

ARTICLE

Received 17 Dec 2014 | Accepted 1 Jul 2015 | Published 10 Aug 2015

DOI: 10.1038/ncomms8971

OPEN

# Genome-wide analysis of the genetic regulation of gene expression in human neutrophils

Anand Kumar Andiappan<sup>1</sup>, Rossella Melchiotti<sup>1</sup>, Tuang Yeow Poh<sup>1</sup>, Michelle Nah<sup>1</sup>, Kia Joo Puan<sup>1</sup>, Elena Vigano<sup>1</sup>, Doreen Haase<sup>1</sup>, Nurhashikin Yusof<sup>1</sup>, Boris San Luis<sup>1</sup>, Josephine Lum<sup>1</sup>, Dilip Kumar<sup>1</sup>, Shihui Foo<sup>1</sup>, Li Zhuang<sup>1</sup>, Anusha Vasudev<sup>1</sup>, Astrid Irwanto<sup>2</sup>, Bernett Lee<sup>1</sup>, Alessandra Nardin<sup>1</sup>, Hong Liu<sup>3,4,5,6</sup>, Furen Zhang<sup>3,4,5,6</sup>, John Connolly<sup>1</sup>, Jianjun Liu<sup>2,7</sup>, Alessandra Mortellaro<sup>1</sup>, De Yun Wang<sup>8</sup>, Michael Poidinger<sup>1</sup>, Anis Larbi<sup>1</sup>, Francesca Zolezzi<sup>1</sup> & Olaf Rotzschke<sup>1</sup>

Neutrophils are an abundant immune cell type involved in both antimicrobial defence and autoimmunity. The regulation of their gene expression, however, is still largely unknown. Here we report an eQTL study on isolated neutrophils from 114 healthy individuals of Chinese ethnicity, identifying 21,210 eQTLs on 832 unique genes. Unsupervised clustering analysis of these eQTLs confirms their role in inflammatory responses and immunological diseases but also indicates strong involvement in dermatological pathologies. One of the strongest eQTL identified (rs2058660) is also the tagSNP of a linkage block reported to affect leprosy and Crohn's disease in opposite directions. In a functional study, we can link the C allele with low expression of the  $\beta$ -chain of IL18-receptor (IL18RAP). In neutrophils, this results in a reduced responsiveness to IL-18, detected both on the RNA and protein level. Thus, the polymorphic regulation of human neutrophils can impact beneficial as well as pathological inflammatory responses.

<sup>1</sup> Singapore Immunology Network (SIgN), Agency for Science Technology and Research (A\*STAR), #04-06, 8A Biomedical Grove, Singapore, Singapore. <sup>2</sup> Department of Human Genetics, Genome institute of Singapore (GIS), Singapore, Singapore. <sup>3</sup> Shandong Provincial Hospital for Skin Diseases, Shandong University, Jinan, Shandong, China. <sup>4</sup> Shandong Provincial Institute of Dermatology and Venereology, Shandong Academy of Medical Sciences, Jinan, Shandong, China. <sup>5</sup> School of Medicine, Shandong University, Shandong Provincial Medical Center for Dermatovenereology, Jinan, Shandong, China. <sup>6</sup> Shandong Provincial Institute of Dermatology and Venereology, Provincial Academy of Medical Science, Jinan, Shandong, China. <sup>7</sup> School of Life Sciences, Anhui Medical University, Hefei, Anhui, China. <sup>8</sup> Department of Otolaryngology, National University of Singapore, Singapore. Correspondence and requests for materials should be addressed to O.R. (email: Olaf\_Rotzschke@immunol.a-star.edu.sg).

**N**eutrophils are key players in the innate immune system and crucial for front-line defence against pathogens. They are the first leukocyte subset to migrate to the site of infection, where they recruit other immune cells and activate both the innate and adaptive branches of the immune system<sup>1,2</sup>. Neutrophils have several features that equip them to contain bacterial and fungal pathogens including high levels of phagocytosis, the ability to form neutrophil extracellular traps and the capacity to secrete cell-toxic and proinflammatory compounds<sup>3,4</sup>. However, these functions must be tightly regulated to avoid excessive destruction of host tissues or the induction of self-reactive immune cells.

On the genetic level this could be achieved by regulatory polymorphisms. genome-wide association studies (GWAS) have found that the majority of allelic variants linked with disease states fall in non-coding regions of the genome<sup>5,6</sup> suggesting that complex pathologies are primarily influenced by genetic control of gene expression. Such non-coding genomic regions with the capacity to affect specific mRNA expression levels are termed 'expression quantitative trait loci' (eQTLs), and their characterization may thus enable us both to understand disease susceptibility within human populations and to identify important molecular drivers of pathology. Recent reports on eQTLs in isolated immune subsets have yielded strong results providing insights into the genetic basis of disease susceptibility<sup>7–10</sup>.

In this study we investigated the allelic control of gene expression in neutrophils on the genome wide level and analysed their potential impact on cellular function. We isolated neutrophils from 114 blood samples taken from healthy donors of Chinese descent and performed a genome-wide eQTL analysis to characterize the impact of *cis* polymorphisms on neutrophil transcriptional regulation. The genome-wide correlation of genotype data with gene expression profile revealed more than 800 genes whose expression was affected by single nucleotide polymorphisms (SNPs) within eQTLs in neutrophils.

## Results

By analysing neutrophils from a total of 114 individuals, we identified 21,210 eQTLs that significantly affected 971 distinct probes within 832 unique genes at a permutation significance threshold of  $P < 0.001$ . The Manhattan plot in Fig. 1a shows the negative logarithm of the association  $P$  value for each eQTL in neutrophils from each donor, across all chromosomes, while the Q-Q plot (Fig. 1b) indicates significant deviations of observed  $P$  values (linear regression analysis, see Methods for more information) from those expected. A complete list is provided in the (Supplementary Data 1).

**Neutrophil eQTLs shared with other immune subsets.** To estimate the extent to which the neutrophil eQTLs were common to other immune cell subsets we compared our data with a published data set containing the *cis* eQTLs of B cells and monocytes isolated from 288 Caucasian individuals<sup>7</sup>. This data set was particularly interesting because neutrophils and monocytes share a myeloid origin, while B lymphocyte differentiation diverges from myeloid cells at the hematopoietic stem cell level. The definition of eQTL by Fairfax *et al.*, was different from our definition of eQTL as they didn't use the additional permutation  $P$  value cutoff and thus the difference in the number of eQTLs. When using a linkage disequilibrium threshold of  $r^2 = 0.8$ , only 248 of the 971 eQTLs (25.54%) are found in all three immune subsets (Fig. 1c). This number was significantly larger than expected by chance with a  $P$  value of  $1.5 \times 10^{-105}$  computed using a hyper geometric distribution.

