Journal of Autoimmunity 72 (2016) 65–72

Contents lists available at ScienceDirect

Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm

Definition of mutations in polyautoimmunity

Angad Johar [a], Juan C. Sarmiento-Monroy [b], Adriana Rojas-Villarraga [b], Maria F. Silva-Lara [a], Hardip R. Patel [c], Ruben D. Mantilla [b], Jorge I. Velez [a], Klaus-Martin Schulte [d], Claudio Mastronardi [a], Mauricio Arcos-Burgos [a, b, **], Juan-Manuel Anaya [b, *]

[a] Genomics and Predictive Medicine, Genome Biology Department, John Curtin School of Medical Research, ANU College of Medicine, Biology & Environment, The Australian National University, Canberra, ACT, Australia
[b] Center for Autoimmune Diseases Research (CREA), School of Medicine and Health Sciences, Universidad del Rosario, Bogota, Colombia
[c] Genome Discovery Unit, Genome Biology Department, John Curtin School of Medical Research, ANU College of Medicine, Biology & Environment, The Australian National University, Canberra, ACT, Australia
[d] Department of Immunology, John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia

ARTICLE INFO

Article history:
Received 25 March 2016
Received in revised form
6 May 2016
Accepted 7 May 2016
Available online 18 May 2016

Keywords:
Polyautoimmunity
Familial autoimmunity
Extreme phenotype
Genetics
Linkage
Network analysis

ABSTRACT

Objectives: Familial autoimmunity and polyautoimmunity represent extreme phenotypes ideal for identifying major genomic variants contributing to autoimmunity. Whole exome sequencing (WES) and linkage analysis are well suited for this purpose due to its strong resolution upon familial segregation patterns of functional protein coding and splice variants. The primary objective of this study was to identify potentially autoimmune causative variants using WES data from extreme pedigrees segregating polyautoimmunity phenotypes.
Methods: DNA of 47 individuals across 10 extreme pedigrees, ascertained from probands affected with polyautoimmunity and familial autoimmunity, were selected for WES. Variant calls were obtained through Genome Analysis Toolkit. Filtration and prioritization framework to identify mutation(s) were applied, and later implemented for genetic linkage analysis. Sanger sequencing corroborated variants with significant linkage.
Results: Novel and mostly rare variants harbored in *SRA1, MLL4, ABCB8, DHX34* and *PLAUR* showed significant linkage (LOD scores are >3.0). The strongest signal was in *SRA1,* with a LOD score of 5.48. Network analyses indicated that *SRA1*, *PLAUR* and *ABCB8* contribute to regulation of apoptotic processes.
Conclusions: Novel and rare variants in genetic linkage with polyautoimmunity were identified throughout WES. Genes harboring these variants might be major players of autoimmunity.

1. Introduction

Recent evidence supports the involvement of rare genetic variants with a major effect underpinning the etiology of complex disorders and that a large proportion (1/3) of deleterious alleles have a frequency of <5% [1]. This distribution of rare variants is likely due to the fact that deleterious variants will be unfavorable due to selection, thereby reducing their minor allele frequency in any given population [1]. The detection of these rare variants has shown to be achievable using severe phenotypes segregating in exceptional pedigrees [1].

Polyautoimmunity is defined as the presence of more than one autoimmune disease (AD) in a single patient [2]. If three or more ADs coexist, the condition is called multiple autoimmune syndrome (MAS), which characterizes one conspicuous and extreme example of polyautoimmunity [2,3] i.e.: ***(i)*** MAS amalgamates signs and symptoms that are present in several ADs, ***(ii)*** the MAS signs and symptoms clustering is not random but linked to subtypes, ***(iii)*** MAS frequently clusters in families, and ***(iv)*** Mendelian segregation and linkage to major loci have been established for MAS [2,3].

* Corresponding author. Center for Autoimmune Diseases Research (CREA), School of Medicine and Health Sciences, Universidad del Rosario, Cra. 26-63-B-51, Bogota, Colombia.
** Corresponding author. Center for Autoimmune Diseases Research (CREA), School of Medicine and Health Sciences, Universidad del Rosario, Cra. 26-63-B-51, Bogota, Colombia.
E-mail addresses: Mauricio.Arcos-Burgos@anu.edu.au (M. Arcos-Burgos), anayajm@gmail.com (J.-M. Anaya).

http://dx.doi.org/10.1016/j.jaut.2016.05.003
0896-8411/

During the last years we have ascertained, through probands affected by polyautoimmunity, several multigenerational and extended pedigrees clustering additional relative members affected by either single or polyautoimmunity syndromes (i.e., familial autoimmunity) [4–6]. Several of these pedigrees were ascertained from the "Paisa" genetic isolate, where several major loci to complex disorders, including MAS have previously been mapped [7–10]. Given the phenotypic characteristics of the polyautoimmunity pedigree members it is fair to hypothesize that autoimmunity segregates in those pedigrees in a Mendelian fashion with different patterns of penetrance (Fig. 1).

In here we report the analysis of whole exome sequencing (WES) of affected and unaffected members of ten pedigrees clustering autoimmune phenotypes (Fig. 1). Using classical and modern techniques of genetic pedigree analyses, we found strong and significant signals of genetic linkage of exonic and regulatory mutations with polyautoimmunity. Given the functional nature of these mutations and the strong linkage to polyautoimmunity it is very likely that genes harboring these mutations are major contributors in causing autoimmunity.

2. Methods

2.1. Families segregating polyautoimmunity

Phenotypes of the individuals that were subject of exome capture from the ascertained pedigrees are presented in Supplementary Table 1. All the individuals were enrolled in the Center for Autoimmune Diseases Research (CREA) at the Universidad del Rosario, in Bogota and Medellin, Colombia. Written informed consent was obtained from all individuals before enrolment in this study, which was approved by the ethics committee of the Universidad del Rosario. Most of the families (n = 7) were recruited from the Paisa community living in Antioquia, Colombia. Historical evidence has documented that individuals from the Paisa community are endogenous, homogenous and have very little population stratification [11]. All patients were diagnosed and followed-up by a single team (JMA, JCS, ARV, RDM) according to international classification criteria (Supplementary Table 2).

