Computational modeling of the cellular interplay in rheumatoid arthritis. Deciphering the role of resident and immune cells in the arthritic joint

Naouel Zerrouk

THESE DE DOCTORAT

NNT : 2023UPASL096

université PARIS-SACLAY

Computational modeling of the cellular interplay in rheumatoid arthritis. Deciphering the role of resident and immune cells in the arthritic joint.

Modélisation computationnelle des interactions cellulaires dans la polyarthrite rhumatoïde. Déchiffrer le rôle des cellules résidentes et des cellules immunitaires dans l'articulation atteinte.

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°577 : Structure et Dynamique des Systèmes Vivants (SDSV)
Spécialité de doctorat : Biologie computationnelle
Graduate School : Life Sciences and Health
Référent : Université d'Évry-Val d'Essonne

Thèse préparée dans l'unité de recherche Laboratoire Européen de Recherche pour la Polyarthrite rhumatoïde – Genhotel (Université Paris-Saclay, Univ Evry) et l'équipe Omics Data Science (Sanofi R&D, Data and Data Science), sous la direction d'**Anna NIARAKIS**, Professeure des universités, et la supervision de **Franck AUGE**, Docteur

Thèse soutenue à Paris Saclay, le 31 Octobre 2023, par

Naouel ZERROUK

Composition du Jury

Membres du jury avec voix délibérative

Liesbet GERIS Professeure, Université de Liège	Présidente
Jasmin FISHER Professeure, University College London	Rapporteure/Examinatrice
Benjamin HALL Docteur, University College London	Examinateur
Gaetane NOCTURNE Maîtresse de conférences - praticienne hospitalière, Université Paris-Saclay	Examinatrice

université PARIS-SACLAY
ÉCOLE DOCTORALE
Structure et dynamique des systèmes vivants (SDSV)

Titre : Modélisation computationnelle des interactions cellulaires dans la polyarthrite rhumatoïde. Déchiffrer le rôle des cellules résidentes et des cellules immunitaires dans l'articulation

Mots clés : Polyarthrite rhumatoïde, Modélisation computationnelle à grande échelle, Modèle de macrophage, Modèle multicellulaire discret, Immuno-inflammation, Crosstalks cellulaires

Résumé : La polyarthrite rhumatoïde (PR) est une maladie auto-immune dont l'étiologie n'est pas totalement comprise. La maladie se manifeste par des douleurs, des raideurs et des gonflements articulaires, ainsi que par une synovite chronique. L'état pathologique du tissu synovial dans la PR est le résultat d'interactions cellulaires entre les synoviocytes résidents et les cellules immunitaires. Ces interactions cellulaires définissent de nombreux aspects de la maladie et pourraient constituer des cibles thérapeutiques potentielles. Il n'existe aujourd'hui aucun traitement curatif de la PR. Les thérapies proposées ne peuvent qu'atténuer les symptômes et sont souvent associées à plusieurs effets indésirables. Dans ce projet de thèse, je souhaite étudier les interactions cellulaires impliquées dans la maladie, comprendre le mécanisme d'action des traitements actuels de la PR et tester de nouvelles combinaisons de médicaments. Je souhaite également identifier de nouvelles cibles thérapeutiques et proposer d'éventuels repositionnements de médicaments. Dans cette optique, j'ai développé une approche computationnelle et systémique afin d'obtenir une représentation multicellulaire des mécanismes impliqués dans la PR. Pour cela, j'ai construit l'atlas de la PR, un atlas qui récapitule les connaissances actuelles sur les cascades intracellulaires impliquées dans la maladie. L'atlas est basé sur une curation manuelle de la littérature et des bases de données, et enrichi avec des données omiques. Il comprend quatre cartes spécifiques au macrophage, de phénotypes M1 et M2, au fibroblaste et au lymphocyte CD4+ de type Th1. Afin d'obtenir une représentation multicellulaire de l'articulation, j'ai créé la première carte multicellulaire de la PR en intégrant les interactions intercellulaires bidirectionnelles identifiées dans la littérature et les données omiques. Cependant, cette carte reste limitée par sa nature statique lorsqu'il s'agit de faire des prédictions. Pour étudier le comportement émergent de ce système, j'ai utilisé l'outil CaSQ pour convertir les cartes en modèles Booléens. Ces modèles logiques ne nécessitent pas de paramètres cinétiques et peuvent intégrer des informations provenant de la littérature et de données omiques, ce qui en fait une option appropriée pour l'analyse des réseaux à grande échelle. La recherche d'attracteurs est un point clé dans l'analyse des modèles Booléens; cependant, elle nécessite une puissance computationnelle importante. De plus, les attracteurs biologiquement cohérents doivent être identifiés d'une manière spécifique aux cellules et aux maladies étudiées. À cet effet, j'ai développé une framework qui analyse les modèles Booléens en mode synchrone à l'aide d'une nouvelle version de l'outil BMA déployé dans un cluster de calcul haute performance. La framework identifie tous les attracteurs, puis teste leur cohérence par rapport aux données omiques et aux connaissances actuelles. Pour identifier de nouvelles options thérapeutiques dans la PR, je me suis intéressée aux macrophages. J'ai effectué une recherche exhaustive de cibles thérapeutiques dans les modèles des macrophages M1 et M2. J'ai ensuite simulé des knock-outs (KOs) simples et doubles qui m'ont permis d'identifier NF-κB, JAK1/JAK2 et ERK1/Notch1 comme cibles potentielles pour supprimer les macrophages pro-inflammatoires. J'ai également identifié GSK3β comme cible potentielle pour activer les macrophages anti-inflammatoires. La prochaine étape consiste à étudier de nouvelles options thérapeutiques dans un contexte multicellulaire en simulant des KO simples et doubles sur le modèle multicellulaire. L'objectif est d'étudier comment la suppression et/ou l'activation de certaines populations cellulaires pourrait réguler l'inflammation et la destruction du cartilage dans la synoviale de la PR. Notre framework ouvre la voie pour l'analyse de nombreuses autres cartes d'une manière intégrative et combinatoire.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Structure et dynamique des systèmes vivants (SDSV)

Title : Computational modeling of the cellular interplay in rheumatoid arthritis. Deciphering the role of resident and immune cells in the arthritic joint.

Keywords : Rheumatoid arthritis, Large-scale computational modeling, Macrophage model, Discrete Multicellular model, Immuno-inflammation, Cellular crosstalks.

Abstract : Rheumatoid arthritis (RA) is an autoimmune disease whose etiology is still not fully understood. The disease manifestation includes joint pain, stiffness and swelling, and chronic synovitis. The pathological state of the synovial tissue in RA is the result of cellular crosstalks between the resident synoviocytes and the immune cells. These cell-cell interactions define many aspects of RA pathogenesis and could offer potential targets for therapeutic interventions. Today there exists no curative treatment for RA. The proposed therapies can only alleviate symptoms and are often associated with several adverse effects. In my thesis, I was interested in studying the cellular interplay involved in the disease, comprehending the mechanism of action of current RA treatments, and testing new drug combinations. Also, I wanted to identify new potential therapeutic intervention points, and perform drug repurposing analysis for the identified targets. In this direction I developed an efficient computational system biology approach to gain a multicellular system-level understanding of the intra- and intercellular cascades involved in RA pathogenesis. I started with building the RA-Atlas, a state-of-the-art atlas that recapitulates existing knowledge about the intracellular cascades involved in RA pathogenesis and evidence of their relationship with pathophysiological causes and effects. The RA-Atlas is based on manual curation of the literature and pathway databases and enriched with omics data. It includes four molecular interaction maps specific to the macrophage, including the M1 and M2 phenotypes, the fibroblast and the CD4+ T helper 1. To provide a multicellular-level representation of the arthritic joint, I created the first RA multicellular map through the integration of bidirectional intercellular interactions that were identified using literature mining and omics data. However, this multicellular map remains limited by its static nature when it comes to predictions. To study the emergent behavior of this biological system, I used the CaSQ tool to convert the maps to Boolean models. These logic-based models do not require kinetic parameters and can integrate information from literature and high-throughput technologies, providing a suitable option for analyzing large-scale networks. Attractor search is a key point in Boolean models' analysis; however, it is computationally challenging in large-scale systems, and the identified attractors need to be validated as biologically consistent states in a cell- and disease-specific manner. In this regard, I developed an efficient computational framework that analyzes Boolean models in a synchronous scheme using a new version of the BMA tool deployed to a high-performance computing cluster. The framework identifies all the attractors of the models, filters out the cycles, then tests the coherence of the stable states against omics datasets and prior knowledge. To investigate new therapeutic options in RA, I focused on the synovial macrophages. I conducted an exhaustive search for therapeutic targets present in the RA M1 and M2 macrophage models. Using single and double knockouts (KOs) simulations, I identified NF-κB, JAK1/JAK2, and ERK1/Notch1 as potential targets to selectively suppress pro-inflammatory RA macrophages. I also identified GSK3β as a promising target to promote anti-inflammatory RA macrophages. The next step would be to investigate new therapeutic options in a multicellular setting by simulating single and double KOs on the RA multicellular model. The goal is to investigate how the suppression and/or promotion of certain cell populations could regulate chronic inflammation and cartilage destruction in the RA synovium. My framework, applied on the RA multicellular map, sets the path to many other disease maps to be analyzed in an integrative and combinatorial manner.

Ackowledgements

First and foremost I would like to thank Pr. Jasmin Fisher and Liesbet Geris for agreeing to serve as reviewers of my PhD manuscript and for their constructive feedback. I would also like to thank Dr. Benjamin Hall and Gaetane Nocturne for for agreeing to be examiners on my jury.

I'd also like to thank the members of my thesis committee, Dr. Benjamin Hall, Dr. Nadine Biesemann, and Dr. Marek Ostaszewski, for all of their suggestions and support throughout the years, which contributed significantly to the advancement of my project. A special thanks to Ben for answering so patiently the countless questions I asked him.

I would like to express my profound gratitude to my incredible supervisors Pr. Anna Niarakis and Dr. Franck Augé for their invaluable advice, unwavering support and constant encouragement and patience during this PhD. I have been extremely lucky to have supervisors who were truly committed to guiding me while giving me all the freedom to explore things on my own. I'll always be grateful for all the great opportunities Anna gave me from my Master internship to the final submission of this thesis. She made immeasurable contributions to my development and growth (even if she still calls me “ma petite”!). I’m also very thankful for Franck’s supervision. He made me feel part of the team at Sanofi and always made room for me in his busy schedule. I will miss the delicious macarons he used to bake for the team!

I would like to thank Pr. Elisabeth Petit Teixeira, my laboratory director, for welcoming me into her lab. She made sure I had everything I needed to carry out my work. I thank her for her kindness, for her help whenever I needed it and for the stimulating discussions (especially the ones about movies) we had over lunch or coffee.

I would also like to thank Dr. Valerie Chaudru for being so kind and caring and for her warm smile in all circumstances.

I am indebted to my exceptional lab mates, whose support has been a constant source of motivation. I'll treasure my memories of our cheat meals and walks around the G2 building. I would like to thank Dr. Vidisha Singh with whom I've shared the same room since my first day at the Genhotel. She has been an unfailing support, both professionally and personally. I’ll always be grateful for her warmth and thoughtfulness (and her great head massages!). I would like to thank Dr. Sahar Aghakhani, my road partner during the last three years, for all the brainstorming sessions we had and for giving me straightforward and honest opinions whenever I needed it. It was really great to walk through the PhD journey

with her. I would like to thank Sacha E Silva-Saffar for his encouragement and for the kind messages he sent me each time I needed support. His eagerness to learn new things was refreshing and I wish him all the best for the rest of his PhD project. I would also like to thank Pilar for her good mood, her cheerfulness and her kindness. I would like to thank my former colleague Dr. Quentin Miagoux for being available whenever I needed help. I enjoyed his funny jokes and the stimulating discussions topics he used to bring out during lunch. I would like to thank Dr. Maeva Veyssiere, a former lab member, for her kindness, for taking the time to help me and for sharing with me her invaluable Python scripts.

I would also like to thank the current and former ODS team members at Sanofi with whom I shared lunch and coffee breaks and had weekly meetings. I would like to thank Dr. Deimante Simaite, Dr. Tommaso Andreani, Dr. Francesca Frau, Dr. Faryad Sahneh, Dr. Marianyela Petrizzelli, Dr. Daniel Margerie, Dr. Franck Rapaport, Clara Renou, Samy Ayed and Dr. Sheng Zhao for their insightful feedback on my work. A special thanks to Dr. Galina Boldina and Dr. Milad Rafiee Vahid who contributed to the advancement of my PhD project. I would like to thank Aurélien Beaude with whom I shared the open space at Chilly Mazarin. I'd also like to thank everyone who has been a part of my Mondays at Sanofi: Dr. Charles Bettembourg, Dr. Daria Madejak, Dr. Olivier Gigonzac, Dr. Véronique Onado, Dr. Hervé Minoux and Dr. Camille de Cevins.

I'd like to thank Dr. Sylvain Soliman for his support whenever I had questions related to computational modeling, and Dr. Zsolt Böcskei for the helpful discussions.

I would also like to express my gratitude to my mom Thouria. She has always believed in me and encouraged me to pursue my goals. In spite of the distance, she always senses when something is wrong and provides me with unfailing love and continuous encouragement. For all this, I'll be eternally grateful.

I'd also like to thank my sister and best friend Karima for being so caring and understanding, for always being on my side, and for always finding the right words to make me laugh.

A special thanks to my dad Chaabane, who inspired me more than he'll ever know. He would have been extremely proud of me.

I would like to thank my close friends Salima and Samira for motivating me and supporting me throughout the years and for making my PhD thesis more pleasant.

Finally, I would like to thank my dear Rahim for being the awesome partner that

he is, for giving me unconditional love and support and for taking care of me. I'm eager to start this new chapter of my life by his side.

Synthèse en français

La polyarthrite rhumatoïde (PR) est une maladie auto-immune inflammatoire chronique dont les mécanismes sous-jacents à son initiation ne sont pas encore totalement élucidés. La pathogenèse de la PR implique une combinaison de facteurs génétiques, épigénétiques et environnementaux et résulte d'interactions complexes entre la signalisation biologique, la régulation génique et le métabolisme. L'activation dérégulée des voies telles que JAK/STAT, SAPK/MAPK et de PI-3K/AKT/mTOR, entre autres, conduisent à une expression accrue des métalloprotéinases et des cytokines pro-inflammatoires, ce qui entraîne une dégradation chronique et une inflammation du tissu synovial. La cascade inflammatoire entraîne une hyperplasie, une dégradation du cartilage et une destruction osseuse dans les articulations. Les articulations touchées perdent leur forme et leur alignement, ce qui entraîne des déformations et une perte de fonction. La douleur, le gonflement, et la raideur des articulations des mains et des pieds sont les premiers symptômes de la PR. Des manifestations extra-articulaires peuvent également apparaître, telles que des maladies gastro-intestinales, respiratoires et cardiovasculaires.

Le comportement pathogène du tissu synovial chez les patients atteint de la PR n’est pas dû à un seul type de cellule. Il résulte de la communication cellulaire entre les synoviocytes locaux et les cellules du système immunitaire inné et adaptatif. Ces interactions intercellulaires déterminent de nombreux aspects de la physiopathologie de la maladie. Elles peuvent activer ou inhiber des populations spécifiques de cellules synoviales. Elles régissent également les cascades de signalisation intracellulaire et l'expression des substances inflammatoires ainsi que les enzymes de remodelage de la matrice dans les articulations. Ce qui régule l'inflammation chronique, la dégradation de la matrice et l'érosion osseuse.

Il n'existe actuellement aucune thérapie curative pour la PR. Le large éventail de molécules associé à la PR, les réseaux cellulaires complexes et l'hétérogénéité de la maladie sont autant d'obstacles à la guérison de la PR. De nombreuses thérapies ont été développées pour soulager les symptômes de la maladie, améliorer la survie et éviter l'invalidité à long terme. Cependant, elles sont également associées à plusieurs effets indésirables et une grande partie de patients atteints de PR ne répondent pas à ces traitements. Nous devons donc mieux comprendre les communications cellulaires et les cascades intracellulaires impliquées dans la pathogenèse de la maladie afin d'élucider le mode d'action des médicaments proposés actuellement contre la PR et d'identifier de nouvelles options thérapeutiques.

La biologie des systèmes peut aider à obtenir une compréhension holistique des maladies complexes telles que la PR grâce à la construction de cartes d'interactions moléculaires. Ces cartes sont des réseaux d'interaction détaillés décrivant les processus moléculaires ainsi que leurs associations avec les phénotypes des maladies décrites. Les cartes d'interactions moléculaires sont construites manuellement et annotées via des recherches bibliographes extensives. Elles sont également standardisées pour permettre une représentation, une visualisation, un stockage, un échange et une réutilisation efficace et précise des connaissances. Les cartes peuvent être utilisées pour l'identification de biomarqueurs, de cibles thérapeutiques et de sous-types de maladies. En outre, elles fournissent un support complet pour la visualisation, l'analyse et l'interprétation des données omiques. Cependant, ces cartes sont statiques et ne peuvent pas être utilisées pour effectuer des simulations et générer des prédictions.

La modélisation dynamique est un outil puissant pour comprendre les caractéristiques et les comportements émergents des systèmes complexes décrits dans les cartes d'interactions moléculaires. Parmi les techniques de modélisation disponibles, le formalisme Booléen est le mieux adapté aux larges systèmes biologiques. En effet, la modélisation Booléenne ne requière pas de paramètres cinétiques, difficiles à déterminer dans la plupart des systèmes. Elle permet également d'utiliser la littérature, les données expérimentales et les données de séquençage afin d'extraire une grande quantité de données qualitatives décrivant les composants et les interactions qui composent ces larges systèmes biologiques. Dans ces modèles Booléens, les nœuds représentent des entités biologiques (gènes, protéines, métabolites...). Chaque nœud peut avoir deux valeurs : zéro (0) pour l'inactivité et un (1) pour l'activité. Les différents composants du système modélisé sont reliés par des arcs qui représentent leurs interactions régulatrices. Ces interactions sont décrites par des fonctions Booléennes : "AND", "OR" et "NOT", et la valeur de chaque composant est déterminée par ces fonctions Booléennes.

Lorsqu'il est simulé, le modèle change d'état global (c'est-à-dire l'état du système entier) et finit par atteindre une configuration particulière appelée attracteur. Les attracteurs sont associés à des phénotypes biologiques, ce qui fait de leur calcul un point clé dans l'analyse des modèles Booléens. Étant donné que le nombre d'états globaux d'un modèle Booléen dépend exponentiellement du nombre de ses nœuds (2^n états globaux pour n nœuds), la recherche d'attracteurs dans les systèmes biologiques à grande échelle nécessite une grande puissance computationnelle. De plus, comme les attracteurs diffèrent en fonction des perturbations externes que le modèle reçoit de son environnement, le nombre total d'attracteurs peut être très élevé. Il est donc nécessaire d'identifier

correctement les états biologiquement cohérents et de les interpréter d'une manière spécifique aux types cellulaires et à la maladie étudiés.

Dans le cadre de ce projet de thèse, nous nous sommes intéressés au développement d'une approche computationnelle et systémique afin d'obtenir une meilleure compréhension des mécanismes impliqués dans la PR. Notre objectif était de fournir une représentation détaillée de la structure et de la dynamique des interactions intercellulaires complexes impliquées dans la PR, ainsi que des cascades intracellulaires et de leur régulation dans cette maladie. Nous voulions également mieux comprendre le mécanisme d'action des thérapies proposées actuellement pour la PR et identifier de potentielles nouvelles options thérapeutiques en mono- et bi-thérapies.

La première étape de ce travail consistait à construire un atlas de la PR en intégrant à la fois la littérature via des recherches bibliographiques et les données omiques. Nous avons cherché à rassembler les données existantes mais fragmentées (expression génique, voies de signalisation, phénotypes cellulaires) sur la pathogenèse de la PR d'une manière spécifique aux types cellulaires étudiés. Dans cette optique, nous avons construit quatre cartes d'interactions moléculaires spécifiques aux populations cellulaires les plus abondantes dans la synoviale de la PR : les macrophages synoviaux de phénotype M1 et M2, les fibroblastes synoviaux et le sous-type CD4+ T helper 1 (Th1).

Ces cartes ont été construites en utilisant les informations provenant des articles publiés et des bases de données telles que KEGG. Ces informations ont fait l'objet d'une curation et d'une annotation manuelles selon la norme MIRIAM. La première étape a consisté en une étude approfondie des données publiques afin d'identifier les voies, les composants et les interactions impliquées dans la pathogenèse de la PR. Nous avons utilisé des études expérimentales à petite échelle chez l'homme pour confirmer l'implication de composants ou d'interactions dans la maladie. Nous avons suivi des critères de curation strictes lors de l'élaboration des cartes. Seuls les composants et les interactions ayant une expression démontrée dans la pathogenèse de la PR et dans le type de cellule concerné ont été intégrés.

La représentation graphique de ces informations s'est faite à l'aide du logiciel CellDesigner, un outil d'édition de diagrammes de réseaux biochimiques. Les nœuds des cartes représentent des biomolécules telles que des protéines, des gènes et des complexes et sont décrits par des identifiants HGNC. Les arcs représentent les interactions entre les nœuds, telles que l'activation, l'inhibition et la catalyse. La syntaxe et la sémantique des symboles utilisés pour représenter les nœuds et les interactions sont conformes au format SBGN PD. Les cartes sont stockées dans le format SBML (Systems Biology Markup Language), un format

standard ouvert basé sur le langage de balisage extensible (XML) pour représenter les réseaux de réactions biochimiques.

Toutes les cartes ont été construites en utilisant la même structure. En effet, les compartiments cellulaires ont été conçus pour refléter la transduction du signal dans une cellule depuis l'espace extracellulaire sous la forme d'un ligand, jusqu'à la membrane plasmique, où il se lie à son récepteur, en passant par le cytoplasme et le noyau, jusqu'à la production de protéines. Ces protéines peuvent être soit sécrétées dans l'espace extracellulaire, soit transportées jusqu'à la membrane. Les phénotypes sont des nœuds particuliers dans les cartes. Ils représentent l'état de la cellule, résultat de multiples processus cellulaires qui définissent et façonnent la morphologie et la fonction de la cellule. Ils décrivent des états biologiques connus pour être actifs ou inactifs dans la PR.

Les cartes d'interactions moléculaires ont été enrichies de voies de signalisations par l'intégration de données d'expression génique spécifiques à la maladie et aux types cellulaires étudiés à l'aide de logiciels d'analyse tels que IPA et MetaCore.

Toutes les cartes construites sont disponibles sous forme de cartes interactives en ligne sur le serveur web autonome MINERVA à l'adresse https://ramap.uni.lu/minerva/. La plateforme MINERVA permet l'exploration visuelle, l'analyse et la gestion de réseaux moléculaires encodés dans des formats de biologie systémique, y compris CellDesigner, SBML et SBGN. Toutes les références des articles scientifiques utilisés pour la construction des cartes sont disponibles dans la section d'annotation des cartes. Des listes de gènes sont aussi disponibles pour visualiser les signatures de chacun des types cellulaires sur les différentes cartes de l'Atlas de la PR.

La deuxième étape de ce projet était d'obtenir une représentation multicellulaire de l'articulation atteinte de la PR. En utilisant à la fois les connaissances préalables et l'intégration des données omiques, nous avons identifié les interactions intercellulaires bidirectionnelles entre les macrophages, les fibroblastes et les types de cellules Th1 dans la synoviale de la polyarthrite rhumatoïde. Ces interactions ont été utilisées pour relier les cartes spécifiques aux cellules correspondantes dans l'Atlas de la PR et construire la première carte multicellulaire de la PR.

La recherche d'interactions intercellulaires dans la littérature et les bases de données s'est faite via différents outils. Le premier est Causaly, un outil de recherche biomédicale qui répond aux questions de recherche auxquelles il peut être difficile de répondre par une recherche traditionnelle par mots-clés. Le système améliore le processus humain de recherche de preuves, qui prend généralement beaucoup de temps en raison de la grande quantité de littérature

biomédicale disponible. Causaly repose sur une intelligence artificielle avancée qui lit la littérature à travers des millions de publications académiques et distille les preuves dans un graphe de connaissances. Contrairement à d'autres outils qui se contentent de compter statistiquement la fréquence d'apparition de mots différents dans une même phrase, Causaly comprend la directionnalité des preuves grâce à son graphe de connaissances qui comprend 220 millions de relations directionnelles. Ces relations peuvent être extraites de sources de données non structurées et structurées ainsi que de données publiques, sous licence et propriétaires (PubMed, brevets, protocoles d'essais cliniques, bases de données chimiques, de cibles et de sécurité...).

Nous avons également utilisé PubMed pour rechercher les communications entre les types de cellules qui nous intéressaient. Des mots-clés et des phrases clés pertinents tels que "M1 macrophage fibroblast interactions in rheumatoid arthritis", "macrophage Th1 activation in RA", "M2 macrophage Th1 inhibition RA", parmi beaucoup d'autres, ont été utilisés pour filtrer les articles dans PubMed.

Pour identifier d'autres moyens de communication cellulaire, nous avons utilisé la base de données CellPhoneDB. Il s'agit d'un référentiel public de récepteurs, de ligands et de leurs interactions répertoriées manuellement. CellPhoneDB contient un total de 978 protéines impliquées dans 1 396 interactions. À partir des cartes d'interactions moléculaires correspondant aux quatre types cellulaires d'intérêt, nous avons extrait les protéines transmembranaires et sécrétées, ainsi que les récepteurs. Dans des fichiers texte, nous avons combiné ces récepteurs et ces ligands de manière à couvrir les paires d'interactions bidirectionnelles suivantes : M1 macrophage-fibroblaste, M1 macrophage-Th1, M2 macrophage-fibroblaste, M2 macrophage-Th1 et fibroblaste-Th1. En ce qui concerne la paire M1 macrophage-fibroblaste, par exemple, nous avons généré un premier fichier texte contenant tous les ligands exprimés dans la carte du macrophage M1 et tous les récepteurs exprimés dans la carte du fibroblaste. Un second fichier a également été généré et comprend tous les ligands représentés dans la carte des fibroblastes et tous les récepteurs exprimés dans la carte des macrophages M1. À l'aide de la base de données CellPhoneDB, nous avons récupéré toutes les interactions qui impliquent à la fois les ligands et les récepteurs présents dans chaque fichier texte généré.

Nous avons également intégré des données omiques via ICELLNET et DiSiR afin d'identifier les interactions statistiquement significatives observées dans ces données. Ces outils se basent sur une ressource de connaissances préalables sur les interactions intercellulaires et sur une méthode pour prédire les événements potentiels de communication entre les cellules. Les événements sont représentés

comme une interaction entre une protéine émettrice et une protéine réceptrice exprimée par le type de cellule source et le type de cellule cible. Les événements sont prédits en estimant la probabilité de l'interaction sur la base du niveau d'expression de l'émetteur et du récepteur dans le type de cellule source et cible, respectivement.

Afin de procéder à une validation croisée des résultats et de ne conserver que les interactions les plus fiables, nous les avons filtré en fonction des critères suivants :

- Seules les interactions statistiquement significatives (seuil FDR égal à 0,05) identifiées à la fois dans ICELLNET et dans DiSiR ont été conservées.
- Seules les interactions identifiées avec au moins deux paires différentes de jeux de données ou deux approches différentes (sur la base de la littérature ou des données omiques) ont été conservées.
- En ce qui concerne les interactions provenant de la littérature et des bases de données, mais non inférées des données omiques, seules les interactions provenant d'au moins deux sources d'information différentes (articles publiés différents ou bases de données différentes) ont été conservées.

Nous avons utilisé ces interactions filtrées pour connecter les cartes d'interactions moléculaires spécifiques aux macrophages, fibroblastes et Th1 entre elles et générer la première carte multicellulaire de la PR.

La troisième étape de ce travail consistait à utiliser les cartes d'interactions moléculaires de la PR comme base pour générer automatiquement des modèles Booléens exécutables. Le but était de développer une approche computationnelle efficace qui utilise des clusters de calcul à haute performance pour effectuer d'une manière efficace la recherche d'attracteurs sur des modèles à grande échelle, et de valider ces attracteurs en tant qu'états biologiquement cohérents en utilisant l'analyse de données omiques et des connaissances préalables.

La construction et l'analyse de modèles Booléens pour les systèmes biologiques complexes à grande échelle restent un défi en biologie des systèmes. Les règles logiques, lorsqu'elles sont définies manuellement, aboutissent à des modèles plus petits qui ne peuvent pas décrire les systèmes biologiques étudiés dans leur entièreté. Tandis que les modèles inférés automatiquement à partir de cartes sont beaucoup plus grands et permettent d'avoir des représentations plus précises des systèmes. Ces modèles à grande échelle, d'autre part, sont plus complexes, avec un nombre considérablement plus élevé de variables, ce qui rend l'identification de leurs attracteurs difficile sur le plan computationnel. De plus, un nombre d'attracteurs considérablement grand entrave l'identification et l'interprétation d'états biologiquement cohérents, en particulier dans un contexte spécifique à un

type cellulaire ou à une maladie donnée.

Nous avons développé une approche computationnelle efficace pour construire, analyser et valider le comportement de modèles Booléens à grande échelle. La première étape de notre approche consiste à utiliser les cartes d'interactions moléculaires comme point de départ pour inférer automatiquement leurs modèles Booléens exécutables via l'outil CaSQ. Pour chaque nœud du fichier XML de départ, CaSQ déduit automatiquement des règles logiques basées sur la topologie et la sémantique du réseau. Tout d'abord, la carte est réduite par le biais d'une série de règles de réécriture de graphes. Certains composants de la carte de départ sont fusionnés en un seul composant ou certaines formes inactives afin d'éviter toute redondance dans le modèle logique. Ensuite, des règles logiques sont automatiquement inférées sur la base de la topologie de la carte réduite. L'outil produit des fichiers exécutables en format SBML-qual ou JSON. En utilisant ce dernier format, il est également possible de produire des réseaux qualitatifs où les nœuds peuvent varier sur une large gamme de valeurs discrètes, appelée granularité. La granularité définit la valeur la plus élevée que les nœuds peuvent prendre dans le modèle. Puisque nous utilisons le formalisme Booléen, la granularité utilisée dans l'analyse de nos modèles est égale à un.

Les modèles Booléens générés sont analysés dans un schéma synchrone à l'aide d'une nouvelle version de l'outil BMA déployé dans un cluster de calcul à haute performance. En effet, BMA a été initialement développé sur le Microsoft .NET et le standard .NET, qui lient les outils aux environnements Windows. Pour permettre l'analyse de modèles à haut débit et tirer parti de la parallélisation sur des clusters à haute performance (généralement basées sur Linux), nous avons développé un prototype de l'outil basé sur le noyau open-source .NET 3.1. Notre approche utilise cette nouvelle version de BMA pour identifier tous les attracteurs possibles des modèles.

Les attracteurs dépendent des stimuli externes que le modèle reçoit de son environnement. Dans les modèles Booléens, les stimuli sont modélisés sous la forme de nœuds externes. Ce sont des nœuds sans régulation en amont. Ils ne sont associés à aucune règle logique dans le modèle ; les valeurs qu'ils prennent sont donc définies par l'utilisateur. Dans l'outil BMA, l'utilisateur peut attribuer des valeurs à ces nœuds externes à l'aide de l'indicateur -ko. En fonction des valeurs que prennent ces nœuds, le modèle atteint différents attracteurs. Pour identifier tous les attracteurs du modèle, nous avons généré toutes les combinaisons possibles de valeurs et les avons attribué aux nœuds externes. Pour chaque combinaison, nous avons calculé l'attracteur correspondant.

Le temps de calcul pour ce type d’analyse est exponentiel. En effet, le nombre de toutes les combinaisons possibles est égal à 2^n, n étant le nombre de nœuds

externes qui varient dans le modèle. Compte tenu du nombre élevé de nœuds externes dans les modèles Booléens, nous avons déployé BMA dans un cluster de calcul à haute performance pour accélérer le processus.

On filtre ensuite les attracteurs qu'on a identifié pour ne conserver que les états stables puis on teste leur cohérence par rapport à des données d'expression génique et aux connaissances préalables. Pour ce faire, nous filtrons davantage les états stables du modèle et sélectionnons ceux qui sont capables de reproduire ce qui est connu dans la littérature ou ce qui est observé dans les données transcriptomiques. Tout d'abord, nous identifions les biomolécules différentiellement exprimées présentes dans le modèle en utilisant à la fois l'exploration de la littérature et l'analyse de l'expression différentielle. Ensuite, nous discrétisons les expressions des biomolécules identifiées pour les convertir en un vecteur binaire de valeurs Booléennes. Nous comparons ensuite les états stables filtrés à ce vecteur binaire à l'aide d'un score de similarité. Sur la base de ce score, nous sélectionnons les états ayant le score le plus élevé et calculons leur vecteur moyen. Le vecteur résultant représente l'état calibré du modèle. Nous avons d'abord appliqué notre méthodologie aux quatre cartes de l'Atlas et qui sont spécifiques au macrophage M1, au macrophage M2, au fibroblaste RA et au CD4+ Th1 RA. Les états calibrés de ces modèles ont ensuite été combinés pour générer et valider le premier modèle multicellulaire de la PR.

La dernière étape de ce travail consistait à utiliser les modèles validés et calibrés pour étudier de nouvelles options thérapeutiques dans la PR par le biais de simulations *in-silico* de simple et double knockouts. Le but était de comprendre le mécanisme d'action des traitements actuels de la PR, de tester de nouvelles combinaisons de médicaments, d'identifier de nouveaux points d'intervention thérapeutique et d'effectuer une analyse de repositionnement des médicaments pour les cibles identifiées.

Pour cela, nous nous somme tout d'abord intéressés aux macrophages synoviaux. En effet, de récentes étudies ont révélé que le système immunitaire inné est un acteur majeur dans l'initiation et le développement de la pathogenèse de la PR. Les macrophages sont l'une des populations de cellules immunitaires innées les plus abondantes dans la PR et leur nombre est en corrélation significative avec la sévérité de la maladie. Les populations de macrophages sont hétérogènes et peuvent se différencier en divers phénotypes en réponse aux stimuli du microenvironnement local. Les phénotypes M1 et M2 représentent les deux extrêmes de leur spectre d'activation.

Par conséquent, en fonction de leur phénotype (M1 ou M2), les macrophages jouent un rôle à la fois dans l'initiation et la résolution de l'inflammation. Les macrophages M1 sont responsables de la surproduction de cytokines

inflammatoires et de la libération d'enzymes de dégradation de la matrice, conduisant à la destruction du cartilage. Ils peuvent également attirer les cellules T pro-inflammatoires et induire leur hyperactivation. D'autre part, les macrophages M2 atténuent l'inflammation par 1) la production de cytokines anti-inflammatoires, y compris l'IL-10 et le TGF-β, 2) l'homéostasie et la réparation des tissus 3) l'activation des cellules T régulatrices. En raison de leur activation et de leur prolifération excessive ainsi que de leur capacité renforcée à résister à l'apoptose, la proportion de macrophages M1 est supérieure à celle des macrophages M2.

Il existe actuellement deux approches pour cibler les macrophages : la régulation à la baisse du phénotype M1 et l'expansion du phénotype M2, ou la repolarisation des macrophages M1 en macrophages M2. Malgré cela, aucun médicament ciblant spécifiquement les macrophages n'est actuellement utilisé en clinique. Par conséquent, mieux comprendre l'approche qui permettrait la déplétion ciblée du macrophage inflammatoire tout en épargnant les autres sous-types de macrophages pourrait constituer une approche thérapeutique efficace dans la PR.

Bien que l'hétérogénéité des macrophages dans la PR n'est pas encore entièrement comprise, les modèles de macrophages M1 et M2 de la PR que nous avons construits et validés visent à couvrir la diversité phénotypique des macrophages grâce à une représentation spécifique de leurs cytokines/chimiokines sécrétées, de leurs stimuli, de leurs récepteurs et de leurs facteurs de transcription. Nous avons donc utilisé ces modèles pour investiguer les mono- et bithérapies qui permettraient de réduire spécifiquement les macrophages pro-inflammatoires et de favoriser les macrophages anti-inflammatoires dans la synoviale de la polyarthrite rhumatoïde. Ceci passe par l'induction du phénotype d'apoptose et l'inhibition du phénotype de prolifération dans le modèle de macrophage M1, et par l'activation du phénotype de prolifération et l'inhibition du phénotype d'apoptose dans le modèle de macrophage M2. Dans cette optique, nous avons testé les réponses de ces modèles aux traitements actuels de la PR afin d'évaluer l'effet de ces traitements sur les macrophages et d'acquérir une meilleure compréhension de leur mécanisme d'action dans ce type de cellules. Nous avons également testé de nouvelles combinaisons de médicaments et des points d'intervention thérapeutique potentiels.