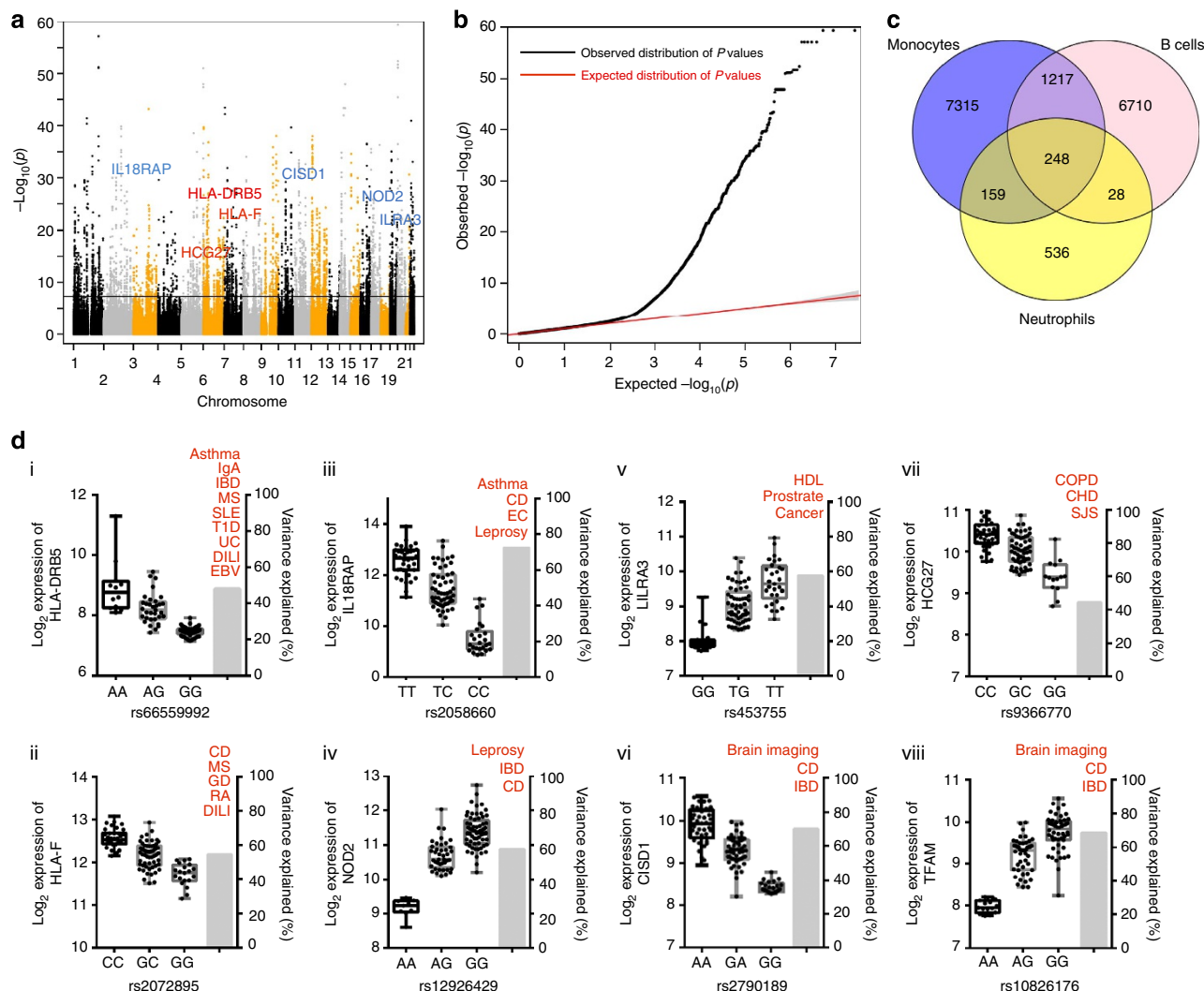
The neutrophils shared a greater number of eQTLs with monocytes (159 or 16.37%) as expected, than with the less-related B cells (28 or 2.88%). However, majority of the detected eQTLs (536 or 55.20%), were only differentially expressed in neutrophils and not in B cells or monocytes. The information on the overlapping probes and their corresponding SNPs are provided in (Supplementary Data 3 and 4).

**In silico pathway analysis of neutrophil eQTLs.** To explore the potential functional implications of the eQTLs detected in neutrophils we used QIAGEN's Ingenuity Pathway Analysis (IPA) tool to identify the biological pathways and processes that were most represented in our eQTL data set (Table 1). In line with the crucial role in pathogen defence, neutrophil eQTLs were significantly enriched in genes participating in the induction of the inflammatory response and/or the manifestation of inflammatory diseases Table 1. This enrichment was in comparison with the entire gene set of all human genes as catalogued by IPA. A significant enrichment was also detected for genes associated with dermatological diseases and disorders as estimated from the  $P$  value of  $1.46 \times 10^{-4}$ – $2.48 \times 10^{-2}$  (69 eQTLs).

While the enrichment of neutrophil-specific eQTLs in genes involved in immune-related pathways was expected, their marked abundance in skin disease-associated genes was somewhat unexpected and intriguing. We therefore further investigated the association of neutrophil eQTLs with dermatological syndromes: IPA revealed that 43 of the neutrophil eQTLs were associated with psoriasis, 35 with dermatitis, 26 with atopic dermatitis and 2 with dermatitis of ear (Table 2). The enrichment of neutrophil eQTLs was particularly marked for both psoriasis and dermatitis, where association  $P$  values (IPA based enrichment analysis) reached  $8.1 \times 10^{-4}$  and  $1.5 \times 10^{-4}$ , respectively. Taking this analysis further, we interrogated gene expression data comparing individuals with psoriasis with healthy controls (GEO data set GSE13355 for psoriasis<sup>11</sup>); interestingly, 41 of the 43 neutrophil eQTL associated genes linked with psoriasis in our data were also differentially expressed in psoriasis versus healthy skin samples of the two groups (Supplementary Fig. 1). The corresponding  $P$  values and  $\log_2$  fold change in gene expression data are provided in Supplementary Table 2. Similar results were also obtained using a second cohort (Gene Expression Omnibus data set GSE14905 ref. 12) where 34 probes out of the 43 were significantly differentially expressed in psoriasis after false discovery rate correction ((Supplementary Figure 2 and Supplementary Table 3). The analysis of a small gene expression data set from doctor-diagnosed dermatitis patients (GSE5667 ref. 13) did not reveal any association with our neutrophil eQTLs, but confirmation using a larger cohort would be necessary to draw any firm conclusions.