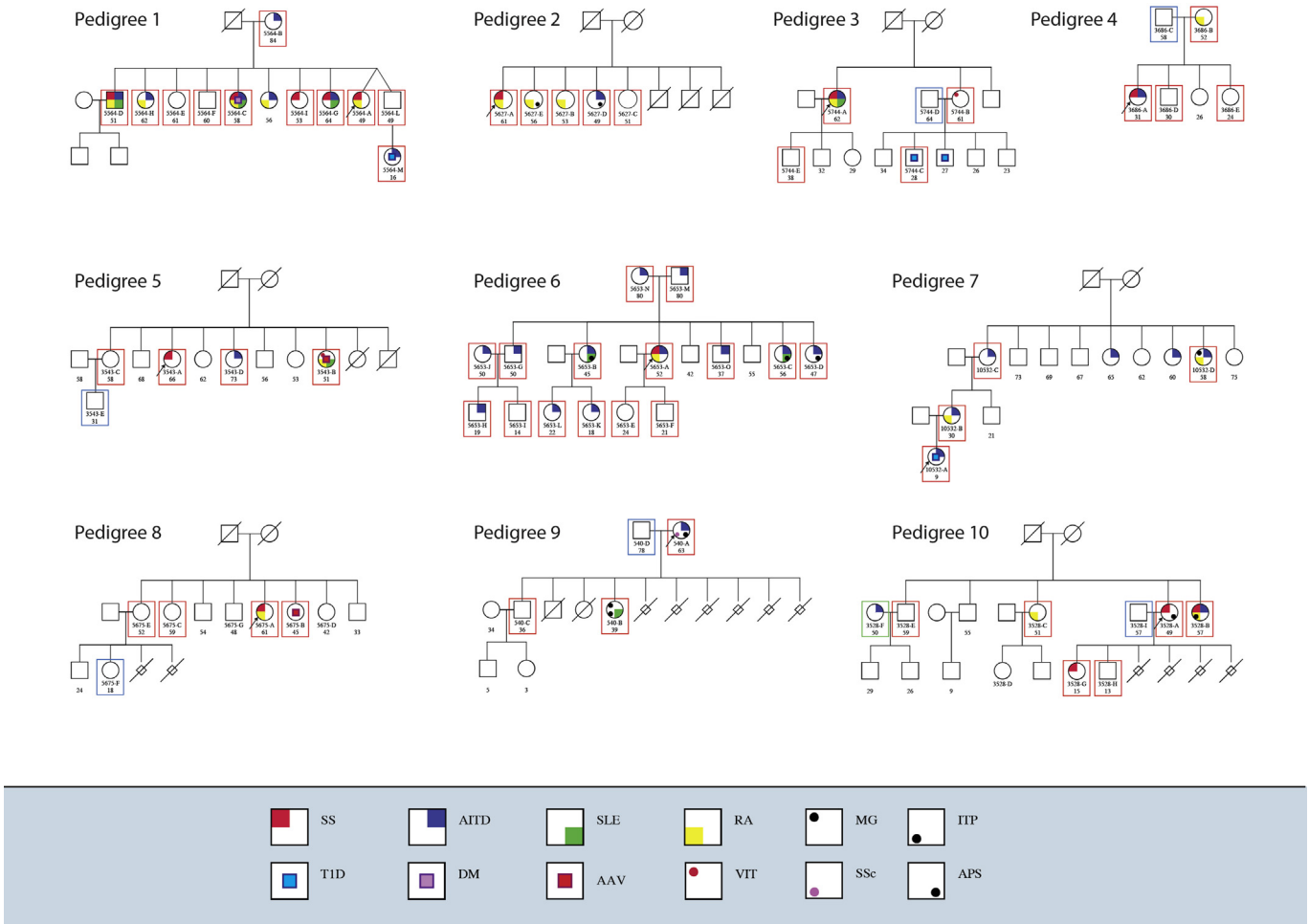

Fig. 1. Ten pedigrees segregating autoimmune phenotypes (MAS, Polyautoimmunity, or single autoimmune phenotypes). Autoimmune phenotypes are denoted in the foot caption e.g. Sjogren's Syndrome (SS), Autoimmune Thyroid Disease (AITD), Systemic Lupus Erythematous (SLE), Rheumatoid Arthritis (RA), Myasthenia Gravis (MG), Antiphospholipid Syndrome (APS), Idiopathic Thrombocytopenic Purpura (ITP), Type 1 Diabetes (T1D), Dermatomyositis (DM), ANCA-associated vasculitis (AAV), Vitiligo (VIT), and Systemic Sclerosis (SSc). Individuals that were subject to WES are presented in Supplementary Table 1.

2.2. DNA library preparation, exome capture and sequencing protocol

Libraries were constructed from 1 μg of genomic DNA using an Illumina TruSeq genomic DNA library kit at the Biomolecular Resource Facility (BRF), John Curtin School of Medical Research (JCSMR). Libraries were multiplexed with 6 samples pooled together (500 ng of each library). Exons were enriched from the pooled 3 μg of library DNA using the Nimblegen Exome enrichment kit (BRF). Each exome-enriched pool was run on a 100-base-pair paired end run on an Illumina HiSeq 2000 sequencer (BRF, JCSMR) [12–14].

2.3. Sequence read processing, alignment and variant calling

The sequencing image data was converted to FASTQ files containing DNA base calls (A, C, G and T) and quality scores using the Illumina CASAVA pipeline (a software programs that converts raw image data into sequences). The resulting FASTQ files were further processed for variant analysis.

The workflow for data curating and analysis for variant calling was developed by the Genome Discovery Unit (GDU), at the Australian National University. Key components of the workflow include: i) Quality assessment; ii) Read alignment using the BOWTIE aligner; iii) Local realignment around the known and novel insertions/deletions (indel) regions to refine indel boundaries; iv) Recalibration of base qualities; v) Variant calling using the GATK (Genome Alignment Tool Kit) algorithm; and vi) Assigning quality scores to variants (see more detailed workflow information published elsewhere) [12–14].

