

SCIENTIFIC REPORTS

OPEN

Common genetic variation in the autoimmune regulator (*AIRE*) locus is associated with autoimmune Addison's disease in Sweden

Daniel Eriksson^{1,2}, Matteo Bianchi³, Nils Landegren^{1,4}, Frida Dalin^{1,4}, Jakob Skov⁵, Lina Hultin-Rosenberg³, Argyri Mathioudaki³, Jessika Nordin³, Åsa Hallgren¹, Göran Andersson⁶, Karolina Tandre⁴, Solbritt Rantapää Dahlqvist⁷, Peter Söderkvist⁸, Lars Rönnblom⁴, Anna-Lena Hulting⁵, Jeanette Wahlberg^{8,9,10}, Per Dahlqvist⁷, Olov Ekwall^{11,12}, Jennifer R. S. Meadows³, Kerstin Lindblad-Toh^{3,13}, Sophie Bensing^{2,5}, Gerli Rosengren Pielberg³ & Olle Kämpe^{1,2,14}

Autoimmune Addison's disease (AAD) is the predominating cause of primary adrenal failure. Despite its high heritability, the rarity of disease has long made candidate-gene studies the only feasible methodology for genetic studies. Here we conducted a comprehensive reinvestigation of suggested AAD risk loci and more than 1800 candidate genes with associated regulatory elements in 479 patients with AAD and 2394 controls. Our analysis enabled us to replicate many risk variants, but several other previously suggested risk variants failed confirmation. By exploring the full set of 1800 candidate genes, we further identified common variation in the *autoimmune regulator* (*AIRE*) as a novel risk locus associated to sporadic AAD in our study. Our findings not only confirm that multiple loci are associated with disease risk, but also show to what extent the multiple risk loci jointly associate to AAD. In total, risk loci discovered to date only explain about 7% of variance in liability to AAD in our study population.

The predominating cause of primary adrenal insufficiency is the autoimmune destruction of the adrenal cortex, known as autoimmune Addison's disease (AAD)¹. Affected patients suffer from loss of the essential adrenal hormones cortisol and aldosterone². Although the symptoms typically develop slowly beginning with increasing fatigue, hyperpigmentation and weight loss, patients often present with acute adrenal insufficiency with abdominal pain, nausea and vomiting^{1,3}. Prompt diagnosis and steroid replacement therapy is essential in order to avoid the otherwise fatal course of disease¹. Autoantibodies targeting the adrenal enzyme 21-hydroxylase are highly specific diagnostic markers, and also confirm an autoimmune pathoetiology^{1,4,5}. The underlying causes of AAD are largely unknown, although the strong heritability and clear phenotype are incentives to genetic studies⁶.

AAD has a reported prevalence of 87–221 per million in European countries and the rarity of the disease has long rendered extensive genetic association studies unfeasible^{7–14}. Most patients with AAD develop additional tissue-specific autoimmune diseases such as type 1 diabetes and autoimmune thyroid disease, both of which

¹Department of Medicine (Solna), Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden.

²Department of Endocrinology, Metabolism and Diabetes Karolinska University Hospital, Stockholm, Sweden.

³Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden. ⁴Science for Life Laboratory, Department of Medical Sciences, Uppsala University, Uppsala, Sweden.

⁵Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. ⁶Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden. ⁷Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden. ⁸Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden. ⁹Department of Endocrinology, Linköping University, Linköping, Sweden. ¹⁰Department of Medical and Health Sciences, Linköping University, Linköping, Sweden. ¹¹Department of Pediatrics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.

¹²Department of Rheumatology and Inflammation Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. ¹³Broad Institute of MIT and Harvard, Cambridge, Massachusetts, United States of America. ¹⁴K.G. Jebsen Center for Autoimmune Diseases, Bergen, Norway. Correspondence and requests for materials should be addressed to D.E. (email: daniel.eriksson@ki.se)

have been studied in genome-wide association studies¹⁵. Many risk genes that were identified for common autoimmune diseases have subsequently also been investigated via candidate-gene studies in AAD, and in this way, genes such as *HLA-DRB1* and *CTLA4* have been found to be associated with AAD^{16–20}. Candidate-gene studies have also connected AAD with *PTPN22*²¹ and *CLEC16A*^{22,23} that are risk loci in several autoimmune diseases with complex inheritance.

AAAD does not display a Mendelian inheritance pattern and the trait is considered complex². However, in a small subset of patients, AAD appears as a component of the monogenic syndrome Autoimmune Polyglandular Syndrome type 1 (APS1) (Online Mendelian Inheritance in Man, # 240300). APS1 is caused by mutations in the autoimmune regulator gene (*AIRE*). The three main components of APS1 are AAD, chronic mucocutaneous candidiasis, and hypoparathyroidism, two of which are required for the clinical diagnosis^{24,25}. The vast majority of patients with APS1 display autoantibodies against interferon- α , interferon- ω , and interleukin-22^{26–29}. These biomarkers are highly specific for APS1 and can be used for identifying undiagnosed APS1 patients among patients with AAD³⁰. In addition to recessively inherited APS1, dominant missense mutations in *AIRE* have also been more recently described^{31–35}. Carriers of dominant *AIRE* mutations typically present a less severe phenotype with a later onset. Overall, APS1 accounts for only a minor fraction of all cases with AAD³⁰.

We recently reported a targeted sequencing study that enrolled a large case group of 479 AAD patients from the Swedish Addison Register (SAR) and 1394 healthy controls; the SAR-Seq study³⁶. Exons and regulatory regions of 1853 genes of interest were covered, enabling novel discoveries in the *BACH2* locus with genome-wide significance³⁶. To date, this has been the most comprehensive genetic study of AAD. An important advantage of studies in rare diseases is that the general population is a sensible control group in genetic case-control studies. The SweGen Variant Frequency Database (1kSWE) harbours whole-genome variant frequencies from 1000 Swedish individuals³⁷. Since no disease information has been collected, the 1kSWE represents a cross-section of the Swedish population rather than a set of healthy individuals. Still, with an AAD prevalence ranging from 87 to 221 per million in European studies^{7–14}, the probability of including a significant number of AAD cases among a thousand Swedes is negligible (see Supplementary Table 1). Therefore, adding the allele counts from 1000 additional Swedish genomes could contribute to a more sensitive detection of disease risk loci with small effect sizes.

Here we present a study combining multiple AAD risk loci to evaluate their additive effects on disease risk and the age of disease onset. To expand the original SAR-Seq dataset and increase the number of detected variants severalfold, we imputed additional genotypes using haplotypes from the international 1000 genomes project³⁸. With imputation and additional controls from the 1kSWE, our analysis could replicate several previous associations. By exploring potentially novel risk loci, we could also associate common variants in the *AIRE* gene to sporadic AAD.

Results

Imputation expands the dataset threefold. To date, the SAR-Seq study has generated the most comprehensive dataset on AAD risk loci. With imputation, we expanded the original data with three times as many common single-nucleotide polymorphisms (SNPs)³⁶. By including allele counts from the 1kSWE³⁷, we were able to compare allele frequencies from 479 AAD cases and 2394 controls (see Supplementary Figs S1 and S2). We set the study-wide statistical significance level to 1.2×10^{-6} , a Bonferroni correction from $\alpha 0.05$ for the number of independent variants ($r^2 < 0.8$)³⁹. We then applied the Cochran-Armitage test for trend to associate risk alleles with AAD and examined *P* value distributions using quantile–quantile plots (see Supplementary Fig. S3). Both the SAR-Seq subjects and the subjects in 1kSWE had been thoroughly investigated with principal component analysis to exclude non-European samples. Nevertheless, we adjusted the overdispersion of the test statistic ($\lambda = 1.19$) with genomic control to normalize λ before further analyses (see Supplementary Fig. S3).

