



HAL
open science

A Mechanistic Cellular Atlas of the Rheumatic Joint

Naouel Zerrouk, Sahar Aghakhani, Vidisha Singh, Franck Augé, Anna Niarakis

► **To cite this version:**

Naouel Zerrouk, Sahar Aghakhani, Vidisha Singh, Franck Augé, Anna Niarakis. A Mechanistic Cellular Atlas of the Rheumatic Joint. *Frontiers in Systems Biology*, 2022, 10.3389/fsysb.2022.925791 . hal-03699702

HAL Id: hal-03699702

<https://hal.science/hal-03699702>

Submitted on 21 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Mechanistic Cellular Atlas of the Rheumatic Joint

1 Naouel Zerrouk^{1,2} †, Sahar Aghakhani^{1,3} †, Vidisha Singh¹, Franck Auge² and Anna Niarakis^{1,3} *

2 ¹ Univ. Paris-Saclay, Univ. Evry, GenHotel – Laboratoire Européen de Recherche pour la
3 Polyarthrite rhumatoïde, Evry, France

4 ² Sanofi R&D, AI & Deep Analytics – Omics Data Science, Chilly-Mazarin, France

5 ³ Lifeware Group, Inria Saclay, Palaiseau, France

6 † These authors contributed equally.

7 * **Correspondence:**

8 Anna Niarakis anna.niaraki@univ-evry.fr

9 **Keywords: rheumatoid arthritis, autoimmune disease, molecular interaction maps, systems**
10 **biology graphical notation, bio-curation**

11 Abstract

12 Rheumatoid Arthritis (RA) is an autoimmune disease of unknown aetiology involving complex
13 interactions between environmental and genetic factors. Its pathogenesis is suspected to arise from
14 intricate interplays between signalling, gene regulation and metabolism, leading to synovial
15 inflammation, bone erosion and cartilage destruction in the patients' joints. In addition, the resident
16 synoviocytes of macrophage and fibroblast types can interact with innate and adaptive immune cells
17 and contribute to the disease's debilitating symptoms. Therefore, a detailed, mechanistic mapping of
18 the molecular pathways and cellular crosstalks is essential to understand the complex biological
19 processes and different disease manifestations. In this regard, we present the RA-Atlas, an SBGN-
20 standardized, interactive, manually curated representation of existing knowledge related to the RA's
21 onset and progression. This state-of-the-art RA-Atlas includes an updated version of the global RA-
22 map covering relevant metabolic pathways and cell-specific molecular interaction maps for CD4+
23 Th1 cells, fibroblasts, and M1 and M2 macrophages. The molecular interaction maps were built
24 using information extracted from published literature and pathway databases and enriched using omic
25 data. The RA-Atlas is freely accessible on the webserver MINERVA (<https://ramap.uni.lu/minerva/>),
26 allowing easy navigation using semantic zoom, cell-specific or experimental data overlay, gene set
27 enrichment analysis, pathway export or drug query.

28 1. Introduction

29 RA is an autoimmune, systemic disease affecting approximately 1% of the world population. It is the
30 most common type of autoimmune arthritis and causes pain, swelling, and stiffness in the joints
31 (McInnes and Schett, 2011; Smolen, Aletaha and McInnes, 2016). RA is considered a complex
32 disease and is suspected to arise from several factors (Glocker *et al.*, 2006), including genetic
33 (Korczywska, 2014), environmental (Deane *et al.*, 2017; Sigaux *et al.*, 2019), epigenetic (Frank-
34 Bertonecelj, Klein and Gay, 2017), but also infectious (Balandraud, Roudier and Roudier, 2004) or
35 hormonal factors (Talsania and Scofield, 2017). RA's pathogenesis is governed by complex
36 interactions between components at different scales of the immune system, affecting signalling,

37 metabolic and regulatory networks. Summarizing this scattered and fragmented knowledge in a
38 formal and standardized representation can significantly facilitate understanding the underlying
39 disease mechanisms.

40 In this direction, some efforts have already been made to assemble the available knowledge in
41 molecular interaction maps. In 2010, a comprehensive RA-specific map (Wu *et al.*, 2010) was
42 published based on high throughput experiments (mRNA, miRNA), literature, publicly available
43 datasets and the pathway database KEGG (Ogata *et al.*, 1999) to establish the connections between
44 different components of the map. A decade later, in 2020, a second effort to formalize RA knowledge
45 was published using Wu *et al.*'s map as a basis, the RA-map (Singh *et al.*, 2020). For this map, the
46 researchers used the Systems Biology Graphical Notation (SBGN) standard and, more precisely, the
47 Process Description (PD) language (Le Novère, 2015) to represent the molecular pathways
48 implicated in the disease. The map was built as a global map using information from human-specific
49 studies related to RA, focusing on small-scale experiments in various cell and tissue types (e.g.,
50 fibroblasts, macrophages, synovial tissue, and peripheral blood mononuclear cells (PBMC)). The RA-
51 map was manually curated for each component and reaction, with defined cellular pathways and
52 molecular signatures under expert validation. In addition to manual curation of the scientific
53 literature, other sources such as the Ingenuity Pathway Analysis (IPA) (Krämer *et al.*, 2014),
54 DISNOR (Lo Surdo *et al.*, 2018), or KEGG databases were used to obtain RA related information
55 regarding molecular pathways and phenotypic signatures.

56 While the previously published RA-map offers a formalized representation of the major pathways
57 implicated in the disease, it does not cover metabolic pathways. It has been shown experimentally
58 that unbalanced metabolic pathways accumulate metabolic intermediates, driving immune response
59 to aggravate chronic inflammation (Pucino *et al.*, 2020). Targeting specific metabolic pathways has
60 been seen to reduce inflammation in both in-vitro and in-vivo studies of arthritis (Fearon *et al.*,
61 2019). Recent studies have also highlighted metabolic reprogramming as a possible mechanism
62 involved in disease pathogenesis (Aghakhani, Zerrouk and Niarakis, 2020). Considering these
63 advances, adding metabolic information to the RA-map becomes necessary. In this perspective, we
64 identified four metabolic pathways to include in the RA-map for their involvement in the disease
65 (Fearon *et al.*, 2019), namely glycolysis, pentose phosphate pathway, citric acid cycle and oxidative
66 phosphorylation.

67 Moreover, the RA-map, a global map constructed as a generic blueprint of a cell, contains
68 information from multiple cellular types. In RA, immune cells such as T cells and B cells invade the
69 synovial tissue and produce high amounts of cytokines such as tumour necrosis factor α (TNF α),
70 interferon γ (IFN γ), interleukin (IL)-1 β , IL-6, IL-17, and cause inflammation (Guo *et al.*, 2018).
71 Immune cells also interact with tissue-resident cells (e.g., fibroblasts and macrophages), leading to
72 osteoclasts' proliferation, bone erosion, and cartilage damage (Guo *et al.*, 2018). In addition, dendritic
73 cells produce vascular endothelial growth factors and promote extensive angiogenesis in RA (Bosisio
74 *et al.*, 2018). The representation of cell-specific molecular interactions helps understand the complex
75 interplay between various cells and factors in the inflamed joint. In this direction, we selected three
76 cell types of interest: fibroblast, macrophage (including the pro-inflammatory M1 and the anti-
77 inflammatory M2), and CD4+ T helper cell 1 (Th1). These cell types are abundant in RA patients'
78 joints (Huber *et al.*, 2006; Li *et al.*, 2017), and produce high amounts of cytokines (Bartok and
79 Firestein, 2010; Schulze-Koops and Kalden, 2001), resist apoptosis and proliferate at high rates
80 (Krämer *et al.*, 2014; Yang, Chang and Wei, 2020). The cell type selection was also based on
81 literature availability, including already existing cell-specific maps, specific signalling pathways, and
82 gene expression datasets.

83 This paper introduces an attempt to create a multicellular atlas of the rheumatic joint to recapitulate
84 existing knowledge related to the disease's onset, progression, and pathogenesis. The RA-Atlas
85 includes an updated version of the RA-map: the RA-map V2 (with metabolic pathways' addition), a
86 fibroblast-specific map, two maps for the M1 and M2 macrophages, and a CD4+ Th1 map. Our
87 ambition is to expand the RA-Atlas in the future by including other relevant cell types like B cells,
88 CD4+ regulatory T cells, dendritic cells, chondrocytes, osteoblasts, and osteoclasts.

