LETTERS

Genome-wide association study in alopecia areata implicates both innate and adaptive immunity

Lynn Petukhova[1], Madeleine Duvic[2], Maria Hordinsky[3], David Norris[4], Vera Price[5], Yutaka Shimomura[1], Hyunmi Kim[1], Pallavi Singh[1], Annette Lee[6], Wei V. Chen[7], Katja C. Meyer[8], Ralf Paus[8,9], Colin A. B. Jahoda[10], Christopher I. Amos[7], Peter K. Gregersen[6] & Angela M. Christiano[1,11]

Alopecia areata (AA) is among the most highly prevalent human autoimmune diseases, leading to disfiguring hair loss due to the collapse of immune privilege of the hair follicle and subsequent autoimmune attack[1,2]. The genetic basis of AA is largely unknown. We undertook a genome-wide association study (GWAS) in a sample of 1,054 cases and 3,278 controls and identified 139 single nucleotide polymorphisms that are significantly associated with AA ($P \leq 5 \times 10^{-7}$). Here we show an association with genomic regions containing several genes controlling the activation and proliferation of regulatory T cells (T_{reg} cells), cytotoxic T lymphocyte-associated antigen 4 (*CTLA4*), interleukin (*IL*)*-2/IL-21*, IL-2 receptor A (*IL-2RA; CD25*) and *Eos* (also known as Ikaros family zinc finger 4; *IKZF4*), as well as the human leukocyte antigen (HLA) region. We also find association evidence for regions containing genes expressed in the hair follicle itself (*PRDX5* and *STX17*). A region of strong association resides within the *ULBP* (cytomegalovirus UL16-binding protein) gene cluster on chromosome 6q25.1, encoding activating ligands of the natural killer cell receptor NKG2D that have not previously been implicated in an autoimmune disease. By probing the role of ULBP3 in disease pathogenesis, we also show that its expression in lesional scalp from patients with AA is markedly upregulated in the hair follicle dermal sheath during active disease. This study provides evidence for the involvement of both innate and acquired immunity in the pathogenesis of AA. We have defined the genetic underpinnings of AA, placing it within the context of shared pathways among autoimmune diseases, and implicating a novel disease mechanism, the upregulation of ULBP ligands, in triggering autoimmunity.

AA affects about 5.3 million people in the United States alone, including males and females across all ethnic groups, with a lifetime risk of 1.7% (refs 1, 2). Autoimmunity develops against the hair follicle, resulting in non-scarring hair loss that may begin as patches that can coalesce and progress to cover the entire scalp (alopecia totalis) or eventually the entire body (alopecia universalis) (Supplementary Fig. 1). The phenomenon of 'sudden whitening of the hair' is ascribed to the acute onset of AA at times of profound grief, stress or fear[3], in which the pigmented hair is selectively shed while the white hair persists. AA spares the stem cell compartment and attacks only the base of the hair follicle, which is surrounded by infiltrating lymphocytes. Despite these marked perturbations in the hair follicle, there is no permanent organ destruction, and regrowth of the hair remains possible. The concept of an autoimmune mechanism as the basis for AA emerged during the twentieth century from multiple lines of evidence[4]. AA hair follicles are surrounded by an immune infiltrate with activated T-helper cells (T_H cells), cytotoxic T cells (T_C cells) and natural killer (NK) cells, characterized as a T_H1-type inflammatory response[5,6]. The notion of a collapse of immune privilege is thought to be a key event in triggering AA[4,7].

Evidence supporting a genetic basis for AA stems from multiple lines of research, including the observed heritability in first-degree relatives[8,9], twin studies[10] and, most recently, from our family-based linkage studies[11]. Although a number of candidate-gene association studies were performed over the past two decades, the informativeness of these studies was inherently limited by small sample sizes and preselection of candidate genes.

To determine the genetic architecture of AA, we genotyped or used publicly available data for up to 1,054 AA cases and 3,278 controls with a combination of Illumina 610K and 550K arrays. We performed association tests adjusted for residual population stratification ($\lambda = 1.051$) and found 139 single nucleotide polymorphisms (SNPs) significant at 5×10^{-7} (Fig. 1 and Supplementary Table 1).

Our analysis identified several susceptibility loci for AA, most of which clustered into eight genomic regions and fell within discrete linkage disequilibrium (LD) blocks (Fig. 2 and Table 1). These include loci on chromosome 2q33.2 containing *CTLA4*, chromosome 4q27 containing *IL-2/IL-21*, chromosome 6p21.32 containing the HLA, chromosome 6q25.1 harbouring the *ULBP* genes, chromosome 10p15.1 containing *IL-2RA* (*CD25*), and chromosome 12q13 containing *Eos* (*IKZF4*) and *ERBB3*. One SNP resides on chromosome 9q31.1 within syntaxin 17 (*STX17*), and one resides on

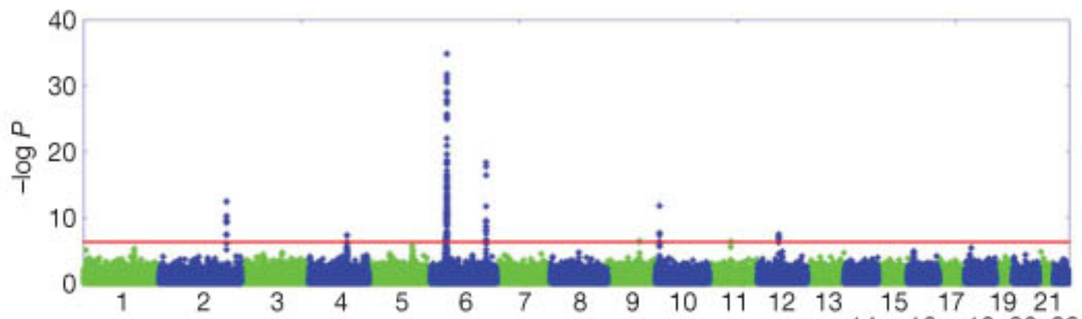

Figure 1 | Manhattan plot of the joint analysis of the discovery GWAS and the replication GWAS. Results are plotted as negative log-transformed *P* values from a genotypic association test controlled for residual population stratification as a function of the position in the genome. Odd chromosomes are in green and even chromosomes in blue. Eight genomic regions contain SNPs that exceed the genome-wide significance threshold of 5×10^{-7} (red line).

