Inference of an integrative, executable network for Rheumatoid Arthritis combining data-driven machine learning approaches and a state-of-the-art mechanistic disease map

Quentin Miagoux[1], Vidisha Singh[1], Dereck de Mézquita[1], Valerie Chaudru[1], Mohamed Elati[2], Elisabeth Petit-Teixeira[1], Anna Niarakis[1,3*]

1 Université Paris-Saclay, Univ Evry, Laboratoire Européen de Recherche pour la Polyarthrite rhumatoïde - Genhotel, 91057, Evry, France.

2 CANTHER, University of Lille, CNRS UMR 1277, Inserm U9020, Lille, France

3 Lifeware Group, Inria, Saclay-île de France, 91120, Palaiseau, France

* Corresponding author

E-mail: anna.niaraki@univ-evry.fr (AN)

Abstract

Rheumatoid arthritis (RA) is a multifactorial, complex autoimmune disease that involves various genetic, environmental and epigenetic factors. Systems biology approaches provide the means to study complex diseases through the integration of different layers of biological information. Combining multiple data types can help compensate for missing or conflicting information and limit the possibility of false positives. In this work, we aim to unravel mechanisms governing the regulation of key transcription factors in RA and derive patient-specific models to gain more insights into the disease heterogeneity and the response to treatment.

We first make use of publicly available transcriptomic datasets (peripheral blood) relative to RA and machine learning to create an RA specific transcription factor (TF) co-regulatory network. The TF cooperativity network is subsequently enriched in signalling cascades and upstream regulators using a state-of-the-art, RA-specific molecular map. The integrative network is then used as a template to analyze patients' data regarding their response to anti-TNF treatment and identify master regulators and upstream cascades affected by the treatment. We use the Boolean formalism to simulate in silico subparts of the integrated network and identify combinations and conditions that can switch on or off the identified TFs, mimicking the effects of single and combined perturbations.

Keywords:

Network inference, Integrative Biology, Rheumatoid Arthritis, Signalling cascades, Gene regulation, Transcription factors, Boolean simulations, Systems Biology

Introduction

Rheumatoid Arthritis (RA) is an inflammatory, autoimmune disease that affects the joints of the body. While the exact aetiology is unknown, it involves a combination of environmental and genetic factors such as smoking and the presence of susceptibility genes, along with sex and age factors. RA affects 0.5-1% of the world population with women three times more susceptible to develop RA than men [1,2]. The onset of the disease is set around the fourth to fifth decade of one’s life [3] and if left untreated can be debilitating for the individual.

Symptoms of RA include synovial inflammation, joint stiffness and pain, cartilage destruction and bone erosion. In early RA, leukocytes invade the synovial joints followed

by other pro-inflammatory mediators which instigate an inflammatory cascade and provoke synovitis [2]. Activated monocytes and T cells, both a source of proinflammatory cytokines such as TNF-a, can be found in peripheral blood [4] and many RA studies have used peripheral blood cells to identify disease-related genes [5–8]. The common therapy for the RA includes the use of Disease-modifying anti-rheumatic drugs (DMARDs). Conventional DMARDs include drugs that target the entire immune system whereas biologic DMARDs are monoclonal antibodies (mAbs) and soluble receptors that target protein messenger molecules or cells. Patients who do not respond to conventional DMARDs usually initiate therapy with TNF inhibitors. However, approximately 30–40% of RA patients fail to respond to antiTNF therapy and are usually obliged to undergo several rounds of drug combinations [9].

Due to the complex nature of RA, systems biology and integrative approaches are needed in order to gain insight into the disease pathogenesis and progression. It is evident that focusing only on one aspect of the disease provides a limited understanding of the multifactorial nature of RA. Recently, many computational approaches, mainly network-based which rely on the integration of multi-omics data (proteomics, genomics, transcriptomics and metabolomics) have succeeded in unravelling key mechanisms in complex diseases [10–13]. In this direction, machine learning is a promising bioinformatics field that allows the use and integration of a variety of biomedical data with inherent complexity and large size. Studies have shown that incorporating prior knowledge to data-driven methodologies improves the quality and the biological relevance of the outcome [14–16]. One such machine learning tool is CoRegNet, an R/Bioconductor package that infers co-regulatory networks of transcription factors (TFs) and target genes by analyzing transcriptomic data and estimating TFs activity profiles. Moreover, the software also allows for network enrichment by integrating regulation

evidence for TF binding sites, protein-protein interaction data and ChiP data from various databases to support cooperative TFs [17].

In this work, we present a framework for the integration of signalling and transcriptional regulation cascades with genomic mutations, combining data-driven approaches with prior knowledge in the form of an integrative RA-specific network. To do so, we make use of publicly available transcriptomic data of white blood cells from patients suffering from RA, and the tool CoRegNet to infer a co-regulatory network. Next, we develop an integration pairing method to couple the RA co-regulatory network with a state-of-the-art disease map for RA [18], to enrich the cooperativity network with upstream, signalling regulators. Disease maps are comprehensive, knowledge-based representations of disease mechanisms including disease-related molecular interactions supported by literature-based evidence [19,20].

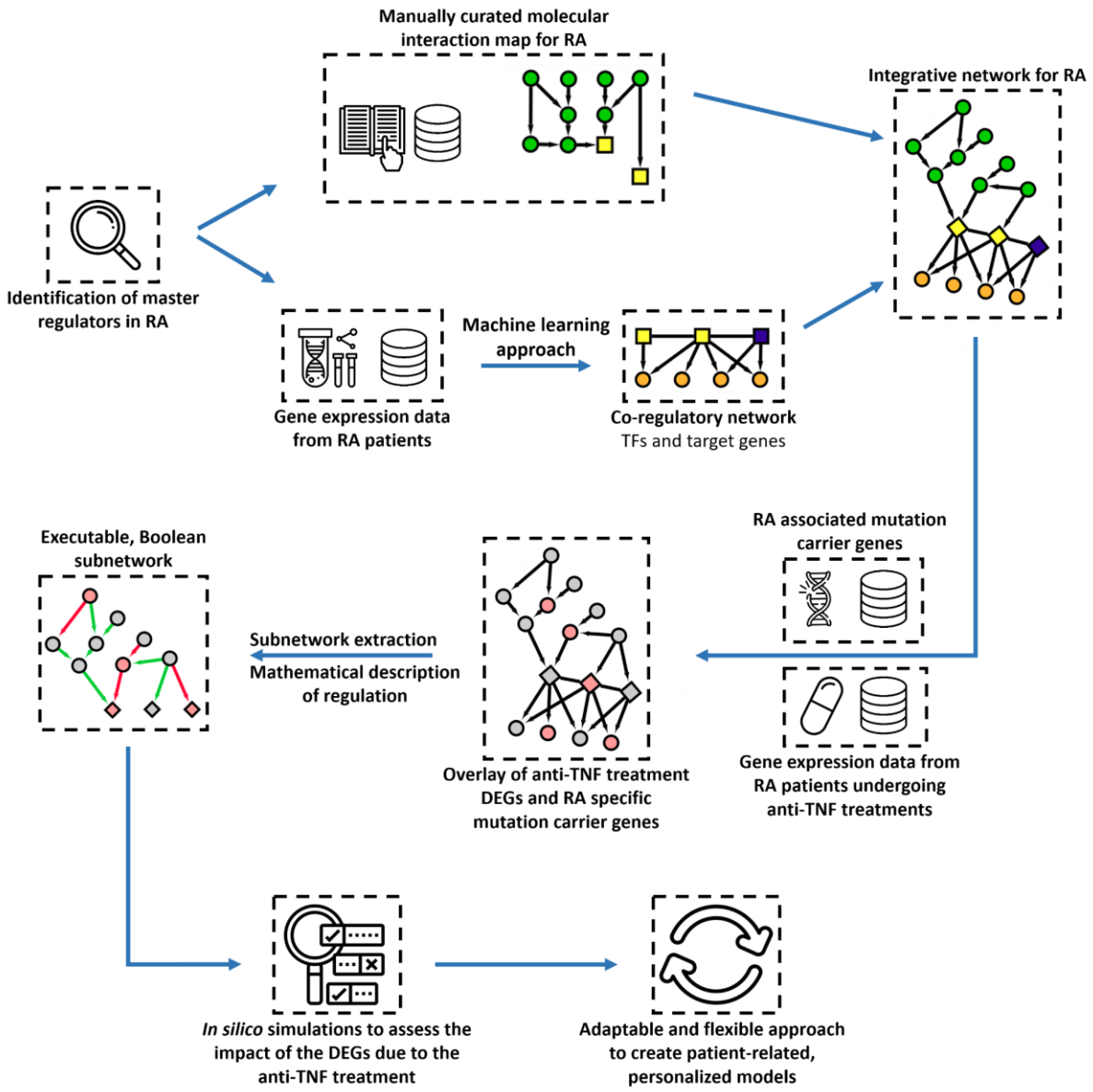


Figure 1. Workflow of an integrative and executable network for Rheumatoid Arthritis. From the identification of master regulators in Rheumatoid Arthritis to an adaptable and flexible personalized model, combining data-driven machine learning approaches and a state-of-the-art mechanistic disease map.

Next, we project on the integrative RA network public genomic data and transcriptomic data from treated RA patients, highlighting key mutation carriers and differentially expressed genes associated with the response to anti TNF treatment (Figure 1). The goal is to unravel mechanisms governing the regulation of key transcription factors and

genes identified as mutation carriers or DEGs in RA patients undergoing anti TNF treatment. Lastly, we study the dynamic behaviour of the system using the Boolean formalism to simulate subparts of the integrated network [21,22]. This includes real time simulations, sensitivity analysis and dose-response analyses to study the impact of other signaling cascades on the expression of the identified TFs, and steady state analysis revealing combinations and conditions that can switch on or off the identified TFs, mimicking the effects of the treatment [23].

Materials and Methods

Data description and pre-processing

We used a transcriptomic dataset from Illumina HiSeq 2500 of white blood cells from RA patients and healthy donors (GSE117769). The data set describes 120 subjects in total (51 RA and 50 controls, and 19 patients with either Ankylosing spondylitis or Psoriatic arthritis). From this dataset, we extracted 96 samples (46 RA and 50 controls). For the 46 RA samples, 43 were of Caucasian ancestry and 3 of unknown ancestry. Moreover, two samples consisted of duplicates of the same origin (female, control, unknown ancestry) and their expression matrix average was used for the analysis. We conducted a preliminary analysis on the expression matrix data with DESeq2 [24], using normalization and variance stabilizing transformation on the matrix expression. Principal Component Analysis (PCA) revealed 5 outliers (4 RA and 1 control) that were removed from the dataset (Figure S1). The final dataset used for further analysis comprised 90 samples (42 RA and 48 controls). Finally, we performed a normalisation on the raw expression matrix after removal of low read counts (>10).