89 The maps can be used to provide high-quality curated information for disease-related pathways, a
90 template for omic data analysis, and a starting point for dynamic computational models. The
91 following sections outline our efforts to integrate low and high throughput experiments, combining
92 prior knowledge, manual curation of scientific literature, omic data analysis, and systems biology
93 standards to create state-of-the-art mechanistic representations of disease-related pathways.

94 **2. Methodology and Data**

95 **2.1. A Common Strategy for Global and Cell-Specific Maps**

96 **2.1.1. Map Construction**

97 The original RA-map paper thoroughly describes the general methodology for constructing the
98 different RA-Atlas maps (Singh *et al.*, 2020). Briefly, we used the CellDesigner software (Funahashi
99 *et al.*, 2003) to construct the molecular interaction maps in the PD format, one of the three standard
100 languages of SBGN (Le Novère, 2015). The various components of the maps (i.e., phenotype,
101 protein, gene, RNA, simple molecule, ion, complex) were distinguished in the form of specific
102 glyphs and notations according to the said standards. Cellular compartments featured in each map
103 differ according to their specificities. They are designed to reflect the molecular architecture of each
104 map (e.g., extracellular space, plasma membrane, cytoplasm, nucleus, mitochondrion, endoplasmic
105 reticulum, phenotypes, secreted compartment, and transmembrane domain). In addition, HUGO Gene
106 Nomenclature Committee identifiers (HGNC) (Tweedie *et al.*, 2021) were used for signalling and
107 gene regulatory pathways components. The CellDesigner maps are available in Systems Biology
108 Markup Language (SBML) format (Hucka *et al.*, 2003). We provide a schematic representation of
109 the major steps in Figure 2.

110 **2.1.2. Curation Criteria**

111 All the RA-Atlas maps underwent the first step of manual curation based on a broad study of
112 literature and pathway databases to search for components and interactions involved in RA's
113 pathogenesis. We focused on small-scale experimental studies and added animal-based studies where
114 the human-related information was not adequate to validate the role of components or interactions.
115 However, the information based on animal studies remains limited and completely traceable (four
116 molecules are involved in the Th1 map, four in macrophage maps and two in the fibroblast map).
117 Including a component or a reaction in a map followed precise curation criteria: first, we included all
118 components experimentally proven to be expressed in RA's pathogenesis (and in a cell-specific
119 manner for the cell-specific maps). Then, we included all reactions experimentally proven to occur in
120 RA's pathogenesis (and cell-specifically for the cell-specific maps). General pathway interactions
121 were then added to complete specific pathways where disease-specific information was unavailable.

122 **2.1.3. Annotations**

123 Annotations were added to provide references for all species, reactions, and compartments present in
124 the maps using MIRIAM (Minimal Information Requested In the Annotation of Models) (Novère *et*
125 *al.*, 2005), a standard for annotating and curating computational models and maps. MIRIAM
126 annotations are added through the dedicated section of CellDesigner with the relation "bqbiol: is
127 describedby", which is used to link a component or a reaction to the literature or data that describes it
128 (e.g., PubMed references (PMIDs), DOI, GEO, KEGG identifier). Annotated maps provide
129 information about the various sources of information and help assess the maps' specificity.

130 **2.1.4. Visualization and Accessibility**

131 All RA-Atlas maps presented in this work are available as online interactive maps on the standalone
132 web server MINERVA (Molecular Interaction NETwoRks VisuAlization) (Gawron *et al.*, 2016).
133 MINERVA platform allows for visual exploration, analysis and management of molecular networks
134 encoded in systems biology formats, including CellDesigner, SBML and SBGN. MINERVA also
135 provides automated content annotation and verification and overlaying experimental data. Our maps
136 integrate data from various sources. We provide cell-specific overlays (cf. the overlay section)
137 extracted from the literature and publicly available datasets to visualize cellular signatures on the
138 different RA-Atlas maps.

139 **2.2. RA-map V2**

140 **2.2.1. Evaluation of the Prior RA-map**

141 The components and reactions present in the previous version of the RA-map were assessed, and
142 bibliographical references were added wherever required to increase the overall confidence level of
143 the map and limit false positives.

144 **2.2.2 Metabolic Pathways Addition**

145 Metabolic pathways of interest (i.e., glycolysis, pentose phosphate pathway, citric acid cycle and
146 oxidative phosphorylation) were extracted from the PANTHER pathway database (Mi and Thomas,
147 2009), adapted to SBGN PD standards when required and added to the original RA-map. Metabolites
148 and metabolic enzymes were named following BiGG IDs (King *et al.*, 2016). In addition,
149 considerable bibliographical work was conducted to find evidence linking the newly added metabolic
150 pathways to the already present signalling and gene regulatory pathways.

151 **2.2.3. Annotation Score**

152 Thorough annotations throughout the manual curation process allowed us to calculate an annotation
153 score for each compound present in the RA-map V2 based on the number of bibliographic references
154 describing it.

155 **2.3. Cell-specific Maps**

156 To build cell-specific maps for M1 and M2 macrophages, fibroblasts, and Th1, we screened literature
157 (small scale experiments, GWAS) for signalling pathways, genes or transcription factors associated
158 with RA. If available, we also integrated pre-existing maps. We used cell/disease-specific markers or
159 gene expression datasets to retrieve signalling pathways involved in RA pathogenesis. To define gene
160 signatures and primary RA cellular outcomes, we used databases such as the MSig database
161 (Liberzon *et al.*, 2015), KEGG and biocuration. The retrieved gene signatures were used to connect

162 secreted molecules to the cellular phenotypes included in the map. The cellular phenotypes include
163 angiogenesis, apoptosis, bone erosion, cell chemotaxis, recruitment and infiltration, cell growth,
164 survival and proliferation, hypoxia, inflammation, matrix degradation and osteoclastogenesis.

165 Two proprietary software were used to enrich the maps: IPA and MetaCore (Clarivate) version 22.1
166 build 70800, both displaying vast map databases. In MetaCore, we mapped the list of DEG to the
167 internal map database using the "Pathway maps" option through the "One-click analysis" tab. The
168 number of overlapping DEGs is displayed next to each identified map. The same analysis was done
169 using IPA's "canonical pathways" analysis. Highlighted DEG and their associated pathways were
170 identified, added to the corresponding RA-Atlas map and further annotated with PMIDs or GEO
171 accession numbers.

172 **2.3.1. Datasets**

173 Regarding the datasets, for the RA macrophage map, we used GSE97779, a publically available
174 microarray dataset from the GEO database (Edgar, 2002). The dataset contains nine RA synovial
175 macrophages sample from nine patients and five peripheral blood monocyte-derived macrophages
176 samples from five healthy donors. Gene expression was quantile normalized using the
177 preprocessCore package (Bostald, 2021). GSE164498, a RNA-seq single-cell dataset from the GEO
178 database, was used to provide M1 and M2 macrophage signatures. It contains 1766 M1 macrophage
179 cells and 2063 M2 macrophage cells. M1 macrophages were polarized using LPS, and IFN-Gamma
180 and M2 macrophage cells were polarized using IL-4 and IL-13. We used GSE109449, a RNA-seq
181 single cell dataset available in the GEO database for the RA fibroblast map. It contains 384 freshly
182 isolated fibroblasts in two RA and two osteoarthritis patients. For the RA Th1 map, we used
183 SDY998, a single-cell RNA-seq dataset from the Immport database (Bhattacharya *et al.*, 2018)
184 containing 19 samples from RA patients and two synovial samples from osteoarthritis patients with
185 three cell types: 1142 B cells, 1844 fibroblasts, 750 monocytes, and 1529 T cells. The gene
186 expression was normalized using log₂ (CPM). For the GSE analysis, we used four datasets: SDY998,
187 GSE32901, GSE107011 and GSE135390. The first dataset is the same one used for DEA. The
188 second is a microarray gene expression dataset containing PBMC samples from 5 healthy donors. It
189 includes four CD4⁺ T cells subtypes: naive CD4⁺ T cells, Th1, Th17 rich, and Th17 poor memory
190 CD4⁺ T cells. Gene expression was normalized using the log₂ RMA method. The third dataset is
191 bulk RNA-seq transcriptome profiling of 29 immune cell types, including Th1 and extracted from
192 PBMC sorted from 4 healthy individuals. The counts were normalized using the Transcripts Per
193 Million (TPM) method. The last dataset is RNA-seq gene expression profiling. It contains PBMC
194 samples coming from 3 healthy donors. Several subtypes of CD4⁺ T cells are in the samples: Th1,
195 TH2, TH22, Treg, naive CD4⁺ and Th17. Count data were normalized using TMM. In addition, we
196 used GSE172188, a microarray gene expression dataset of synovial samples from ten patients with
197 active RA treated with Abatacept. For each patient, synovial biopsies were obtained from the same
198 affected joint before (W0) and 16 weeks after (W16) starting treatment with Abatacept.