2.4. Linkage analysis and composite likelihood ratio test of pedigrees

We used the Pedigree Variant Annotation Analysis and Search Tool (pVAAST) [15–17] to identify candidate genetic variants enriched in affected family members co-segregating (in genetic linkage) with polyautoimmunity.

In brief, the pVAAST algorithm allows the identification of potential disease causing variants through a method that combines classical linkage analysis (Elston-Stewart algorithm), with case-control association, and functional impact of the tested variant using a Composite Likelihood Ratio Test (CLRTv). This produces a gene-based LOD score and CLRTp (pVAAST test statistic) value. Dominant, codominant, and recessive models were maximized under the assumption of upper bound penetrance proximal to 100%.

The significance of these test statistics is derived using a combination of permutation tests (in this case 10,000) without replacement and gene-drop simulation. This was achieved using a total of 1366 individuals as controls. This control dataset includes 1057 individuals from the 1000 genomes project (Phase 1), 184 Danish exomes, 10 individuals from the 10Gen database and 62 additional unaffected controls from the same geographical and cultural region where the extreme pedigrees were ascertained (the 1000 genomes −phase1− has a set of 60 individuals recruited from Medellin, Colombia; the same area of ascertainment of most of these pedigrees).

2.5. Further refinement and curating of candidate gene list generated via pVAAST algorithm

Candidate genes reported as containing a significant pVAAST test statistic (based on LOD score and case-control allele frequencies) were further filtered as follows: of all the genes that were significant, those containing a CLRTp score greater than the mean and with a LOD score greater than 3.0 were retained for further downstream analysis. This is conditional on the fact that variants within these genes had an allelic frequency of < 5% in the ExAC (Exome Aggregation Consortium), dbSNP and the 1000 Genomes database. Afterwards, additional filtration was conducted based upon annotations from variant effect predictors (including: Polyphen2, SIFT, MutationTaster, and FATHMM), as implemented in the DNA-seq analysis package SVS (SNP and Variation Suite) Version 8.3.4 (Golden Helix, Bozeman, USA), as well as PROVEAN (Protein Variant Effect Analyser) [12]. Mutations predicted to be damaging from at least one of these variant effect predictor algorithms are considered as plausible disease candidates for further downstream analysis.

2.6. Sanger sequencing and variant quality scores

Sanger sequencing was performed for the validation of the gene variants reported herein as previously described [12]. Briefly, a flanking region around each sequence variant site was amplified by PCR. We used the following PCR conditions for the amplification of all the amplicons: (1) initial activation step of 3 min at 94 °C, (2) 40 cycles as follows: 45 s of denaturing at 94 °C, 30 s of annealing between 53 and 62 °C (according to the primer pair, Supplementary Table 3), 60 s of extension at 72 °C, and (3) final extension step of 10 min at 72 °C. A 15-μL aliquot of the PCR product was analysed by electrophoresis in a 1.5% agarose gel to confirm the expected size of the amplicons. Afterwards, 85 μL of each PCR product was purified with the Qiaquick nucleotide removal kit (Qiagen, Valencia, CA, USA) following the manufacturer's guidelines. Thereafter, the purified PCR products were spectrophotometrically quantified with a NanoDrop ND-1000 (Wilmington, DE, USA), and sent for Sanger sequencing to the Australian Cancer Research Foundation—Biomolecular Resource Facility (BRF) at the John Curtin School of Medical Research. Bidirectional sequencing of PCR amplicons were carried out by using Big DyeTM chemistry (Big Dye Terminator, Version 3.1; Applied Biosystems, Foster City, CA, USA) with the sequencing primers reported in Supplementary Table 3. The sequencing protocol was followed according to the BRF standard operative procedures. In addition to the Sanger sequencing, variant Phred and Mapping Quality scores are also given in Supplementary Table 4. For the variant under consideration, the Phred score will quantify the probability of an incorrect base call. Meanwhile, mapping quality determines the likelihood of an incorrect alignment of the sequence read containing the identified variant(s).

2.7. Metacore pathway and network analysis

To identify potential enriched polyautoimmunity related physiological pathways involving candidate genes from the pVAAST data, network analyses were performed. For constructing networks and pathways, were examined with the 'Shortest Paths' algorithm implemented within the MetaCore software suite (Version 6.24, Build 67895, Thomson Reuters, New York, USA). Details regarding some the differences between the various algorithms can be found in the MetaCore Manual. These procedures allowed us to obtain a heuristic integration of maps and networks and rich ontologies for diseases based on the potential biological functions of these candidate genes.

3. Results

Significant signals of linkage were obtained for variants harbored in five genes: *SRA1, MLL4, ABCB8, PLAUR,* and *DHX34*

(Table 1).

When conducting permutation tests (number of permutations = 10,000), each of these genes ascertained a p-value of < 0.005 for the CLRTp test statistic. *SRA1, ABCB8* and *PLAUR* exhibited a maximisation of the LOD score (and consequently the CLRTp test statistic) under the recessive model for linkage analysis. In the case of *DHX34* and *MLL4,* the greatest maximisation occurred under the dominant model. Sanger sequencing, as shown by Fig. 2, corroborated these variants. In addition, the input parameters suggest that LOD scores are maximized for candidate variants within these genes at a disease allele penetrance between 70 and 99%. This result can account for the presence of potential phenocopies within this cohort.

A description of the numbers of affected and unaffected individuals clustering in those pedigrees carrying alternate alleles linked to polyautoimmunity is shown in Table 2. Genotypes of individuals containing mutations at the variant sites harboured in candidate genes identified from linkage and CLRT analysis are presented in Supplementary Table 5 (affected and unaffected individuals as well as the genotypes corresponding to the individual ID codes are given).

