



OPEN ACCESS

EDITED BY

Emiliano Giardina,
University of Rome Tor Vergata, Italy

REVIEWED BY

Félicie Costantino,
Université de Versailles Saint-Quentin-en-
Yvelines, France
Andrea Latini,
University of Rome Tor Vergata, Italy

*CORRESPONDENCE

Valérie Chaudru,
✉ valerie.chaudru@univ-evry.fr

RECEIVED 23 January 2024

ACCEPTED 18 April 2024

PUBLISHED 13 May 2024

CITATION

Veyssiere M, Rodriguez Ordonez MdP,
Chalabi S, Michou L, Cornelis F, Boland A,
Olaso R, Deleuze J-F, Petit-Teixeira E and
Chaudru V (2024), *MYLK*FLNB* and
*DOCK1*LAMA2* gene–gene interactions
associated with rheumatoid arthritis in the focal
adhesion pathway.

Front. Genet. 15:1375036.

doi: 10.3389/fgene.2024.1375036

COPYRIGHT

© 2024 Veyssiere, Rodriguez Ordonez, Chalabi,
Michou, Cornelis, Boland, Olaso, Deleuze,
Petit-Teixeira and Chaudru. This is an open-
access article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

*MYLK*FLNB* and *DOCK1*LAMA2* gene–gene interactions associated with rheumatoid arthritis in the focal adhesion pathway

Maëva Veyssiere¹, Maria del Pilar Rodriguez Ordonez²,
Smahane Chalabi², Laetitia Michou³, François Cornelis⁴,
Anne Boland⁵, Robert Olaso⁵, Jean-François Deleuze⁵,
Elisabeth Petit-Teixeira² and Valérie Chaudru^{1,2*}

¹Institut National de la Santé et de la Recherche Médicale, Université de Paris, Paris, France, ²GenHotel—Univ Evry, University of Paris Saclay, Evry, France, ³Division of Rheumatology, Department of Medicine, CHU de Québec—Université Laval, Québec City, QC, Canada, ⁴Génétique-Oncogénétique Adulte-Prévention, Institut National de la Santé et de la Recherche Médicale, Clermont-Auvergne University and CHU, Clermont-Ferrand, France, ⁵Commissariat à l’Energie Atomique, Centre National de Recherche en Génomique Humaine (CNRGH), Université Paris-Saclay, Evry, France

Rheumatoid arthritis (RA) is a chronic, systemic autoimmune disease caused by a combination of genetic and environmental factors. Rare variants with low predicted effects in genes participating in the same biological function might be involved in developing complex diseases such as RA. From whole-exome sequencing (WES) data, we identified genes containing rare non-neutral variants with complete penetrance and no phenocopy in at least one of nine French multiplex families. Further enrichment analysis highlighted focal adhesion as the most significant pathway. We then tested if interactions between the genes participating in this function would increase or decrease the risk of developing RA disease. The model-based multifactor dimensionality reduction (MB-MDR) approach was used to detect epistasis in a discovery sample (19 RA cases and 11 healthy individuals from 9 families and 98 unrelated CEU controls from the International Genome Sample Resource). We identified 9 significant interactions involving 11 genes (*MYLK*, *FLNB*, *DOCK1*, *LAMA2*, *RELN*, *PIP5K1C*, *TNC*, *PRKCA*, *VEGFB*, *ITGB5*, and *FLT1*). One interaction (*MYLK*FLNB*) increasing RA risk and one interaction decreasing RA risk (*DOCK1*LAMA2*) were confirmed in a replication sample (200 unrelated RA cases and 91 GBR unrelated controls). Functional and genomic data in RA samples or relevant cell types argue the key role of these genes in RA.

KEYWORDS

rare variants, pathway enrichment analysis, gene–gene interaction, rheumatoid arthritis, familial sample

Introduction

Rheumatoid arthritis (RA) is a chronic, systemic autoimmune disease affecting approximately 0.5 per cent of the population worldwide (Almutairi et al., 2021). Although the exact etiology of the disease remains unclear, a combination of genetic and environmental factors seems crucial for its development. Currently, over a hundred genetic loci have been associated with increased RA risk, including the *HLA-DRB1* gene and its shared epitope (SE) alleles (McAllister et al., 2011; van Dongen and Holoshitz, 2017). These discoveries have been made possible, thanks to the evolution of genome analyses, from genome-wide association studies (GWAS) to next-generation sequencing (NGS) technologies, accompanied by a shift in the tested hypothesis from frequent variants alone (McAllister et al., 2011; Ha et al., 2021) to a combination of all frequency variants including rare variants (Diogo et al., 2013; Bang et al., 2014; Mitsunaga et al., 2015; Veyssiere et al., 2019a). However, most of these studies identified association signals due to variants with high deleterious effects. Yet rare variants' accumulation with less important predicted effects in genes participating in the same biological function might be involved in the development of complex diseases such as RA.

These functions might be identified using two types of approaches: (1) a candidate approach for which the user targets a specific biological pathway identified from an *a priori* knowledge (such as pathophysiological mechanisms involved in the studied disease) and (2) a whole-genome approach where all the pathways present in a database are analyzed without *a priori*. Thanks to the small number of pathways to evaluate with the first approach, it is possible to manually curate each gene to ensure the annotation quality. Moreover, the time and resources needed to analyze such a low number of pathways are reduced compared to those required by the whole-genome approach. However, the second approach allows detecting new associations between a biological function and disease or between diseases.