Inference of the co-regulatory network (CoRegNet)

We used CoRegNet [17] R package version 1.26.0 to infer the co-regulatory network with the normalised gene expression matrix of the pre-processing step. The CoRegNet package implements the H-LICORN algorithm, allowing the identification of cooperative regulators of genes [25]. We enriched the inferred network with protein-protein and regulation evidence and further refined it with an unsupervised method using the unweighted mean. The co-regulatory network inferred with CoRegNet is composed of the significant edges between TF with a False Discovery Rate (FDR) [26] of 5%.

RA map upstream protein extraction

The RA map is a state-of-the-art interactive knowledge base for the RA disease [18]. The RA map is organized in the form of a cell representing the flow of information from the extracellular space to the plasma membrane and then to the cytoplasm, the nucleus and the secreted compartment or cellular phenotypes. For our analysis, as we were focused on upstream regulators of identified TFs, we used mainly the signalling part of the RA map.

More specifically, the list of TFs was uploaded as an overlay to the RA map, and the matching TFs were identified. The matching TFs were subsequently used as seeding nodes for the upstream plugin [27], setting the mode of extraction as upstream and selecting non-blocking modifiers. The obtained file was extracted as an XMLCellDesigner file. With this method we extracted upstream signalling cascades, and also seven translation reactions for which the mRNA was directly linked with the protein in the cytoplasm or the membrane.

The RA map, and consequently the extracted network, is written in the Process Description Systems Biology Graphical Notation scheme [28]. To obtain a more

simplified, Activity Flow (AF) - like representation of the network, the CellDesigner XML file of the previous step was used as an input to the tool CaSQ [29] to create an Activity Flow like executable network. CaSQ provides SBML-qual files for performing *in silico* simulations, but in our case, we used only the SIF file that contains information about the source, interaction type, and target of the Boolean network.

The obtained SIF file was further modified to address the issue of complexes - for every complex represented as a single node in the AF network, we recreated the reactants. This way we would not miss interactions and overlaps between nodes existing inside complexes. Regarding entities represented multiple times (as genes, proteins or mRNAs), for simplification purposes we kept only one entity and merged the corresponding interactions.

Global RA Network inference

We used the R package igraph [30] to convert the CoRegNet object (co-regulatory network) and the RA map SIF file into separate graphs. Then, we merged both networks using igraph, and imported the network into Cytoscape using the RCy3 R/ Bioconductor package [31] forming the global RA-specific network.

Differential expression analysis (DEA) using independent datasets

We conducted multiple differential expression analyses (DEA) using two different datasets. One dataset contained normalised counts from RNA sequencing data of CD4+ T cells including different responses to anti-TNF on RA patients (GSE138747). 37 and 41 RA patients treated with Adalimumab and Etanercept compose this dataset respectively, which were analysed independently. The second is composed of raw counts from RNA sequencing data of whole blood cells of biologic naive rheumatoid arthritis patients from baseline and after 3 months treatment with Infliximab or

Adalimumab (GSE129705). This dataset contains two different cohorts of 40 and 36 RA patients, which were also analysed independently. Thus, using DESeq2, we conducted two DEA on the comparison of responders and non-responders gene expression level for both drugs (Adalimumab and Etanercept) and two DEA on the comparison of baseline and after 3 months of anti-TNF treatment gene expression level. Therefore, from all these analyses, we considered a gene as a differentially expressed gene (DEG) with a corrected p-value (FDR) < 0.1. Finally, DEGs from both analyses were used as an overlay for the global RA-specific network.

List of variants

DisGeNET [32] contains the largest publicly available collection of genes and variants associated with human diseases. From this database, we extracted 2387 variants associated with RA. Then, we filtered out variants that had a variant disease association (VDA) and evidence index (EI) score lower than 0.7. VDA score is computed using the number of curated and non-curated publications supporting the variant disease association, while EI score is computed using contradictory results in publications supporting the variant. The use of a 0.7 threshold gives us at least one curated publication supporting the variant and the disease association resulting in a total of 1635 variants. Within these 1635 variants, we identified 731 associated genes that were subsequently used as an overlay for the global RA-specific network.

Subnetwork extraction

The subnetwork, based on the global network for RA, is focused on TNF, IL6 and TGFB1, 3 molecules highly implicated in RA (see Discussion). These 3 proteins were extracted along with their downstream cascades up to the first TF, to reduce complexity

and focus on the upstream regulators. From the global network for RA displayed in Cytoscape, we selected TNF, IL6 and TGFB1 simultaneously and using the Biological Network Manager (BiNoM) plugin, we selected in a stepwise manner the downstream neighbors of TGFB1, IL6 and TNF up to the first affected TF(s).

Shiny app

The co-regulatory network inferred with CoRegNet, the RA map Activity Flow extracted network and the merged global RA network, along with their overlays, were integrated to a web-based Shiny application using R [33]. The web application uses a Cytoscape viewer based on the R package cyjShiny [34] and is freely available (https://quentin-miagoux.shinyapps.io/global_ra_network/).

Subnetwork logic-based dynamical

The subnetwork previously obtained from the global network for RA, displayed on Cytoscape was extracted in SBML format using the BiNoM plugin and its function "export to SBML". Then, we removed co-regulatory interactions between TFs, which could have been misleading and interpreted as activation in further boolean simulation. We adjusted the overlay using CellDesigner and converted the SBML file into SMBL-Qual using CaSQ (v0.9.11).

The SMBL-Qual file was then imported on CellCollective to perform real time simulation. Using the "Simulation" tab on CellCollective, we mimicked the downregulation according to each subnetwork conditions (before/after treatment, responders/non-responders, mutation carriers), and activated one by one each input (TNF, IL6 and TGFB1). Furthermore, we also performed a sensitivity and dose response analysis with CellCollective, using 5 different initial states, described in Table 3.

The same file was used for analysis with the software GINsim [35] after a post processing modification step for node name recognition by the software. We used the nightly build version 3.0.0b-SNAPSHOT and the functions *Reduce model* for the reduced version, *Compute stable states* to obtain the stable states of the model and *Run simulation* with configurations of perturbations for the *in silico* KO experiments.

Results

Inference of the co-regulatory network

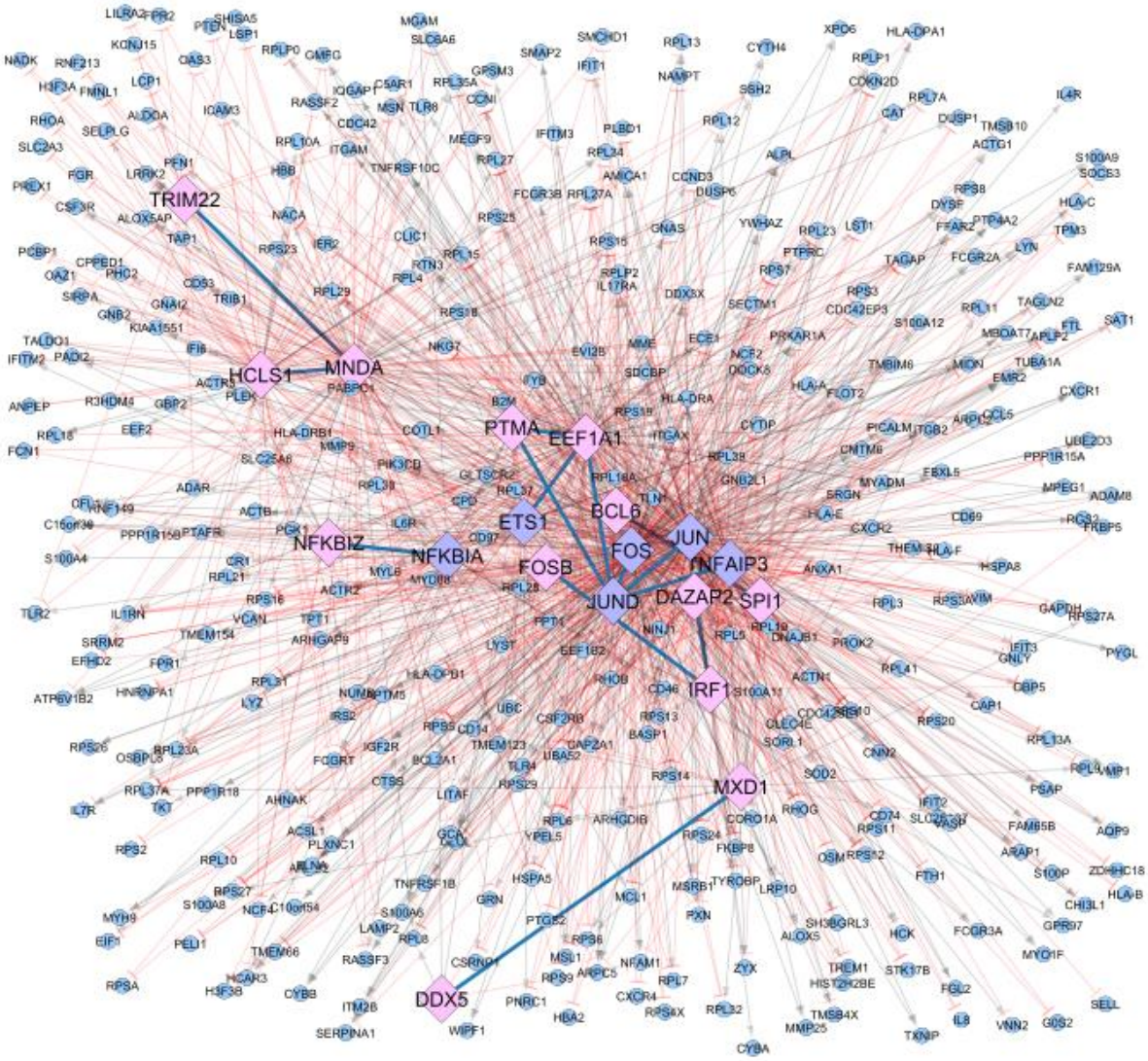
To infer the co-regulatory network, we selected a transcriptomic dataset from GEO database (GSE117769), including 120 samples (51 RA and 50 control, and 19 patients with either Ankylosing spondylitis or Psoriatic arthritis). After a series of pre-processing checks including the sample origins, duplicates, and quality of the data using a PCA on the matrix expression with a normalization and variance stabilizing transformation (shown in Figure S1), we ended up with 90 patients (48 Controls and 42 RA patients). Then, of the remaining samples we obtained normalized counts using DESeq2, on which we finally applied CoRegNet to infer the co-regulatory network, which is presented in Figure 2.

