The American Journal of Human Genetics, Volume *89*

Supplemental Data

Variants Near *FOXE1* Are Associated with Hypothyroidism and Other Thyroid Conditions: Using Electronic Medical Records for Genome- and Phenome-wide Studies

Joshua C. Denny, Dana C. Crawford, Marylyn D. Ritchie, Suzette J. Bielinski, Melissa A. Basford, Yuki Bradford, High Seng Chai, Lisa Bastarache, Rebecca Zuvich, Peggy Peissig, David Carrell, Andrea H. Ramirez, Jyotishman Pathak, Russell A. Wilke, Luke Rasmussen, Xiaoming Wang, Jennifer A. Pacheco, Abel N. Kho, M. Geoffrey Hayes, Noah Weston, Martha Matsumoto, Peter A. Kopp, Katherine M. Newton, Gail P. Jarvik, Rongling Li, Teri A. Manolio, Iftikhar J. Kullo, Christopher G. Chute, Rex L. Chisholm, Eric B. Larson, Catherine A. McCarty, Daniel R. Masys, Dan M. Roden, and Mariza de Andrade

Table of Contents

Table S1: Minor Allele Frequency (MAF) for SNPs near *FOXE1* by site in Unmatched Analysis.

			GHC		Marshfield		Mayo		Northwestern		Vanderbilt	
Chr	SNP	Postion	Case MAF N=233	Control MAF N=1884	Case MAF N=514	Control MAF N=1187	Case MAF N=233	Control MAF N=1884	Case MAF N=92	Control MAF N=470	Case MAF N=81	Control MAF N=352
9	rs925489	99586421	0.3098	0.3582	0.2724	0.3353	0.3026	0.3546	0.2473	0.3347	0.2346	0.3366
9	rs7850258	99588834	0.3098	0.3582	0.2724	0.3353	0.3004	0.3547	0.2446	0.3326	0.2346	0.3366
9	rs965513	99595930	0.3086	0.3586	0.2743	0.3357	0.3026	0.3551	0.2473	0.333	0.2346	0.3381
9	rs10759944	99596793	0.3111	0.3582	0.2724	0.3336	0.3026	0.3536	0.2473	0.3337	0.2346	0.3352

Table S2: Characteristics of hypothyroidism cases and controls.
Plus-minus values are mean±SD.

	Unmatched Analysis		Matched Analysis		Replication Set	
	Cases N=1317	Controls N=5053	Cases N=1314	Controls N=3150	Cases N=263	Controls N=1616
Median birth decade	1930	1930	1930	1930	1930	1940
Age (yr)[a]	68.7±14.0	60.7±12.5	68.3±14.0	63.1±12.1	58.9±14.2	53.9±13.5
Female (%)	73.0	48.3	73.0	67.8	78.7	43.9
Follow-up (yr)[b]	19.3±9.4	15.5±9.5	19.0±9.5	17.2±9.8	11.9±4.9	10.1±5.6

[a]Age is calculated as the first age matching case or control definition (e.g., a billing code, laboratory value, or medication).
[b]Follow-up is calculated as the number of years the patients was observed by billing codes.

Table S3: SNPs associated with hypothyroidism at $p<10^{-6}$ in the unmatched and matched analyses.
Logistic regression adjusted for birth decade, sex, and site of ascertainment.

SNP	Chr	Position	Minor allele	Nearest Gene	Minor Allele Frequency		Unmatched Analysis*		Matched Analysis	
					Case	Control	OR (95% CI)	p-value	OR (95% CI)	p-value
rs7850258	9	99588834	A	*FOXE1*	0.285	0.348	0.74 (0.67, 0.82)	$3.96x10^{-9}$	0.75 (0.67, 0.83)	$1.42x10^{-8}$
rs965513	9	99595930	A	*FOXE1*	0.286	0.348	0.74 (0.67, 0.82)	$4.19x10^{-9}$	0.75 (0.67, 0.83)	$1.52x10^{-8}$
rs925489	9	99586421	C	*FOXE1*	0.286	0.348	0.74 (0.67, 0.82)	$4.68x10^{-9}$	0.75 (0.67, 0.83)	$1.58x10^{-8}$
rs10759944	9	99596793	A	*FOXE1*	0.286	0.347	0.75 (0.68, 0.83)	$8.19x10^{-9}$	0.75 (0.68, 0.83)	$2.53x10^{-8}$
rs4979402	9	116262496	G	*DFNB31*	0.288	0.247	1.29 (1.16, 1.42)	$1.23x10^{-6}$	1.24 (1.12, 1.38)	$6.76x10^{-5}$
rs4979397	9	116259709	T	*DFNB31*	0.286	0.246	1.28 (1.16, 1.42)	$1.91x10^{-6}$	1.23 (1.11, 1.37)	$9.56x10^{-5}$
rs1877432	9	99583701	A	9q22.3	0.437	0.382	1.25 (1.14, 1.37)	$1.99x10^{-6}$	1.25 (1.14, 1.38)	$3.90x10^{-6}$
rs1408528	9	116260594	C	*DFNB31*	0.286	0.246	1.28 (1.16, 1.42)	$2.07x10^{-6}$	1.23 (1.11, 1.37)	$1.01x10^{-4}$
rs17827152	4	55173443	A	4q12; closest to *KIT*	0.263	0.219	1.28 (1.15, 1.42)	$3.20x10^{-6}$	1.29 (1.16, 1.44)	$5.20x10^{-6}$
rs1535971	9	116269221	T	*DFNB31*	0.299	0.260	1.27 (1.15, 1.40)	$3.65x10^{-6}$	1.22 (1.10, 1.35)	$1.73x10^{-4}$
rs17043990	3	72963160	C	*SHQ1*	0.0091	0.0024	3.71 (2.08, 6.60)	$8.34x10^{-6}$	4.37 (2.16, 8.83)	$1.58x10^{-6}$