Using a candidate approach, Mitsunaga et al. highlighted an aggregation of rare variants, with moderate-to-high impact, in the Japanese population affected by RA, in genes participating in the mitochondrial respiratory chain (Mitsunaga et al., 2015). A burden test combining all variants showed the association with severe erosion (OR = 2.16 [1.43–3.28] and p -value = 1.56×10^{-4}). More recently, an analysis of 58 cases and 66 controls from the Chinese population highlighted the enrichment of rare and frequent variants associated with RA in several biological pathways, including extracellular matrix receptor interaction, protein digestion and absorption, and focal adhesion (Li et al., 2017).

In this study, we used a whole-exome approach to identify biological functions enriched in rare variants co-segregating with RA in multiplex families. In addition, we hypothesized that interactions between the genes participating in these functions would increase the risk of developing the disease. A discovery and a replication step allowed validating this hypothesis.

Material and methods

Sample collection

Blood samples were obtained from 30 individuals belonging to nine French families with multiple cases of autoimmune diseases

(Supplementary Figure S1). Among the 30 individuals, 19 were RA cases (26.3% had Sjögren's syndrome in addition to RA; the others were RA cases without any other autoimmune disease), and 11 were unaffected with RA (27.3% of them had another autoimmune disease—details in Supplementary Figure S1). Data of additional healthy controls were downloaded from the International Genome Sample Resource (IGSR): 98 CEU (Utah residents with Northern and Western European ancestry) and 91 GBR (British in England and Scotland) individuals sequenced on an Illumina platform (Auton et al., 2015; Sudmant et al., 2015). The mean sequencing depth for the CEU and GBR samples is 60X. For the replication step, 200 RA samples from trio families of French European ethnicity (Michou et al., 2007) were used. The study was approved by the Ethics Committees of Hôpital Bicêtre and Hôpital Saint Louis (Paris, France; CPPRB 94–40). All subjects provided written informed consent for participation in the study.

Variant identification

Whole-exome sequencing (WES) performed on the genomic DNA samples obtained from the nine French multiplex families, and subsequent data processing was detailed in a previous study (Veyssiere et al., 2019a). The mean sequencing depth for these samples was 100X. To summarize, the exons were captured with Agilent SureSelect Human All Exon kit (V5), sequenced on an Illumina HiSeq2000 platform, and mapped to the human reference genome hg19 (Lander et al., 2001) using the Burrows–Wheeler alignment-maximal exact match (BWA-MEM) algorithm (H. Li, 2013). Duplicated reads were removed using the Picard toolkit, and variants were called with the Haplotype Caller (HC) algorithm from the Genome Analysis ToolKit (GATK) suite (McKenna et al., 2010; Van der Auwera et al., 2013). Only single-nucleotide variants (SNVs) and small indels (maximum length of 50 bp) with a total DP (read depth) ≥ 12 , mapping quality (MQ) ≥ 30 , quality by depth (QD) ≥ 2 , FS score (phred-scaled p -value using Fisher's exact test to detect strand bias) ≤ 25 , and call-rate $\geq 95\%$ were kept.

Variants were classified as rare when their MAF was below 1% in population of European ancestry in the 1000 Genomes Project (Phase 3, 2015 August), the Exome Aggregation Consortium project (ExAC), the Exome Sequencing project (ESP6500–6500 exomes), and the Complete Genomics project (CG69–69 individuals).

Pathway identification

Over-representation analysis was applied to genes containing rare non-neutral (SNPEFF annotation $\in \{\text{HIGH, MODERATE, and MODIFIER}\}$ (Cingolani et al., 2012) and/or Phred CADDscore ≥ 15 (Rentzsch et al., 2019)) variants with complete penetrance and no phenocopy in at least one of the nine multiplex families. It was performed using ClueGO (Bindea et al., 2009) in Cytoscape version 3.6 (Shannon et al., 2003) for two types of annotations: the Gene Ontology “biological processes” (GO BPs) and pathway databases including Reactome (Gillespie et al., 2022), KEGG pathways (Kanehisa and Goto, 2000), and WikiPathways (Kutmon et al., 2016). Gene/term annotations inferred by computational

processes only (IEA for Inferred from Electronic Annotation) were removed from the analysis. p -values were calculated using a unilateral hypergeometric test and corrected for multiple tests with the Benjamini–Hochberg procedure. Default clustering parameters in ClueGO were kept.

Statistical epistasis analyses

Interactions were first tested for each pair of variants fulfilling the following quality control criteria: (1) missing genotype rate below 5%, (2) no deviation from Hardy-Weinberg equilibrium in controls (p -value >0.01), and (3) no linkage disequilibrium between selected variants ($R^2 < 0.2$). The pairwise interactions were tested using model-based multifactor dimensionality reduction (MB-MDR) (Calle et al., 2010) by applying an additive genetic model while adjusting for sex (as this variable is the only one available for all samples). This method allows identifying genotypic combinations that significantly increase RA risk, grouped under the label H, and genotypic combinations that decrease RA risk, grouped under the label L. The genotypic combinations are pooled together for each category, H and L, to test their association with RA. For pairs of variants having at least one category with a false discovery rate (FDR) below 5%, the significance of the categories was evaluated with a permutation test (up to one million permutations). To assess if an interaction model was more probable than a model including the variant's proper effect only (model "VAR1" for variant 1 and model "VAR2" for variant 2) or a model considering an independent effect of the two variants (model "VAR 1 + VAR2"), all the models were tested and compared using the Akaike information criterion (AIC). The AIC estimates the quality of statistical models and highlights the one with the best combination between the maximum likelihood and the minimum number of parameters k , according to the parsimony principle. The AIC was corrected (AIC_c) to consider the small number of samples in the dataset by applying the following formula, where n is the sample size:

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1}$$

The models "VAR1," "VAR2," and "VAR1 + VAR2" were evaluated using a logistic regression adjusted on the sex with R software version 3.2.3 (R Core Team, 2022). The interaction model was retained if the differences of the AIC with each of the three other models (ΔAIC) were higher than 10 according to the Burnham and Anderson rule for interpreting the ΔAIC scores (Burnham and Anderson, 2002).