[1]Department of Dermatology, Columbia University, New York, New York 10032, USA. [2]Department of Dermatology, M. D. Anderson Cancer Center, Houston, Texas 77030, USA. [3]Department of Dermatology, University of Minnesota, Minneapolis, Minnesota 55455, USA. [4]Department of Dermatology, University of Colorado, Denver, Colorado 80010, USA. [5]Department of Dermatology, UCSF, San Francisco, California 94115, USA. [6]The Feinstein Institute for Medical Research, North Shore LIJHS, Manhasset, New York 11030, USA. [7]Department of Epidemiology, M. D. Anderson Cancer Center, Houston, Texas 77030, USA. [8]Department of Dermatology, University of Lübeck, 23538 Lübeck, Germany. [9]School of Translational Medicine, University of Manchester, Manchester M13 9PT, UK. [10]Department of Biological Sciences, University of Durham, Durham DH1 3LE, UK. [11]Department of Genetics and Development, Columbia University, New York, New York 10032, USA.

Table 1 | Genes with significant association to AA

Region	Gene	Function	Strongest association (*P* value)	Maximum odds ratio	Involved in other autoimmune disease
2q33.2	*CTLA4*	Co-stimulatory family	3.55×10^{-13}	1.44	T1D, RA, CeD, MS, SLE, GD
	ICOS	Co-stimulatory family	4.33×10^{-8}	1.32	
4q27	*IL-21/IL-2*	T-, B- and NK-cell proliferation	4.27×10^{-8}	1.34	T1D, RA, CeD, PS
6q25.1	*ULBP6*	NKG2D activating ligand	4.49×10^{-19}	1.65	None
	ULBP3	NKG2D activating ligand	4.43×10^{-17}	1.52	None
9q31.1	*STX17*	Premature hair greying	3.60×10^{-7}	1.33	None
10p15.1	*IL-2RA*	T-cell proliferation	1.74×10^{-12}	1.41	T1D, MS, GD, GV
11q13	*PRDX5*	Antioxidant enzyme	4.14×10^{-7}	1.33	MS
12q13	*Eos* (*IKZF4*)	T_{reg} transcription factor	3.21×10^{-8}	1.34	T1D, SLE
	ERBB3	Epidermal growth factor receptor	1.27×10^{-7}	1.34	T1D, SLE
6p21.32	*MICA*	NKG2D activating ligand	1.19×10^{-7}	1.44	T1D, RA, CeD, UC, PS, SLE
(HLA)	*NOTCH4*	Haematopoietic differentiation	1.03×10^{-8}	1.61	T1D, RA, MS
	C6orf10	Unknown	1.45×10^{-16}	2.36	T1D, RA, PS, GV
	BTNL2	Co-stimulatory family	2.11×10^{-26}	2.70	T1D, RA, UC, CD, SLE, MS, GV
	HLA-DRA	Antigen presentation	2.93×10^{-31}	2.62	T1D, RA, CeD, MS, GV
	HLA-DQA1	Antigen presentation	3.60×10^{-17}	2.15	T1D, RA, CeD, MS, SLE, PS, CD, UC, GD
	HLA-DQA2	Antigen presentation	1.38×10^{-35}	5.43	T1D, RA
	HLA-DQB2	Antigen presentation	1.73×10^{-13}	1.60	RA

Each of the eight regions implicated in our study contains multiple significant SNPs, which are detailed in Supplementary Tables 1 and 2. Here we display candidate genes within the implicated regions, and include the *P* value of the most significant SNP, and the odds ratio for the SNP with the largest effect estimate. Diseases are listed for which a GWAS or previous candidate gene study identified the same region (http://www.genome.gov/gwastudies, http://www.cdc.gov/genomics/hugenet): Crohn's disease (CD), celiac disease (CeD), Graves disease (GD), generalized vitiligo (GV), multiple sclerosis (MS), psoriasis (PS), rheumatoid arthritis (RA), system lupus erythematosus (SLE), type I diabetes (T1D), and ulcerative colitis (UC).

chromosome 11q13, upstream from peroxiredoxin 5 (*PRDX5*). Several of these LD blocks coincide with regions of linkage that we reported previously on chromosomes 6p (HLA), 6q (*ULBP*s), 10p (*IL-2RA*) and 18p (*PTPN2*)[11]. We also identified an additional 163 SNPs that were nominally significant ($10^{-4} < P < 5 \times 10^{-7}$), which include 12 regions containing genes involved in the immune response, notably *IL-13*, *IL-6*, *IL-26*, *IFNG*, *SOCS1* and *PTPN2* (Supplementary Tables 2 and 3). Finally, imputation analysis from HapMap release 2.2 identified additional statistically significant SNPs within each of the eight regions and one additional SNP in *PTPN2* that raised it above statistical significance (Supplementary Tables 3 and 4).

We next assessed the extent to which these genetic risk factors contribute to AA. First, we decreased redundancy in our association evidence by using conditional analysis to determine which SNPs represent independent risk factors within each region (Supplementary Fig. 2), thus identifying a set of 16 risk haplotypes (Fig. 2 and Supplementary Table 1). The distribution of risk haplotypes was significantly different between cases and controls ($P = 1.1 \times 10^{-107}$) (Fig. 2q, r). To determine the relative contribution of different alleles to the genetic burden of AA, population-attributable risks were calculated for genotypes of individual SNPs and showed large contributions from individual alleles (ranging from 16 to 69%) (Supplementary Table 5). Together with the high concordance in siblings[8,9], these findings demonstrate an overwhelming contribution of risk from genetic factors in AA, which awaits confirmation in a validation study.

