


RESEARCH ARTICLE

Open Access



# Genome-wide pathway analysis identifies VEGF pathway association with oral ulceration in systemic lupus erythematosus

Adrià Aterido<sup>1,2</sup>, Antonio Julià<sup>1\*</sup> , Patricia Carreira<sup>3</sup>, Ricardo Blanco<sup>4</sup>, José Javier López-Longo<sup>5</sup>, José Javier Pérez Venegas<sup>6</sup>, Àlex Olivé<sup>7</sup>, José Luís Andreu<sup>8</sup>, María Ángeles Aguirre-Zamorano<sup>9</sup>, Paloma Vela<sup>10</sup>, Joan M. Nolla<sup>11</sup>, José Luís Marenco-de la Fuente<sup>12</sup>, Antonio Zea<sup>13</sup>, José María Pego<sup>14</sup>, Mercedes Freire<sup>15</sup>, Elvira Díez<sup>16</sup>, María López-Lasanta<sup>1</sup>, Mireia López-Corbeto<sup>1</sup>, Núria Palau<sup>1</sup>, Raül Tortosa<sup>1</sup>, Josep Lluís Gelpi<sup>17</sup>, Devin Absher<sup>18</sup>, Richard M Myers<sup>18</sup>, Antonio Fernández-Nebro<sup>19,20\*</sup> and Sara Marsal<sup>1</sup>

## Abstract

**Background:** Systemic lupus erythematosus (SLE) is a genetically complex rheumatic disease characterized by heterogeneous clinical manifestations of unknown etiology. Recent studies have suggested the existence of a genetic basis for SLE heterogeneity. The objective of the present study was to identify new genetic variation associated with the clinically relevant phenotypes in SLE.

**Methods:** A two-stage pathway-based approach was used to identify the genetic variation associated with the main clinical phenotypes in SLE. In the discovery stage, 482 SLE patients were genotyped using Illumina Human Quad610 microarrays. Association between 798 reference genetic pathways from the Molecular Signatures Database and 11 SLE phenotypes was tested using the set-based method implemented in PLINK software. Pathways significantly associated after multiple test correction were subsequently tested for replication in an independent cohort of 425 SLE patients. Using an *in silico* approach, we analyzed the functional effects of common SLE therapies on the replicated genetic pathways. The association of known SLE risk variants with the development of the clinical phenotypes was also analyzed.

**Results:** In the discovery stage, we found a significant association between the vascular endothelial growth factor (VEGF) pathway and oral ulceration ( $P$  value for false discovery rate ( $P_{FDR}$ ) < 0.05), and between the negative regulation signaling pathway of retinoic acid inducible gene-1/melanoma differentiation associated gene 5 and the production of antinuclear antibodies ( $P_{FDR}$  < 0.05). In the replication stage, we validated the association between the VEGF pathway and oral ulceration. Therapies commonly used to treat mucocutaneous phenotypes in SLE were found to strongly influence VEGF pathway gene expression ( $P = 4.60e-4$  to  $5.38e-14$ ). Analysis of known SLE risk loci identified a strong association between *PTPN22* and the risk of hematologic disorder and with the development of antinuclear antibodies.

**Conclusions:** The present study has identified VEGF genetic pathway association with the risk of oral ulceration in SLE. New therapies targeting the VEGF pathway could be more effective in reducing the severity of this phenotype. These findings represent a first step towards the understanding of the genetic basis of phenotype heterogeneity in SLE.

**Keywords:** Genome-wide pathway analysis, Genetics, Systemic lupus erythematosus, Oral ulceration, Clinical phenotypes

\* Correspondence: tonijulia@vhir.org; afnebro@gmail.com

<sup>1</sup>Rheumatology Research Group, Vall d'Hebron Research Institute, Barcelona 08035, Spain

<sup>19</sup>Rheumatology Department, Hospital Regional Universitario de Málaga, Málaga 29010, Spain

Full list of author information is available at the end of the article



## Background

Systemic lupus erythematosus (SLE) is a disabling multi-system rheumatic disease with substantial epidemiological variation [1]. SLE is characterized by the dysregulation of the immune system and high phenotypical diversity [2]. This phenotypical heterogeneity includes a wide range of clinical manifestations that are exemplified by the current use of multiple clinical phenotypes as criteria to diagnose the disease [3]. So far, however, little is known about the causes of this phenotypic variation. Understanding the molecular mechanisms associated with the pathogenesis of SLE phenotypes could therefore be of high relevance to develop more efficient therapeutic approaches and preventive strategies.

SLE is characterized by a strong genetic component, with a sibling recurrence rate ( $\lambda_s$ ) of 8–29 and estimated heritability of approximately 66% [4]. To date, nine genome-wide association studies (GWAS) of SLE risk have been performed in European and Asian populations [5]. Together these studies have led to the identification of >40 loci associated with SLE susceptibility. Despite this extraordinary success, there is still a lack of understanding of the genetic variation that is relevant for the development of specific phenotypes within the disease. There is evidence, however, that the main clinical phenotypes in SLE aggregate in families [6], suggesting a genetic basis underlying disease heterogeneity.

To date, only a few candidate gene studies have been performed to uncover the genetics of clinical heterogeneity in SLE [7]. These studies have identified immunity-related genes associated with clinically relevant SLE phenotypes [8]. From these, the most significant associations have been detected between renal disorder and genetic variation in the *ITGAM* and *STAT4* genes, which have been also associated with discoid rash and oral ulceration, respectively [9, 10]. Other significant findings include the association between renal disorder and *TNFSF4*, malar rash and *FCGR2A* and hematological disorder and variation in the *IL21* gene. So far, however, the genetic component for most SLE phenotypes has been only partially explained. Therefore, the analysis of genetic variation at a genome-wide scale is needed to identify additional variation associated with SLE clinical heterogeneity.

Complex traits like disease risk or clinical phenotypes have been shown to be caused by multiple genes of small effect size [11]. The identification of these small-effect genes is currently one of the major challenges in the characterization of the genetic background for disease phenotypes [12]. Importantly, single-marker GWAS do not allow the identification of genetic variants with small effect sizes, unless extremely large sample sizes are used [13]. In this common type of GWAS, a large number of markers are tested for association and, consequently, stringent significance thresholds are applied, which makes

the identification of small-effect variants very difficult [14]. In addition, single-marker GWAS ignore the joint contribution of multiple genes that act coordinately in the same biological process [15]. The characterization of the genetic basis of many complex traits will therefore require the development of new powerful methods that are able to leverage biological knowledge and efficiently integrate the evidence from multiple loci with moderate to small effect sizes.

Recently, novel statistical methodologies that are able to test genetic risk associations at the pathway level have been developed [16]. Pathway-based approaches test whether sets of functionally related genes are jointly associated with a particular phenotype [17]. This methodology strongly reduces the number of association tests and, therefore, it can substantially increase the power to identify new genetic variation compared to single-marker GWAS [18]. The genome-wide pathway analysis (GWPA) has been recently used to characterize the genetic basis of several complex diseases like cancer [19]. Very recently, using the GWPA approach we have identified new genetic variation associated with psoriasis, an autoimmune disease of the skin [20]. This result confirms the utility of GWPA in the study of the genetics of autoimmune diseases.