SNP genotyping in the replication dataset

Applied Biosystems TaqMan™ SNP Genotyping Assays were used to amplify and detect polymorphisms in purified genomic DNA samples during replication. References of each assay for genotyping are indicated in Supplementary Table S3. For quality control, 10% of samples were randomly chosen to be genotyped a second time.

Results

Accumulation of RA-specific rare variants in a pathway regulating focal adhesion

To identify biological processes that may be relevant for RA pathogenesis, we performed enrichment analysis using rare variants with complete penetrance and no phenocopy in nine multiplex families. From 1,730 rare RA-specific SNVs, 403 were non-neutral (386 SNVs and 17 indels). These variants were significantly enriched for six biological pathways (Figure 1; Supplementary Table S1). The most significant group referred to the focal adhesion pathway (FDR = 0.0054). However, it should be noted that some genes are more tolerant to non-neutral variants (e.g., with high CADD scores). Thus, the latest can accumulate even in a healthy population. To verify if the identified functions were not obtained randomly, we drew 100 sets of 403 non-neutral variants from the GnomAD healthy population and performed an enrichment analysis for each of them. Pathways associated with signaling by Rho GTPases and lipid metabolism were the only ones with significant enrichment of non-neutral variants (with, respectively, 4/100 and 1/100 significant enrichments), reinforcing our results on RA patients. The enrichment analysis performed on the GO database highlighted 22 groups of annotations, including 49 terms, significantly enriched in the list of selected genes (FDR $<5\%$). Multiple functions linked to the regulation of T lymphocytes (differentiation, proliferation, and activation) and focal adhesion regulation were identified (Supplementary Table S2).

Epistatic interactions associated with RA in the focal adhesion pathway

We showed an accumulation of RA rare variants in the focal adhesion pathway and hypothesized that the presence of some combination of variants in the genes participating in this pathway may increase, or reduce, the risk of developing RA. To verify this, we performed a statistical interaction analysis. We tested pairwise interactions between variants, rare and frequent, localized on the focal adhesion pathway genes using our sequenced samples and 98 CEU controls. The effect of each combination of variants on RA was estimated using logistic regression models. Considering the definition of the focal adhesion pathway in WikiPathways, composed of 202 genes, we identified 1,027 SNVs present in at least one patient or control. For the 1,000 most significant H or L categories, we applied one million permutations to assess the significance of the interaction. One hundred and forty-four pairs of variants were identified as significantly associated with RA (FDR $<5\%$). For each of these pairs, we checked if the interaction model was better than a model where only each variant's specific effect was considered and a model where a combined effect of the two variants was considered without interaction. We identified 52 interaction models which were better than the other models. Among those, nine interaction models had a $\Delta AIC >10$ with all the competing models. These interactions involved 14 SNVs (Table 1).

TABLE 1 Significant pairwise interactions of variants associated with RA in the focal adhesion pathway using 30 individuals (19 RA and 11 controls) from multiplex families and 98 CEU controls from IGSR.

Combination		Genotype			Frequency		MB-MDR		ΔAIC with		
Gene1 × Gene2	SNV1 × SNV2	SNV1	SNV2 ^a	Cat	Cases	Controls	OR [95% CI]	FDR ^b	SNV1	SNV2	SNV1 + SNV2
<i>MYLK*FLNB</i>	rs2605418*rs1658338	1	1	H	0.53	0.092	18.89 [5.89–60.54]	7.62E-04	-36.85	-39.78	-38.56
		0	2		0.21	0.046					
<i>DOCK1*LAMA2</i>	rs11017150*rs2244008	0	0	L	0.21	0.78	0.07 [0.022–0.24]	6.00E-06	-26.21	-33.89	-26.49
		0	1	H	0.47	0.094	18.3 [5.36–61.40]	3.13E-03	-30.81	-38.49	-31.09
		1	0		0.032	0.074					
<i>RELN*MYLK</i>	rs2229864*rs55956729	2	0	H	0.32	0.037	14.88 [4.72–46.87]	3.87E-03	-33.86	-33.73	-35.75
		0	1		0.26	0.046					
<i>PIP5K1C*FLNB</i>	rs542690*rs1658338	0	1	H	0.53	0.14	19.38 [5.92–65.62]	1.88E-03	-37.01	-27.58	-24.11
		0	2		0.26	0.028					
<i>TNC*LAMA2</i>	rs2989519*rs2244008	0	1	H	0.16	0.018	13.37 [4.3–41.35]	6.47E-03	-28.15	-32.62	-24.71
		1	1		0.21	0.0092					
		2	0		0.26	0.11					
		0	0	L	0.052	0.29	0.078 [0.023–0.26]	4.00E-06	-27.49	-31.96	-24.05
		1	0		0.16	0.47					
<i>PRKCA*RELN</i>	rs3803821*rs262338	0	1	H	0.42	0.073	14.23 [4.54–44.56]	4.95E-03	-32.17	-27.93	-27.34
		1	0		0.32	0.10					
<i>VEGFB*LAMA2</i>	rs11603042*rs2244008	0	1	H	0.32	0.37	17.05 [4.68–62.02]	1.47E-02	-21.58	-29.91	-22.98
		2	1		0.11	0.0092					
<i>ITGB5*FLNB</i>	rs2291079*rs1658338	0	1	H	0.58	0.17	15.24 [4.59–50.65]	8.21E-03	-29.81	-30.70	-29.72
		0	2		0.2	0.046					
<i>FLT1*LAMA2</i>	rs2296189*rs3778137	1	0	H	0.47	0.073	12.09 [3.77–38.76]	2.20E-02	-23.45	-26.31	-22.63