Our GWAS study in AA implicates a new class of NKG2D ligands in autoimmune disease. The *ULBP* genes reside in a 180-kilobase MHC class I-related cluster on human chromosome 6q25.1 (Fig. 2g, h) that arose through duplication of the MHC locus[12]. Each of the *ULBP* genes has been shown to function as an NKG2D-activating ligand[13,14]. NKG2D ligands, including *MICA/B* and *ULBPs*, are stress-induced molecules that act as danger signals to alert NK, natural killer T, $\delta\gamma$ T and $CD8^+$ T lymphocytes through the engagement of the receptor NKG2D[13].

We next considered whether perturbations in the hair follicle microenvironment contribute to the initiation of AA. It has been postulated that NKG2D ligands, if overexpressed in genetically susceptible individuals, can trigger an autoimmune response against the tissue expressing the ligand[15]. To probe this hypothesis in the setting of AA, we examined ULBP3 expression within the hair follicle of unaffected scalp (Fig. 3a, b) and patients with AA (Fig. 3c, d). Whereas ULBP3 is expressed at low levels within the dermal papilla in normal hair follicle (Fig. 3a, b), in two different patients with early active AA lesions, we observed marked upregulation of ULBP3 expression in the dermal sheath as well as the dermal papilla (Fig. 3c, d), but not in control individuals or those with other inflammatory scalp disorders (data not shown). Quantitative immunohistochemistry corroborated a significantly increased number of $ULBP3^+$ cells in 16 additional AA samples (Fig. 3p). We also noted a massive inflammatory cell infiltrate characterized by $CD8^+CD3^+$ T cells (Fig. 3g–l), but only rare NK cells (data not shown). Double immunostaining with an anti-CD8 and an anti-NKG2D antibody revealed that most $NKG2D^+$ cells were $CD8^+$ T cells (Fig. 3m–o). These results suggest that the autoimmune destruction in AA may be mediated in part by $CD8^+NKG2D^+$ cytotoxic T cells, whose activation may be induced by upregulation of ULBP3 in the dermal sheath of the hair follicle.

The localization of an NK-activating ligand in the outermost layer of the hair follicle places it in an ideal position to express a danger signal[16] and engage NKG2D on immune cells in the local milieu. Inducible overexpression of the *ULBP* homologues *Rae1*ε and *Rae1*α in mouse epidermis and pulmonary epithelium were previously shown to markedly alter the immune microenvironment within the skin and lung, respectively[17,18]. We postulate that in genetically susceptible individuals, upregulation of *ULBP3* may have a similar effect on initiating the immune response in AA, and/or may become induced as part of an inflammatory cascade. Consistent with these findings, upregulation of the NK ligand MICA has recently been demonstrated in the hair follicle of patients with AA[19]. Taken together with the increased numbers of perifollicular $NKG2D^+CD8^+$ cells that we and others observed in lesional skin of patients with AA (Fig. 3)[19], these data implicate a mechanism involving the expression of NKG2D ligands, including ULBPs, and infiltration of NKG2D-expressing cells in the aetiology of AA.

In addition to *ULBP3/ULBP6*, we identified other genes that are expressed in the hair follicle and may provide insight into the initiating events (Supplementary Figs 3 and 4). For example, *STX17* (rs10760706, $P = 3.60 \times 10^{-7}$) is expressed in the hair follicle[20] and is associated with the grey hair phenotype in horses, which is of interest because AA preferentially attacks pigmented hairs[21]. *PRDX5* (rs694739, $P = 4.14 \times 10^{-7}$) is an antioxidant enzyme involved in the cellular response to oxidative stress, a process which is dysregulated in AA scalp[22]. PRDX5 has been implicated in the degeneration of target cells in several autoimmune disorders[23–25], and other PRDX family members can serve as autoantigens[26] (Supplementary Table 7). We found evidence for several genes whose expression in the hair follicle may contribute to a disruption in the local milieu, resulting in the onset of autoimmunity.

Our data implicate several factors that may act together to induce and promote immune dysregulation in the pathogenesis of AA. We found strong evidence for genes involved in the differentiation and