DFBB31 SNPs are in the gene
FOXE1 SNPs are from 58kb to 71kb to the gene
DIRAS2 SNP is in the gene

Table S4: Association of variants near *FOXE1* by eMERGE site in the matched analysis.

		All Sites		GHC		Marshfield		Mayo		Northwestern		Vanderbilt	
SNP	Coded/ Non-Coded	Odds Ratio (95% CI)	P-value	Odds Ratio (95% CI)	P-value	Odds Ratio (95% CI)	P-value	Odds Ratio (95% CI)	P-value	Odds Ratio (95% CI)	P-value	Odds Ratio (95% CI)	P-value
rs7850258	A/G	0.74 (0.67, 0.83)	1.42E-08	0.79 (0.66, 0.95)	0.011	0.74 (0.62, 0.89)	0.0011	0.77 (0.61, 0.96)	0.020	0.54 (0.35, 0.84)	0.0039	0.69 (0.49, 0.97)	0.030
rs965513	A/G	0.75 (0.67, 0.83)	1.52E-08	0.78 (0.65, 0.94)	0.0088	0.75 (0.62, 0.90)	0.0015	0.77 (0.61, 0.97)	0.023	0.54 (0.35, 0.83)	0.0031	0.69 (0.49, 0.97)	0.030
rs925489	C/T	0.75 (0.67, 0.83)	1.58E-08	0.79 (0.66, 0.95)	0.011	0.74 (0.62, 0.89)	0.0011	0.77 (0.62, 0.97)	0.024	0.54 (0.35, 0.83)	0.0031	0.69 (0.49, 0.97)	0.030
rs10759944	A/G	0.75 (0.68, 0.83)	2.53E-08	0.80 (0.66, 0.95)	0.013	0.75 (0.62, 0.9)	0.0014	0.78 (0.62, 0.97)	0.027	0.54 (0.35, 0.83)	0.0031	0.69 (0.49, 0.97)	0.028
rs1877432	A/G	1.25 (1.14, 1.38)	3.90E-06	1.14 (0.96, 1.35)	0.123	1.35 (1.15, 1.60)	0.00033	1.13 (0.91, 1.40)	0.27	1.73 (1.18, 2.53)	0.0038	1.33 (0.94, 1.86)	0.104

Table S5: Strongest associations for HLA region in unmatched analysis for European Americans. Analysis is adjusted by birth decade, sex, site, and the first principle component. All associations with $p<1x10^{-2}$ are shown below.