^aCategory assigned to the genotype combination; H: increased RA risk and L: reduced RA risk.

^bFDR calculated by applying the Benjamini–Hochberg procedure.

Replicated interactions between MYLK and FLNB and DOCK1 and LAMA2

We selected the five most significant gene–gene interactions (FDR <0.004—Table 1) for replication. For this purpose, we genotyped 8 participating SNVs on 200 RA patients. After adding 91 GBR controls from the IGSR, we applied the same strategy used in the discovery phase. Three interactions, *MYLK*FLNB*, *DOCK1*LAMA2*, and *PIP5K1C*FLNB*, were significantly associated with RA in the replication dataset (Table 2). One high-risk combination for *MYLK*FLNB* and two for *DOCK1*LAMA2* were more probable than the concurrent models (variant proper effect and additive without interaction) in both initial and replication sets (bold in Table 2). No genotypic combination was associated with RA in the replication dataset (minimum *p*-value = 0.14) for *RELN*MYLK*. The significant combinations identified for *PIP5K1C*FLNB* differed from those identified in the multiplex families of the discovery set. As all replicated interactions harbored a ΔAIC >10, it strengthens our findings.

Discussion

Using whole-exome sequences from patients and healthy relatives belonging to multiplex families, we showed an accumulation of rare RA-associated variants in the genes participating in four biological functions: regulation of focal adhesions, the response to interferons, signaling through Rho GTPases, and lipid homeostasis. In addition, we report epistatic relationships between four genes whose products are involved in focal adhesion, the most significant pathway.

Although some uncertainty could be raised regarding the enrichment analysis results because of the nature of the variants considered (for example, intronic variant or SNV with a low prediction effect), the identification of functions previously established as participating in RA physiopathology supports the validity of the findings. Previous studies have described an increase in the type I interferon response in RA patients even at an early stage of the disease (Castañeda-Delgado et al., 2017; Muskardin and Niewold, 2018) and an association with response to therapy (Cooles et al., 2018;

TABLE 2 Pairwise interactions in 200 RA cases and 91 GBR controls.

Combination		Genotype			Frequency		MB-MDR		ΔAIC with		
Gene1 × Gene2	SNV1 × SNV2	SNV1	SNV2	Cat ^a	Cases	Controls	OR [95% CI]	p-value ^b	SNV1	SNV2	SNV1 + SNV2
<i>MYLK*FLNB</i>	rs2605418*rs1658338	1	1	H	0.38	0.15	2.70 [1.36–5.38]	0.0184	–36.47	–35.01	–34.28
		0	0	L	0.15	0.25	0.38 [0.20–0.71]	0.0302	–36.42	–34.97	–34.23
		0	2		0.02	0.08					
<i>DOCK1*LAMA2</i>	rs11017150*rs2244008	0	0	L	0.38	0.71	0.21 [0.12–0.38]	4.12E-06	–29.93	–5.401	–30.08
		1	0	H	0.30	0.14	4.93 [2.48–9.78]	1.98E-05	–25.92	–49.99	–26.06
		1	1		0.17	0.01					
<i>PIP5K1C*FLNB</i>	rs542690*rs1658338	1	1	H	0.34	0.18	2.72 [1.39–5.34]	0.0185	–37.67	–35.28	–36.19
		2	0	L	0.06	0.12	0.29 [0.13–0.63]	0.0333	–37.94	–35.55	–36.46
		1	2		0.03	0.08					

^aCategory assigned to the genotype combination; H: increased RA risk and L: reduced RA risk.

^bp-value obtained from one million of permutations.

In bold: interactions replicated with categories identified in the initial sample.

2022; de Jong et al., 2019). In addition, Rho GTPase signaling is essential for several mechanisms involved in RA, from the cytokine production by immune cells to the dynamics of joint structural cells such as synoviocytes and osteoclasts (Nakayamada et al., 2005; Zeng et al., 2022). Recently, in a study analyzing the relationship between non-synonymous SNVs identified by WES in RA patients and disease activity, Chen and others identified key genes related to lipid metabolism through protein–protein interaction (PPI) results (Chen et al., 2023). In addition to the four biological functions, the enrichment analysis identified pathways associated with other disorders, including diseases associated with protein glycosylation and pathways involved in the pathogenesis of cardiovascular diseases (CVDs). Interestingly, CVD risk is substantially increased in RA (van Halm et al., 2009). Furthermore, to assess the causality of RA association with coronary artery disease (CAD), a subtype of CVD, a recent Mendelian randomization approach performed using European data showed that genetic liability to RA was associated with an increased risk of CAD (Qiu et al., 2021; Yuan et al., 2022). These results support the interest in studying shared risk factors between RA and CVDs.