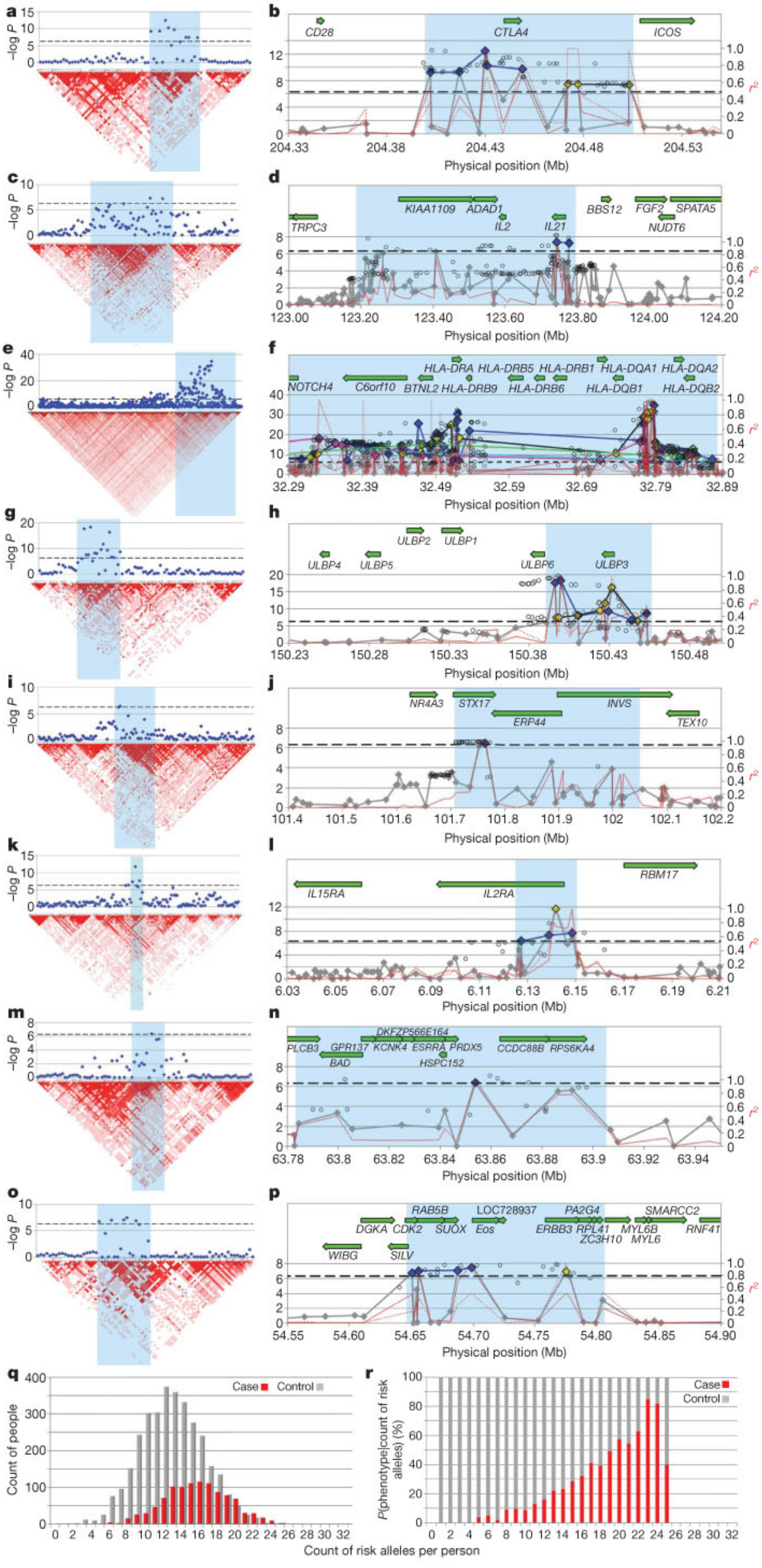

Figure 2 | LD structure and haplotype organization of the implicated regions from GWAS. **a**–**p**, In all graphs the genome-wide significance threshold (5×10^{-7}) is indicated by a black dotted line. Results from the eight regions are aligned with LD maps (**a**, **c**, **e**, **g**, **i**, **k**, **m**, **o**) and transcript maps (**b**, **d**, **f**, **h**, **j**, **l**, **n**, **p**): chromosomes 2q33 (**a**, **b**), 4q26–27 (**c**, **d**), 6p21.3 (**e**, **f**), 6q25.1 (**g**, **h**), 9q31.1 (**i**, **j**), 10p15–p16 (**k**, **l**), 11q13 (**m**, **n**) and 12q13 (**o**, **p**). For the plots with the LD maps, red indicates high LD as measured by D'. For the plots with the transcript maps, results for imputed SNPs are indicated by open circles. Typed SNPs that do not reach significance are in grey; significantly associated typed SNPs are in colour, coded by the risk haplotypes. For example in **b**, conditioning on any of the blue SNPs will decrease evidence for association of the other blue SNPs but will not affect evidence of any of the yellow-coded SNPs. On chromosome 6p in the HLA, significantly associated SNPs can be organized into at least five distinct haplotypes (blue, green, yellow, pink and light blue). The red lines show pairwise LD as measured by r^2, for the most significant SNP in each haplotype, and define the LD block that is demonstrating association. **q**, **r**, The cumulative effect of risk haplotypes is indicated by the distribution of the genetic liability index in cases and controls. We chose the most significantly associated SNP from each haplotype to serve as a proxy for the haplotype, and show in **q** that the distribution of independent genetic risk factors changes as a function of phenotype, with an average of 13 risk alleles found in controls (grey) and 16 found in cases (red). As the number of risk alleles in an individual increases, the proportion affected by AA increases. The conditional probability of phenotype given count of risk alleles is shown in **r** (AA in red, control in grey).