This network includes a total of 19 TF, 14 co-regulatory interactions and a total of 373 regulated target genes. Table 1 summarizes the top five TFs with the highest number of regulatory and co-regulatory interactions. Literature search for the nineteen TF identified from CoRegNet as the master regulators in the dataset, showed their potential implication to RA. Supplementary Table S1 summarises key roles of the TFs and the corresponding literature reference.

Table 1. Top 5 of the CoRegNet identified TFs with the highest number of regulatory interactions (TF - target gene) and co-regulatory interactions (TF-TF).

Top 5 TFs	**No of regulatory interactions**
FOS	288
JUN	211
EEF1A1	155
MNDA	136
TNFAIP3	125
Top 5 TFs	**No of co-regulatory interaction(s)**
JUND	5 (EEF1A1, FOS, JUN, PTMA, TNFAIP3)
EEF1A1	3 (ETS1, JUND, PTMA)
IRF1	2 (DAZAP2, FOSB)
MNDA	2 (HCLS1, TRIM22)
PTMA	2 (EEF1A1, JUND)

Figure 2. Co-regulatory network inferred using the tool CoRegNet and the matrix of normalized counts from the transcriptomic dataset (GSE117769). The dataset used included data from 90 patients (48 Controls and 42 RA patients). Matching TF from CoRegNet with the RA map are coloured in



purple (6) while non-matching are colored in pink (13). CoRegNet target genes are coloured in blue (373). Inhibitions are represented with red blunt arrows, and activations with grey arrows.

RA map upstream regulators of the TFs identified from CoRegNet

Six out of the 19 TFs namely ETS1, FOS, JUN, JUND, NFKBIA and TNFAIP3 from the co-regulatory network were present in the RA map and were used as seeds to extract their upstream regulators. The extracted network comprising the RA map upstream regulators of the matching TFs includes 244 nodes and is shown in Figure 3.

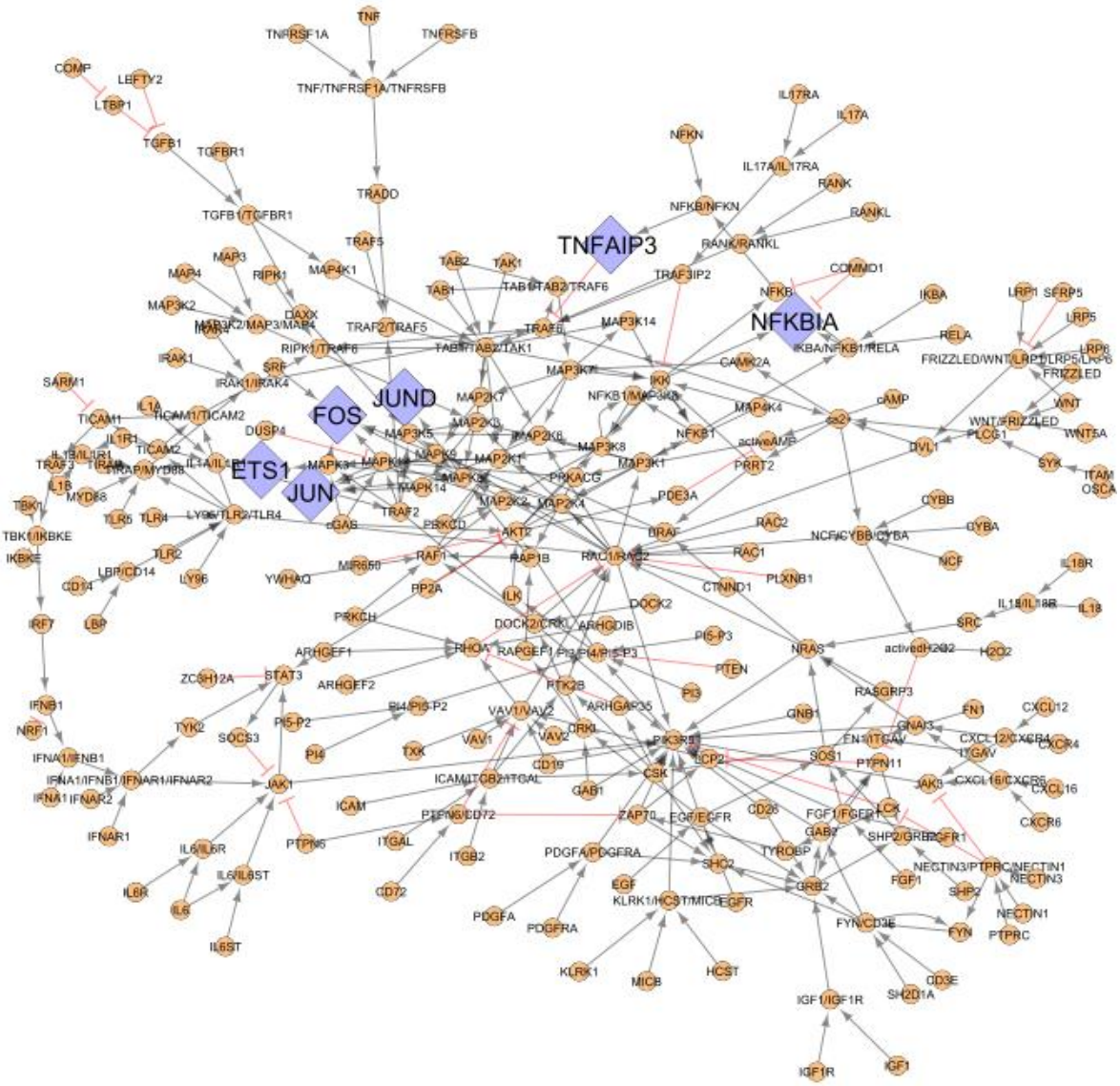


Figure 3. The RA map upstream regulators of the matching TFs between the RA map and the CoRegNet co-regulatory network. Matching TFs from the inferred network with CoRegNet are coloured in purple (6) and upstream regulators are coloured in orange (238). Inhibitions are represented with red blunt arrows, and activations with grey arrows.

Coupling gene co regulation with signalling cascades to obtain a global, integrative RA network

The global, integrative RA network is the result of the merging of the RA map signalling cascades and the CoRegNet object, using as interface the matching TFs. It comprises 614 nodes and 1736 interactions (848 inhibitions, 874 activations, and 14 co-regulatory

interactions shared among TFs) including genes, proteins, complexes, and simple molecules shown in Figure 4. In this network, six TFs were common between the CoRegNet network and the RA map (seeding TFs). In addition, 16 target genes identified with CoRegNet overlapped with the RA map upstream regulators.

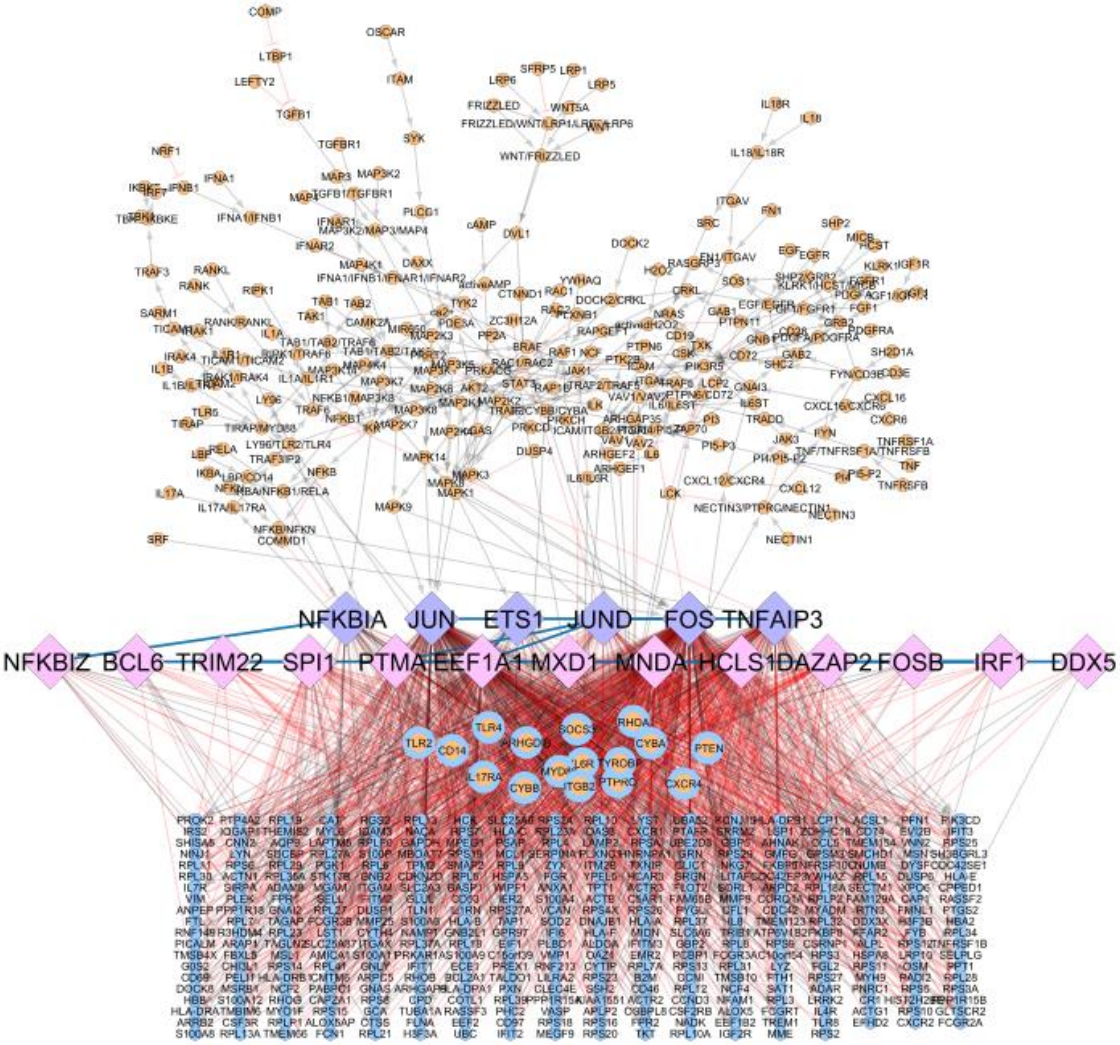


Figure 4. Integrative global network for RA. RA map upstream regulators are bound to TFs identified in the CoRegNet co-regulatory network. Matching TFs are coloured in purple (6), matching genes/proteins are coloured in blue and orange (16), upstream regulators from RA map are coloured in orange (222), CoRegNet TFs are coloured in pink (13) and CoRegNet target genes are coloured in blue (357). Inhibitions are represented with red blunt arrows, and activations with grey arrows.