The most significant result revealed the aggregation of RA rare variants in the pathway regulating the dynamics of focal adhesions. This pathway has already been identified from WES comparing 56 RA and 66 healthy controls in the Han Chinese population (Y. Li et al., 2017). These structures connecting the proteins from the extracellular matrix (ECM) to actin filaments of the cell cytoskeleton are involved in cell migration, differentiation, and survival. Their role in the pathogenesis of the disease has been supported by different studies that showed, for example, the requirement of focal adhesion kinase for synovial fibroblast invasion in RA patients (Shelef et al., 2014) and the role of vascular cell adhesion molecules in the interaction of synovial fibroblasts with immune cells (Postigo et al., 1992; Yoon and Moon, 2021). These observations and the results of our analysis lead us to further study this pathway and search for possible genetic interactions modulating the risk of developing RA. Note that this pathway has also been identified in other autoimmune diseases: in osteoarthritis and systemic lupus erythematosus from differentially expressed genes (Tang

et al., 2012; He et al., 2016) and in type 1 diabetes from genes targeted by circulating miRNAs (Satake et al., 2018).

The MB-MDR approach was used to test the GxG interaction in the latter pathway since it has some advantages. It has been shown that MB-MDR has sufficient power to detect epistasis even if the data suffer from drawbacks such as missing data, genotyping error, genetic heterogeneity, population structure, or low sample size (Cattaert et al., 2011; Mahachie John et al., 2011; Abegaz et al., 2021). It also allows taking into account rare variants in the interaction test (Fouladi et al., 2015). Recently, this approach has been extended to empower individual risk prediction in personalized medicine. Simulation studies have demonstrated that this new algorithm outperforms random forest and elastic net approaches (Gola and König, 2021). We thus decided to use this method for our familial sample. Due to the small number of relatives in each family, relatedness was not considered in the analysis of the discovery set. Nevertheless, the replication set, consisting of unrelated cases and controls, leads to the validation of some interactions.

This statistical method allowed us to identify epistasis between two pairs of genes: *MYLK*FLNB* and *DOCK1*LAMA2*. The question of relation between *HLA-DRB1* shared epitope alleles and these interactions is arising in the context of RA patient carriers. When using different bioinformatics tools—HLA reporter (Huang et al., 2015) and SNP2HLA (Jia et al., 2013)—to investigate the *HLA-DRB1* status in controls, no conclusive results were obtained, essentially due to data quality (phasing issues and insufficient read depth). Consequently, the question could not be addressed (data not shown).

Concerning the SNVs involved in significant interactions, rs2244008, located in *LAMA2*, is a missense variant, whose effect is predicted to be low. Variants rs2605418 and rs1658338, respectively, located in *MYLK* and *FLNB*, are intronic. SNV rs11017150 is in a splice region of the *DOCK1* gene. In reviewing the functional impact of these four variants, data collected through the GTEx project highlighted significant quantitative trait loci (QTLs) for two SNVs out of four. However, correlated expression profiles are not suggestive of RA-specific tissue or functions. SNV rs2605418 is the only one of the

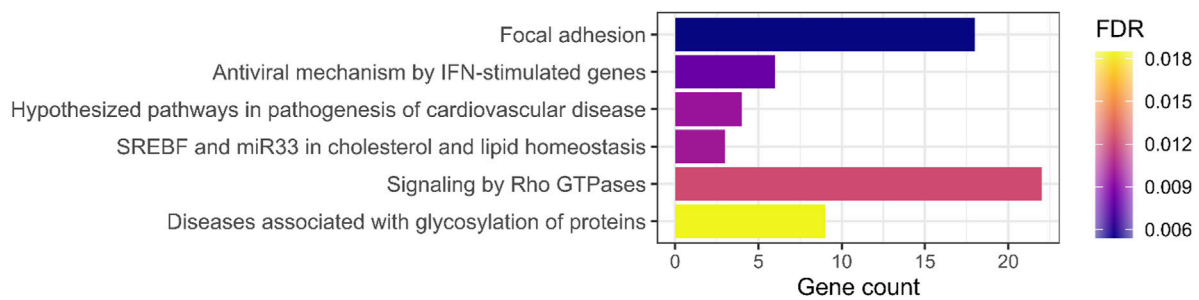


FIGURE 1

Enrichment analysis results of rare variants carried by all RA cases in at least one multiplex family. Barplot representing the number of genes carrying a rare variant (with complete penetrance and no phenocopy in nine multiplex families) by the significant biological pathway (FDR <5%). The color represents the false discovery rate (FDR) value (Benjamini–Hochberg procedure) of the enrichment analysis.

four to be in a region of seven SNVs with strong linkage disequilibrium in the CEU population. The functional impact reported in the GTEx database for each SNV in this LD block is like rs2605418's one (<https://www.ebi.ac.uk/gwas/publications/36224396>). Finally, none of them were associated with RA or a related phenotype through GWAS results. However, concerning *FLNB* (*filamin B*) gene, frequent variants have been associated with characteristics of bone structure in women (Wilson et al., 2009), which could be related to RA. From a functional point of view, the *LAMA2* gene encodes for the sub-unit α of laminins 2 and 4, which are proteins of the ECM mainly expressed in epithelial cells. The role of these proteins in RA pathophysiology has not been defined yet. Nevertheless, *LAMA2* is hypo-methylated (Whitaker et al., 2013) and over-expressed (Sun et al., 2019; Zhang et al., 2019) in fibroblast-like synoviocytes of RA patients compared to patients affected by other arthritic or autoimmune diseases. Interestingly, the interaction between the *LAMA2* protein and another protein of focal adhesions, namely, dedicator of cytokine 1 (*DOCK1*), promotes the migration of fibroblast-like synoviocytes, essential cells in RA, via *RAC1* and *ELMO1* gene's product (Whitaker et al., 2015). This functional relation between *DOCK1* and *LAMA2* could be an illustration of the genetic interaction we observed.