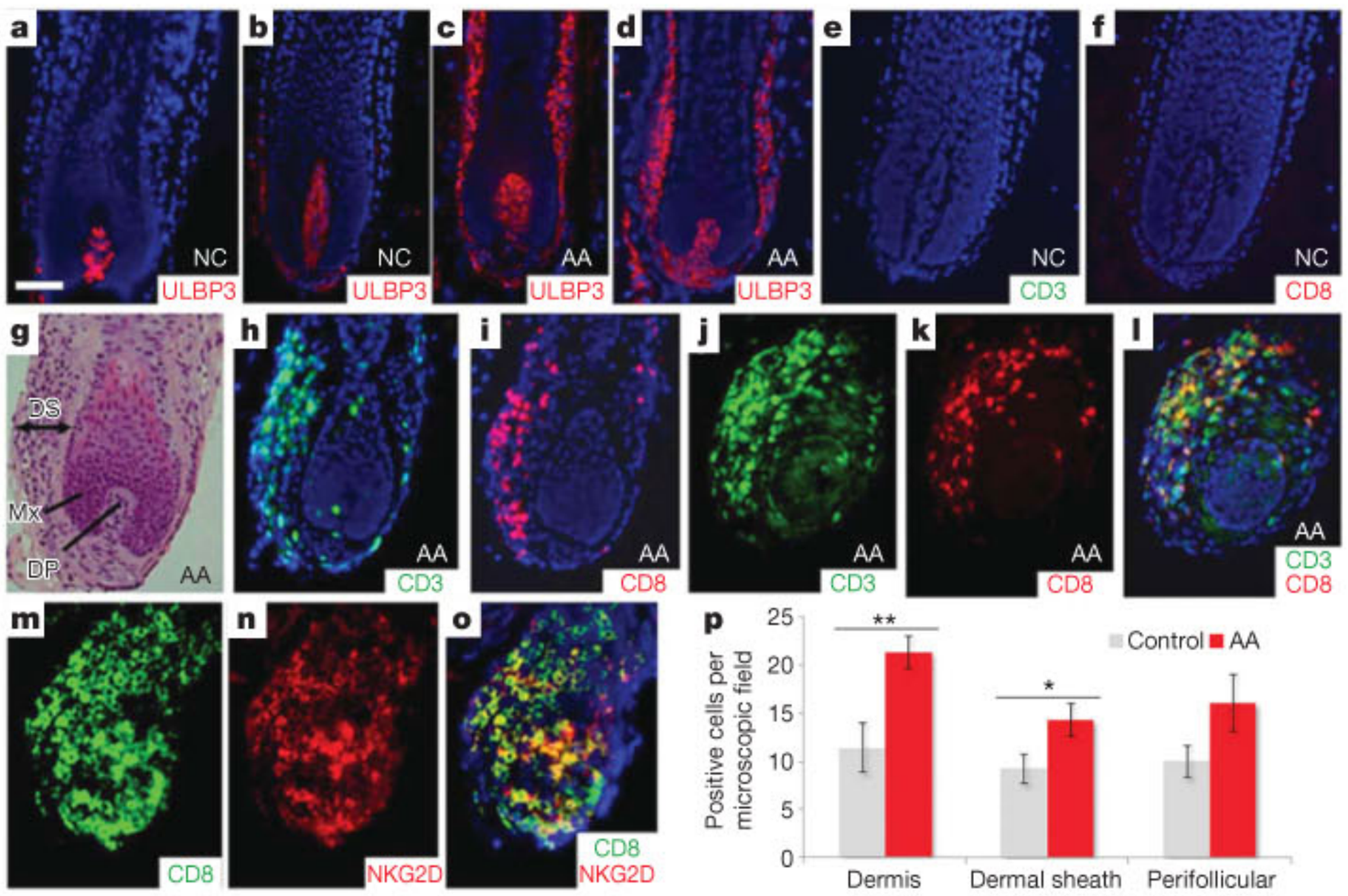

Figure 3 | ULBP3 expression and immune cell infiltration of AA hair follicles. **a**, **b**, Low levels of expression of ULBP3 in the dermal papilla of hair follicles from two unrelated, unaffected individuals. **c**, **d**, Massive upregulation of ULBP3 expression in the dermal sheath of hair follicles from two unrelated patients with AA in the early stages of disease. **e**, **f**, Absence of immune infiltration in two control hair follicles. **g**, Haematoxylin and eosin staining of AA hair follicle. DS, dermal sheath; Mx, matrix; DP, dermal papilla. **h**, **i**, Immunofluorescence analysis using CD3 (**h**) and CD8 (**i**) cell-surface markers for T-cell lineages. There is a marked inflammatory infiltrate in the dermal sheath of two affected AA hair follicles. **j**–**l**, Double-immunofluorescence analysis with anti-CD3 (**j**) and anti-CD8 (**k**) antibodies. **l**, The merged image clearly shows infiltration of $CD3^+CD8^+$ T cells in the dermal sheath of the AA hair follicle. Panels **d** and **g**–**l** are serial sections of the same hair follicle of an affected individual. Counterstaining with 4′,6-diamidino-2-phenylindole is shown in blue (**a**–**f**, **h**, **i**, **l**). Scale bar, 50 μm. AA, alopecia areata patients; NC, normal control individuals. **m**–**o**, Double immunostainings with anti-CD8 (**m**) and anti-NKG2D (**n**) antibodies revealed that most $NKG2D^+$ cells co-expressed $CD8^+$; **o**, merged image. **p**, Quantification of immunohistochemical staining (positive cells per microscope field at a magnification of ×200) for ULBP3 in 16 patients with AA and in 7 controls showed a significantly increased number of $ULBP3^+$ cells in the dermis and dermal sheath in patients with AA compared with control skin. In addition, positive cells were also upregulated in perifollicular regions in AA samples, although this was not statistically significant. Data were analysed by Mann–Whitney test for unpaired samples and are expressed as means ± s.e.m.; asterisk, $P < 0.05$; two asterisks, $P < 0.01$.

maintenance of both immunosuppressive T_{reg} cells, as well as their functional antagonists, pro-inflammatory T helper cells (T_H17). T_{reg} cells have a critical function in preventing immune responses against autoantigens, and their differentiation depends on the early expression of IL-2RA (CD25) (rs3118470, $P = 1.74 \times 10^{-12}$), as well as a key lineage-determining transcription factor, Foxp3. Foxp3-mediated gene silencing is critical in determining that T_{reg} cells effectively suppress immune responses. Both IL-2 (rs7682241, $P = 4.27 \times 10^{-8}$) and its high-affinity receptor IL-2RA (rs3118470, $P = 1.74 \times 10^{-12}$) are central in controlling the survival and proliferation of T_{reg} cells. It was recently found that Eos (IKZF4) (rs1701704, $P = 3.21 \times 10^{-8}$), a member of the Ikaros family of transcription factors, is a key co-regulator of FoxP3-directed gene silencing during T_{reg} differentiation[27]. Although T_{reg} cells probably use several different mechanisms to suppress immune responses, the high expression of CTLA4 (rs1024161, $P = 3.55 \times 10^{-13}$) has been proposed as a major determinant of their suppressive activity, particularly because CTLA4 is essential for the inhibitory activity of T_{reg} cells on antigen-presenting cells[28]. The *IL-2* locus is tightly linked with *IL-21* (rs7682241, $P = 4.27 \times 10^{-8}$), which has pleiotropic effects on multiple cell lineages, including $CD8^+$ T cells, B cells, NK cells and dendritic cells. IL-21 is a major product of proinflammatory T_H17 (IL-17-producing $CD4^+$ T_H cells) and has been shown to have a key role in both promoting the differentiation of T_H17 cells and limiting the differentiation of T_{reg} cells[29]. Collectively, the constellation of immunoregulatory genes implicated in AA clearly point to T_{reg} cells and T_H17 cells as avenues for future studies and novel therapies.

The common-cause hypothesis of autoimmune diseases has received robust validation from GWAS in recent years[30]. This hypothesis evolved initially from epidemiological studies that demonstrated the aggregation of autoimmune diseases within families, and was further supported by findings of shared susceptibility regions in linkage studies. Our GWAS upheld the previously reported associations of HLA genes in AA and other autoimmune disorders (reviewed in ref. 4), whereas we did not find strong evidence for the other loci previously tested in AA using the candidate gene approach (Supplementary Table 6). In accordance with the common-cause hypothesis, our GWAS revealed several risk loci in common with other forms of autoimmunity, such as rheumatoid arthritis, type I diabetes, coeliac disease, systemic lupus erythematosus, multiple sclerosis and psoriasis: in particular, *CTLA4*, *IL-2/IL-21*, *IL-2RA* and genes critical to T_{reg} maintenance (Table 1 and Supplementary Table 1). The commonality with rheumatoid arthritis, type I diabetes and coeliac disease is particularly noteworthy in view of the significance of the NKG2D receptor in the pathogenesis of each of these three diseases[15].

