allele, G-allele, CHOP and c/EBP $\beta$  binding sites. C) Identification of complexes binding to CHOP and C/EBP $\beta$  using known binding sites for these proteins. Abbreviations: Probe: 32P-labeled ds-oligo as indicated. COMP: Competition using unlabelled ds-oligo, Ab.: Antibody used for super-shift of complexes. N.E.: Nuclear Extract. CHOP: ds-oligo having the CHOP binding site from the TRIB3 (tribbles pseudokinase 3) gene 31. CEBP: ds-oligo having the c/EBP $\beta$  binding site from the c/EBP $\alpha$  gene promoter 30. Labels to the left of images indicate formed complexes: c/EBP $\beta$  homodimer, CHOP/c/EBP $\beta$  heterodimer and arrows indicate rs881858 specific complexes not binding CHOP or c/EBP $\beta$ . N.s.: non-specific. Shown are representative blots (n=2-4 of each).

Fig. 3

209x297mm (300 x 300 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Confidential: For Review Only

Figure 4

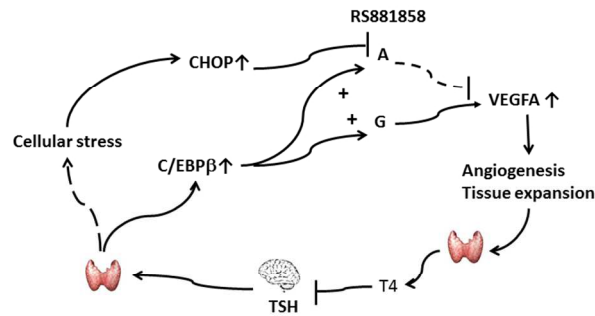
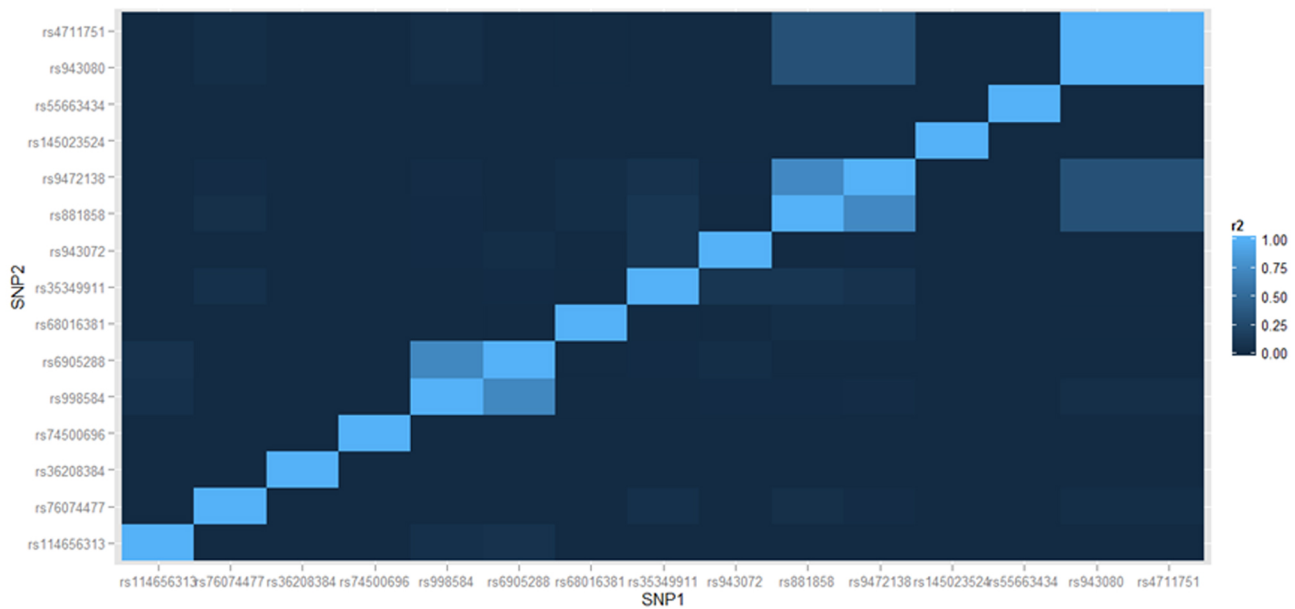


Figure 4: A model illustrating the possible involvement of VEGFA rs881858 in regulating thyroid function and the set point of TSH. The rs881858 A-allele is repressed by CHOP, which is activated by the cellular stress, resulting in less VEGFA enhancer activity, with predicted less thyrocyte VEGF produced. The result is predicted to be reduced angiogenesis and less thyroid compensatory expansion. This result will be less T4 production, less feedback inhibition of TRH and TSH release and therefore increased circulating TSH levels. A higher TSH level will result in increased stimulation of the TSH receptor and a higher degree of activation of the ER-stress response. Dashed lines illustrate rs881858 A-allele related effects.

Fig. 4

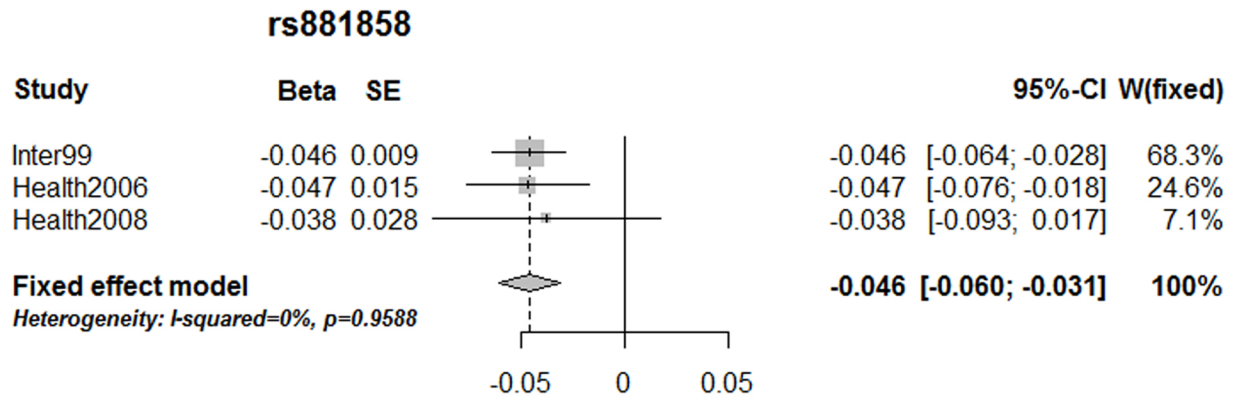
338x190mm (96 x 96 DPI)