199 **2.3.2 Macrophage Maps**

200 **2.3.2.1. Macrophage Map Integration**

201 The macrophage map was constructed as a PD graph on CellDesigner by integrating three maps as a
202 base. The initial map was the RA-specific macrophage map from IPA including 44 molecules and 47
203 interactions associated with PMIDs. The second and third macrophage-specific maps (but not RA
204 specific) were retrieved from the published literature (Oda *et al.*, 2004; Wentker *et al.*, 2017).

205 As the second and third macrophage maps were not RA-specific, we used the RA-specific marker
206 gene lists from IPA as an overlay in MINERVA to identify disease-specific molecules and
207 submodules. The Stream export plugin (Hoksza *et al.*, 2019) in MINERVA allowed us to export
208 these submodules of interest from the macrophages maps. We identified corresponding signalling
209 pathways in the extracted submodules and completed them by adding molecules from the literature
210 and pathway databases like KEGG.

211 **2.3.2.2. RA Macrophage Map Splitting into M1 and M2 Macrophage Maps**

212 We split the RA macrophage map into two maps according to the M1 or M2 phenotype. We included
213 Msig gene signatures for M1 and M2 macrophages and information available in the literature for
214 each signalling pathway in the map. However, it was impossible to assign a phenotype for some
215 pathways due to both phenotypes' lack of information or expression. These pathways were kept in
216 both M1 and M2 maps.

217 **2.3.2.3. M1 and M2 Macrophages Map Enrichment**

218 Differential Expression Analysis (DEA) was performed using the preprocessCore and Limma
219 packages (Ritchie *et al.*, 2015) in R on the GSE97779 dataset (cf. dataset section). The list of DEG
220 was used for enrichment with IPA and MetaCore. To assign the newly identified pathways to the M1
221 or the M2 phenotype, DEA was performed on GSE164498 (cf. dataset section) using Bioturing
222 software (Le *et al.*, 2020) with the Venice method (Vuong *et al.*, 2020) and M1 versus M2
223 macrophage cells to identify signature genes for both M1 and M2 phenotypes. The percentage of
224 shared genes between each newly identified signalling pathway and M1 and M2 signatures was
225 calculated. Based on these criteria, new pathways were assigned to the M1, M2 maps, or both.

226 **2.3.3. Fibroblast Map**

227 To build the RA fibroblast-specific map, we filtered non-fibroblast-specific components from the
228 initial RA-map based using the up-to-date RA fibroblast overlay available in MINERVA (cf. overlay
229 section). Non-fibroblast-specific components were only kept if necessary to transduce the signal. We
230 used the BioTuring software and the Venice method to perform DEA on GSE109449 (cf. dataset
231 section). The filtered DEG list was used as an overlay in MINERVA to retrieve more RA fibroblast-
232 specific components from the RA global map.

233 **2.3.3.1. Fibroblast Map Enrichment**

234 The same DEG list was used to enrich the fibroblast map with new signalling pathways using IPA
235 and MetaCore software.

236 **2.3.4. Th1 Map**

237 Regarding the Th1 map, we did not integrate any pre-existing cell-specific map. Alternatively, we
238 first searched the literature for the main signalling pathways expressed in the Th1 subtype and
239 involved in RA's pathogenesis (using queries like "Th1" and "rheumatoid arthritis"). In addition, we
240 used the multicellular map of inflammation (Serhan *et al.*, 2020) and its not-RA-specific Th1 submap
241 to identify transcription factors and cell-specific receptors involved in RA pathogenesis. We
242 integrated the SDY998 gene expression dataset (cf. dataset section) to enrich the map with new
243 pathways. Clustering was performed on the T cells using the Seurat package (Hao *et al.*, 2021) in R
244 (Find_neighbors and Find_clusters functions). Based on the expression of the marker genes CXCR3,

245 CCR7 and CCR6, the Th1 cluster was identified. We performed DEA using CXCR3+ CCR7- CCR6-
246 Th1 cluster versus CCR7+ naive cells from RA samples using Seurat and Find_marker_genes
247 function. The filtered DEG list was used to identify new pathways using IPA and MetaCore tools
248 described above. To expand the Th1 map, we also performed Gene Set Enrichment Analysis (GSEA)
249 using GSEA software (Subramanian *et al.*, 2005), IPA gene sets and the SDY998 gene expression
250 dataset. The list of enriched gene sets was filtered with a False Discovery Rate (FDR) less or equal to
251 5% to keep only the statistically significant ones. To ensure that the results are Th1 specific, we
252 performed GSEA on three other datasets: GSE32901, GSE107011 and GSE135390 (cf. dataset
253 section), and filtered the enriched gene sets with FDR less or equal to 5%. Only gene sets identified
254 with the SDY998 dataset shared at least by one of the GSE32901, GSE107011 or GSE135390
255 datasets were kept. Only gene sets sharing less than 80% of their core enrichment genes were kept to
256 avoid redundancy.

257 **2.3. Overlays**

258 We provide nine different sample-specific overlays, namely fibroblasts, synovial fluid, synovial
259 tissue, PBMC, blood components, serum, chondrocytes, macrophages, and Th1 to calculate the cell-
260 specificity of the RA-Atlas maps or visualize the cell-specific pathways in MINERVA.

261 Most overlays (i.e. synovial fluid, synovial tissue, PBMC, blood components, and chondrocytes) are
262 updated versions of the overlays provided in the initial RA-map with the addition of new
263 bibliographical references. The fibroblast overlay is also an updated version of the overlay provided
264 in the initial paper with the addition of bibliographical references and the list of DEG obtained from
265 DEA (see RA fibroblast map enrichment section). The overlay file now consists of 2409 fibroblast-
266 specific components. The macrophage overlay includes the RA macrophage-specific components
267 identified in the literature, the DEG list obtained from DEA (cf. M1/M2 macrophages map
268 enrichment with gene expression datasets section), and the 44 molecules present in the IPA RA-
269 specific macrophage map (cf. Macrophage maps integration. It consists of 648 RA macrophage-
270 specific molecules. The Th1 overlay consists of 523 molecules, including the list of DEG obtained
271 previously, the core enrichment genes from the GSEA (cf. Th1 map section), and information found
272 in the literature. Finally, the serum overlay was created to account for a new source of information
273 used in the maps.

274 **3. Results**

275 **3.1. RA-map V2**

276 The RA-map V2 illustrates signalling pathways, gene regulation, metabolic pathways, molecular
277 mechanisms and phenotypes involved in RA's pathogenesis. The map is compartmentalized in a way
278 to represent the flow of information from the extracellular space (ligands) to the plasma membrane
279 (ligands-receptors or metabolic transporters complexes) through the cytoplasm (signalling and
280 metabolism), the nucleus (gene regulation) and the secreted compartments (phenotype activation).

281 The map is fully compliant with SBGN PD standards. It includes 720 species (329 proteins, 135
282 genes, 136 RNAs, 54 simple molecules, 1 ion and 65 molecular complexes), 9 phenotypes
283 (angiogenesis, apoptosis, bone erosion, cell chemotaxis/recruitment/infiltration, cell
284 growth/survival/proliferation, hypoxia, inflammation, matrix degradation and osteoclastogenesis) and
285 602 reactions. This new version of the map represents an update regarding the number of
286 components. We have added 158 species, primarily involved in metabolic pathways, 156 reactions,
287 primary interconnecting metabolic pathways with signalling and gene regulation pathways, and the

288 hypoxic phenotype. We have included 225 additional references that make a total of 575 PMIDs for
289 the whole map. The high number of references enhances confidence in the presented mechanisms.
290 Indeed, 87% of the components included are supported with more than one experimental evidence
291 and 65% more than 2 (Figure 3).