***CTLA4*, *BACH2*, *PTPN22* and *CLEC16A* are replicated risk loci in AAD.** AAD has been associated with several genes, some of which have never been replicated since their first discovery in a candidate-gene study. In our previous targeted sequencing study, we systematically included genes that had previously been linked with AAD or other autoimmune disorders³⁶. Our new imputed dataset contained 15 SNPs outside of the HLA region implicated in AAD in previous studies and allowed us to reinvestigate those suggested loci (Table 1). Except for *BACH2*, we were underpowered to associate the tabulated SNPs with AAD at our explorative study-wide significance level ($P < 1.2 \times 10^{-6}$). However, a number of these SNPs would have been significant in a candidate-gene study with just the few variants tested. Therefore, we regarded SNPs as replicated if they passed the nominal significance level ($\alpha 0.05$) Bonferroni corrected for the 15 SNPs investigated (Table 1). At this level of significance (P value $< 3.3 \times 10^{-3}$), we replicated the well-established associations of *PTPN22*²¹, *CTLA4*^{19,20,40,41}, and *CLEC16A*^{22,23}. *BACH2* had already been associated with AAD in this case group³⁶, but we could impute and confirm the SNP reported by Pazderskaet *al.*⁴². However, the association to variants in *PD-L1*⁴³, *GATA3*⁴⁴, and *CYP27B1*⁴⁴ could not be replicated even at the nominal significance level.

Common genetic variation in *AIRE* is associated to AAD. Our survey of the known risk loci consistently replicated several of the well-established associations. We next investigated the full range of variants detected in our array of captured exons, introns, and regulatory regions. Besides the strong peaks of association on chromosome 6, representing the HLA complex and the *BACH2* locus, we detected association to a novel locus on chromosome 21q22.3 (see Supplementary Fig. S4). A detailed description of the associated SNPs in the novel locus is presented in Table 2. The four tabulated SNPs were all study-wide significant and the top SNP exceeded the traditional cut off for genome-wide significance, 5×10^{-8} (rs9983695-C, OR = 0.37 (0.27–0.52), $P = 2.1 \times 10^{-8}$, MAF 0.04/0.11 in cases/controls). A closer look at the region revealed that all four SNPs were in strong linkage disequilibrium and located within the *AIRE* gene (Fig. 1, see Supplementary Fig. S5). The four associated SNPs were all located in introns 5 and 8, but did not co-localize with any known transcription factor

Locus	Variant	Chr:Position ^a	Minor Allele Frequency		Association	
			SAR-Seq Cases	All Controls	OR ^b	P value ^c
<i>PTPN22</i>	rs2476601-A	1:114377568	0.18	0.12	1.52 (1.26–1.84)	$6.1 \times 10^{-5*}$
<i>CTLA4</i>	rs231806-C	2:204709349	0.30	0.36	0.74 (0.64–0.87)	$5.8 \times 10^{-4*}$
<i>CTLA4</i>	rs231775-G	2:204732714	0.53	0.46	1.09 (0.95–1.25)	$5.5 \times 10^{-5*}$
<i>CTLA4</i>	rs11571302-T	2:204742934	0.34	0.41	0.77 (0.66–0.89)	$1.6 \times 10^{-3*}$
<i>CTLA4</i>	rs7565213-A	2:204743409	0.34	0.41	0.77 (0.66–0.89)	$1.6 \times 10^{-3*}$
<i>CTLA4</i>	rs11571297-C	2:204745003	0.36	0.42	0.77 (0.66–0.90)	$2.4 \times 10^{-3*}$
<i>BACH2</i>	rs3757247-T	6:90957463	0.55	0.40	1.20 (1.04–1.38)	$3.4 \times 10^{-15*}$
<i>BACH2</i>	rs62408233-A	6:90976609	0.46	0.30	2.04 (1.77–2.36)	$1.8 \times 10^{-19*}$
<i>PD-L1</i>	rs1411262-T	9:5459419	0.22	0.25	0.85 (0.71–1.00)	0.07
<i>GATA3</i>	rs3802604-G	10:8102272	0.35	0.37	0.93 (0.80–1.08)	0.37
<i>GATA3</i>	rs569421-C	10:8108592	0.23	0.23	0.98 (0.83–1.16)	0.86
<i>CYP27B1</i>	rs4646536-G	12:58157988	0.35	0.37	0.94 (0.81–1.09)	0.44
<i>CYP27B1</i>	rs10877012-T	12:58162085	0.35	0.37	0.93 (0.80–1.07)	0.34
<i>CYP27B1</i>	rs703842-G	12:58162739	0.35	0.37	0.92 (0.79–1.06)	0.30
<i>CLEC16A</i>	rs12917716-C	16:11189148	0.35	0.41	0.78 (0.67–0.91)	$2.3 \times 10^{-3*}$

Table 1. Summary of allele frequencies and associations for variants outside the HLA region known or suggested to be associated with autoimmune Addison's disease in previous studies. ^aAll positions are given on reference genome hg19. ^bThe odds ratio (OR) represents the effect of the minor allele on the odds of observing AAD, calculated across SAR-Seq cases and all controls. ^cThe presented *P* values are adjusted with genomic control. *Asterisk marks statistical significance after Bonferroni correction for the 15 investigated variants ($P < 3.3 \times 10^{-3}$).

Locus	Variant	Position ^a	Minor/ Major Allele	Minor Allele Frequency				Association	
				AAD-Seq Cases	AAD-Seq Controls	1kSWE Controls	All Controls	OR ^b	P value ^c
<i>AIRE</i>	rs2075875	45709141	C/T	0.04	0.10	0.10	0.10	0.40 (0.29–0.55)	1.7×10^{-7}
<i>AIRE</i>	rs2075876	45709153	A/G	0.04	0.10	0.10	0.10	0.39 (0.28–0.55)	1.6×10^{-7}
<i>AIRE</i>	rs62220374	45709171	G/C	0.04	0.10	0.10	0.10	0.40 (0.29–0.55)	1.7×10^{-7}
<i>AIRE</i>	rs9983695	45711808	C/T	0.04	0.11	0.10	0.11	0.37 (0.27–0.52)	3.5×10^{-8}

Table 2. Single-nucleotide variants at the *AIRE* locus associated with autoimmune Addison's disease in single-variant association analysis. ^aAll positions are given on chromosome 21 (hg19). ^bThe odds ratio (OR) represents the effect of the minor allele on the odds of observing AAD, calculated across AAD-Seq cases and all controls. ^cThe presented *P* values are adjusted with genomic control.

binding sites (see Supplementary Fig. S6)^{45,46}. For the newly associated SNPs in *AIRE*, the minor allele frequencies (MAF) were congruent across our control groups: SAR-Seq controls (10–11%) and the 1kSWE (10%). MAFs were also consistent with the European population of the international 1000 genomes project (11%), and the gnomAD non-Finnish European population (10–11%) (see Supplementary Table S2)^{38,47}. To confirm the significantly lower MAF in our AAD cases, we resequenced the samples using an alternative method. Using single base primer extension of the top associated SNP in *AIRE* (rs9983695), genotypes were identical to the genotypes obtained from sequencing data. We could therefore confirm that the MAF in cases was indeed 4%.

AIRE mutations are well-established causes of monogenic Addison's disease when it appears as a component in APS1, and thus explain a minor fraction of AAD cases. However, using both clinical criteria, screening for cytokine autoantibodies, and *AIRE* gene sequencing, our case group had already been thoroughly screened for APS1 and all suspected patients excluded³⁰. In line with that, we could not detect any carriers of known dominant or novel protein-altering mutations in *AIRE*, but identified a single heterozygous carrier of the recessive R257* (rs121434254-T) variant. In conclusion, common genetic variation in *AIRE* appeared to be reliably associated to AAD even after strict removal of cases with suspected APS1.