CHR	SNP	Chromosomal Position	Minor Alleles	Odds Ratio	P
6	rs2735076	30051469	T	1.20	$3.61x10^{-4}$
6	rs5025708	30063178	G	1.26	$3.95x10^{-4}$
6	rs3094170	29937244	A	1.23	$5.38x10^{-4}$
6	rs9394167	33862374	A	0.86	$9.70x10^{-4}$
6	rs3025643	29677934	T	1.16	$9.76x10^{-4}$
6	rs3806109	33878348	G	1.16	$1.17x10^{-3}$
6	rs9380167	30382440	G	0.85	$1.34x10^{-3}$
6	rs3130253	29741991	A	1.28	$1.35x10^{-3}$
6	rs9277912	33232636	T	1.20	$1.46x10^{-3}$
6	rs4711348	33844273	G	0.87	$1.51x10^{-3}$
6	rs2296748	33853049	T	1.15	$1.67x10^{-3}$
6	rs9266409	31444547	C	0.84	$1.75x10^{-3}$
6	rs6933050	31451611	C	0.84	$1.76x10^{-3}$
6	rs1063355	32735692	A	0.87	$1.77x10^{-3}$
6	rs3130252	29739723	C	1.20	$1.80x10^{-3}$
6	rs1535950	33893805	G	1.15	$1.86x10^{-3}$
6	rs7770216	31448590	T	0.84	$2.11x10^{-3}$
6	rs2281820	33876875	T	0.87	$2.23x10^{-3}$
6	rs3135050	29736416	T	1.19	$2.37x10^{-3}$
6	rs29230	29684372	C	0.83	$2.38x10^{-3}$
6	rs29225	29689020	C	0.79	$2.45x10^{-3}$
6	rs9261387	30169340	T	1.26	$2.49x10^{-3}$
6	rs740882	29683435	A	0.79	$2.54x10^{-3}$
6	rs2477232	34001130	A	0.80	$2.64x10^{-3}$
6	rs3132679	30183822	A	1.20	$2.68x10^{-3}$
6	rs9264869	31379609	G	1.22	$2.70x10^{-3}$
6	rs3948793	32867426	T	1.15	$2.71x10^{-3}$
6	rs9277965	33364375	A	1.19	$2.80x10^{-3}$
6	rs9357097	30393100	T	0.86	$2.87x10^{-3}$
6	rs9276711	32865275	T	1.15	$2.92x10^{-3}$
6	rs3117073	29652804	G	1.16	$2.92x10^{-3}$
6	rs3132726	29931585	T	1.16	$3.04x10^{-3}$
6	rs4713610	33215933	G	1.19	$3.16x10^{-3}$
6	rs1633068	29832289	A	1.18	$3.19x10^{-3}$
6	rs241425	32912887	T	0.88	$3.23x10^{-3}$
6	rs1610628	29836188	G	1.18	$3.25x10^{-3}$
6	rs9276644	32853021	C	1.14	$3.40x10^{-3}$
6	rs12665339	30709211	G	0.84	$3.51x10^{-3}$
6	rs6457769	34339548	C	1.22	$3.65x10^{-3}$
6	rs942496	33896033	T	0.84	$3.65x10^{-3}$

6	rs932338	29814912	T	1.18	$3.75x10^{-3}$
6	rs7765379	32788906	G	1.22	$3.83x10^{-3}$
6	rs3957165	33892842	G	0.85	$3.84x10^{-3}$
6	rs9368781	33976857	A	0.81	$4.00x10^{-3}$
6	rs2523710	31558888	T	0.84	$4.23x10^{-3}$
6	rs942510	33985605	T	0.81	$4.45x10^{-3}$
6	rs2235498	33238408	T	1.16	$4.83x10^{-3}$
6	rs2844673	31069905	A	0.83	$5.12x10^{-3}$
6	rs9258883	29945167	C	1.15	$5.18x10^{-3}$
6	rs2517409	31072372	G	0.83	$5.18x10^{-3}$
6	rs2252926	31074283	G	0.83	$5.18x10^{-3}$
6	rs9366829	33881177	G	1.13	$5.31x10^{-3}$
6	rs9264868	31379580	G	1.20	$5.72x10^{-3}$
6	rs1002985	32960426	T	0.74	$5.75x10^{-3}$
6	rs6913294	31379103	A	0.80	$5.85x10^{-3}$
6	rs9380376	33832772	T	1.14	$5.90x10^{-3}$
6	rs2523898	31101512	T	1.13	$6.14x10^{-3}$
6	rs9391714	31353059	A	0.70	$6.17x10^{-3}$
6	rs12211633	34158895	C	1.14	$6.30x10^{-3}$
6	rs9380374	33785534	A	1.14	$6.34x10^{-3}$
6	rs4713447	31270942	G	0.89	$6.47x10^{-3}$
6	rs3868075	31275794	C	0.89	$6.47x10^{-3}$
6	rs12206652	34173960	G	0.87	$6.49x10^{-3}$
6	rs1264420	30683582	T	1.14	$6.62x10^{-3}$
6	rs9405048	30778271	T	0.85	$6.62x10^{-3}$
6	rs9266773	31460327	C	1.22	$6.70x10^{-3}$
6	rs9357152	32772938	G	1.15	$6.77x10^{-3}$
6	rs1906953	34144424	T	1.19	$7.10x10^{-3}$
6	rs2524044	31364732	C	1.17	$7.62x10^{-3}$
6	rs9277920	33233703	A	1.13	$7.75x10^{-3}$
6	rs3778624	30410579	G	0.87	$7.80x10^{-3}$
6	rs753725	30998850	A	1.12	$7.81x10^{-3}$
6	rs2532919	31041954	G	0.84	$7.84x10^{-3}$
6	rs9405015	31263782	T	0.59	$7.93x10^{-3}$
6	rs2523864	31126525	A	1.13	$7.96x10^{-3}$
6	rs2524077	31351582	A	1.17	$8.53x10^{-3}$
6	rs2532934	31002738	C	1.12	$8.94x10^{-3}$
6	rs2844651	31008643	A	1.12	$9.02x10^{-3}$
6	rs2442749	31460019	G	1.13	$9.27x10^{-3}$
6	rs2844498	31584833	A	1.12	$9.49x10^{-3}$
6	rs9366755	30394708	T	0.85	$9.60x10^{-3}$
6	rs1061535	30045903	A	1.15	$9.98x10^{-3}$