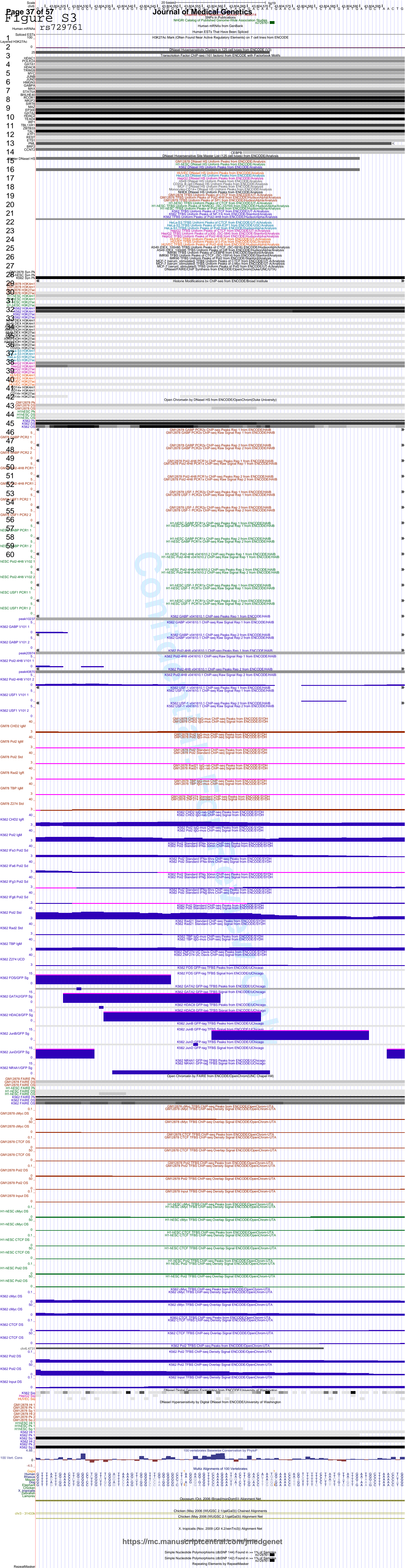
Figure S1

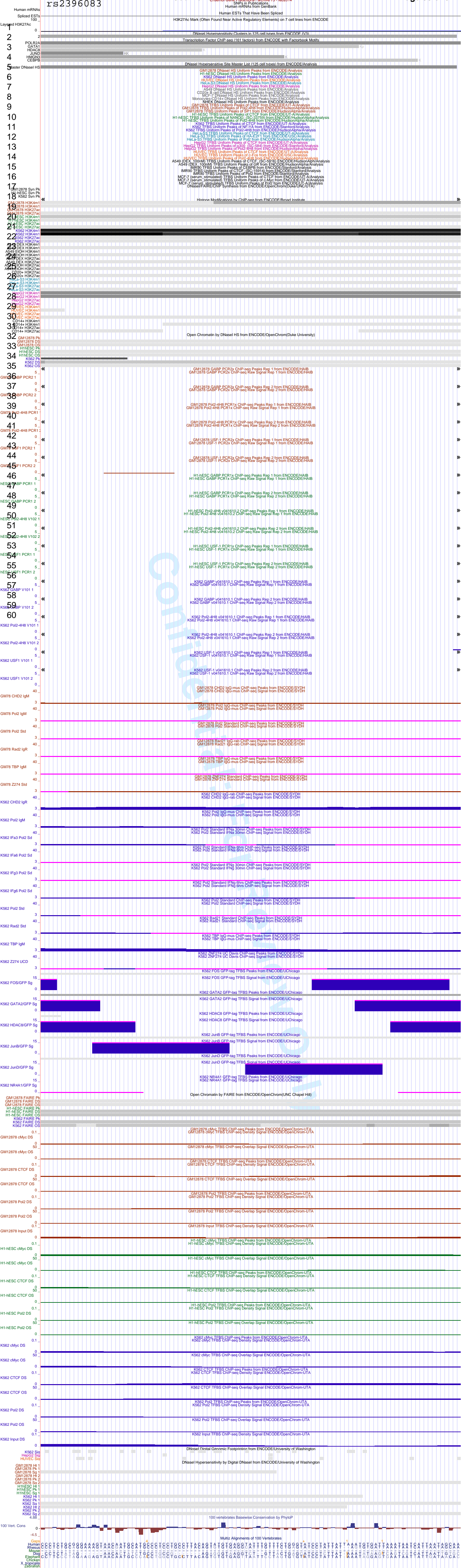


1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Figure S2



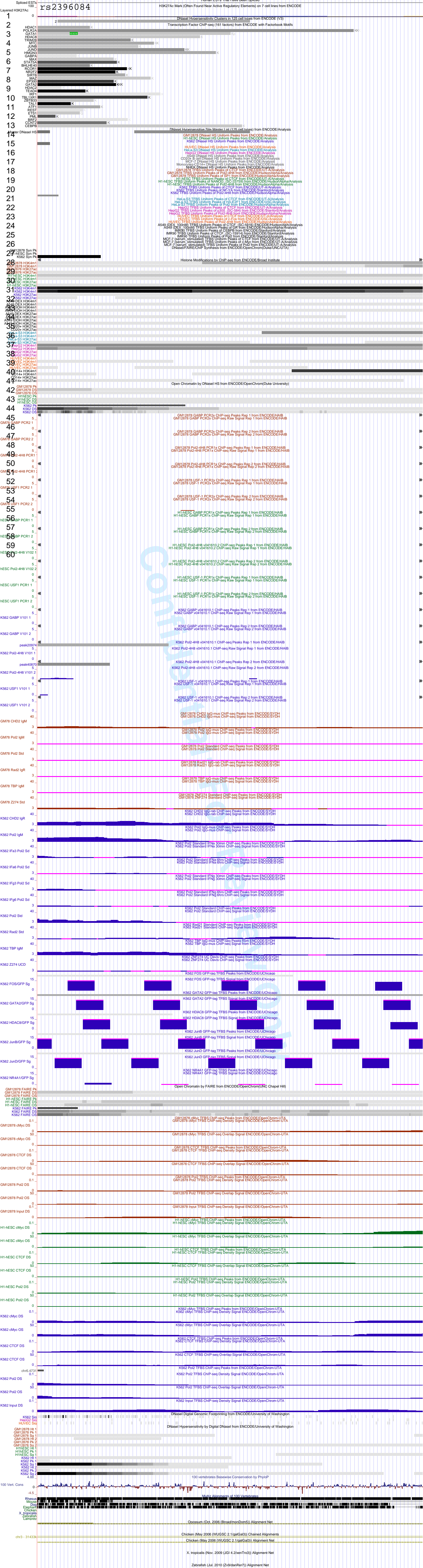




https://mc.manuscriptcentral.com/jmedgenet



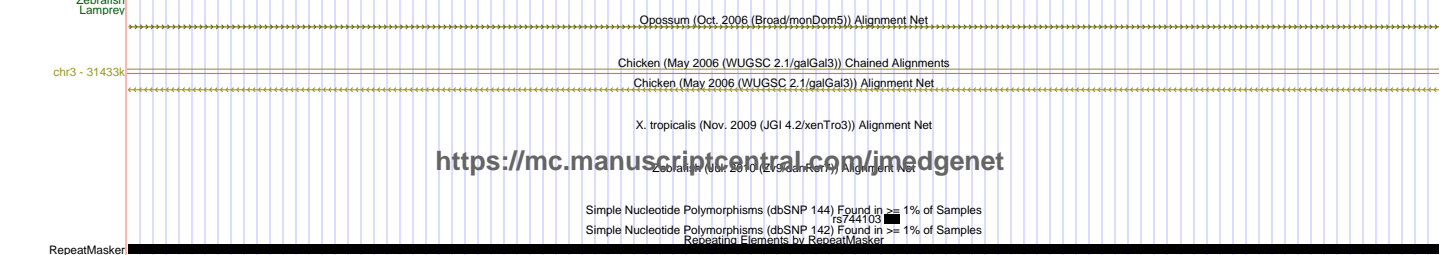
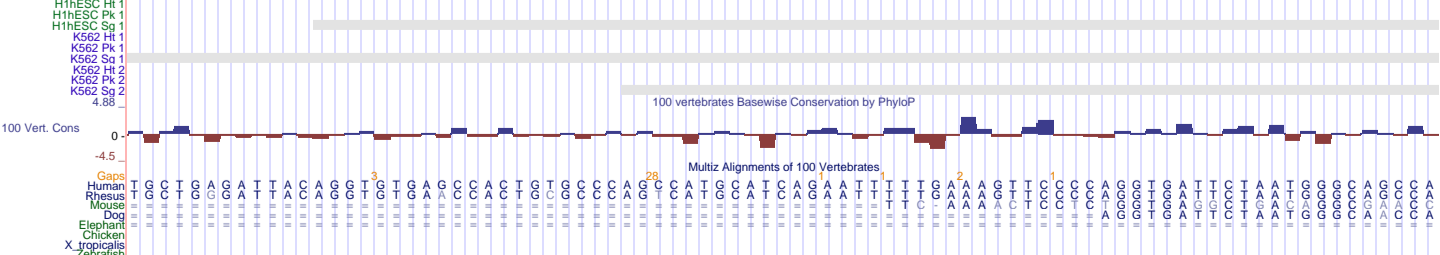
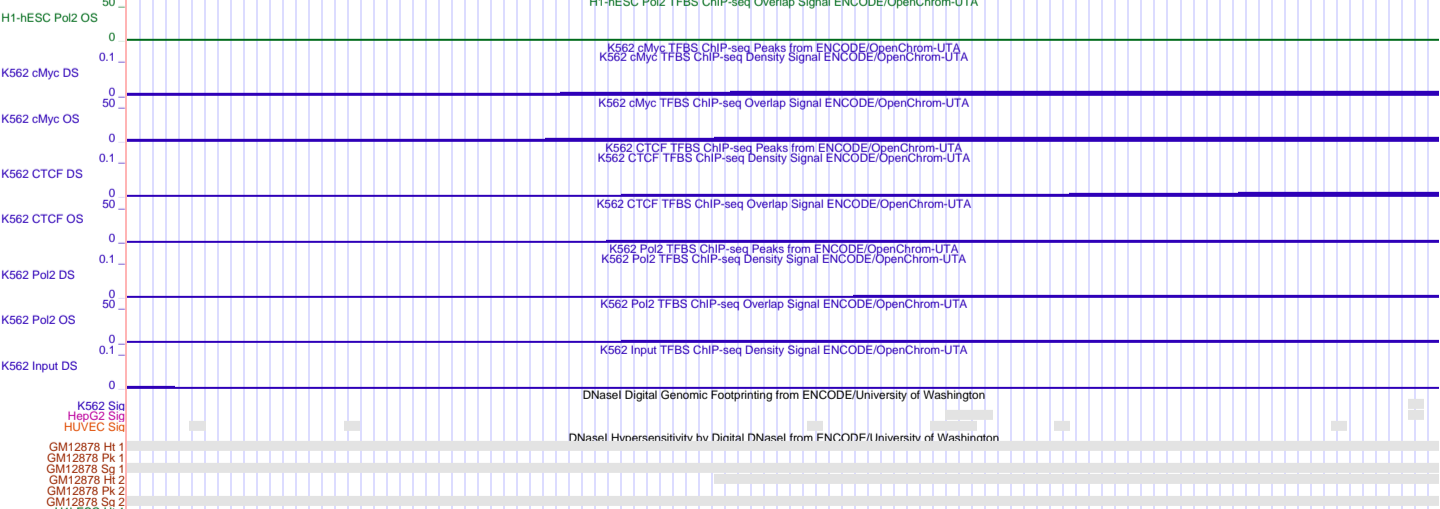