292 In several cases, the map includes compounds with no bibliographic reference. For example, many
293 simple molecules act as products or reactants of well-known biological reactions whose expressions
294 are rarely highlighted in disease-specific experimental studies (e.g., ATP, ADP, NADH, NADPH,
295 H₂O, O₂, CoA, and FADH). Also, molecules that act as pathway intermediates do not present
296 bibliographical references. In this case, we have evidence that a pathway is expressed in a disease-
297 specific manner, but not all intermediates have been studied experimentally. For instance, specific
298 metabolic pathways are found to be expressed, however, experimental evidence is not available for
299 every component.

300 Beyond the considerable effort of the initial RA-map to include the disease hallmarks (e.g.,
301 cytokines, chemokines, growth factors, toll-like receptors) and molecular pathways (e.g., JAK-STAT
302 pathway, NF-KB pathway, MAPK pathway), we aimed to expand the coverage in disease-specific
303 metabolic dysregulated pathways. In particular, we focused on adding glycolysis, citric acid cycle,
304 pentose phosphate pathway, and oxidative phosphorylation pathways as their interplay with
305 inflammation and immunity mechanisms, among others, have been proven in RA (Weyand and
306 Goronzy, 2017; Pucino *et al.*, 2020). In addition, non-metabolic functions of metabolites and
307 metabolic enzymes were illustrated, such as their transcription regulation function (e.g., Hexokinase
308 2 or Phosphoglycerate Kinase 1 (Kim and Dang, 2005)) or their involvement in disease-specific
309 signalling pathways or phenotypes (e.g., Lactate (Yi *et al.*, 2022), Glucose-6-Phosphate Isomerase
310 (Zong *et al.*, 2015) or Pyruvate (CHANG and WEI, 2011)).

311 The RA-map V2 integrates information from several sources and cell types (Figure 4). The user can
312 take advantage of its extensive annotation and assess the map's coverage by opting for a specific
313 representation of a particular cell type in MINERVA. The map is mainly composed of fibroblast
314 specific information (73%) but other cell types, tissues and fluids involved in RA's pathogenesis
315 (Fang, Zhou and Nandakumar, 2020) are also represented, such as synovial tissue, synovial fluid,
316 blood components, serum components, PBMC, chondrocytes and macrophages. Therefore, when
317 interpreting those results, one must consider that a specific component can be common to several cell
318 types or that this component is present only in one cell type.

319 **3.2. RA Fibroblast Map**

320 The RA fibroblast map results from an extensive effort to make the RA-map fibroblast-specific. It is
321 based on manual curation of the RA-map, using literature and gene expression integration. Using the
322 list of DEG, 23 components were identified as fibroblast-specific from the RA-map V2. In addition,
323 new pathways were identified using IPA and MetaCore. The top 5 identified signalling pathways are
324 shown in Table 1.

325 The enriched fibroblast-specific map comprises ten compartments from the extracellular space where
326 the biological signal is induced to the secreted components' and transmembrane proteins'
327 compartments. As the RA-map V2, it is SBGN PD compliant. It contains 853 species, including 411
328 proteins, 115 genes, 115 RNAs, 96 molecular complexes, 2 simple molecules and 9 phenotypes (8 of
329 them coming from the initial RA-map, the 9th one being T cells activation). The components interact
330 together via 509 reactions.

331 77% of the map's components are RA fibroblast-specific (Figure 5). Some of the molecules that are
332 not fibroblast-specific were added to the map to keep the signal transduction going from the
333 extracellular part to the nucleus. Indeed, in some RA fibroblast-specific pathways, these molecules
334 are needed to activate or inhibit intermediate proteins. Other molecules that are not RA fibroblast-
335 specific are found in the extracellular space. These molecules can come from the synovial fluid and
336 be secreted by other cell types to activate specific signalling pathways in the RA fibroblast. In this
337 case, the secreted molecules do not have to be RA fibroblast-specific. Supporting literature was not
338 always available for small molecules or adaptor proteins.

339 **3.3. RA Macrophage Maps**

340 We obtained the global macrophage map by converting the IPA cell and disease-specific map into an
341 SBGN PD compliant map in CellDesigner. New pathways were added from the RA-specific
342 submodules exported from the previously published maps (Oda *et al.*, 2004; Wentker *et al.*, 2017)
343 like TLRs pathway, NFkB pathway, PI3K-AKT pathway, MAPK signalling pathway, and IL-18.
344 Other pathways were identified from the literature, such as the IFN pathway, Notch pathway, IL10
345 and IL23 pathways.

346 KEGG, MSig database and literature screening allowed to add 8 phenotypes to the map. The
347 phenotypes include inflammation, apoptosis, angiogenesis, extracellular matrix degradation, cell
348 chemotaxis/migration, T cells activation, proliferation/survival and osteoclastogenesis.

349 **3.3.1. RA Macrophage M1 Map**

350 The RA macrophage M1-specific map includes all the signalling pathways from the RA macrophage
351 map leading to the macrophage polarization into the pro-inflammatory M1 phenotype. It was
352 enriched using gene expression dataset integration to identify new RA M1 macrophage-specific
353 pathways using IPA and MetaCore. The top 5 identified signalling pathways are shown in Table 2.

354 The RA macrophage M1-specific map includes 640 species: 278 proteins, 76 genes, 76 RNAs, 114
355 molecular complexes, 8 simple molecules and the same phenotypes as the RA macrophage map.
356 These components interact together via 448 reactions.

357 61% of the components of this map are RA macrophage-specific (Figure 5). In addition, all the
358 receptors and secreted components are RA macrophage-specific. Although less literature is available
359 compared to the RA Fibroblast map, gene expression datasets integration enabled us to keep a high
360 RA macrophage specificity in the map.

361 **3.3.2. RA Macrophage M2 Map**

362 The M2 macrophage map includes all the signalling pathways from the RA macrophage map leading
363 to the macrophage polarization into the anti-inflammatory M2 phenotype. After the enrichment step
364 with IPA and MetaCore, additional pathways were added. The top 5 signalling pathways are shown
365 in table 2

366 The enriched map includes 520 species: 243 proteins, 59 genes, 59 RNAs, 90 molecular complexes,
367 three simple molecules and 7 of the 8 phenotypes from the RA macrophage map. Indeed,
368 osteoclastogenesis is not associated with any of the secreted molecules in the RA M2 macrophage
369 map. The Macrophage M2 map's components interact together via 342 reactions.

370 55% of the components of this map are RA macrophage-specific (Figure 5). All the receptors and
371 secreted components are RA macrophage-specific. We can see that fewer pathways are specific to the
372 M2 phenotype in RA because RA disease is mainly associated with the M1 phenotype.

373 **3.4. RA Th1 Map**

374 We built the RA Th1 cell map through literature mining to identify the main cell-specific signalling
375 pathways involved in RA's pathogenesis. However, less information about the Th1 subtype was
376 available in the literature than fibroblasts and macrophages. Hence, it was essential to expand the
377 map using new methods. New signalling pathways were identified using IPA and MetaCore and the
378 enriched gene sets from GSEA (Tables 3 and 4).

379 The map consists of 321 species, including 167 proteins, 29 genes, 29 RNAs, 64 molecular
380 complexes, and 7 phenotypes. The phenotypes include inflammation, apoptosis, osteoclastogenesis,
381 cell chemotaxis, angiogenesis, matrix degradation and proliferation. RA Th1 map's components are
382 connected through 179 reactions.

383 58% of the components of this map are RA Th1-specific, including all the transcription factors, most
384 of the receptors, and secreted components (Figure 5).