Table S6: *FOXE1* SNPs associated with hypothyroidism in the Replication Analysis.
The Hardy-Weinberg equilibrium for each SNP in cases and controls was p>0.44.
MAF=Minor allele frequency.

SNP	Position	Nearest Gene	Coded/ NonCoded	Odds Ratio (95% CI)	MAF control	MAF case	p-value
rs7850258	99588834	*FOXE1*	A/G	0.60 (0.48-0.74)	0.35	0.25	$5.7x10^{-6}$
rs965513	99595930	*FOXE1*	A/G	0.59 (0.47-0.74)	0.35	0.25	$4.8x10^{-6}$
rs925489	99586421	*FOXE1*	C/T	0.59 (0.46-0.74)	0.35	0.25	$1.1x10^{-5}$
rs10759944	99596793	*FOXE1*	A/G	0.60 (0.47-0.75)	0.34	0.25	$7.6x10^{-6}$

Figure S1: Algorithm for identifying primary hypothyroidism cases and controls.
Of the PH cases, 204 (15% of the cases) had testing for anti-thyroperoxidase antibodies in their EMR, and 111 (54%) individuals had positive antibody tests as defined by EMR-specific reference ranges.

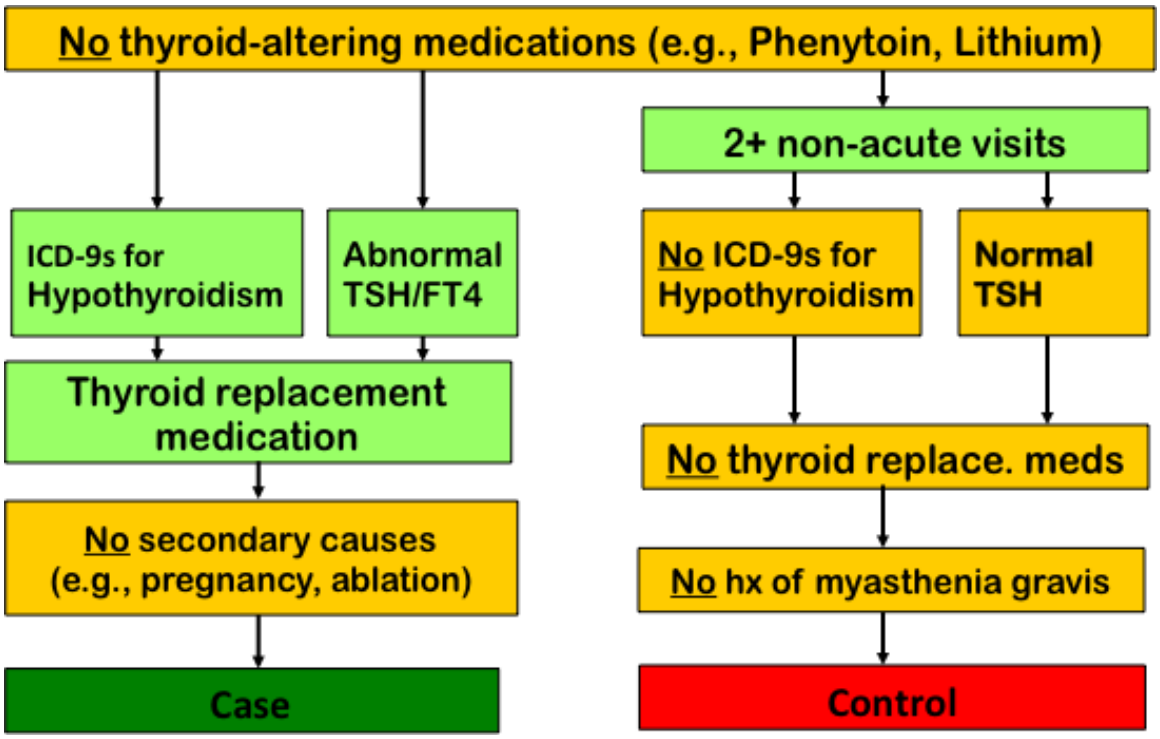

Figure S2: Analysis of Linkage Disequilibrium from 1000genomes Pilot 1 CEU samples to *FOXE1* locus (rs7850258).