## Neutrophil eQTLs linked to GWAS defined disease markers.

Given the diverse immunological roles of neutrophils we next asked more broadly whether neutrophil eQTLs were associated with any other disease states in the human population. We exploited the complete Catalog of Published Genome-Wide Association Studies (GWAS) ([www.genome.gov/gwastudies](http://www.genome.gov/gwastudies)) that reports 14947 SNPs taken from 2071 publications considering varied pathological conditions. To facilitate maximum retrieval of possible associations and to limit the impact of differences in ethnicity and varying use of tagSNPs in the genotyping platforms, we extended our set of 971 neutrophil eQTLs by including SNPs that were linked by an  $r^2$  value exceeding 0.8. This combined analysis of the GWAS data sets revealed 89 diseases/traits that were associated with one or more genes linked to neutrophil



**Figure 1 | Summary of eQTL identified in neutrophils.** (a) Manhattan plot of genome-wide distribution, and (b) Quantile-Quantile (Q-Q) plot, of the negative logarithm of the association  $P$  values for eQTLs in neutrophils. Candidates in blue on the panel (a) such as IL18RAP, HLA-DRB5, CISD1 and LILRA3 were neutrophil specific, while HCG27, TFAM, HLA-F and NOD2 are shared across other immune cells and displayed in red. (c) Extent of sharing of eQTLs in neutrophils, B cells and monocytes, using comparative data from (Fairfax *et al.*<sup>7</sup>). (d) Selected neutrophil eQTL candidates associated with SNP markers that have been linked with the noted disease phenotypes.  $\log_2$  expression, error bars for median  $\log_2$  expression and standard deviation are shown (left y axis) by SNP genotype (x axis) for each of eight genes significant as neutrophil eQTLs together with variance explained (right y axis). Crohn's disease (CD), immunoglobulin A (IgA), inflammatory bowel disease (IBD), multiple sclerosis (MS), systemic lupus erythematosus (SLE), type 1 diabetes (T1D), ulcerative colitis (UC), drug-induced liver injury (DILI), epstein-barr virus (EBV) status, total cholesterol measured as HDL, coronary heart disease (CHD), chronic obstructive pulmonary disease (COPD), Stevens-Johnson syndrome (SJS), rheumatoid arthritis (RA), graves disease (GD), eosinophil counts (EC).

eQTLs (Supplementary Table 1 and Supplementary Data 2). This was confirmed as significant by enrichment analysis using a hypergeometric distribution for the probability of observing the overlap of the genes affected by the eQTLs against the frequencies reported in the GWAS catalogue (Supplementary Table 4). Among the identified conditions, autoimmune diseases were particularly well-represented: Crohn's disease, multiple sclerosis, ulcerative colitis, inflammatory bowel disease (IBD), as well as systemic lupus erythematosus, Grave's disease, type 1 diabetes and vitiligo were all significantly associated with genes linked to the set of neutrophil eQTLs (Table 3).

Several neutrophil eQTLs were associated with GWAS variants that had been linked to more than one disease/trait, suggesting a more widespread functional impact. The contribution to disease manifestation as well as the genotype-effect on gene expression for some of the more prominent neutrophil eQTLs is shown in

Fig. 1d. With the exception of the two MHC genes, *HLA-DRB5* and *HLA-F*, all of these eQTLs were expressed in neutrophils but not monocytes or B cells (compare Fig. 1c). The targeted genes are involved in immune-recognition/regulation and mitochondrial function: *HLA-DRB5* encodes the  $\beta$ -chain of one of the common MHC class II molecules; *HLA-F* is a non-classical class I-like molecule that was recently identified as a ligand for killer inhibitory receptors on NK cells<sup>14</sup>; *IL18RAP* is the  $\alpha$ -chain of the IL18-receptor; *NOD2* is a pattern recognition receptor for bacterial peptidoglycans; *LILRA3* can bind to MHC class I heavy chains and *HCG27* is a long non-coding RNA gene located in the HLA gene locus. *CISD1* is a mitochondrial membrane protein, while *TFAM* represents a mitochondrial transcription factor. The latter had recently been associated with the modulation of innate immune responses by regulating mitochondrial DNA stress<sup>15</sup>.

Functional characterization of the IL18RAP eQTL rs2058660.

A common feature of all the examples in Fig. 1d is that the eQTL has a strong impact on the respective gene: each account for 40–80% of the observed expression variance of its target gene, implying a major contribution to the disease susceptibility associated with the gene. A particularly interesting example is rs2058660, which accounts for 70% of the variance in *IL18RAP* expression. *IL18RAP*’s ligand, IL18, belongs to the IL-1 cytokine super-family and has a key role in the initiation of proinflammatory immune responses. Accordingly, *IL18RAP* has been identified as a GWAS candidate for asthma and other immune-related diseases including Crohn’s disease<sup>16</sup> and leprosy<sup>17</sup>. A comparison of the negative logarithm association *P* values for SNPs in the gene locus shows that a strong GWAS disease association coincides with a strong effect on neutrophil *IL18RAP* expression (Fig. 2). SNPs exhibiting a low association *P* value in our expression analysis (Fig. 2a), also tended to exhibit low association *P* values in the leprosy GWAS (Fig. 2b) as well as the IBD-GWAS (Fig. 2c) data sets used. This applies particularly for the two SNPs linked to *IL18RAP* expression, rs2058660 and rs917997, which were identified as top candidates based on *P* values in the respective GWAS-studies. For the leprosy GWAS the *P* values displayed in Fig. 2b relate only to the discovery GWAS data and hence the *P* values do not reach genome-wide significance. However, rs2058660 was replicated at genome-wide significance<sup>17</sup> along with the replication cohort with a *P* value of  $1.59 \times 10^{-23}$ .

Therefore, we went on to investigate the functional implications of the allelic state of rs2058660 on IL18-mediated downstream effects in neutrophils. The C allele is associated with low expression of *IL18RAP* (Fig. 1d), and according to the GWAS data, the ‘TT’ genotype is linked with a higher incidence of leprosy, while ‘CC’ is the risk genotype for IBD. Considering that an elevated immune response promotes pathogen defence, but also increases the risk of autoimmune reactions, we

hypothesize that the risk pattern observed might be explained by a compromised immune-response in low-IL18RAP-expressing T allele individuals. Fortin *et al.*<sup>18</sup> had also indicated that IL-18 can contribute to generation of several inflammatory cytokines in neutrophils. To test whether this inflammatory cytokine production was dependent on our disease-associated variant, we isolated neutrophils from a total of nine individuals with the ‘TT’ genotype and eight individuals with ‘CC’. We then stimulated the neutrophils overnight with IL-18 before measuring the abundance of proinflammatory gene transcripts (Fig. 3). The real-time PCR for *CCL3*, *CCL20* and *TNF-α* transcripts revealed significantly different expression levels between the two genotypes: in line with lower IL18RAP expression, exposing neutrophils from ‘CC’ genotype individuals to IL-18 resulted in smaller increases in transcription of the respective cytokines and chemokines. The same trend was evident for IL-8 secretion, as ‘CC’ individuals exhibited significantly smaller increases in IL-8 release in response to IL-18 than did ‘CT’ or ‘TT’ individuals. Thus, the *IL18RAP* eQTL clearly correlates with the induction proinflammatory response genes possibly by varying the amount of the IL18-receptor on the surface of neutrophils.