Multiple AAD risk loci display additive effects. Complex diseases are typically influenced by a large number of genes⁴⁸. To study the joint effects of AAD risk variants, we investigated the distribution of risk variants among cases and controls. At six loci, we counted the number of risk alleles carried in each of the 479 cases and 1394 controls from the SAR-Seq study. We included the confirmed risk alleles rs62408233-A in *BACH2*, rs231806-G in *CTLA4*, rs2476601-A in *PTPN22*, rs12917716-G in *CLEC16A*, and the alleles 03:01, 04:04, and 04:03 in *HLA-DRB1*. In addition, we included the newly identified rs9983695-T in *AIRE*. When we stratified the cases and controls by their number of risk alleles, we observed that the fraction of cases tends to rise with an increasing number of risk alleles (Fig. 2A). It was intriguing to note the low proportion of cases emerging in

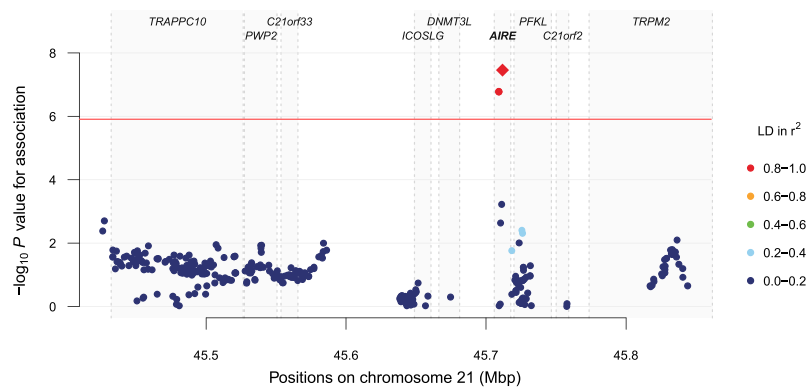


Figure 1. Association between genetic variants and autoimmune Addison's disease at the 21q22.3 locus. P values are calculated with a Cochran-Armitage test for trend, adjusted with genomic control and plotted as the negative \log_{10} . The horizontal red line denotes the level of study-wide statistical significance, 1.2×10^{-6} . Variants are coloured according to their correlation (r^2) with the top SNP (red diamond, rs9983695), which has a P value of 3.5×10^{-8} . Due to adjacent positions for three of the associated variants, their red data points overlap and only one red point is distinguishable in the figure in addition to the red diamond. The approximate location of the genes is indicated by gene names at the top of the plot and gray boxes in the background.

strata with less than three risk alleles, and at the other extreme, the high proportion of cases among subjects with more than nine risk alleles. With reference to subjects with less than three risk alleles, we could estimate with a linear model that the odds ratio for AAD more than doubles with every additional AAD risk allele (Fig. 2B). For instance, the odds for disease is 200 (95% CI: 26–2000) times higher in cases with 9 risk alleles, compared to cases with 0, 1 or 2 risk alleles (see Supplementary Table S3). For subjects with 10, 11 or 12 risk alleles, the odds of disease are even higher. In order to visualize the distribution of risk alleles in detail, we constructed a heatmap showing the number of risk alleles each study sample carried in the same 6 risk loci (Fig. 2C). *HLA-DRB1* risk alleles are known to confer the highest risk, especially when inherited in homozygous form, or heterozygous with particular *HLA-DRB1* risk alleles^{18,36}. But even within different HLA strata, effect from the other risk loci was evident, perfectly in line with joint effects from independent loci (Fig. 2C).

In genetic interaction, risk alleles at different loci act in synergy, producing a joint effect exceeding the sum of their separate effects⁴⁹. To investigate whether AAD risk alleles displayed any statistical interaction, we selected the significant risk alleles in Table 1, 03:01, 04:04 and 04:03 in *HLA-DRB1*, and applied a logistic regression model with interaction terms, as implemented in the PLINK software and with default parameterization. Beyond the additive, however, no synergistic effects on disease risk could be detected (see Supplementary Table S4). We next assessed whether a subject's age at disease onset was dependent on their combined risk allele load (see Supplementary Fig. S7). On average, subjects with more than nine risk alleles acquired their AAD more than eight years earlier than subjects with less than five risk alleles ($P = 0.018$, 95% CI: -15.8 to -1.5) (see Supplementary Table S5). However, this difference was outweighed by the variance within each stratum. No single loci could be associated to age of onset and the top SNP (rs1569495-C on chromosome 22, $P = 5.0 \times 10^{-6}$, MAF 0.2356) had a false discovery rate of 23% (see Supplementary Fig. S7).

Only a minor fraction of heritability is explained. Several organ-specific autoimmune diseases such as type 1 diabetes and autoimmune thyroid disease, are co-occurring in families and individuals, reflecting their shared risk factors¹⁵. AAD has many risk loci in common with other autoimmune diseases (Fig. 3)⁵⁰. Yet, the SNPs significantly associated to AAD and covered in this study explain a mere ~7% (range 5–9% depending on prevalence, SE range 0.7–1%) of the variance in liability in our study population. With all independent SNPs in our final dataset, including SNPs not associated to AAD, 28% (SE 4%) of liability is explained. Hence, most of disease heritability remains to be explained in future studies.

In the past, sample sizes have been large enough to make successful candidate-gene studies in AAD, yet too small to discover the same genes in a GWAS setting. With increasing sizes of AAD sample collections, we wanted to address the question whether we have already reached the sample sizes required to discover novel risk loci in a GWAS. We therefore estimated the number of cases and controls required to discover risk loci with allele frequencies and odds ratios comparable to the well-established risk loci given a traditional genome-wide significance level, 5.0×10^{-8} . Established risk alleles in *CTLA4* and *PTPN22* do not require any additional confirmation but their frequencies and impacts can exemplify what to expect of AAD risk alleles. Assuming five healthy controls for every AAD case, around 2000 cases would suffice to discover risk alleles with properties comparable to *PTPN22* and *CTLA4* in our Swedish dataset, each with 80% statistical power (see Supplementary Table S6). To discover risk loci with smaller effect sizes, in the range of *CLEC16A*, would require approximately 3000 cases. The estimated sample sizes and alternative proportions of cases and controls are presented in Supplementary Table S6.