Two use-cases: Identification of key TFs using DEG from RA patients undergoing anti-TNF treatment

Two datasets of RNAseq expression data coming from RA patients undergoing anti-TNF treatment were analysed to obtain differentially expressed genes (DEG). The first dataset focuses on responders and non-responders to RA treatment, including 37 and 41 RA patients treated with Adalimumab and Etanercept respectively. The second one involves untreated and treated (Infliximab or Adalimumab) RA patients including two cohorts of 40 and 36 RA patients, respectively. DEGs from these analyses were mapped to the global network for RA (presented in Figure S2 and Figure S3 respectively).

DEGs from responders/non-responders RA patients data mapping shows a total of 15 matching nodes, including 4 Etanercept DEGs and 11 Adalimumab DEGs. Also, 4 matching nodes are CoRegNet and RA map TFs (NFKBIA, JUN, FOS and TNFAIP3) and 1 CoRegNet TF only (FOSB), in the global network for RA. DEGs from untreated and treated RA patients mapping show a total of 101 matching nodes, including 2 CoRegNet and RA map TFs (NFKBIA and FOS) and 4 CoRegNet TF only (BCL6, MXD1, MNDA and DAZAP2).

Finally, cross-analysis revealed that a total of 9 over 19 TFs included in the global network for RA overlapped with a DEG from at least one analysis (presented in Table 2). Among these 9 TFs, two of them, NFKBIA and FOS overlapped with a DEG in both analyses.

Table 2. Transcription factors from the global network for RA with at least one differentially expressed gene (DEG) overlapping from the analysis of a) responders/non responders RA patients to anti-TNF treatment and b) before/after anti-TNF treatment of RA patients.↓ denotes downregulation.

	TF	Responders / non-responders	After treatment / before treatment
CoRegNet and RA map TF	*NFKBIA*	↓	↓
	JUN		↓
	FOS	↓	↓
	TNFAIP3		↓
CoRegNet TF	*BCL6*	↓	
	MXD1	↓	
	MNDA	↓	
	DAZAP2	↓	
	FOSB		↓

Subnetwork logic-based dynamical analysis

While the analyses highlight TFs differentially expressed (downregulated) after treatment or response to treatment with anti-TNF drugs, it is evident from the global network that the identified TFs can be regulated by a variety of other upstream cascades, besides those implicated in the TNF signalling.

To study further the interconnections with other pathways, we focused on IL6 and TGF-beta signalling (mentioned as TGFB1 in the network). Interleukin-6 (IL6) is the target of Tocilizumab (TCZ), an IL6 inhibitor frequently used in the treatment of RA. The inhibitor was developed in 2008, and its efficacy in treating the disease is quite similar to those of tumor necrosis factor (TNF) inhibitors [36]. TGF-beta signalling has been shown to be activated in RA synovium, however TGF beta blockage did not seem to have an

effect on experimental arthritis [37]. To study further the impact of these cascades on the expression of the identified TFs, we constructed a subnetwork, by selecting the molecules TGFB1, IL6 and TNF in the global network along with their downstream neighbors until reaching an identified TF.

The subnetwork contains 38 nodes, including 4 TFs and is highly enriched in MAPKs, as seen in Figure 5. By projecting the DEGs and known genomic variants associated with the disease we can see that intermediate nodes and TFs are downregulated by the anti-TNF treatment, while the inputs (IL6, TNF and TGFB1) along with a few intermediate nodes are characterized as mutation carriers.

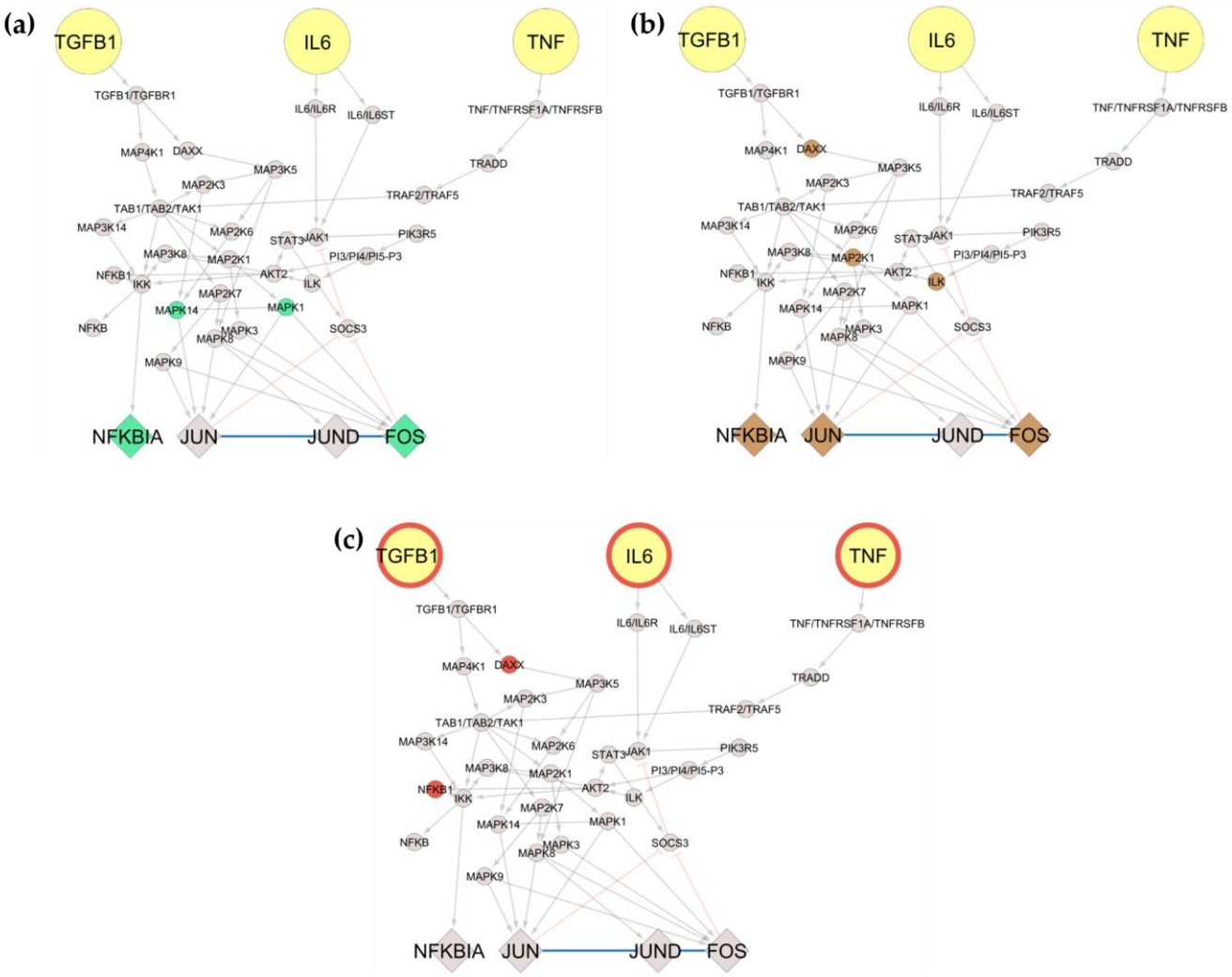


Figure 5. Subnetwork extraction of target genes (TGFB1, IL6 and TNF) from the global network for RA including differentially expressed genes and genomic variant overlap. We extracted three target

genes (TGFB1, IL6 and TNF) with their downstream neighbors until the first transcription factor matches. Then, we overlapped over the subnetwork three analyses including **(a)** Matching differentially expressed genes (DEG) from responders/non responders RA patients to anti-TNF treatment shown in green (4). **(b)** Matching DEG before/after anti-TNF treatment of RA patients shown in brown (6). **(c)** Matching DisGeNET variants shown in red (5).

In the next step, we wanted to evaluate the impact of single and combined perturbations of the network inputs, on the expression of the TFs. To do so, we used the possibility of adding Boolean rules to the network with the tool CaSQ [29]. Boolean models have been long used to describe biological mechanisms in health and disease [38], and they are an optimal approach for modelling signalling and gene regulation, when kinetic parameters are scarce. Boolean models are qualitative by nature, based on the assignment of binary values to the variables and the use of the logical operators AND, OR and NOT to describe the regulation of all molecules in the system [21]. The CaSQ tool receives as an input an SBML CellDesigner file [39] and produces a Boolean network with preliminary logical rules.

The Boolean model produced from our subnetwork has 38 nodes (three inputs, six outputs and 29 intermediate nodes) and 59 interactions. The SBML qual file was imported to Cell Collective [40] to perform real time simulations and sensitivity analysis and to GINsim [35] to calculate stable states and perform *in silico* KO (knock out) simulations. For the *in silico* KO simulations, a reduced version of 23 nodes was used.

Real time simulations using the Cell Collective platform

First we wanted to see the impact of the molecules either affected by the treatment or identified as mutation carriers, to the model outputs. The analysis corresponding to before and after anti-TNF treatment, showed that MAPK14 and MAPK1 were downregulated. Accordingly, for the responders and non responders analysis, MAP2K1,

ILK, DAXX were also identified as downregulated. Lastly, DAXX and NFKB1 were identified as mutation carriers. To mimic the effects of the downregulation of these molecules on the model outputs, we performed *in silico* simulations setting their activation level to zero.

For the dataset of before and after anti-TNF treatment of RA patients, MAPK14 and MAPK1 activity levels were set to zero, and simulations turning the inputs sequentially active revealed that when setting either TNF, TGFB1 or Il6 on, all TFs are expressed (Figure 6 a-c). When mimicking the downregulation of MAP2K1, ILK, DAXX for the dataset of responders/non responders to anti-TNF treatment, we observed that when setting TNF and TGFB1 on, all TFs are expressed (Figure 6 d,e). However, when IL6 is set on, only the TF NFKBIA is expressed (Figure 6 f). Finally, when mimicking the downregulation of DAXX and NFKB1 for the mutation carrier from DisGeNET, we observe that when setting either TNF, TGFB1 or IL6 on, all TFs are expressed (Figure 6 g-i).