To gain a better understanding of the genetic basis underlying phenotype heterogeneity in SLE we have performed, for the first time, a GWAS of clinical phenotypes using the GWPA approach. In this study we have analyzed a discovery cohort of 482 SLE patients of European ancestry to determine the association between 798 reference biological pathways and the main clinical phenotypes of SLE represented that are used as diagnosis criteria. Using an independent cohort of 425 SLE patients from the same ancestry, we have then performed a validation study of the most significant genetic pathways. Based on these results, we have performed an *in silico* validation analysis to evaluate the functional impact of drugs commonly used to treat the associated phenotype. Our findings provide new insights into the biological mechanisms associated with clinical phenotypes of SLE.

## Methods

### Study population

In the discovery stage, a total of 482 SLE patients were recruited. SLE patients were collected from the out-patient clinics of the rheumatology departments of 15 Spanish University Hospitals belonging to the Immune-Mediated Inflammatory Disease (IMID) Consortium [21]. All patients were diagnosed by a rheumatologist. Only those patients with SLE that fulfilled  $\geq 4$  of the 1982 revised American College of Rheumatology (ACR) diagnosis criteria were included in the present study [3]. All patients included in this study were >16 years old at

the time of sample collection and had >3 years of evolution from the diagnosis date. SLE patients with psoriasis, inflammatory bowel disease (Crohn's disease or ulcerative colitis) or other rheumatic diseases like rheumatoid arthritis, or multiple sclerosis were excluded from the study. All SLE patients were Caucasian European with all four grandparents born in Spain.

In the validation stage, an independent cohort of 425 SLE patients was used to replicate the genetic pathways that were significantly associated with the SLE phenotypes in the discovery stage. All patients from the validation cohort fulfilled the ACR diagnostic criteria for SLE and were also collected from the IMID Consortium, following the same inclusion and exclusion criteria as for the discovery cohort. All the procedures were followed in compliance with the principles of the Declaration of Helsinki.

The main epidemiological and clinical variables of the discovery and validation cohorts are summarized in Table 1. The distribution of each variable was compared between the discovery and validation cohorts using Fisher's exact test or Student's *t* test for categorical and quantitative variables, respectively.

### SLE phenotypes

The diagnosis of SLE is of major importance to guide both the disease classification and the patient therapy [22]. Given the high phenotypic heterogeneity of SLE, in order to analyze the most relevant clinical manifestations for disease diagnosis we defined the SLE phenotypes

according to the established ACR diagnostic criteria for SLE [3]. Consequently, the 11 SLE phenotypes represented by the ACR diagnostic criteria were analyzed using the GWPA approach. These criteria include malar rash, discoid rash, photosensitivity, oral ulcers, arthritis, serositis, renal disorder, neurologic disorder, hematologic disorder, immunologic disorder and antinuclear antibodies. The distribution of each clinical phenotype in the discovery and replication cohorts is shown in Table 1.

### DNA extraction and genome-wide genotyping in the discovery and validation patients

In the discovery stage, the genome-wide genotyping of the 482 SLE patients was performed using the Illumina Quad610 Beadchips (Illumina, San Diego, CA, USA) at the Centro Nacional de Genotipado (CeGen, Madrid, Spain). The genotyping quality control analysis was performed using PLINK software (Additional file 1: Figure S1) [23]. To evaluate the presence of potential population stratification in the SLE patient cohorts, we used the principal component analysis (PCA) implemented in the EIGENSOFT (v4.2) software [24]. Using the first 10 PCs of variation over 10 iterations we identified 14 samples showing an outlier genetic background and were excluded from downstream analysis (Additional file 1: Figure S1). After the quality control analysis, a final dataset of 507,051 single nucleotide polymorphisms (SNPs) and 395 SLE patients was available for the GWPA.

The validation of the two genetic pathways associated with SLE in the discovery stage required the genotyping

**Table 1** Main epidemiological and clinical features of the discovery and validation patient cohorts

Clinical and epidemiological variables <sup>a</sup>	Discovery cohort	Validation cohort	<i>P</i>
Genetic background	CEU	CEU	-
Number of individuals ( <i>N</i> )	482	425	-
Gender (% females)	443 (91.91%)	398 (93.65%)	0.37
Average age at onset ( <i>m</i> ± <i>SD</i> )	33.50 ± 14.09	32.39 ± 13.32	0.22
Malar rash ( <i>N</i> <sup>+</sup> / <i>N</i> )	185/174	183/175	1.00
Discoid rash ( <i>N</i> <sup>+</sup> / <i>N</i> )	46/313	49/309	0.74
Photosensitivity ( <i>N</i> <sup>+</sup> / <i>N</i> )	191/168	173/185	0.20
Oral ulcers ( <i>N</i> <sup>+</sup> / <i>N</i> )	153/206	172/186	0.15
Arthritis ( <i>N</i> <sup>+</sup> / <i>N</i> )	275/84	267/91	0.54
Serositis ( <i>N</i> <sup>+</sup> / <i>N</i> )	110/249	122/236	0.34
Renal disorder ( <i>N</i> <sup>+</sup> / <i>N</i> )	120/239	108/250	0.38
Neurologic disorder ( <i>N</i> <sup>+</sup> / <i>N</i> )	25/334	35/293	0.10
Hematologic disorder ( <i>N</i> <sup>+</sup> / <i>N</i> )	341/18	337/21	0.63
Immunologic disorder ( <i>N</i> <sup>+</sup> / <i>N</i> )	289/70	292/66	0.77
Antinuclear antibodies ( <i>N</i> <sup>+</sup> / <i>N</i> )	346/13	347/11	0.84

**Abbreviations:** CEU Caucasian European, *M* mean, *N* sample size, *N*<sup>+</sup> sample size of positive individuals for the clinical variable, *N*<sup>-</sup> sample size of negative individuals for the clinical variable, *P* *P* value, *SD* standard deviation

<sup>a</sup>Number of individuals shown in the table represents the total number of patients that were initially recruited for the present study. Conversely, the sample size of positive/negative individuals for the indicated clinical phenotype represents the final number of SLE patients having data on both genotype and phenotype available for association analysis

and analysis of a total 1347 SNPs. Given the large number of variants to be tested and the utility of genome-wide data for accurate genetic ancestry identification, the 425 SLE patients in the validation cohort were genotyped using the same microarray platform. Genotyping for the validation stage was performed at the HudsonAlpha Institute for Biotechnology (Huntsville, AL, USA). The same quality control analysis as in the discovery stage was performed (Additional file 1). A total of 394 SLE patients and all 1347 SNPs from the two genetic pathways passed the quality control and were available for the pathway-based analysis of the validation stage.

### Analysis of association between established SLE risk SNPs and SLE phenotypes

#### *Genetic variants associated with SLE risk*

The list of established genetic variants ( $P < 5e-8$ ) for SLE risk was obtained from a recent GWAS meta-analysis in a case-control cohort of European ancestry [5]. A total of 43 genetic variants associated with SLE risk were identified and selected for the analysis of association with SLE clinical phenotypes (Additional file 1: Table S6).

#### *Imputation of genetic variants associated with SLE risk*

From the established autosomal SLE risk SNPs ( $N = 41$  SNPs), the genetic variants that were not directly genotyped by the GWAS Quad610 genotyping array ( $N = 17$  SNPs) were imputed (Additional file 1). Those SNPs that did not pass the stringent imputation quality control filter ( $N = 1$  SNP, information quality metric  $< 0.8$ ) were excluded from the study. Therefore, after excluding two non-autosomal variants and a low-quality imputed SNP, a total of 40 from the initial 43 established SLE risk SNPs were finally available for analysis of association with SLE phenotypes in the discovery cohort. In the validation cohort, the same procedure was followed to obtain the genotypes of the risk variants to be tested for replication.