Discussion

Recent years have seen a dramatic increase in the number of reported genetic variations and SNPs have been implicated in susceptibility to a wide variety of diseases. The majority of the SNPs fall into non-coding regions of the genome, implying that the main function of these variations is the control of gene expression as opposed to variations in the gene product itself. Interestingly, the majority of eQTLs also appear to be cell type-specific<sup>7,19</sup>. Albeit the latter may be partly be affected by different thresholds and is possibly also compounded by the ‘winners curse’<sup>20</sup> more than half of the neutrophil eQTLs reported here were not present in either monocytes or B cells, and had not been reported before. Some of the neutrophil-specific eQTL genes discovered here have already been validated in a meta-analysis of full blood-eQTL in an computational bioinformatics analysis<sup>21</sup>, which detected 237 of our neutrophil eQTLs. Another whole blood transcriptome study<sup>22</sup> had identified 10,914 cis eQTLs which overlapped with 699 of the 832 genes identified in our study based on purified neutrophils. While on one hand it confirms our data, it also revealed 133 new genes (16%) that were uniquely defined in our data set. Moreover, as the full blood sample contains many other leukocytes, such monocytes and lymphocytes, a fair number of the eQTLs could initially not be related to neutrophils. Hence, a definite assignment to neutrophils is only possible when purified cells are used. Interestingly a mouse eQTL study by Mostafavi *et al.*<sup>10</sup> detected 84 (10.3%) of our 818 neutrophil eQTLs which could be

Table 1   Diseases or disorders enriched in the significant neutrophil eQTL data set as identified by ingenuity pathway analysis (IPA).		
Disease or disorder	P value	Number of eQTL
Inflammatory response	$2.39 \times 10^{-6}$ – $3.50 \times 10^{-2}$	130
Inflammatory disease	$1.46 \times 10^{-4}$ – $3.23 \times 10^{-2}$	94
Dermatological diseases and conditions	$1.46 \times 10^{-4}$ – $2.48 \times 10^{-2}$	69

Table 2   Genes reported by IPA for the category dermatological diseases and conditions.			
Disease	P value	Neutrophil eQTL genes	Number of eQTL
Psoriasis	$8.05 \times 10^{-4}$	ACPP,ACTA2,ANXA1,APOBEC3A,ARG1,ATXN3,C3AR1,CLEC7A,CSTB,CTSB,CTSK,CTSS, DBN1,EIF5,EIF5A,FAM26F,GM2A,GNA15,GSN,HAL,HLA-G,HMOX1,IL16,KDELR2,KRT10, KYNU,LILRB1,MBP,PTPRC,RGS1,RNASE3,S100A12,S100P,SEC23B,SIGIRR,TEK,THBD, TYMP,UBE2I,VNN1,VWF,YWHAB,ZNF91	43
Dermatitis	$1.46 \times 10^{-4}$	ACADVL,ANXA1,ANXA5,C3AR1,CAPZB,CCR3,CPNE1,CSK,CYTIP,F2RL1,FAS,FCER1G, FLOT1,GNB1L,GSN,GSR,HLA-DRB5,HNRNPR,IL16,KRT10,NFKB1,PGLYRP1,POLE4,RABGEF1, RNASE3,RNF138,RPS6KA4,SELL,SIGIRR,SPAG1,SPI1,STAT6,SYK,THBD,UBE2I	35
Atopic Dermatitis	$1.78 \times 10^{-2}$	ACADVL,ANXA1,ANXA5,CAPZB,CCR3,CPNE1,F2RL1,FCER1G,FLOT1,GNB1L,GSN, HLA-DRB5,HNRNPR,KRT10,RNF138,SPAG1,SPI1,SYK,UBE2I	19
Dermatitis of ear	$2.48 \times 10^{-2}$	CYTIP,STAT6	2

eQTL, expression quantitative trait loci; IPA, ingenuity pathway analysis



**Table 3 | Specific list of genes with neutrophil eQTLs which also had a significant GWAS P value for an autoimmune condition.**

Autoimmune/inflammatory condition	Neutrophil eQTLs
Crohn's disease	CISD1*, IL18RAP*, TRPT1*, TUFM*, DNLZ, ERAP2, HLA-F, RNASET2, TFAM
Inflammatory bowel disease	CISD1*, HLA-DRB5*, TFAM, UTS2
Ulcerative colitis	HLA-DRB5*, LOC650557*, PIM3*, NFKB1, MANBA, UTS2
Rheumatoid arthritis	ARAP1*, CENTD2*, HLA-F, LOC253039, LY6G5C, RNASET2
Type 1 diabetes	HLA-DRB5*, HOXA5*, LOC650557*, SULT1A2*, TUFM*
Multiple sclerosis	CPT1B*, ECGF1, HLA-DRB5*, HLA-F, TYMP*, XRCC6BP1
Graves disease	HCG4*, HLA-F, RNASET2
Vitiligo	HCG4*, RNASET2
Systemic lupus Erythematosus	HLA-DRB5*, UBE2L3
Behcet's disease	CCR3*
Ankylosing spondylitis	CAST*

eQTL, expression quantitative trait loci; GWAS, genome-wide association studies.

Combining data from the neutrophil eQTLs identified with the GWAS catalogue variations for disease associations at GWAS threshold we found a significant enrichment for autoimmune and inflammatory disease. Refer to supplementary table 1 and supplementary online file 1 for comprehensive list.

\*refers to the neutrophil eQTLs not discovered in the B cells or monocytes.

mapped between human and mouse. This further validates our results on the relevance of these neutrophil eQTLs in other species too. More information on the analysis of comparison of the various data sets has been provided in the methods section and in the online data set file 5.

Since the overlap of eQTLs even between related subsets was somewhat limited, the number of known eQTLs can represent only the tip of the iceberg; this applies even more so as many take effect only under certain functional conditions, such as in response to specific environmental stimuli<sup>23–26</sup>. Considering further the huge variety of different tissues and cellular subtypes forming the human body, a rather large fraction of genes is likely to be regulated by eQTLs and thus the collection of eQTL data for a selected cell subset can be considered highly informative on its function in health and disease.

We conducted an unbiased pathway analysis (IPA) which revealed an enrichment of neutrophil eQTLs in natural inflammatory responses and in inflammatory diseases consistent with the established role of neutrophils in host defence against pathogens and upon tissue damage<sup>27</sup>. Less expected, however, was the enrichment of neutrophil eQTLs in pathways of dermatological diseases. Further analysis of the neutrophil eQTL-targeted genes in two public databases of psoriasis gene expression revealed that a large proportion was significantly differentially expressed between psoriatic cases and controls. Psoriasis is an immune-mediated disorder characterized by epidermal hyper-proliferation<sup>28</sup>, and neutrophils are known to accumulate under the stratum corneum of highly inflamed psoriatic lesions, where they activate T cells and promote the growth of epidermal keratinocytes<sup>29,30</sup>. Our eQTL analysis strongly supports a major contribution of the neutrophil immune cell subset to the pathology of psoriasis, and may be informative for further studies on the underlying mechanisms. A recent meta-analysis of psoriasis-related transcriptome and GWAS data sets also supports this claim. A significant fraction of the psoriasis GWAS candidates showed maximal expression in neutrophils thereby emphasizing its role in immunity<sup>31</sup>.