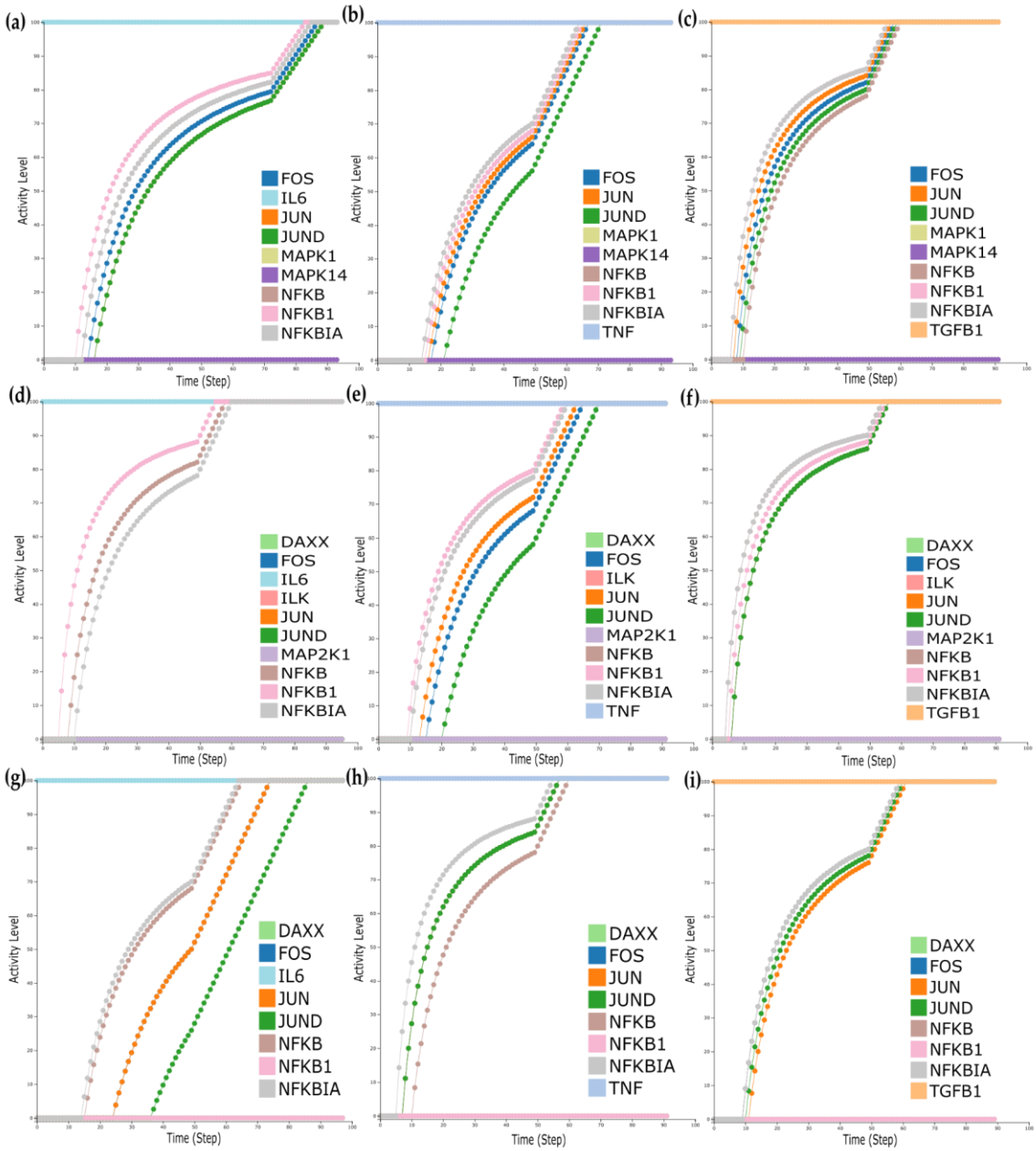


Figure 6. Real time simulations with Cell Collective. 1) For the dataset of before and after anti-TNF treatment of RA patients, where MAPK14 and MAPK1 were found downregulated **(a-c)**. **(a)** Simulations with IL6 activity set to 100%. **(b)** Simulations with TNF activity set to 100% **(c)** Simulations with TGFB1 activity set to 100%. **2)** For the dataset of responders/non responders where MAP2K1, ILK, and DAXX were found downregulated **(d-f)**. **(d)** Simulations with IL6 activity set to 100%. **(e)** Simulations with TNF activity set to 100%. **(f)** Simulations with TGFB1 activity set to 100%. 3) For the dataset of the mutation carrier. **(g-i)**. **(g)** Simulations with IL6 activity set to 100%. **(h)** Simulations with TNF activity set to 100%. **(i)** Simulations with TGFB1 activity set to 100%.

Dose response and Sensitivity analysis

For the dose response we studied five different initial conditions shown in Table 3 that mimic the effect of different scenarios in combination to TNF activity status. The first condition corresponds to TNF blockage and simultaneous impairment of IL6 and TGFB1 signalling, the second corresponds to having the IL6 cascade active, the third to having the TGFB1 active, the fourth both IL6 and TGFB1 active and lastly, the fifth condition mimics what happens to the system when only the TNF input is active.

Table 3. Initial conditions for dose response analysis.

Initial conditions	Input state
1	All inputs inactive
2	IL6 active
3	TGFB1 active
4	IL6+TGFB1 active
5	TNF active

We performed dose response analysis for all aforementioned conditions and observed that the expression of the TFs is dose dependent for TNF, TGFB1, and IL6 (Figure 7 b, e and f), while simultaneous activation of IL6 and TGFB1 cascades has a synergistic effect causing an increase of the activation levels of the TFs even for lower doses of IL6 and TGFB1 (Figure 7 c and d).

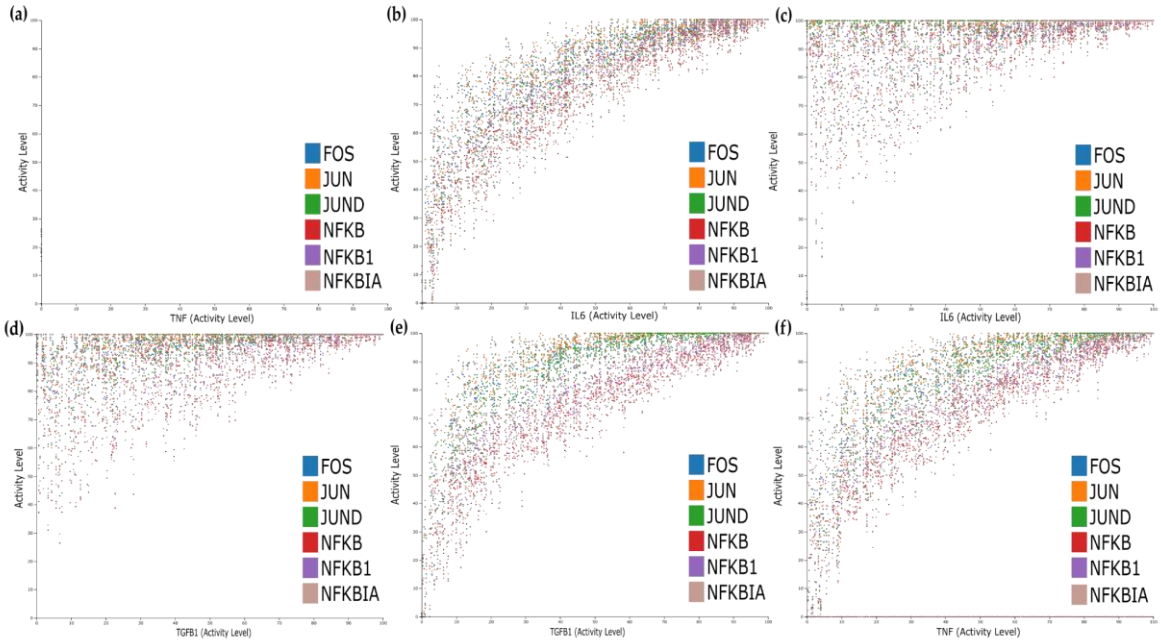


Figure 7. Dose response analysis. (a) All input inactives. **(b)** IL6 active. (**c)** IL6 and TGFB1 active (TGFB1 view). **(d)** IL6 and TGFB1 active (IL6 view). **(e)** TGFB1 active. **(f)** TNF active.

Next, we wanted to see how the downregulation of the TFs observed after anti TNF treatment could be counterbalanced by the other pathways, given that in non responders the expression of these TFs was kept intact, despite the treatment. We performed environmental sensitivity analysis to identify which of the two model inputs (IL6 and TGFB1) has the greatest impact on the up regulation of the TFs included in the model, when TNF activity is blocked. The results depicted in Supplementary files, showed that indeed the TFs could be upregulated in the absence of TNF activity, for a combination of activity ranges of the other two inputs (Suppl Figures S4).

Wild type stable state analysis and KO simulations

We performed stable state analysis for the model using the software GINsim. The analysis for the wild type (no perturbations) revealed five steady states (fixed points) and no complex attractor. The configurations of these five stable states as far as the

molecules of interest are concerned (grey, blue and pink nodes of Figure 8) are shown in Table 4.