In addition to the high LOD scores and significant pVAAST statistics generated from the aforementioned analysis, these genes may also have substantial relevance in biological pathways and networks, thereby influencing key processes that underpin the physiological basis for autoimmunity in these patients (Table 3). In particular *SRA1* and *PLAUR* are seemingly both involved in the negative regulation of apoptosis (*P*-value = 1.791e-3), as well as negative regulation of cysteine type endopeptidase activity (*P*-value = 1.287e-5). As indicated in the Metacore analysis in Fig. 3 and Table 3, these processes seemingly occur as a consequence of the SF1 transcription factor binding to and subsequently activating the SRA1 protein. This in turn activates the ESR1, which then subsequently has a seemingly inhibitory effect on the protein encoded by *PLAUR*. (Table 3 and Fig. 3).

4. Discussion

As a whole, we are presenting significant evidence of linkage of functional exonic variants harbored in five genes that cosegregate with polyautoimmunity in extreme pedigrees clustering autoimmune phenotypes. The power of exome sequencing is further enhanced by the fact that this study was done in families, most of them belonging to a genetic isolate, a circumstance that increases genetic and environmental homogeneity.

The strongest candidate of the five genes identified is *SRA1* with a LOD score of 5.48. No unaffected individual was heterozygous for both of the identified variants, harbored in this locus. The 3 base insertion at chr5:139931629 (which was denoted to be damaging by the Provean variant effect prediction algorithm, as a consequence of an Arginine insertion between the Valine and Alanine at this position) was not present in any unaffected individual. The fact that *SRA1* has been enriched according to the linkage analysis and pVAAST test statistics in affected pedigree members is further supported by the aforementioned network analysis.

The relationship between *SRA1* and the immune system remains to be studied in more detail. However, it is noteworthy to mention that in mouse studies SRA1 levels of mRNA expression were found to be relatively high in mouse spleen, which presumably exerts proinflammatory actions [17]. In fact, *SRA1* KO mice showed significantly reduced expression of a subset of inflammatory-related genes, including tumor necrosis factor-α (TNF-α) and monocyte chemotactic protein-1 (MCP-1), that was accompanied by decreased levels of blood TNF-α [18].

Another candidate gene is *PLAUR* (LOD = 3.62), which encodes the urokinase plasminogen activator receptor (uPAR). When the endogenous ligand urokinase plasminogen activator (uPA) binds to uPAR, it triggers the conversion of plasminogen into plasmin, an active serine protease that is involved in key pathophysiological mechanisms occurring in cancer [19]. Interestingly, uPA and uPAR occur in immune cells, in particular on activated T-cells [20] and monocytes [21], and several of the uPA–induced effects are independent from plasminogen such as regulation of cell migration, angiogenesis, and adhesion [22]. Also, uPAR expression is enhanced by proangiogenic as well as proatherogenic growth factors and cytokines such as IL-1, suggesting its involvement in inflammatory and proliferative processes [23].

The Metacore network analysis of candidate genes from the pVAAST analysis is depicted in Fig. 3. The fact that the analysis revealed both *PLAUR* and *SRA1* are involved in the negative regulation of apoptosis and cysteine type endopeptidase activity suggests they have a potentially important contribution to the autoimmune pathophysiology within these patients. Hypothetically, the occurrence of the rare variants harbored in *SRA1* and *PLAUR* could contribute to the development of autoimmunity by dysregulating apoptosis. Thus, a putative explanation of this network would be that the transcription factor stereoidogenic factor-1 (SF1) increases the expression of SRA1, which in turn can positively modulate anti-apoptotic pathways mediated by estrogen receptor-1 (ESR1), via a binding interaction [24,25]. For instance, downstream anti-apoptotic pathways activated by ESR1 could counter PLAUR-induced apoptosis in line with the negative interaction depicted between ESR1 and PLAUR in the network analysis described in Fig. 3. Additionally, in this network, the tumor suppressor protein p53 is hypothesized to negatively regulate PLAUR, presumably by its non-apoptotic actions [26]. On the other hand PLAUR receives stimulatory feedback from its ligand (PLAU), and two transcription factors named transcription factor 7-like 2 (TCF7L2) and nuclear factor-kappa B (NF-κB), which in turn could positively influence PLAUR-mediated apoptosis. The *P*-values indicate that the association between the corresponding nodes and these GeneGo processes is highly unlikely to have arisen by chance. Furthermore, based on the results from the pVAAST and Metacore test statistics, it can be hypothesized that the negative regulation of apoptosis may be disrupted as a consequence of deficient cysteine type endopeptidase activity due to one or both mutations in each of

Table 1
LOD score and pVAAST test statistics of genes harbouring genetic variants with significant values of linkage to polyautoimmunity in the 10 families. LOD scores are calculated based on the Elston-Stewart algorithm, pVAAST scores are gene-based and *P*-values are derived from permutation (without replacement) and gene-drop simulation.

Gene	Variant (rs id)	LOD score	Linkage model	pVAAST CLRTp score	Permutation based P-value	Alleles Ref/Alt
SRA1	5:139931629-Ins (rs5871740)	5.5	Recessive	214.030	0.0001	C/CTCG
SRA1	5:139931629-SNV (rs202193903)	5.5	Recessive	214.030	0.0001	C/G
ABCB8	7:150744528	4.1	Recessive	27.359	0.00903	G/T
ABCB8	7:150744370	4.1	Recessive	27.359	0.00903	CGT/–
DHX34	19:47883126-SNV (rs151213663)	3.8	Dominant	46.873	0.0003	C/T
PLAUR	19:44153100 (rs4760)	3.6	Recessive	20.031	0.004	A/G
MLL4	19:36218440-SNV (rs186268702)	3.4	Dominant	137.17	0.0001	G/A