385 **3.5. Applications**

386 Several options are available in MINERVA for the user to dive into the different RA-Atlas maps.
387 Users can explore the maps using the search bar and get the locations of their compounds of interest
388 and their associated references (Figure 6A) or visualize experimental data and cell-specific
389 components using the appropriate overlays (Figure 6B). Users can exploit online queries for drug
390 targets via DrugBank (Wishart *et al.*, 2008) and ChEMBL (Gaulton *et al.*, 2017).

391 Users can also use plugins available in MINERVA, such as the Stream-export, to export parts of the
392 network upstream and downstream of a node. GSEA plugin can be used to calculate enrichment for
393 uploaded user-provided data overlays. We performed GSEA by using a list of DEG from the
394 GSE172188 dataset (cf. dataset section) between the paired post- and pre-treatment samples as an
395 overlay to identify enriched signalling pathways from the fibroblast map in RA patients after
396 treatment with Abatacept compared to untreated RA patients. The list of DEG was downloaded from
397 the supplementary materials of the associated publication (Triaille *et al.*, 2021). Results of the GSE
398 analysis are shown in Figure 7.

399 **4. Discussion**

400 Knowledge assembly in the form of networks is an active field in Systems Biology. Many public
401 repositories such as KEGG (Ogata *et al.*, 1999), Reactome (Gillespie *et al.*, 2022), WikiPathways
402 (Martens *et al.*, 2021), Panther (Mi and Thomas, 2009), and proprietary platforms such as IPA
403 (Krämer *et al.*, 2014) and MetaCore (Clarivate) offer mechanistic insights into a variety of biological
404 processes. However, most of the diagrams available can be generic, lack cell specificity, and without
405 proper annotations and access to the literature used for their assembly. A key aspect of current
406 research is to move beyond generic characterizations of biological processes toward more specific
407 pathways considering cell, tissue, organ, and disease states.

408 International projects such as the Human Cell Atlas (Regev *et al.*, 2017) and the HapMap (Tanaka,
409 2003) have significantly contributed to our understanding of the role of different cells in the human

410 body and the specificities linked to the anatomical or functional localization of these cells. Moreover,
411 advancements in single-cell techniques offer insights into unprecedented specificity for various cell
412 types and subpopulations in normal and pathological conditions.

413 RA is a disease involving several cells and their crosstalks. Unfortunately, there is no cure, and the
414 proposed therapies can only alleviate symptoms and increase survival. Hence, understanding the
415 cellular interplay and intracellular cascades and their regulation could be advantageous in identifying
416 novel therapeutic targets that could help relieve the heavy burden of the disease's debilitating
417 symptoms.

418 With this aim, we built the RA-Atlas, an interactive, manually curated and enriched with extensive
419 omic data analysis representation of molecular mechanisms involved in RA's pathogenesis. It
420 includes an updated version of the global RA-map (Singh *et al.*, 2020) with the addition of metabolic
421 pathways, and cell-specific molecular interaction maps for resident fibroblasts, M1 and M2 resident
422 macrophages and CD4+ Th1 cells. The diagrams depicting the biological mechanisms are formalized
423 using the SBGN PD language (Le Novère, 2015). They are primarily based on high-quality manual
424 curation of the scientific literature, enrichment, and cross-validation using expression data with either
425 bulk or single-cell resolution. Furthermore, our group is an active player in the Disease Map
426 consortium. This initiative aims at fostering collaborations for the creation of disease-specific maps
427 such as The Parkinson's map (Fujita *et al.*, 2014), the AsthmaMap (Mazein, Knowles, *et al.*, 2018),
428 the Atlas of Cancer Signalling Network (Kuperstein *et al.*, 2015), The Atlas of Inflammation
429 Resolution (Serhan *et al.*, 2020), the Cystic Fibrosis Map: CyFi-MAP (Pereira *et al.*, 2021), and more
430 recently, the COVID19 Disease Map (Ostaszewski *et al.*, 2021). In addition, the community works
431 on developing best practices for annotation, curation, the use of systems biology standards, and the
432 development of platforms and technology for handling complexity (Ostaszewski *et al.*, 2019).

433 **5. Adopting FAIR principles**

434 The RA-Atlas is a project that tries to implement the FAIR principles (Wilkinson *et al.*, 2016) for
435 findability, accessibility, interoperability and reproducibility. The Atlas being a part of the Disease
436 Map community project (Mazein, Ostaszewski, *et al.*, 2018; Ostaszewski *et al.*, 2019), it makes its
437 dissemination among the community easier. The RA-Atlas is also freely accessible via a web
438 browser and the platform MINERVA (Gawron *et al.*, 2016). Its content is compliant with SBGN PD
439 (Le Novère, 2015) for the representation, SBML (Hucka *et al.*, 2003) for the construction, MIRIAM
440 for the annotations (Novère *et al.*, 2005), PMIDs and stable identifiers for the references. All entities
441 are annotated using HGNC symbols (Tweedie *et al.*, 2021) for signaling and gene regulation
442 components and BiGG IDs (King *et al.*, 2016) for metabolic compounds. Lastly, the extensive
443 annotations and the access to content allow for transparent and facilitated reuse of the resource.

444 **6. Perspectives**

445 The long-term objective of this project is to construct a computational repository to decipher the
446 interplay between cells of the innate and adaptive immunity in RA, which eventually leads to bone
447 and cartilage breakdown. We wish to expand the RA-Atlas with molecular interaction maps of B
448 cells, chondrocytes, osteoblasts, and osteoclasts and enrich existing cell-specific maps with relevant
449 metabolic pathways. We will also create a top-level view focusing on mapping how the different
450 cells interact to cope with complexity. Furthermore, intercellular interactions must be clearly defined,
451 mapped, and described to provide a comprehensive view of the cellular interplay in RA. Lastly, our
452 maps can serve as a basis for constructing executable disease networks (Singh *et al.*, 2018, Aghamiri

453 *et al.*, 2020; Niarakis and Helikar, 2021; Hall and Niarakis, 2021, Miagoux *et al.*, 2021), allowing for
454 in-silico simulations, hypotheses formation and predictions.

455

456 **Conflict of Interest**

457 FA is employed by SANOFI-AVENTIS R&D, AN collaborates with SANOFI-AVENTIS R&D via a
458 public-private partnership grant (CIFRE contract, n° 2020/0766), and NZ is the recipient of the
459 CIFRE doctoral contract. All other authors declare no conflict of interest.

460 **Author Contributions**

461 Conceptualization, A.N.; methodology, A.N. and F.A.; formal analysis, N.Z., S.A., F.A. and A.N.;
462 data curation, N.Z., S.A. and V.S.; writing—original draft preparation, N.Z., S.A., V.S. and A.N.;
463 writing—review and editing, N.Z., S.A., V.S., F.A., and A.N.; visualization, N.Z. and S.A.;
464 supervision, F.A., and A.N.; project administration, F.A., and A.N.; All authors have read and agreed
465 to the published version of the manuscript.

466 **Funding**

467 The work is supported by the public-private partnership grant (CIFRE contract, n° 2020/0766)

468 **Acknowledgements**

469 The authors would like to thank the Luxembourg Elixir hub, Dr Marek Ostaszewski, Dr Piotr Gauron
470 of the LCSB for their technical support, and Dr Zsolt Bocskei for helpful discussions.

471 **Licenses**

472 The tools IPA, MetaCore and BioTuring, were accessible via SANOFI Paris.

473 **Data Availability**

474 The RA-Atlas is freely accessible at <https://ramap.uni.lu/minerva/>. All PMIDs of scientific articles
475 used for the map construction are available in the annotation section of the maps. All datasets used
476 for the map construction are publicly available (cf. dataset section). All data and code used to
477 generate results, including XML files of the maps, cell-specificity and confidence score calculations,
478 overlays, and gene expression analysis are available on a GitLab repository at
479 <https://gitlab.com/genhotel/a-mechanistic-multicellular-atlas-of-the-rheumatic-joint> (accessed on
480 June, 8th 2022).

481 **References**

482 Aghakhani, S., Zerrouk, N. and Niarakis, A. (2020) 'Metabolic Reprogramming of Fibroblasts as
483 Therapeutic Target in Rheumatoid Arthritis and Cancer: Deciphering Key Mechanisms Using
484 Computational Systems Biology Approaches', *Cancers*, 13(1), p. 35. doi: 10.3390/cancers13010035.