#### *Statistical association analysis*

The statistical association analysis between the allele dosage of the established SLE risk SNPs and the SLE clinical phenotypes was performed using the logistic regression model implemented in the SNPTEST v2 software (Oxford, UK) [25]. In this model, the allele dosage was defined as follows:

$$\text{Allele Dosage}_i = \sum_{g=0}^2 \Pr(G = g) * g$$

Where  $g$  represents each genotype of a particular genetic variant  $i$  and  $\Pr(G = i)$  is the marginal posterior probability obtained by imputation. The allele dosage takes values between 0 and 2. SLE patients without phenotypical data available for the phenotype analyzed

were excluded from the association analysis. Finally, the  $P$  values obtained from the discovery and replication stages were combined using the METAL software [26].

### GWPA

#### *GWPA method*

The gene-set analysis, also referred to as pathway analysis, is a very powerful methodology to analyze the genetic architecture of complex diseases using GWAS data [27, 28]. An important advantage of this approach is that the hypothesis space is significantly reduced compared to single-marker GWAS. While extremely large number of markers ( $> 500,000$  to several millions) are independently tested for association in single-marker GWAS, the number of simultaneous tests is several orders of magnitude lower in pathway-based GWAS (typical range 500–2000 pathways). Consequently, the threshold for significance is much less stringent than the consensus threshold used for single-marker GWAS ( $P < 5e-8$ ) [18, 27]. In addition, pathway-based studies integrate the effects of multiple genetic risk variants that participate in the same biological processes. For these reasons, pathway-based studies can have high statistical power to discover new susceptibility genetic variants provided that they operate within the analyzed pathways.

In the present study, the GWPA was performed using the set-based test implemented in the PLINK software as described previously [20, 23]. Compared to other methods, this set-based test uses genotype data to estimate pathway association instead of  $P$  values for significance. Importantly, this approach accounts for the linkage disequilibrium between SNPs and therefore avoids an increase in false positive results due to genes with multiple, highly correlated markers. For each pathway, independent SNPs are first identified (linkage disequilibrium of  $r^2 < 0.2$  here), and from these an average statistic is calculated. Finally, the statistical significance of the pathway is computed using permutation, thereby efficiently correcting by the number of SNPs within the pathway (Additional file 1). As described for the analysis of association with established SLE risk SNPs, those patients with missing data for the phenotype tested for association were excluded from the analysis. In order to account for multiple testing, the false discovery rate (FDR) method was used. The corrected empirical  $P$  values from the discovery and validation stages were combined using Fisher's method.

#### *Gene set definition*

Reference biological pathway annotation databases BioCarta, Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome were used for the present study [29]. A total of 217, 186 and 674 curated biological pathways from the Biocarta, KEGG and Reactome databases, respectively, were included, respectively (5th October

2015). Very small uninformative pathways (i.e.  $\leq 15$  genes) were excluded from the analysis. As described previously, we also excluded large genetic pathways (i.e.  $> 300$  genes) [17]. The SNP-gene mapping was performed using the NCBI RefSeq database release 63 (12th October 2015) and an SNP-gene distance window of 20 Kb [30]. The final gene set used for the present GWPA was composed of 211,724 SNPs mapping to 798 different pathways. The list of genetic pathways included in the GWPA is shown in Additional file 1: Table S7.

#### ***In silico* analysis of VEGF pathway genes after treatment with topical immunotherapies for cutaneous SLE**

In the GWPA, we identified significant genetic association between oral ulceration and the VEGF pathway. The VEGF pathway plays a crucial role in angiogenesis, and there is increasing evidence supporting the implication of this biological process in the pathogenesis of SLE cutaneous phenotypes [31]. These disease phenotypes are commonly treated with steroid and non-steroid topical immunotherapies in SLE [32, 33]. Consequently, we hypothesized that the topical immunotherapies prescribed for cutaneous SLE mediate their therapeutic effect in this tissue through the VEGF pathway and, therefore, should induce significant transcriptional changes in the pathway genes. In order to test this hypothesis, we used transcriptional data from microarray experiments in the NCBI Gene Expression Omnibus microarray database (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). In this database, we searched for whole genome expression profiling datasets generated from cutaneous/mucocutaneous human samples or cell cultures (5 November 2015). From these, we looked for tissue or cell cultures treated with any of the common steroid and non-steroid topical immunotherapies most widely used in SLE (Additional file 1). We found a total of three datasets analyzing the transcriptional variation after treatment with four common immunotherapies: betamethasone valerate and pimecrolimus (GSE32473), diphencyprone (GSE52360) and imiquimod (GSE68182). The first two transcriptional datasets (i.e. GSE32473 and GSE52360) were obtained from skin biopsies and the latter (GSE68182) from an *in vitro* study on vaginal mucosal cells (i.e. cell line Vk2/E6E7). For each gene expression dataset, we performed quality control analysis and subsequent normalization on the log<sub>2</sub> scale using the quantile normalization method. The analysis of differential expression of the VEGF pathway genes between treated and non-treated samples was performed using Student's *t* test. The statistical significance of the global perturbation of the VEGF pathway was assessed using the binomial test. All analyses were performed using the R statistical software [34].

## **Results**

### **Phenotypic characterization of the studied cohorts**

The epidemiological and phenotypical characteristics of the discovery and replication SLE populations are shown in Table 1. There were no significant differences between the discovery and replication cohorts in the distribution of the epidemiological and phenotypical variables ( $P > 0.05$ , Table 1).

### **Identification of SLE risk genetic variants associated with SLE phenotypes**

In the discovery stage, we found that 19 out of the 43 SNPs previously associated with SLE risk were also significantly associated with one or more clinical phenotypes ( $P_{\text{Discovery}} < 0.05$ , Table 2). The association results between each established genetic variant and SLE phenotype are summarized in Additional file 1: Table S1. However, in the independent validation cohort, only the association between *PTPN22* and hematologic disorder ( $P_{\text{Replication}} = 0.043$ ,  $P_{\text{Combined}} = 8.25e-4$ ) and between *PTPN22* and the production of antinuclear antibodies ( $P_{\text{Replication}} = 0.028$ ,  $P_{\text{Combined}} = 0.001$ ) were significantly replicated. Combining the statistical evidence from the two cohorts, an additional seven loci were found to be associated with SLE phenotypes at the nominal level ( $P_{\text{Combined}} < 0.05$ , Table 2).

### **Identification of genetic pathways associated with SLE phenotypes**

In the GWPA, two genetic pathways were found to be significantly associated with an SLE phenotype after multiple test correction ( $P$  value for false discovery rate ( $P_{\text{FDR}} < 0.05$ , Table 3). The VEGF pathway was associated with the presence of oral ulcers ( $P_{\text{FDR}} = 0.044$ ) and the RIG-I/MDA5 negative regulation signaling pathway was associated with the production of antinuclear antibodies ( $P_{\text{FDR}} = 0.016$ ). The results for analysis of association between each genetic pathway and SLE phenotype are shown in Additional file 1: Table S2.

Using the independent validation cohort, the association between the VEGF genetic pathway and oral ulcers in SLE was significantly replicated ( $P_{\text{FDR}} = 0.026$ , Table 3). The details of association between the VEGF pathway and oral ulcers are shown in Additional file 1: Table S3.