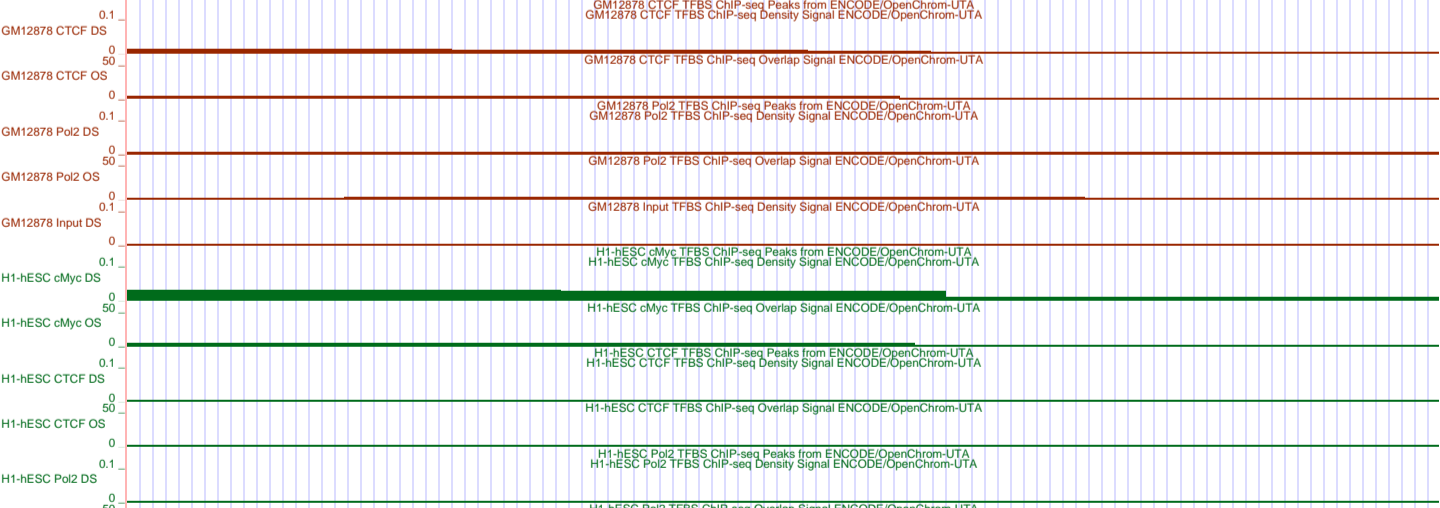
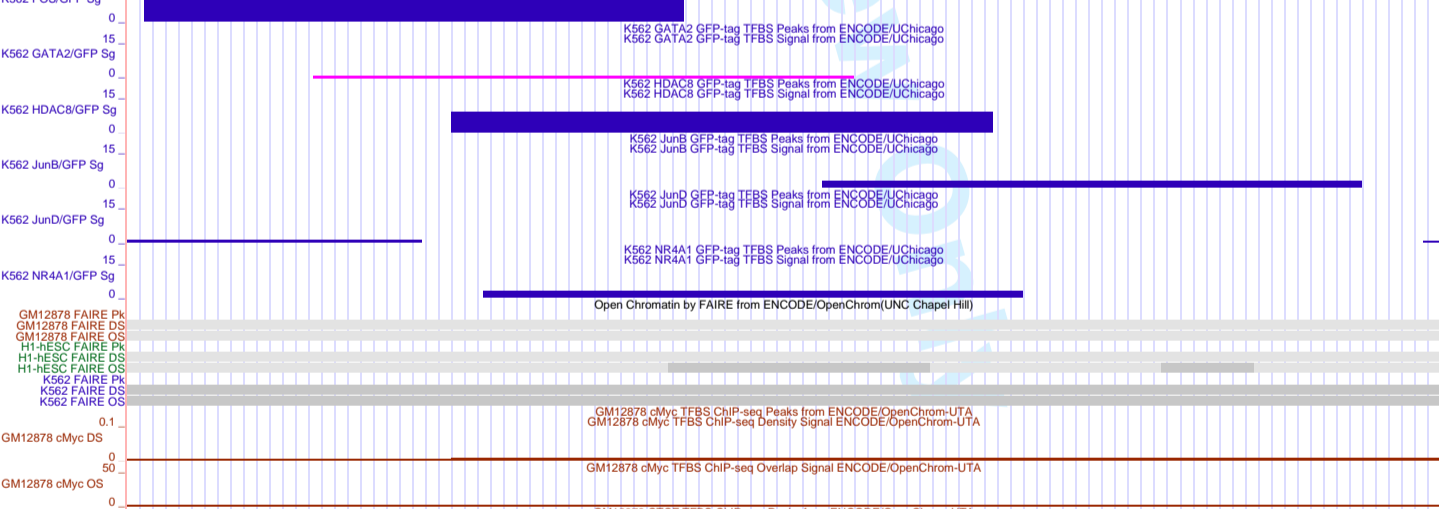
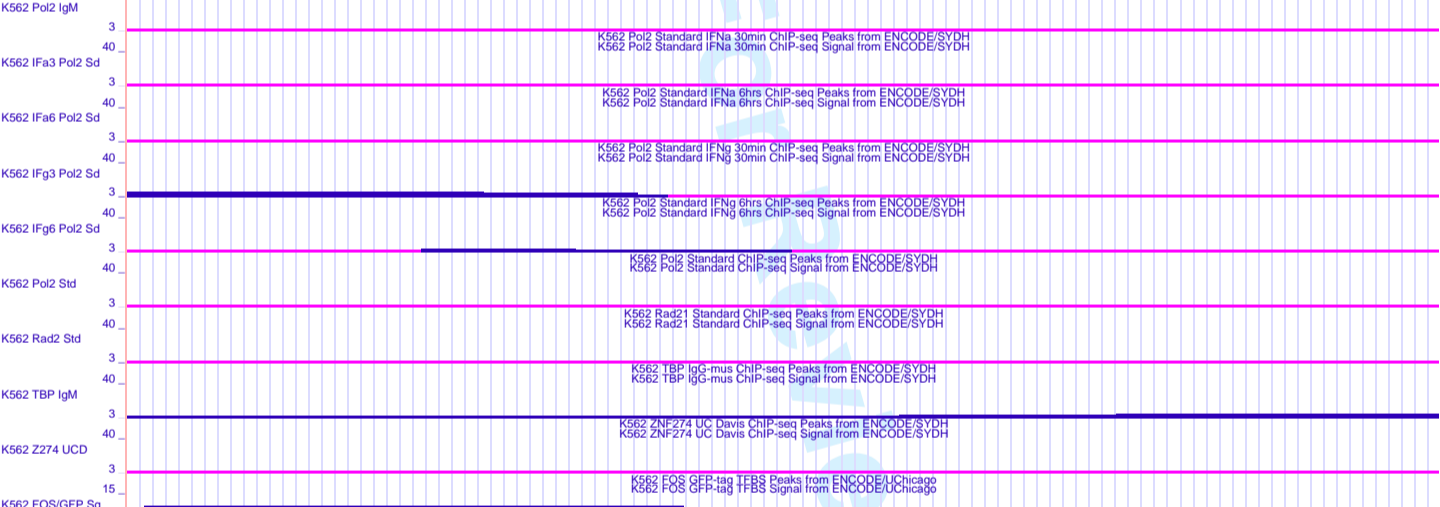
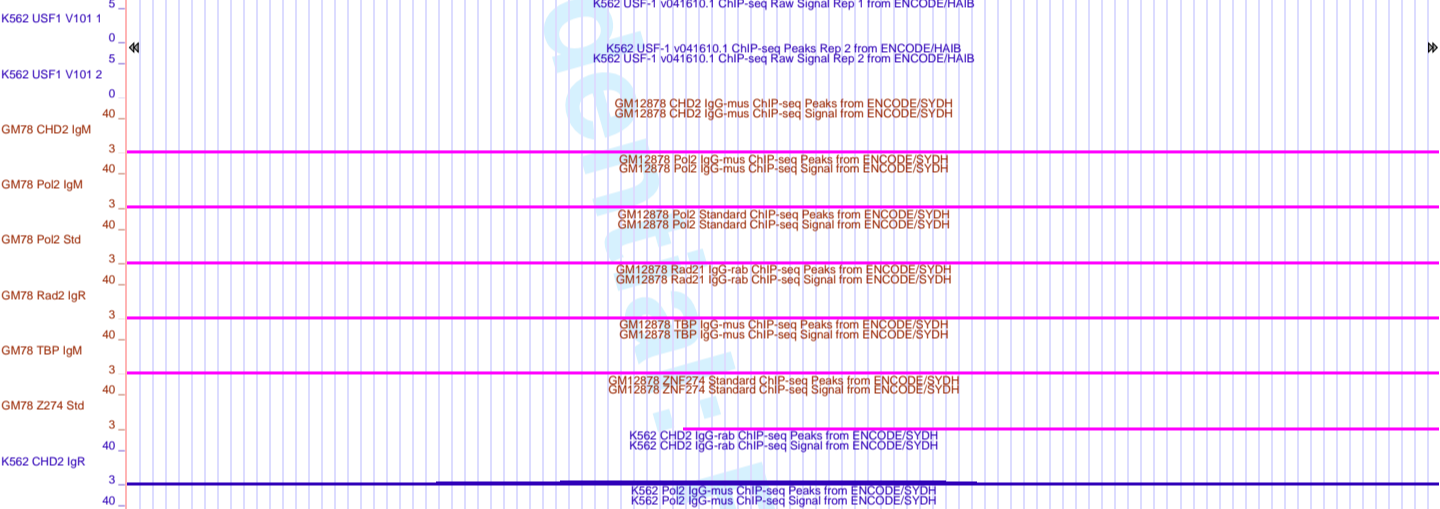
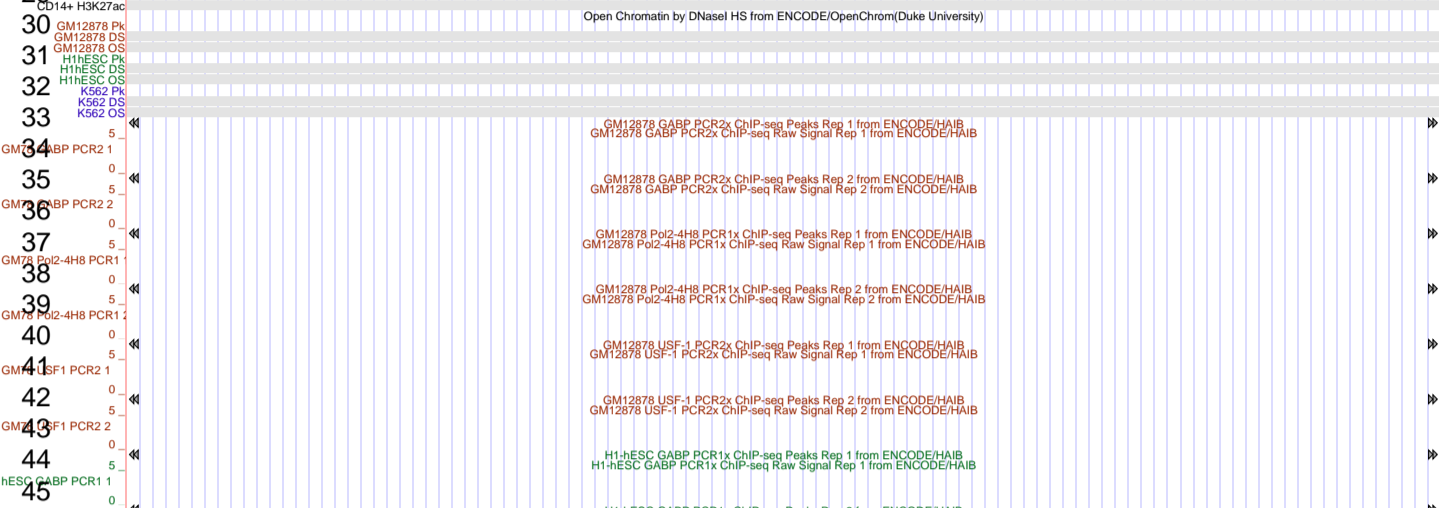
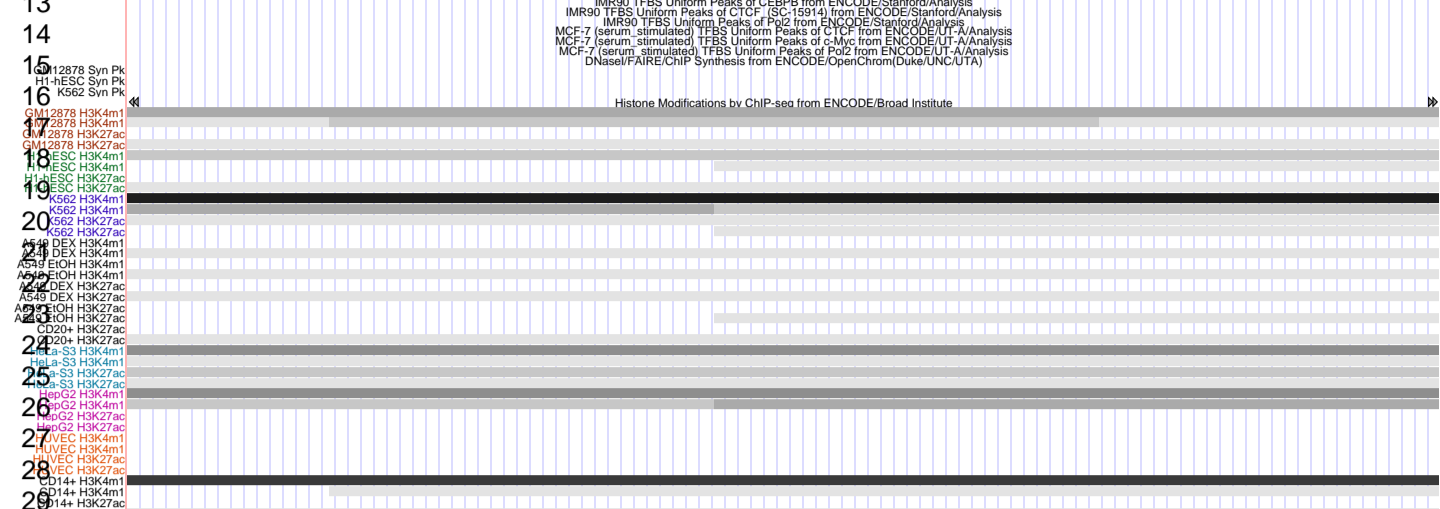
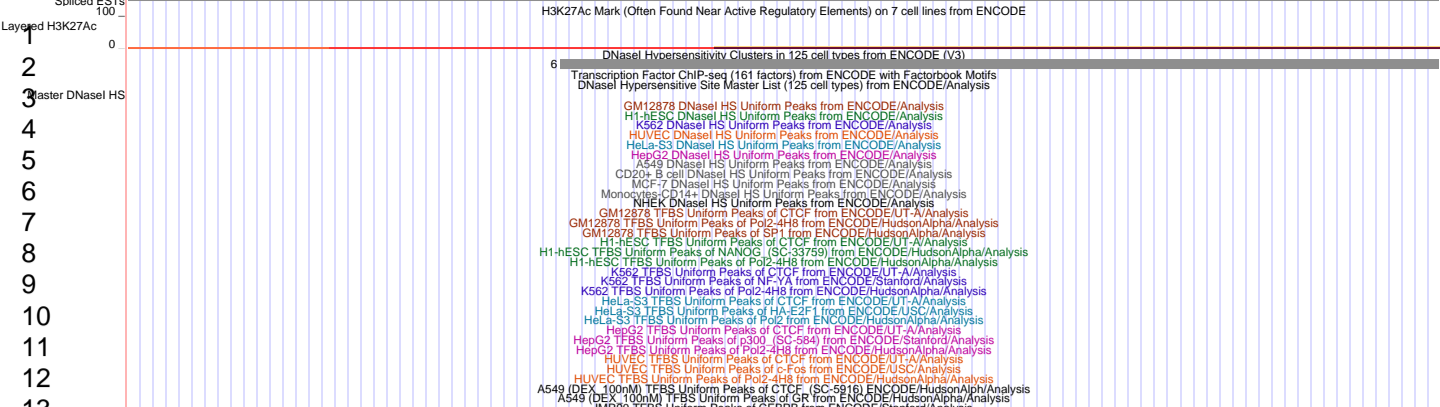
Journal of Medical Genetics 14 Human mRNAs from GenBank Human ESTs That Have Been Spliced



