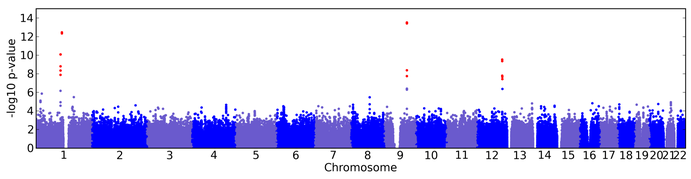
Novel associations for hypothyroidism include known autoimmune risk loci

B.T. Naughton[1], N. Eriksson[1], J.Y. Tung[1], A.K. Kiefer[1], D.A. Hinds[1], U. Francke[1,2], J.L. Mountain[1], C.B. Do[1].

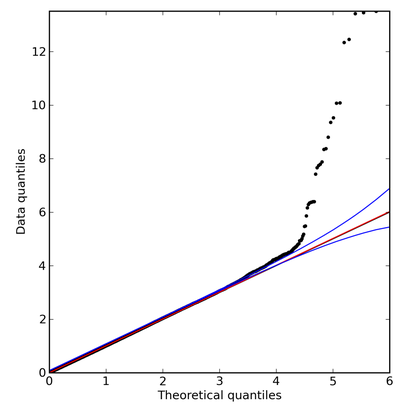
[1]23andMe, Inc, Mountain View, CA; [2]Stanford University, Stanford, CA.

Introduction

Hypothyroidism is the most common thyroid disorder, affecting about 5% of the general population. Here we present the first large genome-wide association study of hypothyroidism, in 2,564 cases and 24,448 controls from the customer base of 23andMe, Inc., a personal genetics company. We identify four genome-wide significant associations, two of which are well known to be involved with a large spectrum of autoimmune diseases: rs6679677 near **PTPN22** and rs3184504 in **SH2B3** (p-values 3.5e-13 and 3.0e-11, respectively). We also report associations with rs4915077 near **VAV3** (p-value 8.3e-11), another gene involved in immune function, and rs965513 near **FOXE1** (p-value 3.1e-14). Of these, the association with PTPN22 confirms a recent small candidate gene study, and FOXE1 was previously known to be associated with thyroid-stimulating hormone (TSH) levels. Although SH2B3 has been previously linked with a number of autoimmune diseases, this is the first report of its association with thyroid disease. The VAV3 association is novel. These results suggest heterogeneity in the genetic etiology of hypothyroidism, implicating genes involved in both autoimmune disorders and thyroid function. Using a genetic risk profile score based on the top association from each of the four genome-wide significant regions in our study, the relative risk between the highest and lowest deciles of genetic risk is 2.1.



Manhattan plot. Negative log p-values for SNPs by genome position. Genome-wide significant SNPs are shown in red.



Quantile-quantile plot. Observed p-values versus theoretical p-values under the null hypothesis of no association. The genomic control inflation factor for the study was 0.98 and is indicated by the red line.

Results

In these plots, circles represent unannotated SNPs, upside-down triangles represent non-synonymous variants, and boxes with an x are SNPs in regions that are highly conserved across 44 placental mammals. Colors depict the squared correlation (r^2) of each SNP with the most associated SNP (e.g., rs4915077, shown in purple). Gray indicates SNPs for which r^2 information was missing. Plots were produced using the LocusZoom program

	Number	Female	Male
Control	24448	8917 (36.5%)	15531 (63.5%)
Case	2564	1891 (73.8%)	673 (26.2%)
	< 45	46–55	56–65	> 65
Control	11850 (48.5%)	3903 (16.0%)	4673 (19.1%)	4022 (16.5%)
Case	580 (22.6%)	419 (16.3%)	801 (31.2%)	764 (29.8%)
	V1	V2	V3
Control	260 (1.1%)	16482 (67.4%)	7706 (31.5%)
Case	27 (1.1%)	1845 (72.0%)	692 (27.0%)

Participants broken down by sex, age, and genotyping platform. V1, V2, and V3 refer to the three platforms used in this study, see Methods.

Table 2. Statistics for genome-wide significant SNPs and selected replications.

SNP	Chr.	Pos.	Region	Alleles	MAF	HWE	*p*-value	OR
rs965513	9	99595930	*FOXE1*	G/A	0.331	0.23	$3.1 \cdot 10^{-14}$	0.778 (0.73 - 0.83)
rs6679677	1	114105331	*PTPN22*	C/A	0.090	0.74	$3.5 \cdot 10^{-13}$	1.445 (1.31 - 1.59)
rs2476601	1	114179091	*PTPN22*	G/A	0.091	0.66	$4.6 \cdot 10^{-13}$	1.439 (1.31 - 1.58)
rs4915077	1	108167539	*VAV3*	T/C	0.084	0.32	$8.3 \cdot 10^{-11}$	1.397 (1.27 - 1.54)
rs3184504	12	110368991	*SH2B3*	T/C	0.499	0.31	$3 \cdot 10^{-10}$	0.823 (0.77 - 0.87)
rs378836	14	35561627	*NKX2-1*	G/A	0.432	0.96	$3.2 \cdot 10^{-5}$	1.137 (1.07 - 1.21)
rs944289	14	35718997	*NKX2-1*	T/C	0.427	0.61	0.096	1.053 (0.99 - 1.12)
rs1472565	1	19627617	*CAPZB*	T/C	0.476	0.78	$7.9 \cdot 10^{-6}$	1.147 (1.08 - 1.22)
rs10799824	1	19713761	*CAPZB*	G/A	0.153	0.31	0.066	0.853 (0.72 - 1.01)
rs4704397	5	76554198	*PDE8B*	G/A	0.388	0.5	0.021	1.152 (1.02 - 1.30)
rs755109	9	99736024	*FOXE1*	T/C	0.367	0.11	0.73	0.976 (0.86 - 1.10)
rs2235544	1	54148158	*DIO1*	A/C	0.496	0.16	0.76	0.991 (0.93 - 1.05)

All genomic positions are given with respect to NCBI build 36.3. Alleles are listed as major/minor and are specified for the forward strand. Odds ratios are per copy of the minor allele. Two SNPs are listed for *PTPN22*; of these, rs2476601 is the non-synonymous change R620W.

Conclusion

We have presented two novel associations with hypothyroidism: the non-synonymous change R262W in SH2B3, and the SNP rs4915077 near VAV3. SH2B3 has been associated with a host of autoimmune diseases and thus is a good candidate for hypothyroidism. VAV3 has not yet been associated with autoimmune disease or thyroid function. However, it has been proposed as the candidate gene in the Idd18.1 region linked with type 1 diabetes in mouse, and the Vav1/Vav2/Vav3 family is necessary for adaptive immune function in mouse. VAV3 is also expressed in the thyroid and is down-regulated in a subset of thyroid tumors.

We have found one more association with a mutation well known to be involved in autoimmune disease: R620W in PTPN22. This association has been observed before in a small candidate gene study of Hashimoto thyroiditis. That study found, in a sample of 194 cases and 2064 controls, an OR of 1.77 (1.31–2.40) for rs2476601/R620W. We observe an OR for this SNP of 1.42 (1.29-1.57), which is within their confidence interval, despite slightly different phenotypes.

Acknowledgments

We thank the customers of 23andMe who answered surveys and participated in this research. Thanks to all the employees of 23andMe, who together have made this research possible, especially Kimberly Barnholt, Geoffrey Benton, Arnab Chowdry, Emily Drabant, Michael Macpherson.