485 Aghamiri, S. S. *et al.* (2020) 'Automated inference of Boolean models from molecular interaction
486 maps using CaSQ', *Bioinformatics*. Edited by J. Xu, 36(16), pp. 4473–4482. doi:
487 10.1093/bioinformatics/btaa484.

- 488 Balandraud, N., Roudier, J. and Roudier, C. (2004) 'Epstein–Barr virus and rheumatoid arthritis',
489 *Autoimmunity Reviews*, 3(5), pp. 362–367. doi: 10.1016/j.autrev.2004.02.002.
- 490 Bartok, B. and Firestein, G. S. (2010) 'Fibroblast-like synoviocytes: Key effector cells in rheumatoid
491 arthritis', *Immunological Reviews*, 233(1), pp. 233–255. doi: 10.1111/j.0105-2896.2009.00859.x.
- 492 Bhattacharya, S. *et al.* (2018) 'ImmPort, toward repurposing of open access immunological assay
493 data for translational and clinical research', *Scientific Data*, 5, pp. 1–9. doi: 10.1038/sdata.2018.15.
- 494 Bosisio, D. *et al.* (2018) 'Dendritic cells in inflammatory angiogenesis and lymphangiogenesis',
495 *Current Opinion in Immunology*, 53, pp. 180–186. doi: 10.1016/j.coi.2018.05.011.
- 496 Bostald, B. (2021) *preprocessCore: A collection of pre-processing functions. R package version*
497 *1.56.0*. Available at: <https://github.com/bmbolstad/preprocessCore> (Accessed: 12 April 2022).
- 498 CHANG, X. and WEI, C. (2011) 'Glycolysis and rheumatoid arthritis', *International Journal of*
499 *Rheumatic Diseases*, 14(3), pp. 217–222. doi: 10.1111/j.1756-185X.2011.01598.x.
- 500 Clarivate (no date) *MetaBase & MetaCore, a Cortellis solution*. Available at:
501 [https://clarivate.com/products/biopharma/discovery-clinical-regulatory/early-research-intelligence-](https://clarivate.com/products/biopharma/discovery-clinical-regulatory/early-research-intelligence-solutions/)
502 [solutions/](https://clarivate.com/products/biopharma/discovery-clinical-regulatory/early-research-intelligence-solutions/) (Accessed: 12 April 2022).
- 503 Deane, K. D. *et al.* (2017) 'Genetic and environmental risk factors for rheumatoid arthritis', *Best*
504 *Practice & Research Clinical Rheumatology*, 31(1), pp. 3–18. doi: 10.1016/j.berh.2017.08.003.
- 505 Edgar, R. (2002) 'Gene Expression Omnibus: NCBI gene expression and hybridization array data
506 repository', *Nucleic Acids Research*, 30(1), pp. 207–210. doi: 10.1093/nar/30.1.207.
- 507 Fang, Q., Zhou, C. and Nandakumar, K. S. (2020) 'Molecular and Cellular Pathways Contributing to
508 Joint Damage in Rheumatoid Arthritis', *Mediators of Inflammation*, 2020, pp. 1–20. doi:
509 10.1155/2020/3830212.
- 510 Fearon, U. *et al.* (2019) 'Altered metabolic pathways regulate synovial inflammation in rheumatoid
511 arthritis', *Clinical and Experimental Immunology*, 197(2), pp. 170–180. doi: 10.1111/cei.13228.
- 512 Frank-Bertoncelj, M., Klein, K. and Gay, S. (2017) 'Interplay between genetic and epigenetic
513 mechanisms in rheumatoid arthritis', *Epigenomics*, 9(4), pp. 493–504. doi: 10.2217/epi-2016-0142.
- 514 Fujita, K. A. *et al.* (2014) 'Integrating Pathways of Parkinson's Disease in a Molecular Interaction
515 Map', *Molecular Neurobiology*, 49(1), pp. 88–102. doi: 10.1007/s12035-013-8489-4.
- 516 Funahashi, A. *et al.* (2003) 'CellDesigner: a process diagram editor for gene-regulatory and
517 biochemical networks', *BIOSILICO*, 1(5), pp. 159–162. doi: 10.1016/S1478-5382(03)02370-9.
- 518 Gaulton, A. *et al.* (2017) 'The ChEMBL database in 2017', *Nucleic Acids Research*, 45(D1), pp.
519 D945–D954. doi: 10.1093/nar/gkw1074.
- 520 Gawron, P. *et al.* (2016) 'MINERVA—a platform for visualization and curation of molecular
521 interaction networks', *npj Systems Biology and Applications*, 2(1), p. 16020. doi:
522 10.1038/npjbsa.2016.20.

- 523 Gillespie, M. *et al.* (2022) 'The reactome pathway knowledgebase 2022', *Nucleic Acids Research*,
524 50(D1), pp. D687–D692. doi: 10.1093/nar/gkab1028.
- 525 Glocker, M. O. *et al.* (2006) 'Rheumatoid arthritis, a complex multifactorial disease: On the way
526 toward individualized medicine', *Medicinal Research Reviews*, 26(1), pp. 63–87. doi:
527 10.1002/med.20045.
- 528 Guo, Q. *et al.* (2018) 'Rheumatoid arthritis: pathological mechanisms and modern pharmacologic
529 therapies', *Bone Research*, 6(1), p. 15. doi: 10.1038/s41413-018-0016-9.
- 530 Hall, B.A. and Niarakis, A. (2021) Data integration in logic-based models of biological mechanisms,
531 *Current Opinion in Systems Biology*, Volume 28, 100386, doi.org/10.1016/j.coisb.2021.100386.
- 532 Hao, Y. *et al.* (2021) 'Integrated analysis of multimodal single-cell data', *Cell*, 184(13), pp. 3573-
533 3587.e29. doi: 10.1016/j.cell.2021.04.048.
- 534 Hoksza, D. *et al.* (2019) 'MINERVA API and plugins: opening molecular network analysis and
535 visualization to the community', *Bioinformatics*. Edited by L. Cowen, 35(21), pp. 4496–4498. doi:
536 10.1093/bioinformatics/btz286.
- 537 Huber, L. C. *et al.* (2006) 'Synovial fibroblasts: Key players in rheumatoid arthritis', *Rheumatology*,
538 45(6), pp. 669–675. doi: 10.1093/rheumatology/kel065.
- 539 Hucka, M. *et al.* (2003) 'The systems biology markup language (SBML): a medium for
540 representation and exchange of biochemical network models', *Bioinformatics*, 19(4), pp. 524–531.
541 doi: 10.1093/bioinformatics/btg015.
- 542 Kim, J. and Dang, C. V. (2005) 'Multifaceted roles of glycolytic enzymes', *Trends in Biochemical
543 Sciences*, 30(3), pp. 142–150. doi: 10.1016/j.tibs.2005.01.005.
- 544 King, Z. A. *et al.* (2016) 'BiGG Models: A platform for integrating, standardizing and sharing
545 genome-scale models', *Nucleic Acids Research*, 44(D1), pp. D515–D522. doi: 10.1093/nar/gkv1049.
- 546 Kinne, R. W. *et al.* (2000) 'Macrophages in rheumatoid arthritis', *Arthritis Research*, 2(3), pp. 189–
547 202. doi: 10.1186/ar86.
- 548 Korczowska, I. (2014) 'Rheumatoid arthritis susceptibility genes: An overview', *World Journal of
549 Orthopedics*, 5(4), p. 544. doi: 10.5312/wjo.v5.i4.544.
- 550 Krämer, A. *et al.* (2014) 'Causal analysis approaches in Ingenuity Pathway Analysis.', *Bioinformatics
551 (Oxford, England)*, 30(4), pp. 523–30. doi: 10.1093/bioinformatics/btt703.
- 552 Kuperstein, I. *et al.* (2015) 'Atlas of Cancer Signalling Network: a systems biology resource for
553 integrative analysis of cancer data with Google Maps', *Oncogenesis*, 4(7), pp. e160–e160. doi:
554 10.1038/oncsis.2015.19.
- 555 Le, T. *et al.* (2020) 'BBrowser: Making single-cell data easily accessible', *bioRxiv*, (Figure 1), p.
556 2020.12.11.414136. Available at:
557 [https://www.biorxiv.org/content/10.1101/2020.12.11.414136v1%0Ahttps://www.biorxiv.org/content/
558 10.1101/2020.12.11.414136v1.abstract](https://www.biorxiv.org/content/10.1101/2020.12.11.414136v1%0Ahttps://www.biorxiv.org/content/10.1101/2020.12.11.414136v1.abstract).