Our GWAS establishes the genetic basis of AA, revealing several loci that contribute to disease susceptibility. These findings open new avenues of exploration for therapy based on the underlying mechanisms of AA with a focus not only on T-cell subsets and mechanisms common to other forms of autoimmunity, but also on unique mechanisms that involve signalling pathways downstream of the NKG2D receptor.

METHODS SUMMARY

Cases were ascertained through the National Alopecia Areata Registry (NAAR) with approval from institutional review boards and genotyped on the Illumina HumanHap 610 chip. Three sets of previously published control data sets were used for comparison of allele frequencies. These had been genotyped on the Illumina HumanHap 550v2. All samples were confirmed to be of European ancestry by principal component analysis with ancestry informative markers. Stringent quality control measures were used to remove samples and markers that did not exceed predefined thresholds. Tests of association were run with and without measures to control for residual population stratification. Tissue specimens and RNA from human scalp biopsies were obtained with approval from institutional review boards. All experiments were performed according to the Helsinki guidelines.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 January; accepted 22 April 2010.

1. Cooper, G. S., Bynum, M. L. & Somers, E. C. Recent insights in the epidemiology of autoimmune diseases: improved prevalence estimates and understanding of clustering of diseases. *J. Autoimmun.* **33,** 197–207 (2009).
2. Safavi, K. H., Muller, S. A., Suman, V. J., Moshell, A. N. & Melton, L. J. III. Incidence of alopecia areata in Olmsted County, Minnesota, 1975 through 1989. *Mayo Clin. Proc.* **70,** 628–633 (1995).
3. Jelinek, J. E. Sudden whitening of the hair. *Bull. N. Y. Acad. Med.* **48,** 1003–1013 (1972).
4. Gilhar, A., Paus, R. & Kalish, R. S. Lymphocytes, neuropeptides, and genes involved in alopecia areata. *J. Clin. Invest.* **117,** 2019–2027 (2007).
5. Gilhar, A. *et al.* Transfer of alopecia areata in the human scalp graft/Prkdc(scid) (SCID) mouse system is characterized by a TH1 response. *Clin. Immunol.* **106,** 181–187 (2003).
6. Gilhar, A., Shalaginov, R., Assy, B., Serafimovich, S. & Kalish, R. S. Alopecia areata is a T-lymphocyte mediated autoimmune disease: lesional human T-lymphocytes transfer alopecia areata to human skin grafts on SCID mice. *J. Investig. Dermatol. Symp. Proc.* **4,** 207–210 (1999).
7. Ito, T., Meyer, K. C., Ito, N. & Paus, R. Immune privilege and the skin. *Curr. Dir. Autoimmun.* **10,** 27–52 (2008).
8. McDonagh, A. J. & Tazi-Ahnini, R. Epidemiology and genetics of alopecia areata. *Clin. Exp. Dermatol.* **27,** 405–409 (2002).
9. Van der Steen, P. *et al.* The genetic risk for alopecia areata in first degree relatives of severely affected patients. An estimate. *Acta Derm. Venereol.* **72,** 373–375 (1992).
10. Jackow, C. *et al.* Alopecia areata and cytomegalovirus infection in twins: genes versus environment? *J. Am. Acad. Dermatol.* **38,** 418–425 (1998).
11. Martinez-Mir, A. *et al.* Genomewide scan for linkage reveals evidence of several susceptibility loci for alopecia areata. *Am. J. Hum. Genet.* **80,** 316–328 (2007).
12. Radosavljevic, M. *et al.* A cluster of ten novel MHC class I related genes on human chromosome 6q24.2–q25.3. *Genomics* **79,** 114–123 (2002).
13. Eagle, R. A. & Trowsdale, J. Promiscuity and the single receptor: NKG2D. *Nature Rev. Immunol.* **7,** 737–744 (2007).
14. Eagle, R. A., Traherne, A. J., Hair, J. R., Jafferji, I. & Trowsdale, J. ULBP6/RAET1L is an additional human NKG2D ligand. *Eur. J. Immunol.* **39, 3207–3216** (2009).
15. Caillat-Zucman, S. How NKG2D ligands trigger autoimmunity? *Hum. Immunol.* **67,** 204–207 (2006).
16. Matzinger, P. The danger model: a renewed sense of self. *Science* **296,** 301–305 (2002).
17. Strid, J. *et al.* Acute upregulation of an NKG2D ligand promotes rapid reorganization of a local immune compartment with pleiotropic effects on carcinogenesis. *Nature Immunol.* **9,** 146–154 (2008).
18. Borchers, M. T. *et al.* Sustained CTL activation by murine pulmonary epithelial cells promotes the development of COPD-like disease. *J. Clin. Invest.* **119,** 636–649 (2009).
19. Ito, T. *et al.* Maintenance of hair follicle immune privilege is linked to prevention of NK cell attack. *J. Invest. Dermatol.* **128,** 1196–1206 (2008).
20. Zhang, Q., Li, J., Deavers, M., Abbruzzese, J. L. & Ho, L. The subcellular localization of syntaxin 17 varies among different cell types and is altered in some malignant cells. *J. Histochem. Cytochem.* **53,** 1371–1382 (2005).
21. Rosengren Pielberg, G. *et al.* A *cis*-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. *Nature Genet.* **40,** 1004–1009 (2008).
22. Akar, A, *et al.* Antioxidant enzymes and lipid peroxidation in the scalp of patients with alopecia areata. *J. Dermatol. Sci.* **29,** 85–90 (2002).
23. Holley, J. E., Newcombe, J., Winyard, P. G. & Gutowski, N. J. Peroxiredoxin V in multiple sclerosis lesions: predominant expression by astrocytes. *Mult. Scler.* **13,** 955–961 (2007).
24. Gerard, A. C., Many, M. C., Daumerie, C., Knoops, B. & Colin, I. M. Peroxiredoxin 5 expression in the human thyroid gland. *Thyroid* **15,** 205–209 (2005).
25. Wang, M. X. *et al.* Expression and regulation of peroxiredoxin 5 in human osteoarthritis. *FEBS Lett.* 531 359–362 (2002).
26. Karasawa, R., Ozaki, S., Nishioka, K. & Kato, T. Autoantibodies to peroxiredoxin I and IV in patients with systemic autoimmune diseases. *Microbiol. Immunol.* **49,** 57–65 (2005).
27. Pan, F. *et al.* Eos mediates Foxp3-dependent gene silencing in $CD4^+$ regulatory T cells. *Science* **325,** 1142–1146 (2009).
28. Wing, K. *et al.* CTLA-4 control over $Foxp3^+$ regulatory T cell function. *Science* **322,** 271–275 (2008).
29. Monteleone, G., Pallone, F. & Macdonald, T. T. Interleukin-21 as a new therapeutic target for immune-mediated diseases. *Trends Pharmacol. Sci.* **30,** 441–447 (2009).
30. Gregersen, P. K. & Olsson, L. M. Recent advances in the genetics of autoimmune disease. *Annu. Rev. Immunol.* **27,** 363–391 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the many patients and their family members who participated in the National Alopecia Areata Registry from which the patient cohort was derived; S. Schwartz, D. A. Greenberg, S. E. Hodge, R. Ottman, K. Kiryluk, J. Lee, J. D. Terwilliger and R. Plenge for discussions about statistical methodology; A. Bowcock, M. Girardi, R. Clark, J. Trowsdale, R. Clynes, S. Ghosh and R. Bernstein for critical insights and perspectives on genetics, hair and immunobiology; C. Higgins, M. Kurban, M. Kiuru, H. Lam and M. Zhang for expert assistance in the laboratory; A. Martinez-Mir, M. Peacocke, A. Zlotogorski, M. Grossman, P. Schneiderman, D. Gordon and J. Ott for their critical input in the early phases of this study. We are grateful to the National Alopecia Areata Foundation (NAAF) for support of funding the initial studies, and to V. Kalabokes and her staff at NAAF for their efforts on our behalf. The patient cohort was collected and maintained by the National Alopecia Areata Registry (N01AR62279) (to M.D.). This work was supported in part by the DFG Cluster of Excellence, Inflammation at Interfaces (to R.P.) and by the National Institutes of Health grants R01AR44422 (to C.I.A. and P.K.G.), R01CA133996 and P30CA016772 (to C.I.A.) and R01AR52579 and R01AR56016 (to A.M.C.).