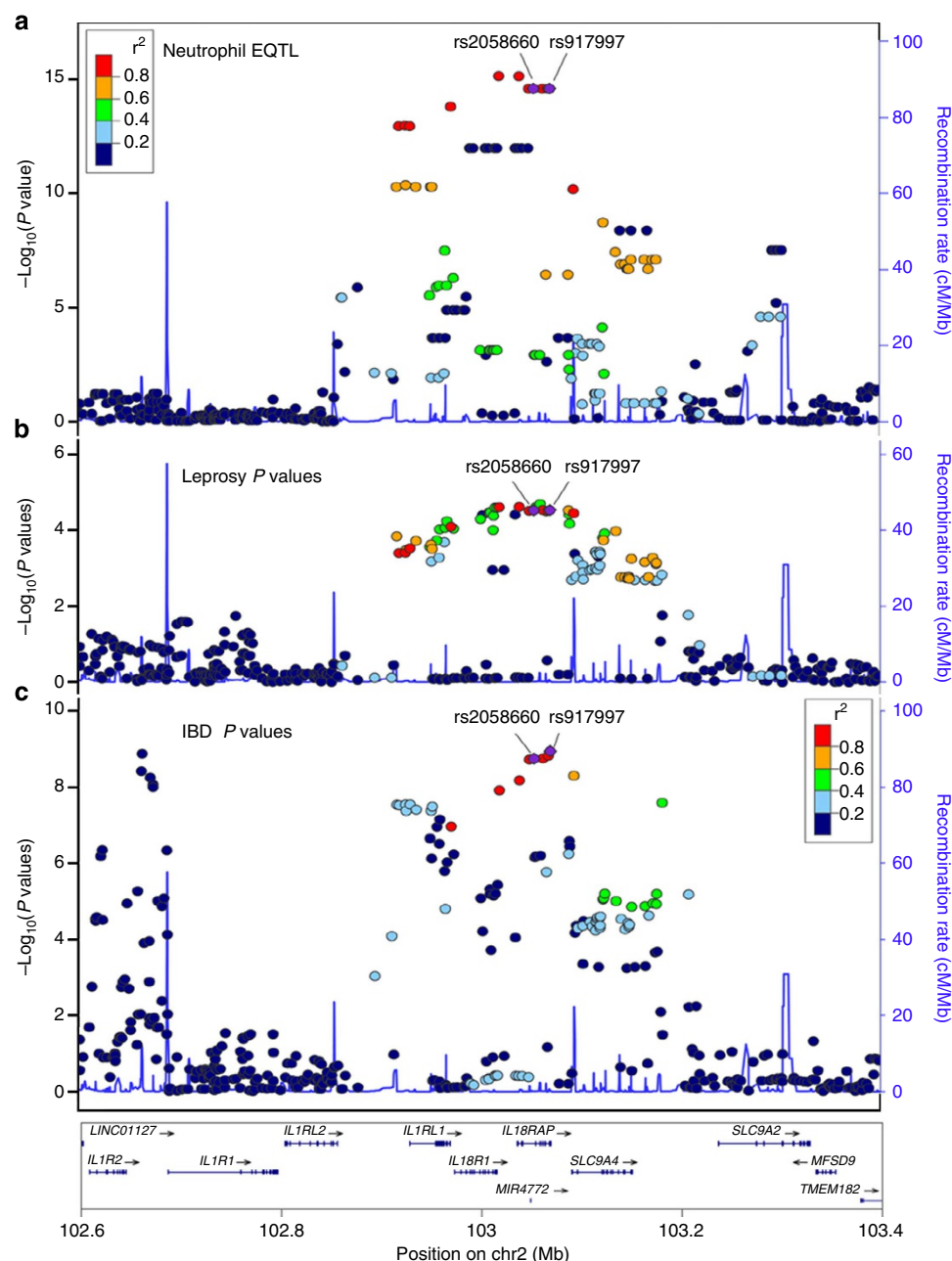
Besides pathway analysis, the correlation of neutrophil eQTLs with existing GWAS data is another approach to determine function and disease-implication of the subset in an unbiased way. Within a GWAS catalogue covering more than 2,000 syndromes and conditions<sup>32</sup>, significant overlap was detected between neutrophil eQTLs and genes associated with 89 phenotypes. A significant enrichment was also seen in GWAS for autoimmune conditions such as Crohn's disease, ulcerative colitis, systemic lupus erythematosus, Graves's disease, type 1

diabetes and vitiligo. In this study, we took *IL18RAP* as a candidate gene and showed that the differential expression levels that were associated with specific genotypes also had functional consequences in terms of modulating neutrophils' ability to produce the inflammatory cytokine profiles crucial for their immune functions<sup>18</sup>. Context-specific effects of the *IL18RAP* genotype on response to various stimuli have also been recently reported for monocytes and myeloid cells<sup>23,33</sup>, adding an additional layer of complexity to our understanding of the regulatory processes important for overall functioning of the immune system. Mostafavi *et al.*<sup>10</sup> demonstrated that some cis QTL are even shared between humans and mice. This implies that allelic regulation of key genes is crucial for the population to maintain evolutionary flexibility to effectively respond to selection pressures. Our observations may therefore help understand how variations in gene expression associate with disease risk for individuals based both on their haplotype and its effect on specific immune cell subsets.

In summary, using a genome-wide cis eQTL analysis we identified numerous genes whose expression was affected by the presence of polymorphic loci in neutrophils. This data set constitutes a valuable resource both for linking disease-associated genetic variants with function, as demonstrated by the strong overlap between eQTLs and GWAS signals, and for identifying diseases where a dysregulation of key neutrophil genes might play a role. Recent literature already expanded the role of neutrophils from simple suicidal killers to sophisticated players in host defense<sup>34–36</sup>, and our study now adds another layer of understanding of their functional complexity as well as highlighting their involvement in dermal diseases. As they are part of an integrated network of immune cells, eQTLs in neutrophils may have important implications for immune function and barrier protection and could contribute to inter-individual differences in disease susceptibility and pathogenesis.