- 559 Li, S. *et al.* (2017) 'Effector T helper cell populations are elevated in the bone marrow of rheumatoid
560 arthritis patients and correlate with disease severity', *Scientific Reports*, 7(1), p. 4776. doi:
561 10.1038/s41598-017-05014-8.
- 562 Liberzon, A. *et al.* (2015) 'The Molecular Signatures Database Hallmark Gene Set Collection', *Cell*
563 *Systems*, 1(6), pp. 417–425. doi: 10.1016/j.cels.2015.12.004.
- 564 Martens, M. *et al.* (2021) 'WikiPathways: connecting communities', *Nucleic Acids Research*, 49(D1),
565 pp. D613–D621. doi: 10.1093/nar/gkaa1024.
- 566 Mazein, A., Knowles, R. G., *et al.* (2018) 'AsthmaMap: An expert-driven computational
567 representation of disease mechanisms', *Clinical & Experimental Allergy*, 48(8), pp. 916–918. doi:
568 10.1111/cea.13211.
- 569 Mazein, A., Ostaszewski, M., *et al.* (2018) 'Systems medicine disease maps: community-driven
570 comprehensive representation of disease mechanisms', *npj Systems Biology and Applications*, 4(1), p.
571 21. doi: 10.1038/s41540-018-0059-y.
- 572 McInnes, I. B. and Schett, G. (2011) 'The Pathogenesis of Rheumatoid Arthritis', *New England*
573 *Journal of Medicine*, 365(23), pp. 2205–2219. doi: 10.1056/NEJMra1004965.
- 574 Mi, H. and Thomas, P. (2009) 'PANTHER Pathway: An Ontology-Based Pathway Database Coupled
575 with Data Analysis Tools', in *Methods in molecular biology (Clifton, N.J.)*, pp. 123–140. doi:
576 10.1007/978-1-60761-175-2_7.
- 577 Miagoux, Q., Singh, V., *et al.* (2021) 'Inference of an Integrative, Executable Network for
578 Rheumatoid Arthritis Combining Data-Driven Machine Learning Approaches and a State-of-the-Art
579 Mechanistic Disease Map', *Journal of Personalized Medicine*, 11(8):785.
580 doi.org/10.3390/jpm11080785
- 581 Niarakis, A. and Helikar, T. (2021) 'A practical guide to mechanistic systems modeling in biology
582 using a logic-based approach', *Briefings in Bioinformatics*, 22(4), pp. 1–19. doi:
583 10.1093/bib/bbaa236.
- 584 Novère, N. Le *et al.* (2005) 'Minimum information requested in the annotation of biochemical models
585 (MIRIAM)', *Nature Biotechnology*, 23(12), pp. 1509–1515. doi: 10.1038/nbt1156.
- 586 Le Novère, N. (2015) 'Quantitative and logic modelling of molecular and gene networks', *Nature*
587 *Reviews Genetics*, 16(3), pp. 146–158. doi: 10.1038/nrg3885.
- 588 Oda, K. *et al.* (2004) 'Molecular Interaction Map of a Macrophage', *AfCS Research Reports*, 2(14),
589 pp. 1–12.
- 590 Ogata, H. *et al.* (1999) 'KEGG: Kyoto Encyclopedia of Genes and Genomes.', *Nucleic acids*
591 *research*, 27(1), pp. 29–34. doi: 10.1093/nar/27.1.29.
- 592 Ostaszewski, M. *et al.* (2019) 'Community-driven roadmap for integrated disease maps', *Briefings in*
593 *Bioinformatics*, 20(2), pp. 659–670. doi: 10.1093/bib/bby024.
- 594 Ostaszewski, M. *et al.* (2021) 'COVID19 Disease Map, a computational knowledge repository of

- 595 virus–host interaction mechanisms', *Molecular Systems Biology*, 17(10), pp. 1–22. doi:
596 10.15252/msb.202110387.
- 597 Pereira, C. *et al.* (2021) 'CyFi-MAP: an interactive pathway-based resource for cystic fibrosis',
598 *Scientific Reports*, 11(1), p. 22223. doi: 10.1038/s41598-021-01618-3.
- 599 Pucino, V. *et al.* (2020) 'Metabolic Checkpoints in Rheumatoid Arthritis', *Frontiers in Physiology*,
600 11(April), pp. 1–12. doi: 10.3389/fphys.2020.00347.
- 601 Regev, A. *et al.* (2017) 'The Human Cell Atlas', *eLife*, 6, pp. 1–30. doi: 10.7554/eLife.27041.
- 602 Ritchie, M. E. *et al.* (2015) 'limma powers differential expression analyses for RNA-sequencing and
603 microarray studies', *Nucleic Acids Research*, 43(7), pp. e47–e47. doi: 10.1093/nar/gkv007.
- 604 Schulze-Koops, H. and Kalden, J. R. (2001) 'The balance of Th1/Th2 cytokines in rheumatoid
605 arthritis', *Best Practice & Research Clinical Rheumatology*, 15(5), pp. 677–691. doi:
606 10.1053/berh.2001.0187.
- 607 Serhan, C. N. *et al.* (2020) 'The Atlas of Inflammation Resolution (AIR)', *Molecular Aspects of
608 Medicine*, 74(September), p. 100894. doi: 10.1016/j.mam.2020.100894.
- 609 Sigaux, J. *et al.* (2019) 'Air pollution as a determinant of rheumatoid arthritis', *Joint Bone Spine*,
610 86(1), pp. 37–42. doi: 10.1016/j.jbspin.2018.03.001.
- 611 Singh, V. *et al.*, (2018) 'Computational Systems Biology Approach for the Study of Rheumatoid
612 Arthritis: From a Molecular Map to a Dynamical Model', *Genomics and computational biology*, 4(1),
613 e100050. doi.org/10.18547/gcb.2018.vol4.iss1.e100050.
- 614 Singh, V. *et al.* (2020) 'RA-map: building a state-of-the-art interactive knowledge base for
615 rheumatoid arthritis', *Database*, 2020, pp. 1–18. doi: 10.1093/database/baaa017.
- 616 Smolen, J. S., Aletaha, D. and McInnes, I. B. (2016) 'Rheumatoid arthritis', *The Lancet*, 388(10055),
617 pp. 2023–2038. doi: 10.1016/S0140-6736(16)30173-8.
- 618 Subramanian, A. *et al.* (2005) 'Gene set enrichment analysis: a knowledge-based approach for
619 interpreting genome-wide expression profiles.', *Proceedings of the National Academy of Sciences of
620 the United States of America*, 102(43), pp. 15545–50. doi: 10.1073/pnas.0506580102.
- 621 Lo Surdo, P. *et al.* (2018) 'DISNOR: a disease network open resource', *Nucleic Acids Research*,
622 46(D1), pp. D527–D534. doi: 10.1093/nar/gkx876.
- 623 Talsania, M. and Scofield, R. H. (2017) 'Menopause and Rheumatic Disease', *Rheumatic Disease
624 Clinics of North America*, 43(2), pp. 287–302. doi: 10.1016/j.rdc.2016.12.011.
- 625 Tanaka, T. (2003) 'The International HapMap Project', *Nature*, 426(6968), pp. 789–796. doi:
626 10.1038/nature02168.
- 627 Triaille, C. *et al.* (2021) 'Common Transcriptomic Effects of Abatacept and Other DMARDs on
628 Rheumatoid Arthritis Synovial Tissue', *Frontiers in Immunology*, 12(August), pp. 1–9. doi:
629 10.3389/fimmu.2021.724895.