Author Contributions L.P. performed technical aspects in preparation of samples for genotyping, the statistical analysis and preparation of the manuscript. M.D., V.P., M.H. and D.N. participated in phenotyping, diagnosis, and access to patient samples from the National Alopecia Areata Registry. Y.S., P.S. and H.K. provided expertise in RT–PCR and immunofluorescence. K.C.M. and R.P. provided expertise in immunhistochemistry. A.L. and P.K.G. provided control samples and performed genotyping as well as insight into autoimmune diseases. W.V.C. and C.I.A. provided additional statistical analysis and control samples from a distinct cohort. C.A.B.J. performed hair follicle microdissection and provided indispensable scientific expertise on the dermal sheath. A.M.C. provided oversight and conceptual guidance to the project, input into the functional significance of candidate genes, supervision of laboratory personnel, management of collaborations, preparation of the manuscript and all reporting requirements for granting agencies.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.M.C. (amc65@columbia.edu).

METHODS

Sample ascertainment. Cases were ascertained through the National Alopecia Areata Registry[31] with approval from institutional review boards. Written informed consent was obtained from all participants. To reduce the possibility of bias resulting from population stratification, only patients who self-reported European ancestry were selected for genotyping. Three sets of previously published control data sets were used for comparison of allele frequencies. These were obtained from the New York Cancer Project[32,33] and the Cancer Genetic Markers of Susceptibility breast[34] and prostate[35] cancer studies. All cases were genotyped on the Illumina HumanHap 610 chip and all controls were genotyped on the Illumina HumanHap 550v2. All samples were confirmed to be of European ancestry by principal component analysis with ancestry-informative markers. Stringent quality control measures were used to remove samples and markers that did not exceed predefined thresholds. Tests of association were run with and without measures to control for residual population stratification. Tissue specimens and RNA from human scalp biopsies were obtained with approval from institutional review boards. All experiments were performed in accordance with the Helsinki guidelines.

Genotyping. Quality control was performed with Helix Tree software (Golden Helix) or PLINK (http://pngu.mgh.harvard.edu/purcell/plink/)[36]. We removed SNPs that lacked more than 5% data, that did not follow Hardy–Weinberg equilibrium in controls ($P < 0.0001$), that were not present in both Illumina 550Kv2 and Illumina 610K, or that were indicated to have discordant clustering between platforms by Illumina, leaving 463,301 SNPs for analysis. Next, 19 samples with more than 10% missing genotype data were removed. In addition, three case and eight control samples that shared more than 25% inferred identity by descent were removed. Principal component analysis using a subset of 3,568 ancestry informative markers[37] identified 5 cases and 12 controls as ethnic outliers, and these were removed before analysis. Samples more than six standard deviations units from five components were excluded from subsequent analysis. Visual inspection of a plot of the first two eigenvectors identified 141 controls for which matched cases did not exist. These were excluded from further analysis.