To conclude, this interaction analysis allows identifying new genetic risk factors contributing to missing heritability in RA. In particular, we identified interactions implicating four genes of the focal adhesion pathways. Additional functional and omic evidence of the potential role of *FLNB*, *LAMA2*, and *DOCK1* exists to argue for further investigations.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: <https://www.ebi.ac.uk/eva/?eva-study=PRJEB57722>.

Ethics statement

The studies involving humans were approved by the Hôpital Bicêtre and Hôpital Saint Louis (Paris, France; CPPRB 94–40). The studies were conducted in accordance with the local legislation and

institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MV: investigation, validation, writing–original draft, writing–review and editing, conceptualization, data curation, formal analysis, and methodology. MR-O: formal analysis, investigation, and writing–review and editing. SC: formal analysis, investigation, and writing–review and editing. LM: data curation, resources, and writing–review and editing. FC: data curation, resources, and writing–review and editing. AB: data curation, resources, and writing–review and editing. RO: data curation, resources, and writing–review and editing. JD: funding acquisition, resources, and writing–review and editing. EP-T: funding acquisition, investigation, resources, supervision, validation, writing–original draft, and writing–review and editing. VC: conceptualization, funding acquisition, investigation, methodology, supervision, validation, writing–original draft, and writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was partially supported by the ARTHRITIS Fondation COURTIN (<http://www.fondation-arthritisme.org/>) and by the Genopole (<https://www.genopole.fr/>). Both were granted to EP-T. This work was partially supported by the Fondation pour la Recherche Médicale, grant number SPF201909009302, to MV. The sequencing platform was supported by the France Génomique National infrastructure, funded as part of the « Investissements d'Avenir » program managed by the Agence Nationale pour la Recherche (contract ANR-10-INBS-09).

Acknowledgments

The authors thank Fatiha Benmebkhout for her technical support, and they are grateful to RA patients and their families for their participation in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2024.1375036/full#supplementary-material>