Afin d'identifier les cibles thérapeutiques potentielles présentes dans les modèles, nous avons effectué une recherche exhaustive à l'aide de la Therapeutic Target Database (TTD). Il s'agit d'une base de données de médicaments conçue pour fournir des informations sur les cibles décrites dans la littérature, les conditions pathologiques ciblées, les informations sur les voies de signalisation perturbées

et les médicaments associés à chacune de ces cibles. La base de données contient actuellement 3578 cibles et 38 760 médicaments. Afin de simuler les knockouts dans les modèles, nous avons sélectionné les cibles sur la base du mode d'action des médicaments qui leur sont associés, et nous n'avons gardé que celles qui peuvent être ciblées par au moins un inhibiteur (1643 cibles). Nous avons ensuite identifié les cibles présentes dans les modèles de macrophages M1 et M2 puis simulé leurs KOs.

Nous avons également testé l'effet synergique que pourraient avoir les cibles thérapeutiques précédemment identifiées sur nos modèles. Pour cela les cibles ont été combinées par paires et les modèles de macrophage RA M1 et RA M2 ont été ensuite utilisés pour prédire l'effet de leurs inhibitions.

Nous avons retrouvé une cible thérapeutique et deux combinaisons de cibles thérapeutiques potentielles, à savoir NFkB, JAK1/JAK2 et ERK1/Notch1, qui permettent d'inhibier sélectivement les macrophages pro-inflammatoires de la PR. Nous avons également retrouvé GSK3B comme cible prometteuse afin de promouvoir les macrophages anti-inflammatoires dans la synoviale de la PR. L'interet potentiel/prouvé de ces cibles dans la PR étant documenté dans la literature, ces résultats nous permettent de valider d'avantage nos modèles de macrophages.

L'inhibition du NF-κB a conduit à une suppression sélective du macrophage M1 de la PR. NF-κB représente une cible thérapeutique potentielle très attrayante car il s'agit d'un facteur de transcription clé des macrophages M1, responsable de l'expression accrue des cytokines produites par les macrophages M1 dans la synoviale de la PR. Plusieurs études soutiennent le concept de l'inhibition de NF-κB pour des interventions thérapeutiques dans les maladies inflammatoires. Dans la PR, l'inhibition in vitro de NF-κB induit l'apoptose des fibroblastes et contribue à une diminution significative des marqueurs M1 et à une augmentation des marqueurs M2. D'autres études ont montré que les effets bénéfiques observés des anti-inflammatoires non stéroïdiens (AINS) et des glucocorticoïdes, tous deux utilisés dans le traitement de la PR, sont également dus à l'inhibition du NF-κB. Toutefois, leur utilisation est limitée en raison d'effets secondaires graves. D'autres inhibiteurs de NF-κB ont été identifiés, mais la majorité d'entre eux ne répondent pas aux normes requises pour rejoindre des programmes d'essais clinique. En effet, l'inhibition non sélective de NF-κB dans tous les types de cellules a de multiples effets néfastes, car il est essentielle au maintien du fonctionnement cellulaire de base. Des traitements biologiques ciblant directement les gènes régulés par NF-κB, tels que TNF, IL-6 et IL-1, ont été mis au point. Cependant, l'inhibition d'une seule cytokine n'est peut être pas optimale car plusieurs cytokines ont des activités biologiques synergiques. Par conséquent,

la découverte de techniques d'inhibition de NF-κB qui ciblerait un type de cellule en particulier est nécessaire pour modifier la balance bénéfices/risques.

Nous avons retrouvé GSK3β comme cible d'intérêt pouvant promouvoir les macrophages M2 dans la PR. GSK3β est impliqué dans la progression de plusieurs maladies, y compris la PR. Des études suggèrent que cette protéine joue un rôle central dans différentes voies de signalisation qui sont pertinentes pour la fonction des macrophages, y compris la polarisation et la réponse inflammatoire. L'inhibition de GSK3β dans la PR supprime les réponses inflammatoires dans les synoviocytes de type fibroblaste. De plus, l'inhibition de GSK3β dans la maladie de la rhinite inflammatoire allergique augmente l'expression des marqueurs phénotypique des macrophages M2. CREB1 est l'une des cibles de GSK3β. Lorsque GSK3β est inhibée, elle induit un gain de fonction de CREB1 qui, à son tour, envoie un signal de survie anti-inflammatoire et anti-apoptotique dans les monocytes et les macrophages. CREB1 augmente également l'expression des marqueurs M2 et favorise le phénotype M2 dans les macrophages provenant de souris.

La prolifération des macrophages M1 a été supprimée par le knock-out d'ERK1. Lorsque celui-ci est associé à l'inhibition de Notch1, il induit également l'apoptose des macrophages M1 dans notre modèle. L'intérêt thérapeutique potentiel du ciblage conjoint d'ERK1 et de Notch1 a déjà été démontré dans le cancer, mais pas dans la PR. Il a été démontré que le ciblage de Notch1 renforce l'efficacité des inhibiteurs d'ERK1 chez les patients atteints de cancer. Dans la PR, les inhibitions d'ERK1 et de Notch1 de manière indépendante réduisent l'inflammation dans la PR chez les souris. Notch1, d'autre part, est connue pour influencer la fonction des macrophages M1 via une régulation transcriptionnelle directe et une régulation métabolique indirecte.

Nous avons également retrouvé la paire JAK1/JAK2 comme une combinaison ayant un effet synergique sur notre modèle et permettant d'induire l'apoptose chez les macrophages M1 dans la PR. Le baricitinib, un inhibiteur oral de JAK sélectif pour JAK1 et JAK2, est approuvé par la Food and Drug Administration (FDA) pour le traitement de la PR. Ce résultat valide donc d'avantage le comportement de notre modèle. Baricitinib empêche l'activation des voies STAT, ce qui, à son tour, empêche les réactions auto-immunes et inflammatoires associées à la PR, y compris la sécrétion d'IFNg. Cependant, la façon dont les inhibiteurs de JAK modulent les phénotypes des macrophages, et si ce mécanisme explique leur bénéfice clinique dans la PR, n'est pas encore complètement comprise. Une étude récente a montré que le baricitinib modulait l'expression des marqueurs phénotypiques membranaires ainsi que la sécrétion de certaines cytokines signatures chez les macrophages sains. Une autre étude

confirme l'effet de l'inhibition des JAK sur les phénotypes des macrophages atteints de PR en modifiant le profil métabolique des macrophages M1 et en rééquilibrant la reprogrammation métabolique en faveur de la phosphorylation oxydative.

Bien que les modèles de macrophages donnent une représentation détaillée de la diversité phénotypique de la population de macrophages et nous permettent d'étudier des processus spécifiques à ce type cellulaire comme l'apoptose et la prolifération, ils ne fournissent aucune information sur l'état biologique global de l'articulation dans la PR, tel que l'inflammation, la dégradation de la matrice et l'angiogenèse. En outre, ces modèles peuvent prédire les réponses des macrophages aux médicaments testés, mais pas les réponses cellulaires d'autres types de cellules présentes dans la synoviale. Par conséquent, pour étudier de nouvelles options thérapeutiques dans la PR, nous devons effectuer des simulations dans un environnement multicellulaire qui prend en compte les différents types de cellules présents dans la synoviale de la PR ainsi que leur communication intercellulaire, tous deux impliqués dans la pathogenèse de la maladie.

Dans cette optique, nous avons étudié de nouvelles options thérapeutiques dans un cadre multicellulaire en simulant des KOs simples et doubles sur le modèle Booléen multicellulaire calibré de la PR. Comme ce modèle offre une description multicellulaire de la synoviale de la PR, il nous permet de mieux comprendre comment la suppression et/ou l'activation de certaines populations cellulaires et/ou de certains processus biologiques pourraient réguler l'inflammation chronique, la destruction du cartilage et l'érosion osseuse dans l'articulation atteinte.

De la même manière que pour les simulations sur les modèles de macrophages, nous avons identifié les cibles thérapeutiques potentielles présentes dans le modèle multicellulaire en utilisant la base donnée Therapeutic Target Database (TTD. Afin de simuler les knockouts dans les modèles, nous n'avons gardé que les cibles qui sont associées à au moins un inhibiteur (1643 cibles). Nous avons ensuite identifié les cibles présentes dans le modèle puis simulé leurs KOs. Nous avons utilisé l'état calibré du modèle multicellulaire comme condition initiale pour les simulations. Les états phénotypiques du modèle après inhibition des cibles ont ensuite été comparés à leurs états calibrés.

Nous avons également testé l'effet synergique potentiel que pourraient avoir les cibles thérapeutiques précédemment identifiées sur notre modèle. Pour cela les cibles ont été combinées par paires et le modèle multicellulaire de la PR a ensuite été utilisé pour prédire l’effet de leurs inhibitions.

Nous avons retrouvé AKT2, qui appartient à la famille des sérines/thréonines kinases, comme cible potentielle pour l'inhibition de la prolifération et de la migration des fibroblastes dans la PR. Il a été montré que la voie de signalisation PI3K/AKT est corrélée à l'apparition et au développement de la PR. Plusieurs études soutiennent le concept de l'inhibition de l'AKT2 pour une intervention thérapeutique dans la PR. Elles ont montré que le blocage de la voie AKT inhibe la progression de la PR. Elles ont également démontré que le silençage in vitro d'AKT2 par siRNA empêche de manière significative la prolifération cellulaire et la migration des fibroblastes humains dans la PR.

Nous avons également retrouvé cavéoline 1 (CAV1) comme cible potentielle pour induire l'apoptose et inhiber la prolifération et la migration dans les fibroblastes de la PR. CAV1 est l'un des principaux composants structurels des cavéoles et possède un certain nombre de fonctions de signalisation. En ce qui concerne les troubles inflammatoires, CAV1 pourrait soit prévenir soit induire l'inflammation, en fonction du contexte cellulaire. Dans la PR, le silencing in vitro de CAV1 réduit considérablement la prolifération cellulaire et favorise l'apoptose dans les fibroblastes humains. En revanche, l'expression forcée de CAV1 dans les fibroblastes de la PR rétablit la prolifération cellulaire et atténue l'apoptose.

L'inhibition de CREB1 dans le modèle a conduit à l'inhibition de la prolifération des fibroblastes de la PR tout en maintenant leur capacité à résister à l'apoptose. Cet effet inhibiteur a été démontré dans une étude expérimentale où la suppression in vitro de l'activité de CREB1 a réduit les fonctions aberrantes des cellules synoviales chez les patients atteints de PR via la suppression de la prolifération des fibroblastes synoviaux.

L'inhibition de NF-κB dans le modèle multicellulaire a bloqué la différenciation des macrophages M1 en ostéoclastes. Il a également supprimé la migration des fibroblastes, la prolifération et la migration des Th1 et la dégradation de la matrice extracellulaire dans l'articulation. Des recherches ont démontré que NF-κB médiait l'activation transcriptionnelle de plusieurs métalloprotéinases matricielles (MMP), responsables de la destruction de la matrice extracellulaire et du cartilage articulaire dans la PR. Elles ont également montré que l'inhibition de NF-κB bloquait l'expression de ces MMPs. Le blocage de l'expression des MMP dans la synoviale de la PR entraîne la suppression de la migration et de l'invasion des fibroblastes synoviaux de la PR.

L'importance critique de NF-κB dans le renouvellement osseux a également été mise en évidence expérimentalement. En effet, les souris NF-κB1-/- NF-κB2-/- développent un défaut d'ostéoclastes, montrant les fonctions critiques de NF-κB dans le développement de cette lignée cellulaire. En outre, il a été démontré que l'inhibition de NF-κB constituait une approche efficace pour inhiber la formation

d'ostéoclastes et la résorption osseuse, et que le blocage de l'activation nucléaire de NF-κB présentait des avantages anti-inflammatoires et anti-ostéolytiques.

De plus, plusieurs études ont décrit le rôle de NF-κB dans la différenciation des cellules Th1. Elles ont montré que les réponses Th1 étaient significativement altérées et que la production d'IFN-g était abrogée en raison d'une diminution de l'activation du NF-κB. Une autre étude menée sur des souris transgéniques dont les cellules T étaient dépourvues de la voie de signalisation NF-κB/Rel a montré que NF-κB jouait un rôle dans les réponses d'hypersensibilité Th1-dépendantes.

Nous avons également retrouvé deux cibles thérapeutiques potentielles pour réguler les CD4+ Th1 dans la PR, à savoir mTOR et TBX21. TBX21 (ou T-bet) est un facteur de transcription spécifique aux Th1 qui contrôle l'expression de la cytokine phare des Th1, l'IFNg. Des études ont démontré que l'inhibition de T-bet supprimait de manière significative l'expression des gènes IFNg et IL-17 et diminuait les symptômes de l'arthrite chez les souris. En outre, il a été montré que les cellules CD4+ des souris T-bet-/- s'orientaient vers une différenciation Th2 anti-inflammatoire et exprimaient des niveaux élevés de GATA-3. Par conséquent, la régulation du rapport T-bet/GATA-3, par la déplétion de T-bet, pourrait réduire les fonctions pro-inflammatoires des cellules T dans la PR via la régulation du rapport Th1/Th2.

Des études ont montré que mTOR jouait également un rôle important dans le développement des cellules Th1. En effet, les cellules T CD4+ déficientes en mTOR ne parviennent pas à se différencier en cellules Th1 effectrices. mTOR régule cette différenciation par l'intermédiaire de STAT4 et SOCS3, qui régulent ensuite l'expression de T-bet. Dans le cas de la PR, l'inhibition de mTOR s'est révélée efficace pour réduire l'inflammation dans les articulations de souris mais aussi chez des patients atteints de PR.

Nous avons également retrouvé des cibles thérapeutiques potentielles qui permettraient de réduire l'inflammation et l'angiogenèse dans l'articulation atteinte de PR, à savoir JAK1/JAK2 et MIR221. Le double KO JAK1/JAK2 a inhibé l'inflammation chronique de la synoviale et la différenciation des macrophages en ostéoclastes dans notre modèle. Le baricitinib, un inhibiteur oral de JAK sélectif pour JAK1 et JAK2, est déjà approuvé par la Food and Drug Administration (FDA) pour le traitement de la PR et permet de diminuer l'inflammation systémique et la progression de la destruction osseuse associée à la PR. MIR221, quant à lui, a permis la baisse de l'angiogenèse dans le modèle. Il a été démontré que MIR221 inhibait THBS1, qui agit comme un facteur anti-angiogénique sur les cellules endothéliales. Ainsi, l'inhibition de MIR221 rétablirait l'expression de THBS1 qui, dans la PR, contribue à restaurer l'homéostasie tissulaire au cours de la résolution

de l'inflammation.

Une nouvelle combinaison a également été identifiée via les simulation *in silico*. Le double KO de NF-κB et de STAT3 a inhibé la différenciation des cellules précurseurs d'ostéoclastes dans le modèle multicellulaire. Qu'elle soit normale ou pathologique, l'ostéoclastogenèse dépend strictement de la présence des cytokines nécessaires à la différenciation des ostéoclastes. On a constaté que CSF1, RANKL, TNF-α et IL-6 induisent la différenciation des ostéoclastes et l'activité de résorption osseuse chez les patients atteints de PR. STAT3 et NF-κB sont des facteurs de transcription essentiels pour l'expression de ces facteurs ostéoclastogéniques, par conséquent, cibler à la fois STAT3 et NF-κB pourrait être une stratégie prometteuse pour inhiber l'érosion osseuse dans la PR.

Pour conclure, grâce à ce travail, nous avons contribué à une meilleure compréhension de la PR au niveau multicellulaire grâce à une approche computationnelle et systémique efficace. Nous avons obtenu une représentation détaillée de la structure et de la dynamique de la communication intercellulaire dans la PR ainsi que des cascades intracellulaires et de leur régulation dans cette maladie. Nous avons également validé notre approche via l'identification de cibles thérapeutiques et la mise en évidence de nouvelles combinaisons de médicaments. Même si l'intérêt potentiel/prouvé des cibles qu'on a identifié a déjà été documenté dans la PR, notre approche apporte une meilleure compréhension des voies de signalisation associées à ces cibles ainsi que de leurs mécanismes d'action d'une manière spécifique à la PR et au type cellulaire étudié.

Pour résumer, durant cette thèse de doctorat:

- Nous avons construit les premières cartes d'interactions moléculaires qui décrivent d'une manière spécifique les populations cellulaires les plus abondantes dans la synoviale de la PR : les macrophages (y compris les phénotypes M1 et M2), les fibroblastes et le sous-type CD4+ T helper 1 (Th1). Ces cartes récapitulent les connaissances existantes relatives aux interactions intracellulaires impliquées dans la pathogenèse de la PR d'une manière spécifique à chaque type de cellule. Elles ont été construites à l'aide d'informations extraites de la littérature et de bases de données, et ont été enrichies via l'intégration de données omiques. Les cartes d'interactions moléculaires spécifiques aux cellules de la PR sont accessibles à l'adresse https://ramap.uni.lu/minerva/ et peuvent être utilisées comme base de connaissances interactive, à l'aide de la plateforme MINERVA.

- Nous avons construit la première carte d'interactions moléculaires multicellulaire de la PR, un atlas qui décrit les interactions intra- et intercellulaires impliquées dans la pathogenèse de la maladie. Pour cela, nous avons connecté les cartes citées précédemment par l'ajout d'interactions intercellulaires bidirectionnelles identifiées à l'aide de de la littérature et des bases de données, et de l'intégration des données omiques spécifiques à la maladie et au type cellulaire étudiés.

- Pour comprendre les caractéristiques et les comportements émergents des systèmes biologiques décrits dans ces cartes, nous avons converti ces cartes d'interactions moléculaires en modèles Booléens exécutables à l'aide de CaSQ. Nous avons développé un framework qui analyse ces modèles Booléens en mode synchrone à l'aide d'une nouvelle version de l'outil BMA déployé dans un cluster de calcul haute performance. Le framework identifie tous les attracteurs, puis teste leur cohérence par rapport aux données omiques et aux connaissances actuelles. Nous avons utilisé ce framework pour calibrer nos modèles Booléens, ouvrant ainsi la voie à l'exploration de nombreuses autres cartes.

- Pour étudier de nouvelles options thérapeutiques dans la PR, nous nous sommes tout d'abord concentrés sur les macrophages synoviaux. Le nombre de macrophages dans le tissu synovial enflammé augmente rapidement au cours de la PR, et leur polarisation en un phénotype pro-inflammatoire M1 ou anti-inflammatoire M2 joue un rôle critique dans la progression pathologique de la PR. Ainsi, la suppression sélective des macrophages M1 ou l'augmentation des macrophages M2 pourrait être une stratégie prometteuse pour le traitement de la PR. Dans cette optique, nous avons simulé des KO simples et doubles dans les modèles calibrés de macrophages M1 et M2. Nous avons retrouvé NF-κB, JAK1/JAK2 et ERK1/Notch1 comme cibles pouvant potentiellement supprimer sélectivement les macrophages pro-inflammatoires de la PR. Nous avons également retrouvé GSK3β comme une cible ayant un intérêt potentiel pour la promotion des macrophages anti-inflammatoires.

- Pour étudier de nouvelles options thérapeutiques dans un cadre multicellulaire, nous avons simulé des KO simples et doubles dans le modèle multicellulaire de la PR calibré. Nous avons retrouvé AKT2, CAV1 et CREB1 comme des cibles prometteuses pour limiter le comportement invasif des fibroblastes. Nous avons montré que dans notre modèle, l'inhibition de

mTOR et de TBX21 pouvait réguler les cellules CD4+ Th1 hyperactives, que l'inhibition de MIR221 pouvait réduire l'angiogenèse dans l'articulation et que l'inhibition de la paire NF-κB /STAT3 pouvait supprimer l'érosion osseuse. D'autres effets des inhibitions de NF-κB et de JAK1/JAK2 au niveau multicellulaire ont également été révélés. L'inhibition de NF-κB a permis de supprimer la différenciation des macrophages M1 en ostéoclastes, diminuer l'hyperactivation des fibroblastes et des cellules Th1 et réguler la dégradation de la matrice extracellulaire dans l'articulation. L'inhibition de JAK1/JAK2, de l'autre côté, a permis de réduire l'inflammation et la destruction osseuse.

Nous proposons, pour ce projet de thèse, les perspectives suivantes :

- D'autres populations cellulaires importantes, notamment les lymphocytes B, les cellules dendritiques, les plasmocytes, les mastocytes, les ostéoclastes et les chondrocytes, sont impliquées dans la pathogenèse de la PR. Il serait donc intéressant d'agrandir la carte multicellulaire de la PR en ajoutant d'autres cartes spécifiques a ces types cellulaires. Ces cartes pourraient être reliées aux cartes existantes via l'identification d'interactions intercellulaires supplémentaires. En outre, comme plusieurs études ont mis en évidence la présence de divers points de contrôle métaboliques dans la PR, nous aimerions intégrer les voies métaboliques pertinentes dans nos cartes afin de mieux comprendre les interactions complexes entre les voies métaboliques et les réponses inflammatoires et immunitaires.

- Nous avons intégré divers jeux de données omiques pour construire les cartes d'interaction moléculaire de la PR, identifier la interactions cellule-cellule et valider les modèles Booléens. Cependant, les échantillons de macrophages, de fibroblastes et de Th1 provenant de la synoviale de patients atteints de PR étaient très difficiles à trouver et rarement disponibles dans les mêmes jeux de données. Par conséquent, l'accès à un plus grand nombre d'échantillons synoviaux de patients atteints de la PR pour les types cellulaires qui nous intéressent améliorerait la robustesse de notre approche. La PR étant une maladie très hétérogène, nous devrions également mettre en œuvre une stratégie pour gérer les incohérences entre ces jeux de données omiques.

- L'intégration de données multi-omiques avec des couches transcriptomiques et protéomiques, par exemple, contribuerait à améliorer la méthodologie proposée. En outre, des modèles qualitatifs à valeurs multiples peuvent être envisagés pour prendre en compte les biomolécules qui ne sont pas exprimées de manière différentielle dans ces données omiques, tels que les gènes de ménage. Toutefois, une puissance de calcul plus importante serait nécessaire pour faire face à la complexité exponentielle de ces modèles.

- Lors de la calibration des modèles Booléens, un mode synchrone a été utilisé pour effectuer la recherche d'attracteurs, et seuls les états stables ont été testés et validés. Il serait intéressant d'étudier des attracteurs plus complexes en utilisant les modes synchrone et asynchrone. Même s'il est difficile d'estimer comment les états calibrés du modèle changeraient lorsque le mode est modifié, l'identification des attracteurs en mode synchrone et asynchrone, et la comparaison de ces attracteurs aux connaissances existentes ainsi qu'aux données omiques pourraient aider à comprendre les effets de ces changements sur les prédictions des modèles.

- Afin d'identifier d'avantages de cibles thérapeutiques et de combinaisons médicamenteuses, nous pourrions élargir notre approche à tous les nœuds des modèles, et non pas uniquement à ceux répértoriés dans TTD. Nous pourrions effectuer des simulations *in silico* de KOs simples et doubles de tous les nœuds présents dans nos modèles. Ceci aura pour conséquence de générer une quantité considérable de données à analyser mais pourrait potentiellement nous permettre d'identifier de nouvelles cibles qui n'ont pas été documentées dans la PR auparavant.

- Pour valider davantage le comportement des modèles calibrés et tester les prédictions générées, ce travail bénéficierait grandement de l'intégration de données transcriptomiques de type single cell supplémentaires, en particulier de patients avant et après le traitement. Il bénéficierait également d'expériences in vitro qui cibleraient les voies/facteurs prédits tout en mesurant les processus cellulaires tels que l'apoptose, la prolifération ou l'inflammation.

- De nombreuses autres cartes d'interactions moléculaires ont déjà été publiées pour plusieurs maladies, notamment la maladie de Parkinson, la maladie d'Alzheimer, le cancer, l'asthme et, plus récemment, le COVID19.

Notre approche, appliqué à la carte multicellulaire de la polyarthrite rhumatoïde, ouvre la voie à l'exploration de ces cartes d'une manière intégrative et combinatoire.

- D'autres approches prometteuses utilisant des cartes d'interactions moléculaires sont en cours de développement. Elles sont basées sur des méthodes d'apprentissage automatique qui utilisent des réseaux biologiques et des données de séquençage pour prédire, par exemple, les interactions entre protéines, les fonctions des protéines, les interactions entre médicaments et cibles et le diagnostic des maladies. Les cartes que nous avons créées pourraient être utilisées dans ce type d'approche et pourraient apporter des résultats complémentaires à ceux déjà obtenus.

Table of Contents

List of Figures i

List of Tables vii

List of Abbreviations ix

Introduction 1

Chapter 1. The pathogenesis of rheumatoid arthritis 6

1.1 Etiology 6
1.1.1 Genetic susceptibility 7
1.1.2 Epigenetic factors 9
1.1.3 Environmental factors 10
1.2 Epidemiology 11
1.3 Diagnosis 12
1.4 Pathophysiology 14
1.5 Cellular interplay and crosstalks 16
1.1.4 Synovial fibroblasts 17
1.1.5 Synovial macrophages 19
1.1.6 CD4+ T helper 1 (Th1) 21
1.6 Treatments 22

Chapter 2. Systems Biology to better understand complex diseases 24

2.1 Molecular interaction maps 24
2.1.1 Community developed standards for knowledge representation 25
2.1.2 From a cellular to a multicellular level 26
2.2 Computational modeling in biology 27
2.2.1 Quantitave models in biology 28
2.2.2 Agent-based models in biology 29
2.2.3 Qualitative models in biology 29
2.2.4 Boolean models 31
2.2.5 Long-term behavior of Boolean models 33

Chapter 3. Description of data and data analysis 36

3.1 GSE97779 dataset 36
3.2 GSE164498 dataset 36
3.3 GSE109449 dataset 37
3.4 SDY998 dataset 38
3.5 GSE32901 dataset 42
3.6 GSE107011 dataset 42
3.7 GSE135390 dataset 42
3.8 E_MTAB_8322 dataset 43

Chapter 4. Construction of RA cell-specific molecular interaction maps 45

4.1 Introduction 45
4.2 Materials and methods 45
4.2.1 Molecular interaction maps construction 46
4.2.2 Annotations and curation criteria 47

4.2.3 Integration of pre-existing maps....48
4.2.3.1 RA macrophage map....48
4.2.3.2 RA fibroblast map....49
4.2.3.3 RA CD4+ Th1....50
4.2.4 Cell-specific maps enrichment using gene expression datasets....51
4.2.4.1 RA macrophage map....52
4.2.4.2 RA fibroblast map....53
4.2.4.3 RA CD4+ Th1 map....53
4.2.5 Cell-specific overlays....53
4.3 Results....54
4.3.1 RA fibroblast map....55
4.3.2 RA macrophage maps....57
4.3.3 RA M1 macrophage map....58
4.3.4 RA M2 macrophage map....58
4.3.5 RA Th1 Map....59
4.3.6 Visualization and accessibility....62
4.3.7 Applications....62
4.4 Discussion....65

Chapter 5. Construction of the RA multicellular map....66

5.1 Introduction....66
5.2 Materials and methods....66
5.2.1 Identification of cell-cell interactions using literature and databases mining....68
5.2.2 Identification of cell-cell interactions using omics data....70
5.2.2.1 ICELLNET framework....71
5.2.2.2 Dimer Signal Receptor Analysis (disir) framework....73
5.2.2.3 Selected omics datasets....75
5.2.3 Cell-cell interactions filtering criteria....78
5.2.4 Construction of the RA multicellular map in celldesigner....79
5.3 Results....80
5.3.1 Bidirectional interactions between RA macrophage and RA fibroblast....80
5.3.2 Bidirectional interactions between RA macrophage and RA CD4+ Th1....84
5.3.3 Bidirectional interactions between RA fibroblast and RA CD4+ Th1....86
5.3.4 RA multicellular map....88
5.4 Discussion....89

Chapter 6. Generation and calibration of large-scale Boolean models....92

6.1 Introduction....92
6.2 Materials and methods....93
6.2.1 Generation of Boolean models from molecular interaction maps....94
6.2.2 Stabilization proof using Bio Model Analyzer (BMA)....95
6.2.3 BMA architecture and underlying technologies....96
6.2.4 Parallel computing for the calculation of all possible attractors....96
6.2.5 Filtering the model's attractors....98
6.2.6 Validation of the model's behavior....98
6.2.6.1 Identification of differentially expressed biomolecules using literature search and gene expression data analysis....99
6.2.6.2 Computation of similarity scores and data discretization....99
6.2.6.3 Selection of the steady states with the highest similarity score....101
6.3 Results....101
6.3.1 Updating the RA cell-specific molecular interaction maps....101

6.3.2 Generation of the RA cell-specific Boolean models 102
6.3.3 Identification of differentially expressed molecules using literature search and transcriptomic data analysis 103
6.3.4 Computation of all the possible attractors of the RA cell-specific models 103
6.3.5 Validation of the cell-specific models' behaviors 104
6.3.6 Calibration of the RA multicellular model 106
6.3.6.1 Combination of the calibrated states of the RA cell-specific models via intercellular communication 107
6.3.6.2 Identification of differentially expressed molecules using literature search and transcriptomic data analysis 109
6.3.6.3 Computation of all the possible attractors of the multicellular model 109
6.3.6.4 Validation of the multicellular model's behavior 110
6.4 Discussion 110

Chapter 7. Selective M1 macrophage depletion and M2 macrophage promotion in the RA synovium 112

7.1 Introduction 112
7.2 Methods and results 113
7.2.1 Single knockouts of the therapeutic targets present in the RA M1 and M2 macrophage models 113
7.2.2 Double knockouts of the therapeutic targets present in the RA M1 and M2 macrophage models 116
7.2.3 Double knockouts of the receptors expressed by the RA M1 and M2 macrophages 118
7.3 Discussion 119

Chapter 8. Investigation of new therapeutic options in RA using the RA multicellular model 122

8.1 Introduction 122
8.2 Methods and results 123
8.2.1 Single knockouts of the therapeutic targets present in the RA multicellular model 123
8.2.2 Double knockouts of the therapeutic targets present in the RA multicellular model 130
8.3 Discussion 134

General discussion 138

Bibliography 145

ANNEX A 181

Supplementary Data 181
CIFRE phd project proposal 255

ANNEX B 264

Publications 264
Scientific Presentations 264

List of Figures

Figure 1. Deformities of the RA joints in the hand (taken from MedlinePlus (3)).. 1

Figure 2. Boolean toy model to showcase different examples of Boolean formulas. A) A, B, C, D, and E are the network nodes. The directed edge → or —| denotes activation or inhibition. B) Boolean rules governing the nodes' states for the network given in A. Figure taken from Montagud et al., 2022 (16). 3

Figure 3. Multistep progression towards RA disease development. Figure taken from McInnes et al., 2011 (17). 7

Figure 4. Manhattan plots of a GWAS meta-analysis showing the strong association of HLA-DRB1 (in the red square) with RA in the Japanese population. The y axis shows the −log10 P values of the SNPs in the meta-analysis. The SNPs for which the P values were smaller than 1.0 × 10−15 are indicated at the upper limit of the plot. This figure was adapted from Okada et al., 2012 (23). 8

Figure 5. Global prevalence of RA. The figure displays the reported prevalence ranges for rheumatoid arthritis (RA) per country (per 100,000 of the population), as denoted by the key. The figure was taken from Finckh et al., 2022 (65) 12

Figure 6. Differences between the joint structure in rheumatoid arthritis and a healthy joint. Figure adapted from Castro-Sánchez and Roda-Navarro, 2017 (89) 16

Figure 7. RA cellular interplay. The figure is taken from Walsh and Choi, 2014 (97). 17

Figure 8. The role of fibroblasts in RA. Figure taken from Bottini and Firestein, 2013 (111). 19

Figure 9. The role of macrophages in RA. Figure taken from Udalova et al., 2016 (127). 20

Figure 10. Action of biologic DMARDs in rheumatoid arthritis. Figure taken from Huang et al., 2021 (160) 23

Figure 11. Three different types of SBGN networks are used to represent protein phosphorylation catalyzed by an enzyme and modulated by an inhibitor. (a) Process description. (b) Entity Relationship. (c) Activity Flow. Figure adapted from Le Novère et al., 2009 (10). 26

Figure 12. Different representations of a sample Boolean network consisting of three nodes. (A) Graph representation, (B) logical Boolean rules, (C) state transition table and (D) state transition graph. Figure taken from Kaderali and Radde, 2008. (234) 32

Figure 13. Updating schemes in Boolean models. Figure adapted from Schwab et al., 2020 (225). 33

Figure 14. Possible types of attractors in Boolean models. Figure taken from Garg et al., 2008 (226). 34

Figure 15. A) GSE109449 dataset visualization using BBrowser® in Bioturing tool. B) DEA outcomes using the Venice method performed on the GSE109449 dataset.....38

Figure 16. Screenshot of ABIS Shiny application https://giannimonaco.shinyapps.io/ABIS/. A) shows the expression of CXCR3 within the peripheral blood mononuclear cell (PBMC). B) shows the expression of CCR7 within the peripheral blood mononuclear cell (PBMC). C) shows the expression of CCR6 within the peripheral blood mononuclear cell (PBMC). 40

Figure 17. A) Umap plot showing CCR7 expression among the CD4+ T cells clusters identified using Seurat. B) Violin plot showing the expressions of both CXCR3 and CCR6 in the CD4+ T cell clusters. C) CD4+ T cells annotation based on markers expression. 41

Figure 18. Workflow for the construction and use of the RA map. First, cell specific maps are constructed with the CellDesigner software (270) using information from published literature and signaling pathway databases. Then, they go through manual curation and annotation following MIRIAM standard. Next, maps are enriched with additional pathways involved in RA pathogenesis via the integration of disease- and cell-specific gene expression datasets using pathway analysis software like IPA and MetaCore. The resulting RA-Atlas is accessible online as an interactive map via MINERVA platform (271). 46

Figure 19. Annotations added to the MIRIAM section of CellDesigner. 48

Figure 20. A) Macrophage maps from (Kanae Oda et al., 2004) (276) in MINERVA. B) Macrophage map from (Pia Wentker et al., 2017) (277) in MINERVA. Green color indicates the overlay with RA specific molecules retrieved from IPA..49

Figure 21. Initial RA map from Singh et al.,2020 (280). Purple color indicates the

overlay with the latest RA fibroblast overlay available in MINERVA. 50

Figure 22. Th1 map within the atlas of inflammation resolution (281) in CellDesigner. 51

Figure 23. A) Screenshot showing "Pathway maps" option in MetaCore in the "One-click analysis" page. B) Screenshot showing IPA's "canonical pathways" tool. 52

Figure 24. RA Atlas in MINERVA. Blue frames indicate clickable links leading to the different cell-specific maps included in the RA-Atlas. 54

Figure 25. Cellular compartments of the RA M1 macrophage map 55

Figure 26. Percentage of cell-specific components per cell-specific maps. Figure taken from Zerrouk et al., 2022 (283). 62

Figure 27. Visualization of RA macrophage M2 map in MINERVA. (A) Snapshot of the search for VEGFA. (B) Snapshot of the visualization of the macrophage overlay. Figure taken from Zerrouk et al., 2022 (283). 64

Figure 28. Workflow for the construction of the RA multicellular map. We used literature and databases mining via several tools like Causaly (287) and CellphoneDB (178) as well as keywords search in PubMed, to retrieve the intercellular interactions already identified in the published literature. We also integrated omics datasets via ICELLNET (177) and DiSiR (179) to identify statistically significant interactions observed in these datasets. We filtered the retrieved interactions to cross validate the results and only keep the most reliable ones. We used those filtered interactions to connect the RA cell-specific maps and generate the first RA multicellular map. 67

Figure 29. Screenshot of the advanced search option in Causaly. We selected fibroblast and macrophage as cell-types and rheumatoid arthritis as the condition of interest in the concept boxes. We didn't specify any directionality. 69

Figure 30. Overview of CellPhoneDB dataset. Figure taken from Efremova et al., 2020 (178). 70

Figure 31. ICELLNET pipeline to study intercellular communication from cell transcriptional profiles. Figure taken from Noël et al., 2021 (177) 73

Figure 32. Schematic overview and illustration of DiSiR framework taken from Vahid et al., 2023 (179). (A) DiSiR uses as input 1) a single-cell gene

expression matrix, 2) cell type annotations, and 3) a list of putative ligands–receptor interactions at subunit level. DiSiR then characterizes the ligand 'L'–receptor 'R' interaction between cell types 'A' and 'B' based on the products of expressions of 'L' subunits by cell type 'A' and expressions of 'R' subunits by cell type 'B'. This interaction is identified by DiSiR as a significant interaction if these products are significantly higher than random shuffling results for all 'L–R' subunit combinations. (B) Diagram that shows the workflow of data processing, filtering, and visualization of the obtained results in DiSiR. Weak interactions are filtered out if the fraction of cells expressing the ligand (or receptor) within its corresponding cell type is less than a threshold equation Th. DiSiR also provides an interactive graph-based representation module and heatmap/bubble plots to show the resulting interactions for complex signaling pathways at different level.....74