## Methods

**Samples.** This study was performed with the approval of the Institutional Review Board (IRB Reference–NUS07-023 and NUS10-445) of the National University of Singapore and is in compliance with the Helsinki declaration. Blood samples for DNA analysis were collected from a cohort of 114 individuals of Chinese ethnicity, collected from random recruitment drives from January to August 2011 from the National University of Singapore with written informed consent. The demographics of the cohort reflects a median age of 21 (Range 18–29) with more 85% of the participants between 18 and 22 years of age and 42% females. All volunteers were free of serious medical conditions as determined by a questionnaire on self-reported medical history for conditions, such as autoimmune



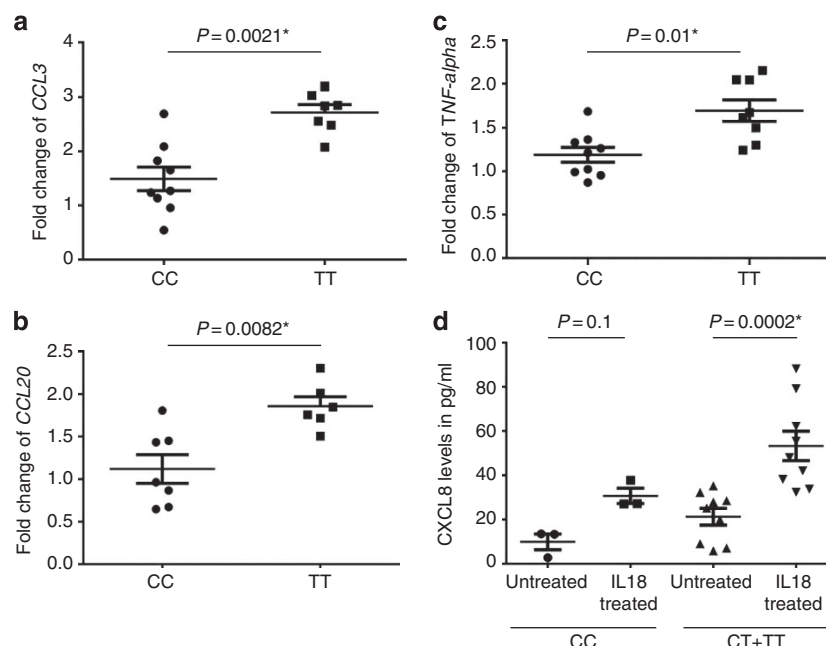
**Figure 2 | Genetic and functional association of IL18RAP locus SNPs with disease phenotype and possible immune responses.** (a) A linkage disequilibrium (LD) plot of the IL18RAP region with association  $P$  values for IL18RAP eQTL in neutrophils observed in this study. (b) LD plot of the IL18RAP region with association  $P$  values for leprosy as observed from Liu *et al* (AJHG, 2012). The disease-associated SNP rs2058660 was replicated at genome-wide significance in the replication cohort with a  $P$  value of  $1.59 \times 10^{-23}$  (c) LD plot of the IL18RAP region with association  $P$  values for IBD as observed from Jostins *et al*, Nature 2012.

and infectious disease. For this study, blood samples were taken from 114 donors of Chinese ethnicity and used for genotype and gene-expression analysis. Genomic DNA was extracted from whole blood samples and SNP genotyping carried out using the Illumina Omni5Quad system. Total RNA was extracted from the purified neutrophils.

**Neutrophil isolation and storage.** Following collection into a BD EDTA Vacutainer, 5 ml of blood was pipetted and diluted with 5 ml of sterile phosphate buffer saline (PBS) and underlaid with 5 ml of Ficoll Paque PLUS in a 15-ml Falcon tube. It was centrifuged at 1,800g at room temperature for 20 min. Peripheral blood mononuclear cells were aspirated into a 15-ml Falcon tube before addition of 6 ml of  $1 \times$  BD Pharm Lyse to remove red blood cells. After incubation for 15 min at room temperature the lysis solution was diluted to 15 ml with RPMI + 10% fetal bovine serum (FBS) and tubes centrifuged at 300g, at 4 °C for 5 min. The cell pellet was resuspended into 10 ml of RPMI + 10% FBS and

centrifuged at 300g, at 4 °C for 5 min. Cells were finally resuspended in 500  $\mu$ l of RPMI + 10% FBS and a 10  $\mu$ l sample taken for counting using a MACSQuant. Cells were then adjusted to the appropriate concentration for use in experiments as per the relevant methods section. Cell concentration was adjusted to  $5 \times 10^6$  cells per ml with RPMI + 10% FBS. The samples were aliquoted into sterile 96-well U-bottom plates and the remaining cell suspensions were centrifuged at 300g, 4 °C, 5 min. The supernatant was aspirated up to  $\sim 100$   $\mu$ l mark and 300  $\mu$ l of Trizol LS reagent was added before storage at  $-80$  °C. This preparation of neutrophil samples possibly has a contamination of eosinophils of  $\sim 1$ –4% and 1–2% of lymphocytes.

**RNA extraction from neutrophils.** Total RNA was extracted using the double extraction protocol in batches of 24 samples: acid guanidinium thiocyanate-phenol-chloroform extraction (Trizol Invitrogen) followed by mirVana miRNA Isolation Kit. Agilent Bioanalyzer was used to evaluate total RNA integrity and the RNA



**Figure 3 | Functional impact of the IL18RAP SNP rs2058660 genotypes.** A total of 19 donors were used for this experiment with  $n = 7$  for TT and  $n = 9$  for CC for the real time experiments. The gene expression of (a) *CCL3* (b) *CCL20* (c) *TNF-α* following treatment of neutrophils with IL-18 for 14 h, are displayed as a fold increase relative to unstimulated samples, with median and error bars for fold change of the respective gene expression values.  $P$  values shown are using a non-parametric Mann-Whitney test (d) CXCL8 levels measured by ELISA in the supernatant of the neutrophils cultured with IL-18 for 14 h. A total of 13 individuals was used for this experiment with  $n = 4$  for CC,  $n = 6$  for CT and  $n = 3$  for TT. The Kruskal-Wallis involving all groups was significant at  $P = 0.001^*$ .  $P$  values displayed in the figure D are pair wise Mann-Whitney test for each group between treated and untreated.

integrity number was calculated; all RNA samples had a RNA integrity number  $\geq 6.5$ . DNase digestion was not required since the reverse transcription (RT) reaction uses Oligo DT primers that bind specifically to the polyA tail of the RNA.

**Neutrophil stimulation experiments.** For the stimulation experiments neutrophils were isolated from blood as described above, within 4 h of its collection. Briefly, whole blood was diluted with an equal volume of PBS and the peripheral blood mononuclear cells and neutrophils were separated using  $1,077 \text{ g ml}^{-1}$  ficoll-plaque plus gradient density. The top layer of interphase cells were removed, leaving behind the neutrophils and erythrocytes. The erythrocytes were then lysed in an ammonium chloride-potassium buffer and the remaining neutrophils resuspended in RPMI1640 supplemented with 10% heat-inactivated FBS and 2 mM L-glutamine ( $1 \times 10^6$  cells/ml). Neutrophils were then treated for 14 h in the absence or presence of  $100 \text{ ng ml}^{-1}$  IL-18 or  $100 \text{ ng ml}^{-1}$  lipopolysaccharide. For the subsequent analysis by real-time PCR or enzyme-linked immunosorbent assay, RNA was isolated using Trizol. Real-time PCR was performed using BIO RAD RT-PCR machine. Real-time PCR data were expressed as delta CT, the difference in Ct between the gene of interest (goi) and the endogenous control (end ctl) for a given sample:  $\Delta\text{Ct} = \text{Ct}(\text{goi}) - \text{Ct}(\text{end.ctl})$ . The delta CT is then converted to linear fold change in gene expression using the following formula  $2^{(-\Delta\text{Ct})}$ . Results are then expressed as fold-change between the IL-18 treated samples against the mocked-treated samples between the genotypes.