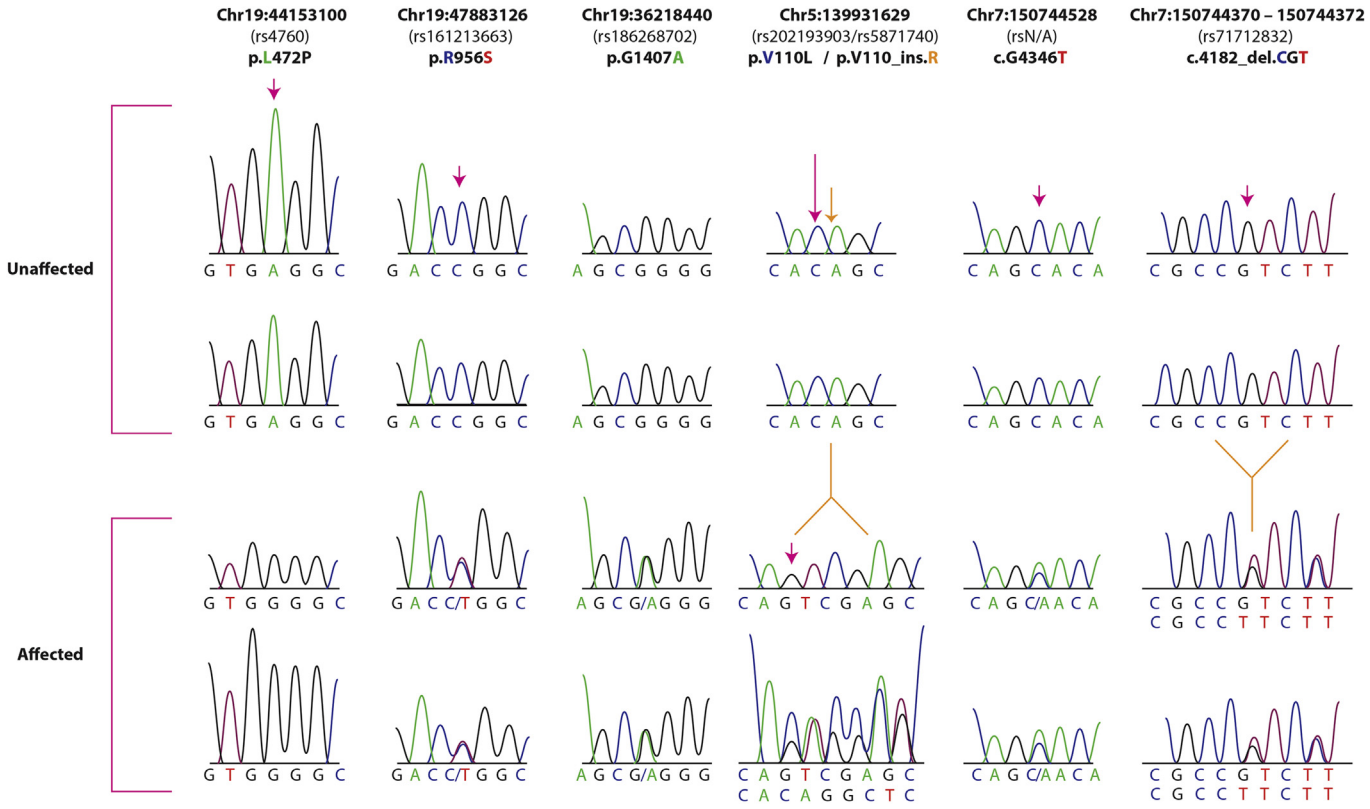

Fig. 2. Chromatograms of the results of the Sanger sequencing of unaffected (top rows) and affected individuals (bottom rows) segregating variants harbored in genes in significant genetic linkage with polyautoimmunity. The bold subheading on the top indicates the chromosome (Chr) number and the variant position for each one of the studied genes; from left to right they represent *PLAUR, DHX34, MLL4, SRA1* and *ABCB8* (last two panels). Also each row represents the DNA of a separate individual. The available reference SNP (rs) numbers are also shown between parentheses. Below, it is shown the position of the changes either within the protein (p) or the cDNA (c) sequence for the 5 coding and 2 non-coding variants, respectively. In each coding variant the left capital letter indicates the aminoacid occurring in the unaffected individuals, and the right capital letter shows either the aminoacid change (rs4760, rs151213663, rs186268702 and rs202193903) or the insertion (ins) of an arginine (R) in the rs5871740 variant that occurs in the affected individuals. With regards to the non-coding variants, the capital letters indicate a nucleotide change (rsN/A) or a trinucleotide deletion (del) in the rs71712832 variant. In both coding and non-coding variants each capital letter color corresponds to that of the nucleotide involved in the variation as shown in Sanger chromatograms. A short sequence is shown under each individual chromatogram. In case of heterozygosity, there are two sequence reads representing the maternal and parental DNA. The red and orange arrows signal the nucleotide position where the variations occur. The orange inverted and non-inverted Y-shaped pointers depicts the expansion and contraction of the sequence occurring as consequence of the insertion and deletion, respectively, in affected individuals.

these respective genes. Therefore, a deficiency in negative regulation of apoptotic signaling, may lead to enhanced destruction of non-foreign cells, underpinning the basis of the ADs carried by these 32 affected patients. The crucial role of apoptosis in autoimmunity has been studied since more than 20 years [27–29]. Thus, the possible functional relationship between *SRA1* and *PLAUR* and autoimmune-related apoptosis should be investigated in future studies.

Although the remaining genes identified from the pVAAST analysis did not show evidence of pathway or physiological

Table 2
Number of affected and unaffected individuals with mutations harboured in candidate genes from the pVAAST analysis, and allele frequencies in external databases (dbSNP, 1000 Genomes and Exome Aggregation Consortium - ExAC).

Variant/Gene	Number of affected individuals (across families)	Number of unaffected individuals (across families)	dbSNP allele frequency (build 141)	1 K allele frequency (phase 3)	ExAC frequency
5:139931629-Ins (rs7871740) *SRA1*	14 homozygous 5 heterozygous	4 heterozygous	Absent	Absent	Absent
5:139931629-SNV (rs202193903) *SRA1*	1 homozygous 11 heterozygous	2 heterozygous	Absent	Absent	Absent
19:36218440-SNV (rs186268702) *MLL4*	5 heterozygous	0	0.184%	0.1%	0.07%
19:47883126-SNV (rs151213663) *DHX34*	4 heterozygous	0	0.587%	0.24%	0.4718%
7:150744528 *ABCB8*	3 heterozygous	0	Absent	Absent	Absent
7:150744370 *ABCB8*	6 heterozygous	2 heterozygous	30.367%	31.73%	Absent
19:44153100-SNV (rs4760) *PLAUR*	2 (homozygous), 8 (heterozygous)	5 heterozygous	9.248%	6.85% (1.2% homozygotes)	12.25% (Homozygotes 1.9%)

Ins: insertion, SNV: single nucleotide variation.