Figure 33. Output plots of ComBat on E_MTAB_8322 and GSE109449 datasets combined. The left plots are density estimates, and the right plots are QQ plots. Top plots represent the means and bottom plots represent the variances. The black plots represent the mean and variance distributions of the combined datasets. The red plots represent the theoretical normal distribution of the mean and variance. If the distribution of the batch effect is normal, black and red plots should overlap. In this case, red and black plots do not overlap, therefore, non-parametric mode of ComBat should be used. 76

Figure 34. Schematic representation of the intercellular interactions in CellDesigner. The black arrow represents the transport of the secreted ligand (Ligand B) from the sending cell (in yellow) to the extracellular space of the receiving cell (in blue) where it binds to its corresponding receptor. The ligand/receptor complex will then transduce the signal via a series of biochemical reactions in the cytoplasm. 79

Figure 35. The RA multicellular map in CellDesigner. Red squares indicate the cell-type of each cell-specific molecular interaction map. Arrows in pink, yellow, blue and purple indicate cell-cell communication.....89

Figure 36. Schematic representation of the workflow we propose to generate and analyze large-scale Boolean models. Molecular interaction maps built in CellDesigner XML format are converted to executable Boolean models using CaSQ. Then, a new version of BMA is deployed to a high-performance computing cluster to identify all the models' attractors. These attractors are filtered to keep only the steady states. Next, the filtered steady states are validated. Differentially expressed biomolecules present in the models are

identified using both literature mining and transcriptomic data analysis. The expressions of the identified biomolecules are discretized and converted to a binary vector of experimentally observed Boolean values. After that, similarity scores are computed to describe the similarity between the steady states and the experimentally observed values. The steady states with the highest score are selected, their mean vector represents the calibrated model's state.....93

Figure 37. A) Toy XML file describing the formation of the complex AB. This reaction is catalyzed by either Activator 1 or Activator 2 (green arrows) and inhibited by Inhibitor 1 (red arrow). B) The Boolean model of the toy XML file depicted in A that we generated using CaSQ. C) The logical rules associated with each node in the Boolean model depicted in B.....95

Figure 38. Plot showing the number of processed inputs' combinations by BMA per hour using a single core machine with a base frequency of 2.3 GHz. 98

Figure 39. Example of gene expression discretization on a volcano plot showing the DEGs between RA and osteoarthritis synovial fibroblasts from GSE109449 dataset. DEGs were filtered using an FDR equal to 0.05 and a logFC equal to 0.584. Overexpressed DEGs are associated with the value one while under expressed molecules are associated with the value zero. 100

Figure 40. RA multicellular model in BMA graphical interface. 107

Figure 41. A) CXCL13 mathematical rule in RA M1 macrophage model before adding the intercellular interactions. B) CXCL13 mathematical rule in RA M1 macrophage model after adding the intercellular interactions. 108

Figure 42. a) In silico simulation of the proliferation and apoptosis phenotypes in the RA M1 macrophage model before the NF-κB KO. b) In silico simulation of the proliferation and apoptosis phenotypes in the RA M1 macrophage model after the NF-κB KO. 115

Figure 43. a) In silico simulation of the proliferation and apoptosis phenotypes in the RA M2 macrophage model before the GSK3β KO. b) In silico simulation of the proliferation and apoptosis phenotypes in the RA M2 macrophage model after the GSK3β KO..... 116

Figure 44. a) In silico simulation of the proliferation and apoptosis phenotypes in the RA M1 macrophage model after ERK1 KO. b) In silico simulation of the proliferation and apoptosis phenotypes in the RA M1 macrophage model after ERK1 and Notch1 KO. 118

Figure 45. A) In silico simulation of the M2 macrophage cell-specific phenotypes before GSK3β KO. B) In silico simulation of the M2 macrophage cell-specific phenotypes after GSK3β KO. 126

Figure 46. A) In silico simulation of the M1 macrophage cell-specific phenotypes before NF-κB KO. B) In silico simulation of the M1 macrophage cell-specific phenotypes after NF-κB KO. 127

Figure 47. A) In silico simulation of the fibroblast cell-specific phenotypes before CAV1 KO. B) In silico simulation of the fibroblast cell-specific phenotypes after CAV1 KO. 128

Figure 48. A) In silico simulation of the Th1 cell-specific phenotypes before TBX21 KO. B) In silico simulation of the Th1 cell-specific phenotypes after TBX21 KO. 129

Figure 49. A) In silico simulation of the Angiogenesis phenotype before MIR221 KO. B) In silico simulation of the Angiogenesis phenotype after MIR221 KO. 129

Figure 50. A) In silico simulation of the Matrix degradation phenotype before NF-κB KO. B) In silico simulation of the Matrix degradation phenotype after NF-κB KO. 130

Figure 51. A) In silico simulation of the Inflammation phenotype before JAK1/JAK2 KO. B) In silico simulation of the Inflammation phenotype after JAK1/JAK2 KO. 133

Figure 52. A) In silico simulation of the Osteoclastogenesis phenotype before NF-κB/STAT3 KO. B) In silico simulation of the Osteoclastogenesis phenotype after NF-κB /STAT3 KO. 134

List of Tables

Table 1. RA classification criteria: domains, categories and point scores. The table was taken from Kay et al., 2012 (81). 14

Table 2. Description of the omics datasets and their associated outcomes. 44

Table 3. IPA and MetaCore maps using the RA fibroblast DEGs list. 56

Table 4. Top 5 IPA and MetaCore maps using RA macrophage DEGs list. 58

Table 5. Top 5 IPA and MetaCore maps using RA Th1 DEGs list. 60

Table 6. List of enriched gene sets in RA Th1 after filtering. 60

Table 7. Pairs of datasets used to infer inward and outward cell-cell communication between RA macrophages and RA fibroblast. 75

Table 8. Pairs of datasets used to infer inward and outward cell-cell communication between RA fibroblasts and RA CD4+ Th1. 77

Table 9. Pairs of datasets used to infer inward and outward cell-cell communication between RA macrophages and RA CD4+ Th1. 78

Table 10. Cell-cell interactions between RA macrophage and RA fibroblast in the synovium. 81

Table 11. Cell-cell interactions taking place between RA macrophage and RA CD4+ Th1 in the synovium. 84

Table 12. Cell-cell interactions taking place between RA CD4+ Th1 and RA fibroblast in the synovium. 86

Table 13. Number of nodes and reactions in the RA cell-specific molecular interaction maps. 101

Table 14. Number of nodes, inputs, and interaction in the RA cell-specific Boolean models 102

Table 15. Number of nodes and inputs upstream of the cell-specific phenotypes in the RA cell-specific Boolean models. 102

Table 16. Expected Boolean states of the phenotypes in the RA cell-specific Boolean models according to the literature. 105

Table 17. Description of the RA cell-specific models' calibrated states. The first row displays the number of steady states with the highest similarity scores. The second row shows similarity scores describing the ability of the calibrated models to reproduce the experimentally observed expressions. The third row indicates the number of nodes fixed at zero or one in the calibrated states of the models. 105

Table 18. Mismatches identified during the validation of the models' behaviors. 106

Table 19. Number of nodes fixed in the RA cell-specific models before and after adding cell-cell interactions. 108

Table 20. Single knockouts of the therapeutic targets from the TTD database that perturb the RA macrophages' phenotypes. 114

Table 21. Combinations of therapeutic targets (from TTD database) that perturb the RA macrophages' phenotypes. 117

Table 22. Therapeutic targets from TTD database that perturb RA phenotypes in the RA multicellular model. Red arrows describe the inhibition of the phenotypes that were active in the calibrated state of the model. Green arrows describe the activation of the phenotypes that were inhibited in the calibrated state of the model. The absence of arrows means that the phenotype state remains unchanged. 124

Table 23. Description of the therapeutic targets perturbing the RA phenotypes in the RA multicellular model. 124

Table 24. Combinations of therapeutic targets (from TTD database) that perturb RA phenotypes in the RA multicellular model. . Red arrows describe the inhibition of the phenotypes that were active in the calibrated state of the model. Green arrows describe the activation of the phenotypes that were inhibited in the calibrated state of the model. The absence of arrows means that the phenotype state remains unchanged. 131

Table 25. Description of the therapeutic targets perturbing the disease phenotypes in the RA multicellular model. 131

List of Abbreviations

AAPA	Anti-Acetylated Peptide Antibody
aCarP	Anti-Carbamylated Protein Antibody
ACPA	Anti-Citrullinated Protein Antibody
ACR	American College of Rheumatology
AF	Activity Flow
AMPA	Posttranslationally Modified Protein Antibody
APC	Antigen Presenting Cell
BMA	Bio Model Analyzer
BsAbs	Bispecific Antibodies
CaSQ	CellDesigner as SBML-qual
CAV1	CAVeolin 1
CCL2	C-C Motif Chemokine Ligand 2
CREB1	CAMP Responsive Element Binding protein 1
DEA	Differential Expression Analysis
DEG	Differentially Expressed Gene
DiSiR	Dimer Signal Receptor
DMARDs	Disease-Modifying Antirheumatic Drugs
ER	Entity Relationship
ERK1	Extracellular signal-Regulated Kinase 1
ES	Enrichment Score
EULAR	European League Against Rheumatism
FC	Fold Change
FDA	Food and Drug Administration

FDR	False Discovery Rate
GEO	Gene Expression Omnibus
GM-CSF	Granulocyte Macrophage Colony-Stimulating Factor
GSEA	Gene Set Enrichment Analysis
GSK3β	Glycogen Synthase Kinase-3 beta
HGNC	HUGO Gene Nomenclature Committee
HLA	Human Leukocyte Antigen
ICAM1	Intercellular Adhesion Molecule 1
IFNG	Interferon Gamma
IgG	Immunoglobulin G
IL-6	InterLeukin 6
IPA	Ingenuity Pathway Analysis
JSON	JavaScript Object Notation
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	KnockOut
MAPK	Mitogen-Activated Protein Kinase
MHC	Major Histocompatibility Complex
MINERVA	Molecular Interaction NEtwoRk VisuAlization
MIRIAM	Minimal Information Requested In the Annotation of biochemical Models
MOA	Mode Of Action
MoAbs	Monoclonal Antibodies
mTOR	mammalian Target Of Rapamycin
NF-kB	Nuclear Factor kappa-light-chain-enhancer of activated B cells

Notch1	Neurogenic locus notch homolog protein 1
NSAIDs	Nonsteroidal Anti-Inflammatory Drugs
OA	OsteoArthritis
PAD	PeptidylArginine Deiminase
PBMC	Peripheral Blood Mononuclear Cell
PD	Process Description
PI3K	PhosphatIdylinositide-3-Kinase
RA	Rheumatoid Arthritis
RF	Rheumatoid Factor
SBGN	Systems Biology Graphical Notation
SBML	Systems Biology Markup Language
SE	Shared Epitope
STAT	Signal Transducers and Activators of Transcription
Tfh	Follicular Helper T cells
Th1	T Helper 1
TLRs	Toll-Like Receptors
TNF	Tumor Necrosis Factor
Treg	Regulatory T cell
TTD	Therapeutic Target Database
VEGF	Vascular Endothelial Growth Factor

Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory autoimmune disease whose causal mechanisms are still not fully understood. RA pathogenesis involves a combination of genetic, epigenetic, and environmental factors and arises from intricate interplays between different biological processes such as signal transduction, gene regulation and metabolism. Deregulated activation of the Janus Kinase/Signal Transducers and Activators of Transcription (JAK/STAT), the stress-activated protein kinase/mitogen-activated protein kinase (SAPK/MAPK) and Phosphatidylinositide-3-Kinase/AKT/mammalian Target of Rapamycin (PI-3K/AKT/mTOR) pathways, among many others, lead to an up-regulated expression of matrix metalloproteinase and proinflammatory cytokines, resulting in the chronic degradation and inflammation of the synovial tissue (1). The inflammatory cascade leads to joint hyperplasia, cartilage damage, and bone destruction. The affected joints lose their shape and alignment, resulting in deformities and loss of function (Figure 1). Pain, swelling, tenderness and stiffness of small joints in the hands and feet are the first RA symptoms, but extra-articular manifestations can also appear, affecting the gastrointestinal, respiratory, and cardiovascular systems (2).

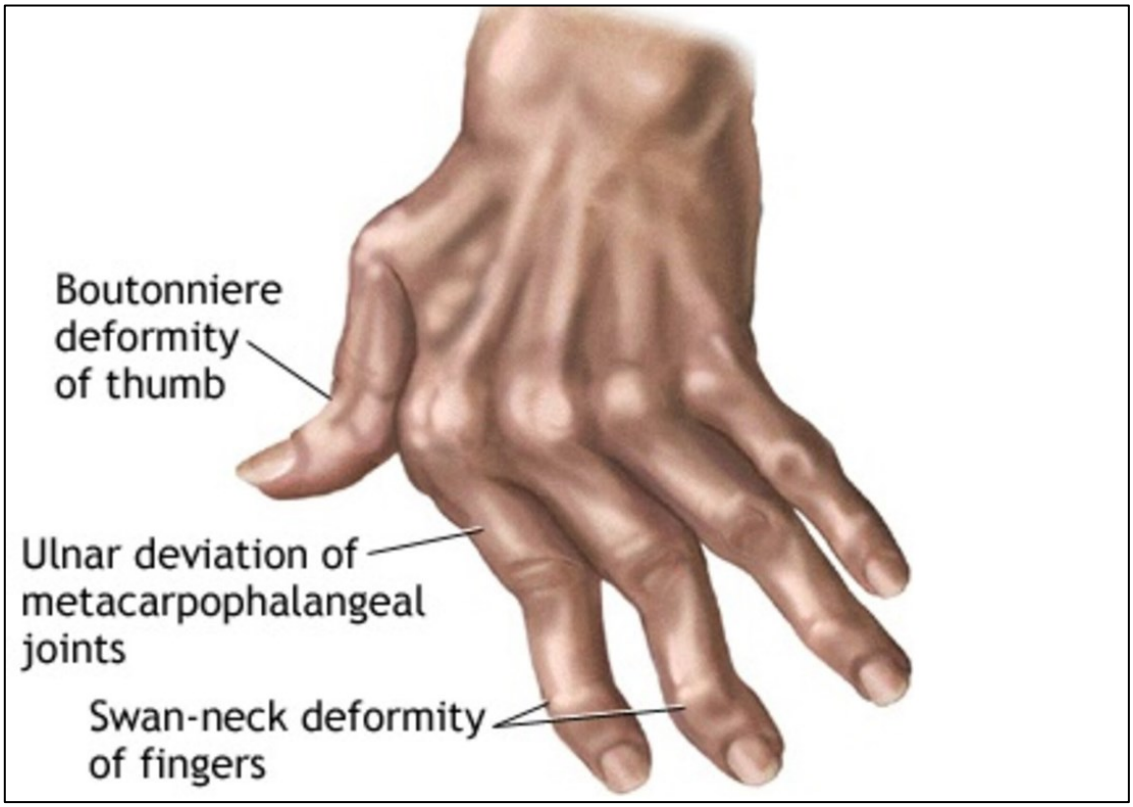


Figure 1. Deformities of the RA joints in the hand (taken from MedlinePlus (3)).

The pathogenic behavior of the RA synovial tissue cannot be associated with a single cell type. It is the result of cellular communication between local synoviocytes and cells from the innate and adaptive immune system. These intercellular interactions determine numerous aspects of the disease's pathophysiology and can activate or downregulate specific synovial cell populations. Cell cell communication can also impact the intracellular signaling and the expression of inflammatory substances and matrix remodeling enzymes in the joints, leading to chronic inflammation, matrix degradation, and bone erosion (4).

There is currently no curative therapy for RA. The wide range of molecules associated with RA, the intricate cellular crosstalks and the disease heterogeneity are all barriers to healing RA. However, many therapies have been developed to relieve disease symptoms, improve survival, and avoid long-term disability (5). However, these therapies have been associated with several adverse effects, and a substantial proportion of RA patients are non-responders (6,7). A better understanding of the cellular communication and the intracellular cascades involved in the disease pathogenesis could help elucidate the mode of action of current RA drugs and identify new therapeutic options.

Systems biology can offer a global view of such complex diseases through the construction of disease maps (8). Disease maps are detailed interaction networks describing molecular processes as well as evidence of their relationship to disease phenotypes (9). Disease maps are manually curated and extensively annotated through broad literature search. They are also standardized to enable an efficient and accurate representation, visualization, storage, exchange, and reuse of knowledge (10,11). These maps can be utilized for the identification of biomarkers, potential drug targets and disease subtypes. Furthermore, they provide a comprehensive template for visualization, analysis, and interpretation of omics datasets. However, these maps are static and cannot be used to generate hypotheses or predictions regarding the behavior of the system under different perturbations.

Computational modeling is a powerful tool to understand the emergent features and behaviors of the complex biological systems described in disease maps. Among the modeling techniques available, Boolean formalism is the most suited for large-scale biological systems. Indeed, Boolean modeling does not include kinetic parameters that can be difficult to determine in most systems. It can make use of experimental literature and high-throughput technologies to retrieve a large amount of qualitative data on individual components and interactions

(12,13). In these models, nodes represent biological entities (genes, proteins, metabolites…). Each node can have two values: zero for inactivity or absence and one for activity or presence. The different components of the modeled system are connected through edges that represent their regulatory interactions. These regulatory dependencies are expressed by Boolean functions: "AND", "OR" and "NOT" (Figure 2), and the value of each component is determined by these Boolean functions.

When simulated, the global state of the model (i.e., the state of the entire system) changes and eventually reaches particular configurations called attractors. Attractors are single states or a reoccurring sequence of states that describe the long-term behavior of the model. They have been connected to biological phenotypes, making their computation a key point in the analysis of Boolean models (14). Considering that the number of global states of a Boolean model is exponentially dependent on its node number (2^n global states for n nodes), attractors' search in large-scale biological systems is computationally demanding (15). Moreover, as attractors change based on the external perturbations that the model receives from its environment, their number can be massive, and the biologically consistent states still need to be identified and interpreted in a cell and disease specific manner.

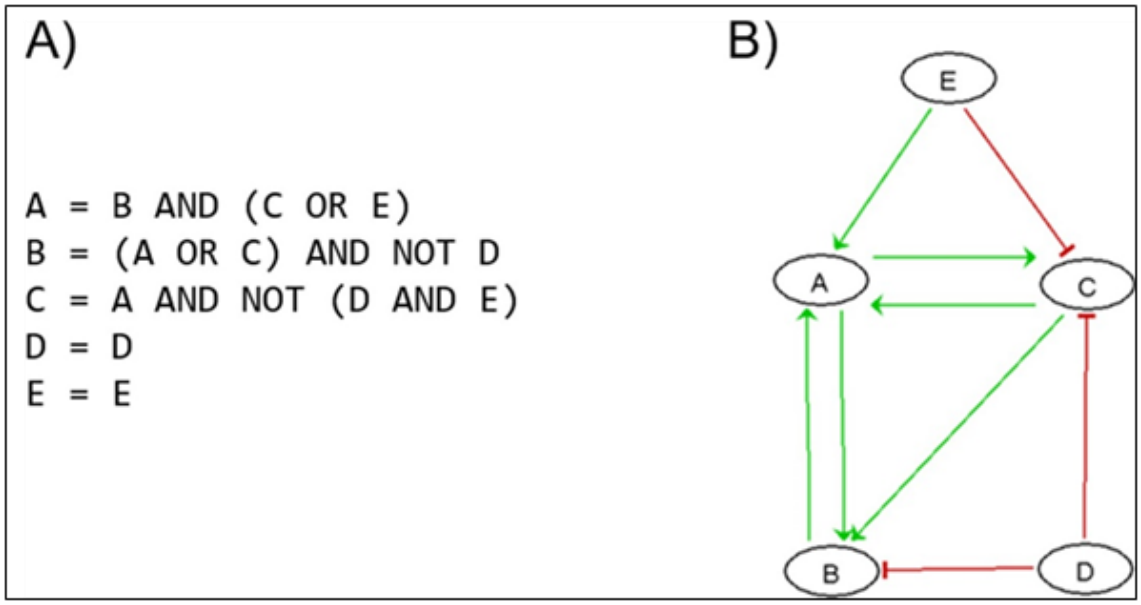


Figure 2. Boolean toy model to showcase different examples of Boolean formulas. A) A, B, C, D, and E are the network nodes. The directed edge → or —| denotes activation or inhibition. B) Boolean rules governing the nodes' states for the network given in A. Figure taken from Montagud et al., 2022 (16).

This Ph.D. project is a collaboration between the Genhotel laboratory and Sanofi in the context of a CIFRE (Conventions Industrielles de Formation par la Recherche) program. CIFRE is a french program that enables companies to benefit from financial assistance to recruit a doctoral student whose research work, supervised by a public research laboratory, will lead to the defense of a thesis. The objective of this program is to reinforce exchanges between public research laboratories and the socio-economic world, encourage the employment of PhDs in companies and contribute to the innovation process of companies established in France. The Cifre program is fully funded by the French Ministry of Higher Education, Research and Innovation, which has entrusted its implementation to the National Association for Research and Technology (ANRT) (17) .

The project proposal that was approved by the ANRT for this PHD thesis is provided in the CIFRE PhD project proposal section. We were interested in developing an efficient computational system biology approach to obtain a multicellular system-level understanding of RA. We wanted to provide a detailed representation of the structure and dynamics of the intricate cell-cell interactions in RA as well as the complex intracellular cascades and their regulation in this disease. Also, we wanted to shed light on the mechanism of action of the current RA therapies and to identify new potential therapeutic options and drug combinations.

The first objective of the Ph.D. thesis was to build a state-of-the-art RA-Atlas that describes the rheumatic joint using both literature search and omics data integration. We aimed at recapitulating the existing and fragmented knowledge (gene expression, signaling pathways, cellular phenotypes) about RA pathogenesis in a cell-specific manner. In this direction, we constructed four molecular interaction maps specific to the most abundant cell populations in RA synovium: synovial macrophages, including the M1 and M2 phenotypes, synovial fibroblasts and CD4+ T helper 1 (Th1) subtype.

The second objective was to obtain a multicellular level description of the RA joint. Using both literature mining and omics data analysis and integration, we identified bidirectional intercellular interactions between macrophages, fibroblasts and Th1 cell-types in the RA synovium. These cell-cell interactions were used to connect the corresponding cell-specific molecular maps within the RA-Atlas and build the first RA multicellular map.

The third objective was to use the RA molecular interaction maps as a scaffold to automatically generate their corresponding executable Boolean models. The goal was to develop an efficient computational framework that uses high-performance computing clusters to efficiently perform the attractors search on such

large-scale models, and to validate these attractors as consistent biological states using omics datasets analysis and prior knowledge.

The fourth objective consisted in using the validated and calibrated models to investigate new therapeutic options in RA via *in-silico* single and double knockout simulations. The goal was to investigate and comprehend the mechanism of action of current RA treatments, test new drug combinations, identify new potential therapeutic intervention points, and perform drug repurposing analysis for the identified targets.

Through this collaboration, the Genhotel laboratory aims at expanding its expertise in maps and models construction and validation for RA via the integration of the multicellular scale. To achieve this objective, the laboratory benefits from Sanofi's calculation services and facilitated access to commercially available databases and softwares as well as its expertise in RA.

Sanofi, on the other hand, is interested in increasing its skills in the domain of disease map and Boolean based model simulation. The expected output from this collaboration for the company is to have access to a multiscale model of RA, fully operational and easy to maintain, that is able to simulate the pathogenesis. This model will allow Sanofi to have a better understanding of the disease with a full up to date knowledge base on RA. It will also be used to integrate and analyse in-house data, perform large-scale screenings of potential therapeutic targets and to better understand the mechanism of action of drugs currently under development at Sanofi.

The thesis manuscript starts with an overview of the RA disease pathogenesis, followed by an introduction to systems biology approaches suitable for studying complex human diseases. Next, we describe the data available for our analyses, and more precisely the omics datasets along with the statistical analyses and associated results. After that, five independent chapters follow, each subdivided into an introduction, method, result and discussion sections, to describe a) the construction of the RA cell-specific molecular interaction maps, b) the construction of the RA multicellular map, c) the generation and calibration of the RA Boolean models, d) the *In-silico* single and double knockouts simulations we performed on the calibrated RA macrophages models and e) the *In-silico* simulations we performed on the RA multicellular model. Finally, we complete this thesis manuscript with a general discussion and future directions for the project.

Chapter 1. The pathogenesis of rheumatoid arthritis

1.1 Etiology

Figure 3 depicts how susceptibility genes, epigenetic alterations, and environmental factors all contribute to the development of RA disease (1). This combination leads to autoantibodies production and loss of tolerance. The presence of autoantibodies results in systemic autoimmunity which, under the influence of infectious/neuroimmune triggers, causes the transition to arthritis. In this section, we describe the genetic, epigenetic and environmental factors associated with the development of RA.

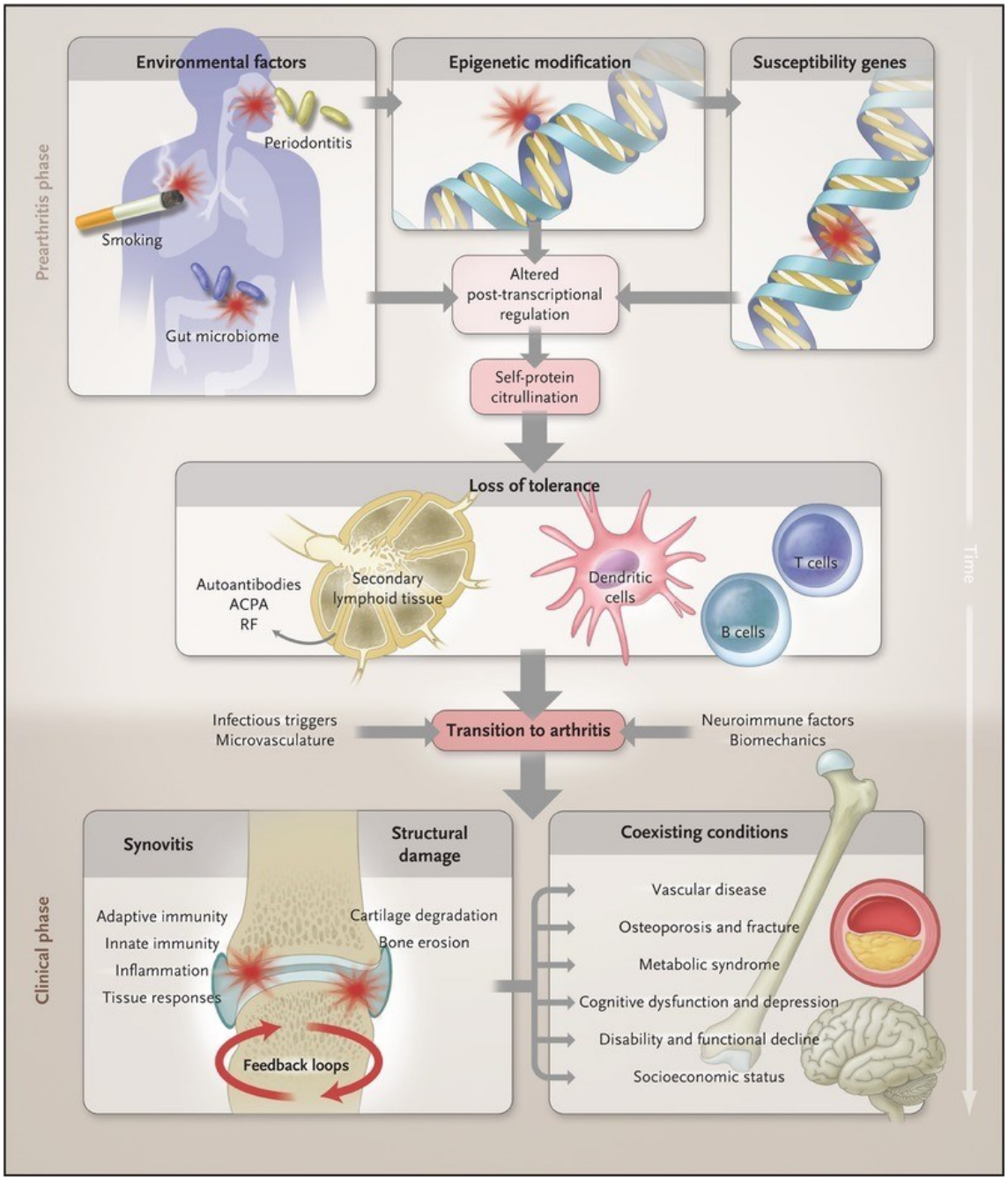


Figure 3. Multistep progression towards RA disease development. Figure taken from McInnes et al., 2011 (18).

1.1.1 Genetic susceptibility

The presence of autoantibodies like Rheumatoid Factor (RF) and Anti-Citrullinated Protein (ACPA) in the RA patients' synovial fluid and serum is an early hallmark for the disease (19). Based on ACPA status, RA patients can be sub-grouped into two major subsets: ACPA-positive RA (ACPA+ RA) and ACPA-

negative RA (ACPA- RA). Genetic predisposition plays a significant role in the development of RA in both groups. Indeed, the overall heritability of RA (the amount of variation in disease susceptibility that can be explained by genetic factors) is estimated to be 66% and is comparable in ACPA positive and ACPA negative RA (20).

Genome-Wide Association Study (GWAS) meta-analysis have identified more than 100 loci associated with RA (21). These loci provide important insights into relevant pathways involved in the disease process but have limited utility in clinics as a prediction tool. Most of them are not unique to RA but rather reflect susceptibility alleles that associate with other autoimmune diseases as well, such as PTPN22, for which a common single nucleotide polymorphism confers increased risk for the development of type 1 diabetes, RA, systemic lupus erythematosus, vitiligo, and Graves' disease (22). The association of HLA region with RA, on the other hand, is distinct and unique. Indeed, HLA explains most of the genetic risk (~13%) with an additional 5% of the genetic risk being explained by the 100 other loci discovered to date. HLA region encodes the major histocompatibility complex (MHC) and contains approximately 250 genes, 60% of which are immune-related (21). HLA-DRB1 gene is the major genetic susceptibility locus for RA (Figure 4). HLA-haplotypes associated with seropositive RA are characterized by HLA-DRB1 alleles having a common amino acid sequence, termed the HLA Shared Epitope (HLA-SE) (22). On the other hand, HLA-haplotypes associated with seronegative RA are non-SE-bearing HLA alleles. Differences between ACPA+ RA and ACPA- RA are further supported by studies demonstrating that many of the identified genetic risk loci in other chromosomal regions (non-HLA genes) only predispose to ACPA-positive RA (23). This indicates that ACPA-positive and ACPA-negative diseases are distinct disease entities.

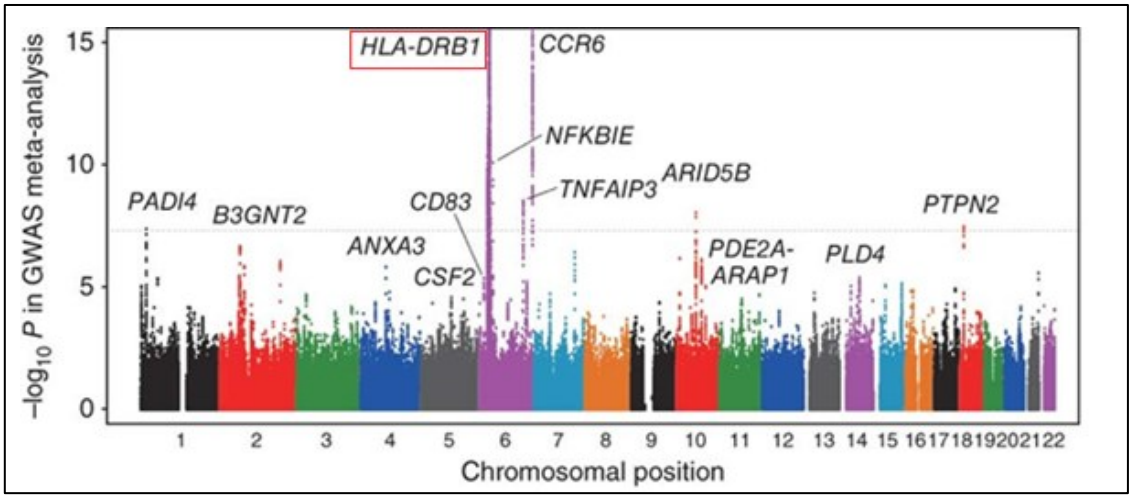


Figure 4. Manhattan plots of a GWAS meta-analysis showing the strong association

of HLA-DRB1 (in the red square) with RA in the Japanese population. The y axis shows the −log10 P values of the SNPs in the meta-analysis. The SNPs for which the P values were smaller than 1.0 × 10−15 are indicated at the upper limit of the plot. This figure was adapted from Okada et al., 2012 (24).

Despite the success to date in identifying risk loci for RA, mostly through meta-analysis of multiple cohorts, about 40% of the genetic component of susceptibility remains unaccounted for. Indeed, genetic heterogeneity, low penetrance of individual disease alleles and the potential for gene–gene interactions hinder the identification of RA genetic factors (25). In this direction, several studies investigated epistasis in RA, and provided evidence of epistatic interactions between polymorphisms in TNFRSF14 and TNFRSF6B (26), BANK1 and BLK (25), HLA-DRB1 and PTPN22 (27), and HTR2A and HLA-DRB1 (28).

1.1.2 Epigenetic factors

Genetic heterogeneity cannot explain all aspects of RA. Thus, investigating the epigenetic impact is becoming increasingly relevant (29). Epigenetic changes in RA have been studied in different types of immune cells. Researchers revealed that inflammatory phenotype of autoimmune monocytes is regulated on the epigenetic level through the increased acetylation of disease-associated genes such as CXCL9, CXCL10, IRF1, and IFNGR (30). They also identified differentially methylated loci in RA B cells (31), including DNA hypermethylation at CD1C, DNA hypomethylation at TNFSF10, both linked to an increased risk of RA, and differential methylation at genes involved in CBL pathways, known to control the B-cell-intrinsic checkpoint of immune tolerance (32). RA CD4+ T cells exhibit hypermethylated genes as well, including JUN, STAT1, PTEN, and CD44. The resulting upregulation of JUN expression activates the MAPK signaling pathway and regulates cell proliferation, cell cycle, and apoptosis. STAT1 activation, on the other hand, is related to IL-6 inflammatory factors and increases the joint infiltration of T cells (33). Furthermore, histone modifications were observed in RA PBMCs through a higher activation of histone deacetylases in RA patients compared to healthy controls. These modifications may decrease expression of cell-cycle related molecules such as p16, p21 and p53, and increase TNF production in RA synovial tissue (34).

Because of the aggressive phenotype of fibroblasts in RA, epigenetic changes were also investigated in this cell type. Fibroblasts have a distinct, non-random methylation pattern that is rearranged specifically throughout the disease's course and differs depending on joint localization (35). Several studies revealed differentially methylated genes that are part of pathways known to be involved in RA pathogenesis (36–38). Other studies showed an overall increase of acetylation

associated with reduced histone deacetylases activity with an abnormal modification of histones involved in the activation fibroblasts (39–41). These epigenetic mechanisms lead to the expression of genes that are implicated in extracellular matrix interactions, adhesion, cell recruitment, and migration in RA fibroblasts. They also result in an increased resistance to apoptosis and promote fibroblasts' proliferation (42).

Other epigenetic modifications include proteins' post-translational modifications. Two of these modifications, citrullination and carbamylation, seem to be involved in the pathogenesis of RA. Citrullination is an enzymatic post-translational modification that is mediated by PeptidylArginine Deiminases (PAD), which transform peptide-bound arginine residues into citrulline, a non-natural amino acid (43). Studies suggest that citrullination is dysregulated in RA through a hyperactivation of PADs, leading to an accumulation of citrullinated proteins in the joint (the RA citrullinome) (44). Dysregulated PAD enzyme activity subsequently promotes pro-inflammatory cytokine and autoantibodies production, and bone destruction in RA joints (45). Carbamylation, on the other hand, is a chemical nonenzymatic post-translational modification in which homocitrulline residues are generated by the reaction of cyanate with the primary amine of lysine residues. Under normal physiological circumstances, cyanate levels are too low to induce substantial carbamylation. However, conditions like inflammation, and the exposure to cigarette smoke enhance cyanate levels (46). Even though increased level of carbamylated proteins is associated with RA, their exact pathogenic role remains unaddressed (43).

1.1.3 Environmental factors

RA has been linked to a variety of environmental, dietary, and lifestyle factors. Multiple studies have estimated that tabaco exposure accounts for 20-30% of environmental risk for RA (47). It is most strongly associated with ACPA positive RA, and especially ACPA positive RA in the setting of the SE. Smoking has been associated with the presence of RFs as well and may lead to increased citrullination which in the setting of the right genetic background, leads to presentation of citrullinated proteins and the generation of ACPA (48). Aside from exposure to tobacco smoke, multiple studies have consistently shown a link between dust exposure, air pollution and ACPA-positive RA. These findings are likely influenced by confounders such as lower income, which is a supposed risk factor for RA by itself (49–51). Several dietary or other factors, such as supplements or medications, have also been linked to RA. Lower intake of vitamin D and antioxidants, as well as higher intake of sugar, sodium, red meats, protein, and iron, are associated with an increased risk of RA (52–56). On the other hand, modest alcohol consumption and increased intake of fatty acids were associated

with decreased risk for RA (57,58).