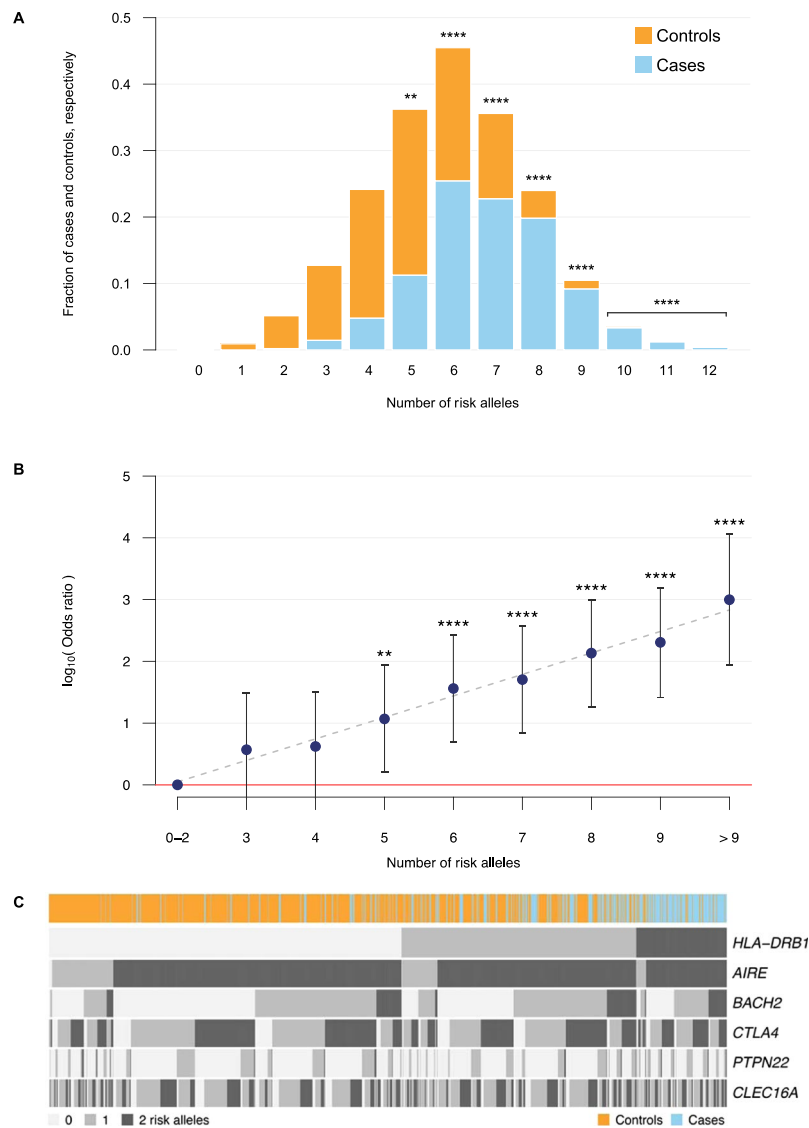


Figure 2. The number of risk alleles in cases and controls across six associated risk loci. The included risk alleles are rs9983695-T in *AIRE*, rs62408233-A in *BACH2*, rs231806-G in *CTLA4*, rs2476601-A in *PTPN22*, rs12917716-G in *CLEC16A*, and 03:01, 04:04, and 04:03 in *HLA-DRB1*. In panel (A), the top bars (orange) represent the fraction of controls and add up to 1, whereas the bottom bars (blue) indicate fraction of cases and also add up to 1. Panel (B) shows the log-transformed odds of AAD dependent on the number of risk alleles. The red line displays the level of the reference group, subjects carrying 0 to 2 risk alleles. The linear model, $\log_{10}(\text{Odds ratio}) = 0.35x - 0.65$, where x is the number of additional risk alleles, is depicted by a dashed grey line and demonstrates that the odds ratio is more than doubled with every additional AAD risk allele ($R^2 = 0.97$). Vertical bars indicate 95% confidence intervals for the point estimates. Panel (C) shows in detail the number of risk alleles across six risk loci and SAR-Seq subjects. Each column represents a subject, and the top bar displays orange colour for controls and blue for cases. The heatmap shows light, medium and dark gray for 0, 1 and 2 risk alleles, respectively. Risk loci have been ordered by descending odds ratio, and subjects sequentially sorted to facilitate interpretation. Furthest to the left, are all subjects with 0 risk alleles, and furthest to the right, all subjects with 12 risk alleles. In panels (A) and (B), ** indicates P value < 0.01 and **** $P < 0.0001$, compared to the odds in the group of samples carrying 0 to 2 risk alleles. Samples with 10 or more risk alleles are grouped together in statistical testing.

Discussion

The cause of autoimmune Addison's disease is largely unknown but the strong heritability in our Swedish population enables insight through genetic studies⁶. Although several risk loci have been studied one by one, they have not been comprehensively investigated together. In this study, we simultaneously investigated established and suggested risk loci, and assessed both their independent and joint effects.

We started this study by revisiting several risk variants that had been established or proposed in previous case-control studies. For the risk variants in *PTPN22* and *CLEC16A*, we confirmed the odds ratios reported in previous studies^{21–23,51}. Likewise, for *CTLA4*, we could replicate findings for non-coding variants, both upstream

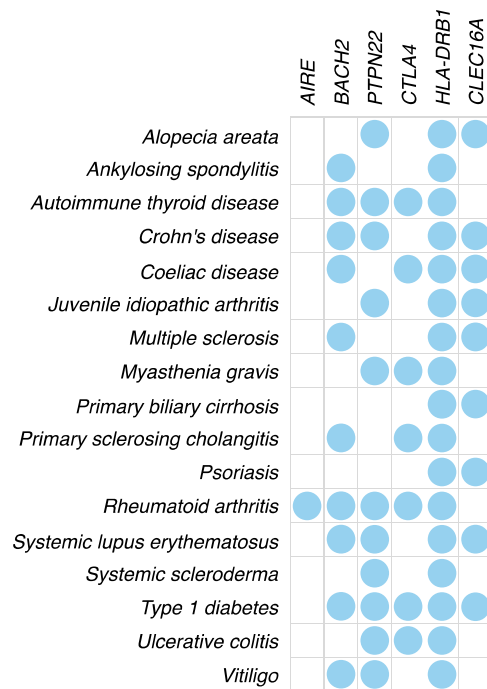


Figure 3. Loci implicated in autoimmune Addison's disease and their associations to other autoimmune diseases in genome-wide association studies. GWAS and ImmunoChip associations from the NHGRI-EBI catalog of published GWAS and *immunobase.org*. Blue dots indicate genome-wide significant associations for the diseases and the loci.

and downstream of the gene²⁰. The common missense mutation in *CTLA4* exon 1 (rs231775, c.49 A > G, p.Thr17Ala) also showed association to AAD, but with a weaker odds ratio compared to the non-coding SNPs^{19,41}. We could not, however, replicate the associations to *PD-L1*⁴³, *GATA3*⁴⁴ or *CYP27B1*⁴⁴ despite covering the suggested loci and previously associated variants.

We next explored our full range of genetic variants and identified *AIRE* as a novel risk locus in AAD with complex inheritance. Interestingly, recessive *AIRE* mutations cause APS1, of which AAD is a major component. Dominant *AIRE* mutations typically give rise to milder disease phenotypes with later onset compared to the recessive APS1, but can also present with AAD^{31,33,34}. Family members carrying dominant *AIRE* mutations can acquire various autoimmune manifestations, some of which often co-occur with sporadic AAD³⁴. These links between *AIRE* and AAD raise the question of whether the association we identified in *AIRE* could be due to the presence of undiagnosed APS1 patients in our case group. However, in addition to excluding all diagnosed cases of APS1, we took several measures to identify suspected APS1 patients prior to association testing. First, we sequenced the *AIRE* exons without detecting any known or novel protein-altering mutations indicative of APS1. We also excluded cases with cytokine autoantibodies targeting type 1 interferons α and ω , as well as IL-22, typical of recessive APS1 and also present in some cases with dominant APS1^{34,52-55}. Furthermore, by excluding cases and controls with high degree of relatedness, we selected independent samples rather than families to avoid including cases of monogenic disease. It therefore appears unlikely that the identified association with *AIRE* was explained by undiagnosed cases of APS1.

Taken together, our results suggest that common genetic variation in the *AIRE* locus may influence the susceptibility to AAD, even in the absence of protein alteration. This is in contrast with a previous Norwegian study that investigated one of our top associated SNPs (rs2075876-A/G) in a smaller Addison's disease case group (n = 311), without finding a disease association⁵⁶. Hence, in the context of previous studies, our results do not infer a final relationship between AAD development and the studied SNPs, but again point to the *AIRE* locus as a factor in autoimmune disease, beyond recessive APS1. Future studies in additional AAD case groups are required to fully establish an association to disease, and functional studies to investigate how a potential causal relationship could be mediated. Since AAD risk loci may also differ between studies of different populations⁴⁴, it would be especially valuable to replicate this novel finding in other populations.