### **Perturbation of the VEGF genetic pathway by current therapies for cutaneous SLE**

Topical immunotherapies are drugs commonly used to treat cutaneous phenotypes of SLE like oral ulceration. Given the observed genetic association between oral ulcers and VEGF pathway, we performed an *in silico* analysis to evaluate the effect of four current topical immunotherapies on the expression of its constituent genes. Using whole genome expression datasets from

**Table 2** SLE risk SNPs association with clinical phenotypes

SNP	Chr	Pos	Gene	RA	Phenotype	$P_D$	$P_V$	$P_C$	OR (CI, 95%)
<b>rs2476601<sup>a</sup></b>	<b>1</b>	<b>114377568</b>	<b>PTPN22</b>	<b>A</b>	<b>Hematologic disorder</b>	<b>0.0039</b>	<b>0.0433</b>	<b>5.20e-4</b>	<b>9.488 (1.310-68.726)</b>
					<b>Antinuclear antibodies</b>	<b>0.0146</b>	<b>0.0281</b>	<b>0.0011</b>	<b>0.168 (0.004-0.990)</b>
rs704840	1	173226195	TNFSF4	G	Renal disorder	0.0407	0.9210	0.1290	1.199 (0.941-1.527)
<b>rs3024505</b>	<b>1</b>	<b>206939904</b>	<b>IL10</b>	<b>T</b>	<b>Arthritis</b>	<b>0.0325</b>	<b>0.1540</b>	<b>0.0170</b>	<b>1.557 (1.091-2.221)</b>
<b>rs3768792</b>	<b>2</b>	<b>213871709</b>	<b>IKZF2</b>	<b>C</b>	Photosensitivity	0.0437	0.3320	0.4580	0.893 (0.674-1.184)
					<b>Hematologic disorder</b>	<b>0.0319</b>	<b>0.1880</b>	<b>0.0143</b>	<b>0.507 (0.301-0.855)</b>
rs9311676	3	58470351	ABHD6,PXK	C	Oral ulcers	0.0338	0.7960	0.9210	0.832 (0.674-1.028)
rs564799	3	159728987	IL12A	C	Discoid rash	0.0411	0.1410	0.6853	0.942 (0.674-1.316)
<b>rs1270942</b>	<b>6</b>	<b>31918860</b>	<b>MHC class III</b>	<b>C</b>	<b>Arthritis</b>	<b>0.0138</b>	<b>0.6670</b>	<b>0.0407</b>	<b>1.530 (1.010-2.346)</b>
<b>rs9462027</b>	<b>6</b>	<b>34797241</b>	<b>UHRF1BP1</b>	<b>A</b>	<b>Malar rash</b>	<b>0.0263</b>	<b>0.1730</b>	<b>0.0137</b>	<b>1.333 (1.071-1.660)</b>
					Oral ulcers	0.0251	0.8950	0.0933	1.210 (0.971-1.508)
rs6932056	6	138242437	TNFAIP3	C	Serositis	0.0193	0.8140	0.1364	1.580 (0.898-2.781)
rs4917014	7	50305863	IKZF1	T	Oral ulcers	0.0280	0.1550	0.5821	0.944 (0.753-1.184)
rs2663052	10	50069395	WDFY4	C	Neurologic disorder	0.0370	0.1590	0.6311	0.963 (0.662-1.401)
rs4948496	10	63805617	ARID5B	C	Malar rash	0.0331	0.4580	0.3258	1.116 (0.906-1.375)
					Photosensitivity	0.0202	0.9480	0.0911	1.203 (0.977-1.483)
rs2732549	11	35088399	CD44	T	Immunologic disorder	0.0163	0.9910	0.0907	1.277 (0.976-1.670)
<b>rs7941765</b>	<b>11</b>	<b>128499000</b>	<b>ETS1,FLI1</b>	<b>C</b>	<b>Arthritis</b>	<b>0.0072</b>	<b>0.0701</b>	<b>0.0015</b>	<b>1.470 (1.151-1.877)</b>
<b>rs10774625</b>	<b>12</b>	<b>111910219</b>	<b>SH2B3</b>	<b>A</b>	<b>Serositis</b>	<b>0.0207</b>	<b>0.0756</b>	<b>0.0038</b>	<b>0.715 (0.571-0.895)</b>
					Renal disorder	0.0423	0.6410	0.0773	0.814 (0.650-1.018)
					<b>Immunologic disorder</b>	<b>0.0051</b>	<b>0.8190</b>	<b>0.0321</b>	<b>0.738 (0.566-0.964)</b>
<b>rs1059312</b>	<b>12</b>	<b>129278864</b>	<b>SLC15A4</b>	<b>C</b>	<b>Serositis</b>	<b>0.0253</b>	<b>0.2370</b>	<b>0.0156</b>	<b>1.321 (1.053-1.658)</b>
rs4902562	14	68731458	RAD51B	A	Renal disorder	0.0130	0.8670	0.1010	0.834 (0.664-1.047)
rs2941509	17	37921194	IKZF3	A	Antinuclear Antibodies	0.0304	0.0413	0.4943	0.676 (0.237-1.930)
rs2304256	19	10475652	TYK2	C	Arthritis	0.0298	0.7690	0.1841	0.826 (0.622-1.096)

A total of 19 SLE risk genetic variants were significantly associated with different systemic lupus erythematosus (SLE) phenotypes in the discovery stage. From these, the association between *PTPN22* and hematologic disorder and between *PTPN22* and the production of antinuclear antibodies (\*) were significantly replicated in the validation cohort. Combining the statistical evidence from the two cohorts, seven additional genetic variants were found to be nominally associated with SLE phenotypes (shown in bold). *Abbreviations:* SNP single-nucleotide polymorphism, Chr chromosome, Pos SNP base pair in build GRCh37/hg19, RA disease risk allele,  $P_D$   $P$  value discovery cohort,  $P_V$   $P$  value validation cohort,  $P_C$   $P$  value combined, OR odds ratio according to the allele associated with disease risk, CI confidence interval (95%)

patients and relevant cell types treated with these therapies, we found that three out of the four analyzed drugs significantly perturb the expression of VEGF pathway genes (Fig. 1, Additional file 1: Table S4). A total of 16, 12 and 7 genes out of the 29 genes from the VEGF pathway were significantly differentially expressed after imiquimod ( $P = 5.38e-14$ ), betamethasone valerate

( $P = 5.69e-9$ ) and diphencyprone ( $P = 4.59e-4$ ) treatment, respectively (Additional file 1: Table S5).