Women constitute approximately two-thirds of RA patients, and many epidemiologic studies point to sex-related factors in RA risk. Female-specific variables that increase RA risk include early menopause (59), the presence of polycystic ovary syndrome (60), and potentially pre-eclampsia (61). On the other hand, several studies indicate the presence of sex-specific protective factors including breast-feeding (62), use of hormone replacement therapy (63) and oral contraception (64).

The impact of environmental factors on the risk of developing RA raises several methodological and practical issues. The literature on environmental risk factors is relatively scarce, with findings frequently lacking reproducibility (65).

1.2 Epidemiology

RA is a relatively common disease, affecting millions of people worldwide. It is the most prevalent systemic autoimmune disease among the rheumatic inflammatory musculoskeletal diseases (66). The prevalence and incidence of RA have been evaluated in several studies. Given the high heterogeneity in the methodologies and in the different definitions of RA that were used, epidemiological estimates from the literature are challenging to compare. The Global Burden of Disease (GBD) study of 2017 revealed a global prevalence at 0.25% (95% CI 0.24–0.3%) (67). More recently, a meta-analysis performed in 2021 using stricter inclusion criteria revealed a pooled prevalence of 0.46% (95% CI 0.37–0.57%) (68).

RA prevalence varies widely depending on geographic region. It is the highest in North America (0.38%; 95% CI 0.36–0.40%), Western Europe (0.35%; 95% CI 0.31–0.38%) and the Caribbean (0.34%; 95% CI 0.30–0.37%) (66) (Figure 5).

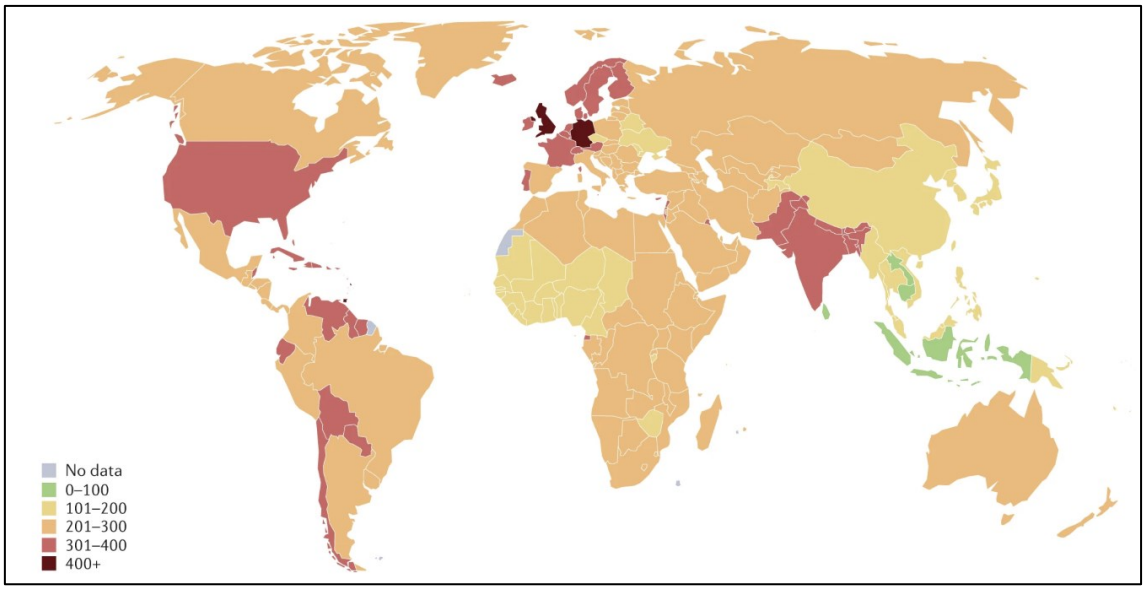


Figure 5. Global prevalence of RA. The figure displays the reported prevalence ranges for rheumatoid arthritis (RA) per country (per 100,000 of the population), as denoted by the key. The figure was taken from Finckh et al., 2022 (66).

The prevalence of RA also varies depending on sex and age. It is 4-5 times higher in females than male below the age of 50, but above 60-70 years the female/male ratio is only about 2 (69).

RA has also increased mortality. Indeed, RA patients have 40% - 50% increased total mortality risk compared to the population without RA, with significantly increased respiratory and cardiovascular disease mortality (70,71). However, this mortality difference appears to be narrowing (72–74), possibly as a result of early aggressive therapy, increased therapeutic alternatives, and a better long-term chronic inflammation control.

1.3 Diagnosis

Both RF and ACPA are used for RA early diagnosis. Their sensitivities are similar (67% for ACPA and 69% for RF), however ACPAs are more specific for RA compared to RF when compared to healthy controls (95% for ACPA and 85% for RF) (75). Furthermore, patients with a variety of non-rheumatic diseases like Sjogren's syndrome, bacterial and viral infections can also have RFs detected (76). The presence of ACPA is associated with RA irrespective of RF status, while the association of RF with disease relies on its interaction with ACPA (77). ACPA alone have then enough predictive power to effectively distinguish high-risk individuals from the background population (78); nevertheless, a third of RA patients are ACPA negative, which makes RA difficult to diagnose in the early stages for this population (79). Recent studies have been conducted in this direction to identify

new markers in ACPA- RA patients like MMP3 (79) and antibodies against carbamylated proteins (aCarP) (80). Imaging tests of the hands and feet, like X-rays and Magnetic Resonance Imaging (MRI), can also be performed to check the joint damage in RA patients and monitor the progression of the disease. However, they are not generally abnormal in the early stages of RA, before joint damage occurs (81).

To allow early interventions in the disease course and to prevent the joint destruction that follows inflammation, the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) put in place revised classification criteria emphasizing RA characteristics that emerge early in the disease course. These classification criteria combine imaging evidence, symptoms duration and serological markers identification. They are grouped into four categories, with point scores for each: joint symptoms; serology (including RF and/or ACPA); symptom duration, whether <6 weeks or >6 weeks; and acute-phase reactants (C-Reactive Protein (CRP) and/or Erythrocyte Sedimentation Rate (ESR)) (Table 1). The calculated scores are used to predict which patients would go on to develop RA (82).

Table 1. RA classification criteria: domains, categories and point scores. The table was taken from Kay et al., 2012 (82).

Domain	Category	Point score
Joint involvement (0–5 points)	1 large joint	0
	2–10 large joints	1
	1–3 small joints (large joints not counted)	2
	4–10 small joints (large joints not counted)	3
	>10 joints including at least one small joint	5
Serology (at least one test needed for classification; 0–3 points)	Negative RF and negative ACPA	0
	Low positive RF or low positive ACPA	2
	High positive RF or high positive ACPA	3
Acute-phase reactants (at least one test needed for classification; 0–1 point)	Normal CRP and normal ESR	0
	Abnormal CRP or abnormal ESR	1
Duration of symptoms	<6 weeks	0
	≥6 weeks	1

1.4 Pathophysiology

The earliest break in self-tolerance in RA is characterized by the production of autoantibodies, including RFs and Anti-post translationally Modified Protein Antibodies (AMPA) (22). RF are primarily directed against antigenic determinants in the Fc portion of the IgG molecules (19,83), while AMPA are directed against

post-translationally modified proteins and comprise antibodies against various modifications such as carbamylation (aCarP), acetylation (AAPA) and citrullination (ACPA)(22). These abnormalities in the cellular and humoral immune response in RA is likely influenced by genetic predisposition and environmental factors including abnormal activation of T cells, infectious/neuroimmune triggers, genetically susceptible B cell repertoire, or failing checkpoints of tolerance control. HLA-DR genes, for instance, appear to regulate autoantibody production and the Shared Epitope (SE) is considered as the most significant risk factor for increased ACPA production in RA (84). Mechanistically, SE-bearing HLA-DR molecules are able to present self-peptides that lead to a tolerized CD4 T cell repertoire, enhancing the reactivity of these cells to citrullinated proteins and subsequently leading to autoreactive B cells (22). Furthermore, RA susceptibility haplotype in PADI4 produced more stable PAD transcripts (PAD is the enzyme that catalyzes peptides citrullination), thus allowing for increased levels of ACPA antibodies in RA patients (85). The presence of autoantibodies in RA leads to systemic autoimmunity which by itself is initially reversible and self-limiting, but under the influence of repetitive triggering and/or other unknown factors, transient systemic autoimmunity become persistent (22). The produced antibodies form immune complexes in the joint, leading to the attraction, through complement or direct activation, of innate and adaptive immune cells, such as dendritic cells, T-lymphocytes, neutrophils, and B-lymphocytes, in the synovium. Migration of B cells and T cells towards the synovial compartment is involved in initiating synovial inflammation through the subsequent activation of tissue-resident cells including macrophages and fibroblasts (19,22). The interactions between these cells cause high production of chemokines and cytokines such as IL-1, IL-6, IL-17 and TNF which ultimately contributes to chronic inflammation of the joint. The accumulation of several types of immune cells in the synovial tissue turns the healthy lining structure into a pannus-like structure and leads to angiogenesis via the high proliferation of blood vessels (86). The excessive number of macrophages and fibroblast causes synovial hyperplasia. The synovial lining, which usually has 1-3 cell layers, thickens noticeably and significant amounts of matrix degradation enzymes are secreted, degrading the cartilage matrix (Figure 6). This environment favors the polarization of macrophages towards the M1 pro inflammatory phenotype and inhibits the proliferation of the M2 anti-inflammatory phenotype. This RA environment also perturbs the delicate balance among T-helper 1 (Th1) and T-helper 2 (Th2), leading to an aberrant and uncontrolled Th1/M1 activation which further causes organ damage (87). Osteoclastogenesis is affected by the inflammatory state in the joint as well. Autoimmune responses in RA lead to increased osteoclastic bone resorption and impaired osteoblastic bone formation (88). Indeed, under specific stimuli from the RA joint, circulating monocytes migrate to a specific location in the bones and

fuse with each other to become mature multinucleated osteoclasts (89).

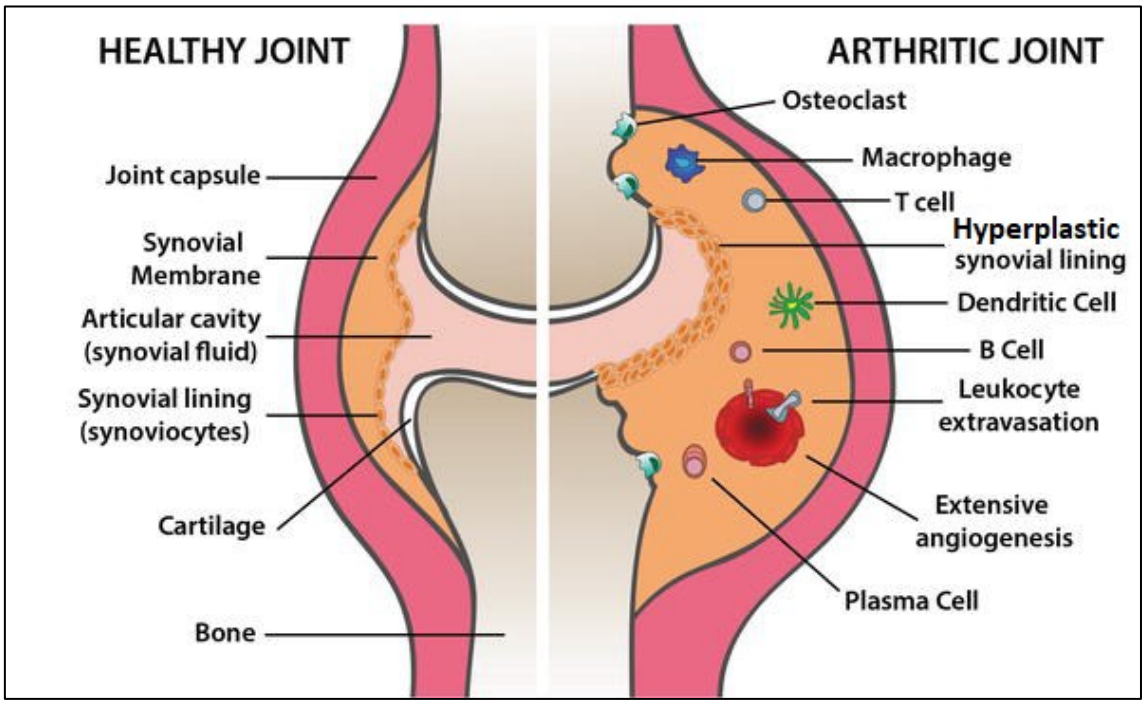


Figure 6. Differences between the joint structure in rheumatoid arthritis and a healthy joint. Figure adapted from Castro-Sánchez and Roda-Navarro, 2017 (90).

1.5 Cellular interplay and crosstalks

The joint synovium of RA patients contains a variety of immune cell-types that are involved in the pathogenesis of the disease, including macrophages, Dendritic Cells (DCs) Natural Killer (NK) cells, T cells and B cells. In addition, some non-immune cells, like fibroblasts, and endothelial cells, can also participate in the development of RA (91). Therefore, RA can no longer be considered as the result of a distinct humoral or cellular autoimmune response to a single autoantigen. Rather, a new perspective on this disease is emerging, one that aims to understand it as the result of pathologic cell-cell interactions that take place inside a distinct and defined environment, the synovium (Figure 7) (4). These cell-cell interactions can be achieved in two general ways: first through secreted mediators, notably inflammatory cytokines such as TNF, IL-6, IL-17, and many others; and second through direct cell-cell contact that is mediated by cell surface receptors and ligands, including some membrane-anchored proteins (92). These mediators alter the activation and differentiation state of one or both cell types. They regulate inflammation, autoimmunity, and articular destruction in the joints by initiating cascades of signaling pathways further resulting in the expression of pro inflammatory molecules. Such cascades trigger disease phenotypes like angiogenesis, cartilage matrix degradation, inflammation, and synovial

hyperplasia (93).

The three most abundant cell populations in RA synovium are synovial macrophages, synovial fibroblasts, and infiltrating T lymphocytes, with a bias towards CD4+ T helper 1 (Th1) subtype in the latter population (4,94). These cell-types play a crucial role in the initiation and progression of RA (95–97). Furthermore, cell-cell interactions among these populations in the RA synovium not only define many aspects of the synovial biology of this disease, but also offer targets for therapeutic interventions (4). The following sections will focus on the pathogenic role of RA macrophages, RA fibroblast and RA Th1 as well as their intercellular interactions in the synovium.

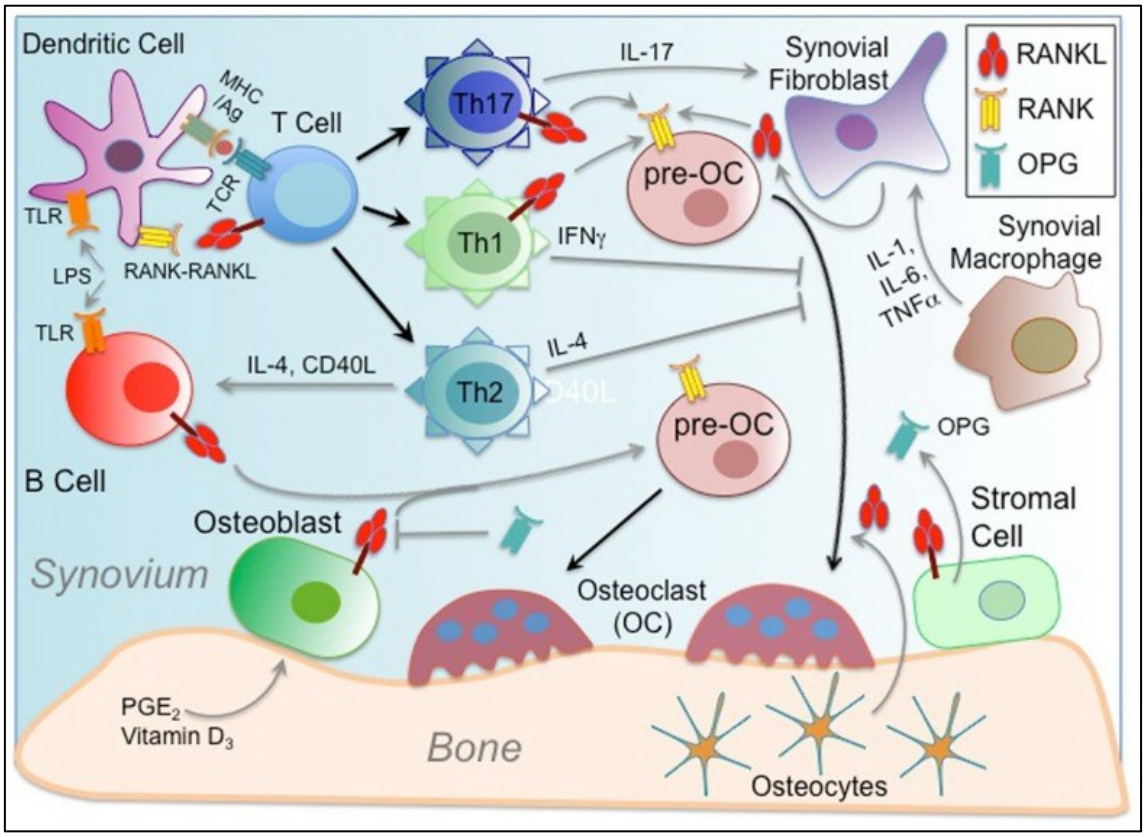


Figure 7. RA cellular interplay. The figure is taken from Walsh and Choi, 2014 (98).

1.1.4 Synovial fibroblasts

Synovial fibroblasts are the main stromal cells of the joint synovium. They are found in the synovial sublining and lining layers (99). In healthy joints, fibroblasts maintain synovial fluid and extracellular matrix homeostasis via the production of articular and synovial fluid components such as lubricin and hyaluronic acid (100). In RA, they are influenced by their environment and exhibit aggressive behavior (Figure 8). Indeed, RA synovial tissue promotes survival of fibroblast, inhibits their

deletion through apoptosis (101) and increases their invasiveness into the extracellular membrane. Aggressive RA fibroblasts play an important part in the development of the pannus by producing matrix metalloproteinases (MMPs), which degrade the cartilage, allowing pannus to expand and invade further (102). They have additional pro-inflammatory effects through production of cytokines such as Tumor Necrosis Factor alpha (TNFα), InterLeukin-6 (IL-6) and Granulocyte Macrophage Colony-Stimulating Factor (GM-CSF), chemotactic factors such as C-C Motif Chemokine Ligand 2 (CCL2) and IL-8 (CXCL8). Several pro-angiogenic factors are also expressed by fibroblasts, in particular Vascular Endothelial Growth Factor (VEGF) and TGF-β. These mediators sustain inflammation in RA and contribute to their interactions with other cell types (103).

The physical association of fibroblasts and macrophages is reflected by an intense cellular communication between both cell types. Fibroblasts have a dual effect on macrophages. They drive the pro-inflammatory differentiation of macrophages by an excessive production of cytokines such as IL-6 and CCL2 within the RA synovial microenvironment. On the other hand, fibroblasts produce GAS6, which in turn induces a pro-resolving and immune-regulatory phenotype in resident synovial macrophages that express the corresponding receptor (104). Fibroblasts are also the most relevant source of CSF1 and CSF2 within the microenvironment of macrophages, promoting the survival, proliferation, and maintenance of the local macrophage pool (105,106).

Fibroblasts interact intensively with T cells as well. Accumulating evidence indicates that bidirectional signaling between T cells and fibroblasts is robust and functionally significant (107,108). Fibroblasts promote T cell survival, chemotaxis, and activation via the secretion of cytokines such as type 1 IFN, IL-15 and IL-7 (109,110). In addition, Fibroblasts can promote the differentiation of proinflammatory T cells subtypes, including Th1, and inhibit the differentiation of anti-inflammatory subtypes of T cells in the RA synovial joints via IL-6, IL-26 and CCL20 secretion notably (111). Fibroblasts can also function as antigen-presenting cells (APCs) and are able to present peptides derived from autoantigens found within joint tissues to activated T cells (107).

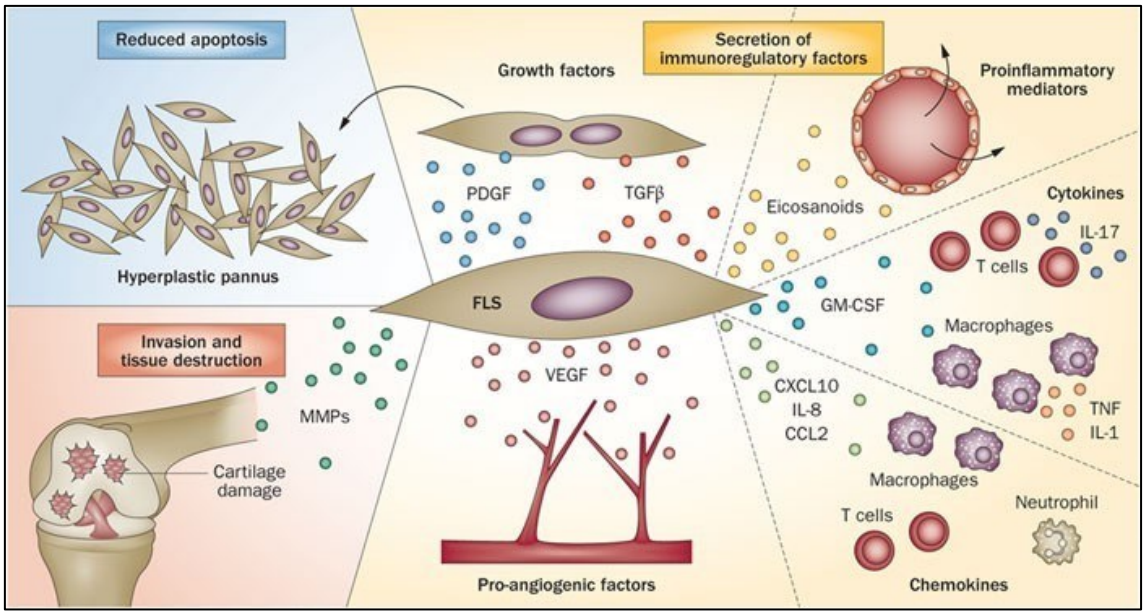


Figure 8. The role of fibroblasts in RA. Figure taken from Bottini and Firestein, 2013 (112).

1.1.5 Synovial macrophages

Tissue resident macrophages are distributed throughout the lining and sparsely in the sub-lining region of the healthy synovium and have been shown to play an essential role in normal tissue physiology, by providing the architecture of the joint, maintaining and reinstating synovial tissue homeostasis notably (113). In RA, macrophages are abundant in the inflammatory synovial membrane and at the cartilage-pannus junction. They are highly activated; the degree of macrophage activation correlates with the joint pain and general inflammatory status of the patient (114). RA macrophages act as APCs by presenting autoantigens through large amounts of major histocompatibility complex class II molecules. They secrete proinflammatory or regulatory cytokines (IL-1, IL-6, IL-10, IL-13), growth factors, chemokines (CXCL1 and CXCL10) to promote inflammation and angiogenesis. They also produce metalloproteinases, contributing to the destruction of the cartilage matrix (93,115).

Plasticity is a key feature of macrophages. They are capable of presenting heterogeneous phenotypes, creating various subpopulations. Depending on their activation state (M1 or M2), these innate immune system cells can have extensive proinflammatory and damaging abilities, contributing significantly to inflammation and joint deterioration. Or, they can have remodeling and inflammation resolution abilities. The M1 and M2 phenotypes represent the extremes of a macrophage activated spectrum, which is characterized by the presence of intermediate phenotypes involved in the immuno-regulation or in tissue repair and defined by

different signaling pathways, surface markers, and cytokine production (87). In RA, the proportion of M1 macrophages is higher than that of M2 macrophages. This increased pro-inflammatory ability of macrophages is related to their excessive activation and proliferation as well as their enhanced resistance to apoptosis (116).

Both M1 and M2 macrophages participate in the disease course by interacting with other cell types (Figure 9). M2 macrophages can inhibit the proliferation, activation, and cytokine production of CD4+ pro inflammatory T cells (117,118). They can also boost CD4+ regulatory T cells function by producing immunoregulatory mediators like IL-10 and TGF-β (119). Inversely, M1 macrophages can attract CD4+ T cells and induce their hyper activation by producing CXCL16 and IL-15 (120,121). The increased CXCL16 expression in RA macrophages promotes recruitment of CXCR6+ T cells and may thereby contribute to synovial inflammation and immunopathology (122). Furthermore, M1 macrophage-derived IL-6, IL-12 and IL-18 contribute to the differentiation of naïve CD4+ T cells into Th1 pro inflammatory subtype (123). These cytokines also participate in the inhibition of regulatory T cells (124).

Macrophages interact with fibroblasts as well. They secrete PDGF, TGF-β, IL-22 and HBEGF (heparin binding EGF-like growth factor) to promote the invasive behavior of resident synovial fibroblasts and the consequent destruction of the articular cartilage (105,125,126). Moreover, cellular communication between macrophages and fibroblasts amplifies the pro-fibrotic response and induce higher production of IL-11, LIF, CSF3, IL-33, Il-6 and growth factor receptors by fibroblasts (125,127).

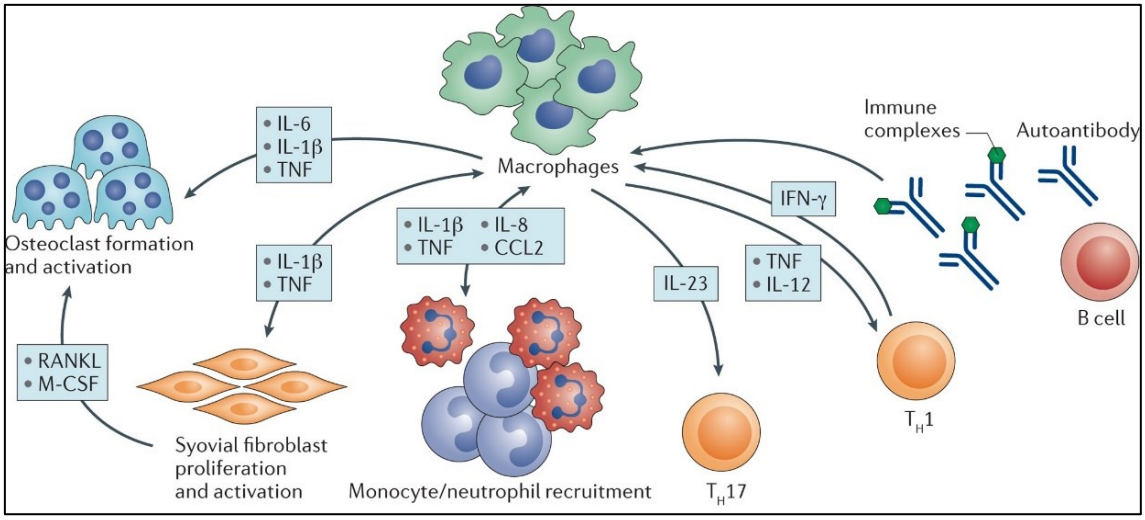


Figure 9. The role of macrophages in RA. Figure taken from Udalova et al., 2016 (128).

1.1.6 CD4+ T helper 1 (Th1)

Although the exact role of T cells in RA remains unclear, there is compelling evidence to support the significant contribution of CD4+ T-cells to the chronic autoimmune response of RA. Indeed, highly expanded CD4+ T cell clones are observed in RA synovium (129) and are characterized by increased apoptosis resistance and higher autophagic flux compared to healthy cells (130).

CD4+ T cells play a central role in regulating the immune response through the secretion of specific cytokines, the activation of innate and adaptive immune cells as well as nonimmune cells, and the suppression of immune reaction (131). They interact with Major Histocompatibility Class II (MHC-II) molecules, expressed on the surface of APCs (132). The key genetic association between the HLA-DR locus and RA further supports that CD4+ T cells are directly implicated in the disease (133). CD4+ T cells interaction with APCs results in CD4+ T cells differentiation into distinct subtypes: Th1, Th2, Th9, Th17, Th22, Treg (regulatory T cells), and Tfh (follicular helper T cells) (134).

RA is considered a Th1-associated disease. Indeed, Th1 is the most abundant CD4+ T cells subtype in synovial fluid of RA patients (135,136). Th1 cells promote the development of a proinflammatory microenvironment in the synovium by inducing the secretion of proinflammatory cytokines, interferon-g (IFN-g), Tumor Necrosis Factor-a (TNF-a), LymphoToxin (LT), and IL-2 (137), leading to synovitis, cartilage destruction and bone erosion (138). The high level of Th1-derived TNF-a is responsible for dendritic cell differentiation, in charge of autoantigen presentation to T cells in the synovium of RA patients (139). It also leads to osteoclastogenesis and bone resorption by activating osteoclast precursors and inhibiting osteoblast differentiation (140). CXCR3 was identified as a surface marker for Th1 cells (141). CXCR3 is expressed in 40% of autoreactive CD4+ T cells in RA synovium (142) and binds to CXCL9 and CXCL10, two chemokines highly expressed in RA synovium (143). Furthermore, T-bet was identified as a master transcription factor in Th1 (144), and is associated with several inflammatory diseases, including RA (145–147).

Th1 cell-type interacts intensively with macrophages and fibroblasts. It contributes to the M1 polarization of macrophages via TNF-a, IFN-g and GM-CSF production (148,149), but also via direct contact through T Cell Receptor (TCR)/CD3 molecules (150). The regulation of macrophages by Th1 cells in RA also reflects in their ability to regulate the differentiation of macrophages to osteoclasts through the expression of RANKL (151). Th1 cell-type interacts with fibroblasts as well and has a strong ability to increase their proliferation, adhesion, invasiveness. Furthermore, Th1-associated cytokines, IFN-g and TNF-a are known to increase the expression of InterCellular Adhesion Molecule 1 (ICAM1) and Vascular Cell Adhesion

Molecule 1 (VCAM1) in fibroblasts, subsequently leading to production of MMPs and inflammatory cytokines including VEGF, IL-6 and IL-8 (152,153). Therefore, these interactions participate in initiating and sustaining pro-inflammatory processes and joint destruction in RA (152).

1.6 Treatments

There is currently no cure for RA. The broad range of molecules associated with RA, the interplay between genetics and environment in RA initiation, the disease heterogeneity, the limited ability to recognize the population at risk, and the inability to make an early diagnosis are all obstacles to curing RA. The proposed therapies merely seek to alleviate disease symptoms, relieve pain, increase survival, and prevent long-term disability (5). Nonsteroidal anti-inflammatory drugs (NSAIDs), including aspirin and ibuprofen, are used to help manage chronic pain, inflammation, and swelling in RA but do not slow down the disease progression and can have adverse effects after prolonged use (154). Glucocorticoids are proposed as well to decrease disease activity but can only be used for a limited period of time and in combination with Disease-Modifying Antirheumatic Drugs (DMARDs) (155).

DMARDs are commonly used treatments in RA. DMARDs target inflammation and thereby prevent further joint damage (6). They are categorized into conventional DMARDs (methotrexate, hydroxychloroquine, and sulfadiazine) or biologic DMARDs (TNF-α inhibitors (anti-TNF mAb), IL-6 inhibitors (anti-IL-6 mAb), B cell depleting antibodies, and inhibitors of co-stimulatory molecules). Each DMARD has a unique mechanism of action, ultimately interfering with critical pathways in the inflammatory cascade. Conventional DMARDs include drugs that target the entire immune system whereas biologic DMARDs are monoclonal antibodies and soluble receptors that target protein messenger molecules or cells (156) (Figure 10).

Both conventional and biologic DMARDs have several side effects. Conventional ones can cause bone marrow suppression, gastrointestinal distress, allergic reactions, hepatotoxicity, liver cirrhosis and many other adverse effects (6). Even though methotrexate monotherapy results in clinical symptoms and joint damage reduction in about 25-40% of patients, which increases to almost 50% with glucocorticoid addition (157), around 10-30% of patients discontinue therapy within a year due to the previously described adverse effects (158). Biologic DMARDs, on the other hand, can cause increased risk of bacterial, fungal, and viral infections, tuberculosis, herpes zoster, hepatitis B/C, skin cancers, congestive heart failure and drug-induced lupus (156).

Although several alternative mechanisms of action therapies are approved as first-line options, approximately 90% of biologic-naive RA patients receive an anti-TNF as their first biologic treatment (159). Furthermore, a substantial proportion of RA patients (30-40%) do not respond to this treatment. Some of these non-responders never achieve a response (primary inefficacy), others may respond and then lose the treatment efficacy over time (secondary inefficacy) (160). Even after multiple DMARD therapy switching, about 20%–30% of these patients remain treatment refractory (7). In addition, despite diverse targets, all biologics have comparable efficacy when used with methotrexate (ACR70 response rates about 35-40%), which diminishes as previous drug exposure increases (157).

These limitations emphasize the urgent need for novel therapeutic options, as well as a better understanding of the mechanism of action of current RA treatments. A deeper comprehension of the cellular interaction and the intracellular cascades involved in RA pathogenesis, as well as their regulation, could help overcome this challenge.

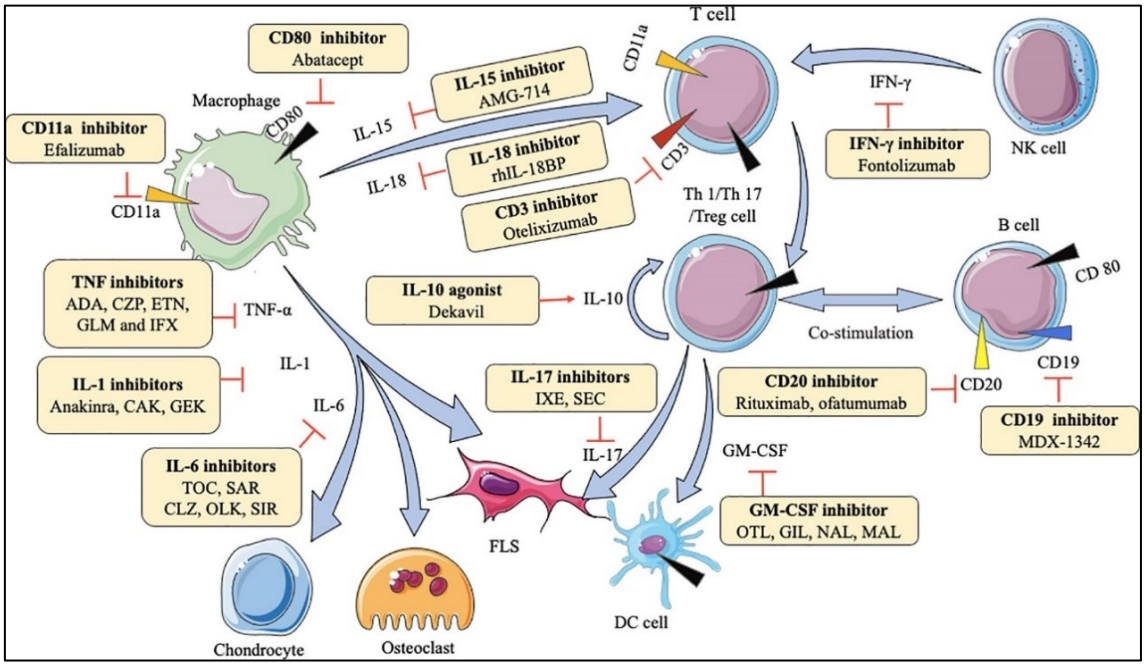


Figure 10. Action of biologic DMARDs in rheumatoid arthritis. Figure taken from Huang et al., 2021 (161).

Chapter 2. Systems Biology to better understand complex diseases

Complex disease mechanisms involve several interconnected biological processes and signaling pathways that result in specific gene expression patterns and cellular phenotypes. Each molecule can rarely be assigned a distinct biological role, as in living organisms nothing acts in isolation (162). Rather, cellular functions and phenotypes arise from the interactions between the subcellular components of a biological system (13). This is further underlined by ongoing developments in high-throughput experimental techniques and the increase of the processing compute power, which provide tremendous amounts of data regarding these molecular interactions and reinforce the importance of considering biological systems as a whole. Therefore, there is an urgent need for a system-level understanding of the structure and dynamics of these intricate interactions and for the development of new methodologies to gain insights through more holistic representations (163).

Systems biology is defined as a field of study that uses a holistic approach combining mathematical modeling with experimental methodologies and bioinformatics tools to better understand the mechanics of complex biological systems and predict their behavior across multiple scales going from cells, tissues to organisms (8). One strategy to understand the complexity of biological systems is their abstraction to complex, multi-layered networks of interactions. Several levels can be identified, depending on the spatial scale of interest. At the cellular level, the networks are composed of subcellular components like proteins and genes, connected by edges, with each edge reflecting the biochemical interactions between these components (162). At the tissue level, the networks are composed of interactions between different cell types and of the cells with their supporting stroma. The networks at the organismic level are comprised of interactions between different organs or body systems (164).