All SNPs in *FOXE1* had r^2<0.42 using 1000 genomes data.

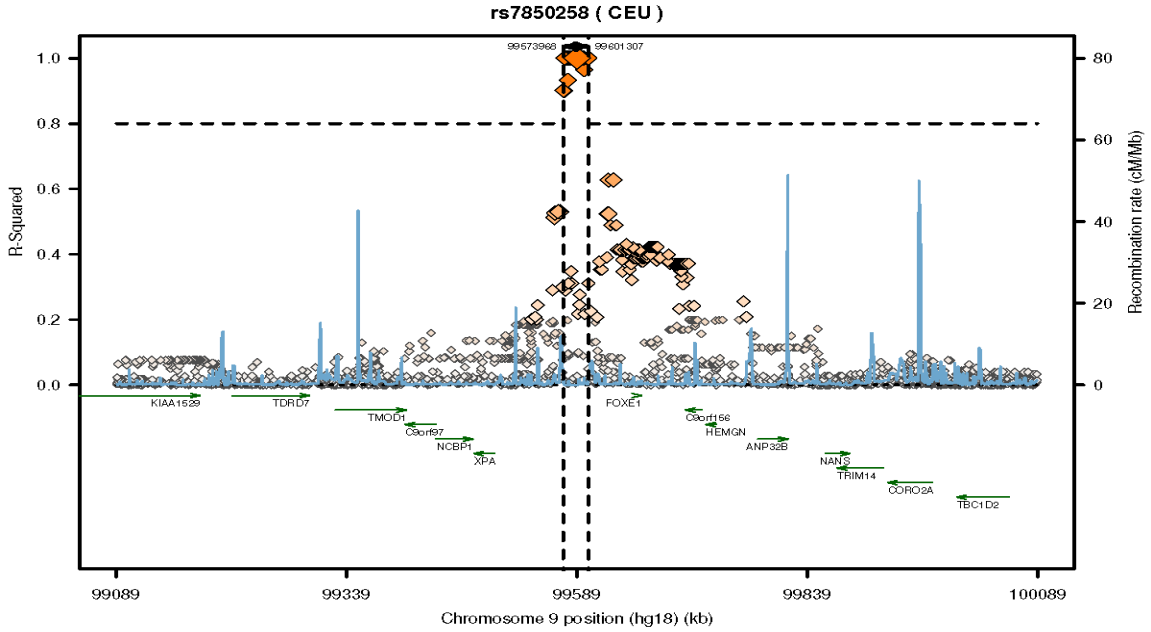

Figure S3: Genome-wide Association Analysis from non-TSH controls.
Single nucleotide polymorphism (SNP) tests of association (logistic regression) assuming an additive genetic model adjusted for sex and age. The red horizontal line indicates $P=5x10^{-8}$, the threshold for genome-wide significance.