https://mc.manuscriptcentral.com/jmedgen

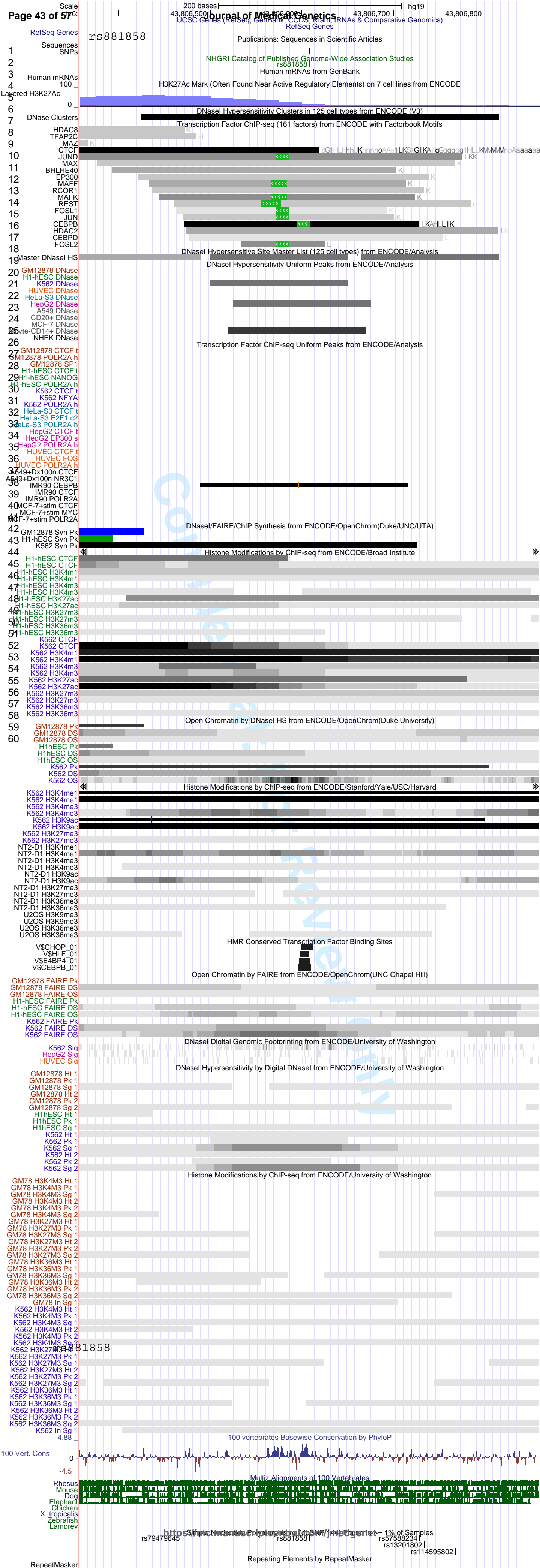
rs9367177 rs2396084 rs9367177 Simple Nucleotide Polymorphisms (dbSNP 144) Found in >= 1% of Samples

Human mRNAs Spliced ESTs H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE

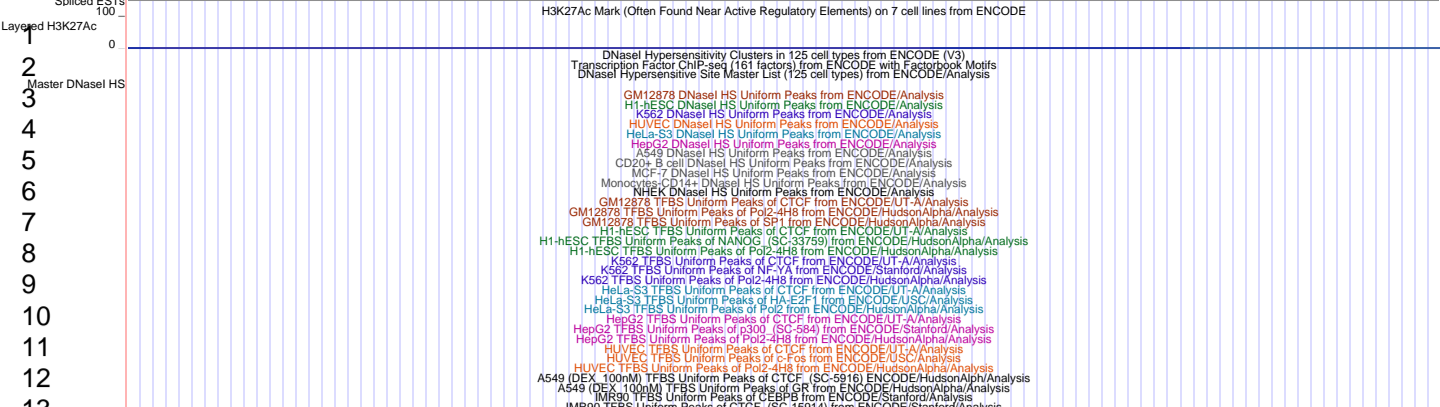






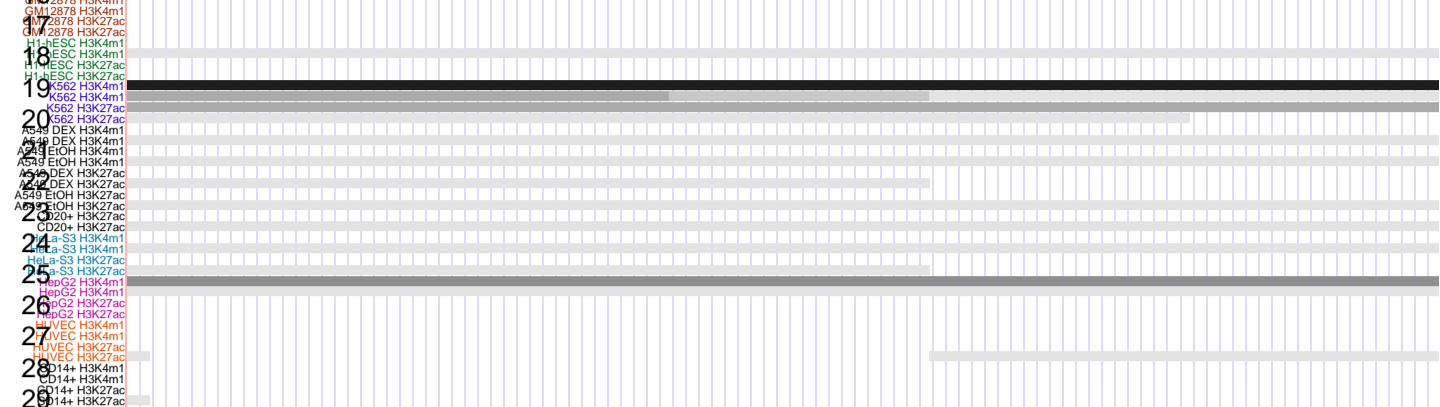


Human mRNAs... Spliced ESTs... H3K27Ac Mark... DNaseI Hypersensitivity Clusters... Transcription Factor ChIP-seq... DNaseI Hypersensitive Site Master List



Histone Modifications by ChIP-seq from ENCODE/Broad Institute

Open Chromatin by DNaseI HS from ENCODE/OpenChrom/Duke University



Open Chromatin by FAIRE from ENCODE/OpenChrom/UNC Chapel Hill

DNaseI Hypersensitivity by Digital DNaseI from ENCODE/University of Washington

100 vertebrates Basewise Conservation by PhyloP

Multiz Alignments of 100 Vertebrates

Possum (Oct. 2006 (Broad/monDom5)) Alignment Net

Chicken (May 2006 (WUGSC 2.1/galGal3)) Alignment Net

X. tropicalis (Nov. 2009 (JGI 4.2/xenTro3)) Alignment Net

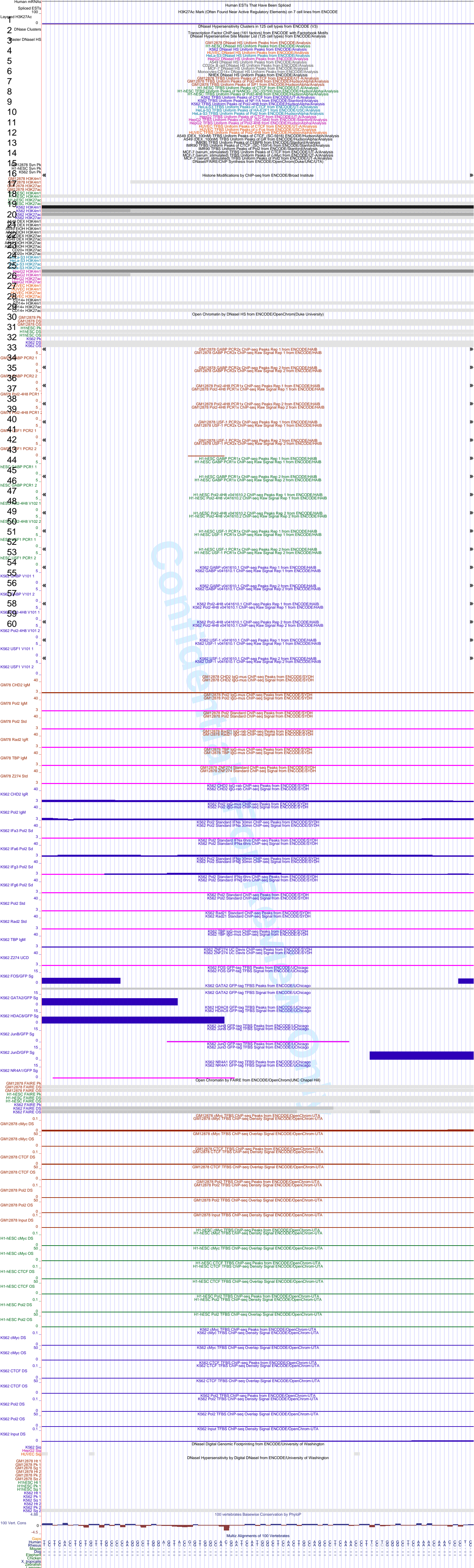
Simple Nucleotide Polymorphisms (dbSNP 142) Found in >= 1% of Samples