2.1 Molecular interaction maps

To describe disease mechanisms and better understand their biological processes at the cellular level, efforts have already been made to assemble the available and fragmented knowledge in molecular interaction maps. A disease-specific map is a detailed, knowledge-based network that describes disease mechanisms and disease-related signaling, metabolic, and gene regulatory pathways. It also provides evidence connecting these pathways to the disease's causes and effects (9). Disease maps are manually curated and extensively annotated through broad literature search and an active involvement of domain experts, making them a

more reliable and solid source of information compared to automatically inferred biological networks. Indeed, such networks are often noisy, and rely on putative, not manually curated interactions. They are associated with high uncertainty and likely riddled with false positive and false negative interactions. Furthermore, they do not provide mechanistic details regarding the diseases processes (165).

State-of-art mechanistic molecular interaction maps have already been published for several diseases including Parkinson's (166), Alzheimer's (167), rheumatoid arthritis (168), cancer (169), asthma (170), and more recently COVID19 (171). They consist of hundreds of nodes and interactions and illustrate manually curated molecular mechanisms and signaling pathways in a disease-specific manner. These maps can be used for identification of disease biomarkers, potential drug targets and disease subtypes for better diagnosis and stratification of patients. Furthermore, they provide a comprehensive template for visualization, analysis, and interpretation of omics datasets. They can be analyzed in terms of the underlying network structure as well.

Disease maps are both human and machine-readable. Visual exploration of the maps enables clinical and life sciences researchers to easily investigate complex and interconnected diseases pathways, while machine-readable formats of the underlying information create an interface to a broad range of bioinformatic workflows (172). Therefore, maps play a critical role in bridging bioinformatics and molecular biology.

2.1.1 Community developed standards for knowledge representation

To harmonize biological networks and cellular processes representation, Systems Biology Graphical Notation (SBGN) standard has been developed (10). It allows for an efficient and accurate representation, visualization, storage, exchange, and reuse of information. Information in SBGN is represented using graphical objects (glyphs) structured in a graph, and the regulated Systems Biology Ontology language is used. SBGN consists of three complementary languages, covering different granularity of biological processes (Figure 11): Process Description (PD), Entity Relationship (ER) and Activity Flow (AF). PD is a directed, sequential representation of the mechanistic and temporal details of the underlying process. It illustrates how entities change form as a response to various influences (173). ER diagrams emphasize the effects that entities have on one another's changes rather than the transformations themselves. AF diagrams permit modulatory arcs to directly link different activities, rather than entities and processes or relationships. It is often used for coping with biochemical network complexity or with incomplete or indirect knowledge (10).

Standards regarding the manual curation of biological networks can also help harmonizing the content in various disease maps. In this direction, Minimal Information Requested In the Annotation of biochemical Models (MIRIAM) has been developed as a set of guidelines for the proper annotation and curation processes. MIRIAM is composed of two parts. The first is a qualifier describing the relationship between a model component and the resource used to annotate it, whereas the second is a proposed annotation scheme that specifies the documentation of the model by external knowledge (11). Annotations using MIRIAM are in a machine-readable format which facilitates their import and exchange and that can be unambiguously parsed by software to perform simulations and analysis.

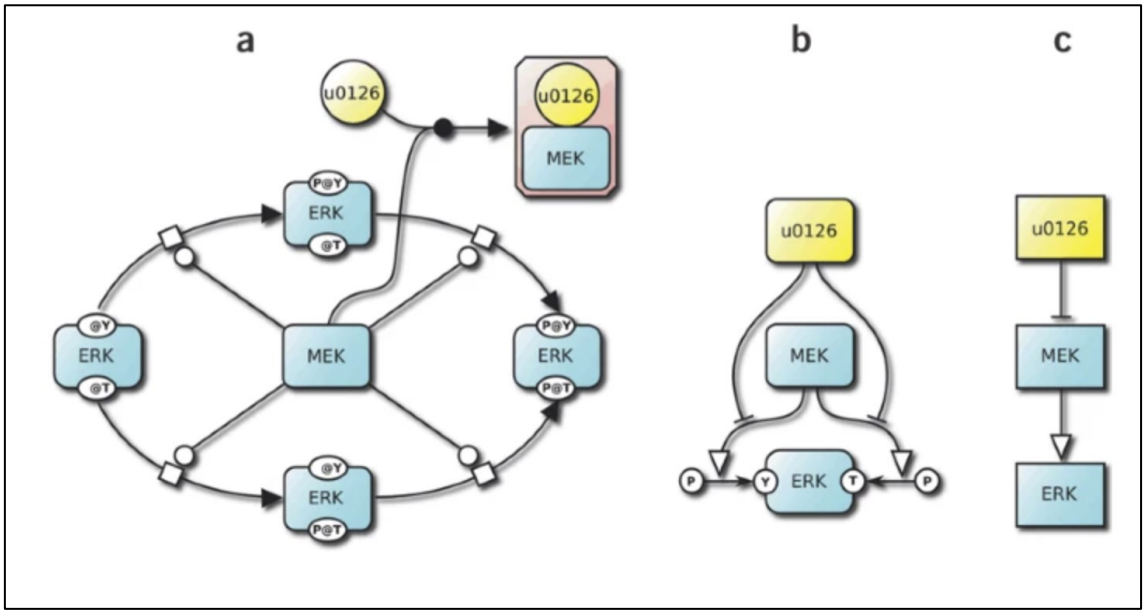


Figure 11. Three different types of SBGN networks are used to represent protein phosphorylation catalyzed by an enzyme and modulated by an inhibitor. (a) Process description. (b) Entity Relationship. (c) Activity Flow. Figure adapted from Le Novère et al., 2009 (10).

2.1.2 From a cellular to a multicellular level

Multiscale knowledge management is at the heart of the disease map concept. This means developing and exploiting protocols for the high-quality representation of information at different levels of granularity including subcellular, cellular, tissue, organ, and organism levels (9). Changes in communication between cells and their surrounding microenvironment are a feature of numerous human diseases including RA. Signaling crosstalk via soluble and membrane-bound factors is critical for informing diverse cellular decisions, including decisions to activate cell cycle or programmed cell death, undergo

migration or differentiate along the lineage (174). However, these intercellular interactions are missing from most disease maps that were built at the cellular level. Therefore, an accurate representation of cell-cell signaling links and effective systems-level analyses of those links is needed for understanding global communications among cells and their contribution to diseases pathogenesis.

Biochemical tests including yeast two-hybrid screening, co-immunoprecipitation, proximity labelling proteomics, and X-ray crystallography have been traditionally used to identify interactions between proteins that are secreted or displayed extracellularly to mediate intercellular communication (175). Such information can be retrieved through extensive literature search and curation. Another approach is to make use of RNA-seq, single cell RNA-seq or spatial transcriptomic technologies to dissect cell-cell communication. Many tools have been developed for this purpose (174,176–180) . They consist of a resource of intercellular interactions prior knowledge and a method to predict potential cell-cell communication events. Events are represented as a one-to-one interaction between a transmitter and receiver protein expressed by the source and target cell type respectively. Events are predicted by estimating the likelihood of crosstalk based on the expression level of the transmitter and the receiver proteins (181). Such interactions may be incorporated into disease maps to describe communication between different cell-types involved in the diseases' pathogenesis.

Disease maps, whether they are built at the cellular or multicellular level, represent dense knowledge bases and comprehensive representations of complex biological systems. Although maps remain limited by their static nature when it comes to predictions and hypotheses testing, they can be used as the starting point for different types of mathematical modelling and in-silico experiments.

2.2 Computational modeling in biology

One of the primary goals of dynamical modeling is to understand the emergent features and behaviors of complex biological systems. It seeks to explain how cellular functions such as mobility or proliferation are caused by interactions between subcellular components such as proteins and small molecules (13). The purpose of a model is to imitate the behavior of the system it mimics, based on actual, known attributes of the system components. To achieve this purpose, the model may need to span different length scales and contain data from various fields of research (182). Developing a model entails formulation of a conceptual model, constructing the model by translating the conceptual model to a list of components, selecting mathematical expressions to characterize the relationships between its components, determining parameter values, and performing simulations and other mathematical analyses to reproduce observations or lead

to predictions. Models' attributes can be modified as well as the rules that control their interactions, allowing researchers to test hypotheses, determine which features caused specific outcomes, and conduct nearly any type of virtual experiment (183).

To account for the different quality of information that is available for a network under study, modeling formalisms of different levels of complexity have been developed over the last years.

2.2.1 Quantitave models in biology

The quantitative modeling approach is based on differential equations. It describes the dynamics of signaling pathways by translating them into rate equations. They can be based on ordinary differential equations (ODEs), Partial Differential Equations (PDEs) or on probabilistic stochastic equations. The most widely used ones are sets of coupled ordinary differential equations (ODEs) that describe the system's development over time using mass-action kinetics for the rates of production and consumption of the biomolecular species (184). These types of models require shorter simulation time, and can be parametrised using experimental lab data or clinical patients data because the large majority of collected data describes the temporal changes in some variable of interest; e.g., levels of pro-inflammatory cytokines. These models can also be investigated analytically via the identification of possible steady states and their stability. Several ODEs models have developed to better understand several diseases like Alzheimer's (185), cancer (186–188) and Parkison's (189).

In RA, several ODEs models have been published to describe joint erosion, cytokines interactions, interactions of immune cells within the RA environment, or the circadian dynamics involved in the progression of RA (190–193). Even if PDEs are less commonly used to describe RA specific processes in comparison to ODEs, a model describing the spatio-temporal interactions between immune cells, cytokines and drugs in RA was also published (194).

ODEs and PDEs being deterministic models, stochastic mathematical models for RA have been developed. Such models have been mainly applied in the context of treatment decisions, to analyse RA incidence rates, or to predict radiological progression within RA (195–197).

Quantitative models are more precise and specific about a system but require a large effort in model construction as it demands knowledge of kinetic information for each interaction, as well as a high number of parameters. Because many of these characteristics are unknown and difficult to determine in most biological processes, these models are limited to small and well-characterized networks (an

ODE model is considered as large if it has more than 40 parameters to be inferred) and cannot be used to describe large-scale systems (12,13).

2.2.2 Agent-based models in biology

Individual-based (IB) (or agent-based) modelling approaches have been used to incorporate stochasticity at a cellular level within mathematical models (198). This approach allows each cell to be described as an individual agent which follows a set of predetermined rules. These agents reside in a defined spatial domain called lattice. One strength of IB models is their ability to model heterogeneous populations. Another advantage is that the researcher does not need to have an understanding of the aggregate or big picture behavior of the phenomenon. These behaviors emerge from the rules that are developed by the researcher. However, such models can be computationally expensive especially when modelling a large number of agents. Indeed, Each agent is moving about and demanding calculations from the Central Processing Unit (CPU). When this number gets too large, the CPU slows down. This expense can be reduced by using a hybrid modelling approach that combines IB models with classical methods such as differential equations. These hybrid approaches have been used to model various biological phenomena and diseases (199–201). However, such models remain limited in term of size as they require a huge number of parameters that are often difficult to find in the literature. Adding only one parameter in the model can substantially increase the amount of data required to constrain it (202). Even simple models tend to contain a considerable number of parameters that grows linearly with the number of agents in the model and results in an exponential expansion of the parameter space. Moreover, the runtime of an IB model does not only depend on how many agents are in the model, but also on how complex their decision-making rules are, and how often they interact with each other and with their environment (203).

2.2.3 Qualitative models in biology

Qualitative modeling approaches are more suitable for characterizing large-scale systems for which the majority of kinetic parameters is unavailable. They do not require kinetic parameters and can provide a qualitative dynamic description of the system. Furthermore, experimental literature and high-throughput technologies can be used to provide qualitative data on individual components and interactions at the molecular level. In these models, variables can have a finite number of states. A variable's state is defined by a logical combination of the states of other variables. These models are scalable and suitable for systems containing hundreds of components. They are increasingly being used to model biological networks (204–206).

The class of qualitative modeling approaches comprises various formalisms of different complexity. More refined techniques include Petri nets (207) and logical modeling (208–210). As graph models, these frameworks solely rely on the network structure, yet they enable the analysis of important functional properties of large-scale signal transduction networks such as input–output relationships and feedback loops, and they also allow certain predictions (184).

The application of Petri nets to biological systems was first proposed by Reddy et al. (211). Some of the main advantages of Petri nets are that they are visual, have different flavors and can be designed and analyzed by a range of tools (212). While the Petri net approach is particularly suited for modeling mass flows as they arise in metabolic networks, the description of signal or information flows as characteristics of gene regulatory and signal transduction networks is less straightforward.

In contrast, the logical modeling approaches have directly been introduced as qualitative descriptions of signaling and regulatory networks. Logic-based network models were pioneered in the biomedical sciences by Kauffman (213,214), followed by Thomas (215) and represent a compromise between structural analysis and ODE methods in terms of precision and complexity.

Several Logic-based models have been developed to study complex diseases such as COVID-19 (216), cancer (16,217,218), diabetes (219) and RA (220,221). In their simplest form, logic-based models permit each biochemical species to be in one of two discrete states: ON or OFF. The state of a logic network evolves in a dynamic fashion as nodes in the network are switched ON and OFF according to the state of other nodes in the network (222). More complex logic-based methods have been developed, such as multi-state and fuzzy logic methods, which permit nodes to be in more than two discrete states (210,223). Although theoretically able to more precisely simulate biochemical regulation, these more complicated approaches require parameter value estimates that are rarely known and, in some cases, are difficult to correlate with biophysical chemistry theory. Thus, discrete two-state logic models (Boolean models) are an intuitive and predictive method for describing biochemical interactions without requiring prior knowledge of complex mechanistic details of reaction kinetics (needed for ODE systems) or degrees of membership (needed for multi-state fuzzy logic systems).

Based on the advantages and disadvantages of all the previously described modelling approaches, we can say that Boolean models are the most suited option for this PhD project as they can handle large-scale systems and the regulation of these model's components is given in a parameter-free way, without the need for kinetic parameters and precise quantitative data.

2.2.4 Boolean models

Boolean models are the simplest qualitative models describing an abstraction of the biological system where each biomolecule can have two values: zero or one. This indicates if a gene is expressed or not expressed, a transcription factor is active or inactive, and a molecule's concentration is above or below a certain threshold. Changes in the biomolecules' values depend on their interactions in the system and are defined by logical rules using the Boolean operators "AND", "OR" and "NOT" (Figure 12). The regulation of this state variable is given in a parameter-free way, making Boolean modeling a viable option for large-scale systems with unknown kinetic parameters and precise quantitative data (14,224). Several examples from models of genetic regulatory networks show that Boolean approaches give meaningful biological information (16,220,225,226).

The state of a simulated Boolean model changes at each execution of its logical rules. These states can be visualized at different timeframes using state transition tables and state transition graphs where each node corresponds to one state of the network, while edges represent transitions from one state to its successor (Figure 12).

The order of executions is defined by the chosen updating scheme (Figure 13). In the synchronous update, all functions are applied simultaneously at each execution. The dynamics of the models are deterministic, and each state has one successor. Therefore, this approach does not consider the stochastic nature of biological processes.

In the asynchronous update approach, only one random component is selected and modified at each time step. This stochastic update process results in several successor states for each state, depending on the selected component (227). Even though asynchronous models are usually considered as a closer representations of the biological processes, the run time for asynchronous simulations can become a strong limitation when analyzing large-scale Boolean models (228).

Many tools to perform simulations and analysis of logical models already exist, including ADAM (229), BoolNet (230) , BooleanNet (231) , Cell Collective (232), CellNetAnalyzer (233), GINsim (234) , and BioModel Analyzer (235).

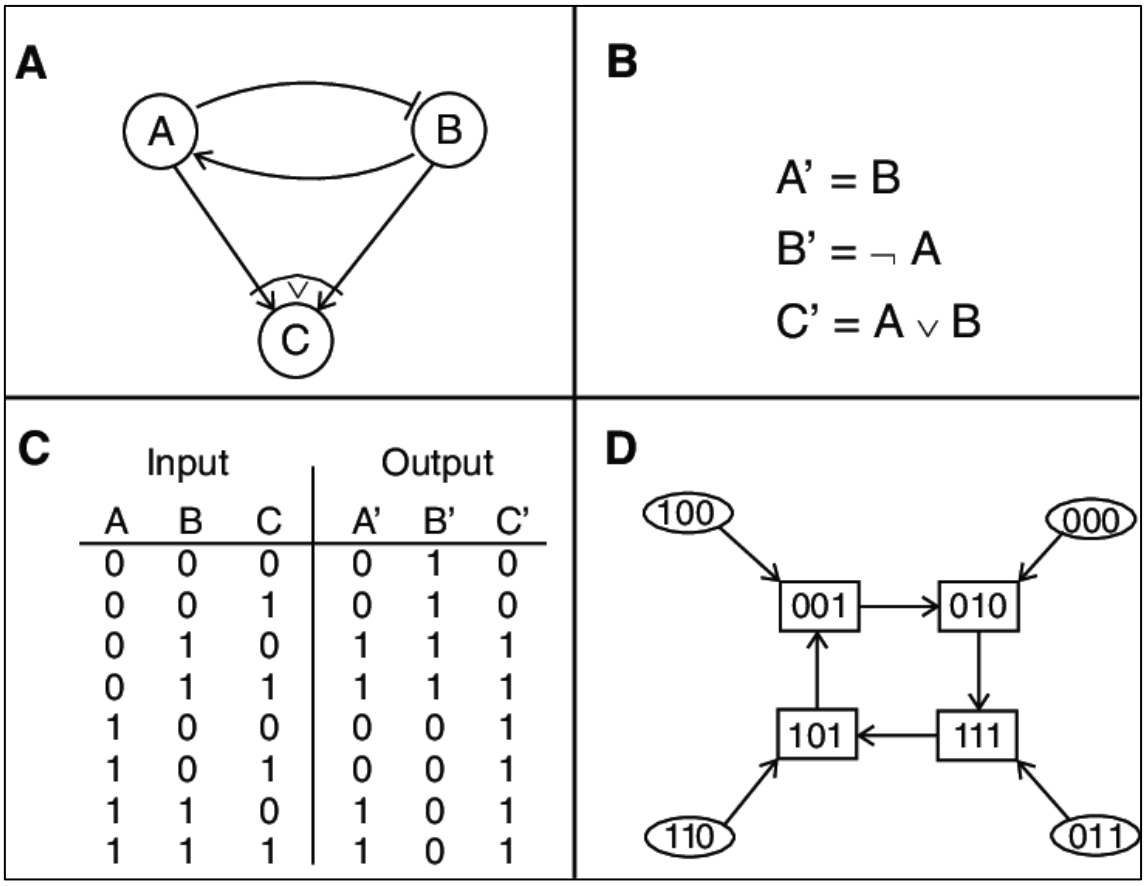


Figure 12. Different representations of a sample Boolean network consisting of three nodes. (A) Graph representation, (B) logical Boolean rules, (C) state transition table and (D) state transition graph. Figure taken from Kaderali and Radde, 2008. (236)

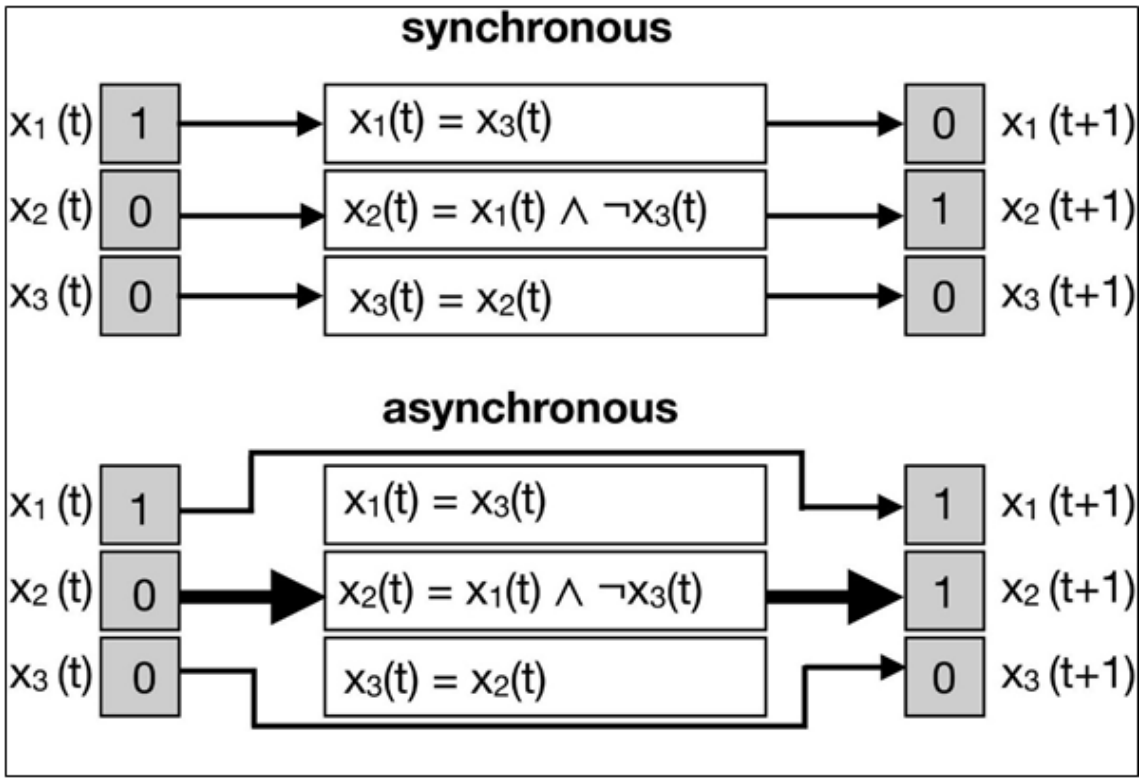


Figure 13. Updating schemes in Boolean models. Figure adapted from Schwab et al., 2020 (227).

2.2.5 Long-term behavior of Boolean models

When simulated, the biomolecules states may converge to periodic sequences of states, called attractors. Once reached, they cannot be left unless an external perturbation occurs. Attractors represent the model's long-term behavior and have been connected to biological phenotypes, making their computation a key point in Boolean models' analysis (14). Attractors can be steady states comprising only one state. These attractors occur in both synchronous and asynchronous updating schemes. Attractors can also have more than one state. In synchronous models, they can be simple cycles, a fixed sequence of states that are periodically replicated. In asynchronous models, complex attractors may be reached (Figure 14). Complex attractors are formed by overlapping loops which origin from the possibility of reaching more than one successor state in the asynchronous update scheme (237).

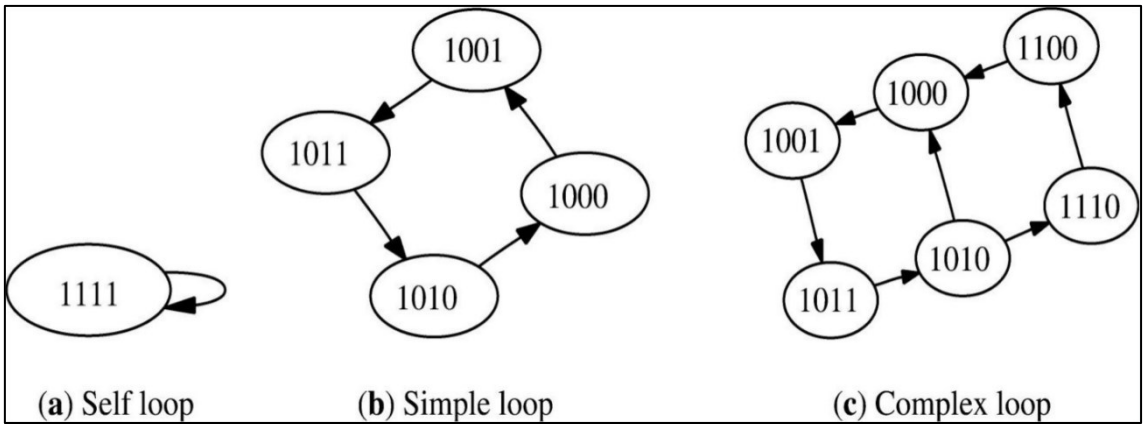


Figure 14. Possible types of attractors in Boolean models. Figure taken from Garg et al., 2008 (228).

It is still difficult to construct and analyze Boolean models for large-scale biological systems. When Boolean models are constructed manually, researchers typically focus on a small number of nodes; hence, the resulting models fail to reflect the whole range of a biological system. When models are inferred from maps, they are much larger and are more accurate representations of the system of interest (238). These large-scale models, on the other hand, are more complex, with a far higher number of nodes. Therefore, computing and enumerating all attractors in such complex models are still challenging tasks (15). Furthermore, it is difficult to compose larger models from smaller building blocks using Boolean networks. Hierarchical structuring, which makes the design and analysis of models simpler, is not possible in Boolean networks (212).

One of the most basic approaches to identify attractors is to enumerate state transitions with all potential states as initial conditions and see which attractor each initial condition eventually reaches. This work can be accomplished by creating and evaluating the corresponding state transition graph of the Boolean model. Given that the size of a Boolean model's state space is proportional to the number of nodes (2^n states for n nodes), if n is large, constructing and analyzing the state transition graph is impossible (15). One approach to overcome this limitation is the use of efficient constraint-solving algorithms, like SAT, without explicitly calculating transitions of the model (15,239). Some of these logical modelling tools also allow properties of the state transition graphs to be verified by means of existing model-checking tools, such as NuSMV (240–243). The properties are formulated in terms of temporal logic or in a suitable high-level query template capturing recurrent biological questions. The model checker tests if the state transition graph, which may be explicitly generated or implicitly encoded in a symbolic description of the model, satisfies the property. Another approach is

to perform model reduction or model decomposition to allow for in-depth dynamic analysis while preserving the important dynamical properties and topological features of the network (244,245).

Although these approaches enable the attractors' search in large-scale Boolean models, the resulting attractors still need to be validated as biologically consistent states in a cell- or disease-specific manner.

Chapter 3. Description of data and data analysis

In this section we describe each omics dataset we used during this PhD work for building the RA cell-specific maps, identifying the intercellular interactions between fibroblasts, macrophages and Th1 cell-types, and calibrating the cell-specific and multicellular models. We also discuss the statistical tools and the analysis we performed on the selected datasets and their corresponding outcomes.

3.1 GSE97779 dataset

GSE97779 dataset is a publicly available microarray dataset from the GEO database (246). The dataset contains nine RA synovial macrophages samples from nine patients and five peripheral blood monocyte-derived macrophages samples from five healthy donors. In the original study, researchers aimed to test whether a gene expression signature corresponding to IFN-g-mediated suppression of transcription is present in the RA synovial macrophages (247). They performed Gene Set Enrichment Analysis (GSEA) and Differential Expression Analysis (DEA) to compare RA synovial and control macrophages. They found out that the expression of IFN-g-induced genes was increased in RA synovial macrophages, whereas the expression of IFN-g-repressed genes was decreased in RA synovial macrophages relative to control macrophages. They also demonstrated that the pathophysiological relevance of IFN-g in RA is mediated by the suppression of MAF and downstream gene expression.

In this work, we quantile normalized the gene expression using the preprocessCore package (248). We annotated the genes using BioMart package (249). Then, we performed DEA using the Limma package (250) to identify the Differentially Expressed Genes (DEGs) between RA and healthy macrophage samples. We filtered the DEGs using a False Discovery Rate (FDR) adjusted p-value threshold equal to 0,05 and an absolute fold change threshold equal to 1,5. 5430 statistically significant DEGs were identified and used in the next steps of this work.

3.2 GSE164498 dataset

GSE164498 dataset is a publicly available RNA-seq Single-Cell dataset from the GEO database. It contains 1766 HL-60-derived M1 macrophage cells and 2063 HL-60-derived M2 macrophage cells. M1 macrophages were polarized using LPS, and IFN-g and M2 macrophages were polarized using IL-4 and IL-13. The objective of the original study was to uncover the gene regulatory networks underlying macrophage polarization through a comparative analysis of bulk, single-cell and ATAC-seq data (251). Researchers constructed *de novo* gene

regulatory networks and revealed subtype-specific transcription factor interactions and distinct activation trajectories during the macrophage polarization process.

In this work, we used Bioturing software (252) with the Venice method (253) to identify M1 and M2 macrophage marker genes. Venice method is a non-parametric approach that identifies the marker genes of a cell population. It searches for genes that differentiate the two selected groups by constructing a classifier that tries to identify the population of each cell given the expression level of a specific gene. The classifier's accuracy and its p-value are then computed and used as metrics to quantify the "marker quality" of the gene. Using the Venice method and an FDR cut-off of 0.05, 2306 genes were identified as marker genes. Since an M1 versus M2 macrophage comparison was performed, up-regulated marker genes were assigned to the M1 macrophage gene signature, and down-regulated marker genes were assigned to the M2 macrophage gene signature.

3.3 GSE109449 dataset

GSE109449 is a single cell (sc) RNA-seq dataset available in the GEO database (246). It contains 384 freshly isolated synovial fibroblasts in two RA and two OsteoArthritis (OA) patients. The objective of the original study was to identify functional and transcriptional differences between targeted subpopulations of fibroblast in the RA synovium (254). Researchers performed DEA and GSEA on this dataset, in combination with microarray and bulk RNA-seq data and revealed that a distinct subset of PDPN+CD34–THY1+ fibroblasts is expanded in RA and may be pathogenic. They also demonstrated that this subset is enriched around blood vessels in RA synovium, and its expression profile reveals potential pathogenic role in matrix invasion, immune cell recruitment, and osteoclastogenesis.

In this work, we performed DEA between RA and OA cells using the BioTuring software and the Venice method (Figure 15). We filtered the DEGs using an FDR adjusted p-value threshold equal to 0,05 and an absolute fold change threshold equal to 1,5. 1998 statistically significant DEGs were identified.

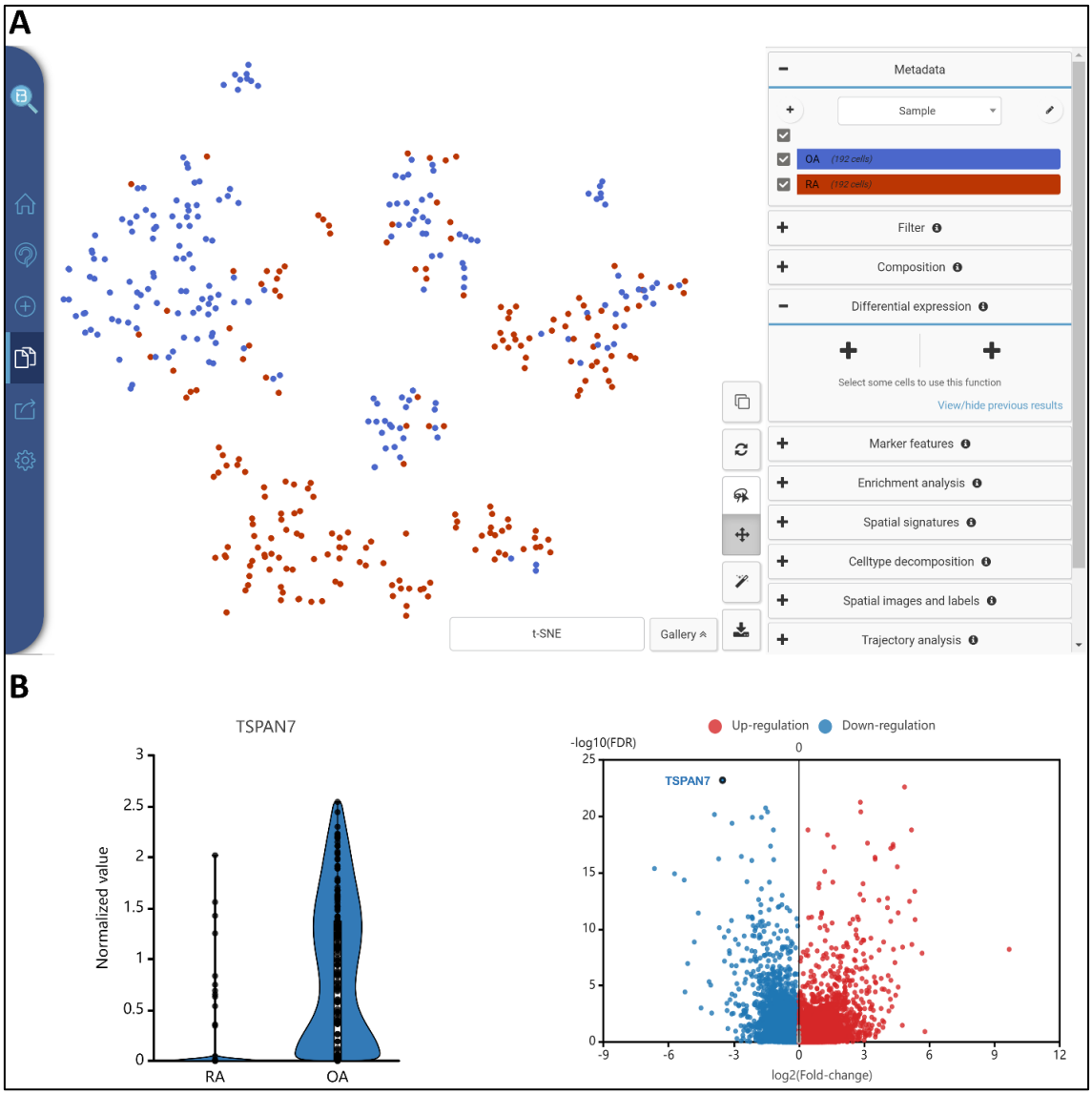


Figure 15. A) GSE109449 dataset visualization using BBrowser® in Bioturing tool. B) DEA outcomes using the Venice method performed on the GSE109449 dataset.

3.4 SDY998 dataset

SDY998 is a single cell RNA-seq dataset from the Immport database (255) containing 19 samples from RA patients and two synovial samples from OA patients including four cell types: 1142 B cells, 1844 fibroblasts, 750 monocytes, and 1529 T cells. In the initial study, researchers aimed to define the cell populations that drive joint inflammation in RA (256). They analyzed 5,265 scRNA-seq cells and identified unique cell populations that are expanded in RA synovia. They also mapped inflammatory mediators to their source cell populations to identify the subsets that are potentially key mediators of RA pathogenesis.

In this project, we used Seurat package (257) to perform dimensionality reduction on the CD4+ T cells via the *Find_neighbors* and *Find_clusters* functions. Ten CD4+ T cell clusters were identified. We used ABIS shiny app to select the relevant Th1 subtype markers (258) (Figure 16). Based on the expression of the marker genes CXCR3, CCR7 and CCR6, the Th1 cluster was identified. Figure 17A illustrates how we used the expression of CCR7 marker to identify CD4+ naïve T cell clusters. Figure 17B shows the expressions of both CXCR3 and CCR6 markers among the ten CD4+ T cell clusters. We observe a high expression of CXCR3 in cluster 3, 5 and 9. Cluster 5 is the only one not expressing CCR6 marker at all (CCR6 is specific to Th17 subtype).

We also performed DEA using CXCR3+ CCR7- CCR6- Th1 cluster versus CCR7+ naive CD4+ T cells from RA samples using Seurat and *Find_marker_genes* function. We filtered the DEGs using an FDR adjusted p-value threshold equal to 0,05 and an absolute fold change threshold equal to 1,5. 102 statistically significant DEGs were identified.