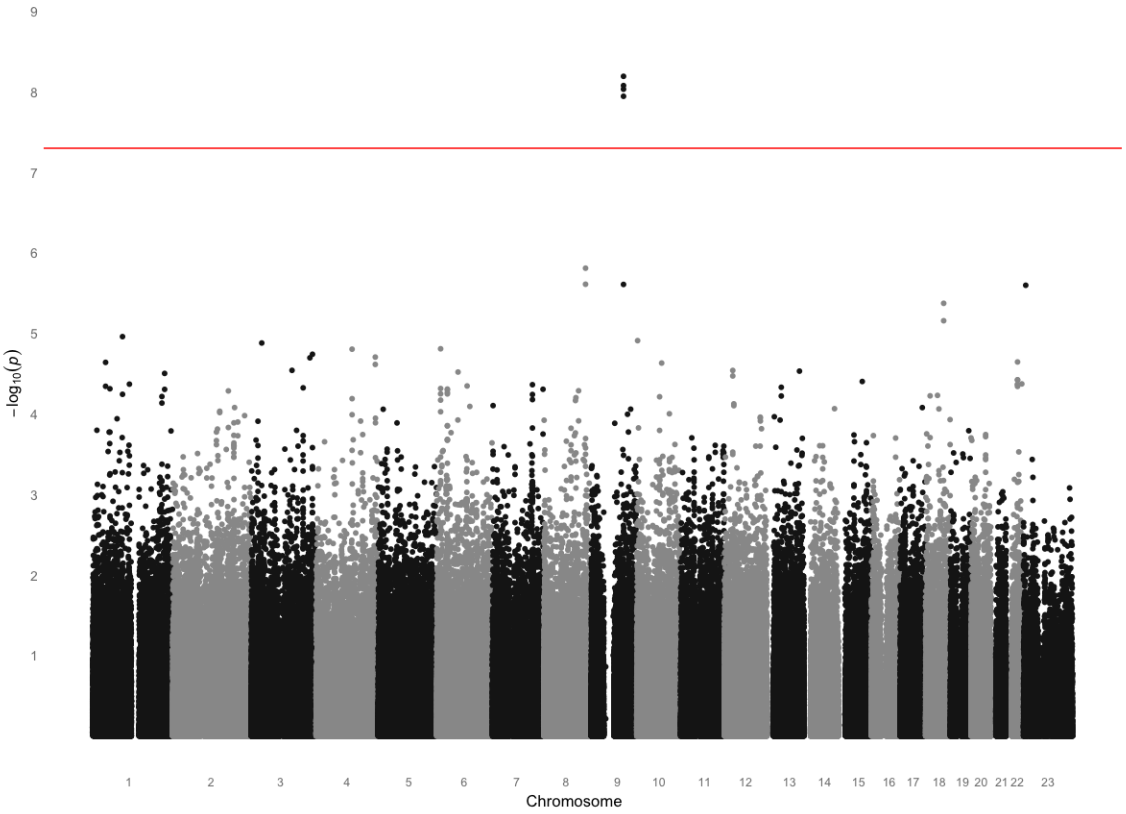

Figure S4: Result of Genome-wide Association Analysis from Matched Analysis.
Single nucleotide polymorphism (SNP) tests of association (logistic regression) assuming an additive genetic model adjusted for sex, birth decade, and study site. The red horizontal line indicates $P=5x10^{-8}$, the threshold for genome-wide significance.