Repeating Elements by RepeatMasker

RepeatMasker

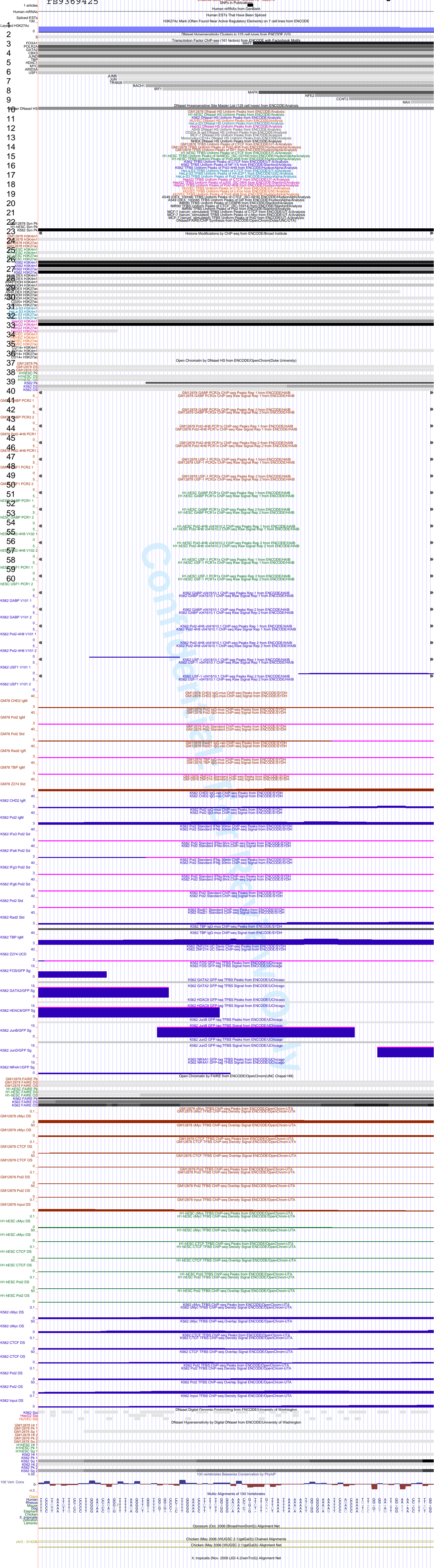
rs3792602

https://mc.manuscriptcentral.com/jmedgenet

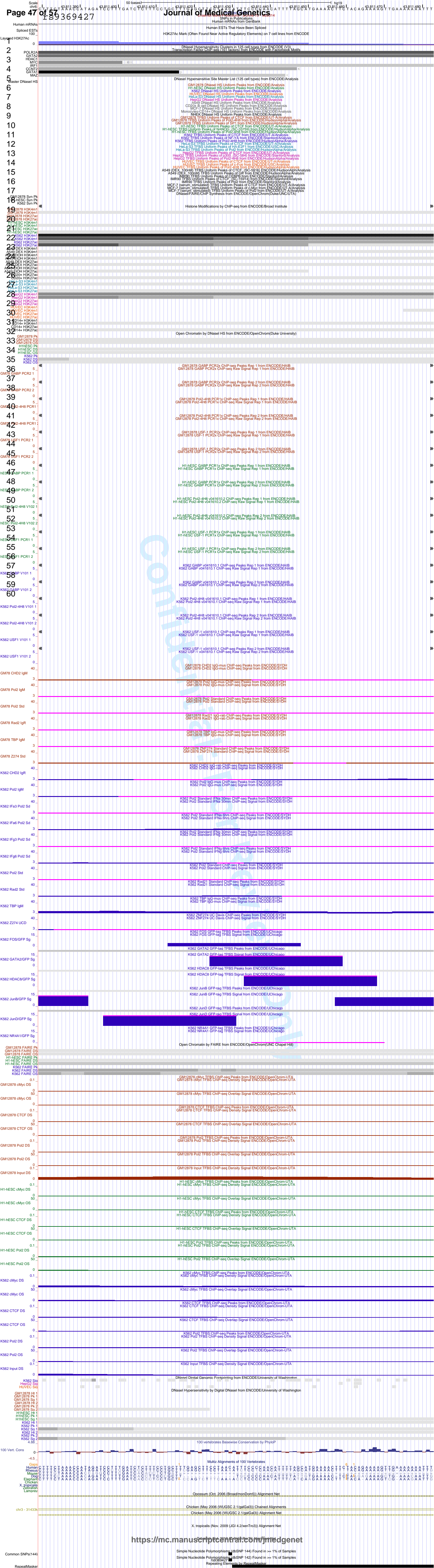


rs9369425

# Journal of Medical Genetics







Scale chr6: 13,811,291 13,811,400 13,811,410 13,811,420 13,811,430 13,811,440 13,811,450 hg19

Human mRNAs Spliced ESTs Human ESTs That Have Been Spliced

Layovered H3K27Ac H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE

2 DNaseI Hypersensitivity Clusters in 125 cell types from ENCODE (V3)

3 Transcription Factor ChIP-seq (161 factors) from ENCODE with Factorbook Motifs

4 Master DNaseI HS DNaseI Hypersensitive Site Master List (125 cell types) from ENCODE/Analysis

5-18 GM12878 DNaseI HS Uniform Peaks from ENCODE/Analysis
H1-hESC DNaseI HS Uniform Peaks from ENCODE/Analysis
K562 DNaseI HS Uniform Peaks from ENCODE/Analysis
HUVEC DNaseI HS Uniform Peaks from ENCODE/Analysis
HeLa-S3 DNaseI HS Uniform Peaks from ENCODE/Analysis
HepG2 DNaseI HS Uniform Peaks from ENCODE/Analysis
A549 DNaseI HS Uniform Peaks from ENCODE/Analysis
MDAMB231 DNaseI HS Uniform Peaks from ENCODE/Analysis
MCF7 DNaseI HS Uniform Peaks from ENCODE/Analysis
NHEK DNaseI HS Uniform Peaks from ENCODE/Analysis
GM12878 TFBS Uniform Peaks of CTCF from ENCODE/UT-A/Analysis
GM12878 TFBS Uniform Peaks of SP1 from ENCODE/HudsonAlpha/Analysis
H1-hESC TFBS Uniform Peaks of CTCF from ENCODE/UT-A/Analysis
H1-hESC TFBS Uniform Peaks of NANOG (SC-3759) from ENCODE/HudsonAlpha/Analysis
H1-hESC TFBS Uniform Peaks of Pol2-4H8 from ENCODE/HudsonAlpha/Analysis
K562 TFBS Uniform Peaks of CTCF from ENCODE/UT-A/Analysis
K562 TFBS Uniform Peaks of NF1YA from ENCODE/Stanford/Analysis
K562 TFBS Uniform Peaks of Pol2-4H8 from ENCODE/HudsonAlpha/Analysis
HeLa-S3 TFBS Uniform Peaks of CTCF from ENCODE/UT-A/Analysis
HeLa-S3 TFBS Uniform Peaks of Pol2 from ENCODE/HudsonAlpha/Analysis
HepG2 TFBS Uniform Peaks of CTCF from ENCODE/UT-A/Analysis
HepG2 TFBS Uniform Peaks of p300 (SC-584) from ENCODE/Stanford/Analysis
HepG2 TFBS Uniform Peaks of Pol2-4H8 from ENCODE/HudsonAlpha/Analysis
HUVEC TFBS Uniform Peaks of CTCF from ENCODE/UT-A/Analysis
HUVEC TFBS Uniform Peaks of c-Fos from ENCODE/USC/Analysis
HUVEC TFBS Uniform Peaks of Pol2-4H8 from ENCODE/HudsonAlpha/Analysis
A549 (DEX 100nM) TFBS Uniform Peaks of CTCF (SC-5916) ENCODE/HudsonAlpha/Analysis
A549 (DEX 100nM) TFBS Uniform Peaks of GR from ENCODE/HudsonAlpha/Analysis
IMR90 TFBS Uniform Peaks of CEBPB from ENCODE/Stanford/Analysis
IMR90 TFBS Uniform Peaks of Pol2 from ENCODE/Stanford/Analysis
MCF7 (serum stimulated) TFBS Uniform Peaks of CTCF from ENCODE/UT-A/Analysis
MCF7 (serum stimulated) TFBS Uniform Peaks of c-Myc from ENCODE/UT-A/Analysis
MCF7 (serum stimulated) TFBS Uniform Peaks of Pol2 from ENCODE/UT-A/Analysis
DNaseI/FAIRE/ChIP Synthesis from ENCODE/OpenChrom(Duke/UNC/UTA)

19-28 Histone Modifications by ChIP-seq from ENCODE/Broad Institute

29-32 Open Chromatin by DNaseI HS from ENCODE/OpenChrom(Duke University)

33-36 GM12878 GABP PCR2x ChIP-seq Peaks Rep 1 from ENCODE/HAIB

37-40 GM12878 GABP PCR2x ChIP-seq Peaks Rep 2 from ENCODE/HAIB

41-44 GM12878 Pol2-4H8 PCR1x ChIP-seq Peaks Rep 1 from ENCODE/HAIB

45-48 GM12878 Pol2-4H8 PCR1x ChIP-seq Peaks Rep 2 from ENCODE/HAIB

49-52 GM12878 USF-1 PCR2x ChIP-seq Peaks Rep 1 from ENCODE/HAIB

53-56 GM12878 USF-1 PCR2x ChIP-seq Peaks Rep 2 from ENCODE/HAIB

57-60 K562 GABP v041610.1 ChIP-seq Peaks Rep 1 from ENCODE/HAIB

61-64 K562 GABP v041610.1 ChIP-seq Peaks Rep 2 from ENCODE/HAIB

65-68 K562 Pol2-4H8 v041610.1 ChIP-seq Peaks Rep 1 from ENCODE/HAIB

69-72 K562 Pol2-4H8 v041610.1 ChIP-seq Peaks Rep 2 from ENCODE/HAIB

73-76 K562 USF-1 v041610.1 ChIP-seq Peaks Rep 1 from ENCODE/HAIB

77-80 K562 USF-1 v041610.1 ChIP-seq Peaks Rep 2 from ENCODE/HAIB

81-84 GM12878 CHD2 IgG-mus ChIP-seq Peaks from ENCODE/SVDH

85-88 GM12878 Pol2 Standard ChIP-seq Peaks from ENCODE/SVDH

89-92 GM12878 Rad21 IgG-rab ChIP-seq Peaks from ENCODE/SVDH

93-96 GM12878 TBP IgG-mus ChIP-seq Peaks from ENCODE/SVDH

97-100 K562 CHD2 IgG-rab ChIP-seq Peaks from ENCODE/SVDH

101-104 K562 Pol2 Standard IFNa 30min ChIP-seq Peaks from ENCODE/SVDH

105-108 K562 Pol2 Standard IFNa 6hrs ChIP-seq Peaks from ENCODE/SVDH

109-112 K562 Pol2 Standard IFNg 30min ChIP-seq Peaks from ENCODE/SVDH

113-116 K562 Pol2 Standard IFNg 6hrs ChIP-seq Peaks from ENCODE/SVDH

117-120 K562 Rad21 Standard ChIP-seq Peaks from ENCODE/SVDH

121-124 K562 TBP IgG-mus ChIP-seq Peaks from ENCODE/SVDH

125-128 K562 ZNF274 UC Davis ChIP-seq Peaks from ENCODE/SVDH

129-132 K562 FOS GFP-tag TFBS Peaks from ENCODE/UCicago

133-136 K562 HDAC8 GFP-tag TFBS Peaks from ENCODE/UCicago

137-140 K562 JunB GFP-tag TFBS Peaks from ENCODE/UCicago

141-144 K562 JunD GFP-tag TFBS Peaks from ENCODE/UCicago

145-148 K562 NR4A1 GFP-tag TFBS Peaks from ENCODE/UCicago

149-152 Open Chromatin by FAIRE from ENCODE/OpenChrom(UNC Chapel Hill)

153-156 GM12878 cMyc TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

157-160 GM12878 CTCF TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

161-164 GM12878 Pol2 TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

165-168 GM12878 Input TFBS ChIP-seq Density Signal ENCODE/OpenChrom-UTA

169-172 H1-hESC cMyc TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

173-176 H1-hESC CTCF TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

177-180 H1-hESC Pol2 TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

181-184 K562 cMyc TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

185-188 K562 CTCF TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

189-192 K562 Pol2 TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

193-196 K562 Input TFBS ChIP-seq Density Signal ENCODE/OpenChrom-UTA

197-200 DNaseI Digital Genomic Footprinting from ENCODE/University of Washington

201-204 100 vertebrates Base-wise Conservation by PhyloP

205-208 Multiz Alignments of 100 Vertebrates

209-212 Opposum (Oct. 2006 (Broad/monDom5)) Alignment Net

213-216 Chicken (May 2006 (WUGSC 2.1/galGal3)) Chained Alignments

217-220 Chicken (May 2006 (WUGSC 2.1/galGal3)) Alignment Net

221-224 X. tropicalis (Nov. 2009 (JGI 4.2/xenTro3)) Alignment Net

225-228 RepeatMasker

229-232 Simple Nucleotide Polymorphisms (dbSNP 144) Found in >= 1% of Samples

233-236 Repeating Elements by RepeatMasker

237-240 Common SNPs (144)