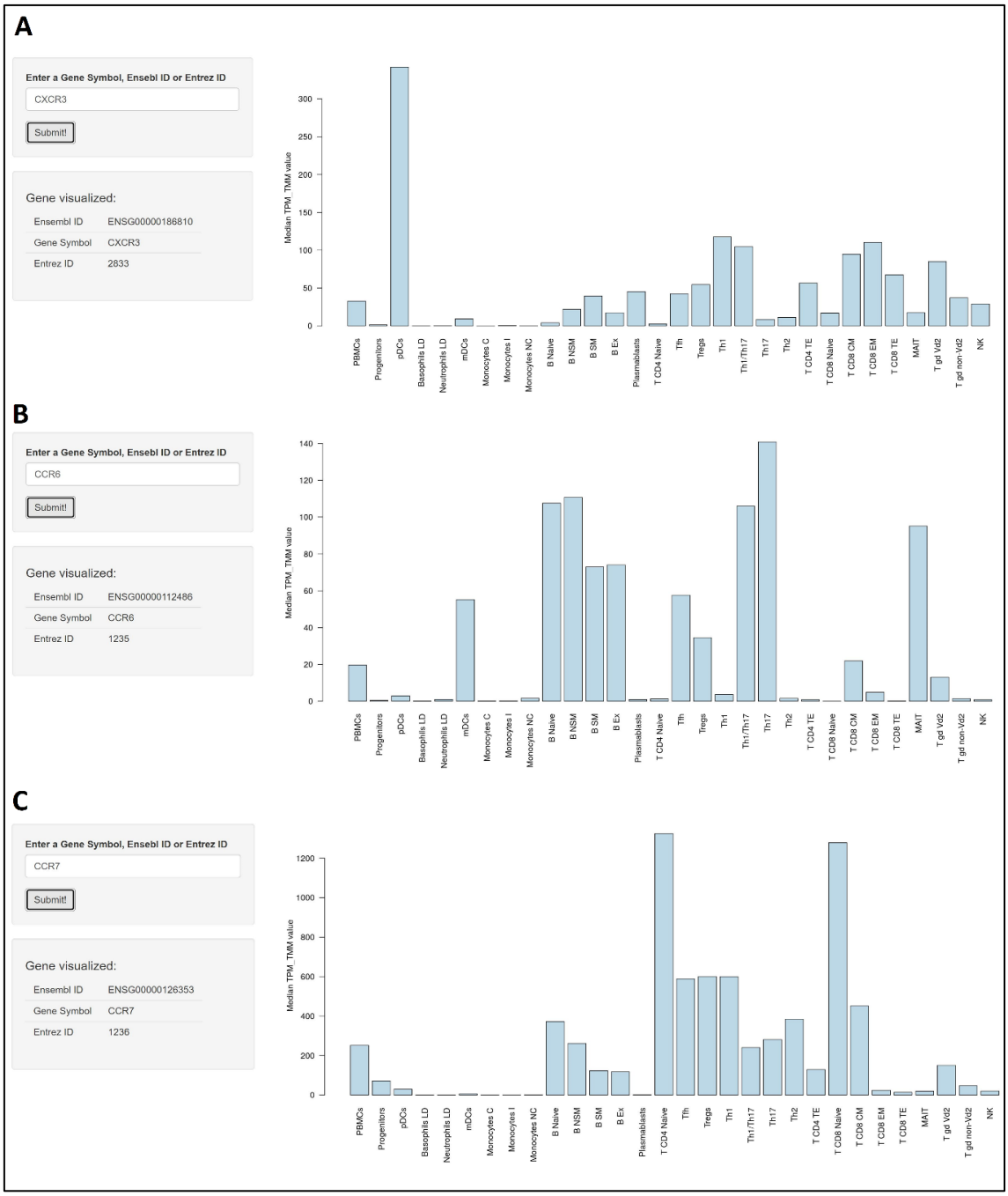


Figure 16. Screenshot of ABIS Shiny application https://giannimonaco.shinyapps.io/ABIS/. A) shows the expression of CXCR3 within the peripheral blood mononuclear cell (PBMC). B) shows the expression of CCR7 within the peripheral blood mononuclear cell (PBMC). C) shows the expression of CCR6 within the peripheral blood mononuclear cell (PBMC).

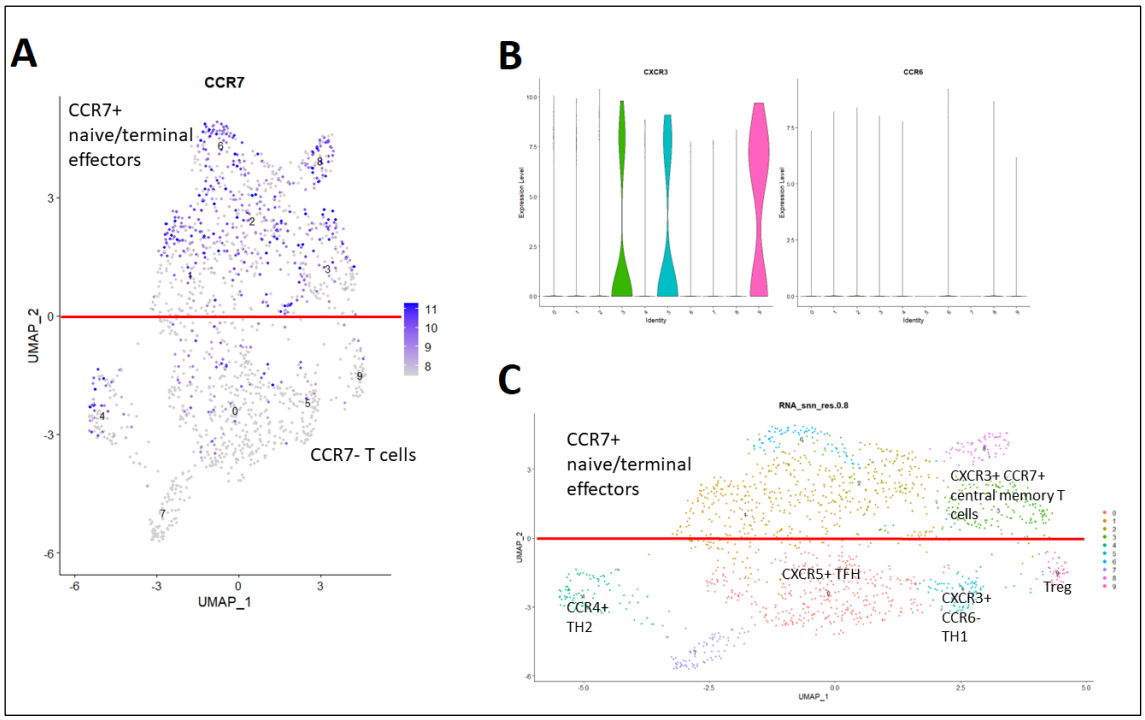


Figure 17. A) Umap plot showing CCR7 expression among the CD4+ T cells clusters identified using Seurat. B) Violin plot showing the expressions of both CXCR3 and CCR6 in the CD4+ T cell clusters. C) CD4+ T cells annotation based on markers expression.

We also used the SDY998 dataset to perform GSEA. GSEA reflects the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes. A positive Enrichment Score (ES) indicates gene set enrichment at the top of the ranked list while a negative ES indicates gene set enrichment at the bottom of the ranked list. We used the 664 IPA gene sets (gmt format) and GSEA software (259). The tool takes as input the gene expression matrix in a text format of the phenotypes under study. In our case, the phenotypes correspond to Th1 and naïve CD4+ T cells. 186 gene sets were enriched in the Th1 phenotype (gene sets with positive enrichment scores). The list of enriched gene sets in the Th1 phenotype was filtered with an FDR less or equal to 5% to keep only the statistically significant ones. 101 gene sets were kept. To ensure that the results are Th1-specific, we performed GSEA on three other datasets: GSE32901, GSE107011 and GSE135390 described below. Only gene sets identified with the SDY998 dataset and shared by at least one of the GSE32901, GSE107011 or GSE135390 datasets were kept (41 gene sets). Then, to avoid redundancy, only gene sets sharing less than 80% of their core enrichment genes were kept (24 gene sets).

3.5 GSE32901 dataset

GSE32901 is a publicly available microarray gene expression dataset containing PBMC samples from 5 healthy donors. It includes four types of samples: naive CD4+ T cells, Th1, Th17 rich, and Th17 poor memory CD4+ T cells. Researchers in the original study examined pairwise differential gene expression between the different types of samples available in the dataset (260). They demonstrated an increased expression in memory CD4+ compared to naïve CD4+ T cells of genes contained among immune-mediated disease loci. They also identified a differential regulation pattern for genes solely expressed in Th17 cells (IL17A and CCL20) compared to genes expressed in both Th17 and Th1 cells (IL23R and IL12RB2).

In this work, we normalized the gene expression using the log2 RMA method (261). We performed GSEA using the 664 IPA gene sets (gmt format), GSEA software, and the Th1 cells and naïve CD4+ T cells gene expression matrices. The list of enriched gene sets in the Th1 phenotype (with positive enrichment scores) was filtered with an FDR equal to 10% to keep only the statistically significant ones.

3.6 GSE107011 dataset

GSE107011 dataset is bulk RNA-seq transcriptome profiling of 29 immune cell types, including Th1 and naive CD4+ T cells, and extracted from PBMC sorted from 4 healthy individuals. The data was used first in a study that aimed to identify sets of genes that are specific, are co-expressed, and have housekeeping roles across the 29 cell types (258). Researchers examined differences in gene expressions revealing cell type specificity. They also performed absolute deconvolution using the identified transcriptomics signatures.

In this project, we normalized the counts using the Transcripts Per Million (TPM) method. We performed GSEA using the 664 IPA gene sets (gmt format), GSEA software, and the Th1 cells and naïve CD4+ T cells gene expression matrices. The list of enriched gene sets in the Th1 phenotype (with positive enrichment scores) was filtered with an FDR less or equal to 10% to keep only the statistically significant ones.

3.7 GSE135390 dataset

GSE135390 is an RNA-seq gene expression dataset. It contains PBMC samples coming from 3 healthy donors. Several subtypes of CD4+ T cells are available in the samples: Th1, TH2, TH22, Treg, naive CD4+ and Th17. The aim of the original study was to provide a better understanding of the functional specialization and differentiation of Th and Treg cell populations (262). Researchers identified in this

dataset core transcriptional signatures shared across all Th and Treg cell populations and unique signatures that define each of the Th or Treg populations.

In this work, we normalized the count data using Trimmed Mean of M-values (TMM). We performed GSEA using the 664 IPA gene sets (gmt format), GSEA software, and the Th1 cells and naïve CD4+ T cells gene expression matrices. The list of enriched gene sets in the Th1 phenotype (with positive enrichment scores) was filtered with an FDR less or equal to 10% to keep only the statistically significant ones.

3.8 E_MTAB_8322 dataset

E_MTAB_8322 is an RNA-seq single cell dataset publicly available in ArrayExpress database (263). It contains synovial samples coming from five treatment-naïve RA, six treatment-resistant RA, six in sustained remission and four patients with UPA (Early undifferentiated arthritis). Four healthy donor synovial tissues were included as control. The macrophage population was identified using FACS sorting. Quality control of the dataset involved removal of cells with less than 500 expressed genes and removal of weakly expressed genes. The aim of the original study was to identify the phenotypic changes in patients with early/active RA, treatment-refractory/active RA and RA in sustained remission (104). Researchers demonstrated that each clinical state was characterized by different frequencies of nine discrete phenotypic clusters within four distinct macrophage subpopulations with diverse homeostatic, regulatory and inflammatory functions. They revealed two subpopulations (MerTKposTREM2high and MerTKposLYVE1pos) with unique remission transcriptomic signatures, enriched in negative regulators of inflammation that induced the repair response of synovial fibroblasts in vitro.

In this work, we normalized gene expressions using Seurat's *NormalizeData* function. We only used the gene expression matrix of the five treatment naïve RA patients (5815 macrophage cells) in the next steps of the work.

Table 2 summarizes the datasets we analyzed for this PhD project, as well as the main outcomes of their analysis.

Table 2. Description of the omics datasets and their associated outcomes.

Dataset	Technology	Database	Cell type(s)	Outcomes
GSE97779	Microarray	GEO	-RA synovial macrophages -Healthy macrophages	-DEGs between RA and healthy macrophages
GSE164498	Single cell RNA seq	GEO	-Healthy M1 macrophages -Healthy M2 macrophages	-M1 and M2 marker genes
GSE109449	Single cell RNA seq	GEO	-RA synovial fibroblasts -OA synovial fibroblasts	-DEGs between RA and OA fibroblasts
SDY998	Single cell RNA seq	Immport	-RA synovial fibroblasts -RA synovial monocytes -RA synovial T cells -RA synovial B cells	- Synovial CD4+ Th1 cluster -DEGs between Th1 cluster and naive CD4+ T cells -Gene sets enriched in the synovial Th1 cluster
GSE32901	Microarray	GEO	-Healthy PBMCs (including naive CD4+ T cells and Th1 cells)	-Gene sets enriched in the Th1 cells from PBMCs samples
GSE107011	RNA seq	GEO	-Healthy PBMCs (including naive CD4+ T cells and Th1 cells)	-Gene sets enriched in the Th1 cells from PBMCs samples
GSE135390	RNA seq	GEO	-Healthy PBMCs (including naive CD4+ T cells and Th1 cells)	-Gene sets enriched in the Th1 cells from PBMCs samples
E_MTAB_8322	Single cell RNA seq	ArrayExpress	-RA synovial macrophages	-Normalized gene expression in the RA synovial macrophages

Chapter 4. Construction of RA cell-specific molecular interaction maps

4.1 Introduction

Knowledge assembly in the form of networks is of great interest in the field of systems biology as it can help elucidate the complexity of human diseases. Many public repositories of pathway assemblies are available, including KEGG (Kyoto Encyclopedia of Genes and Genomes) (264), Reactome (265), Wiki Pathways (266), Panther (267) as well as proprietary platforms, such as IPA (Ingenuity Pathway Analysis) (268) and MetaCore (269). However, most of the networks and pathways provided are generic, lack cell specificity, and are assembled without sufficient annotations or access to the literature used for their construction.

Moving beyond generic characterizations of biological processes to more detailed pathways that consider cell, tissue, organ, and disease states is a key element of current research. International projects such as the Human Cell Atlas (270) and the HapMap (271) have significantly contributed to our understanding of the role of different cells in the human body and the specificities linked to the anatomical or functional localization of these cells. Moreover, advancements in high-throughput sequencing and proteomics techniques offer insights into unprecedented specificity and resolution for various cell types and subpopulations in normal and pathological conditions.

In this direction, we decided to work on building the first cellular atlas of the rheumatic joint to recapitulate existing knowledge related to the intracellular interactions involved in the disease's pathogenesis in a cell-type and disease specific manner. We constructed cell-specific maps using a combination of information extracted from published literature or pathway databases, and omics data analysis and integration. The RA atlas includes for the moment four molecular interaction maps specific to the most abundant cell populations in RA synovium: synovial macrophages (including the M1 and M2 phenotypes), synovial fibroblasts and CD4+ T helper 1 (Th1) subtype, and it can be expanded to include a variety of other cells relevant to the RA pathology, such as chondrocytes and osteoblasts, or osteoclasts.

4.2 Materials and methods

Figure 18 illustrates the three steps process we followed to build the RA cell-specific maps.

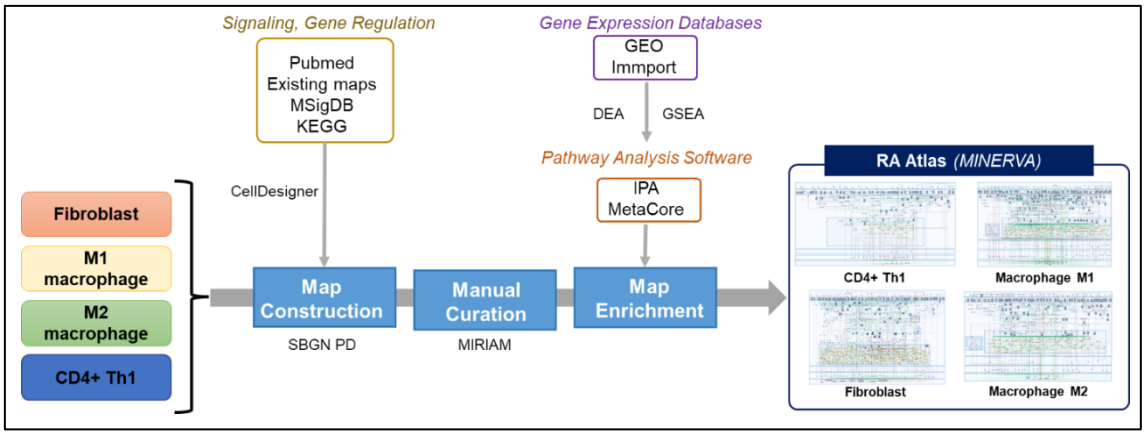


Figure 18. Workflow for the construction and use of the RA map. First, cell specific maps are constructed with the CellDesigner software (272) using information from published literature and signaling pathway databases. Then, they go through manual curation and annotation following MIRIAM standard. Next, maps are enriched with additional pathways involved in RA pathogenesis via the integration of disease- and cell-specific gene expression datasets using pathway analysis software like IPA and MetaCore. The resulting RA-Atlas is accessible online as an interactive map via MINERVA platform (273).

4.2.1 Molecular interaction maps construction

To construct the molecular interaction maps, we used the CellDesigner software (272), a diagram editor tool of gene-regulatory and biochemical networks. Nodes in the maps represent biomolecules like proteins, genes, complexes and other molecules, and the edges represent the interactions between the nodes, such as activation, inhibition, catalysis, and state transition. Syntax and semantics of the symbols used to represent both species and interactions are compliant with SBGN PD format (274). In addition, HUGO Gene Nomenclature Committee identifiers (HGNC) (275) were used for naming signaling and gene regulatory pathways components. Maps are stored in the Systems Biology Markup Language (SBML) format, a standard open extensible markup language (XML)–based format for representing biochemical reaction networks (276).

All cell-specific maps were built using the same structure. Cellular compartments were designed to reflect the signal transduction in a cell from the extracellular space in the form of ligand, going to the plasma membrane, where it binds to its corresponding receptor, passing through the cytoplasm and nucleus and leading to the production of proteins that can be either secreted in the extracellular space or transported to the membrane. Phenotypes are particular nodes in the maps. They represent the endpoint of multiple cellular processes that define and shape the morphology and function of the cell. They describe biological states known to be active or inactive in RA. To define the phenotypes of each cell-specific map

and their corresponding regulators, we used databases such as the MSig database (277), KEGG (264) and bio curation. The retrieved gene signatures were used to connect molecules to the relevant cellular phenotypes included in the maps.

4.2.2 Annotations and curation criteria

All the cell-specific maps were manually curated. The first step was a thorough investigation of public data including published articles and databases to identify pathways, components and interactions implicated in RA pathogenesis. We used small-scale experimental investigations in human and included animal-based studies where human-related data was insufficient to confirm the involvement of components or interactions. Data from animal studies remains limited and completely traceable. We followed strict curation criteria when constructing the maps. Only components and interactions experimentally demonstrated to be expressed in RA pathogenesis and in the relevant cell-type were added. When disease- or cell-specific information was unavailable, general pathway interactions were added to complete specific pathways. MIRIAM (11), a standard for annotating and curating computational models and maps, was used to add annotations and provide references for all the components and interactions. References were added using the tag "bqbiol: isDescribedBy", available in the MIRIAM section of CellDesigner. This tag is used to link a species or a parameter to the literature that describes the concentration of the species or the value of the parameter. We used it to connect a component or an interaction to the literature or data that describe it (e.g., PubMed references (PMIDs), DOI, GEO, KEGG identifiers) (Figure 19).

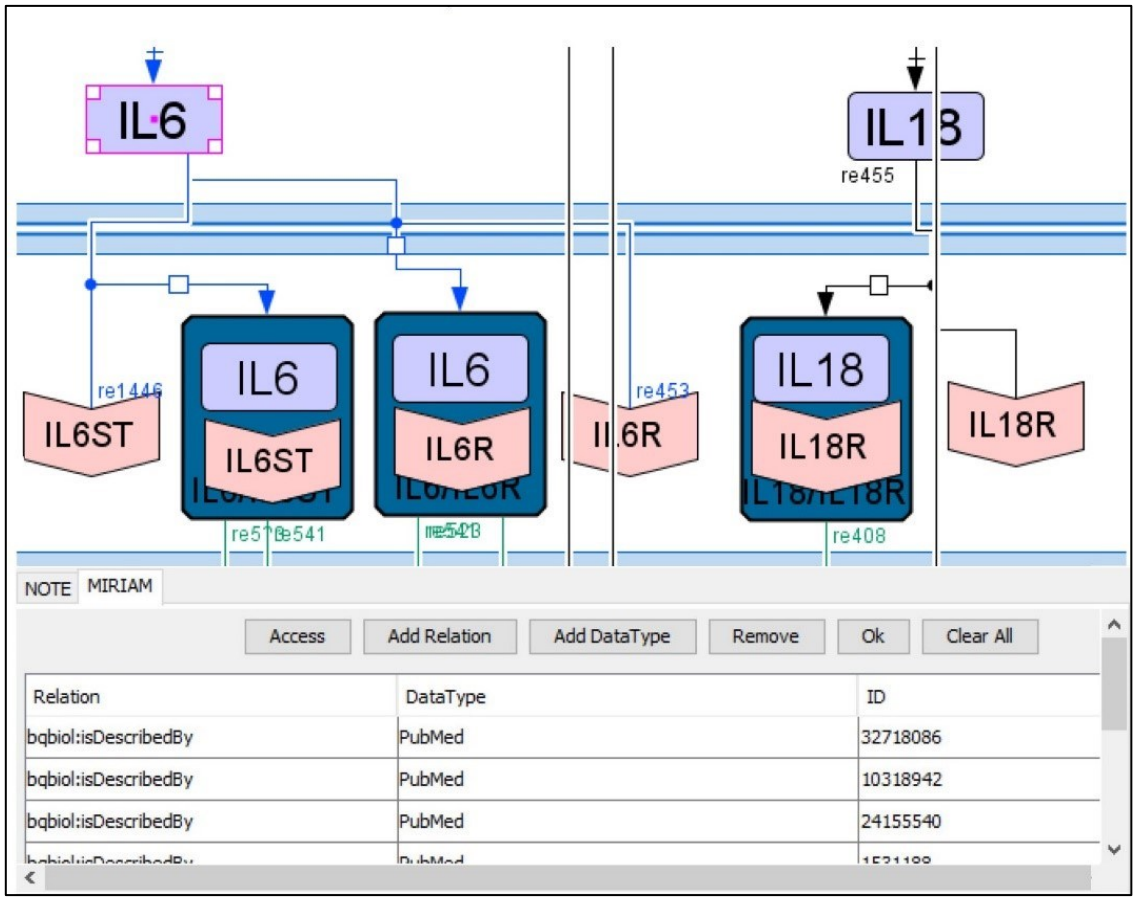


Figure 19. Annotations added to the MIRIAM section of CellDesigner.

4.2.3 Integration of pre-existing maps

We integrated publicly available pre-existing maps where available for the cell-types of interest. We used cell and disease-specific markers or gene expression datasets to retrieve signaling pathways involved in RA pathogenesis and filter out irrelevant components.

Some of these maps were not built in an SBGN PD compliant format and had to be converted. To do this, relevant mechanistic details that we retrieved from published literature and signaling pathway databases were used to complete the missing parts in the pathways of interest before adding them to the maps.

4.2.3.1 RA macrophage map

The RA macrophage map was constructed by integrating three pre-existing maps. The first map was the RA-specific macrophage map from IPA including 44 molecules and 47 interactions associated with PMIDs. The second and third macrophage-specific maps (but not RA specific) were retrieved from the published literature (278,279). As the second and third macrophage maps were not RA-specific, we used the RA-specific marker gene lists from IPA as an overlay

in MINERVA (Molecular Interaction NEtwoRks VisuAlization) (273) to identify disease-specific molecules and submodules (Figure 20). We were able to extract these submodules of interest from the macrophages' maps using the Stream export plugin (280). We identified the signaling pathways present in the retrieved submodules and completed them by incorporating molecules from the literature and pathway databases such as KEGG.

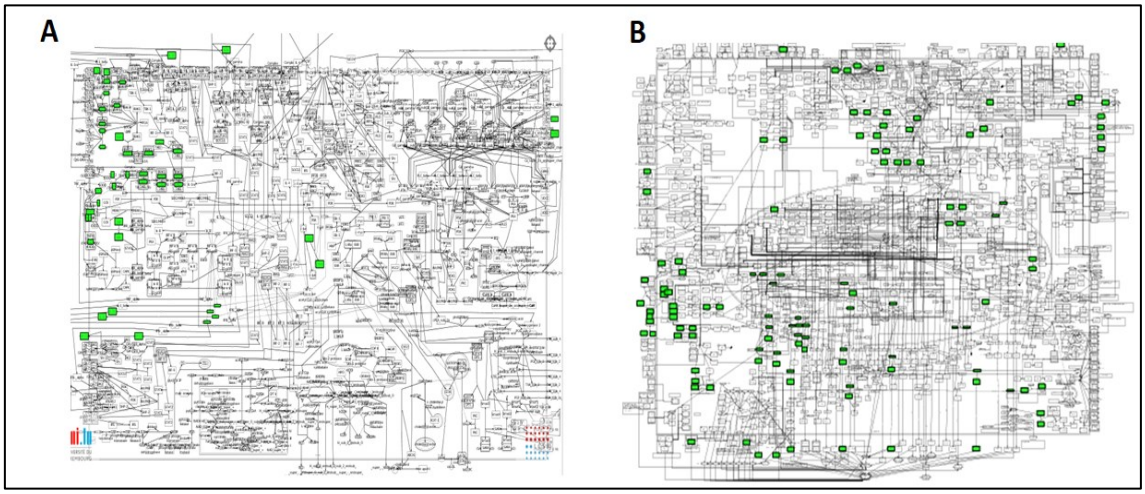


Figure 20. A) Macrophage maps from (Kanae Oda et al., 2004) (278) in MINERVA. B) Macrophage map from (Pia Wentker et al., 2017) (279) in MINERVA. Green color indicates the overlay with RA specific molecules retrieved from IPA.

4.2.3.2 RA fibroblast map

We constructed the RA fibroblast-specific map by filtering non-fibroblast-specific components from the initial RA-map (281) (Figure 21) using the most recent RA fibroblast overlay available in MINERVA (see Cell-specific overlays section). Non-fibroblast-specific components were only conserved if they were essential for signal transduction.

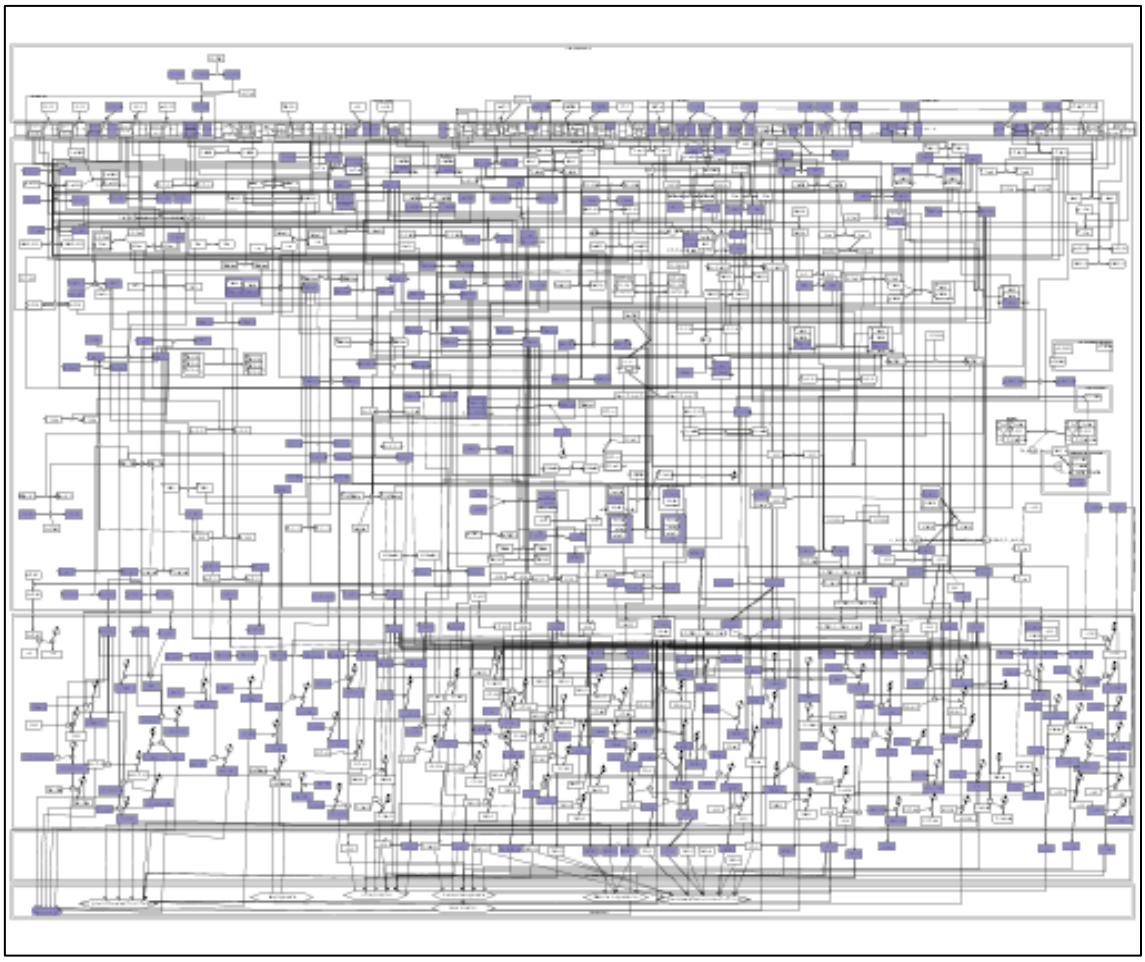


Figure 21. Initial RA map from Singh et al.,2020 (282). Purple color indicates the overlay with the latest RA fibroblast overlay available in MINERVA.

4.2.3.3 RA CD4+ Th1

We used the multicellular map of inflammation (283) that provides a non-RA specific Th1 submap. As we can see in Figure 22, the map is smaller in size compared to the previous ones. This is due to the lack of information about gene expression and signaling pathways that are specific to the Th1 subtype. Indeed, the majority of public data merely investigate biological mechanisms involved in the functions of the overall T lymphocytes population.

We were able to identify transcription factors and cell-specific receptors involved in RA pathogenesis from this map through extensive literature mining. Only components identified to be expressed in RA Th1 subtype in published literature were added to the map.

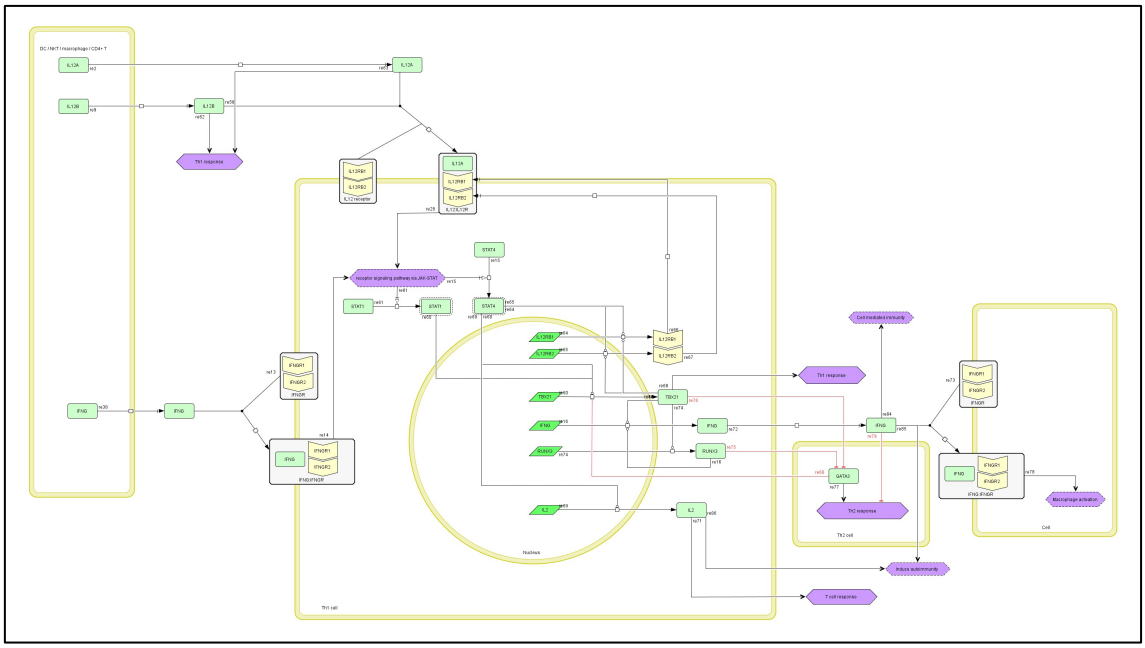


Figure 22. Th1 map within the atlas of inflammation resolution (283) in CellDesigner.

4.2.4 Cell-specific maps enrichment using gene expression datasets

To expand the cell-specific maps, we utilized disease and cell-specific gene expression datasets that are publicly available combined with pathways analysis tools, namely IPA and MetaCore version 22.1 build 70800 (269). Both provide massive map databases that we utilized to identify new pathways. We used the "Pathway maps" option in MetaCore in the "One-click analysis" page (Figure 23A) to map cell and disease specific lists of differentially expressed genes (DEGs) to the internal Metacore map database. Next to each retrieved map, the number of overlapping DEGs is indicated. IPA's "canonical pathways" tool was used to perform the same analysis (Figure 23B). The highlighted DEGs and pathways were identified, added to the relevant cell-specific map, and annotated with PMIDs or GEO accession numbers (284).

Some maps in IPA and MetaCore are not SBGN compliant. Before integrating the highlighted pathways identified in these maps, we converted them into an SBGN PD format by adding mechanistic details that we retrieved from published literature and signaling pathway databases.

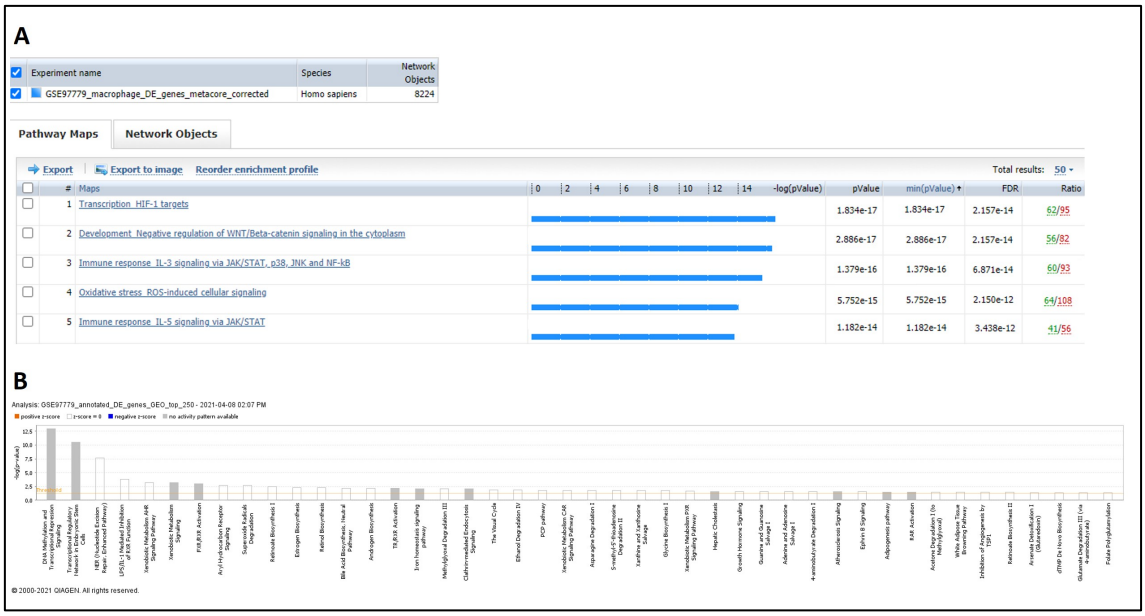


Figure 23. A) Screenshot showing "Pathway maps" option in MetaCore in the "One-click analysis" page. B) Screenshot showing IPA's "canonical pathways" tool.

4.2.4.1 RA macrophage map

DEA was performed between RA and healthy synovial macrophages samples in the GSE97779 microarray dataset (see Chapter 3. Description of data and data analysis). The resulting list of DEG was mapped to existing pathways and biological functions that are part of IPA and MetaCore knowledgebases. Highlighted signaling pathways were retrieved from the overlapping maps, manually curated, and annotated with PMIDs or GEO accession, and added to the RA macrophage map.

To obtain separate maps for the M1 and M2 phenotypes, first we identified M1 and M2 gene signatures. DEA was performed between M1 macrophage and M2 macrophage cells present in the GSE164498 dataset (see Chapter 3. Description of data and data analysis). We combined the resulting list of DEGs with M1 and M2 macrophage Msig gene signatures. We also went through an extensive literature-search to retrieve information regarding the expression of the biomolecules and pathways present in the RA macrophage map in each phenotype. Biomolecules found in the literature to be upregulated in the M1 or M2 phenotype were added to the signatures. The percentage of shared components between each pathway in the map and the M1 and M2 signatures was calculated and based on these percentages, pathways were assigned to the M1 or M2 phenotype. However, it was impossible to assign a phenotype for some pathways due to lack of information or expression. These pathways were kept in both M1 and M2 maps.

4.2.4.2 RA fibroblast map

DEA was performed to identify differentially expressed gene between RA and osteoarthritis fibroblast cells in the GSE109449 single-cell RNA-seq dataset (see Chapter 3. Description of data and data analysis). The resulting DEG list was mapped to IPA and MetaCore knowledgebases. Highlighted signaling pathways were retrieved from the overlapping maps, manually curated, annotated with PMIDs or GEO accession, and added to the RA fibroblast map.