**Discussion**

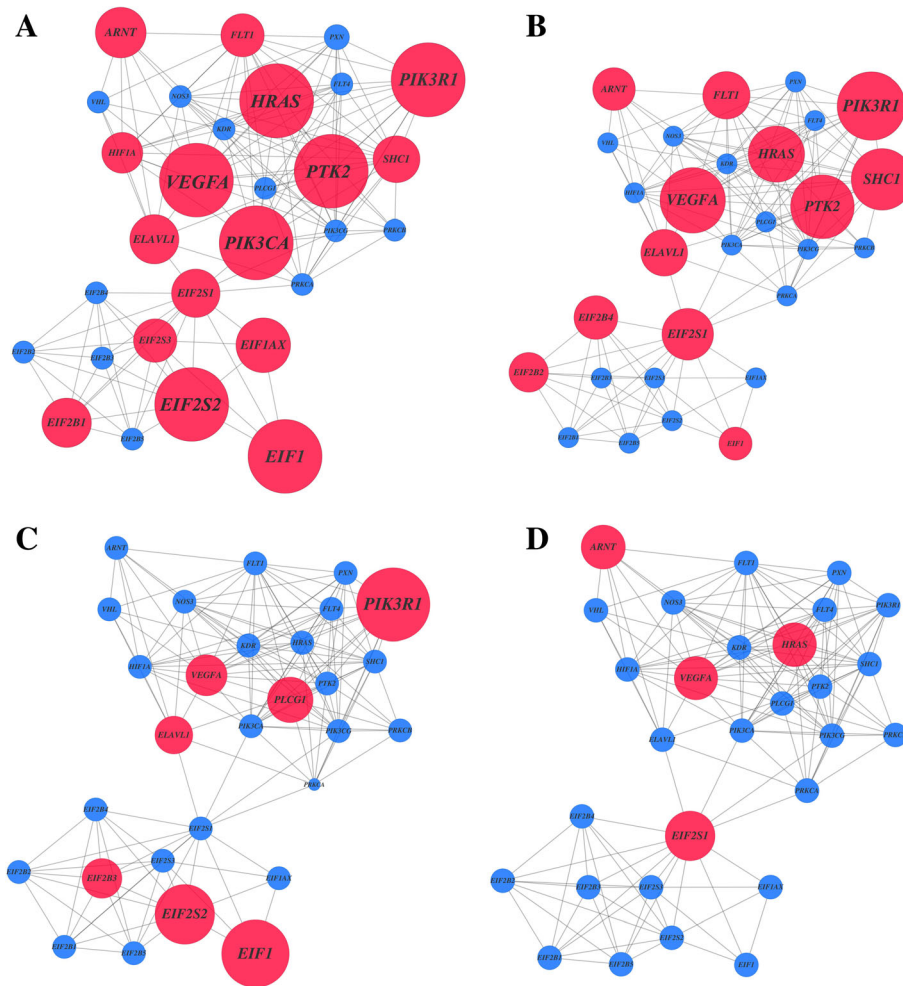
One of the major challenges in the pathogenesis of SLE is to understand the biological mechanisms responsible for its phenotypic heterogeneity. Although single-marker

**Table 3** Genetic pathways associated with the SLE phenotypes in the discovery stage

Pathway	Phenotype	#Genes	SNPs	$P_D$	FDR <sub>D</sub>	$P_V$	FDR <sub>V</sub>	$P_C$
RIG-I/MDA5 negative regulation signaling	Antinuclear antibodies	31	464	2.00e-5	0.016	0.234	0.234	-
VEGF <sup>a</sup>	Oral ulcers	29	883	5.58e-5	0.044	0.013	0.026	2.08e-5

Two genetic pathways were significantly associated with systemic lupus erythematosus (SLE) phenotypes in the discovery stage *Abbreviations:* SNP single-nucleotide polymorphism,  $P_D$   $P$  value discovery cohort, FDR<sub>D</sub> false discovery rate discovery cohort,  $P_V$   $P$  value validation cohort, FDR<sub>V</sub> false discovery rate validation cohort,  $P_C$   $P$  value combined

<sup>a</sup>The association between the vascular endothelial growth factor (VEGF) pathway and oral ulcers was significantly replicated in the validation cohort



**Fig. 1** Vascular endothelial growth factor (VEGF) pathway perturbation after topical immunotherapy. Network representation of VEGF genes according to the differential gene expression after treatment with common topical immunotherapies: imiquimod (**a**), betamethasone valerate (**b**), diphencyprone (**c**) and pimecrolimus (**d**). Genes are represented as nodes and are connected by edges according to experimental or computational evidence of interaction between their encoded proteins. The diameter of each node is proportional to the significance of differential expression, with significant genes ( $P < 0.05$ ) in red and non-significant genes ( $P \geq 0.05$ ) in blue

GWAS have successfully identified a large number of genetic variants associated with SLE risk, the genetic basis of SLE phenotypes has so far been analyzed only in a few candidate gene studies. In order to identify new genetic variation, we have here performed the first GWAS on SLE phenotypes using a pathway-based approach. Using a discovery cohort of individuals with European ancestry and a validation cohort with the same ancestry, we have identified and validated the association between VEGF genetic pathway and oral ulcers, a common manifestation of SLE. The results of this study provide new insights into the genetic basis and the biological mechanisms associated with the clinical heterogeneity of SLE.

The VEGF pathway is a network of genes that are involved in the transduction of different intracellular

signals and act coordinately to modulate inflammatory and angiogenic processes [35, 36]. The dysregulation of angiogenesis has been described as an important biological mechanism in the pathogenesis of SLE [37]. In particular, there is growing evidence that angiogenesis is also involved in the development of skin manifestations in SLE patients [38, 39]. The serum levels of VEGFA protein itself have been suggested as a useful marker for disease activity monitoring in SLE patients [40]. Importantly, the serum levels of VEGFA have also been found to be significantly elevated in SLE patients with cutaneous manifestations [41]. Despite this clinical evidence, the VEGF pathway had not been previously associated with SLE at the genetic level. Our study provides the first evidence of a genetic association between the VEGF pathway and oral ulceration in SLE.

Oral ulcers are frequently chronic mucocutaneous lesions that affect up to 54% of patients with SLE [42, 43]. This clinical manifestation is characterized by high angiogenic activity and a loss of the epithelial and connective tissue in the oral mucosa [44, 45]. Accordingly, anti-angiogenic therapies like thalidomide have been suggested to promote oral ulcer healing and to control ulcer recurrence [46, 47]. From a clinical perspective, oral ulceration has been associated with an increase in the disease activity and a worse prognosis in SLE [48, 49]. The early detection of oral ulcers is therefore highly valuable as it contributes to an earlier diagnosis of SLE and, consequently, to faster initiation of treatment.

The genetic association identified in this study is consistent with previous evidence from other ulcer-related diseases like Behçet disease (BD), recurrent aphthous ulceration (RAU) and gastroduodenal ulcers. BD is an inflammatory disorder characterized by an extremely high frequency of oral ulcers (>95%) [50]. Clinical evidence suggests that VEGF cytokine could be directly implicated in the formation of oral ulcers in BD [51]. In RAU, the most common oral mucosal disease, the salivary levels of VEGF have been also associated with oral ulceration [52]. Finally, genetic variation in the *VEGF* gene has been also associated with the risk of developing gastroduodenal ulcers [53]. Evidence from these studies implicates angiogenesis in the pathophysiological development of oral ulcers, at both the genetic and at the functional level. Accordingly, the genes in the VEGF pathway are strong candidates for susceptibility in diseases with a high prevalence of oral ulcers like BD or RAU. Future studies aimed at testing the association between VEGF pathway genes and these diseases are therefore warranted.

Topical steroid and non-steroid immunotherapies have been successfully used for the treatment of the cutaneous manifestations in SLE. However, topical immunotherapies are not exempt from side effects, including skin atrophy or telangiectasias with corticosteroids and the exacerbation of the inflammatory processes with non-steroid immunotherapies like imiquimod [32]. In this study, we have demonstrated that topical steroid and non-steroid immunomodulators significantly affect the expression of the VEGF pathway genes. The results of this *in silico* analysis indicate that the VEGF genetic pathway could be a key mediator of the benefits of topical immunotherapy to reduce oral ulceration. Therefore, our findings suggest that the VEGF pathway is a relevant source of new drug targets for oral ulceration that could allow more specific treatment while reducing the undesirable side effects of current therapies. In order to confirm the VEGF pathway as new source for drug discovery in the treatment of SLE oral ulceration, further prospective studies using oral ulcer samples from SLE patients are clearly needed.

