Table 1 | Summary statistics of significantly associated regions identified in the genome-wide association analysis of AD case-control status, AD-by-proxy phenotype, and meta-analysis

Region			Case-control status (phase 1)		AD-by-proxy (phase 2)		Overall (phase 3)							
Locus	Chr	Gene	SNP	*P*	SNP	*P*	SNP	bp	A1	A2	MAF	*Z*	*P*	Direction
1	1	***ADAMTS4***	rs4575098	1.57×10^{-4}	rs4575098	6.88×10^{-8}	rs4575098	161155392	A	G	0.240	6.36	**2.05×10^{-10}**	?+++
2	1	*CR1*	rs6656401	**1.39×10^{-17}**	rs679515	**8.85×10^{-10}**	rs2093760	207786828	A	G	0.205	8.82	**1.10×10^{-18}**	++++
3	2	*BIN1*	rs4663105	**3.58×10^{-29}**	rs4663105	**5.46×10^{-26}**	rs4663105	127891427	C	A	0.415	13.94	**3.38×10^{-44}**	?+++
4	2	*INPPD5*	rs10933431	1.67×10^{-6}	rs10933431	2.51×10^{-6}	rs10933431	233981912	G	C	0.235	−6.13	**8.92×10^{-10}**	?---
5	3	***HESX1***	NA		rs184384746	**1.24×10^{-8}**	rs184384746	57226150	T	C	0.002	5.69	**1.24×10^{-8}**	???+
6	4	***CLNK***	rs6448453	0.024	rs6448451	**1.19×10^{-8}**	rs6448453	11026028	A	G	0.252	6.00	**1.93×10^{-9}**	?+-+
--	4	*HS3ST1*	rs7657553	**2.16×10^{-8}**	rs7657553	0.790	rs7657553	11723235	A	G	0.291	1.95	0.051	?++-
7	6	*HLA-DRB1*	rs9269853	**2.66×10^{-8}**	rs6931277	1.78×10^{-7}	rs6931277	32583357	T	A	0.153	−6.49	**8.41×10^{-11}**	?---
8	6	*TREM2*	NA		rs187370608	**1.45×10^{-16}**	rs187370608	40942196	A	G	0.002	8.26	**1.45×10^{-16}**	???+
9	6	*CD2AP*	rs9381563	**5.35×10^{-9}**	rs9381563	8.10×10^{-6}	rs9381563	47432637	C	T	0.355	6.33	**2.52×10^{-10}**	?+++
10	7	*ZCWPW1*	rs1859788	**6.05×10^{-9}**	rs7384878	**2.38×10^{-10}**	rs1859788	99971834	A	G	0.310	−7.93	**2.22×10^{-15}**	----
11	7	*EPHA1*	rs11763230	**2.58×10^{-11}**	rs7810606	1.01×10^{-6}	rs7810606	143108158	T	C	0.500	−6.62	**3.59×10^{-11}**	?---
12	7	***CNTNAP2***	NA		rs114360492	**2.10×10^{-9}**	rs114360492	145950029	T	C	2.59×10^{-4}	5.99	**2.10×10^{-9}**	???+
13	8	*CLU/PTK2B*	rs4236673	**6.36×10^{-20}**	rs1532278	**7.45×10^{-9}**	rs4236673	27464929	A	G	0.391	−8.98	**2.61×10^{-19}**	----
14	10	*ECHDC3*	rs11257242	**2.38×10^{-8}**	rs11257238	5.84×10^{-5}	rs11257238	11717397	C	T	0.375	5.69	**1.26×10^{-8}**	?+++
15	11	*MS4A6A*	rs7935829	**8.21×10^{-13}**	rs1582763	**4.72×10^{-9}**	rs2081545	59958380	A	C	0.381	−7.97	**1.55×10^{-15}**	----
16	11	*PICALM*	rs10792832	**1.12×10^{-17}**	rs3844143	**5.31×10^{-11}**	rs867611	85776544	G	A	0.314	−8.75	**2.19×10^{-18}**	?---
17	11	*SORL1*	rs11218343	**5.57×10^{-11}**	rs11218343	2.81×10^{-6}	rs11218343	121435587	C	T	0.040	−6.79	**1.09×10^{-11}**	?---
18	14	*SLC24A4*	rs12590654	**1.98×10^{-8}**	rs12590654	3.70×10^{-6}	rs12590654	92938855	A	G	0.344	−6.39	**1.65×10^{-10}**	?---
19	15	***ADAM10***	rs442495	3.09×10^{-4}	rs442495	2.65×10^{-7}	rs442495	59022615	C	T	0.320	−6.07	**1.31×10^{-9}**	?---
20	15	***APH1B***	rs117618017	0.022	rs117618017	2.64×10^{-7}	rs117618017	63569902	T	C	0.132	5.52	**3.35×10^{-8}**	++++
21	16	***KAT8***	rs59735493	8.25×10^{-4}	rs59735493	3.72×10^{-6}	rs59735493	31133100	A	G	0.300	−5.49	**3.98×10^{-8}**	?---
22	17	*SCIMP*	rs113260531	3.21×10^{-6}	rs9916042	**4.73×10^{-8}**	rs113260531	5138980	A	G	0.120	6.12	**9.16×10^{-10}**	?+++
23	17	*ABI3*	rs28394864	7.29×10^{-5}	rs28394864	6.80×10^{-6}	rs28394864	47450775	A	G	0.473	5.62	**1.87×10^{-8}**	?+++
--	17	*BZRAP1-AS1*	rs2632516	**1.42×10^{-9}**	rs2632516	0.005	rs2632516	56409089	C	G	0.455	−4.90	9.66×10^{-7}	?---
--	18	*SUZ12P1*	rs8093731	**4.63×10^{-8}**	rs8093731	0.766	rs8093731	29088958	T	C	0.010	−2.17	0.030	?-?-
24	18	***ALPK2***	rs76726049	0.039	rs76726049	1.83×10^{-7}	rs76726049	56189459	C	T	0.014	5.52	**3.30×10^{-8}**	?+++
25	19	*ABCA7*	rs4147929	**8.64×10^{-9}**	rs3752241	**2.87×10^{-8}**	rs111278892	1039323	G	C	0.161	6.50	**7.93×10^{-11}**	?+++
26	19	*APOE*	rs41289512	**2.70×10^{-194}**	rs75627662	**9.51×10^{-296}**	rs41289512	45351516	G	C	0.039	35.50	**5.79×10^{-276}**	?+++
27	19	***AC074212.3***	rs76320948	1.54×10^{-5}	rs76320948	1.80×10^{-5}	rs76320948	46241841	T	C	0.046	5.46	**4.64×10^{-8}**	?+?+
28	19	*CD33*	rs3865444	**4.25×10^{-8}**	rs3865444	4.97×10^{-5}	rs3865444	51727962	A	C	0.320	−5.81	**6.34×10^{-9}**	?---
29	20	*CASS4*	rs6014724	8.72×10^{-8}	rs6014724	6.32×10^{-6}	rs6014724	54998544	G	A	0.089	−6.18	**6.56×10^{-10}**	?---