241-244 RepeatMasker

245-248 RepeatMasker

249-252 RepeatMasker

253-256 RepeatMasker

257-260 RepeatMasker

261-264 RepeatMasker

265-268 RepeatMasker

269-272 RepeatMasker

273-276 RepeatMasker

277-280 RepeatMasker

281-284 RepeatMasker

285-288 RepeatMasker

289-292 RepeatMasker

293-296 RepeatMasker

297-300 RepeatMasker

301-304 RepeatMasker

305-308 RepeatMasker

309-312 RepeatMasker

313-316 RepeatMasker

317-320 RepeatMasker

321-324 RepeatMasker

325-328 RepeatMasker

329-332 RepeatMasker

333-336 RepeatMasker

337-340 RepeatMasker

341-344 RepeatMasker

345-348 RepeatMasker

349-352 RepeatMasker

353-356 RepeatMasker

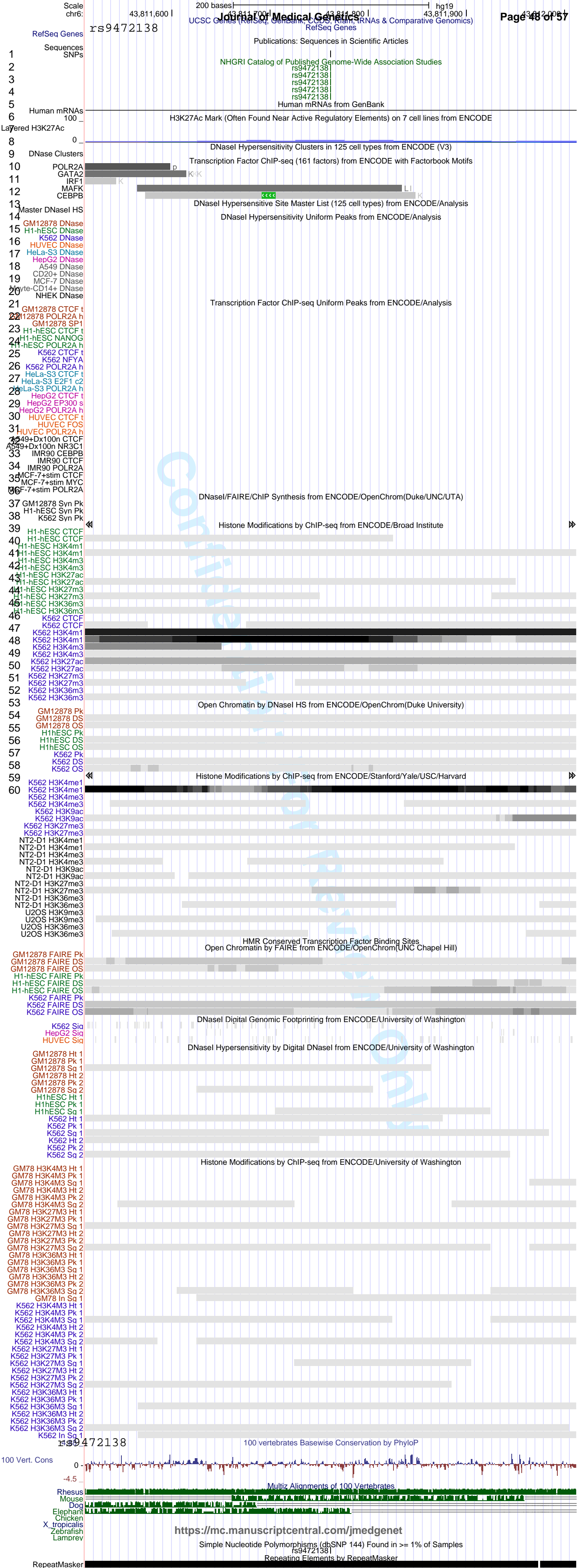
357-360 RepeatMasker

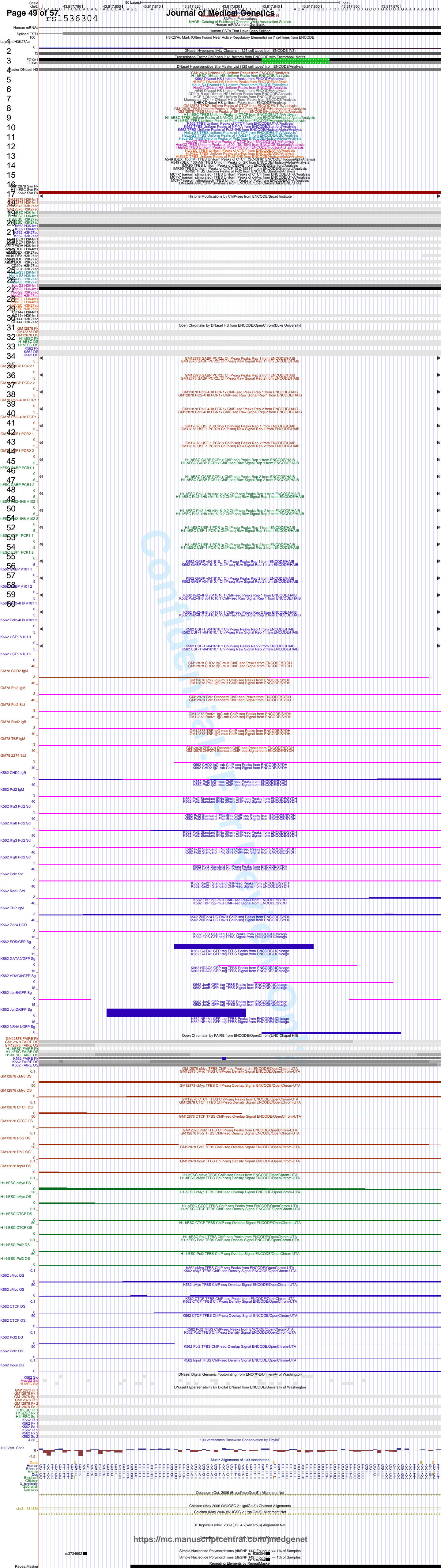
361-364 RepeatMasker

365-368 RepeatMasker

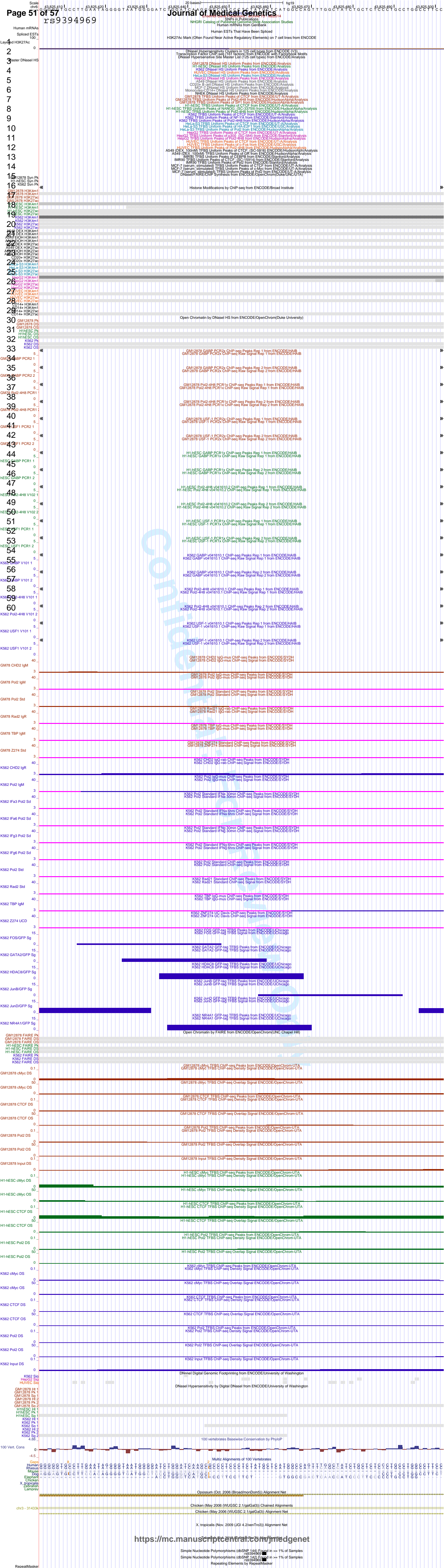
369-372 RepeatMasker

373-376 RepeatMasker





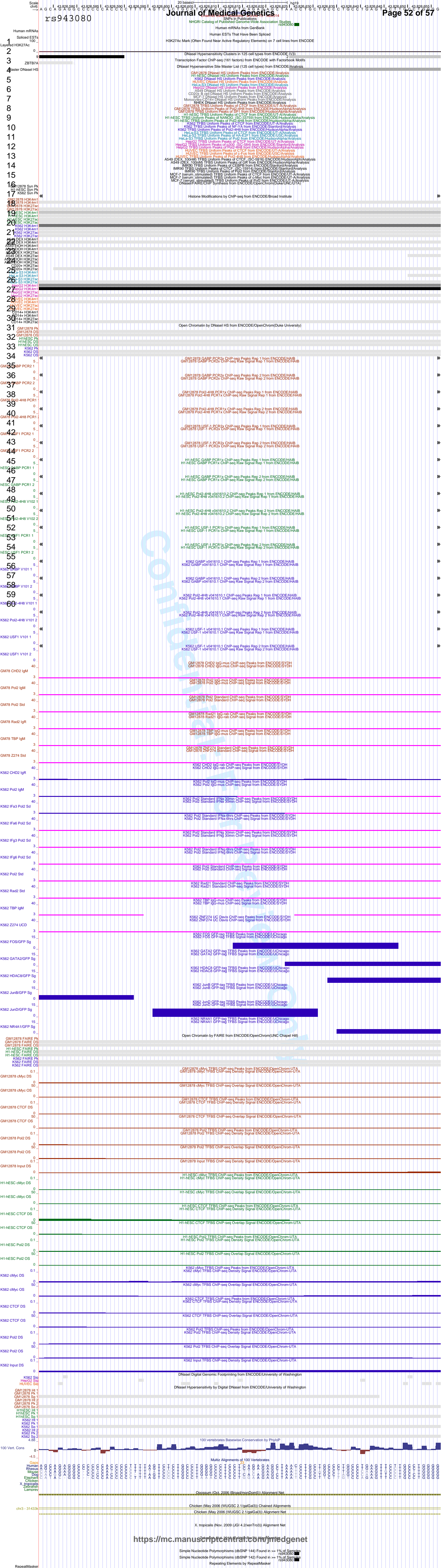




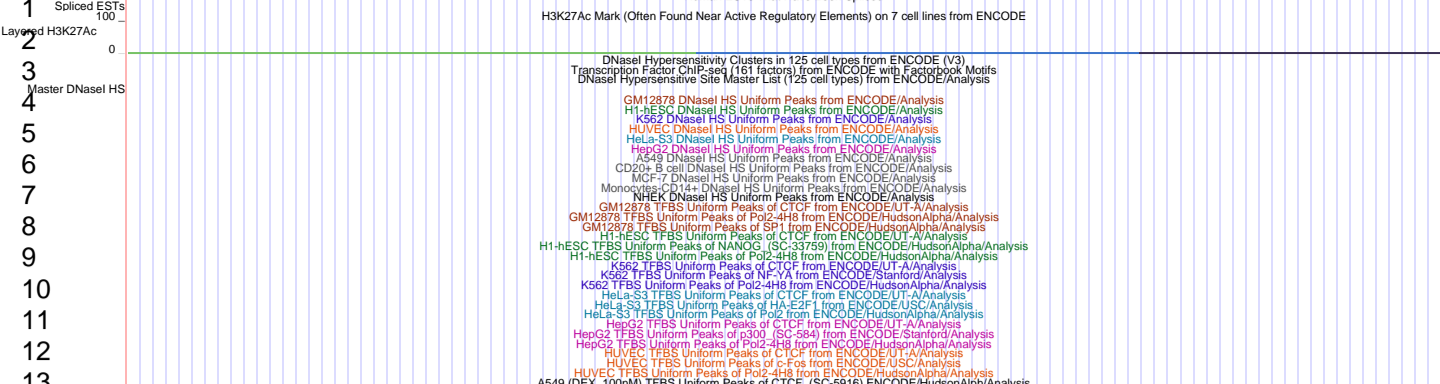
<https://mc.manuscriptcentral.com/jmgenet>