Table 3
GeneGo ontological process descriptions produced by the Metacore network analysis for the candidate gene list. Table gives the nodes corresponding to the process descriptions as well as the *P*-value.

Gene(s)	Network algorithm	GeneGo process annotation	Annotation process P-value	Network nodes
PLAUR SRA1	Shortest paths	Negative regulation of apoptosis	1.791e-13	SF1 SRA1 ESR1 (nuclear) PLAUR
PLAUR SRA1	Shortest paths	Negative Regulation of cysteine type endopeptidase activity	1.287e-5	SF1 SRA1 ESR1 (nuclear) PLAUR
PLAUR SRA1	Shortest paths	Positive regulation of phosphorylation	1.592e-20	SF1 SRA1 ESR1 (nuclear) PLAUR
PLAUR	Shortest paths	Negative regulation of proteolysis	1.761e-4	ESR1 (nuclear) PLAUR
PLAUR	Shortest paths	Urokinase plasminogen activator signalling	2.404e-3	ESR1 (nuclear) PLAUR

significance in the Metacore analysis, this does not rule out their potential relevance in autoimmunity. A brief description of their biological relevance is provided below.

MLL4 encodes an enzyme named histone methyl transferase that methylates 'Lys-4' of histone H3, which is a key epigenetic modification involved in gene activation. This enzyme can cause di- and trimethylation to H3 histones interacting with transcription start sites during gene transcription [30,31] and monomethylations to H3 histones associated with enhancer sequences [32]. *MLL4* appears to be involved in different types of cancers such as endometrial, large intestine, glioma and liver carcinomas [33]. Mutations harboured in its paralog gene, namely *MLL2,* have been associated with Kabuki syndrome, which is a complex multisystem developmental disorder characterized by craniofacial, intellectual and cardiac defects [34]. Remarkably, it has been recently reported that patients with Kabuki syndrome carry mutations within *MLL2* presented in humoral immune diseases and in some cases, autoimmunity [35]. In the case of *MLL4*, empirical data is also present for its potential relevance in autoimmunity. It has been found that the methyltransferase encoded by *MLL4* seemingly binds to PTIP (Pax Transactivation Domain Interacting Protein) [36]. PTIP in turn controls for class switch recombination on the immunoglobulin heavy chain locus. Thus a mutation in *MLL4,* may affect the binding of the methyltransferase with PTIP, thereby potentially resulting in defects in class switching. Consequently, this evidence suggests that *MLL4* might also be involved in key pathophysiological aspects underlying autoimmunity.

ABCB8 encodes a multi-pass protein (ABCB8) that is an integral component of the inner mitochondrial membrane. ABCB8 is an ATP-dependent transporter involved in mitochondrial iron export, which is crucial for normal cardiac function. Mouse with deficient expression of ABCB8 in their heart showed aberrant iron homeostasis, increased mitochondrial damage and developed cardiomyopathy [37]. *ABCB8* is a member of the MDR/TAP subfamily. Members of this subfamily are involved in antigen processing and presentation by pumping degraded cytosolic peptides across the endoplasmic reticulum into the membrane-bound compartment where class I molecules assemble [38,39]. The identified variants chr7:150744528 and chr7:150744370 both are located in the E2F1 binding site (http://htd.cbi.pku.edu.cn). By themselves the mutations likely decrease ABCB8 expression, reduce mitochondrial potential, render the cells susceptible to redox stress, and promote cell death. The effects of the mutations may hence converge with cell cycle arrest of infected cells (e.g., by Parvovirus B19) [40]. Thus, the possibility exists that mutations of *ABCB8* could contribute to autoimmunity by altering cell death and antigen processing and presentation to the adaptive immune system.

Finally, *DHX34* encodes an RNA helicase (DHX34). *In vitro* studies suggested that DHX34 is involved in nonsense-mediated decay (NMD), a surveillance process that degrades aberrant mRNAs that harbour premature stop codons and also might regulate the abundance of RNAs [41]. DHX34 belongs to the DEAD box protein family; another RNA helicase from this family such as DHX33 has been reported to play a crucial role in sensing viral RNA in myeloid dendritic cells [42]. The possible biological relationships between DHX34 and autoimmunity remains to be elucidated.

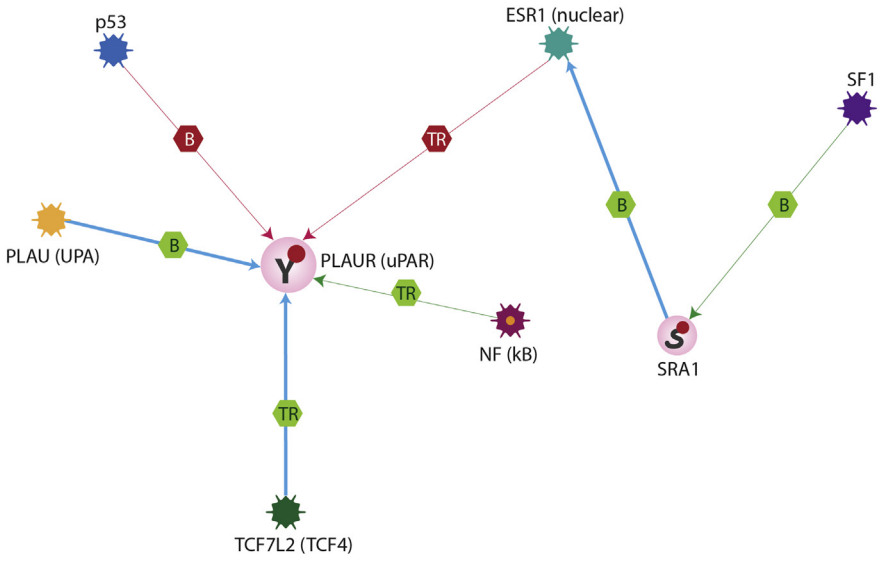

Fig. 3. Metacore network analysis of candidate genes from the pVAAST analysis. Network is generated by the 'Shortest Paths' algorithm. Proteins encoded by genes from the candidate list are denoted by a circular red dot. Abbreviations for the mechanism of interaction occurring between the respective proteins on each node are as follows: TR denotes transcriptional regulation whilst B represents binding. In addition, the effects and mechanisms of these interactions are also represented via the following color coordination: red = inhibitory effect, and green = activation.