In the present study, we have also identified and validated the association between SLE risk locus *PTPN22* and the production of antinuclear antibodies and with the presence of a hematologic disorder. It is the first time that this coding SNP (rs2476601) has been associated with the development of hematologic disease in SLE. A previous study reported a non-significant trend for association between *PTPN22* and antinuclear antibody positivity [54]. The results from our study provide the evidence to confirm the association between this susceptibility gene and the production of a common autoantibody in SLE. Also, the previously identified association between genetic variation in the *TNFSF4* gene and renal disorder was replicated in the discovery cohort [9]. Conversely, the reported association between renal disorder and *ITGAM* and *STAT4* genes was not replicated. The renal disorder phenotype encompasses different clinical manifestations (e.g. persistent proteinuria or cellular casts) and, therefore, differences in frequencies in any of these sub-phenotypes could have prevented the replication. Additional studies performing specific analysis of association for each renal disease subtype could help to further refine this genetic association. Oral ulceration was also previously found to be associated with variation in *STAT4*. This association was not replicated in the present single-marker analysis. The lack of replication could be explained by the comparatively smaller sample size of our study cohorts and by the small effect size reported for the association (OR ~ 1.12). For example, in the European cohort, *STAT4* association was not statistically significant despite analyzing >4,000 individuals [9]. Finally, after combining both patient cohorts, we also found seven other SLE risk loci - *IL10*, *IKZF2*, *MHC class III*, *UHRF1BP1*, *ETS1-FLI1*, *SH2B3* and *SLC15A4* - to be nominally associated with different phenotypes. These also represent new genetic associations with SLE heterogeneity. Further studies using independent cohorts of phenotypically well-characterized SLE patients like the present one are needed to corroborate these associations.

GWPA represents a new and powerful approach to identify genetic variation associated with complex phenotypes. However, this study has limitations. First, the sample size of the discovery and replication cohorts is moderate compared to recent case-control GWAS and this could have led to missed pathway associations with SLE clinical phenotypes. Second, the statistical power to detect significant associations is lower for those clinical phenotypes with more extreme frequencies (e.g. 7% of SLE patients have a neurologic disorder). Therefore, larger cohorts of well-characterized SLE patients will be needed to identify additional genetic variation associated with clinical phenotypes. Finally, the GWPA methodology also has intrinsic limitations, mainly related to the



current knowledge of biological processes and subsequent definition of the pathways [55]. For many human genes, the functional annotation is far from complete, which precludes their mapping to reference biological pathways. Also, genomic variation located far from the transcribed region itself could be relevant for the regulation of the gene expression. Conversely, variants within genes could be influencing the activity of other distant genes or genes in other chromosomes (i.e. trans-regulation). With the increase in knowledge of genomic regulation [56], the mapping of SNPs to their functionally related genes will clearly improve. Consequently, the integration of this knowledge into GWPA will likely increase the power of this approach to identify new pathways associated with human diseases.

## Conclusions

In conclusion, we have performed, for the first time, a genome-wide association analysis to identify genetic risk factors for the main phenotypes in SLE. To do this, we have used a pathway-based analysis approach. Using this approach, we have identified and validated the association between the VEGF genetic pathway and the presence of oral ulcers in SLE. These findings show the existence of a genetic basis underlying SLE heterogeneity that is independent from the genetic component associated with disease risk. The results of the present study could contribute to the development of more efficient therapies to treat cutaneous manifestations of SLE in the near future.

## Additional file

**Additional file 1:** Supplementary information is available in the online *Supplementary Material* file. (PDF 13965 kb)

## Abbreviations

ACR: American College of Rheumatology; BD: Behçet disease; CEU: Caucasian European; CHR: Chromosome; CI: Confidence interval; *ETS1*: ETS Proto-Oncogene 1 gene; FDR: False discovery rate; FDR<sub>D</sub>: False discovery rate discovery cohort; FDR<sub>V</sub>: False discovery rate validation cohort; *FLI1*: Fli-1 Proto-Oncogene gene; GEO: Gene Expression Omnibus; GWAS: Genome-wide association study; GWPA: Genome-wide pathway analysis; *IKZF2*: Zinc finger DNA binding protein helios, subfamily 1A, 2 gene; *IL10*: Interleukin-10 gene; IMID: Immune-mediated inflammatory disease; KEGG: Kyoto Encyclopedia of Genes and Genomes; M: Mean; *MDA5*: Melanoma differentiation-associated protein 5 gene; MHC: Major histocompatibility complex; N: Sample size; N<sup>-</sup>: Sample size of negative individuals for a particular clinical variable; N<sup>+</sup>: Sample size of positive individuals for a particular clinical variable; OR: Odds ratio; *P*: *P* value; *P*<sub>C</sub>: *P* value combined; *P*<sub>D</sub>: *P* value discovery cohort; POS: single nucleotide polymorphism base pair in build GRCh37/hg19; *PTPN22*: Protein tyrosine phosphatase, non-receptor type 22 gene; *P*<sub>V</sub>: *P* value validation cohort; RA: Disease risk allele; RAU: Recurrent aphthous ulceration; *RIG-I*: Retinoic acid-inducible gene 1 gene; SD: Standard deviation; *SH2B3*: Lymphocyte-specific adapter protein Lnk 3 gene; *SLC15A4*: Solute carrier family 15 member 4 gene; SLE: Systemic lupus erythematosus; SNP: Single nucleotide polymorphism; *UHRF1BP1*: Ubiquitin-like containing PHD and RING finger domains 1-binding protein 1 gene; VEGF: Vascular endothelial growth factor;  $\lambda_s$ : Sibling recurrence rate

## Acknowledgements

The authors thank the patients and clinical specialists collaborating in the IMID Consortium for participation.

## Funding

This study was funded by the Spanish Ministry of Economy and Competitiveness (grant numbers: PSE-010000-2006-6 and IPT-010000-2010-36) and by the "Agència de Gestió d'Ajuts Universitaris i de Recerca" (AGAUR, Generalitat de Catalunya, FI-DGR 2016, grant number: 00587). The study sponsor had no role in the collection, analysis or interpretation of the data.

## Availability of data and materials

Not applicable.

## Authors' contributions

Study design: AA, AJ, AFN and SM. Data acquisition: PC, RB, FJLL, JJPV, AO, JLA, MAAZ, PV, JMN, JLMF, AZ, JMP, MF, ED, MLL, MLC, JLG, AFN and SM. Statistical analysis and data interpretation: AA, AJ, AFN and SM. Sample processing: NP and RT. Genotyping: RMM and DA. Manuscript drafting: all authors. All authors read and approved the final manuscript.