It is not uncommon that monogenic and complex diseases with similar phenotypes share risk loci. For instance, *CTLA4* is mutated in the monogenic autoimmune lymphoproliferative syndrome type V (OMIM #616100), which has bowel inflammation as an important component^{57,58}, and is associated with sporadic forms of inflammatory bowel disease (IBD)⁵⁹. Another example is the genes in the interleukin-10 pathway, in which rare mutations cause severe IBD in neonates (OMIM #613148), and common variation is linked to sporadic IBD^{59,60}. Mutations in the *INS* (*insulin*) and *GLIS3* (*GLIS Family Zinc Finger 3*) genes cause monogenic neonatal diabetes (OMIM #606176, #610199), while common variations are linked to sporadic type 1 diabetes⁶¹. The *AIRE* gene, mutated in APS1, appears to harbor common variants associated to the risk of AAD in our Swedish case

group. Common genetic variation in *AIRE* has previously been linked to rheumatoid arthritis in GWA studies in Japanese and Chinese Han populations^{62,63}. However, it is worth noting that the risk of rheumatoid arthritis is linked to the allele rs2075876-A, whereas we associate AAD to the other allele, rs2075876-G. It may seem contradictory, but apparent shared risk loci and shared risk alleles can reflect different underlying causal variants or be influenced by other factors. The risk allele rs2075876-A has failed in validation studies of rheumatoid arthritis in European cohorts⁶⁴.

Sporadic AAD is considered a complex trait where no single gene is strictly required and no single gene is sufficient to cause the disease. Therefore, we investigated the joint effects of the confirmed risk loci known today. It was striking to note the high proportion of cases in strata with more than nine AAD risk alleles. Since we only observed a few subjects with more than nine risk alleles in our case group, these combinations of alleles could only explain a fraction of the overall disease heritability. However, it was interesting to note that subjects carrying 12 risk alleles appear to have a large proportion of their genetic propensity of AAD explained by the investigated six loci. Looking in detail at which risk alleles were present in each of our subjects, the results were in line with independent segregation of risk alleles with additive effects. With the studied risk alleles, we found a subtle effect on the age of disease onset. The number of risk alleles was not a strong predictor, but the overall effect is clear: the higher the burden of risk alleles, the earlier the disease onset.

In total, the significant risk loci *HLA*, *AIRE*, *BACH2*, *CTLA4*, *PTPN22* and *CLEC16A* explained about 7% of variance in liability of AAD in our case group. Considering that the heritability statistic is dependent on the study population and the given environment, it is ill-suited for comparison between studies. It has, however, become clear that GWAS in autoimmune diseases tend to explain only a minor fraction of heritability, for instance 9% in studies of Graves's disease and 16% in ulcerative colitis^{65–67}. By including the full range of variants detected in our array of captured exons, introns, and regulatory regions, an additional 23% of liability in our case group was explained. Consequently, 77% (23/(7 + 23)) of SNP variation has remained undetected in our study because the effect sizes have been too small to pass our statistical significance threshold. The discrepancy between the 23% accounted for by all SNPs and the total disease heritability (>90%), suggests that additional genetic variation is left to discover outside of our investigated genomic regions and/or in other classes of genetic variation, for instance, insertions, deletions and copy number variations.

Prior to SAR-Seq, most genetic studies in AAD had been limited to candidate gene approaches². These candidate genes were typically selected based on previous associations with autoimmune disorders. In fact, all currently known AAD risk loci are associated with other autoimmune diseases. This probably reflects how these loci were selected for investigation, but may also point towards a high degree of sharing of genetic risk factors in autoimmunity. However, as the number of Addison's disease samples contained within biobanks continues to get larger, we have reached the time when hypothesis-free explorations of common genetic variation are feasible with reasonable statistical power. Our calculations show that around 2000 AAD cases would suffice to replicate many of the well-established risk loci with genome-wide significance. Considering that most of the genetics underlying AAD is still uncharted, even smaller case groups would provide opportunities for novel findings.

Even though the SweGen Variant Frequency Database had made allele counts public, access to individual genotypes was still pending at the time of analysis. Therefore, the main association testing was made without population substructure covariates. Rather, the hundreds of thousands of variants tested enabled correction of the test statistics by means of genomic control. Replication in other populations and larger case groups could strengthen the association to *AIRE* and will be a natural part of large-scale genetic studies investigating the whole genome of AAD patients. Characterized by high heritability in our Swedish population and high co-inheritance with other autoimmune diseases, AAD is an attractive subject for genetic studies of sporadic autoimmunity. Furthermore, as pointed out in this study, the rarity of Addison's disease can paradoxically be an advantage in genetic studies, as more and more healthy individuals are genotyped worldwide, and the AAD diagnosis can be made with high specificity.

Methods

Study subjects. This study included the final 479 AAD cases and 1394 controls previously described in detail in the SAR-Seq study³⁶. In brief, all case subjects fulfilled diagnostic criteria for primary adrenal insufficiency¹ and were required to test positive in two independent assays for 21-hydroxylase autoantibodies, performed at two separate laboratories. To focus the study on polygenic autoimmune Addison's disease, case subjects where other causes of adrenal failure could be suspected were excluded; for example, APS1, adrenal metastases, adrenalectomy, congenital adrenal hyperplasia, hypogonadotropic hypogonadism, isolated ACTH deficiency, tuberculosis or adrenal failure secondary to corticosteroid treatment. Carriers of autoantibodies targeting the cytokines interferon ω , interferon α and/or interleukin 22 were excluded because undiagnosed APS1 could be suspected. SAR-Seq controls were included from blood donors ($n = 848$, Uppsala Bioresource) and directed sampling of healthy volunteers ($n = 653$). All subjects were carefully investigated with regards to ancestry and kinship, excluding non-European subjects and first-degree relatives. For the present study, allele counts from the SAR-Seq study and the 1kSWE were merged to augment the control population (see Supplementary Fig. S1)³⁷.

All study subjects were collected in Sweden and gave their informed consent. The SAR-Seq study was performed in accordance with the Declaration of Helsinki and was approved by the local ethics committees 2008/296-31/2 in Stockholm, 2009/013 in Uppsala, 02-253 in Umeå, and 98110 in Linköping, Sweden.

Sequencing of target genes. Targeted sequencing was described in detail in the SAR-Seq study³⁶. In brief, a selection of 1853 genes was targeted by our custom designed SeqCap EZ Choice XL capture library (06266517001; Roche NimbleGen, Basel, Switzerland). Besides genes causing monogenic adrenal disease, genes associated with autoimmune disease in genome-wide association studies were systematically included. Exons, promoters and surrounding conserved elements were targeted⁶⁸. DNA samples were sonicated to 400-bp

fragments and sequenced with 100-bp paired-end reads using Illumina HiSeq2500 (Illumina, San Diego, CA, US).

Bioinformatic processing of targeted sequencing data. The processing from sequencing reads to a high quality call set was comprehensively explained in the SAR-Seq study³⁶. In essence, for each sample, sequencing reads were mapped to the hg19 human reference genome⁶⁹. Samples with mean target coverage less than 10X were excluded. Variant calling was carried out using HaplotypeCaller (GATK 3.2.2) in gVCF mode, according to the procedures recommended for single samples all-sites calling [34, 35]. Final genotyping was performed collectively for all SAR-Seq samples with GenotypeGVCFs. Thereafter, only SNPs were considered in further analyses. Genotype calls with depth less than 8 reads or a genotype quality score less than 20, were excluded^{70,71}. Remaining SNPs were assigned probability scores using the VariantRecalibrator (GATK 3.2.2) and filtered with sensitivity tranche level 99.0.