Note: independent lead SNPs are defined by r^2 <0.1; distinct genomic loci are >250 kb apart. The locus column indicates the loci number based on phase 3 (-- indicates that this locus is non-significant). The gene symbols are included to conveniently compare the significant loci with previously discovered loci. The bolded genes correspond to the novel loci indicating the genes in closest proximity to the most significant SNP, while emphasizing that this is not necessarily the causal gene. Allele1 is the effect allele for the meta-association statistic. The directions of effect of the distinct cohorts are in the following order: ADSP, IGAP, PGC-ALZ, and UKB; note that the first cohort is often missing as this concerns exome sequencing data. Corrected *P* value for significance equals 5×10^{-8} (marked as bold and underlined values). Note that the lead SNP can differ between the distinct analyses, while it tags the same locus.

of these loci excluding ADSP resulted in similar association signals (Supplementary Table 5), implying that we have correctly adjusted for partial sample overlap between IGAP and ADSP. The lead SNPs in 3 loci (with nearest genes *HESX1*, *TREM2*, and *CNTNAP2*) were only available in the UKB cohort (Table 1), but were of good quality (imputation quality INFO score >0.91, Hardy-Weinberg equilibrium P >0.19, missingness <0.003). These SNPs were all rare (minor allele frequency (MAF) <0.003), meaning that they will require future confirmation in another similarly large sample. However, variants in *TREM2* have been robustly linked to AD in previous research[9].

Verifying the 13 novel loci against other recent genetic studies on AD[9,12,16–18], 4 loci (*TREM2*, *ECHDC3*, *SCIMP*, and *ABI3*) have been previously discovered in addition to the 16 identified by Lambert et al., leaving 9 novel loci at the time of this writing (*ADAMTS4*, *HESX1*, *CLNK*, *CNTNAP2*, *ADAM10*, *APH1B*, *KAT8*, *ALPK2*, and *AC074212.3*). The *ADAMTS4* and *KAT8* loci have also since been identified in a recent analysis in a partially overlapping sample[13]. Comparing our meta-analysis results with all loci of Lambert et al.[4] to determine differences in associated loci, we were unable to observe 4 loci (*MEF2C*, *NME8*, *CELF1*, and *FERMT2*) at a GWS level (observed P values were 1.6×10^{-5} to 0.0011), which was mostly caused by a lower association signal in the UKB data set (Supplementary Table 6). By contrast, Lambert et al.[4] were unable to replicate the *DSG2* and *CD33* loci in the second stage of their study. In our study, *DSG2* was also not supported (meta-analysis $P=0.030$; UKB analysis $P=0.766$), implying invalidation of this locus, while the *CD33* locus (rs3865444 in Table 1) was significantly associated with AD (meta-analysis $P=6.34\times10^{-9}$; UKB analysis $P=4.97\times10^{-5}$), implying a genuine genetic association with AD risk.

Next, we aimed to find further support for the novel findings by using an independent Icelandic cohort (deCODE[19,20]), including 6,593 AD cases and 174,289 controls (Fig. 1 and Supplementary Table 7), to test replication of the lead SNP or an LD-proxy of the lead SNP (r^2 >0.9) in each locus. We were unable to test two loci