Simple Nucleotide Polymorphisms (dbSNP 144) Found in >= 1% of Samples

Repeating Elements by RepeatMasker

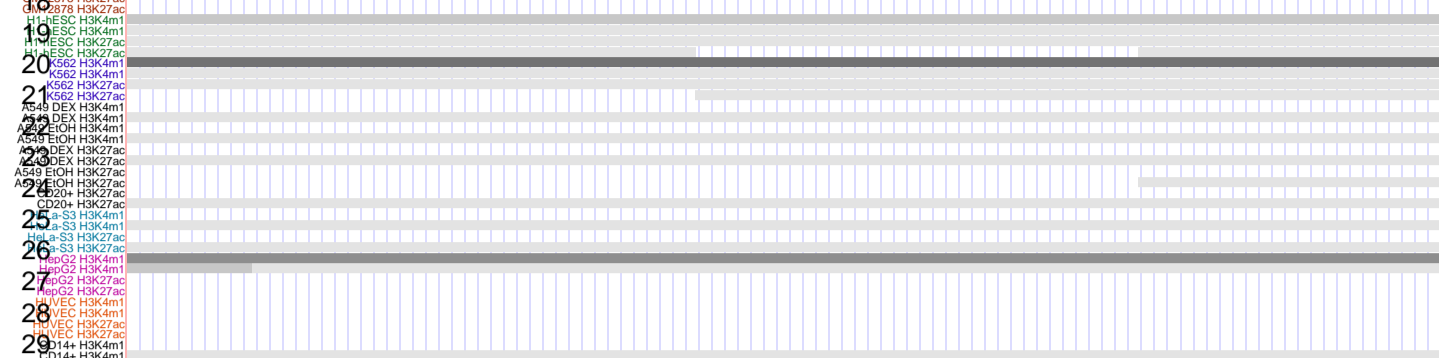


Human mRNAs, Spliced ESTs, H3K27Ac Mark, DnaseI Hypersensitivity Clusters, Transcription Factor ChIP-seq, DnaseI Hypersensitive Site Master List, Human ESTs That Have Been Spliced



Histone Modifications by ChIP-seq from ENCODE/Broad Institute

Open Chromatin by DNaseI HS from ENCODE/OpenChrom(Duke University)



Open Chromatin by FAIRE from ENCODE/OpenChrom(UNC Chapel Hill)

GM12878 cMyc TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

GM12878 CTCF TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

GM12878 Pol2 TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

GM12878 Input TFBS ChIP-seq Density Signal ENCODE/OpenChrom-UTA

H1-hESC cMyc TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

H1-hESC CTCF TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

H1-hESC Pol2 TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

K562 cMyc TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

K562 CTCF TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

K562 Pol2 TFBS ChIP-seq Peaks from ENCODE/OpenChrom-UTA

K562 Input TFBS ChIP-seq Density Signal ENCODE/OpenChrom-UTA

DnaseI Digital Genomic Fingerprinting from ENCODE/University of Washington

DnaseI Hypersensitivity by Digital DNaseI from ENCODE/University of Washington

100 vertebrates Basewise Conservation by PhyloP

Multiz Alignments of 100 Vertebrates

Opposum (Oct. 2006 (Broad/monDom5)) Alignment Net

Chicken (May 2006 (WUGSC 2.1/galGal3)) Chained Alignment Nets

X. tropicalis (Nov. 2009 (JGI 4.2/xenTro3)) Alignment Net

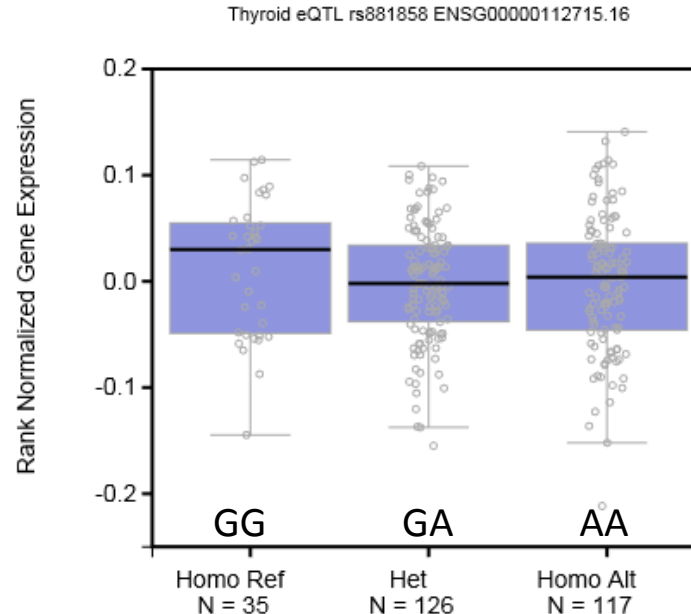
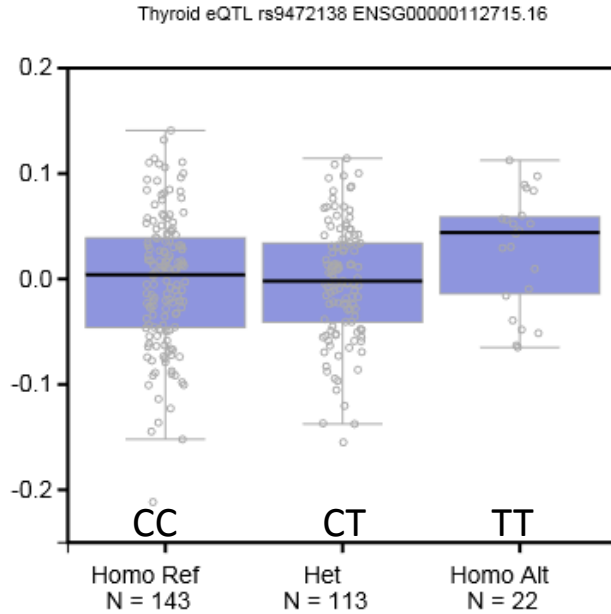
RepeatMasker

Simple Nucleotide Polymorphisms (dbSNP 142) Found in >= 1% of Samples

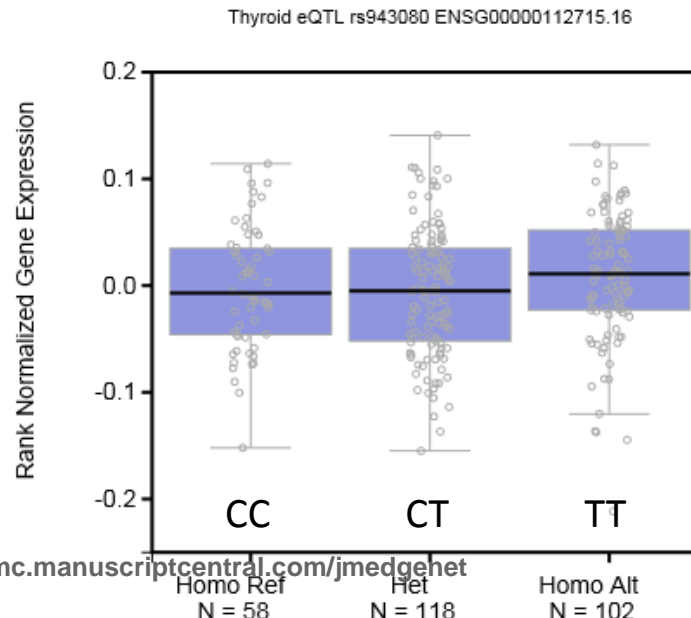
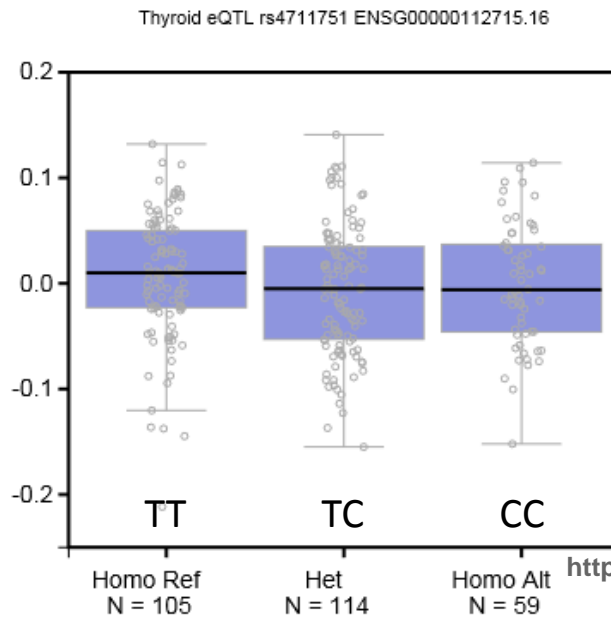
Simple Nucleotide Polymorphisms (dbSNP 142) Found in >= 1% of Samples

Revealing Elements by RepeatMasker

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43



SNP	P-Value	Effect Size
rs9472138	0,15	0,067
rs881858	0,22	-0,058
rs943080	0,24	0,049
rs4711751	0,26	-0,047





## Supplementary tables

**Table S1: Baseline characteristics of the three Danish cohorts participating in discovery analyses.**

	<b>Inter99</b>	<b>Health2006</b>	<b>Health2008</b>
<i>N</i>	5645	2711	601
Men (%)	50.1	44.8	43.7
Age (yrs)	46.1 ± 7.9	48.9 ± 13.1	46.4 ± 8.1
BMI (kg/m <sup>2</sup> )	26.2 ± 4.5	25.8 ± 4.6	25.6 ± 4.3
TSH (mIU/L)	1.39 ± 0.67	1.53 ± 0.73	1.15 ± 0.55

Data are mean ± standard deviation.