- 630 Tweedie, S. *et al.* (2021) 'Genenames.org: the HGNC and VGNC resources in 2021', *Nucleic Acids*
631 *Research*, 49(D1), pp. D939–D946. doi: 10.1093/nar/gkaa980.
- 632 Vuong, H. *et al.* (2020) 'Venice: A new algorithm for finding marker genes in single-cell
633 transcriptomic data', *bioRxiv*, p. 2020.11.16.384479. doi: 10.1101/2020.11.16.384479.
- 634 Wentker, P. *et al.* (2017) 'An Interactive Macrophage Signal Transduction Map Facilitates
635 Comparative Analyses of High-Throughput Data', *The Journal of Immunology*, 198(5), pp. 2191–
636 2201. doi: 10.4049/jimmunol.1502513.
- 637 Weyand, C. M. and Goronzy, J. J. (2017) 'Immunometabolism in early and late stages of rheumatoid
638 arthritis', *Nature Reviews Rheumatology*, 13(5), pp. 291–301. doi: 10.1038/nrrheum.2017.49.
- 639 Wilkinson, M. D. *et al.* (2016) 'The FAIR Guiding Principles for scientific data management and
640 stewardship', *Scientific Data*, 3(1), p. 160018. doi: 10.1038/sdata.2016.18.
- 641 Wishart, D. S. *et al.* (2008) 'DrugBank: a knowledgebase for drugs, drug actions and drug targets',
642 *Nucleic Acids Research*, 36(suppl_1), pp. D901–D906. doi: 10.1093/nar/gkm958.
- 643 Wu, G. *et al.* (2010) 'A comprehensive molecular interaction map for rheumatoid arthritis', *PLoS*
644 *ONE*, 5(4). doi: 10.1371/journal.pone.0010137.
- 645 Yang, X., Chang, Y. and Wei, W. (2020) 'Emerging role of targeting macrophages in rheumatoid
646 arthritis: Focus on polarization, metabolism and apoptosis', *Cell Proliferation*, 53(7), pp. 1–12. doi:
647 10.1111/cpr.12854.
- 648 Yi, O. *et al.* (2022) 'Lactate metabolism in rheumatoid arthritis: Pathogenic mechanisms and
649 therapeutic intervention with natural compounds', *Phytomedicine*, 100(February), p. 154048. doi:
650 10.1016/j.phymed.2022.154048.
- 651 Zong, M. *et al.* (2015) 'Glucose-6-phosphate isomerase promotes the proliferation and inhibits the
652 apoptosis in fibroblast-like synoviocytes in rheumatoid arthritis', *Arthritis Research & Therapy*,
653 17(1), p. 100. doi: 10.1186/s13075-015-0619-0.
- 654
- 655
- 656
- 657
- 658 Figure 1. Schematic representation of the RA joint depicting the variety of actors implicated in RA's
659 pathogenesis. Blue frames indicate the different maps included in the RA-Atlas.
- 660 Figure 2. Stepwise construction of the RA-Atlas.
- 661 Figure 3. Annotation score of the RA-map V2. (A) Pie-chart of the annotation score's distribution
662 among the map's components. (B) Overlay of the annotation scores on the RA-map V2. Each
663 component is colored according to its annotation score.

664 Figure 4. Cell-specificity of the RA-map V2 components. (A) Distribution of the RA-map V2 cell-
 665 specificity. (B) Visualization of the fibroblast-specific overlay on the RA-map V2.

666 Figure 5. Percentage of cell-specific components per cell-specific maps.

667 Figure 6. Visualization of RA macrophage M2 map in MINERVA. (A) Snapshot of the search for
 668 VEGFa. (B) Snapshot of the visualization of the macrophage overlay.

669 Figure 7. Gene set enrichment analysis on the RA-Atlas using the GSEA plugin. Enriched pathways
 670 are presented in the right panel. Results can also be visualized in the corresponding maps (here, a
 671 zoom of the RA-map V2) to illustrate the enriched components.

672

673

674

675

676 Table 1. Top 5 of IPA and MetaCore maps using the RA fibroblast DEG list.

IPA map	Overlap	MetaCore map	Overlap
Antigen presentation	48.7%	Immune response_Induction of the antigen presentation machinery by IFN-gamma	45.2%
Inhibition of matrix metalloproteinases	39%	Cell adhesion_ECM remodeling	41.8%
Hepatic fibrosis/Hepatic stellate cell activation	22.7%	IL-1 beta- and Endothelin-1-induced fibroblast/ myofibroblast migration and extracellular matrix production in asthmatic airways	47.5%
GP6 signaling pathway	20.5%	TGFbeta-induced fibroblast/ myofibroblast migration and extracellular matrix production in asthmatic airways	36.6%
Axonal guidance signaling	15%	Glucocorticoid-induced elevation of intraocular pressure as a glaucoma risk factor	32.2%

677

678

679

680

681 Table 2. Top 5 of IPA and MetaCore maps using RA macrophage DEG list.

IPA map	Overlap	MetaCore map	Overlap
DNA methylation and transcriptional repression signaling	32.3%	Development_Role of proteases in hematopoietic stem cell mobilization	22.2%
transcriptional regulatory network in embryonic stem cells	19.6%	Colorectal cancer (general schema)	20%
NER (nucleotide excision repair , enhanced pathway)	10.1%	Cigarette smoke components TCDD and Benzo[a]pyrene and receptor AHR signaling in lung epithelial cells	13.8%
Xenobiotic metabolism AHR signaling pathway	6.6%	Cigarette smoke-mediated regulation of NRF2-antioxidant pathway in airway epithelial cells	13.7%
LPS/IL1 mediated inhibition of RXR function	4.1%	Inhibition of Ephrin receptors in colorectal cancer	13.3%

682

683

684

685

686

687

688

689

690

691

692

693

694

695 Table 3. Top 5 of IPA and MetaCore maps using RA Th1 DEG list.

IPA map	Overlap	MetaCore map	Overlap
Antigen presentation pathway	30.8%	Generation of cytotoxic CD8+ T cells in COPD	24%
B cell development	22.7%	Maturation and migration of dendritic cells in skin sensitization	19.5%
Th1 pathway	10.7%	Immune response_Induction of the antigen presentation machinery by IFN-gamma	18.8%
TH2 pathway	10.2%	COVID-19: immune dysregulation	17%
Th1 and TH2 pathway	9.3%	Chemokines in inflammation in adipose tissue and liver in obesity, type 2 diabetes and metabolic syndrome X	14.5%

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710 Table 4. List of enriched gene sets in RA Th1 after filtering.

Gene set	FDR
ICOS-ICOSL SIGNALLING IN T HELPER CELLS	0.003
CDC42 SIGNALLING	0.003
COMMUNICATION BETWEEN INNATE AND ADAPTIVE IMMUNE CELLS	0.004
Th1 ACTIVATION PATHWAY	0.004
PKC SIGNALLING IN T LYMPHOCYTES	0.005
ANTIGEN PRESENTATION PATHWAY	0.005
PROTEIN UBIQUITINATION PATHWAY	0.007
ALTERED T CELL AND B CELL SIGNALLING IN RHEUMATOID ARTHRITIS	0.007
CROSSTALK BETWEEN DENDRITIC CELLS AND NATURAL KILLER CELLS	0.009
INTERFERON SIGNALLING	0.01
OXIDATIVE PHOSPHORYLATION	0.01
REGULATION OF CELLULAR MECHANICS BY CALPAIN PROTEASE	0.013
REMODELLING OF EPITHELIAL ADHERENS JUNCTIONS	0.014
GRANULOCYTE ADHESION AND DIAPEDESIS	0.016

POLYAMINE REGULATION IN COLON CANCER	0.023
ROLE OF IL-17F IN ALLERGIC INFLAMMATORY AIRWAY DISEASES	0.027
HYPOXIA SIGNALLING IN THE CARDIOVASCULAR SYSTEM	0.031
PTEN SIGNALLING	0.032
ACTIVATION OF IRF BY CYTOSOLIC PATTERN RECOGNITION RECEPTORS	0.034
IL-8 SIGNALLING	0.038
TCA CYCLE II (EUKARYOTIC)	0.038
ROLE OF CYTOKINES IN MEDIATING COMMUNICATION BETWEEN IMMUNE CELLS	0.039
PD-1, PD-L1 CANCER IMMUNOTHERAPY PATHWAY	0.049