Proportion of variance in liability explained by SNPs. In PLINK 1.9, SNPs with MAF larger than 1% were pruned such that no pairs of markers within a 150-variant window had a squared correlation stronger than 0.872. GCTA (Genome-wide Complex Trait Analysis)⁷³ v.0.93.9 was then used to estimate a genomic relationship matrix, and to perform a restricted maximum likelihood analysis, assuming an AAD prevalence ranging from 87 to 221 per million.

Imputation, dataset merging and quality control. Imputation with IMPUTE2 version 2.3.2⁷⁴ was performed using the full international 1000 genomes Phase 3 (October 2014), based on haplotypes obtained from 2504 individuals. To maximize imputation accuracy, no pre-phasing was used. Default and recommended settings were applied for effective population size (20 000) and the region size (5×10^6 bp). Imputed genotypes with certainty less than 95% were flagged as missing, and subsequently positions with more than 20% missing calls were definitively excluded before merging the SAR-Seq dataset with allele counts from 1kSWE. The merging and quality control procedures are outlined in the supplementary flowchart, where cut-off values are specified in detail (see Supplementary Fig. S1). The merging with 1kSWE controls was done in R v3.2.1⁷⁵. Biallelic SNPs overlapping the SAR-Seq dataset were extracted from the 1kSWE. Available quality control metrics were employed to stringently exclude low quality variants from 1kSWE (see Supplementary Fig. S2). Variants with differential missingness between SAR-Seq cases and controls were flagged for exclusion. Hardy-Weinberg equilibrium was tested both in each of the control sets separately, and subsequently in the merged control set. To further ensure a homogeneous merged control set, variants significantly associated to either of the control groups in a Cochran-Armitage test for trend ($P < 0.001$) were flagged for exclusion. Variants with MAF less than 0.05 in the SAR-Seq dataset, the 1kSWE, or the merged dataset were flagged for exclusion, to increase power to associate more common variants. Variant filters were effectuated in parallel and many variants failed on more than one criterion. The number of independent SNPs in downstream analyses was calculated in PLINK 1.9 by pruning the dataset such that no pairs of variants within a 150-variant window were correlated (r^2) more than 0.8^{39,72,76}.

Variant and gene annotation. The impacts of alleles were predicted using SnpEff 4.1 [40]. Curated epigenetic information for genomic positions was retrieved from HaploReg 4.1 and RegulomeDB^{45,46}. Diseases that share potential risk loci with AAD were collected from the NHGRI-EBI catalog of published GWAS and *immuno-base.org*^{77,78}. Gene locations were extracted from the genome browser at UCSC, assembly hg19⁷⁹.

Validation genotyping. The validation genotyping was performed using single base primer extension and fluorescent polarization template dye incorporation. Signal intensities were detected in a Hidex Sense (Hidex, Turku, Finland) fluorescence absorbance reader, and genotypes were called with the AlleleCaller 4.0.0.1 software.

Statistical methods. PLINK 1.9 was used for allele counts, differential missingness, conditional analysis in logistic regression, and merging with the imputed data^{72,76}. Study-wide association was tested with a two-sided Cochran Armitage test for trend in R. Inflation factor λ was estimated from the median χ^2 and used for genomic control⁸⁰. LD structures were measured and visualized by squared correlation (r^2) in Haploview⁸¹. Sample size calculation used the normal distribution without continuity correction as an approximation to the binomial distribution. Odds ratios from contingency tables, and t -tests for different number of risk alleles were calculated in GraphPad Prism v6.0.h. Linear models were calculated with the glm function in R. Screening for gene-gene interaction was performed using the epistasis command in PLINK 1.9, testing a logistic regression model for statistical interaction; $\ln(P(Y = \text{case})/P(Y = \text{control})) = \beta_0 + \beta_1 g_A + \beta_2 g_B + \beta_3 g_A g_B$. Proportions of alleles conferring risk versus protection were compared between control groups using χ^2 test in R.

References

- Husebye, E. *et al.* Consensus statement on the diagnosis, treatment and follow-up of patients with primary adrenal insufficiency. *Journal of Internal Medicine* **275**, 104–115 (2014).
- Mitchell, A. L. & Pearce, S. H. Autoimmune Addison disease: pathophysiology and genetic complexity. *Nat Rev Endocrinol* **8**, 306–316, <https://doi.org/10.1038/nrendo.2011.245> (2012).
- Hsieh, S. & White, P. C. Presentation of primary adrenal insufficiency in childhood. *J Clin Endocrinol Metab* **96**, E925–928, <https://doi.org/10.1210/jc.2011-0015> (2011).
- Winqvist, O., Karlsson, F. A. & Kämpe, O. 21-Hydroxylase, a major autoantigen in idiopathic Addison's disease. *Lancet* **339**, 1559–1562 (1992).
- Falorni, A. *et al.* Determination of 21-hydroxylase autoantibodies: inter-laboratory concordance in the Euradrenal International Serum Exchange Program. *Clinical Chemistry and Laboratory Medicine*. **53**, 1761–1770, <https://doi.org/10.1515/cclm-2014-1106> (2015).
- Skov, J. *et al.* Heritability of Addison's disease and prevalence of associated autoimmunity in a cohort of 112,100 Swedish twins. *Endocrine* **58**, 521–527, <https://doi.org/10.1007/s12020-017-1441-z> (2017).