5. Study's limitations

We acknowledge the lack of functional genetics as a shortcoming of our study. In addition and as we have mentioned elsewhere: hybridization probes are not available for all annotated exons within the gold standard databases. Also, exome sequencing is not able to detect mutations in non-coding DNA that alter gene function by various regulatory mechanisms and enhancer effects. Such variants are emerging as important contributors to genetic disease and they occur in > 98% of the human genome, which is missed by exome capture [3]. Thus, it is recognized that all genes associated with polyautoimmunity were not included in our study and that new associated genes may be discovered as updated information becomes available.

It is worth to mention that the dissection of major genes, harboring mutations predisposing to polyautoimmunity, in this manuscript, does not discard interactions of these loci with environmental noxae (i.e. the autoimmune ecology) [43], neither does it the switch of the polyautoimmunity phenotype on by infectious diseases. With the available methods for analysis of pedigrees, it is very challenging to dissect epigenetic effects. The twins and case-control cohorts are better suited to characterize these non-hereditary factors than the use of a limited number of extended and multigenerational pedigrees that maximize the best strategies of physical gene mapping.

6. Conclusions

In summary, our linkage analysis in combination with the pVAAST composite likelihood ratio test has successfully identified 5 candidate genes that account for the observed autoimmune phenotypes in extreme pedigrees with polyautoimmunity. The strongest candidate from a statistical and biological relatedness point of view was by far *SRA1*. It is hoped that further functional studies can validate the postulation for the contribution of these genes in ADs.

Competing interests

None.

Acknowledgements

We thank the patients and their families for their cooperation, and the members of the Center for Autoimmune Diseases Research (CREA) for fruitful discussions. Supported by grants from Colciencias (373-2011, 0425-2013) and the School of Medicine and Health Sciences, Universidad del Rosario (ABN011), Bogota, Colombia.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jaut.2016.05.003.