4.2.4.3 RA CD4+ Th1 map

To enrich the map with additional pathways, first we used the SDY998 single-cell RNA-seq dataset to identify the Th1 cluster. Then, we performed DEA between the identified Th1 cluster and naïve CD4+ T cells in RA samples (see Chapter 3. Description of data and data analysis). Using the IPA and MetaCore databases and the resulting DEGs list, new signaling pathways were identified and added to the map. We also used Gene Set Enrichment Analysis (GSEA) software (259), IPA gene sets, and the SDY998 gene expression dataset to further enrich the Th1 map. We filtered the retrieved gene sets to only keep the statistically significant ones. We also made sure that the results were Th1-specific by running GSEA on three additional datasets: GSE32901, GSE107011, and GSE135390. Strict filtering criteria were set up to keep gene sets that are at least identified in different datasets and to discard the redundant ones. Details regarding GSEA analysis and datasets used are described in Chapter 3. Description of data and data analysis. Filtered IPA gene sets went through manual curation and were added to the Th1 map.

4.2.5 Cell-specific overlays

To calculate the cell specificity of the RA-Atlas maps or to visualize the cell specific pathways and components, we provide three different cell-specific overlays, namely fibroblasts, macrophages and Th1. They contain lists of biomolecules that have been identified as being expressed in the relevant cell type. The fibroblast overlay is an updated version of the overlay provided in the initial RA map article (282). We added to it new components coming from literature mining and the list of DEGs obtained from DEA (see Cell-specific maps enrichment using gene expression datasets). The overlay file consists of 2,409 fibroblast-specific components. The macrophage overlay includes the RA macrophage-specific components identified in the literature, the DEG list obtained from DEA (see Cell-specific maps enrichment using gene expression datasets) and the 44 molecules present in the IPA RA-specific macrophage map (see Integration of pre-existing maps). It consists of 648 RA macrophage-specific molecules. The Th1 overlay consists of 523 molecules, including the list of DEGs used to enrich the map, the core enrichment genes from GSEA (see Cell-specific maps enrichment using gene expression datasets) and Th1-specific components identified in the literature.

4.3 Results

We present an attempt to create the first cellular atlas of the rheumatic joint to recapitulate existing knowledge related to the intracellular pathogenic mechanisms in four different cell types: the fibroblast, the M1 macrophage, the M2 macrophage and the Th1 (285). Figure 24 illustrates the main page of the RA-Atlas in MINERVA platform.

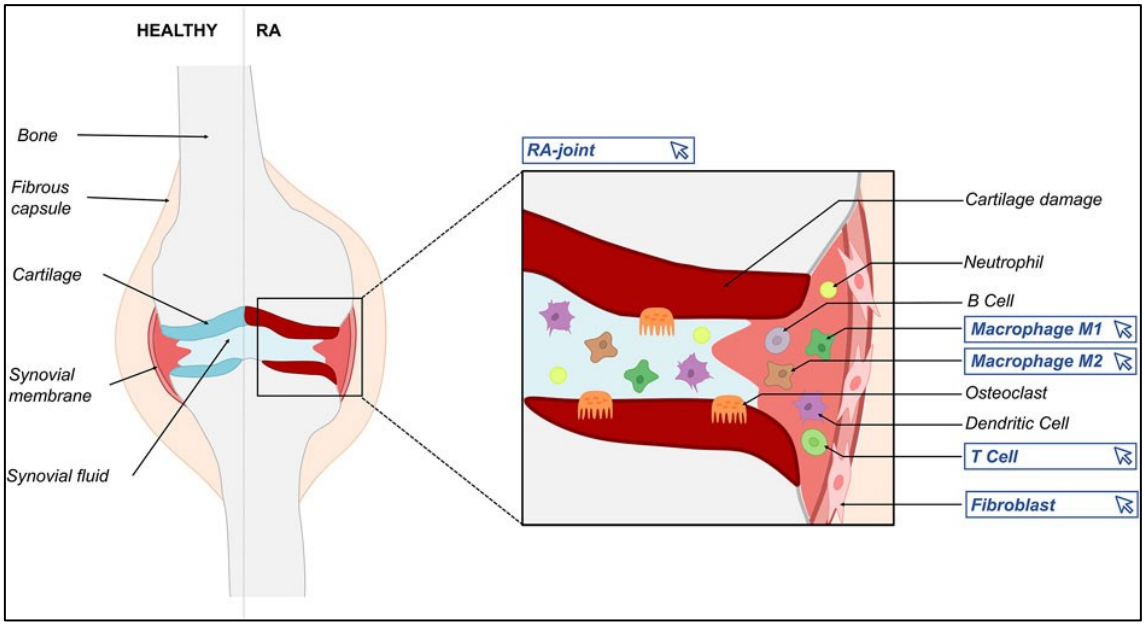


Figure 24. RA Atlas in MINERVA. Blue frames indicate clickable links leading to the different cell-specific maps included in the RA-Atlas.

Full-page landscape images of the RA fibroblast map, the RA M1 macrophage map, the RA M2 macrophage map and the RA CD4+ Th1 maps are provided in Supplementary figure 1, Supplementary figure 2, Supplementary figure 3 and Supplementary figure 4 respectively. However, due to their high complexity, these images are not readable in a A4 size page and are displayed for illustrative purposes only. For an optimal visual exploration of the maps, see Visualization and accessibility section.

All the cell-specific maps are compartmentalized in a way to represent the flow of information from the extracellular space (ligands) to the plasma membrane (ligands-receptors complexes) through the cytoplasm (signaling cascades), the nucleus (gene regulation) and the secreted compartments (phenotype activation) (Figure 25).

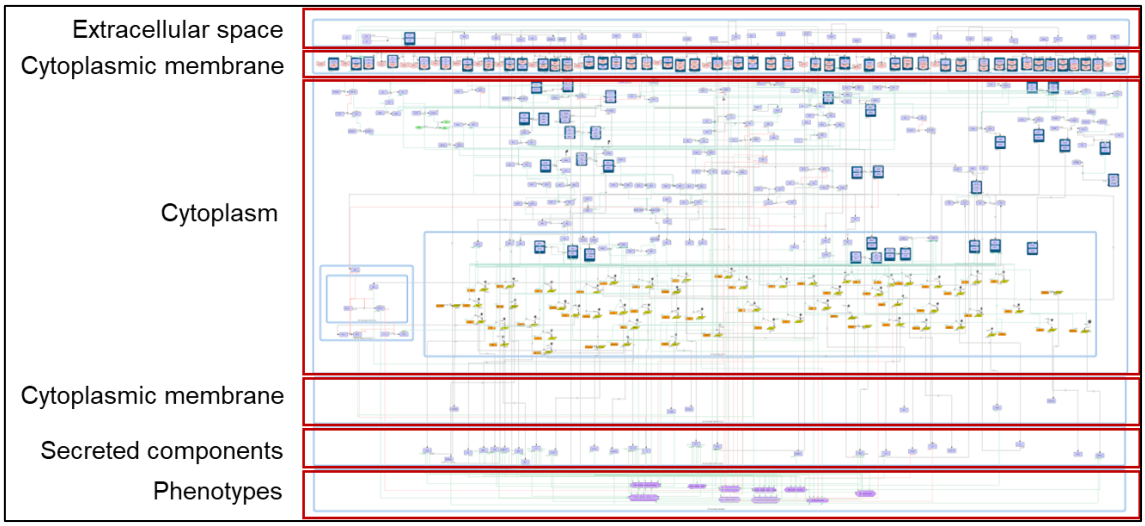


Figure 25. Cellular compartments of the RA M1 macrophage map

4.3.1 RA fibroblast map

After manual curation and gene expression data analysis, 596 species were identified as fibroblast-specific in the initial RA-map (282). We kept these 596 species in the map and filtered out the remaining non-fibroblast-specific components. We also added 257 new species to the RA fibroblast map using IPA and MetaCore. Table 3 displays the top five discovered signaling pathways.

Table 3. IPA and MetaCore maps using the RA fibroblast DEGs list.

IPA maps	Overlap	MetaCore pathways maps	Overlap
Antigen presentation	48,70%	Immune response _Induction of the antigen presentation machinery by IFN-gamma	45,20%
Inhibition of matrix metalloproteases	39%	Cell adhesion_ECM remodeling	41,80%
Hepatic fibrosis/Hepatic stellate cell activation	22,70%	IL-1 beta- and Endothelin-1-induced fibroblast/ myofibroblast migration and extracellular matrix production in asthmatic airways	47,50%
GP6 signaling pathway	20,50%	TGF-beta-induced fibroblast/ myofibroblast migration and extracellular matrix production in asthmatic airways	36,60%
Axonal guidance signaling	15%	Glucocorticoid-induced elevation of intraocular pressure as glaucoma risk factor	32,20%

The map is SBGN PD compliant. It contains 853 species, including 411 proteins, 115 genes, 115 RNAs, 96 molecular complexes, 2 simple molecules and 9 phenotypes (8 of them coming from the initial RA-map, the 9th being T cells activation). The components interact via 509 reactions.

77% of the map's components are RA fibroblast-specific (Figure 26). Some of the molecules that are not fibroblast-specific were added to the map to keep the signal transduction going from the extracellular part to the nucleus. Indeed, in some RA fibroblast-specific pathways, these molecules are needed to activate or inhibit intermediate proteins. Other molecules that are not RA fibroblast-specific are found in the extracellular space. These molecules are secreted by the other cell types present in the synovial fluid and can activate specific intracellular path-

ways in the fibroblasts via ligand/receptors interactions. Therefore, these secreted molecules do not have to be RA fibroblast specific to be kept in the map. Other small molecules and adaptor proteins for which supporting literature was not available were kept on the map.

4.3.2 RA macrophage maps

We constructed a global macrophage map by integrating the IPA cell- and disease-specific map. As the map was built in SBGN AF language, we converted it into an SBGN PD compliant map in CellDesigner by adding the missing mechanistic details. New pathways were added from the RA-specific submodules exported from the previously published macrophages maps (278,279), like TLRs pathway, NF-κB pathway, PI3K-AKT pathway, MAPK signaling pathway, and IL-18. Other pathways were identified from the literature, such as the IFN pathway, Notch pathway, IL10 and IL23 pathways. KEGG and MSig database, and literature screening allowed us to add 8 phenotypes to the map. The phenotypes include inflammation, apoptosis, angiogenesis, extracellular matrix degradation, cell chemotaxis/migration, T cells activation, proliferation/survival and osteoclastogenesis. In total, 492 species were added to the RA macrophage map.

The map was also enriched using the GSE97779 gene expression dataset via the identification of 385 new species involved in RA macrophage-specific pathways using IPA and MetaCore. The top 5 identified signaling pathways are shown in Table 4.

Table 4. Top 5 IPA and MetaCore maps using RA macrophage DEGs list.

IPA maps	Overlap	MetaCore maps	Overlap
DNA methylation and transcriptional repression signaling	32,30%	Development_Role of proteases in hematopoietic stem cell mobilization	22,20%
Transcriptional regulatory network in embryonic stem cells	19,60%	Colorectal cancer (general schema)	20%
NER (nucleotide excision repair, enhanced pathway)	10,10%	Cigarette smoke components TCDD and Benzo[a]pyrene and receptor AHR signaling in lung epithelial cells	13,80%
Xenobiotic metabolism AHR signaling pathway	6,60%	Cigarette smoke-mediated regulation of NRF2-antioxidant pathway in airway epithelial cells	13,70%
LPS/IL1 mediated inhibition of RXR function	4,10%	Inhibition of Ephrin receptors in colorectal cancer	13,30%

4.3.3 RA M1 macrophage map

The RA macrophage M1-specific map includes all the signaling pathways from the global RA macrophage map leading to the macrophage polarization into the pro-inflammatory M1 phenotype. The RA macrophage M1-specific map includes 640 species: 278 proteins, 76 genes, 76 RNAs, 114 molecular complexes, 8 simple molecules and the same phenotypes as the global RA macrophage map. These components interact together via 448 reactions. 61% of the components of this map are RA macrophage-specific (Figure 26). All the receptors and secreted components are RA macrophage specific. Although less literature is available for the construction of the RA macrophage map compared to the RA fibroblast map, gene expression datasets integration enabled us to keep a high RA macrophage specificity in the map.

4.3.4 RA M2 macrophage map

The M2 macrophage map includes all the signaling pathways from the global RA

macrophage map leading to the macrophage polarization into the anti-inflammatory M2 phenotype. The enriched map includes 520 species: 243 proteins, 59 genes, 59 RNAs, 90 molecular complexes, three simple molecules and 7 of the 8 phenotypes from the RA macrophage map. Indeed, osteoclastogenesis is not associated with any of the molecules in the RA M2 macrophage map. The M2 macrophage map's components interact together via 342 reactions. 55% of the components of this map are RA macrophage-specific (Figure 26). All the receptors and secreted components are RA macrophage specific. Because RA disease is mainly associated with the M1 phenotype, we can see that fewer pathways are specific to the M2 phenotype.

4.3.5 RA Th1 Map

We created the RA Th1 cell map by mining the literature for the major cell-specific signaling pathways implicated in the pathogenesis of RA. However, there was less information about the Th1 subtype in the literature than regarding fibroblasts and macrophages. As a result, the map merely included 130 species at that point. Therefore, it was critical to expand it using new ways. IPA, MetaCore, and GSEA enriched gene sets were used to identify new signaling pathways. The top 5 identified signaling pathways using IPA and MetaCore are shown in Table 5. The list of enriched gene sets in RA Th1 after filtering are displayed in Table 6.

Table 5. Top 5 IPA and MetaCore maps using RA Th1 DEGs list.

IPA map	Overlay	MetaCore map	Overlay
Antigen presentation pathway	30,80%	Generation of cytotoxic CD8+ T cells in COPD	24%
B cell developement	22,70%	Maturation and migration of dendritic cells in skin sensitization	19,50%
TH1 pathway	10,70%	Immune response_Induction of the antigen presentation machinery by IFN-gamma	18,80%
TH2 pathway	10,20%	COVID-19: immune dysregulation	17%
TH1 and TH2 pathway	9,30%	Chemokines in inflammation in adipose tissue and liver in obesity, type 2 diabetes and metabolic syndrome X	14,50%

Table 6. List of enriched gene sets in RA Th1 after filtering.

Gene sets	FDR
ICOS-ICOSL SIGNALING IN T HELPER CELLS	0,003
CDC42 SIGNALING	0,003
COMMUNICATION BETWEEN INNATE AND ADAPTIVE IMMUNE CELLS	0,004
TH1 ACTIVATION PATHWAY	0,004
PKC SIGNALING IN T LYMPHOCYTES	0,005
ANTIGEN PRESENTATION PATHWAY	0,005
PROTEIN UBIQUITINATION PATHWAY	0,007

ALTERED T CELL AND B CELL SIGNALING IN RHEUMATOID ARTHRITIS	0,007
CROSSTALK BETWEEN DENDRITIC CELLS AND NATURAL KILLER CELLS	0,009
INTERFERON SIGNALING	0,01
OXIDATIVE PHOSPHORYLATION	0,01
REGULATION OF CELLULAR MECHANICS BY CALPAIN PROTEASE	0,013
REMODELING 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 SIGNALING IN THE CARDIOVASCULAR SYSTEM	0,031
PTEN SIGNALING	0,032
ACTIVATION OF IRF BY CYTOSOLIC PATTERN RECOGNITION RECEPTORS	0,034
IL-8 SIGNALING	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

The resulting map consists of 321 species, including 167 proteins, 29 genes, 29 RNAs, 64 molecular complexes, and 7 phenotypes. The phenotypes are inflammation, apoptosis, osteoclastogenesis, cell chemotaxis, angiogenesis, matrix degradation and proliferation. RA Th1 map's components are connected through 179 reactions. 58% of the components of this map are RA Th1-specific, including all the transcription factors, most of the receptors, and secreted components (Figure 26).

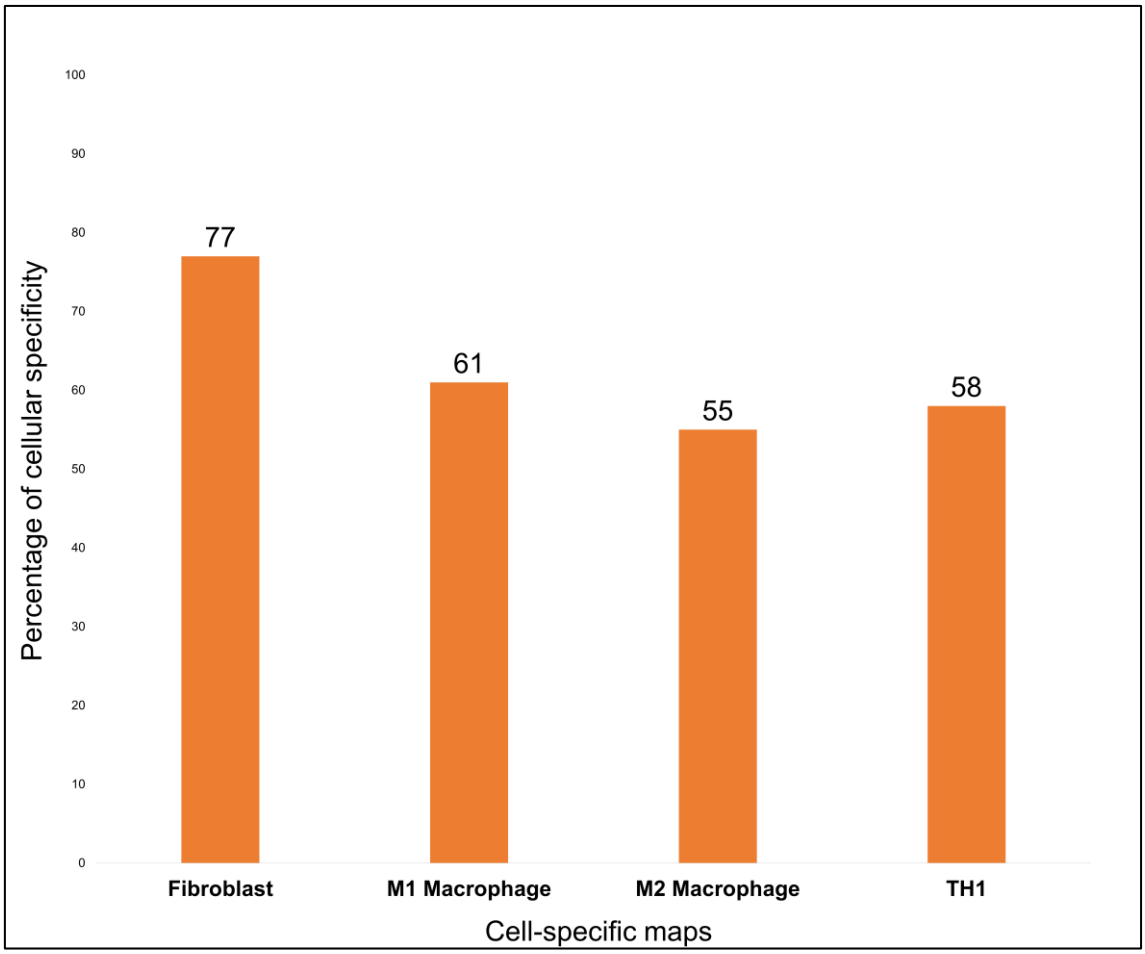


Figure 26. Percentage of cell-specific components per cell-specific maps. Figure taken from Zerrouk et al., 2022 (285).

4.3.6 Visualization and accessibility

All cell-specific maps are available as online interactive maps on the standalone web server MINERVA at https://ramap.uni.lu/minerva/. The MINERVA platform enables the visual exploration, analysis and management of molecular networks encoded in systems biology formats, including CellDesigner, SBML and SBGN. All PMIDs of scientific articles used for the map construction are available in the annotation section of the maps. We also provide cell-specific overlays (see Cell-specific overlays) extracted from the literature and publicly available datasets to visualize cellular signatures on the different RA-Atlas maps. All data and code used to generate results, including XML files of the maps, cell-specificity, overlays, and gene expression analysis are available on a GitLab repository at https://gitlab.com/genhotel/a-mechanistic-multicellular-atlas-of-the-rheumatic-joint.

4.3.7 Applications

MINERVA provides multiple possibilities for users to explore the various RA-Atlas

maps. Users can search the maps for components of interest and their references (Figure 27A) or visualize experimental data and cell-specific biomolecules using the relevant overlays (Figure 27B). Users can search for drug targets online using DrugBank (286) and CHEMBL (287). Users can also use plugins available in MINERVA, such as the Stream export, to export parts of the network upstream and downstream of a node, or GSEA plugin to calculate enrichment for uploaded user-provided data overlays.

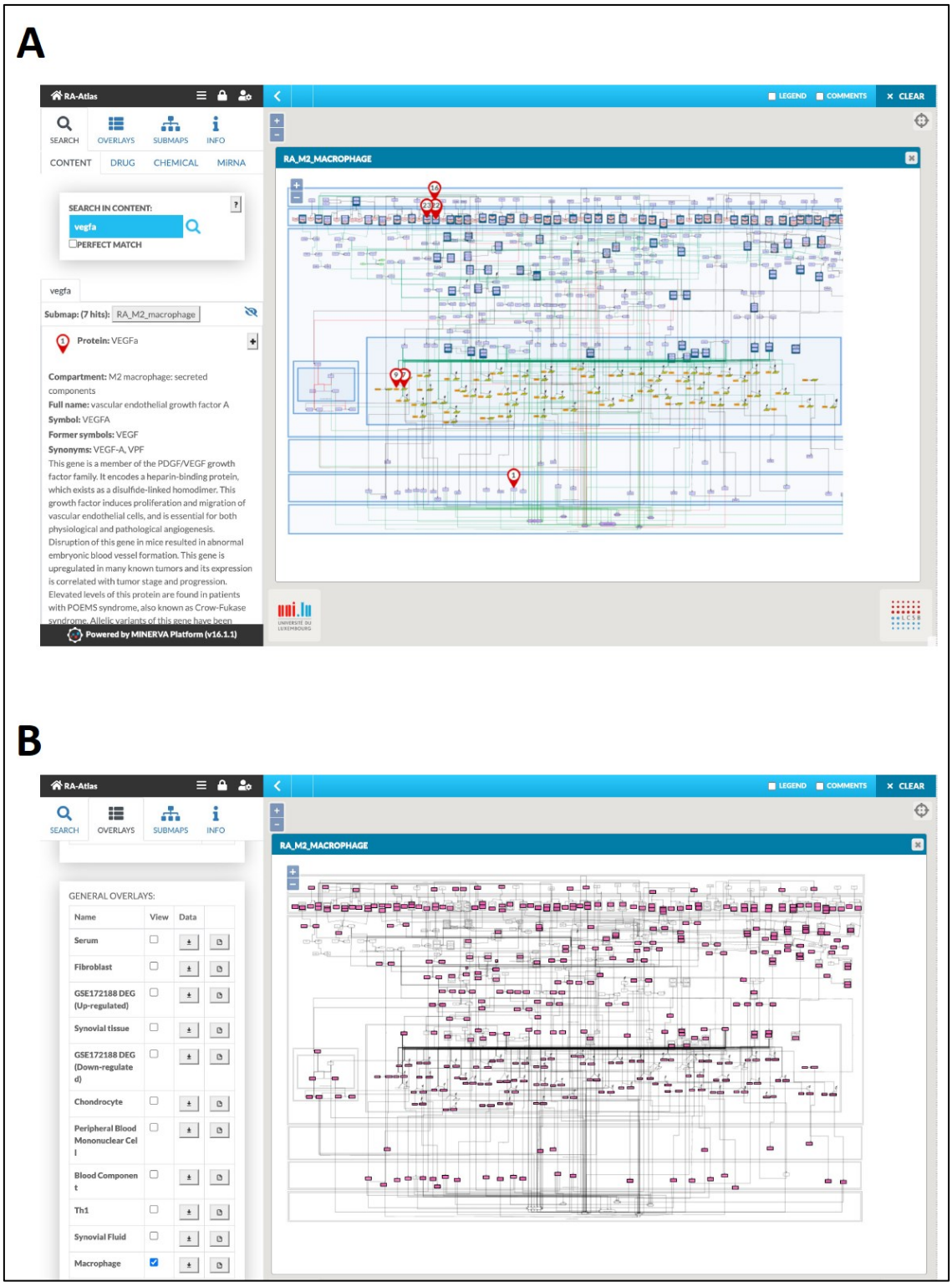


Figure 27. Visualization of RA macrophage M2 map in MINERVA. (A) Snapshot of the search for VEGFA. (B) Snapshot of the visualization of the macrophage overlay. Figure taken from Zerrouk et al., 2022 (285).

4.4 Discussion

We present here the first attempt to create a cellular atlas for RA disease (285), an interactive, manually curated and enriched with extensive omics data analysis representation of molecular mechanisms involved in RA's pathogenesis. It includes cell-specific molecular interaction maps for resident fibroblasts, M1 and M2 resident macrophages and CD4+ Th1 cells. Maps are primarily based on high-quality manual curation of scientific literature and enrichment and cross-validation using expression datasets. All the components and reactions in the RA atlas are annotated using RA and human-specific studies and datasets.

These maps can be used as an interactive knowledge base, accessible online via the platform MINERVA[1], and serve as a template for overlaying multiple datasets. Visualization of experimental data through the maps could help highlight aspects of the affected biological processes and make differences between experimental conditions more evident.

We made a significant effort to implement the FAIR (Findable, Accessible, Interoperable and Reusable) principles in this work (288). Indeed, the atlas content is compliant with SBGN PD for the representation, SBML for the construction, MIRIAM for the annotation, PMIDs and stable identifiers for the references, and HGNC symbols for the annotation of the signaling and gene regulation components. Moreover, we give a thorough description of the methodology used to build the RA-Atlas in our open access publication (285), allowing for transparent and facilitated reuse of the resource.

The RA Atlas is a part of the Disease Map community project (9,171), making its dissemination easier and further contributing to the community effort at fostering collaborations for the creation of disease-specific maps such as The Parkinson's map (166), the AsthmaMap (170), the Atlas of Cancer Signaling Network (169), The Atlas of Inflammation Resolution (283), and more recently, the COVID19 Disease Map (171).

The next steps of this work will be to expand the RA-Atlas with molecular interaction maps of B cells, chondrocytes, osteoblasts, and osteoclasts and enrich existing cell-specific maps with relevant metabolic pathways.

Chapter 5. Construction of the RA multicellular map

5.1 Introduction

In the previous chapter, we presented the first RA cellular atlas of the rheumatic joint that recapitulates existing knowledge related to the intracellular interactions involved in the disease's pathogenesis in a cell-type specific manner. However, in the RA-Atlas, cells are isolated from one another and do not communicate with each other. This representation does not consider the intricate cellular interplay that occurs in the RA synovium. Therefore, to better understand RA, we need a multicellular representation of the synovium, that describes both the intracellular cascades and the cellular crosstalks involved in the disease pathogenesis (4).

Cellular communication can be achieved in two general ways: first through secreted mediators, notably inflammatory cytokines such as TNF, IL-6, IL-17, and many others; and second through direct cell-cell contact that is mediated by cell surface receptors and ligands, including some membrane-anchored proteins (92). These mediators alter the activation or differentiation state of one or both cell types. They regulate inflammation, autoimmunity, and articular destruction in the joints by initiating cascades of signaling pathways further resulting in the expression of pro inflammatory molecules. Such cascades trigger disease phenotypes like angiogenesis, cartilage matrix degradation, inflammation, and synovial hyperplasia (115). Mapping these ligand–receptor interactions is therefore fundamental to understanding cellular behavior and response to neighboring cells.

In this direction, we developed the first RA multicellular molecular interaction map, a state-of-the-art description of the rheumatic joint that recapitulates knowledge related to both intra- and intercellular interactions involved in the disease's pathogenesis. We connected the RA cell-specific maps (see Chapter 4. Construction of RA cell-specific molecular interaction maps) via the addition of bidirectional intercellular interactions that were identified using both literature and database mining, and omics data integration in a disease- and cell-specific manner.

5.2 Materials and methods

Figure 28 illustrates the three steps process we followed to identify disease-specific cell-cell interactions between the cell-types represented in the RA cellular atlas.

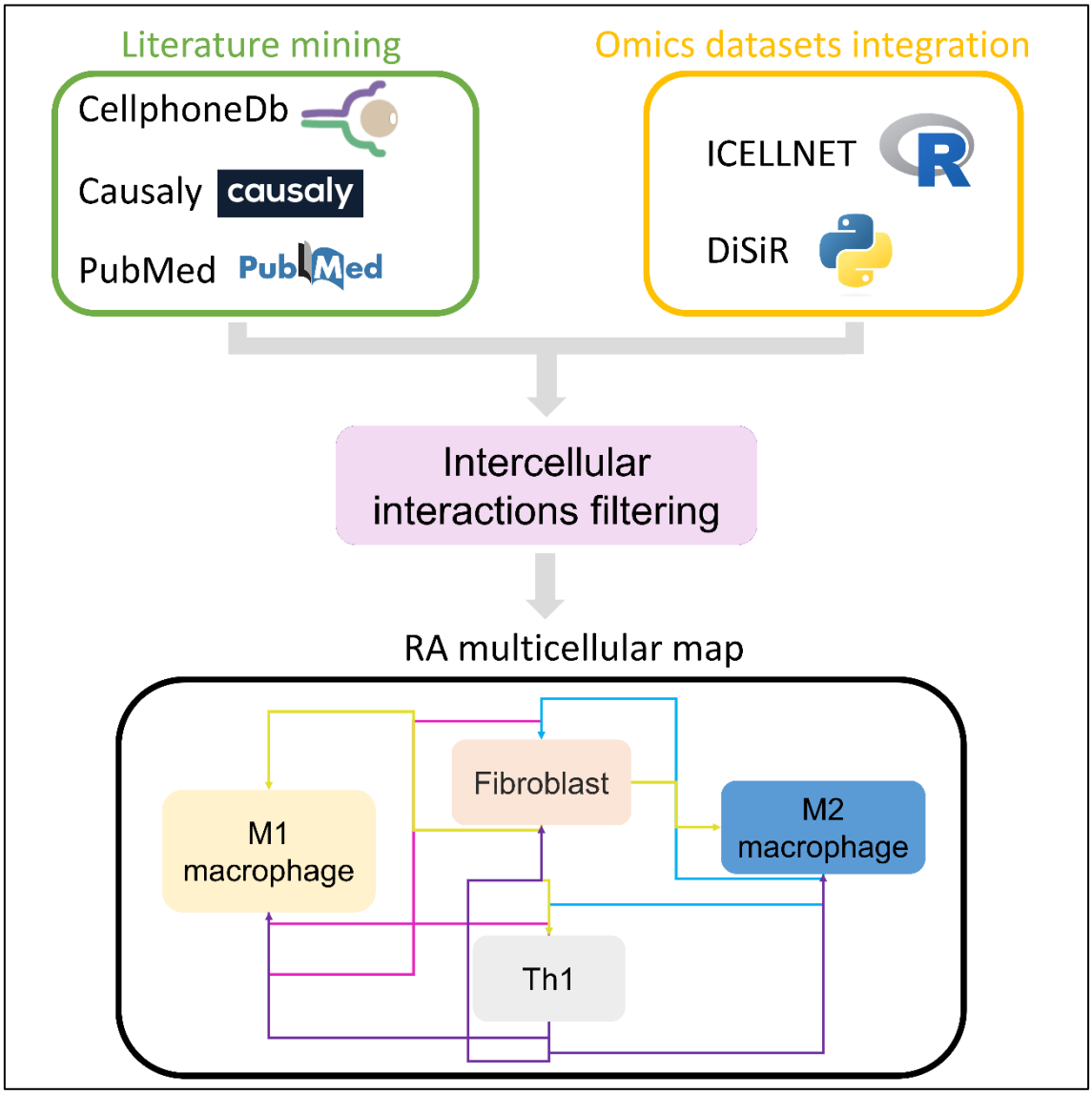


Figure 28. Workflow for the construction of the RA multicellular map. We used literature and databases mining via several tools like Causaly (289) and CellphoneDB (179) as well as keywords search in PubMed, to retrieve the intercellular interactions already identified in the published literature. We also integrated omics datasets via ICELLNET (178) and DiSiR (180) to identify statistically significant interactions observed in these datasets. We filtered the retrieved interactions to cross validate the results and only keep the most reliable ones. We used those filtered interactions to connect the RA cell-specific maps and generate the first RA multicellular map.

5.2.1 Identification of cell-cell interactions using literature and databases mining

We used the tool Causaly (289) to uncover intercellular relationships between RA M1 and RA M2 macrophages, RA fibroblasts, and RA Th1 subtypes in published literature. Causaly is a biomedical discovery research tool that answers research questions that can be hard to answer with traditional keyword search. The system augments the human process of finding evidence which typically is very time-consuming due to the vast amount of biomedical literature available.

Underlying Causaly is an advanced Artificial Intelligence that is machine-reading literature across millions of academic publications and distills the evidence out of it into a knowledge graph. Unlike other tools which just count statistically how often different words appear in the same sentence, Causaly understands the directionality of evidence e.g., the difference between a treatment and a side effect, via its knowledge graph that consists of 220 million directional relationships. These relationships can be extracted from unstructured- and structured data sources as well as public, licensed and Proprietary data (PubMed, Patents, clinical trials protocols, chemical, target, and safety databases…). Causaly uses the Unified Medical Language System ontology (UMLS) to structure all data including MeSH, SNOMED, GO and many others. It also organizes knowledge by linguistic strength from co-occurrence evidence to causal evidence.

Figure 29 shows a screenshot of the advanced search module in Causaly that we used to identify intercellular interactions between our cell-types of interest, namely macrophages, fibroblasts and Th1, in the context of rheumatoid arthritis disease. All the retrieved information from Causaly was carefully and manually verified to validate the quality of evidence.

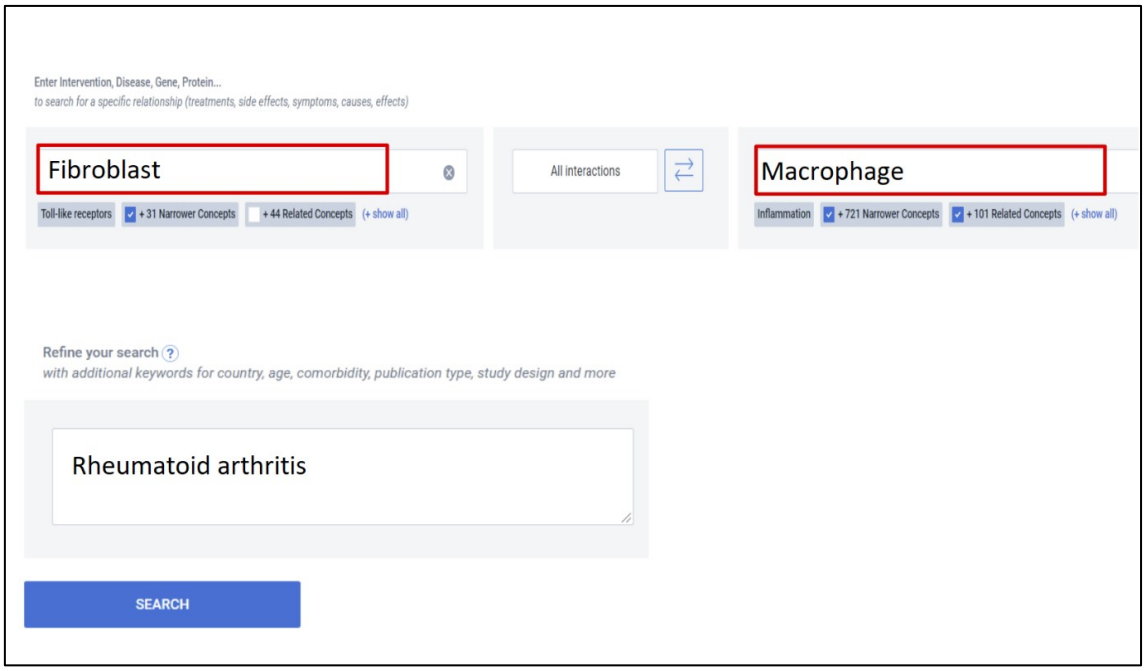


Figure 29. Screenshot of the advanced search option in Causaly. We selected fibroblast and macrophage as cell-types and rheumatoid arthritis as the condition of interest in the concept boxes. We didn't specify any directionality.

We used PubMed as well to retrieve cell-cell communication between the cell-types of interest. Relevant keywords and key sentences like 'M1 macrophage fibroblast interactions in rheumatoid arthritis', 'macrophage Th1 activation in RA', 'M2 macrophage Th1 inhibition RA' among many others were used to filter the literature abstracts and studies in PubMed. Along with it, peer review articles concerning RA and their bibliography were searched, and information was mined.

To identify additional ways of cellular communication, we used CellPhoneDB database (Figure 30). It is a publicly available repository of manually curated receptors, ligands, and their interactions (179). Subunit architecture (i.e., the subunits that form the quaternary structure of a protein) is included for both ligands and receptors, representing heteromeric complexes accurately. It integrates new manually reviewed interactions with evidenced roles in cell-cell communication together with existing datasets that pertain to cellular communication. CellPhoneDB stores a total of 978 proteins involved in 1,396 interactions. 501 of them are secreted proteins and 585 of them are membrane proteins. Out of all proteins stored in CellPhoneDB, 466 are heteromers.