Statistics. Reported association values were obtained with logistic regression assuming an additive genetic model and included a covariate to adjust for any residual population stratification. Statistics unadjusted for residual population stratification were also examined, as well as P values obtained with the false discovery rate method, and were found to be equivalent to reported values. LD was quantified and evaluated with Haploview[38]. SAS was used to perform stratified analysis and logistic modelling to determine whether SNPs shared a common haplotype. If the adjusted OR differed from the crude estimate by more than 10%, a common haplotype was inferred. Assessment of individual genetic liability was performed in Excel (Microsoft). A single marker was chosen as a proxy for each of the independent risk haplotypes. Alleles for the 16 proxy markers were coded 1 if associated with increased risk and 0 otherwise, and then summed for each individual. A two-tailed Student's t-test was used to determine the significance of the difference in the distribution of risk alleles between cases and controls, under an assumption of unequal variance. The population-attributable fraction (F_p) for each SNP was calculated as

$$F_p = \frac{\sum_i F_i(R_i - 1)}{1 + \sum_i F_i(R_i - 1)}$$

where R_i indexes the estimate associated with heterozygous and homozygous carriage of risk-increasing genotypes compared to the normal heterozygote, and F_i denotes the genotype frequencies in the controls. LD-based imputation using the Markov Chain Haplotyping algorithm (MACH 1.0.16; http://www.sph.umich.edu/csg/abecasis/mach/tour/imputation.html) was used to perform genome-wide maximum-likelihood genotype imputation. A weighted logistic regression test on binary trait using mach2dat was used to assess the quality of the imputation, again followed by a logistic regression association test assuming an additive model with top ten principal components as covariates to adjust for any residual population stratification using PLINK.

Tissue specimens. Human skin scalp biopsies were obtained from 18 patients with AA (age range 28–77 years) from a lesional area; control samples were either biopsies of frontotemporal human skin scalp taken from seven healthy women undergoing facelift surgery (age range 35–67 years) or biopsies from the occipital region of human skin scalp of two healthy men. All experiments were performed in accordance with the Helsinki guidelines. Specimens were embedded directly in OCT compound, or fixed in 10% formalin and embedded in paraffin blocks and cut into sections 5 μm thick.

Immunohistology. To detect ULBP3 protein expression *in situ*, a staining method based on labelled streptavidin–biotin was performed. In brief, paraffin sections were deparaffinized and immunostained overnight at 4 °C, after antigen retrieval with citrate buffer and using appropriate blocking steps against endogenous peroxidase, with the use of the rabbit anti-human ULBP3 antibody (1:250 dilution in antibody diluent; DCS). All incubation steps were interspersed by washing three times with Tris-buffered saline (10 mM Tris/HCl pH 7.6, 150 mM NaCl) for 5 min each. This was followed by staining with a biotinylated PolyLink secondary antibody (DCS) for 20 min at room temperature, between 20 and 25 °C, and developed for 20 min at room temperature with the peroxidase–streptavidin conjugate (DCS) method. Finally, the slides were labelled with 3-amino-9-ethylcarbazole substrate (Vector Elite ABC Kit; Vector Laboratories) and counterstained with haematoxylin.

Quantitative immunohistomorphometry. The number of ULBP3$^+$ cells was evaluated in three microscopic fields at 200× magnification in the dermis, in the dermal sheath of the hair follicle and perifollicular around each hair bulb from patients with AA, and in control skin. All data were analysed by a Mann–Whitney test for unpaired samples (results are expressed as means ± s.e.m.; $P < 0.05$ was regarded as significant).

Indirect immunofluorescence. Indirect immunofluorescence on fresh frozen sections of human scalp skin was performed as described previously[39]. The primary antibodies used were mouse monoclonal anti-ULBP3 (clone 2F9; 1:50 dilution; Santa Cruz Biotechnology), rabbit polyclonal anti-CD3 (1:50; DAKO), mouse monoclonal anti-CD8 (clone C8/144B; prediluted; Abcam), rabbit polyclonal anti-CD8 (1:200; Abcam), mouse monoclonal anti-NKG2D (clone 1D11; 1:100; Abcam), rabbit polyclonal anti-STX17 (1:500; Sigma), rabbit polyclonal anti-PRDX5 (1:500; Abnova), guinea-pig polyclonal anti-K74 (1:2,000), and guinea-pig polyclonal anti-K31 (1:8,000). The anti-K74 and anti-K31 antibodies were provided by L. Langbein.

RT–PCR analysis. Total RNA was isolated from scalp skin and whole blood of a healthy control individual by using the RNeasy Minikit (Qiagen) in accordance with the manufacturer's instructions. Total RNA (2 μg) was reverse transcribed with oligo(dT) primers and SuperScript III (Invitrogen). With the use of the first-strand complementary DNAs as templates, PCR was performed with Platinum PCR SuperMix (Invitrogen) and the primer pairs shown in Supplementary Table 8. The amplification conditions were 94 °C for 2 min, followed by 35 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 50 s, with a final extension at 72 °C for 7 min. PCR products were run on 2.0% agarose gels. Real-time PCR was performed on an ABI 7300 (Applied Biosystems). PCR reactions were performed with ABI SYBR Green PCR Master Mix, 300 nM primers and 50 ng of cDNA with the following consecutive steps: 50 °C for 2 min, 95 °C for 10 min, and 40 cycles of 95 °C for 15 s and 60 °C for 1 min. The samples were run in triplicate and normalized to an internal control (the glyceraldehyde-3-phosphate dehydrogenase gene) with the accompanying software.

31. Duvic, M., Norris, D., Christiano, A., Hordinsky, M. & Price, V. Alopecia areata registry: an overview. *J. Investig. Dermatol. Symp. Proc.* **8,** 219–221 (2003).
32. Mitchell, M. K., Gregersen, P. K., Johnson, S., Parsons, R. & Vlahov, D. The New York Cancer Project: rationale, organization, design, and baseline characteristics. *J. Urban Health* **81,** 301–310 (2004).
33. Plenge, R. M. *et al.* TRAF1-C5 as a risk locus for rheumatoid arthritis—a genomewide study. *N. Engl. J. Med.* **357,** 1199–1209 (2007).
34. Hunter, D. J. *et al.* A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nature Genet.* **39,** 870–874 (2007).
35. Yeager, M. *et al.* Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nature Genet.* **39,** 645–649 (2007).
36. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
37. Tian, C. *et al.* European population genetic substructure: further definition of ancestry informative markers for distinguishing among diverse european ethnic groups. *Mol. Med.* **15, 371–383** (2009).
38. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21,** 263–265 (2005).
39. Bazzi, H. *et al.* Desmoglein 4 is expressed in highly differentiated keratinocytes and trichocytes in human epidermis and hair follicle. *Differentiation* **74,** 129–140 (2006).