References

- Abegaz, F., Van Lishout, F., Mahachie John, J. M., Chichoempu, K., Bhardwaj, A., Duroux, D., et al. (2021). Performance of model-based multifactor dimensionality reduction methods for epistasis detection by controlling population structure. *BioData Min.* 14 (1), 16. doi:10.1186/s13040-021-00247-w
- Almutairi, K., Nossent, J., Preen, D., Keen, H., and Inderjeeth, C. (2021). The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. *Rheumatol. Int.* 41 (5), 863–877. doi:10.1007/s00296-020-04731-0
- Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., et al. (2015). A global reference for human genetic variation. *Nature* 526 (7571), 68–74. doi:10.1038/nature15393
- Bang, S.-Y., Na, Y.-J., Kim, K., Joo, Y. B., Park, Y., Lee, J., et al. (2014). Targeted exon sequencing fails to identify rare coding variants with large effect in rheumatoid arthritis. *Arthritis Res. Ther.* 16 (5), 447. doi:10.1186/s13075-014-0447-7
- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinforma. Oxf. Engl.* 25 (8), 1091–1093. doi:10.1093/bioinformatics/btp101
- Burnham, K. P., and Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. New-York: Springer-Verlag.
- Calle, M. L., Urrea, V., Malats, N., and Van Steen, K. (2010). mbmdr: an R package for exploring gene-gene interactions associated with binary or quantitative traits. *Bioinforma. Oxf. Engl.* 26 (17), 2198–2199. doi:10.1093/bioinformatics/btq352
- Castañeda-Delgado, J. E., Bastián-Hernandez, Y., Macias-Segura, N., Santiago-Algarra, D., Castillo-Ortiz, J. D., Alemán-Navarro, A. L., et al. (2017). Type I interferon gene response is increased in early and established rheumatoid arthritis and correlates with autoantibody production. *Front. Immunol.* 8, 285. doi:10.3389/fimmu.2017.00285
- Cattaert, T., Calle, M. L., Dudek, S. M., Mahachie John, J. M., Van Lishout, F., Urrea, V., et al. (2011). Model-based multifactor dimensionality reduction for detecting epistasis in case-control data in the presence of noise. *Ann. Hum. Genet.* 75 (1), 78–89. doi:10.1111/j.1469-1809.2010.00604.x
- Chen, J., Li, S., Zhu, J., Su, W., Jian, C., Zhang, J., et al. (2023). Multi-omics profiling reveals potential alterations in rheumatoid arthritis with different disease activity levels. *Arthritis Res. Ther.* 25 (1), 74. doi:10.1186/s13075-023-03049-z
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6 (2), 80–92. doi:10.4161/fly.19695
- Cooles, F. A. H., Anderson, A. E., Lendrem, D. W., Norris, J., Pratt, A. G., Hilken, C. M. U., et al. (2018). The interferon gene signature is increased in patients with early treatment-naïve rheumatoid arthritis and predicts a poorer response to initial therapy. *J. Allergy Clin. Immunol.* 141 (1), 445–448. doi:10.1016/j.jaci.2017.08.026
- Cooles, F. A. H., Tarn, J., Lendrem, D. W., Naamane, N., Lin, C. M., Millar, B., et al. (2022). Interferon- α -mediated therapeutic resistance in early rheumatoid arthritis implicates epigenetic reprogramming. *Ann. Rheumatic Dis.* 81, 1214–1223. doi:10.1136/annrheumdis-2022-222370
- de Jong, T. D., Snoek, T., Mantel, E., van der Laken, C. J., van Vollenhoven, R. F., and Lems, W. F. (2019). Dynamics of the type I interferon response during immunosuppressive therapy in rheumatoid arthritis. *Front. Immunol.* 10, 902. doi:10.3389/fimmu.2019.00902
- Diogo, D., Kurreeman, F., Stahl, E. A., Liao, K. P., Gupta, N., Greenberg, J. D., et al. (2013). Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis. *Am. J. Hum. Genet.* 92 (1), 15–27. doi:10.1016/j.ajhg.2012.11.012
- Fouladi, R., Bessonov, K., Van Lishout, F., and Van Steen, K. (2015). Model-Based multifactor dimensionality reduction for rare variant association analysis. *Hum. Hered.* 79 (3–4), 157–167. doi:10.1159/000381286
- Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., et al. (2022). The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* 50 (D1), D687–D692. doi:10.1093/nar/gkab1028
- Gola, D., and König, I. R. (2021). Empowering individual trait prediction using interactions for precision medicine. *BMC Bioinforma.* 22 (1), 74. doi:10.1186/s12859-021-04011-z
- Ha, E., Bae, S.-C., and Kim, K. (2021). Large-scale meta-analysis across East Asian and European populations updated genetic architecture and variant-driven biology of rheumatoid arthritis, identifying 11 novel susceptibility loci. *Ann. Rheumatic Dis.* 80 (5), 558–565. doi:10.1136/annrheumdis-2020-219065
- He, P., Zhang, Z., Liao, W., Xu, D., Fu, M., and Kang, Y. (2016). Screening of gene signatures for rheumatoid arthritis and osteoarthritis based on bioinformatics analysis. *Mol. Med. Rep.* 14 (2), 1587–1593. doi:10.3892/mmr.2016.5423
- Huang, Y., Yang, J., Ying, D., Zhang, Y., Shotelersuk, V., Hirankarn, N., et al. (2015). HLAREporter: a tool for HLA typing from next generation sequencing data. *Genome Med.* 7 (1), 25. doi:10.1186/s13073-015-0145-3
- Jia, X., Han, B., Onengut-Gumuscu, S., Chen, W.-M., Concannon, P. J., Rich, S. S., et al. (2013). Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One* 8 (6), e64683. doi:10.1371/journal.pone.0064683
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. Article 1. doi:10.1093/nar/28.1.27
- Kutmon, M., Riutta, A., Nunes, N., Hanspers, K., Willighagen, E. L., Bohler, A., et al. (2016). WikiPathways: capturing the full diversity of pathway knowledge. *Nucleic Acids Res.* 44 (D1), D488–D494. Article D1. doi:10.1093/nar/gkv1024
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409 (6822), 860–921. doi:10.1038/35057062
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. doi:10.48550/ARXIV.1303.3997
- Li, Y., Lai-Han Leung, E., Pan, H., Yao, X., Huang, Q., Wu, M., et al. (2017). Identification of potential genetic causal variants for rheumatoid arthritis by whole-exome sequencing. *Oncotarget* 8 (67), 111119–111129. doi:10.18632/oncotarget.22630
- Mahachie John, J. M., Van Lishout, F., and Van Steen, K. (2011). Model-Based Multifactor Dimensionality Reduction to detect epistasis for quantitative traits in the presence of error-free and noisy data. *Eur. J. Hum. Genet. EJHG* 19 (6), 696–703. doi:10.1038/ejhg.2011.17
- McAllister, K., Eyre, S., and Orozco, G. (2011). Genetics of rheumatoid arthritis: GWAS and beyond. *Open Access Rheumatology Res. Rev.* 3, 31–46. doi:10.2147/OARRR.S14725
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20 (9), 1297–1303. doi:10.1101/gr.107524.110
- Michou, L., Lasbleiz, S., Rat, A.-C., Migliorini, P., Balsa, A., Westhovens, R., et al. (2007). Linkage proof for PTPN22, a rheumatoid arthritis susceptibility gene and a human autoimmunity gene. *Proc. Natl. Acad. Sci. U. S. A.*, 104(5), 1649–1654. doi:10.1073/pnas.0610250104
- Mitsunaga, S., Hosomichi, K., Okudaira, Y., Nakaoka, H., Suzuki, Y., Kuwana, M., et al. (2015). Aggregation of rare/low-frequency variants of the mitochondria respiratory chain-related proteins in rheumatoid arthritis patients. *J. Hum. Genet.* 60 (8), 449–454. doi:10.1038/jhg.2015.50
- Muskardin, T. L. W., and Niewold, T. B. (2018). Type I interferon in rheumatic diseases. *Nat. Rev. Rheumatol.* 14 (4), 214–228. doi:10.1038/nrrheum.2018.31
- Nakayamada, S., Kurose, H., Saito, K., Mogami, A., and Tanaka, Y. (2005). Small GTP-binding protein Rho-mediated signaling promotes proliferation of rheumatoid synovial fibroblasts. *Arthritis Res. Ther.* 7 (3), R476–R484. doi:10.1186/ar1694