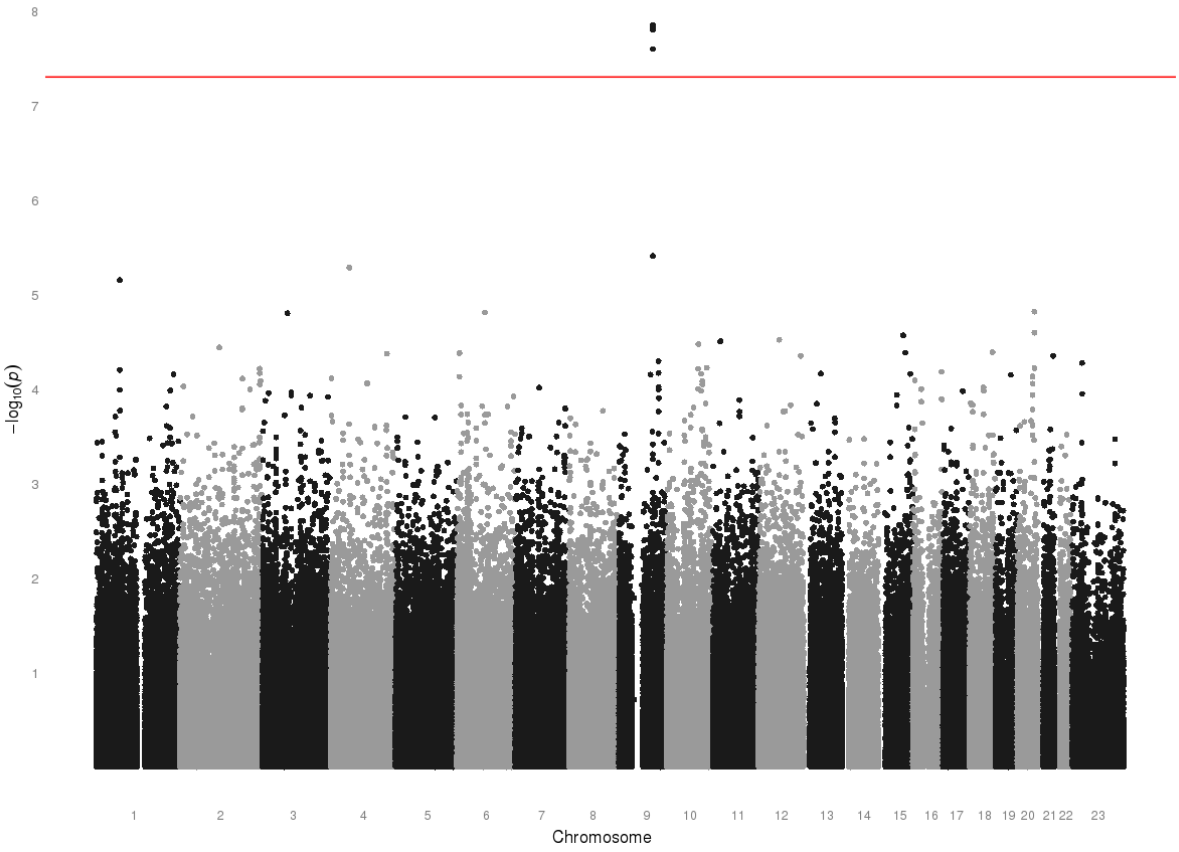

Figure S5: Location and linkage disequilibrium (r^2) for SNPs surrounding *FOXE1* from matched analysis.

Recombination rates were generated using imputation from the eMERGE dataset.

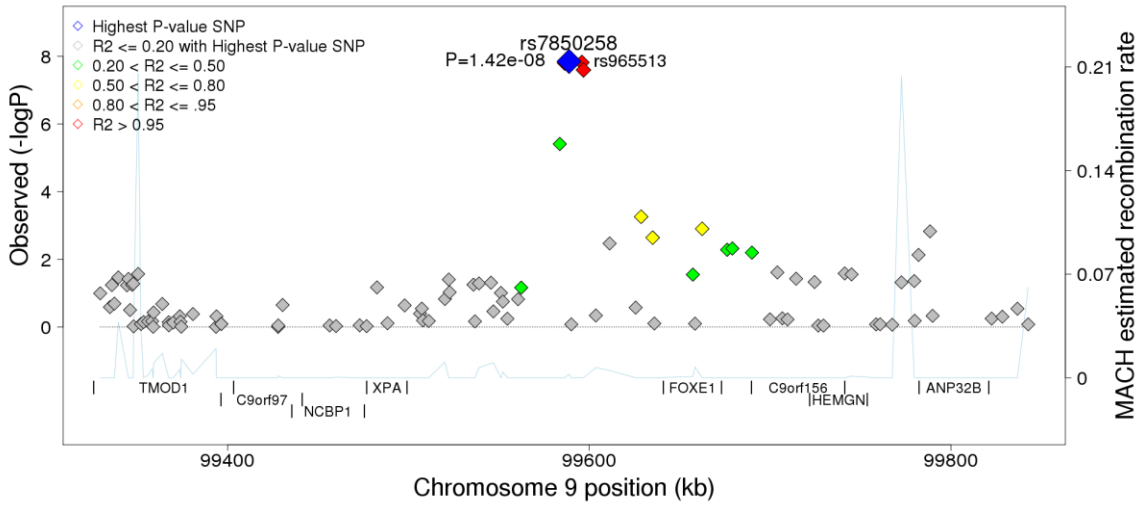

Figure S6: Q-Q plot for unmatched analysis.