References

[1] G. Gibson, Rare and common variants: twenty arguments, Nat. Rev. Genet. 13 (2011) 135–145.
[2] J.M. Anaya, The diagnosis and clinical significance of polyautoimmunity, Autoimmun. Rev. 13 (4–5) (2014) 423–426.
[3] J.M. Anaya, Common mechanisms of autoimmune diseases (the autoimmune tautology), Autoimmun. Rev. 11 (11) (2012) 781–784.
[4] J.M. Anaya, J. Castiblanco, A. Rojas-Villarraga, et al., The multiple autoimmune syndromes. A clue for the autoimmune tautology, Clin. Rev. Allergy Immunol. 43 (2012) 256–264.
[5] J. Castiblanco, J.C. Sarmiento-Monroy, R.D. Mantilla, et al., Familial aggregation and segregation analysis in families presenting autoimmunity, polyautoimmunity, and multiple autoimmune syndrome, J. Immunol. Res. 2015 (2015) 572353.
[6] J. Cárdenas-Roldán, A. Rojas-Villarraga, J.M. Anaya, How do autoimmune diseases cluster in families? A systematic review and meta-analysis, BMC Med. 11 (2013) 73.
[7] J.I. Velez, S.C. Chandrasekharappa, E. Henao, et al., Pooling/bootstrap-based GWAS (pbGWAS) identifies new loci modifying the age of onset in PSEN1 p.Glu280Ala Alzheimer's disease, Mol. Psychiatry 18 (2013) 568–575.
[8] M. Camargo, D. Rivera, L. Moreno, et al., GWAS reveals new recessive loci associated with non-syndromic facial clefting, Eur. J. Med. Genet. 55 (2012) 510–514.
[9] M. Arcos-Burgos, M. Jain, M.T. Acosta, et al., A common variant of the latrophilin 3 gene, LPHN3, confers susceptibility to ADHD and predicts effectiveness of stimulant medication, Mol. Psychiatry 15 (2010) 1053–1066.
[10] M.L. Marazita, A.C. Lidral, J.C. Murray, et al., Genome scan, fine-mapping, and candidate gene analysis of non-syndromic cleft lip with or without cleft palate reveals phenotype-specific differences in linkage and association results, Hum. Hered. 68 (2009) 151–170.
[11] M.L. Bravo, C.Y. Valenzuela, O.M. Arcos-Burgos, Polymorphisms and phyletic relationships of the paisa community from antioquia (Colombia), Gene Geogr. 10 (1996) 11–17.
[12] A.S. Johar, J.M. Anaya, D. Andrews, et al., Candidate gene discovery in autoimmunity by using extreme phenotypes, next generation sequencing and whole exome capture, Autoimmun. Rev. 14 (2015) 204–209.
[13] A.S. Johar, C. Mastronardi, A. Rojas-Villarraga, et al., Novel and rare functional genomic variants in multiple autoimmune syndrome and Sjogren's syndrome, J. Transl. Med. 13 (2015) 173.
[14] G. Paz-Filho, M.C. Boguszewski, C.A. Mastronardi, et al., Whole exome sequencing of extreme morbid obesity patients: translational implications for obesity and related disorders, Genes 5 (2014) 709–725.
[15] T.M. Darlington, R. Pimentel, K. Smith, et al., Identifying rare variants for genetic risk through a combined pedigree and phenotype approach: application to suicide and asthma, Transl. Psychiatry 4 (2014) e471.
[16] H. Hu, J.C. Roach, H. Coon, et al., A unified test of linkage analysis and rare-variant association for analysis of pedigree sequence data, Nat. Biotechnol. 32 (2014) 663–669.
[17] B. Kennedy, Z. Kronenberg, H. Hu, et al., Using VAAST to identify disease-associated variants in next-generation sequencing data, Curr. Protoc. Hum. Genet. 81 (2014), 6 14 1-6 25.
[18] S. Liu, L. Sheng, H. Miao, et al., SRA gene knockout protects against diet-induced obesity and improves glucose tolerance, J. Biol. Chem. 289 (2014) 13000–13009.
[19] M.J. Duffy, The urokinase plasminogen activator system: role in malignancy, Curr. Pharm. Des. 10 (2004) 39–49.
[20] A. Nykjaer, B. Moller, R.F. Todd 3rd, et al., Urokinase receptor. An activation antigen in human T lymphocytes, J. Immunol. 152 (1994) 505–516.
[21] J. Humphries, J.A. Gossage, B. Modarai, et al., Monocyte urokinase-type plasminogen activator up-regulation reduces thrombus size in a model of venous thrombosis, J. Vasc. Surg. 50 (2009) 1127–1134.
[22] T. Chavakis, A.K. Willuweit, F. Lupu, K.T. Preissner, S.M. Kanse, Release of soluble urokinase receptor from vascular cells, Thromb. Haemost. 86 (2001) 686–693.
[23] M. Baran, L.N. Mollers, S. Andersson, et al., Survivin is an essential mediator of arthritis interacting with urokinase signalling, J. Cell Mol. Med. 13 (2009) 3797–3808.
[24] C. Chiueh, S. Lee, T. Andoh, D. Murphy, Induction of antioxidative and anti-apoptotic thioredoxin supports neuroprotective hypothesis of estrogen, Endocrine 21 (1) (2003) 27–31.
[25] S.Y. Lee, T. Andoh, D.L. Murphy, C.C. Chiueh, 17beta-estradiol activates ICI 182,780-sensitive estrogen receptors and cyclic GMP-dependent thioredoxin expression for neuroprotection, FASEB J. 17 (8) (2003) 947–948.
[26] A. Tedeschi, S. Di Giovanni, The non-apoptotic role of p53 in neuronal biology: enlightening the dark side of the moon, EMBO Rep. 10 (6) (2009) 576–583.
[27] N. Ogawa, H. Dang, L. Kong, J.M. Anaya, G.T. Liu, N. Talal, Lymphocyte apoptosis and apoptosis-associated gene expression in Sjögren's syndrome, Arthritis Rheum. 39 (11) (1996) 1875–1885.
[28] W.M. Kuhtreiber, T. Hayashi, E.A. Dale, D.L. Faustman, Central role of defective apoptosis in autoimmunity, J. Mol. Endocrinol. 31 (3) (2003) 373–399.
[29] A. Okuma, K. Hoshino, T. Ohba, S. Fukushi, S. Aiba, S. Akira, et al., Enhanced apoptosis by disruption of the STAT3-IkappaB-zeta signaling pathway in epithelial cells induces Sjogren's syndrome-like autoimmune disease, Immunity 38 (3) (2013) 450–460.
[30] B.E. Bernstein, T.S. Mikkelsen, X. Xie, et al., A bivalent chromatin structure marks key developmental genes in embryonic stem cells, Cell 125 (2006) 315–326.
[31] A. Barski, S. Cuddapah, K. Cui, et al., High-resolution profiling of histone methylations in the human genome, Cell 129 (2007) 823–837.
[32] N.D. Heintzman, R.K. Stuart, G. Hon, et al., Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome, Nat. Genet. 39 (2007) 311–318.
[33] R.C. Rao, Y. Dou, Hijacked in cancer: the KMT2 (MLL) family of methyltransferases, Nat. Rev. Cancer 15 (2015) 334–346.
[34] A.D.C. Paulussen, A.P.A. Stegmann, M.J. Blok, et al., MLL2 mutation spectrum in 45 patients with Kabuki syndrome, Hum. Mutat. 32 (2011) E2018–E2025.
[35] A.W. Lindsley, H.M. Saal, T.A. Burrow, et al., Defects of B-cell terminal differentiation in patients with type-1 Kabuki syndrome, J. Allergy Clin. Immunol. 137 (1) (2016) 179–187.

[36] J.A. Daniel, A. Nussenzweig, The AID-induced DNA damage response in chromatin, Mol. Cell 50 (2013) 309–321.
[37] Y. Ichikawa, M. Bayeva, M. Ghanefar, et al., Disruption of ATP-binding cassette B8 in mice leads to cardiomyopathy through a decrease in mitochondrial iron export, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 4152–4157.
[38] A. Sturm, P. Cunningham, M. Dean, The ABC transporter gene family of Daphnia pulex, BMC Genom. 10 (2009) 170.
[39] K. Katakura, H. Fujise, K. Takeda, et al., Overexpression of LaMDR2, a novel multidrug resistance ATP-binding cassette transporter, causes 5-fluorouracil resistance in Leishmania amazonensis, FEBS Lett. 561 (2004) 207–212.
[40] Z. Wan, N. Zhi, S. Wong, et al., Human parvovirus B19 causes cell cycle arrest of human erythroid progenitors via deregulation of the E2F family of transcriptionfactors, J. Clin. Invest 120 (2010) 3530–3544.
[41] D. Longman, N. Hug, M. Keith, et al., DHX34 and NBAS form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and Caenorhabditis elegans, Nucleic Acids Res. 41 (2013) 8319–8331.
[42] Y. Liu, N. Lu, B. Yuan, et al., The interaction between the helicase DHX33 and IPS-1 as a novel pathway to sense double-stranded RNA and RNA viruses in myeloid dendritic cells, Cell Mol. Immunol. 11 (2014) 49–57.
[43] J.M. Anaya, C. Ramirez-Santana, M.A. Alzate, N. Molano-Gonzalez, A. Rojas-Villarraga, The autoimmune ecology, Front. Immunol. 7 (2016) 139, http://dx.doi.org/10.3389/fimmu.2016.00139.