7. Björnsdóttir, S. *et al.* Drug prescription patterns in patients with Addison's disease: a Swedish population-based cohort study. *The Journal of Clinical Endocrinology & Metabolism* **98**, 2009–2018 (2013).
8. Kong, M. F. & Jeffcoate, W. Eighty-six cases of Addison's disease. *Clinical endocrinology* **41**, 757–761 (1994).
9. Willis, A. C. & Vince, F. P. The prevalence of Addison's disease in Coventry, UK. *Postgraduate medical journal* **73**, 286–288 (1997).
10. Laureti, S., Vecchi, L., Santeusano, F. & Falorni, A. Is the prevalence of Addison's disease underestimated? *J Clin Endocrinol Metab* **84**, 1762, <https://doi.org/10.1210/jcem.84.5.5677-7> (1999).
11. Lovas, K. & Husebye, E. S. High prevalence and increasing incidence of Addison's disease in western Norway. *Clinical endocrinology* **56**, 787–791 (2002).
12. Erichsen, M. M. *et al.* Clinical, immunological, and genetic features of autoimmune primary adrenal insufficiency: observations from a Norwegian registry. *J Clin Endocrinol Metab* **94**, 4882–4890, <https://doi.org/10.1210/jc.2009-1368> (2009).
13. Quinkler, M., Dahlqvist, P., Husebye, E. S. & Kampe, O. A European Emergency Card for adrenal insufficiency can save lives. *European journal of internal medicine* **26**, 75–76, <https://doi.org/10.1016/j.ejim.2014.11.006> (2015).
14. Olafsson, A. S. & Sigurjonsdóttir, H. A. Increasing prevalence of Addison disease: results from a nationwide study. *Endocrine Practice* **22**, 30–35, <https://doi.org/10.4158/EP15754.OR> (2015).
15. Dalin, F. *et al.* Clinical and Immunological Characteristics of Autoimmune Addison Disease: A Nationwide Swedish Multicenter Study. *J Clin Endocrinol Metab* **102**, 379–389, <https://doi.org/10.1210/jc.2016-2522> (2017).
16. Myhre, A. G. *et al.* Autoimmune adrenocortical failure in Norway autoantibodies and human leukocyte antigen class II associations related to clinical features. *J Clin Endocrinol Metab* **87**, 618–623, <https://doi.org/10.1210/jcem.87.2.8192> (2002).
17. Gombos, Z. *et al.* Analysis of extended human leukocyte antigen haplotype association with Addison's disease in three populations. *European Journal of Endocrinology* **157**, 757–761, <https://doi.org/10.1530/eje-07-0290> (2007).
18. Skinningsrud, B. *et al.* Multiple loci in the HLA complex are associated with Addison's disease. *J Clin Endocrinol Metab* **96**, E1703–1708, <https://doi.org/10.1210/jc.2011-0645> (2011).
19. Vaidya, B. *et al.* Association analysis of the cytotoxic T lymphocyte antigen-4 (CTLA-4) and autoimmune regulator-1 (AIRE-1) genes in sporadic autoimmune Addison's disease. *J Clin Endocrinol Metab* **85**, 688–691, <https://doi.org/10.1210/jcem.85.2.6369> (2000).
20. Blomhoff, A. *et al.* Polymorphisms in the cytotoxic T lymphocyte antigen-4 gene region confer susceptibility to Addison's disease. *J Clin Endocrinol Metab* **89**, 3474–3476, <https://doi.org/10.1210/jc.2003-031854> (2004).
21. Roycroft, M. *et al.* The tryptophan 620 allele of the lymphoid tyrosine phosphatase (PTPN22) gene predisposes to autoimmune Addison's disease. *Clinical endocrinology* **70**, 358–362, <https://doi.org/10.1111/j.1365-2265.2008.03380.x> (2009).
22. Skinningsrud, B. *et al.* Polymorphisms in CLEC16A and CIITA at 16p13 are associated with primary adrenal insufficiency. *J Clin Endocrinol Metab* **93**, 3310–3317, <https://doi.org/10.1210/jc.2008-0821> (2008).
23. Skinningsrud, B. *et al.* A CLEC16A variant confers risk for juvenile idiopathic arthritis and anti-cyclic citrullinated peptide antibody negative rheumatoid arthritis. *Annals of the rheumatic diseases* **69**, 1471–1474, <https://doi.org/10.1136/ard.2009.114934> (2010).
24. Esselborn, V. M., Landing, B. H., Whitaker, J. & Williams, R. R. The syndrome of familial juvenile hypoadrenocorticism, hypoparathyroidism and superficial moniliasis. *J Clin Endocrinol Metab* **16**, 1374–1387, <https://doi.org/10.1210/jcem-16-10-1374> (1956).
25. Husebye, E. S., Perheentupa, J., Rautemaa, R. & Kampe, O. Clinical manifestations and management of patients with autoimmune polyendocrine syndrome type I. *J Intern Med* **265**, 514–529, <https://doi.org/10.1111/j.1365-2796.2009.02090.x> (2009).
26. Karner, J. *et al.* Anti-cytokine autoantibodies suggest pathogenetic links with autoimmune regulator deficiency in humans and mice. *Clinical and experimental immunology* **171**, 263–272, <https://doi.org/10.1111/cei.12024> (2013).
27. Laakso, S. M. *et al.* *In vivo* analysis of helper T cell responses in patients with autoimmune polyendocrinopathy - candidiasis - ectodermal dystrophy provides evidence in support of an IL-22 defect. *Autoimmunity* **47**, 556–562, <https://doi.org/10.3109/08916934.2014.929666> (2014).
28. Meager, A. *et al.* Anti-interferon autoantibodies in autoimmune polyendocrinopathy syndrome type 1. *PLoS Med* **3**, e289, <https://doi.org/10.1371/journal.pmed.0030289> (2006).
29. Landegren, N. *et al.* Proteome-wide survey of the autoimmune target repertoire in autoimmune polyendocrine syndrome type 1. *Scientific reports* **6**, 20104, <https://doi.org/10.1038/srep20104> (2016).
30. Eriksson, D. *et al.* Cytokine autoantibody screening in the Swedish Addison Register identifies patients with undiagnosed APS1. *Journal of Clinical Endocrinology and Metabolism* (2017).
31. Cetani, F. *et al.* A novel mutation of the autoimmune regulator gene in an Italian kindred with autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy, acting in a dominant fashion and strongly cosegregating with hypothyroid autoimmune thyroiditis. *J Clin Endocrinol Metab* **86**, 4747–4752, <https://doi.org/10.1210/jcem.86.10.7884> (2001).
32. Ilmarinen, T. *et al.* Functional analysis of SAND mutations in AIRE supports dominant inheritance of the G228W mutation. *Hum Mutat* **26**, 322–331, <https://doi.org/10.1002/humu.20224> (2005).
33. Su, M. A. *et al.* Mechanisms of an autoimmunity syndrome in mice caused by a dominant mutation in Aire. *The Journal of clinical investigation* **118**, 1712–1726, <https://doi.org/10.1172/jci34523> (2008).
34. Oftedal, B. E. *et al.* Dominant Mutations in the Autoimmune Regulator AIRE Are Associated with Common Organ-Specific Autoimmune Diseases. *Immunity* **42**, 1185–1196, <https://doi.org/10.1016/j.immuni.2015.04.021> (2015).
35. Abbott, J. K. *et al.* Dominant-negative loss of function arises from a second, more frequent variant within the SAND domain of autoimmune regulator (AIRE). *J Autoimmun*, <https://doi.org/10.1016/j.jaut.2017.10.010> (2017).
36. Eriksson, D. *et al.* Extended exome sequencing identifies BACH2 as a novel major risk locus for Addison's disease. *J Intern Med* **280**, 595–608, <https://doi.org/10.1111/joim.12569> (2016).
37. Ameur, A. *et al.* SweGen: a whole-genome data resource of genetic variability in a cross-section of the Swedish population. *European Journal of Human Genetics* (Advance online publication 23 August 2017).
38. The Genomes Project, C.. A global reference for human genetic variation. *Nature* **526**, 68–74, <https://doi.org/10.1038/nature15393> (2015).
39. Fadista, J., Manning, A. K., Florez, J. C. & Groop, L. The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. *European journal of human genetics: EJHG* **24**, 1202–1205, <https://doi.org/10.1038/ejhg.2015.269> (2016).
40. Kemp, E. H. *et al.* A cytotoxic T lymphocyte antigen-4 (CTLA-4) gene polymorphism is associated with autoimmune Addison's disease in English patients. *Clinical endocrinology* **49**, 609–613 (1998).
41. Wolff, A. S. *et al.* CTLA-4 as a genetic determinant in autoimmune Addison's disease. *Genes and immunity* **16**, 430–436, <https://doi.org/10.1038/gene.2015.27> (2015).
42. Pazderska, A. *et al.* A Variant in the BACH2 Gene Is Associated With Susceptibility to Autoimmune Addison's Disease in Humans. *J Clin Endocrinol Metab* **101**, 3865–3869, <https://doi.org/10.1210/jc.2016-2368> (2016).
43. Mitchell, A. L. *et al.* Programmed death ligand 1 (PD-L1) gene variants contribute to autoimmune Addison's disease and Graves' disease susceptibility. *J Clin Endocrinol Metab* **94**, 5139–5145, <https://doi.org/10.1210/jc.2009-1404> (2009).
44. Mitchell, A. L. *et al.* Association of autoimmune Addison's disease with alleles of STAT4 and GATA3 in European cohorts. *PLoS One* **9**, e88991, <https://doi.org/10.1371/journal.pone.0088991> (2014).
45. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930–934, <https://doi.org/10.1093/nar/gkr917> (2012).

46. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome research* **22**, 1790–1797, <https://doi.org/10.1101/gr.137323.112> (2012).
47. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291, <https://doi.org/10.1038/nature19057> (2016).
48. Altshuler, D., Daly, M. J. & Lander, E. S. Genetic mapping in human disease. *Science* **322**, 881–888, <https://doi.org/10.1126/science.1156409> (2008).
49. Forsberg, S. K. G., Bloom, J. S., Sadhu, M. J., Kruglyak, L. & Carlborg, O. Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. *Nat Genet* **49**, 497–503, <https://doi.org/10.1038/ng.3800> (2017).
50. Cotsapas, C. *et al.* Pervasive Sharing of Genetic Effects in Autoimmune Disease. *PLoS Genet* **7**, e1002254, <https://doi.org/10.1371/journal.pgen.1002254> (2011).
51. Skinningsrud, B. *et al.* Mutation screening of PTPN22: association of the 1858T-allele with Addison's disease. *European journal of human genetics: EJHG* **16**, 977–982, <https://doi.org/10.1038/ejhg.2008.33> (2008).
52. Zhang, L. *et al.* A robust immunoassay for anti-interferon autoantibodies that is highly specific for patients with autoimmune polyglandular syndrome type 1. *Clinical Immunology* **125**, 131–137, <https://doi.org/10.1016/j.clim.2007.07.015> (2007).
53. Oftedal, B. E. *et al.* Radioimmunoassay for autoantibodies against interferon omega; its use in the diagnosis of autoimmune polyendocrine syndrome type I. *Clinical Immunology* **129**, 163–169, <https://doi.org/10.1016/j.clim.2008.07.002> (2008).
54. Meloni, A. *et al.* Autoantibodies against type I interferons as an additional diagnostic criterion for autoimmune polyendocrine syndrome type I. *J Clin Endocrinol Metab* **93**, 4389–4397, <https://doi.org/10.1210/jc.2008-0935> (2008).
55. Husebye, E. S., Anderson, M. S. & Kämppe, O. Autoimmune Polyendocrine Syndromes. *New England Journal of Medicine* **378**, 1132–1141, <https://doi.org/10.1056/NEJMra1713301> (2018).
56. Boe Wolff, A. S. *et al.* AIRE variations in Addison's disease and autoimmune polyendocrine syndromes (APS): partial gene deletions contribute to APS I. *Genes and immunity* **9**, 130–136 (2008).
57. Schubert, D. *et al.* Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. *Nature medicine* **20**, 1410–1416, <https://doi.org/10.1038/nm.3746> (2014).
58. Kuehn, H. S. *et al.* Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. *Science* **345**, 1623–1627, <https://doi.org/10.1126/science.1255904> (2014).
59. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature genetics* **47**, 979–986, <https://doi.org/10.1038/ng.3359> (2015).
60. Lee, C. H. *et al.* Novel de novo mutations of the interleukin-10 receptor gene lead to infantile onset inflammatory bowel disease. *Journal of Crohn's & colitis* **8**, 1551–1556, <https://doi.org/10.1016/j.crohns.2014.04.004> (2014).
61. Barrett, J. C. *et al.* Genome-wide association study and meta-analysis finds over 40 loci affect risk of type 1 diabetes. *Nature genetics* **41**, 703–707, <https://doi.org/10.1038/ng.381> (2009).
62. Feng, Z. J., Zhang, S. L., Wen, H. F. & Liang, Y. Association of rs2075876 polymorphism of AIRE gene with rheumatoid arthritis risk. *Human immunology* **76**, 281–285, <https://doi.org/10.1016/j.humimm.2015.01.026> (2015).
63. Terao, C. *et al.* The human AIRE gene at chromosome 21q22 is a genetic determinant for the predisposition to rheumatoid arthritis in Japanese population. *Hum Mol Genet* **20**, 2680–2685, <https://doi.org/10.1093/hmg/ddr161> (2011).
64. Garcia-Lozano, J. R. *et al.* Association of the AIRE gene with susceptibility to rheumatoid arthritis in a European population: a case control study. *Arthritis research & therapy* **15**, R11, <https://doi.org/10.1186/ar4141> (2013).
65. Chu, X. *et al.* A genome-wide association study identifies two new risk loci for Graves' disease. *Nat Genet* **43**, 897–901, <http://www.nature.com/ng/journal/v43/n9/abs/ng.898.html#supplementary-information> (2011).
66. Anderson, C. A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nature genetics* **43**, 246–252, <https://doi.org/10.1038/ng.764> (2011).
67. Visscher, P. M. *et al.* Five Years of GWAS Discovery. *American Journal of Human Genetics* **90**, 7–24, <https://doi.org/10.1016/j.ajhg.2011.11.029> (2012).
68. Lindblad-Toh, K. *et al.* A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476–482, <https://doi.org/10.1038/nature10530> (2011).
69. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* **25**, 1754–1760, <https://doi.org/10.1093/bioinformatics/btp324> (2009).
70. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics (Oxford, England)* **27**, 2156–2158, <https://doi.org/10.1093/bioinformatics/btr330> (2011).
71. Carson, A. R. *et al.* Effective filtering strategies to improve data quality from population-based whole exome sequencing studies. *BMC bioinformatics* **15**, 125, <https://doi.org/10.1186/1471-2105-15-125> (2014).
72. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575, <https://doi.org/10.1086/519795> (2007).
73. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76–82, <https://doi.org/10.1016/j.ajhg.2010.11.011> (2011).
74. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529, <https://doi.org/10.1371/journal.pgen.1000529> (2009).
75. Team, R. C. R. *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria., <http://www.R-project.org/> (2015).
76. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7, <https://doi.org/10.1186/s13742-015-0047-8> (2015).
77. MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research* **45**, D896–D901, <https://doi.org/10.1093/nar/gkw1133> (2017).
78. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* **42**, D1001–1006, <https://doi.org/10.1093/nar/gkt1229> (2014).
79. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Research*. **12**, 996–1006, <https://doi.org/10.1101/gr.229102>. Article published online before print in May 2002 (2002).
80. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
81. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics (Oxford, England)* **21**, 263–265, <https://doi.org/10.1093/bioinformatics/bth457> (2005).

Acknowledgements

We would like to thank all case and control subjects participating in this study. Sequencing and genotyping was performed by the SNP&SEQ Technology Platform in Uppsala, supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory. Computational resources were provided by the Swedish National Infrastructure for Computing (SNIC) at Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX), cluster and storage projects b2014026 and snic2016-7-137. We would especially like to thank the

Medical Biobank of Northern Sweden. Sources of Research Support: The Swedish Research Council, Torsten and Ragnar Söderberg Foundations, the European Union Seventh Framework Programme grant 201167 EurAdrenal fp7 consortium, the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Society for Medical Research, the Swedish Society of Medicine, the NovoNordisk Foundation, Tore Nilson's Foundation for Medical Research, Karolinska Institutet, the Swedish Research Council Formas, the Knut and Alice Wallenberg Foundation, the Marianne and Marcus Wallenberg Foundation, the Swedish Rheumatism Foundation, the King Gustaf V's 80-year Foundation and the Åke Wiberg Foundation.

Author Contributions

D.E. and O.K. planned the study. M.B., N.L., F.D., Å.H., and G.R.P. performed the experimental work. D.E. processed and analysed the data. M.B. calculated the proportion of variance in liability explained by SNPs. D.E., M.B., J.S., L.H.-R., A.G., J.N., G.A., L.R., J.M., K.L.-T., and G.R.P. contributed to study design, optimization and interpretation of results. K.T., S.R.D., P.S., and L.R. collected the healthy controls. J.S., A.-L.H., J.W., P.D., O.E., S.B., O.K. characterized the patients. D.E. wrote the manuscript with contributions from all coauthors.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-26842-2>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018