- Postigo, A. A., Garcia-Vicuña, R., Diaz-Gonzalez, F., Arroyo, A. G., Landázuri, M. O. D., Chi-Rosso, G., et al. (1992). Increased binding of synovial T lymphocytes from rheumatoid arthritis to endothelial-leukocyte adhesion molecule-1 (ELAM-1) and vascular cell adhesion molecule-1 (VCAM-1). *J. Clin. Investigation* 89 (5), 1445–1452. doi:10.1172/JCI115734
- Qiu, S., Li, M., Jin, S., Lu, H., and Hu, Y. (2021). Rheumatoid arthritis and cardiovascular disease: a mendelian randomization study. *Front. Genet.* 12, 745224. doi:10.3389/fgene.2021.745224
- R Core Team (2022). R: a language and environment for statistical computing. R foundation for statistical computing. Available at: <https://www.R-project.org/>.
- Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., and Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 47 (D1), D886–D894–D894. doi:10.1093/nar/gky1016
- Satake, E., Pezzolesi, M. G., Md Dom, Z. I., Smiles, A. M., Niewczas, M. A., and Krolewski, A. S. (2018). Circulating miRNA profiles associated with hyperglycemia in patients with type 1 diabetes. *Diabetes* 67 (5), 1013–1023. doi:10.2337/db17-1207
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. Article 11. doi:10.1101/gr.1239303
- Shelef, M. A., Bennin, D. A., Yasmin, N., Warner, T. F., Ludwig, T., Beggs, H. E., et al. (2014). Focal adhesion kinase is required for synovial fibroblast invasion, but not murine inflammatory arthritis. *Arthritis Res. Ther.* 16 (5), 464. doi:10.1186/s13075-014-0464-6
- Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., et al. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature* 526 (7571), 75–81. doi:10.1038/nature15394
- Sun, Z., Wang, W., Yu, D., and Mao, Y. (2019). Differentially expressed genes between systemic sclerosis and rheumatoid arthritis. *Hereditas* 156 (1), 17. doi:10.1186/s41065-019-0091-y
- Tang, Y., Ma, X., Zhang, H., Gu, Z., Hou, Y., Gilkeson, G. S., et al. (2012). Gene expression profile reveals abnormalities of multiple signaling pathways in mesenchymal stem cell derived from patients with systemic lupus erythematosus. *Clin. Dev. Immunol.* 2012, 826182. doi:10.1155/2012/826182
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* 43 (1110), 11. doi:10.1002/0471250953.bi1110s43
- van Drongelen, V., and Holoshitz, J. (2017). Human leukocyte antigen-disease associations in rheumatoid arthritis. *Rheumatic Dis. Clin. N. Am.* 43 (3), 363–376. doi:10.1016/j.rdc.2017.04.003
- van Halm, V. P., Peters, M. J. L., Voskuyl, A. E., Boers, M., Lems, W. F., Visser, M., et al. (2009). Rheumatoid arthritis versus diabetes as a risk factor for cardiovascular disease: a cross-sectional study, the CARRÉ Investigation. *Ann. Rheumatic Dis.* 68 (9), 1395–1400. doi:10.1136/ard.2008.094151
- Veyssiere, M., Perea, J., Michou, L., Boland, A., Caloustian, C., Olasso, R., et al. (2019a). A novel nonsense variant in SUPT20H gene associated with Rheumatoid Arthritis identified by Whole Exome Sequencing of multiplex families. *PLoS One* 14 (3), e0213387. doi:10.1371/journal.pone.0213387
- Whitaker, J. W., Boyle, D. L., Bartok, B., Ball, S. T., Gay, S., Wang, W., et al. (2015). Integrative omics analysis of rheumatoid arthritis identifies non-obvious therapeutic targets. *PLoS ONE* 10 (4), e0124254. doi:10.1371/journal.pone.0124254
- Whitaker, J. W., Shoemaker, R., Boyle, D. L., Hillman, J., Anderson, D., Wang, W., et al. (2013). An imprinted rheumatoid arthritis methylome signature reflects pathogenic phenotype. *Genome Med.* 5 (4), 40. doi:10.1186/gm444
- Wilson, S. G., Jones, M. R., Mullin, B. H., Dick, I. M., Richards, J. B., Pastinen, T. M., et al. (2009). Common sequence variation in *FLNB* regulates bone structure in women in the general population and *FLNB* mRNA expression in osteoblasts *in vitro*. *J. Bone Mineral Res.* 24 (12), 1989–1997. doi:10.1359/jbmr.090530
- Yoon, S.-S., and Moon, E.-Y. (2021). B cell adhesion to fibroblast-like synoviocytes is up-regulated by tumor necrosis factor-alpha via expression of human vascular cell adhesion molecule-1 mediated by B cell-activating factor. *Int. J. Mol. Sci.* 22 (13), 7166. Article 13. doi:10.3390/ijms22137166
- Yuan, S., Carter, P., Mason, A. M., Yang, F., Burgess, S., and Larsson, S. C. (2022). Genetic liability to rheumatoid arthritis in relation to coronary artery disease and stroke risk. *Arthritis & Rheumatology* 74, 1638–1647. doi:10.1002/art.42239
- Zeng, R., Zhuo, Z., Luo, Y., Sha, W., and Chen, H. (2022). Rho GTPase signaling in rheumatic diseases. *iScience* 25 (1), 103620. doi:10.1016/j.isci.2021.103620
- Zhang, F., Wei, K., Slowikowski, K., Fonseka, C. Y., Rao, D. A., Kelly, S., et al. (2019). Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat. Immunol.* 20 (7), 928–942. Article 7. doi:10.1038/s41590-019-0378-1