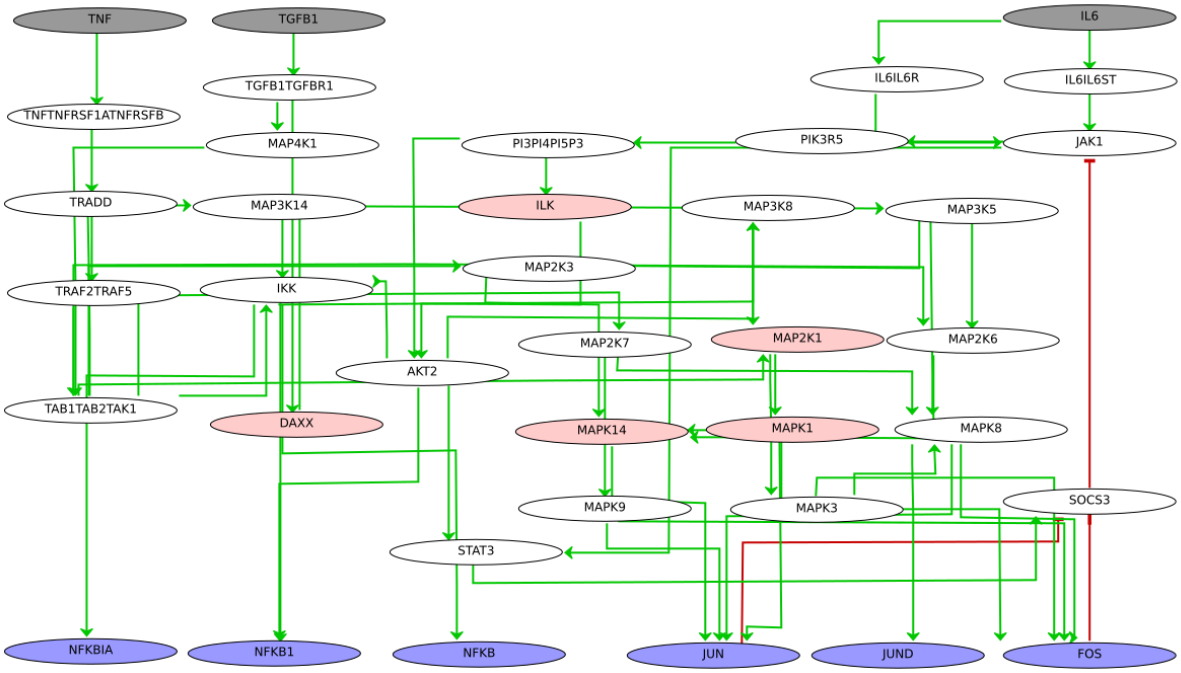


Figure 8. Boolean network comprising the three signalling cascades for TNF, TGFB1 and IL6. Boolean rules were inferred using the tool CaSQ (see Materials and Methods for a step by step model inference). Inputs are depicted in grey, TFs of interest are depicted in purple and intermediate nodes affected by the drug treatment or identified as mutation carriers are depicted in pink. Green arrows denote activation and red, blunt arrows denote inhibition.

Table 4. Stable states of the Boolean network (wild type).

Steady states	TNF	IL6	TGFB1	JUN	FOS	JUND	NFKBIA	DAXX	ILK	NFKB1	MAP2K1	MAPK1	MAPK14
ss1	0	0	0	0	0	0	0	0	0	0	0	0	0
ss2	0	1	0	1	1	1	1	0	1	1	1	1	1
ss3	0	1	1	1	1	1	1	1	1	1	1	1	1
ss4	1	1	0	1	1	1	1	0	1	1	1	1	1
ss5	1	1	1	1	1	1	1	1	1	1	1	1	1

The analysis shows that for the TFs identified as master regulators, the activation of IL6 or IL6 and TGFB1, is capable of positively regulating their expression, even in the

presence of the anti-TNF treatment (TNF=0). Blocking the TNF cascade is able to completely shut down their expression only if combined with the blocking of IL6 and TGFB1 cascades. Regarding DAXX, it is actively expressed only when TGFB1 is activated and ILK is dependent on the activation of IL6. The MAPK molecules are dependent on the activation of IL6 and TGFB1 and do not seem to be impacted by the blocking of TNF.

Next, we created a reduced version of the Boolean modelling using the reduction function of the software GINsim in order to perform *in silico* experiments with combined perturbations. The reduced Boolean model comprised 23 nodes and for the analysis we created virtual KOs (knock outs) for a) MMP14 and MAPK1, b) DAXX, ILK and MAP2K1, and c) DAXX and NFKB1 to mimic the effects of the anti TNF treatment and also the mutation carriers, identified previously. For the simulations we set the initial conditions for the TNF to zero, and let the other inputs free, while setting the initial condition for all intermediate nodes to zero as well.

The results of the *in silico* experiments are shown in Tables 5, 6 and 7, for the molecules of interest and for each of the three conditions, respectively. For each set of conditions, the system was able to reach three steady states. For the first set of conditions, we observe that besides DAXX that is strictly TGFB1 dependent (Table 5, ss2 and ss3), all TFs can get activated with the presence of IL6 or IL6 and TGFB1, despite the TNF blockage and the downregulation of MAPK14 and MAPK1.

Table 5. Stable states of the Boolean network (MAPK1, MAPK14 KO, input TNF=0)

Steady states	TNF	IL6	TGFB1	JUN	FOS	JUND	NFKBIA	DAXX	ILK	NFKB1	MAP2K1	MAPK1	MAPK14
ss1	0	0	0	0	0	0	0	0	0	0	0	0	0
ss2	0	1	0	1	1	1	1	0	1	1	1	0	0
ss3	0	1	1	1	1	1	1	1	1	1	1	0	0

For the second set of conditions, we observe that when TNF is blocked and DAXX, ILK and MAP2K1 are downregulated, IL6 signal alone is not enough to activate the TFs JUN, FOS, JUND and the kinases MAPK14 and MAPK1 (Table 6, ss2). When both IL6 and TGFB1 signals are on, all TFs are activated, and the activity level of kinase MAPK14 restored (Table 6, ss3).

Table 6. Stable states of the Boolean network (DAXX, ILK and MAP2K1 KO, input TNF=0)

Steady states	TNF	IL6	TGFB1	JUN	FOS	JUND	NFKBIA	DAXX	ILK	NFKB1	MAP2K1	MAPK1	MAPK14
ss1	0	0	0	0	0	0	0	0	0	0	0	0	0
ss2	0	1	0	0	0	0	1	0	0	1	0	0	0
ss3	0	1	1	1	1	1	1	0	0	1	0	0	1

Lastly, we simulated the effects of the mutation carriers, as identified by DisGeNet, and the TNF blockage on the activity of the TFs and the kinases in our network. In Table 7 we observe that despite TNF blockage and DAXX and NFKB1 downregulation, all identified TFs and kinases are activated in the presence of IL6 or for IL6 and TGFB1 combined activity.

Table 7. Stable states of the Boolean network (DAXX and NFKB1 KO, input TNF=0)

Steady states	TNF	IL6	TGFB1	JUN	FOS	JUND	NFKBIA	DAXX	ILK	NFKB1	MAP2K1	MAPK1	MAPK14
ss1	0	0	0	0	0	0	0	0	0	0	0	0	0
ss2	0	1	0	1	1	1	1	0	1	0	1	1	1
ss3	0	1	1	1	1	1	1	0	1	0	1	1	1

Discussion

In the present work we make efforts to combine gene co-regulation with mechanistic signalling cascades that provide information about upstream regulation. We make use of the integrative RA network to analyze transcriptomic data regarding anti-TNF treatment and also map information about known disease-associated mutation carriers. Lastly, we make use of the tool CaSQ to add Boolean dynamics to a subnetwork of interest to mimic the effects of the anti-TNF treatment and estimate the impact of IL6 and TGFB1, as well as the impact of the downregulated genes, on the activation profile of the identified TFs .

The nineteen TFs identified as master regulators, have been implicated in RA and/ or autoimmunity as supported by the literature evidence. Six out of the nineteen TFs were also present in the RA map, a state-of-the-art mechanistic network for the disease that was built using manual curation. These six TFs, namely JUN, JUND, FOS, NFKBIA, ETS1 and TNFAIP3 were used as a functional overlap between the co-regulation and the signalling events, enabling us to obtain a network comprising upstream cascades, active TFs and target genes.

We used the integrative network as a template to analyse two independent datasets regarding anti-TNF treatment. We observed downregulation of some of the TFs previously identified as master regulators. To study the impact of the treatment in parallel with the activity of other signalling cascades, we extracted subgraphs from the integrative network. We selected the cascades of TNF, IL6 and TGFB1, up to the first affected TF to reduce complexity and focus on the upstream regulators.

We selected IL6, as it is one of the targets of the biologic treatment in rheumatoid arthritis [36,41–43] and TGF beta because it is an immunomodulatory cytokine highly expressed in RA patients, with a role that is yet to be determined [44–46]. We adjusted

the map to model framework described in Aghamiri et al [29], to obtain an executable Boolean subnetwork to perform *in silico* analysis. As demonstrated from the real time simulations and the dose response analysis, both IL6 and TGFB1 cascades could affect the expression of the TFs and, as seen from the component sensitivity analysis, IL6 and TGFB1 could even counterbalance the downregulation of the studied TFs caused by the TNF blockage.

The steady state analysis confirmed the real time simulation results showing that for the TFs identified as master regulators, the activation of IL6 or IL6 and TGFB1 cascades is capable of positively regulating their expression. Blocking the TNF cascade is able to completely shut down the expression of these TFs only if combined with the blocking of IL6 and TGFB1 cascades. Towards this direction, dual targeted therapies have been proposed, either with the development of dual target agents, blocking simultaneously IL6 and TNF for example, [47] or by administering two biologics at the same time. However, the administration of combined biologics has been linked to increased adverse effects and it is currently under study to evaluate better dosage schemes [48,49].

Simulations with combined KOs mimicking the effect of anti-TNF treatment in combination with the downregulation of genes observed in the analyzed datasets, confirmed the dependency of the TFs activation state on the presence of inputs, and further highlighted specific conditions. For example, when TNF is blocked and DAXX, ILK and MAP2K1 are downregulated, IL6 signal alone is not enough to activate the TFs JUN, FOS, JUND and the kinases MAPK14 and MAPK1 (Table 6, ss2). However, when both IL6 and TGFB1 signals are on, all TFs are activated, and the activity level of kinase MAPK14 restored (Table 6, ss3). MAPK14 (p38a kinase) and MAPK1 (ERK2), are two proteins known to play a pivotal role in RA and are activated by a variety of signals,

including cytokines such as TNF and IL6, but also TGF beta [50]. While p38 had been proposed as a potential target to reduce destruction of bone and cartilage, p38 inhibitors have given disappointing results in regards to therapeutic efficacy [51,52].

As seen in Table 5, suppression of MAPK14 (p38a) does not inhibit the activation of the identified as master regulators TFs, even in the presence of anti-TNF treatment, as other inputs, such as IL6 or TGF beta can counterbalance the effects. Regarding ERK inhibitors, limited data are available which may be due to a lack of efficient pharmacological inhibitors [53]. In older studies, FR180204, an ERK inhibitor, had demonstrated effectiveness against mouse collagen-induced arthritis [54], but there was no significant follow up. In our model, ERK2 inhibition (MAPK1) does not seem to significantly impact the activation of its downstream target TFs, JUN and FOS, as other regulators are also able to activate them.

While resistance to TNF therapy is a common event in the treatment of RA, the reasons behind its mechanisms are still unclear [55]. Currently there is no way of predicting which patient will respond or not to targeted therapy [56]. Executable, integrated networks can help towards personalized models, as mapping dysregulated genes could reveal potential impacted pathways shedding light onto the response to therapy. Dynamic analysis and *in silico* simulations can also inform about the outcome of combined perturbations, predicting the emergent behaviour of the system. Integrating multi-omic data is a key step in understanding pathogenetic mechanisms of multifactorial diseases, where one level of information does not suffice to explain the complex phenotypic traits. Patient-level analysis, by projecting patient-specific data onto a global network, and analysing the effects of patient-specific mutations and DEGs, in combination to treatment effects, could inform about the possibilities of success of a given therapy. Larger patient cohorts and more efficient computational techniques that

would allow simulations in a greater scale could enhance the robustness and the predictive power of such models, helping to understand the response or non response to a given therapy at a patient level.