## Authors' information

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

The study and the consent procedure were reviewed and approved by the Vall d'Hebron Hospital Ethics Committee. Informed consent was obtained from all participants.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Rheumatology Research Group, Vall d'Hebron Research Institute, Barcelona 08035, Spain. <sup>2</sup>Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona 08005, Spain. <sup>3</sup>Rheumatology Department, Hospital Universitario 12 de Octubre, Madrid 28041, Spain. <sup>4</sup>Rheumatology Department, Hospital Universitario Marqués de Valdecilla, Santander 39008, Spain. <sup>5</sup>Rheumatology Department, Hospital Universitario Gregorio Marañón, Madrid 28007, Spain. <sup>6</sup>Rheumatology Department, Hospital del SAS de Jerez de la Frontera, Cádiz 11407, Spain. <sup>7</sup>Rheumatology Department, Hospital Universitari Germans Trias i Pujol, Badalona 08916, Spain. <sup>8</sup>Rheumatology Department, Hospital Universitario Puerta de Hierro, Madrid 28222, Spain. <sup>9</sup>Rheumatology Department, Hospital Universitario Reina Sofía, Córdoba 14004, Spain. <sup>10</sup>Rheumatology Department, Hospital General Universitario de Alicante, Alicante 03010, Spain. <sup>11</sup>Rheumatology Department, Hospital Universitari de Bellvitge, Barcelona 08907, Spain. <sup>12</sup>Rheumatology Department, Hospital de Valme, Sevilla 41014, Spain. <sup>13</sup>Rheumatology Department, Hospital Universitario Ramón y Cajal, 28034 Madrid, Spain. <sup>14</sup>Instituto de Investigación Biomédica de Vigo, Ourense y Pontevedra 36204, Spain. <sup>15</sup>Rheumatology Department, Hospital Universitario A Coruña, A Coruña 15006, Spain. <sup>16</sup>Rheumatology Department, Hospital Complejo Asistencial Universitario de León, León 24001, Spain. <sup>17</sup>Life Sciences, Barcelona Supercomputing Centre, Barcelona 08034, Spain. <sup>18</sup>HudsonAlpha Institute for Biotechnology, Huntsville, Alabama 35806, USA. <sup>19</sup>Rheumatology Department, Hospital Regional Universitario de Málaga, Málaga 29010, Spain. <sup>20</sup>Instituto de Investigación Biomédica de Málaga, Universidad de Málaga, Málaga 29010, Spain.

Received: 26 January 2017 Accepted: 22 May 2017

Published online: 15 June 2017

## References

- Danchenko N, Satia JA, Anthony MS. Epidemiology of systemic lupus erythematosus: a comparison of worldwide disease burden. *Lupus*. 2006; 15(5):308–18.
- Lisnevskaja L, Murphy G, Isenberg D. Systemic lupus erythematosus. *Lancet*. 2014;384(9957):1878–88. doi:10.1016/S0140-6736(14)60128-8. Epub 2014 May 60131.
- Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, Schaller JG, Talal N, Winchester RJ. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum*. 1982;25(11):1271–7.
- Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, Tsao BP, Vyse TJ, Langefeld CD, Nath SK, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PTK, KIAA1542 and other loci. *Nat Genet*. 2008;40(2):204–10. doi:10.1038/ng.81.Epub2008Jan1020.
- Bentham J, Morris DL, Cunnigham Graham DS, Pinder CL, Tomblason P, Behrens TW, Martin J, Fairfax BP, Knight JC, Chen L, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat Genet*. 2015;47(12):1457–64. doi:10.1038/ng.3434.Epub2015Oct1426.
- Olson JM, Song Y, Dudek DM, Moser KL, Kelly JA, Bruner GR, Downing KJ, Berry CK, James JA, Harley JB. A genome screen of systemic lupus erythematosus using affected-relative-pair linkage analysis with covariates demonstrates genetic heterogeneity. *Genes Immun*. 2002;3 Suppl 1:S5–S12.
- Ceccarelli F, Perricone C, Borgiani P, Ciccacci C, Rufini S, Cipriano E, Alessandri C, Spinelli FR, Sili Scavalli A, Novelli G, et al. Genetic factors in systemic lupus erythematosus: contribution to disease phenotype. *J Immunol Res*. 2015;2015:745647. doi:10.1155/2015/745647. Epub742015Dec745621.
- De Azevedo SJ, Addobbati C, Sandrin-Garcia P, Crovella S. Systemic lupus erythematosus: old and new susceptibility genes versus clinical manifestations. *Curr Genomics*. 2014;15(1):52–65. doi:10.2174/138920291501140306113715.
- Sanchez E, Nadig A, Richardson BC, Freedman BI, Kaufman KM, Kelly JA, Niewold TB, Kamen DL, Gilkeson GS, Ziegler JT, et al. Phenotypic associations of genetic susceptibility loci in systemic lupus erythematosus. *Ann Rheum Dis*. 2011;70(10):1752–7. doi:10.1136/ard.2011.154104. Epub152011Jun154130.
- Taylor KE, Remmers EF, Lee AT, Ortmann WA, Plenge RM, Tian C, Chung SA, Nititham J, Hom G, Kao AH, et al. Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus. *PLoS Genet*. 2008;4(5):e1000084. doi:10.1371/journal.pgen.1000084.
- Robinson MR, Wray NR, Visscher PM. Explaining additional genetic variation in complex traits. *Trends Genet*. 2014;30(4):124–32. doi:10.1016/j.tig.2014.02.003.Epub2014Mar1011.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461(7265):747–53. doi:10.1038/nature08494.
- Du Y, Xie J, Chang W, Han Y, Cao G. Genome-wide association studies: inherent limitations and future challenges. *Front Med*. 2012;6(4):444–50. doi:10.1007/s11684-012-0225-3.Epub2012Nov11683.
- Johnson RC, Nelson GW, Troyer JL, Lautenberger JA, Kessing BD, Winkler CA, O'Brien SJ. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics*. 2010;11:724. doi:10.1186/1471-2164-1111-1724.
- Zhang K, Cui S, Chang S, Zhang L, Wang J. i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. *Nucleic Acids Res*. 2010;38(Web Server issue):W90–95. doi:10.1093/nar/gkq1324.Epub2010Apr1030.
- Gui H, Li M, Sham PC, Cherny SS. Comparisons of seven algorithms for pathway analysis using the WTCCC Crohn's disease dataset. *BMC Res Notes*. 2011;4:386. doi:10.1186/1756-0500-1184-386.
- Ramanan VK, Shen L, Moore JH, Saykin AJ. Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends Genet*. 2012;28(7):323–32. doi:10.1016/j.tig.2012.03.004.Epub2012Apr1013.
- Wu MC, Kraft P, Epstein MP, Taylor DM, Chanock SJ, Hunter DJ, Lin X. Powerful SNP-set analysis for case-control genome-wide association studies. *Am J Hum Genet*. 2010;86(6):929–42. doi:10.1016/j.ajhg.2010.05.002.
- Chen D, Enroth S, Ivansson E, Gyllenstein U. Pathway analysis of cervical cancer genome-wide association study highlights the MHC region and pathways involved in response to infection. *Hum Mol Genet*. 2014;23(22):6047–60. doi:10.1093/hmg/ddu304.Epub2014Jun6016.
- Aterido A, Julia A, Ferrandiz C, Puig L, Fonseca E, Fernandez-Lopez E, Dauden E, Sanchez-Carazo JL, Lopez-Esteban JL, Moreno-Ramirez D, et al. Genome-wide pathway analysis identifies genetic pathways associated with psoriasis. *J Invest Dermatol*. 2016;136(3):593–602. doi:10.1016/j.jid.2015.11.026.Epub2015Dec1029.
- Julia A, Tortosa R, Hernanz JM, Canete JD, Fonseca E, Ferrandiz C, Unamuno P, Puig L, Fernandez-Sueiro JL, Sanmarti R, et al. Risk variants for psoriasis vulgaris in a large case-control collection and association with clinical subphenotypes. *Hum Mol Genet*. 2012;21(20):4549–57. Epub 2012 Jul 4519.
- Yu C, Gershwin ME, Chang C. Diagnostic criteria for systemic lupus erythematosus: a critical review. *J Autoimmun*. 2014;48-49:10–3. doi:10.1016/j.jaut.2014.10.01. Epub 2014 Jan 1021.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3):559–75. Epub 2007 Jul 2025.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904–9. Epub 2006 Jul 2023.
- Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007;39(7):906–13.
- Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190–1. doi:10.1093/bioinformatics/btq340. Epub 2010 Jul 2198.
- Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet*. 2010;11(12):843–54. doi:10.1038/nrg2884.
- de Leeuw CA, Neale BM, Heskes T, Posthuma D. The statistical properties of gene-set analysis. *Nat Rev Genet*. 2016;17(6):353–64. doi:10.1038/nrg.2016.29.
- Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27:1739–40. doi:10.1093/bioinformatics/btr260. Epub 2011 May 1735.
- Edwards YJ, Beecham GW, Scott WK, Khuri S, Bademci G, Tekin D, Martin ER, Jiang Z, Mash DC, French-Mullen J, et al. Identifying consensus disease pathways in Parkinson's disease using an integrative systems biology approach. *PLoS One*. 2011;6(2):e16917. doi:10.1371/journal.pone.0016917.
- Lesiak A, Narbutt J, Kobos J, Kordek R, Sysa-Jedrzejowska A, Norval M, Wozniacka A. Systematic administration of chloroquine in discoid lupus erythematosus reduces skin lesions via inhibition of angiogenesis. *Clin Exp Dermatol*. 2009;34(5):570–5. doi:10.1111/j.1365-2230.2008.03006.x. Epub 02008 Dec 03010.
- Chang AY, Werth VP. Treatment of cutaneous lupus. *Curr Rheumatol Rep*. 2011;13(4):300–7. doi:10.1007/s11926-011-0180-z.
- Khandpur S, Sharma VK, Sumanth K. Topical immunomodulators in dermatology. *J Postgrad Med*. 2004;50(2):131–9.
- Ihaka R, Gentleman G. R: A language for data analysis and graphics. *J Comput Graph Stat*. 1996;5(5):299–314.
- Shibuya M, Claesson-Welsh L. Signal transduction by VEGF receptors in regulation of angiogenesis and lymphangiogenesis. *Exp Cell Res*. 2006; 312(5):549–60. Epub 2005 Dec 2005.
- Ferrara N. Vascular endothelial growth factor: basic science and clinical progress. *Endocr Rev*. 2004;25(4):581–611.
- Sakly N, Mirshahi P, Ducros E, Soria J, Ghedira I, Mirshahi M. Angiogenic activity in sera of patients with systemic lupus erythematosus. *Lupus*. 2009; 18(8):705–12. doi:10.1177/0961203309103087.
- Walchner M, Meurer M, Plewig G, Messer G. Clinical and immunologic parameters during thalidomide treatment of lupus erythematosus. *Int J Dermatol*. 2000;39(5):383–8.
- Cuadrado MJ, Karim Y, Sanna G, Smith E, Khamashta MA, Hughes GR. Thalidomide for the treatment of resistant cutaneous lupus: efficacy and safety of different therapeutic regimens. *Am J Med*. 2005;118(3):246–50.
- Kuryliszyn-Moskal A, Klimiuk PA, Sierakowski S, Ciolkiewicz M. Vascular endothelial growth factor in systemic lupus erythematosus: relationship to