**Cell supernatant measurements.** Neutrophils isolated and stimulated with IL-18 overnight for 14 h and then cell-free supernatants were collected. We then performed quantitative measurements of CXCL8 (Ebioscience, San Diego, California) and CXCL6 (Biolegend, San Diego, California) levels by enzyme-linked immunosorbent assay according to manufacturer's instructions.

**Genome-wide gene expression.** Biotinylated cRNA was prepared according to Illumina Ambion TotalPrep 96 RNA Amplification kit (Part Number 4393543) using 40 ng of RNA in batches of 48 samples. The protocol may be summarized into synthesis of first strand, synthesis of second strand, cDNA purification, *in vitro* transcription, cRNA purification. 750 ng of labelled cRNAs were hybridized on the Illumina HumanHT-12-v4 Expression BeadChip Kit (Catalogue number: BD-103-0204) at  $58^\circ\text{C}$  for 16 h. Samples are re-randomized for hybridization and were carried out in batches of 96 samples. Probe sequences provided by Illumina were mapped to genome build hg18 using the pipeline RUM (<http://cbil.upenn.edu/RUM/>). Of the total 47,323 probes found on the gene expression microarray a total of 41,469

probes uniquely mapped by the RUM pipeline to hg18 using the default parameters in the RUM pipeline. A total of 39,596 probes were located on the autosomes and retained for the analysis. We did not remove the probes which had low detection in samples as they did not have a major impact on the conclusions and also could be biologically interesting candidates due to the selective loss of expression in some individuals. However, we have flagged those probes which have expression in  $< 5\%$  samples in the online data set file 1. Quantile normalization without background subtraction was then performed using Genome Studio. Expression values were  $\log_2$  transformed before analysis. Probes containing common SNPs ( $\text{MAF} \geq 1\%$ ) according to the 1000 Genome Pilot Project<sup>37</sup> for CHB/JPT were excluded from the analysis (2,467 probes). Clustering was performed to check for any obvious batch effects and none were observed. Furthermore, a duplicate was detected as it showed an  $r^2 > 0.98$  and hence was removed from further analysis. Principal component analysis of the expression data reveals that the first 4 PC accounts for 26% of the variance and PC and gender correction had minimal change on the detected eQTLs.

**DNA extraction.** DNA was extracted from the blood samples using Qiagen DNeasy Blood & Tissue Kit (Catalogue number: 69504) according to manufacturer's instructions.

**Genome-wide SNP genotyping and quality control.** Illumina Human Omni5Quad that can detect up to 4.3 million SNPs was used for genotyping of SNPs. DNA samples were quantified using PicoGreen assay (Molecular Probes) and 400 ng of DNA were used as starting material. Amplifications and hybridizations were performed following the Infinium HD Assay as per manufacturer's instruction. Genome coordinates provided by Illumina were converted to genome build hg18 (for consistency with respect to probe annotation) using the UCSC's tool liftOver<sup>38</sup>. A total of 1,067 SNPs could not be converted and were removed from the analysis. SNPs located on sex chromosomes, random chromosomes and mitochondrial chromosomes were also excluded from the computations. Finally filtering for monomorphic and low call rate SNPs ( $< 95\%$ ) a total of 2,031,824 SNPs were retained for genome-wide cis eQTL analysis.

**Statistical analysis.** Genome-wide statistical association between SNPs and gene expression in the discovery cohort was evaluated using linear regression as implemented by the Apache Commons Mathematics Library. Genotypes were coded based on allele counts (0, 2 for homozygous genotypes and 1 for heterozygous genotypes).  $P$  values reported as zero were recomputed using the python function linregress from the library scipy to achieve a better precision. For each probe only SNPs located  $\pm 250 \text{ kbp}$  from the probe midpoint were tested for

association. To control the false discovery rate significance was determined by 10,000 phenotype permutations (label swapping) as described in Stranger *et al.*<sup>39</sup>, (originally described by Churchill *et al.*<sup>40</sup>) as against just raw *P* values. A <SNP, probe> pair was considered significant if its nominal *P* value was lower than the 0.001 tail of the distribution of minimal permuted *P* values across all SNPs tested for a particular probe. Each probe was analysed independently. Samples with missing genotypes for a particular SNP were excluded from the analysis of the corresponding SNP-probe pair. Illumina SNP IDs coded as kgp were mapped to the corresponding rs IDs by chromosomal position. Statistical association between genotype and gene expression (or protein expression) in the validation cohort was computed using a Kruskal–Wallis one-way analysis of variance (three genotypes) or Mann–Whitney (two genotypes). Multiple comparisons were performed using Dunn's multiple comparisons tests (Graphpad Prism 6).

**Ingenuity pathway analysis.** To understand the diseases and disorders associated, we performed network analyses of the entire set of neutrophil eQTL using the IPA tool, (Ingenuity Systems, www.ingenuity.com). Enrichment of the 971 unique probes associated with the 21,210 eQTLs with known pathways, processes and diseases was thus performed using IPA. The standard enrichment analysis in IPA was performed with the entire human genome as the background. Enriched pathways, processes and diseases were then extracted from IPA and reported.

For the psoriasis data sets, gene expression data for the 43 genes identified from the previous analysis was included and visualized as networks. The 43 genes associated with psoriasis as detected by IPA was visualized as networks with network connectivity constructed using manually curated data from IPA and gene expression fold changes between normal and disease used to colour the nodes (red for over-expression and green for decreased expression).

**Comparison with other data sets.** Supplementary data from Westra *et al.*, 2014 (ref. 21), Mostafavi *et al.*, 2014 (ref. 10) and Battle *et al.*, 2013 (ref. 21) were retrieved electronically and the eQTLs genes extracted using Biovia Pipeline Pilot (720, 1,222 and 10,914 extracted unique gene symbols). The extract gene names were then matched to the neutrophil eQTLs via their gene names to determine the overlaps between the various studies. In the case of the mouse data from Mostafavi *et al.*, 2014 (ref. 10), the genes were mapped to human gene symbols using the data available in HomoloGene build 68 (1055 mapped human gene symbols). The significance of the overlap was estimated using the hyper geometric distribution in R version 2.15.2.