Supplementary Materials

Supplementary Figures and Table are also available here for review purposes:

https://docs.google.com/document/d/1pW7EuVMr4WOK4qyV2SDVOr8-pMle7Dbp9TK0EXNhs8o/edit

Figure S1. Principal component analysis (PCA) in human blood cells with affected and unaffected Rheumatoid arthritis patients.
The PCA shows 95 samples whose matrix expression data was downloaded from the GEO database (GSE117769). A variance stabilizing transformation was carried on matrix expression data.

Figure S2. Global network for RA with associated DEG from responders/non responders RA patients data to anti-TNF treatment.
Overlapping RA DEG from Adalimumab treatment and Etarnecept treatment data are shown in brown (11) and pink (4) respectively while non-overlapping genes/proteins are shown in grey (599).

Figure S3. Global network for RA with associated DEG from untreated and treated RA patients data to anti-TNF treatment.
Overlapping RA DEG are shown in green (101) while non-overlapping genes/proteins are shown in grey (513).

S4 Fig. Environment sensitivity analysis. TFs could be upregulated in the absence of TNF activity, for a combination of activity ranges of the other two inputs (IL6 and TGFB1).
Upper part of each image corresponds to the correlation (positive above 0.0 and negative below it) of the sensitivity of the external components (shown in blue box) to optimize the activation of chosen internal components (TFs here). Lower part of the image corresponds to the box plots to show the range of activity percentage of external components (IL6 in blue, TGFB1 in orange and TNF as black line as it is chosen to be off) for the optimization of the chosen internal components activation. **(a)** optimize JUND activation. **(b)** optimize JUN activation. **(c)** optimize FOS activation. **(d)** optimize NFKB1 activation. **(e)** optimize NFKBIA activation.

Table S1. Transcription factors identified from CoRegNet and their involvement in RA based on literature evidence.

Author Contributions

Conceptualization, A.N.; methodology, A.N. and M.E.; software, Q.M., D.D.M., V.S., M.E and A.N.; validation, A.N., Q.M. and V.S.; formal analysis, Q.M., D.D.M., M.E. and A.N.; investigation, Q.M., D.D.M. M.E. and A.N.; resources, V.C., E.P.T. and A.N.; data curation, Q.M., D.D.M and V.S.; writing—original draft preparation, Q.M., D.D.M., V.S. and A.N.; writing—review and editing, Q.M., D.D.M., V.S., V.C., M.E., E.P.T. and A.N.; visualization, Q.M. and V.S.; supervision, M.E. and A.N.; project administration, A.N.; funding acquisition, V.C., E.P.T and A.N. All authors have read and agreed to the published version of the manuscript.

Funding

This work is supported by the doctorate program of the University of Paris Saclay, France and Fondagen, Genopole.

Data Availability Statement

Datasets (GSE117769, GSE129705, GSE138747 and DisGeNET variants) used for the analysis are publicly available. All data and code used to generate results, including networks inference, differential expression analysis, network visualisation is available on a GitLab repository at https://gitlab.com/genhotel/inference-of-a-global-integrative-network-for-rheumatoid-arthritis/

The Shiny app login is: **genhotel** and the password is: **3886**

Acknowledgements

We would like to thank Dr Smahane Chalabi, GenHotel, UEVE for her comments on the statistical treatment of the dataset used to infer the CoRegNet object.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Ngo, S.T.; Steyn, F.J.; McCombe, P.A. Gender Differences in Autoimmune Disease. *Frontiers in Neuroendocrinology* **2014**, *35*, 347–369, doi:10.1016/j.yfrne.2014.04.004.
2. Smolen, J.S.; Aletaha, D.; Barton, A.; Burmester, G.R.; Emery, P.; Firestein, G.S.; Kavanaugh, A.; McInnes, I.B.; Solomon, D.H.; Strand, V.; et al. Rheumatoid Arthritis. *Nat Rev Dis Primers* **2018**, *4*, 18001, doi:10.1038/nrdp.2018.1.
3. Deane, K.D.; Demoruelle, M.K.; Kelmenson, L.B.; Kuhn, K.A.; Norris, J.M.; Holers, V.M. Genetic and Environmental Risk Factors for Rheumatoid Arthritis. *Best Pract Res Clin Rheumatol* **2017**, *31*, 3–18, doi:10.1016/j.berh.2017.08.003.
4. Batliwalla, F.M.; Baechler, E.C.; Xiao, X.; Li, W.; Balasubramanian, S.; Khalili, H.; Damle, A.; Ortmann, W.A.; Perrone, A.; Kantor, A.B.; et al. Peripheral Blood Gene Expression Profiling in Rheumatoid Arthritis. *Genes Immun* **2005**, *6*, 388–397, doi:10.1038/sj.gene.6364209.
5. Edwards, C.J.; Feldman, J.L.; Beech, J.; Shields, K.M.; Stover, J.A.; Trepicchio, W.L.; Larsen, G.; Foxwell, B.M.; Brennan, F.M.; Feldmann, M.; et al. Molecular Profile of Peripheral Blood Mononuclear Cells from Patients with Rheumatoid Arthritis. *Mol Med* **2007**, *13*, 40–58, doi:10.2119/2006-000056.Edwards.
6. Micsik, T.; Lőrincz, A.; Gál, J.; Schwab, R.; Peták, I. MDR-1 and MRP-1 Activity in Peripheral Blood Leukocytes of Rheumatoid Arthritis Patients. *Diagn Pathol* **2015**, *10*, 216, doi:10.1186/s13000-015-0447-1.
7. Kuuliala, K.; Kuuliala, A.; Koivuniemi, R.; Kautiainen, H.; Repo, H.; Leirisalo-Repo, M. Baseline JAK Phosphorylation Profile of Peripheral Blood Leukocytes, Studied by Whole Blood Phosphospecific Flow Cytometry, Is Associated with 1-Year Treatment Response in Early Rheumatoid Arthritis. *Arthritis Res Ther* **2017**, *19*, doi:10.1186/s13075-017-1278-0.
8. Li, X.; Lei, Y.; Gao, Z.; Zhang, B.; Xia, L.; Lu, J.; Shen, H. Effect of IL-34 on T Helper 17 Cell Proliferation and IL-17 Secretion by Peripheral Blood Mononuclear Cells from Rheumatoid Arthritis Patients. *Sci Rep* **2020**, *10*, doi:10.1038/s41598-020-79312-z.
9. Farutin, V.; Prod'homme, T.; McConnell, K.; Washburn, N.; Halvey, P.; Etzel, C.J.; Guess, J.; Duffner, J.; Getchell, K.; Meccariello, R.; et al. Molecular Profiling of Rheumatoid Arthritis Patients Reveals an Association between Innate and Adaptive Cell Populations and Response to Anti-Tumor Necrosis

Factor. *Arthritis Research & Therapy* **2019**, *21*, 216, doi:10.1186/s13075-019-1999-3.
10. Eguchi, R.; Karim, M.B.; Hu, P.; Sato, T.; Ono, N.; Kanaya, S.; Altaf-Ul-Amin, Md. An Integrative Network-Based Approach to Identify Novel Disease Genes and Pathways: A Case Study in the Context of Inflammatory Bowel Disease. *BMC Bioinformatics* **2018**, *19*, 264, doi:10.1186/s12859-018-2251-x.
11. Karimizadeh, E.; Sharifi-Zarchi, A.; Nikaein, H.; Salehi, S.; Salamatian, B.; Elmi, N.; Gharibdoost, F.; Mahmoudi, M. Analysis of Gene Expression Profiles and Protein-Protein Interaction Networks in Multiple Tissues of Systemic Sclerosis. *BMC Med Genomics* **2019**, *12*, 199, doi:10.1186/s12920-019-0632-2.
12. Haghjoo, N.; moeini, A.; Masoudi-Nejad, A. Introducing a Panel for Early Detection of Lung Adenocarcinoma by Using Data Integration of Genomics, Epigenomics, Transcriptomics and Proteomics. *Experimental and Molecular Pathology* **2020**, *112*, 104360, doi:10.1016/j.yexmp.2019.104360.
13. Sahu, A.; Chowdhury, H.A.; Gaikwad, M.; Chongtham, C.; Talukdar, U.; Phukan, J.K.; Bhattacharyya, D.K.; Barah, P. Integrative Network Analysis Identifies Differential Regulation of Neuroimmune System in Schizophrenia and Bipolar Disorder. *Brain, Behavior, & Immunity - Health* **2020**, *2*, 100023, doi:10.1016/j.bbih.2019.100023.
14. Greenfield, A.; Hafemeister, C.; Bonneau, R. Robust Data-Driven Incorporation of Prior Knowledge into the Inference of Dynamic Regulatory Networks. *Bioinformatics* **2013**, *29*, 1060–1067, doi:10.1093/bioinformatics/btt099.
15. Zuo, Y.; Cui, Y.; Yu, G.; Li, R.; Ressom, H.W. Incorporating Prior Biological Knowledge for Network-Based Differential Gene Expression Analysis Using Differentially Weighted Graphical LASSO. *BMC Bioinformatics* **2017**, *18*, 99, doi:10.1186/s12859-017-1515-1.
16. Benedetti, E.; Pučić-Baković, M.; Keser, T.; Gerstner, N.; Büyüközkan, M.; Štambuk, T.; Selman, M.H.J.; Rudan, I.; Polašek, O.; Hayward, C.; et al. A Strategy to Incorporate Prior Knowledge into Correlation Network Cutoff Selection. *Nature Communications* **2020**, *11*, 5153, doi:10.1038/s41467-020-18675-3.
17. Nicolle, R.; Radvanyi, F.; Elati, M. CoRegNet: Reconstruction and Integrated Analysis of Co-Regulatory Networks. *Bioinformatics* **2015**, *31*, 3066–3068, doi:10.1093/bioinformatics/btv305.
18. Singh, V.; Kalliolias, G.D.; Ostaszewski, M.; Veyssiere, M.; Pilalis, E.; Gawron, P.; Mazein, A.; Bonnet, E.; Petit-Teixeira, E.; Niarakis, A. RA-Map: Building a State-of-the-Art Interactive Knowledge Base for Rheumatoid Arthritis. *Database (Oxford)* **2020**, *2020*, doi:10.1093/database/baaa017.
19. Mazein, A.; Ostaszewski, M.; Kuperstein, I.; Watterson, S.; Le Novère, N.; Lefaudeux, D.; De Meulder, B.; Pellet, J.; Balaur, I.; Saqi, M.; et al. Systems Medicine Disease Maps: Community-Driven Comprehensive Representation of Disease Mechanisms. *NPJ Syst Biol Appl* **2018**, *4*, 21, doi:10.1038/s41540-018-0059-y.
20. Ostaszewski, M.; Gebel, S.; Kuperstein, I.; Mazein, A.; Zinovyev, A.; Dogrusoz, U.; Hasenauer, J.; Fleming, R.M.T.; Le Novère, N.; Gawron, P.; et al. Community-Driven Roadmap for Integrated Disease Maps. *Brief Bioinform* **2019**, *20*, 659–670, doi:10.1093/bib/bby024.
21. Niarakis, A.; Helikar, T. A Practical Guide to Mechanistic Systems Modeling in Biology Using a Logic-Based Approach. *Briefings in Bioinformatics* **2020**, doi:10.1093/bib/bbaa236.