- disease activity, systemic organ manifestation, and nailfold capillaroscopic abnormalities. *Arch Immunol Ther Exp (Warsz)*. 2007;55:179–85. Epub 2007 Jun 2008.
41. Moneib HA, Salem SA, Aly DG, Khedr HT, Wafaey HA, Hassan HE. Assessment of serum vascular endothelial growth factor and nail fold capillaroscopy changes in systemic lupus erythematosus with and without cutaneous manifestations. *J Dermatol*. 2012;39(1):52. doi:10.1111/j.1346-8138.2011.01322.x. Epub 02011 Sep 01328.
  42. Mustafa MB, Porter SR, Smoller BR, Sitaru C. Oral mucosal manifestations of autoimmune skin diseases. *Autoimmun Rev*. 2015;14(10):930–51. doi:10.1016/j.autrev.2015.06.005.Epub2015Jun1024.
  43. Khatibi M, Shakoopour AH, Jahromi ZM, Ahmadzadeh A. The prevalence of oral mucosal lesions and related factors in 188 patients with systemic lupus erythematosus. *Lupus*. 2012;21(12):1312–5. Epub 2012 Jul 1325.
  44. Munoz-Corcuera M, Esparza-Gomez G, Gonzalez-Moles MA, Bascones-Martinez A. Oral ulcers: clinical aspects. A tool for dermatologists. Part II. Chronic ulcers. *Clin Exp Dermatol*. 2009;34:456–61. doi:10.1111/j.1365-2230.2009.03219.x. Epub 02009 Apr 03214.
  45. Arbiser JL, Johnson D, Cohen C, Brown LF. High-level expression of vascular endothelial growth factor and its receptors in an aphthous ulcer. *J Cutan Med Surg*. 2003;7(3):225–8.
  46. Jacobson JM, Greenspan JS, Spritzler J, Ketter N, Fahey JL, Jackson JB, Fox L, Chernoff M, Wu AW, MacPhail LA, et al. Thalidomide for the treatment of oral aphthous ulcers in patients with human immunodeficiency virus infection. National Institute of Allergy and Infectious Diseases AIDS Clinical Trials Group. *N Engl J Med*. 1997;336(21):1487–93.
  47. Grinspan D. Significant response of oral aphthosis to thalidomide treatment. *J Am Acad Dermatol*. 1985;12(1 Pt 1):85–90.
  48. Urman JD, Lowenstein MB, Abeles M, Weinstein A. Oral mucosal ulceration in systemic lupus erythematosus. *Arthritis Rheum*. 1978;21(1):58–61.
  49. Parodi A, Massone C, Cacciapuoti M, Aragone MG, Bondavalli P, Cattarini G, Rebora A. Measuring the activity of the disease in patients with cutaneous lupus erythematosus. *Br J Dermatol*. 2000;142(3):457–60.
  50. Taylor J, Glenny AM, Walsh T, Brocklehurst P, Riley P, Gorodkin R, Pemberton MN. Interventions for the management of oral ulcers in Behcet's disease. *Cochrane Database Syst Rev*. 2014;9:CD011018. doi:10.1002/14651858.CD14011018.pub14651852.
  51. Yalcin B, Arda N, Tezel GG, Erman M, Alli N. Expressions of vascular endothelial growth factor and CD34 in oral aphthous lesions of Behcet's disease. *Anal Quant Cytol Histol*. 2006;28(6):303–6.
  52. Brozovic S, Vucicevic-Boras V, Mravak-Stipetic M, Jukic S, Kleinheinz J, Lukac J. Salivary levels of vascular endothelial growth factor (VEGF) in recurrent aphthous ulceration. *J Oral Pathol Med*. 2002;31(2):106–8.
  53. Kim YS, Park SW, Kim MH, Jang EJ, Park JS, Park SJ, Baik HW, Chung G, Hahm KB. Novel single nucleotide polymorphism of the VEGF gene as a risk predictor for gastroduodenal ulcers. *J Gastroenterol Hepatol*. 2008;23 Suppl 2:S131–9. doi:10.1111/j.1440-1746.2008.05404.x.
  54. Namjou B, Kim-Howard X, Sun C, Adler A, Chung SA, Kaufman KM, Kelly JA, Glenn SB, Guthridge JM, Scofield RH, et al. PTPN22 association in systemic lupus erythematosus (SLE) with respect to individual ancestry and clinical sub-phenotypes. *PLoS One*. 2013;8(8):e69404. doi:10.1371/journal.pone.0069404. eCollection 0062013.
  55. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol*. 2012;8(2):e1002375. doi:10.1371/journal.pcbi.1002375. Epub 1002012 Feb 1002323.
  56. ENCODE:Project:Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*. 2004;306(5696):636–40.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