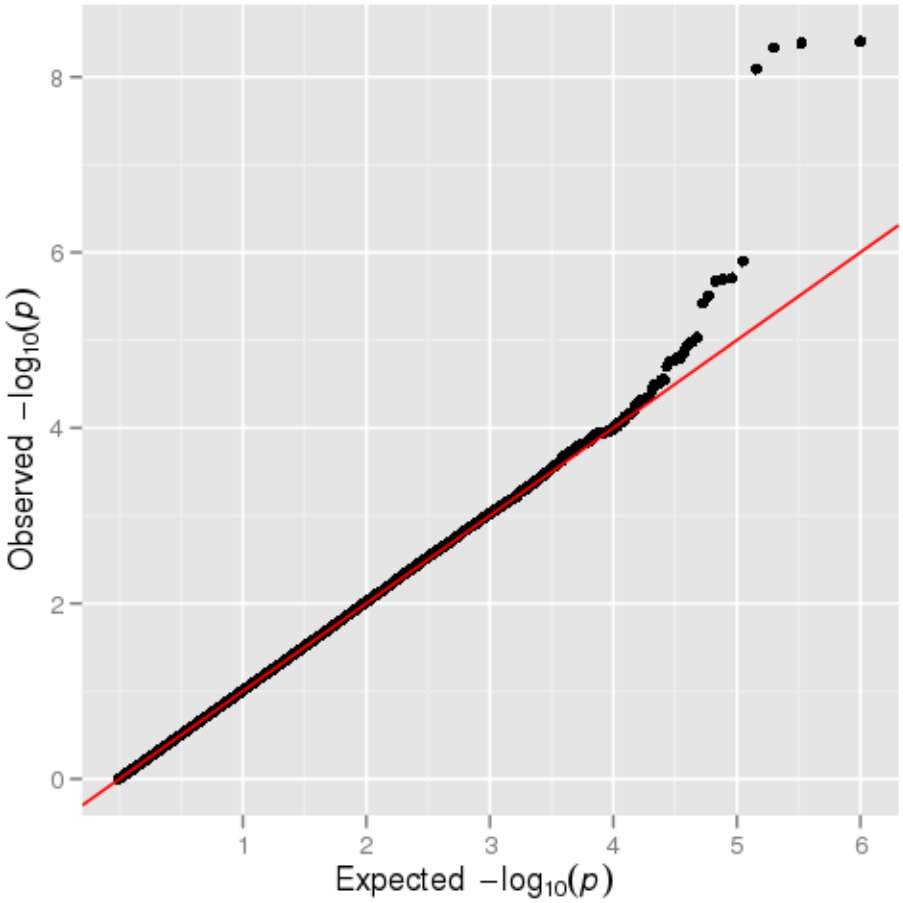

Figure S7: Q-Q plot for comparison to non-TSH controls.

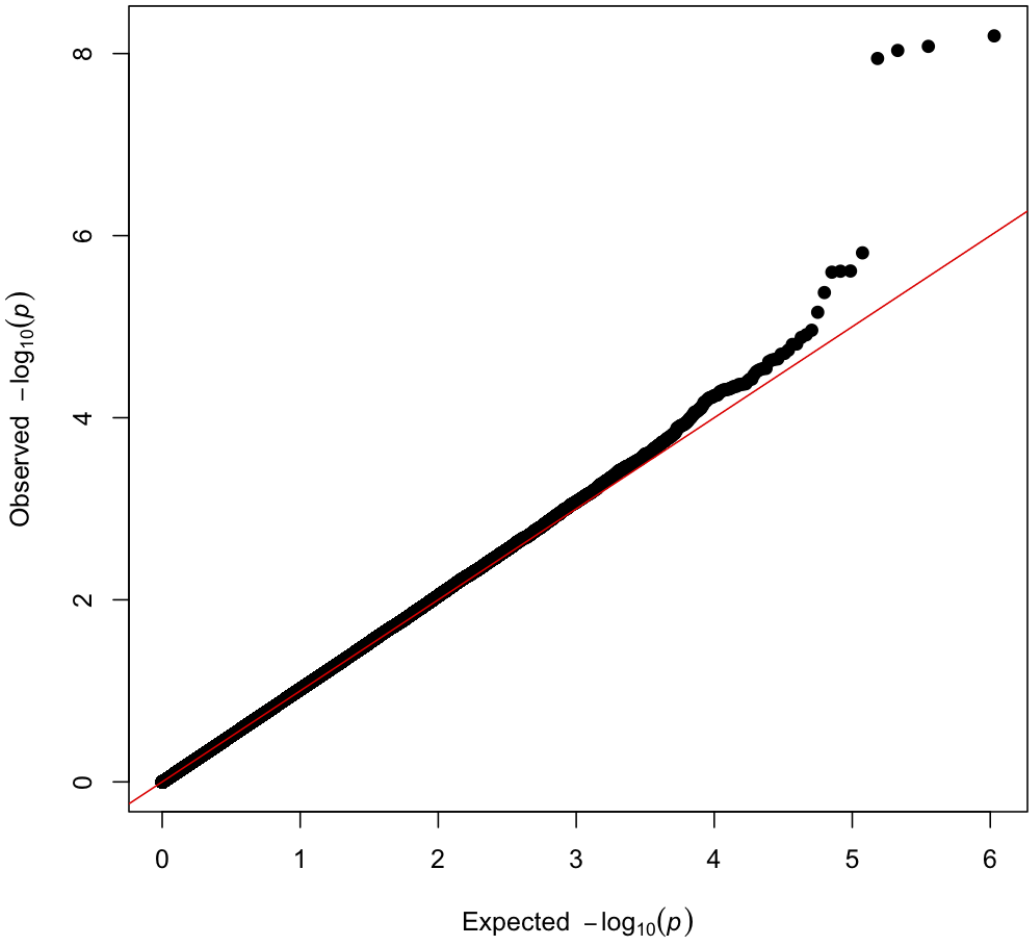

Figure S8: Q-Q plot for matched analysis.