## References

- Lehrer, R. I., Ganz, T., Selsted, M. E., Babior, B. M. & Curnutte, J. T. Neutrophils and host defense. *Ann. Intern. Med.* **109**, 127–142 (1988).
- Kumar, V. & Sharma, A. Neutrophils: cinderella of innate immune system. *Int. Immunopharmacol.* **10**, 1325–1334 (2010).
- Brinkmann, V. *et al.* Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532–1535 (2004).
- Branz, N. *et al.* Neutrophils sense microbe size and selectively release neutrophil extracellular traps in response to large pathogens. *Nat. Immunol.* **15**, 1017–1025 (2014).
- Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. Mapping complex disease traits with global gene expression. *Nat. Rev. Genet.* **10**, 184–194 (2009).
- Knight, J. C. Regulatory polymorphisms underlying complex disease traits. *J. Mol. Med. (Berl)* **83**, 97–109 (2005).
- Fairfax, B. P. *et al.* Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat. Genet.* **44**, 502–510 (2012).
- Zeller, T. *et al.* Genetics and beyond—the transcriptome of human monocytes and disease susceptibility. *PLoS ONE* **5**, e10693 (2010).
- Lappalainen, T. *et al.* Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501**, 506–511 (2013).
- Mostafavi, S. *et al.* Variation and genetic control of gene expression in primary immunocytes across inbred mouse strains. *J. Immunol.* **193**, 4485–4496 (2014).
- Nair, R. P. *et al.* Genome-wide scan reveals association of psoriasis with IL-23 and NF- $\kappa$ B pathways. *Nat. Genet.* **41**, 199–204 (2009).
- Yao, Y. *et al.* Type I interferon: potential therapeutic target for psoriasis? *PLoS ONE* **3**, e2737 (2008).
- Plager, D. A. *et al.* Early cutaneous gene transcription changes in adult atopic dermatitis and potential clinical implications. *Exp. Dermatol.* **16**, 28–36 (2007).
- Goodridge, J. P., Burian, A., Lee, N. & Geraghty, D. E. HLA-F and MHC class I open conformers are ligands for NK cell Ig-like receptors. *J. Immunol.* **191**, 3553–3562 (2013).
- West, A. P. *et al.* Mitochondrial DNA stress primes the antiviral innate immune response. *Nature* **520**, 553–557 (2015).
- Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).
- Liu, H. *et al.* Identification of IL18RAP/IL18R1 and IL12B as leprosy risk genes demonstrates shared pathogenesis between inflammation and infectious diseases. *Am. J. Hum. Genet.* **91**, 935–941 (2012).
- Fortin, C. F., Ear, T. & McDonald, P. P. Autocrine role of endogenous interleukin-18 on inflammatory cytokine generation by human neutrophils. *FASEB J.* **23**, 194–203 (2009).
- Gregersen, P. K. Cell type-specific eQTLs in the human immune system. *Nat. Genet.* **44**, 478–480 (2012).
- Flutre, T., Wen, X., Pritchard, J. & Stephens, M. A statistical framework for joint eQTL analysis in multiple tissues. *PLoS. Genet.* **9**, e1003486 (2013).
- Westra, H. J. *et al.* Cell specific eQTL analysis without sorting cells. *PLoS. Genet.* **11**, e1005223 (2015).
- Battle, A. *et al.* Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Res.* **24** (2013).
- Fairfax, B. P. *et al.* Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* **343**, 1246949 (2014).
- Kim, S. *et al.* Characterizing the genetic basis of innate immune response in TLR4-activated human monocytes. *Nat. Commun.* **5**, 5236 (2014).
- Lee, M. N. *et al.* Common Genetic Variants Modulate Pathogen-Sensing Responses in Human Dendritic Cells. *Science* **343** (2014).
- Ye, C. J. *et al.* Intersection of population variation and autoimmunity genetics in human T cell activation. *Science* **345** (2014).
- Martin, P. & Leibovich, S. J. Inflammatory cells during wound repair: the good, the bad and the ugly. *Trends. Cell. Biol.* **15**, 599–607 (2005).
- Gaspari, A. A. Innate and adaptive immunity and the pathophysiology of psoriasis. *J. Am. Acad. Dermatol.* **54**, S67–S80 (2006).
- Nomura, I. *et al.* Cytokine milieu of atopic dermatitis, as compared to psoriasis, skin prevents induction of innate immune response genes. *J. Immunol.* **171**, 3262–3269 (2003).
- Terui, T., Ozawa, M. & Tagami, H. Role of neutrophils in induction of acute inflammation in T-cell-mediated immune dermatosis, psoriasis: a neutrophil-associated inflammation-boosting loop. *Exp. Dermatol.* **9**, 1–10 (2000).
- Swindell, W. *et al.* Cellular dissection of psoriasis for transcriptome analyses and the post-GWAS era. *BMC. Med. Genomics.* **7**, 27 (2014).
- Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic. Acids. Res.* **42**, D1001–D1006 (2014).
- Hedl, M., Zheng, S. & Abraham, C. The IL18RAP region disease polymorphism decreases IL-18RAP/IL-18R1/IL-1R1 expression and signaling through innate receptor-initiated pathways. *J. Immunol.* **192**, 5924–5932 (2014).
- Mocsai, A. Diverse novel functions of neutrophils in immunity, inflammation, and beyond. *J. Exp. Med.* **210**, 1283–1299 (2013).
- Abram, C. L., Roberge, G. L., Pao, L. I., Neel, B. G. & Lowell, C. A. Distinct roles for neutrophils and dendritic cells in inflammation and autoimmunity in motheaten mice. *Immunity* **38**, 489–501 (2013).
- Nemeth, T. & Mocsai, A. The role of neutrophils in autoimmune diseases. *Immunol. Lett.* **143**, 9–19 (2012).
- Abecasis, G. R. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
- Hinrichs, A. S. *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic. Acids. Res.* **34**, D590–D598 (2006).
- Stranger, B. E. *et al.* Population genomics of human gene expression. *Nat. Genet.* **39**, 1217–1224 (2007).
- Churchill, G. A. & Doerge, R. W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971 (1994).

## Acknowledgements

The authors would like to thank all the volunteers and their family members who participated in this study. We would like to thank Crystal Tan Tze Ying and Tang Suisheng from the Singapore Immunology Network for sample preparation, and Lucy Robinson (Insight Editing London) for commercial assistance in editing of the manuscript.

## Author contributions

O.R. conceived and designed the study while M.P., F.Z. and A.L. were in charge of the sample processing and bioinformatics analysis. A.K.A., R.M. performed analysis and wrote the manuscript. T.Y.P., M.N., K.J.P., E.V., D.H., N.Y. and B.S.L. performed the functional experiments. A.V., J.L., D.K., S.F. and L.Z. performed the neutrophil isolation, RNA extraction and microarray experiments. A.I., B.L., H.L., F.Z. and J.L. helped with analysis from the GWAS data set association data reported. W.D.Y., A.N. and A.M. performed the sample collection and blood processing. All authors have read the manuscript and help with the editing of the manuscript.



## Additional information

**Accession Codes:** All microarray data are deposited into Gene Expression Omnibus at the National Center for Biotechnology Information under the accession number GSE70044.

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Andiappan, A. K. *et al.* Genome-wide analysis of the genetic regulation of gene expression in human neutrophils. *Nat. Commun.* 6:7971 doi: 10.1038/ncomms8971 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>