22. Schwab, J.D.; Kühlwein, S.D.; Ikonomi, N.; Kühl, M.; Kestler, H.A. Concepts in Boolean Network Modeling: What Do They All Mean? *Computational and Structural Biotechnology Journal* **2020**, *18*, 571–582, doi:10.1016/j.csbj.2020.03.001.
23. Abou-Jaoudé, W.; Traynard, P.; Monteiro, P.T.; Saez-Rodriguez, J.; Helikar, T.; Thieffry, D.; Chaouiya, C. Logical Modeling and Dynamical Analysis of Cellular Networks. *Frontiers in Genetics* **2016**, *7*, doi:10.3389/fgene.2016.00094.
24. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biology* **2014**, *15*, doi:10.1186/s13059-014-0550-8.
25. Chebil, I.; Nicolle, R.; Santini, G.; Rouveirol, C.; Elati, M. Hybrid Method Inference for the Construction of Cooperative Regulatory Network in Human. *IEEE Trans Nanobioscience* **2014**, *13*, 97–103, doi:10.1109/TNB.2014.2316920.
26. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **1995**, *57*, 289–300.
27. Hoksza, D.; Gawron, P.; Ostaszewski, M.; Smula, E.; Schneider, R. MINERVA API and Plugins: Opening Molecular Network Analysis and Visualization to the Community. *Bioinformatics* **2019**, *35*, 4496–4498, doi:10.1093/bioinformatics/btz286.
28. Le Novère, N.; Hucka, M.; Mi, H.; Moodie, S.; Schreiber, F.; Sorokin, A.; Demir, E.; Wegner, K.; Aladjem, M.I.; Wimalaratne, S.M.; et al. The Systems Biology Graphical Notation. *Nat Biotechnol* **2009**, *27*, 735–741, doi:10.1038/nbt.1558.
29. Aghamiri, S.S.; Singh, V.; Naldi, A.; Helikar, T.; Soliman, S.; Niarakis, A. Automated Inference of Boolean Models from Molecular Interaction Maps Using CaSQ. *Bioinformatics* **2020**, *36*, 4473–4482, doi:10.1093/bioinformatics/btaa484.
30. Csardi, G.; Nepusz, T. The Igraph Software Package for Complex Network Research. *InterJournal* **2006**, *Complex Systems*, 1695.
31. Gustavsen; A, J.; Pai; Shraddha; Isserlin; Ruth; Demchak; Barry; Pico; R, A. RCy3: Network Biology Using Cytoscape from within R. *F1000Research* **2019**, doi:10.12688/f1000research.20887.3.
32. Piñero, J.; Ramírez-Anguita, J.M.; Saüch-Pitarch, J.; Ronzano, F.; Centeno, E.; Sanz, F.; Furlong, L.I. The DisGeNET Knowledge Platform for Disease Genomics: 2019 Update. *Nucleic Acids Res* **2020**, *48*, D845–D855, doi:10.1093/nar/gkz1021.
33. Chang, W.; Cheng, J.; Allaire, J.J.; Xie, Y.; McPherson, J. *Shiny: Web Application Framework for R*; 2020;
34. Shah, O.; Shannon, P. *CyjShiny: CyjShiny*; 2020;
35. Chaouiya, C.; Naldi, A.; Thieffry, D. Logical Modelling of Gene Regulatory Networks with GINsim. *Methods Mol Biol* **2012**, *804*, 463–479, doi:10.1007/978-1-61779-361-5_23.
36. Ogata, A.; Kato, Y.; Higa, S.; Yoshizaki, K. IL-6 Inhibitor for the Treatment of Rheumatoid Arthritis: A Comprehensive Review. *Mod Rheumatol* **2019**, *29*, 258–267, doi:10.1080/14397595.2018.1546357.
37. Gonzalo-Gil, E.; Criado, G.; Santiago, B.; Dotor, J.; Pablos, J.L.; Galindo, M. Transforming Growth Factor (TGF)-β Signalling Is Increased in Rheumatoid Synovium but TGF-β Blockade Does Not Modify Experimental Arthritis. *Clin Exp Immunol* **2013**, *174*, 245–255, doi:10.1111/cei.12179.

38. Hall, B.; Niarakis, A. Data Integration in Logic-Based Models of Biological Mechanisms. **2021**, doi:10.20944/preprints202105.0219.v1.
39. Funahashi, A.; Matsuoka, Y.; Jouraku, A.; Morohashi, M.; Kikuchi, N.; Kitano, H. CellDesigner 3.5: A Versatile Modeling Tool for Biochemical Networks. *Proceedings of the IEEE* **2008**, *96*, 1254–1265, doi:10.1109/JPROC.2008.925458.
40. Helikar, T.; Kowal, B.; McClenathan, S.; Bruckner, M.; Rowley, T.; Madrahimov, A.; Wicks, B.; Shrestha, M.; Limbu, K.; Rogers, J.A. The Cell Collective: Toward an Open and Collaborative Approach to Systems Biology. *BMC Systems Biology* **2012**, *6*, 96, doi:10.1186/1752-0509-6-96.
41. Hennigan, S.; Kavanaugh, A. Interleukin-6 Inhibitors in the Treatment of Rheumatoid Arthritis. *Ther Clin Risk Manag* **2008**, *4*, 767–775.
42. Woodrick, R.; Ruderman, E.M. Anti-Interleukin-6 Therapy in Rheumatoid Arthritis. *Bull NYU Hosp Jt Dis* **2010**, *68*, 211–217.
43. Choy, E.H.; De Benedetti, F.; Takeuchi, T.; Hashizume, M.; John, M.R.; Kishimoto, T. Translating IL-6 Biology into Effective Treatments. *Nat Rev Rheumatol* **2020**, *16*, 335–345, doi:10.1038/s41584-020-0419-z.
44. Sakuma, M.; Hatsushika, K.; Koyama, K.; Katoh, R.; Ando, T.; Watanabe, Y.; Wako, M.; Kanzaki, M.; Takano, S.; Sugiyama, H.; et al. TGF-β Type I Receptor Kinase Inhibitor down-Regulates Rheumatoid Synoviocytes and Prevents the Arthritis Induced by Type II Collagen Antibody. *International Immunology* **2007**, *19*, 117–126, doi:10.1093/intimm/dxl128.
45. Guo, L. TGF Beta in the Rheumatoid Arthritis Research. *European Journal of BioMedical Research* **2017**, *3*, 5–8, doi:10.18088/ejbmr.3.4.2017.pp5-8.
46. Zhou, G.; Sun, X.; Qin, Q.; Lv, J.; Cai, Y.; Wang, M.; Mu, R.; Lan, H.; Wang, Q.-W. Loss of Smad7 Promotes Inflammation in Rheumatoid Arthritis. *Front. Immunol.* **2018**, *9*, doi:10.3389/fimmu.2018.02537.
47. A Dual Target-Directed Agent against Interleukin-6 Receptor and Tumor Necrosis Factor α Ameliorates Experimental Arthritis | Scientific Reports Available online: https://www.nature.com/articles/srep20150 (accessed on 23 June 2021).
48. Boleto, G.; Kanagaratnam, L.; Dramé, M.; Salmon, J.-H. Safety of Combination Therapy with Two BDMARDs in Patients with Rheumatoid Arthritis: A Systematic Review and Meta-Analysis. *Seminars in Arthritis and Rheumatism* **2019**, *49*, 35–42, doi:10.1016/j.semarthrit.2018.12.003.
49. Biologic Combination Therapy for RA May Increase Risk for Side Effects Available online: https://rheumatology.medicinematters.com/rheumatoid-arthritis-/biologics/biologic-combination-therapy-for-ra-may-increase-risk-for-side-e/16387312 (accessed on 23 June 2021).
50. Canovas, B.; Nebreda, A.R. Diversity and Versatility of P38 Kinase Signalling in Health and Disease. *Nat Rev Mol Cell Biol* **2021**, *22*, 346–366, doi:10.1038/s41580-020-00322-w.
51. Clark, A.R.; Dean, J.L. The P38 MAPK Pathway in Rheumatoid Arthritis: A Sideways Look. *The Open Rheumatology Journal* **2012**, *6*, doi:10.2174/1874312901206010209.
52. Haller, V.; Nahidino, P.; Forster, M.; Laufer, S.A. An Updated Patent Review of P38 MAP Kinase Inhibitors (2014-2019). *Expert Opinion on Therapeutic Patents* **2020**, *30*, 453–466, doi:10.1080/13543776.2020.1749263.
53. Bonilla-Hernán, M.G.; Miranda-Carús, M.E.; Martin-Mola, E. New Drugs beyond Biologics in Rheumatoid Arthritis: The Kinase Inhibitors. *Rheumatology* **2011**,

50, 1542–1550, doi:10.1093/rheumatology/ker192.
54. Ohori, M. ERK Inhibitors as a Potential New Therapy for Rheumatoid Arthritis. *Drug News Perspect* **2008**, *21*, 245–250, doi:10.1358/DNP.2008.21.5.1219006.
55. Sidiropoulos, P.I.; Boumpas, D.T. Differential Drug Resistance to Anti-tumour Necrosis Factor Agents in Rheumatoid Arthritis. *Ann Rheum Dis* **2006**, *65*, 701–703, doi:10.1136/ard.2005.049890.
56. Rheumatoid Arthritis: A Case for Personalized Health Care? - Karsdal - 2014 - Arthritis Care & Research - Wiley Online Library Available online: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.22289 (accessed on 1 July 2021).