**Table S2: Study population characteristics at baseline and at five years follow-up in the Inter99 normal glucose tolerant (NGT) participants.**

<b>Characteristics</b>	<b>Baseline</b>	<b>Follow up</b>	<b>Changes from baseline to follow-up</b>
<i>n</i> (%men)	4,374 (46.3%)	Up to 3,467 (46.4%)	
Age (years)	45.1 ± 7.8	50.5 ± 7.8	5.38 ± 2.3
BMI (kg/m <sup>2</sup> )	25.5 ± 4.0	25.7 ± 4.0	0.44 ± 1.7
HbA1c (%)	5.7 ± 0.3	5.7 ± 0.3	-0.038 ± 0.26
Leptin (ng/ml)	5.3 (2.4-11.1)	-	-
T4 (pmol/L)	15 (13.9-16.3)	14.7 (13.6-15.9)	-0.3 (-1.3-0.6)
TSH (mU/L)	1.26 (0.91-1.75)	1.41 (1.0-1.99)	0.14 (-0.15-0.5)
Fasting plasma glucose (mmol/L)	5.3 ± 0.4	5.2 ± 0.5	-0.046 ± 0.49
2hr glucose (mmol/L)	5.4 ± 1.1	5.5 ± 1.4	0.016 ± 1.48
Fasting serum insulin (pmol/L)	31.0 (22.0 - 45.0)	28.0 (21.0 - 41.0)	-1.0 (-11.0-8.0)
Insulin sensitivity index (ISI <sub>MATSUDA</sub> )	3.16 (2.24-4.44)	3.12 (2.18-4.40)	-
Insulinogenic index	77.1 (49.1-128.0)	77.1 (50.0-124.2)	-
HOMA-IR	1.21 (0.83-1.80)	1.09 (0.77-1.65)	-0.05 (-0.46-0.34)
Disposition index	229.2 (161.7-348.2)	230.8 (161.6-337.2)	-

We use interquartile range for insulin and related measures as the distribution may be skewed. Data are mean ± SD or median (interquartile range)

**Table S3: Oligonucleotides used for cloning *VEGFA* promoter-enhancer constructs and electrophoretic mobility shift assays**

Name	Sequence (italics: tail for fill-in labeling, bold CHOP10 site, red: rs881858)	Purpose
RS881858 G sense >hg19_dna range=chr6:43806593- 43806628	<i>agctTGCTGTTATGCAAT</i> <b>GAT</b> CCCGC	Fill-in labeled probe
RS881858 G a-sense	<i>agctGCGGGAT</i> <b>CATTGCATAACAGCA</b>	Fill-in labeled probe
RS881858 A sense	<i>agctTGCTGTTATGCAAT</i> <b>AAT</b> CCCGC	Fill-in labeled probe
RS881858 A a-sense	<i>agctGCGGGAT</i> <b>TATTGCATAACAGCA</b>	Fill-in labeled probe
c/EBP $\beta$ sense	<i>agctGCGTTGCGCCACGATCTCTC</i>	Fill-in labeled probe
c/EBP $\beta$ a-sense	<i>agctGAGAGATCGTGGCGCAACGC</i>	Fill-in labeled probe
RS881858 G sense f	<i>agctTGCTGTTATGCAAT</i> <b>GAT</b> CCCGC <i>agct</i>	Full length probe for competition assays
RS881858 G a-sense f	<i>agctGCGGGAT</i> <b>CATTGCATAACAGCA</b> <i>agct</i>	Full length probe for competition assays
RS881858 A sense f	<i>agctTGCTGTTATGCAAT</i> <b>AAT</b> CCCGC <i>agct</i>	Full length probe for competition assays
RS881858 A a-sense f	<i>agctGCGGGAT</i> <b>TATTGCATAACAGCA</b> <i>agct</i>	Full length probe for competition assays
c/EBP $\beta$ sense f	<i>agctGCGTTGCGCCACGATCTCTC</i> <i>agct</i>	Full length probe for competition assays
c/EBP $\beta$ a-sense f	<i>agctGAGAGATCGTGGCGCAACGC</i> <i>agct</i>	Full length probe for competition assays
CHOP10 sense	<i>agctTTCTGATGCAACTGGCTGAG</i>	Fill-in labeled probe
CHOP10 a-sense	<i>agctCTCAGCCAGTTGCATCAGAA</i>	Fill-in labeled probe
CHOP10 sense f	<i>agctTTCTGATGCAACTGGCTGAG</i> <i>agct</i>	Full length probe for competition assays
CHOP10 a-sense f	<i>agctCTCAGCCAGTTGCATCAGAA</i> <i>agct</i>	Full length probe for competition assays
<i>VEGFA</i> pro F	CTAACTGGCCGGTACCTTCCCAAAGGAC CCCAGTC	Cloning: PCR primer <i>VEGFA</i> promoter-exon 1, sense
<i>VEGFA</i> pro R	TATCCTCGAGGCTAGGCCCGATTCAAGT GGGGAAT	Cloning: PCR primer <i>VEGFA</i> promoter-exon 1, a- sense
RS881858 sense	AAATCGATAAGGATCCTCTTCCACAGAA GTCAGAGTGC	Cloning: PCR primer RS881858 region, sense
RS881858 a-sense	AAGGGCATCGGTGACCAAAGCCCCTTG CCTCCC	Cloning: PCR primer RS881858 region, a-sense
RS9472138 sense	AAATCGATAAGGATCCACCCTAAGCACG TTCTCCTC	Cloning: PCR primer RS9472138 region, sense
RS9472138 a-sense	AAGGGCATCGGTGACACAACCTACTGA TACATGCCACA	Cloning: PCR primer RS9472138 region, a-sense

Table S4: VEGFA SNPs not reaching study-wide significance for association with circulating levels of thyroid stimulating hormone (TSH)

SNP name	Position (build 37/hg19)	Location wrt VEGFA	Alleles (effect/other)	EAF	Inter99 n=5,420		Health2006 n=-2,442		Health2008 n=583		Combined		
					Effect	P	Effect	P	Effect	P	N	P	I <sup>2</sup> (P <sub>HET</sub> )
rs114656313	43,692,999	Upstream	A/C	0.026	-0.012	0.65	0.03	0.45	0.0007	0.99	8,445	0.95	0 (0.69)
rs76074477	43,713,214	Upstream	A/G	0.043	0.056	0.0098	-0.047	0.63	-0.081	0.21	8,445	0.12	67 (0.05)
rs36208384	43,737,909	5' region	A/C	0.016	0.018	0.56	0.003	0.57	-0.087	0.47	8,440	0.54	0 (0.67)
rs74500696	43,748,845	Intron	A/G	0.012	-0.037	0.36	-0.026	0.95	0.008	0.94	8,445	0.49	0 (0.83)
rs998584	43,757,896	Downstream	G/T	0.45	-0.004	0.62	0.016	0.86	-0.007	0.78	8,426	0.57	0 (0.99)
rs6905288	43,758,873	Downstream	A/G	0.55	-0.004	0.62	-0.016	0.25	-0.008	0.74	8,444	0.27	0 (0.79)
rs68016381	43,761,645	Downstream	T/C	0.047	-0.048	0.021	-0.013	0.12	-0.050	0.46	8,445	0.004	0 (1.0)
rs35349911	43,785,255	Downstream	T/C	0.43	-0.001	0.90	0.028	0.23	-0.013	0.62	8,445	0.54	0 (0.52)
rs943072	43,795,968	Downstream	A/C	0.09	-0.003	0.82	0.003	0.88	-0.016	0.73	8,445	0.86	0 (0.92)
rs145023524	43,819,046	Downstream	A/G	0.006	-0.006	0.91	-0.002	0.55	0.12	0.40	8,445	0.62	0 (0.66)
rs55663434	43,820,609	Downstream	A/G	0.015	-0.031	0.39	-0.015	0.59	0.034	0.73	8,445	0.38	0 (0.83)

#SNPs in LD (r<sup>2</sup>>0.4). EAF: Effect allele frequency. I<sup>2</sup>: heterogeneity at meta-analyses level. P<sub>HET</sub>: P value for heterogeneity

**Table S5: Evaluation of possible regulatory variants in the *VEGFA* locus.**

SNP	Position (hg19/chr 6)	Open chromatin (DNase seq)	Conser- vation	H3K27 acetylation marks	H3K4Me1 marks	Chip-Seq signal	SNP located in TFX BS
rs729761 G>T	43804571	+	-	-	+	+ (RCOR1, NR2F2, TEAD4, GATA2, TAL1)	-
rs2396083 G>C	43804808	(+)	-	-	+	-	-
rs2396084 G>A	43804825	(+)	-	-	+	-	-
rs744103 A>T	43805362	-	-	-	(+)	+ (CMYC)	-
rs10223666 C>G	43805502	(+)	+	-	(+)	-	-
rs1317983* C>T	43806335	-	-	+	+	+ (RAD21, CTCF)	-
<b>rs881858</b> G>A	43806609	+	+	+	+	+ (CEBPB)	+ (CEBPB & CHOP)
rs9472135 T>C	43809802	-	-	-	-	-	-
rs9472137 T>C	43810469	-	-	-	-	-	-
rs9369425 G>A	43810974	+	-	+	+	+ (MAFF)	-
rs9369427 A>C	43811430	-	-	(+)	+	+ (POLR2R, GATA2)	-
<b>rs9472138</b> C>T	43811762	-	-	-	+	+ (MAFK, CEBPB)	-
rs1536304 T>C	43817837	-	-	-	+	+ (FOXAI)	-#
rs7758685 G>A	43825266	-	-	-	-	-	-
rs9394969 G>T	43825459	-	-	-	-	-	-
<b>rs943080</b> T>C	43826627	-	+	-	+	-	-
<b>rs4711751</b> T>C	43828582	-	-	-	-	-	-

\*rs1317983 is located just 274nt 5' of rs881858 and these SNPs share peak for H3K27Ac and H3K4Me1 marks. #FOXAI site is 2 nt 3' of SNP, but position is not important for binding. TFX BS: Transcription factor binding site. This table was compiled based on data in Fig. S3.

**Table S6: Changes from baseline to follow-up (5yr) in the Inter99 cohort among normal glucose tolerant subjects at baseline for GG, GA and AA genotypes of *VEGFA* rs881858**

<i>VEGFA</i> rs881858	GG	GA	AA	N total	*Effect	SE	P
Δ TSH (mIU/L)	0.16 (-0.08, 0.47)	0.12 (-0.17, 0.46)	0.14 (-0.17, 0.53)	2843	-0.024	0.088	0.78
Δ Free T4 (pmol/L)	-0.5 (-1.4, 0.4)	-0.3 (-1.2, 0.8)	-0.3 (-1.3, 0.6)	2856	-0.027	0.067	0.68
Δ BMI (kg/m <sup>2</sup> )	0.61 ± 1.6	0.41 ± 1.8	0.43 ± 1.6	3466	0.033	0.046	0.47
Δ HbA1c (%)	-0.05 (-0.17, 0.13)	-0.05 (-0.16, 0.13)	-0.06 (-0.17, 0.13)	2991	0.002	0.007	0.69
Δ Fasting plasma glucose (mmol/L)	-0.1 (-0.3, 0.2)	-0.1 (-0.3, 0.2)	-0.1 (-0.3, 0.2)	2985	0.005	0.014	0.69
Δ 2-hour glucose during OGTT (mmol/L)	-0.1 (-0.8, 0.7)	0 (-0.9, 0.8)	-0.1 (-1.0, 0.8)	2965	0.072	0.042	0.091
Δ Fasting serum insulin (pmol/L)	-2.0 (-12.7, 8)	-1.0 (-11, 9)	-1.0 (-11, 8)	2981	-0.72	0.68	0.28
Δ HOMA-IR	-0.09 (-0.50, 0.31)	-0.05 (-0.46, 0.34)	-0.04 (-0.46, 0.33)	2979	-0.025	0.028	0.36

\* G allele of *VEGFA* as the effect allele assuming an additive genetic model. N total is the number of normal glucose tolerant subjects at baseline who had genotype and phenotype information available during follow-up as well. Values correspond to median (interquartile range) in non-transformed traits

**Table S7.** Formulas used for calculating insulinogenic index, the Matsuda insulin sensitivity index, the disposition index and HOMA-IR

Trait	Measurement or calculation
Insulinogenic index	$(\text{Serum insulin at 30-min (pmol/l)} - \text{fasting serum insulin (pmol/l)}) / (\text{plasma glucose at 30-min (mmol/l)} - \text{fasting plasma glucose (mmol/l)})$
$ISI_{\text{Matsuda}}$	$(10,000 / \sqrt{(\text{fasting plasma glucose (mmol/l)} \times 18 \times \text{fasting serum insulin (pmol/l)} / 6) \times (\text{mean plasma glucose (mmol/l)} \times 18 \times \text{mean serum insulin (pmol/l)} / 6 \text{ during OGTT})})$
Disposition index	The insulinogenic index $\times ISI_{\text{Matsuda}}$
HOMA-IR	$((\text{fasting serum insulin (pmol/l)} / 6) \times (\text{fasting plasma glucose (mmol/l)})) / 22.5$