We used CellPhoneDB database to identify the intercellular interactions occurring through ligands and receptors expressed by the four cell-types represented in the RA-Atlas: the RA M1 macrophage, the RA M2 macrophage, the RA fibroblast and the RA Th1. From their corresponding molecular interaction maps, we extracted

the proteins they release from the secreted components maps' compartments, and the receptors and transmembrane proteins they express from the cytoplasmic membrane maps' compartments. In text files, we combined these receptors and ligands in a way that covers the following bidirectional cell–cell interaction pairs: M1 macrophage-fibroblast, M1 macrophage-Th1, M2 macrophage-fibroblast, M2 macrophage-Th1, and fibroblast-Th1. Regarding the M1 macrophage-fibroblast pair, for instance, we generated a first text file containing all the ligands expressed in the M1 macrophage map and all the receptors expressed in the fibroblast map. A second file was also generated and consists of all the ligands represented in the fibroblast map and all the receptors expressed in the M1 macrophage map. This way, both outward and inward communications are considered. Using CellPhoneDB database, we retrieved all the manually curated interactions that involve both the ligands and receptors present in each generated text file.

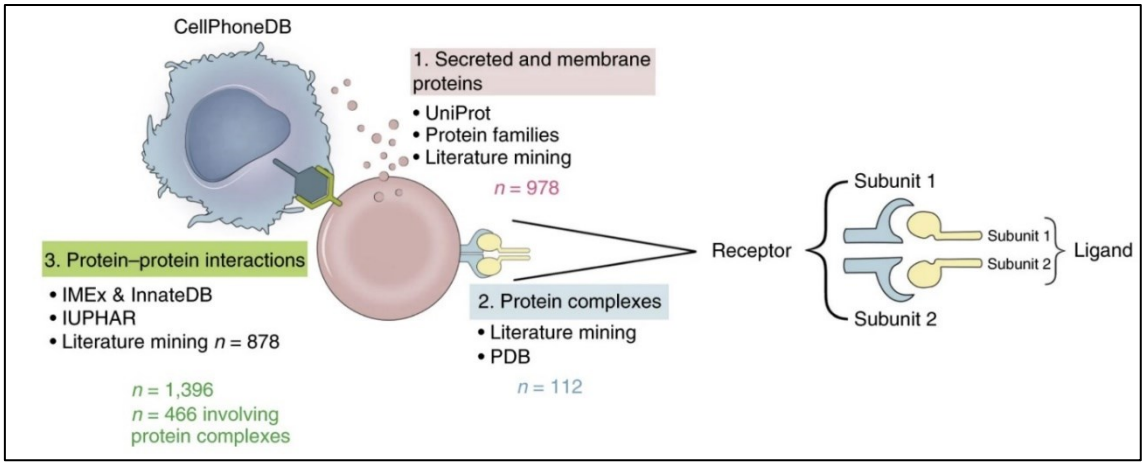


Figure 30. Overview of CellPhoneDB dataset. Figure taken from Efremova et al., 2020 (179).

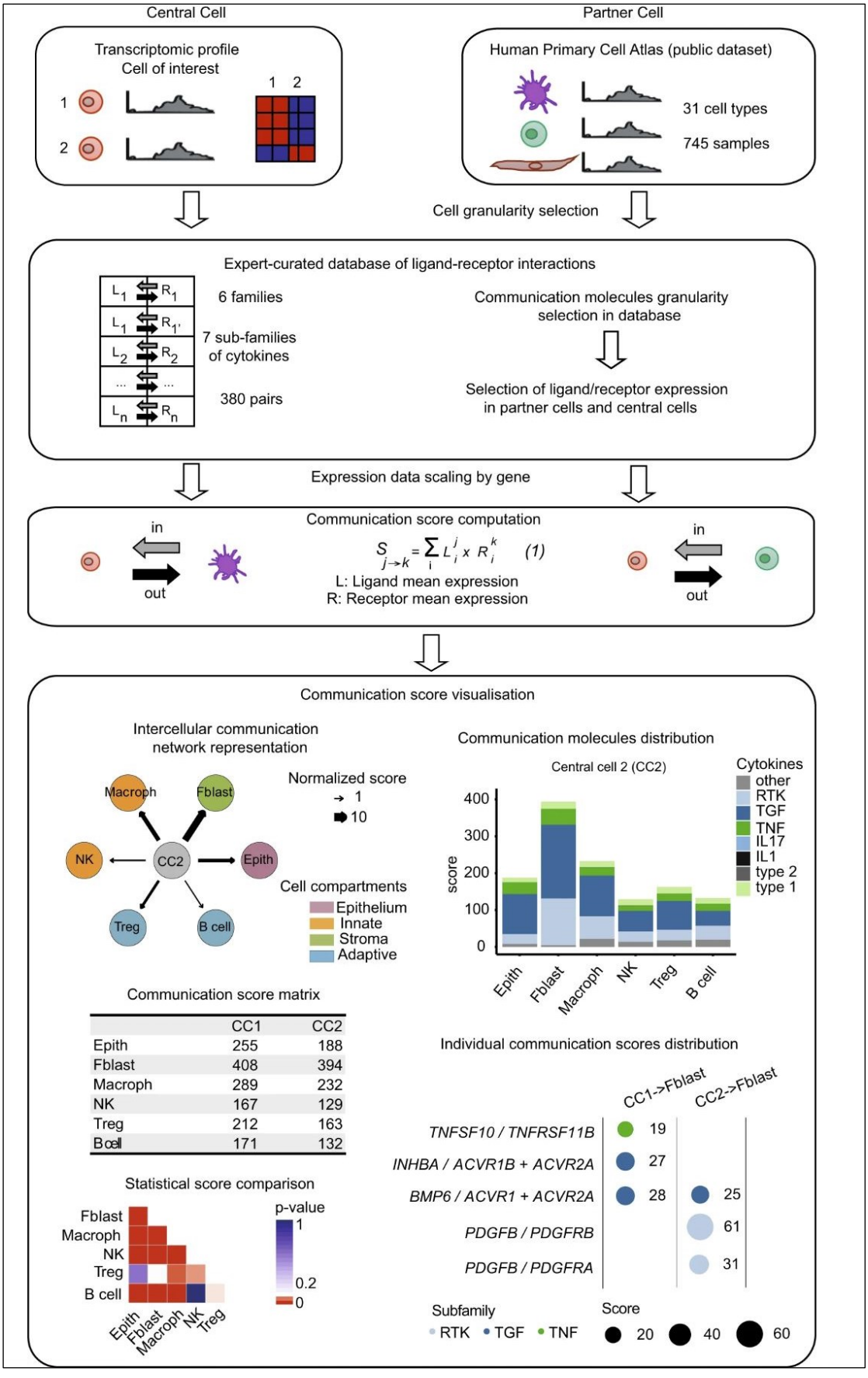
5.2.2 Identification of cell-cell interactions using omics data

Another approach to identify cell-cell interactions is to use transcriptomic data. This approach consists of a resource of intercellular interactions prior knowledge and a method to predict potential cell-cell communication events. Events are represented as a one-to-one interaction between a transmitter and receiver protein expressed by the source and target cell type. Events are predicted by estimating the likelihood of crosstalk based on the expression level of the transmitter and the receiver in the source and target cell type, respectively (181). In this section, we present ICELLNET and DiSiR tool, both were applied on several omics datasets to identify intercellular interactions between the cell-types of interest.

5.2.2.1 ICELLNET framework

First, we used ICELLNET (178), a versatile computational framework to infer cell-to-cell communication from a wide range of microarray, bulk, and single-cell transcriptomic datasets. It integrates an extensive database of 380 ligand-receptor interactions collected from the literature and public databases. Expert manual curation was carried out based on an in-depth investigation of original publications in the literature. Unlike other databases, it doesn't include inferred interactions predicted by protein-protein interactions and it combines biological relevance and experimental validation. ICELLNET is the only database that categorizes expected interactions into biological families: Growth factors, Cytokines, Chemokines, immunological Checkpoints, Notch signaling, and Antigen binding. ICELLNET, like CellPhoneDB, considers the various subunits of ligands and receptors by adding logical principles for protein subunit co-expression. A rigorous examination of cytokine connections identified 14 interactions that were present in ICELLNET but not in CellPhoneDB. Furthermore, Although ICELLNET contains a small number of interactions in comparison to other existing databases, it is very specific and exhaustive for cytokine interactions that play a critical role in RA development.

ICELLNET makes use of its database to infer cell-cell communication via a computational pipeline in R (Figure 31). ICELLNET takes as input the transcriptome profile of the cells of interest. These profiles may be from the same dataset, or from different transcriptomic datasets. All genes or only differentially expressed genes from each transcriptome profile may be used, and no filtering threshold is applied to gene expression. Using the ICELLNET database, the genes coding for ligands/receptors are chosen from all 380 interactions. Since cell-to-cell communication is directional, ligand expression from the central cell, and receptor expression from the partner cell are considered in order to assess outward communication. Conversely, selected receptor expression from the central cell, and ligand expression from partner cell are considered in order to assess inward communication. To avoid a communication score dominated by highly expressed genes, expression levels for each gene are scaled by the maximum gene expression in the dataset. The score of an individual ligand–receptor interaction is computed as the product of their expression levels by the respective source (central) and target (partner) cell. Whenever a communication molecule (ligand or receptor or both) was not expressed by a cell, the score of this particular interaction was set to zero.



	CC1	CC2
Epith	255	188
Fblast	408	394
Macroph	289	232
NK	167	129
Treg	212	163
Bcell	171	132

Figure 31. ICELLNET pipeline to study intercellular communication from cell transcriptional profiles. Figure taken from Noël et al., 2021 (178)

5.2.2.2 Dimer Signal Receptor Analysis (DiSiR) framework

DiSiR (180) is a permutation-based software framework for investigating how individual cells interact with one another by analyzing multi-subunit ligand–multi-subunit receptor interactions from scRNA-seq data, not only for interactions listed in curated databases of ligand-receptor interactions, but also for interactions not listed in these databases.

As described in Figure 32, DiSiR uses a single-cell gene expression matrix, cell type annotations, and a user–defined putative list of ligand–receptor interactions at subunit level as inputs. For each ligand/receptor subunit, it calculates cell-type A and B average expressions over all the cells within cell-type A and cell-type B respectively. If a ligand or receptor subunit is not expressed in at least one of the resulting average expressions, then it will be removed from the input interactions for the cell type A–cell type B communication. Then, it generates k sets of shuffled cell-type A and cell-type B average expressions' by randomly permuting the cell-type assignment of all cells and re-calculating the average expressions of the ligands/receptors subunits over all the cells belonging to the new random cell-types. The default value for K is 1000.

DiSiR identifies the ligand 'L'–receptor 'R' interaction between cell type A and B based on the products of log-normalized expressions of 'L' subunits by cell type A and log-normalized expressions of 'R' subunits by cell type 'B'. This interaction is considered as significant if these products are significantly higher than the result of random shuffling of cell type labels for all 'L–R' subunit combinations, according to a two-sided non-parametric permutation test (results with a P-value < 0.05 are considered significant). DiSiR uses a filtering approach as well to discard weak interactions that might be associated with noise by thresholding the fraction of cells expressing each ligand/receptor subunit within its corresponding cell type. A LR interaction between two cell types is then retained if this score is higher than a certain threshold level *Th* for all ligand/receptor subunits.

To run DiSiR, we used ICELLNET database as a putative list of ligand–receptor interactions, some at the subunit level. We ran DiSiR with k=1000 (permutations) and a *Th* threshold at zero, the default values.

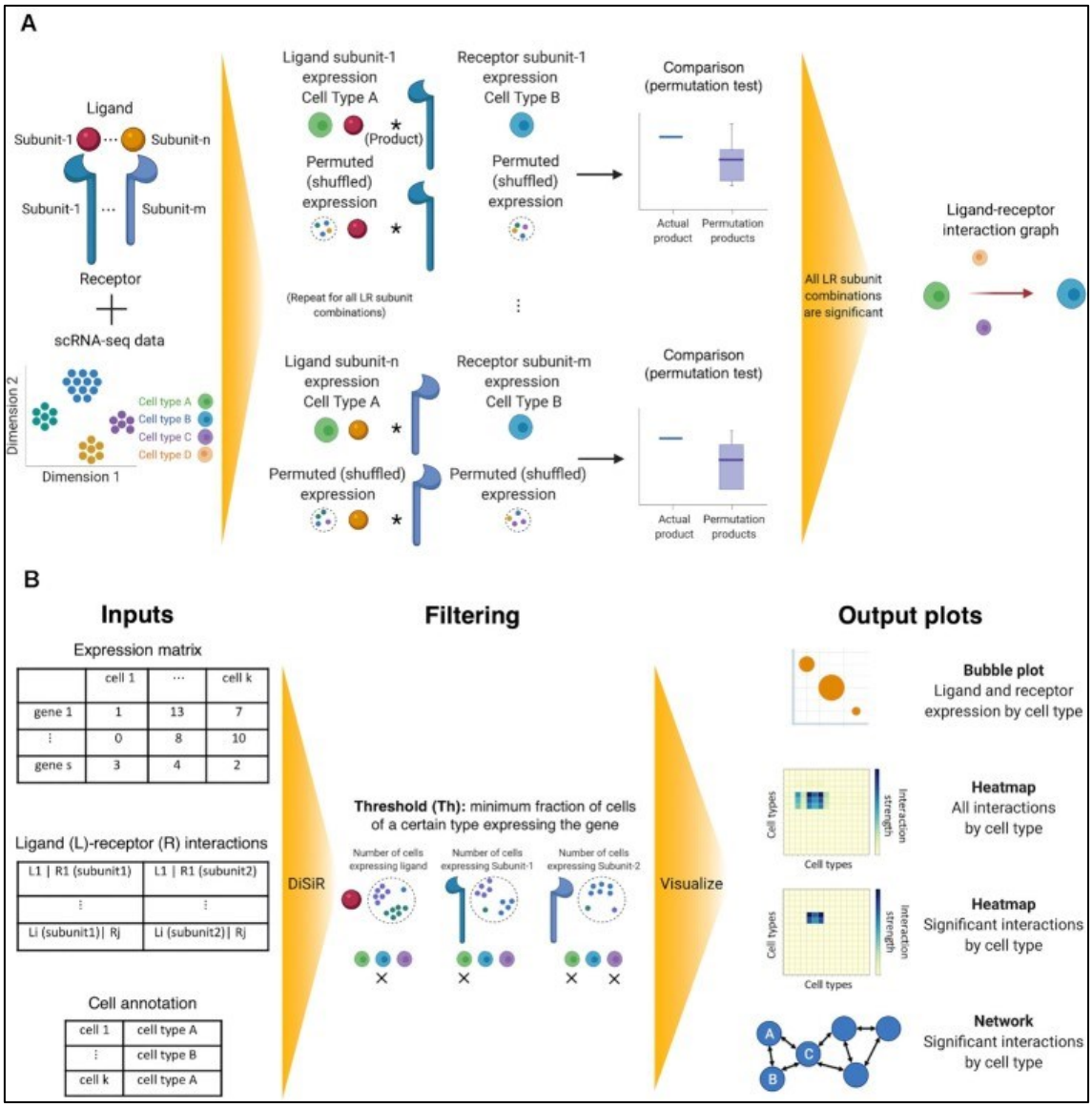


Figure 32. Schematic overview and illustration of DiSiR framework taken from Vahid et al., 2023 (180). (A) DiSiR uses as input 1) a single-cell gene expression matrix, 2) cell type annotations, and 3) a list of putative ligands–receptor interactions at subunit level. DiSiR then characterizes the ligand 'L'–receptor 'R' interaction between cell types 'A' and 'B' based on the products of expressions of 'L' subunits by cell type 'A' and expressions of 'R' subunits by cell type 'B'. This interaction is identified by DiSiR as a significant interaction if these products are significantly higher than random shuffling results for all 'L–R' subunit combinations. (B) Diagram that shows the workflow of data processing, filtering, and visualization of the obtained results in DiSiR. Weak interactions are filtered out if the fraction of cells expressing the ligand (or receptor) within its corresponding cell type is less than a threshold equation Th. DiSiR also provides an interactive graph-based representation module and heatmap/bubble plots to show the resulting interactions for complex signaling pathways at different level.

5.2.2.3 Selected omics datasets

To cover both outward and inward communication taking place between the four cell-types of interest, we selected three single cell RNA-seq datasets. Detailed descriptions of these datasets are available in Chapter 3. Description of data and data analysis. In this section we illustrate how we utilize them to infer cell-cell communication using both ICELLNET and DiSiR tools.

5.2.2.3.1 Bidirectional interactions between RA macrophage and RA fibroblast

Table 7 summarizes the pairs of gene expression datasets that we used to infer cell-cell communication between RA macrophages and RA fibroblasts cell-types.

Table 7. Pairs of datasets used to infer inward and outward cell-cell communication between RA macrophages and RA fibroblast.

Pair of datasets	RA macrophage/monocyte cells source	RA fibroblast cells source
1	SDY998	SDY998
2	E_MTAB_8322	SDY998
3	SDY998	GSE109449
4	E_MTAB_8322	GSE109449

Synovial samples of such cell-types are not available within the same dataset. The SDY998 dataset was the only one containing synovial fibroblasts (1844 cells) and synovial monocytes (750) (pair 1 in Table 7). We used this dataset to identify intercellular interactions, first with RA monocytes as sending cells and subsequently with RA fibroblasts as sending cells. However, because this dataset only contains monocytes and not macrophages, we cross-validated the results using additional pairs of datasets to ensure reliable and macrophage-specific results.

We combined cells that come from different experimental designs (see pair 2, 3, and 4 in Table 7). To reduce the resulting batch effect, we used ComBat (290), one of the most popular batch effect adjustment methods when the effects come from known sources. ComBat was originally developed for microarray gene expression data but had been successfully employed on single cell RNA-seq data (291). ComBat models the batch effect and then subtracts out the modeled effect in the gene expression matrix, creating a batch effect free dataset. To model the batch effect, ComBat assumes that in the presence of batch effect, the mean and variance of gene expression demonstrate systematic differences across batches on a standardized scale and follow a linear model. Expression data is first stand-

ardized so that all genes have similar means and variances. Thereafter, the standardized data is fitted to standard distributions using a Bayesian approach to estimate the batch effects present. The computed batch-effect estimators are then used to correct the original expression matrix. There are two modes to run ComBat, which are parametric and non-parametric adjustments to correct for batch effects. Parametric mode includes scale adjustments whereas non-parametric mode only corrects the mean of the batch effect.

To correct the batch effect in the pairs of independent datasets, we used the *sva* package that supports direct adjustment for known batch effects with the ComBat function (292). To do so, we pass the full gene expression matrix that comprises RA monocyte/macrophage cells and RA fibroblast cells. Batch variables are passed as a separate argument (batch) to the function. The output is a set of corrected measurements, where batch effects have been removed. Plots shown in Figure 33 represent visual outputs of the ComBat parametric mode applied on pair n°4 in Table 7. If the distribution of the batch effect is normal, black and red plots should overlap. In this case, we can see that the black and red plots do not overlap. The same results were obtained with pairs 2 and 3 in Table 7. Therefore, the non-parametric mode was applied on all pairs of independent datasets.

Batch corrected datasets were then utilized to run both ICELLNET and DiSiR tools, first with RA monocytes/macrophages as sending cells, then with RA fibroblasts as sending cells.

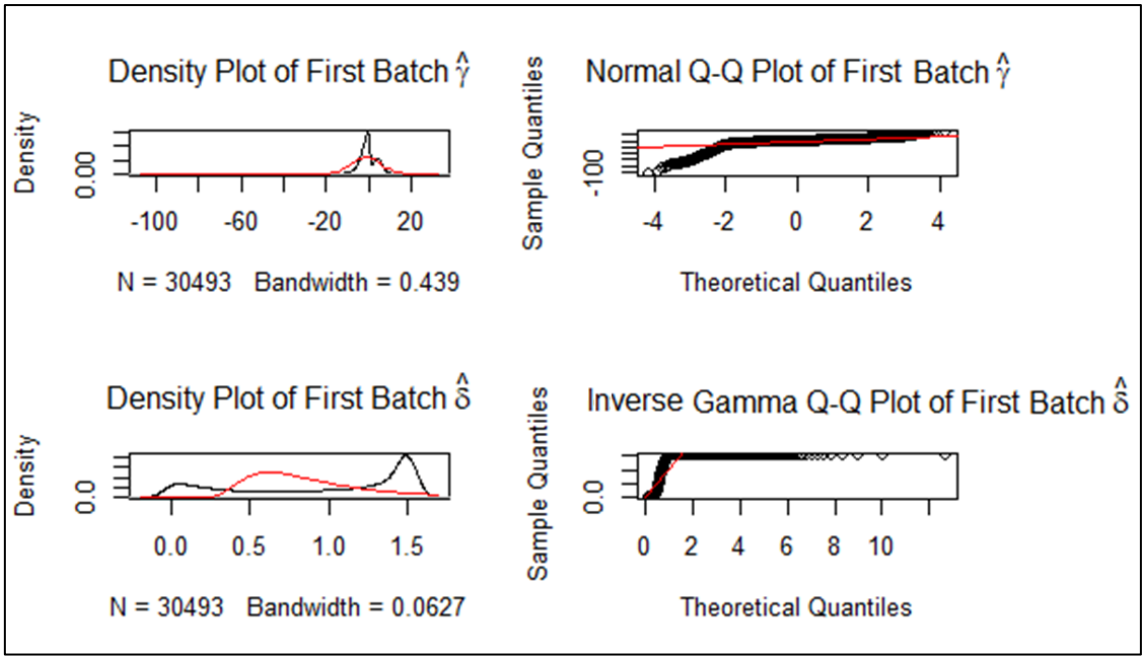


Figure 33. Output plots of ComBat on E_MTAB_8322 and GSE109449 datasets combined. The left plots are density estimates, and the right plots are QQ plots. Top

plots represent the means and bottom plots represent the variances. The black plots represent the mean and variance distributions of the combined datasets. The red plots represent the theoretical normal distribution of the mean and variance. If the distribution of the batch effect is normal, black and red plots should overlap. In this case, red and black plots do not overlap, therefore, non-parametric mode of ComBat should be used.

5.2.2.3.2 Bidirectional interactions between RA Th1 and RA fibroblast

Table 8 summarizes the pairs of gene expression datasets we used to infer cell-cell communication between RA fibroblasts and RA CD4+ Th1 cell-types.

Table 8. Pairs of datasets used to infer inward and outward cell-cell communication between RA fibroblasts and RA CD4+ Th1.

Pair of datasets	RA fibroblast cells source	RA CD4+ Th1 cells source
1	SDY998	SDY998
2	GSE109449	SDY998

RA synovial samples of such cell types, particularly T lymphocytes, are scarce. The SDY998 dataset was the only publicly available one that contains synovial T cells coming from RA patients. First, we applied clustering techniques to identify the CD4+ Th1 subtype (see Chapter 3. Description of data and data analysis). Then, cells within the Th1 cluster were paired with fibroblast cells (pair 1 in Table 8). As both cell-types' samples come from the same dataset, no batch effect adjustment was required.

We also used GSE109449, an RNA-seq Single Cell dataset containing synovial fibroblasts from RA patients (see Chapter 3. Description of data and data analysis) that we combined with the Th1 cluster from the SDY998 dataset (pair 2 in Table 8). To correct the batch effect, we used the sva package with the non-parametric mode of ComBat function (292) as described previously.

Both ICELLNET and DiSiR were applied on pair 1 and pair 2 in Table 8, first with RA CD4+ Th1 as sending cells, then with RA fibroblasts as sending cells.

5.2.2.3.3 Bidirectional interactions between RA CD4+ Th1 and RA macrophage

Table 9 summarizes pairs of gene expression datasets that we used to identify cell-cell communication between RA macrophage and RA CD4+ Th1 cell-types.

Table 9. Pairs of datasets used to infer inward and outward cell-cell communication between RA macrophages and RA CD4+ Th1.

Pair of datasets	RA monocyte/macrophage cells source	RA CD4+ Th1 cells source
1	SDY998	SDY998
2	E_MTAB_8322	SDY998

Cells within the Th1 cluster from the SDY998 dataset were paired with the 750 monocytes from the same dataset (pair 1 in Table 9). Because this dataset only contains monocytes and not macrophages, we cross-validated the results using additional pairs of datasets to ensure reliable and macrophage-specific results. We combined RA macrophage and RA Th1 samples coming from different experimental designs (see pair 2 in Table 9). We adjusted the batch effect using ComBat method through the *sva* package.

ICELLNET and DiSiR were then utilized to infer cell-cell interactions, considering RA CD4+ Th1 as sending cells, then as receiving cells.

5.2.3 Cell-cell interactions filtering criteria

Two approaches were used to identify bidirectional cell-cell communication between RA fibroblast, RA macrophage and RA Th1 cell types. The first one is based on literature and databases mining using several tools like Causaly and CellPhoneDB database. The second approach consists of integrating cell- and disease-specific omics datasets through two different dedicated tools: ICELLNET and DiSiR.

These different approaches, tools and datasets produced different outcomes. To cross-validate the results, we put in place a list of filtering criteria regarding the previously identified intercellular interactions. Indeed, among the interactions identified between each directional cell–cell interaction pair:

- Only interactions with statistically significant p-values (FDR threshold equal to 0.05) in DiSiR were kept.
- Only interactions identified with both ICELLNET and DiSiR were kept.
- Only interactions identified with at least two different pairs of datasets, or two different approaches (literature mining- and omics data-based) were kept.
- Regarding interactions coming from published literature and databases but not inferred from omics datasets, only interactions coming

from at least two different sources of information (different published articles or different databases) were kept.

5.2.4 Construction of the RA multicellular map in CellDesigner

The previously filtered cell-cell interactions were used to connect the RA cell-specific molecular interactions maps described in Chapter 4. Construction of RA cell-specific molecular interaction maps. We used CellDesigner to link interacting ligands and receptors as described in Figure 34. Ligands produced by the sending cell (secreted components compartment in the cell-specific maps) are transported via Transport arrows to the extracellular space of the receiving cell. The ligand will then bind to its corresponding receptor and form a ligand/receptor complex that can induce signal transduction in the cytoplasm of the receiving cell. When cell-cell interactions occur through cell-cell contact, Heterodimer Complex Association is used instead in CellDesigner to bind transmembrane proteins with their corresponding receptors.

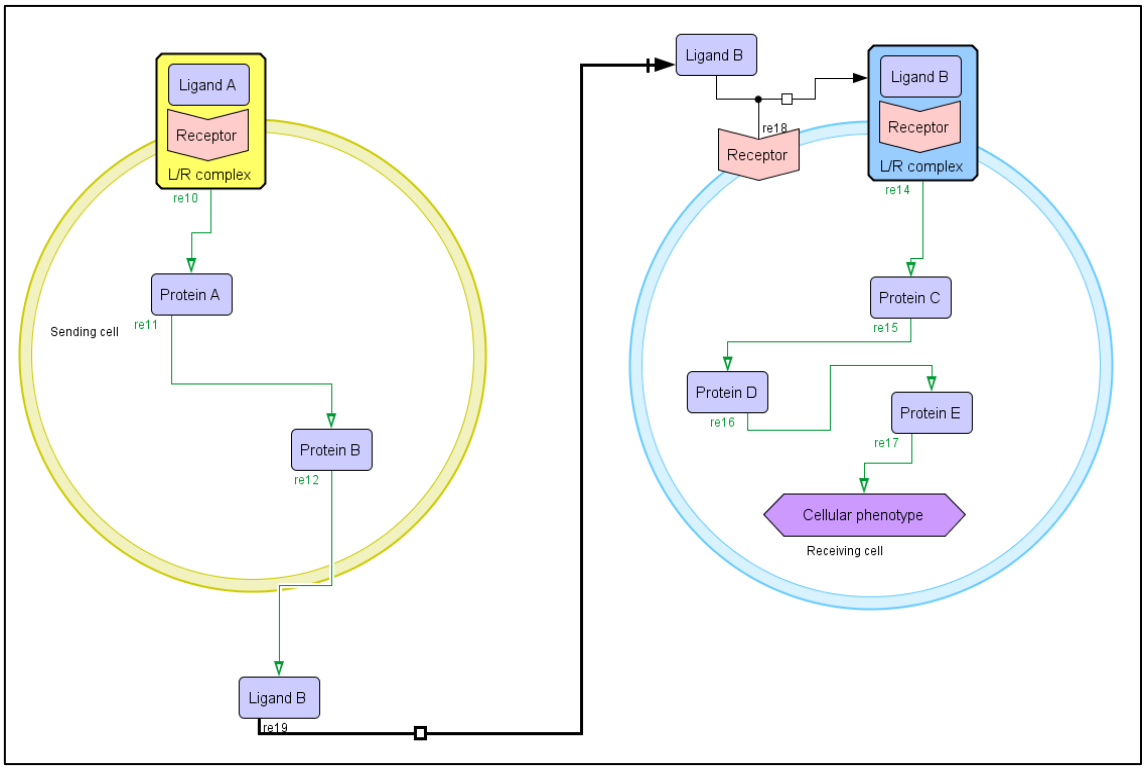


Figure 34. Schematic representation of the intercellular interactions in CellDesigner. The black arrow represents the transport of the secreted ligand (Ligand B) from the sending cell (in yellow) to the extracellular space of the receiving cell (in blue) where it binds to its corresponding receptor. The ligand/receptor complex will then transduce the signal via a series of biochemical reactions in the cytoplasm.

For some of the identified intercellular interactions, ligands, and receptors together with their upstream and downstream regulations were already represented in the RA cell-specific maps. In this case, only transport/complex association arrows were added to connect both maps. For some other intercellular interactions, ligands and/or receptors were missing from the maps. Therefore, we had to search in the published literature and databases for the signaling pathways upstream of these ligands and/or downstream of these receptors. The identified signaling pathways went then through an extensive manual curation to ensure that most of their components were expressed in a diseases and cell-specific manner. For that, we used published articles as well as the previously generated and relevant cell specific DEGs lists (see Chapter 3. Description of data and data analysis). Once curated, the missing ligands and/or receptors together with their corresponding signaling pathways were added to the relevant map(s) in CellDesigner and transport arrows were used to connect the interacting cell-types.

Identified cell-cell communication from published literature and databases that involve macrophages is often not specific to the M1 or M2 phenotype as RA macrophages are usually studied as a single cell-type. This is especially true for the cell-cell interactions obtained from omics datasets. Indeed, no dataset encompassing synovial cells of RA M1 and RA M2 macrophages was found. Instead, RA macrophage synovial samples were employed. Therefore, such cell-cell interactions cannot be directly mapped on the RA M1 macrophage and RA M2 macrophage molecular interactions maps. To overcome this limitation, first we used the previously identified M1 and M2 gene signatures coming from DEA outcomes and Msig gene signatures (Chapter 5. Construction of the RA multicellular map. We also went through an extensive literature search to retrieve information regarding the expression of the biomolecules missing from the maps in each phenotype. The percentage of shared components between each missing pathway in the maps and the M1 and M2 signatures was calculated and, based on these percentages, cell-cell interactions were assigned to the M1 or M2 phenotype.

5.3 Results

In this section, we present the identified cell-cell interactions in literature and databases as well as the ones inferred from omics datasets. These interactions were filtered using the previously described criteria.

5.3.1 Bidirectional interactions between RA macrophage and RA fibroblast

Table 10 summarizes the bidirectional intercellular interactions between RA macrophage and RA fibroblast. Each interaction is associated with diverse resources used to identify it: published literature in PubMed, the internal database of CellPhoneDB and the various pairs of omics datasets used to infer it.

Table 10. Cell-cell interactions between RA macrophage and RA fibroblast in the synovium.

Ligand/ receptor com-plexes	Sending cell	Receiving cell	Source		
			Literature mining (PMID)	CellPhoneDB database	Omics datasets pairs
CXCL1/ CXCR1	Fibroblast	M1 macrophage	33087182	Absent	(SDY998, SDY998) (GSE109449, SDY998)
CCL2/ CCR2	Fibroblast	M1 macrophage		Present	(SDY998, SDY998)
CSF2/ CSF2RA+ CSF2RB	Fibroblast	M1 macrophage	1700731 27813830 29997624	Present	
TNFSF11/ TNFRSF11A	Fibroblast	M1 macrophage		Absent	(GSE109449, SDY998) (SDY998, SDY998) (GSE109449, E_MTAB_8322)
TNFA/ TNFRSF1A	Fibroblast	M1 macrophage	29997624 28807007 27813830 25057003	Present	
IL1B/IL1R	Fibroblast	M1 macrophage	28807007 26883280 29997624	Present	(GSE109449, SDY998) (SDY998, E_MTAB_8322)
JAG1/ NOTCH1	Fibroblast	M1 macrophage		Present	(SDY998, SDY998) (GSE109449, E_MTAB_8322)
GAL/GALR2	Fibroblast	M1 macrophage		Absent	(SDY998, SDY998) (GSE109449, SDY998)
COL4A4/ ITGA1+ ITGB1	Fibroblast	M1 macrophage		Present	(GSE109449, SDY998) (SDY998, SDY998)
HLA-B/ LILRB1	Fibroblast	M1 macrophage		Absent	(SDY998, SDY998) (SDY998, E_MTAB_8322) (GSE109449, E_MTAB_8322)
IFNE/ IFNAR1+ IFNAR2	Fibroblast	M1 macrophage		Present	(SDY998, E_MTAB_8322)
IL6/IL6R+ IL6ST	Fibroblast	M1 macrophage	29997624 35663975	Present	(SDY998, SDY998)
IL12/ IL12RB1+ IL12RB2	Fibroblast	M1 macrophage		Present	(GSE109449, SDY998)
HLA-B/	Fibroblast	M2		Absent	(SDY998, SDY998)

LILRB1		macrophage			(SDY998, E_MTAB_8322) (GSE109449, E_MTAB_8322)
VEGFC/ VEGFR3	Fibroblast	M2 macrophage		Present	(GSE109449, SDY998) (SDY998, SDY998)
GAL/GALR2	Fibroblast	M2 macrophage		Absent	(SDY998, SDY998) (GSE109449, SDY998)
SEMA4D/ PLXNB2+ MET	Fibroblast	M2 macrophage		Present	(GSE109449, E_MTAB_8322)
EFNB1/ EPHB1	Fibroblast	M2 macrophage		Present	(SDY998, SDY998)
WNT5B/ FZD1+LRP5	Fibroblast	M2 macrophage	30022048	Present	(GSE109449, SDY998)
CSF1/CSF1R	Fibroblast	M2 macrophage	27383913	Present	(SDY998, E_MTAB_8322)
IL34/ CSFR1R	Fibroblast	M2 macrophage		Present	(GSE109449, E_MTAB_8322) (GSE109449, SDY998)
TGFB/ TGFBR1+ TGFBR2	Fibroblast	M2 macrophage	29997624 31068444	Present	(GSE109449, SDY998)
PRL/PRLR	Fibroblast	M2 macrophage		Absent	(SDY998, SDY998) (GSE109449, SDY998)
COL4A4/ ITGB1+ ITGA1	Fibroblast	M2 macrophage		Present	(GSE109449, SDY998) (SDY998, SDY998)
SEMA3A/ PLXNA1	Fibroblast	M2 macrophage		Present	(GSE109449, SDY998) (SDY998, E_MTAB_8322)
GAS6/ MERTK	Fibroblast	M2 macrophage		Absent	(SDY998, SDY998) (GSE109449, SDY998)
CD40L/ CD40	M1 macrophage	Fibroblast		Absent	(SDY998, GSE109449) (SDY998, SDY998)
JAG1/ NOTCH3	M1 macrophage	Fibroblast		Present	(SDY998, SDY998)
EDA/EDA2R	M1 macrophage	Fibroblast		Absent	(SDY998, GSE109449) (E_MTAB_8322, SDY998)
ICAM1/ ITGB2+ ITGAL	M1 macrophage	Fibroblast		Present	(SDY998, SDY998) (E_MTAB_8322, SDY998)
IL18/IL18R	M1 macrophage	Fibroblast		Present	(E_MTAB_8322, GSE109449) (E_MTAB_8322, SDY998) (SDY998, GSE109449)
IL1B/ IL1R1	M1 macrophage	Fibroblast	27383913 31178859 31068444	Present	(E_MTAB_8322, GSE109449)

TNFA/ TNFRSF1B	M1 macrophage	Fibroblast	27383913 31178859	Present	
CXCL10/ CXCR3	M1 macrophage	Fibroblast		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998) (E_MTAB_8322, GSE109449)
IFNG/ IFNGR1+ INFGR2	M1 macrophage	Fibroblast		Present	(SDY998, SDY998) (SDY998, GSE109449) (E_MTAB_8322, SDY998)
FASL/FAS	M1 macrophage	Fibroblast		Absent	(SDY998, SDY998) (SDY998, GSE109449)
CCL5/CCR5	M1 macrophage	Fibroblast		Present	(SDY998, GSE109449)
HBEGF/ EGFR	M1 macrophage	Fibroblast	31068444	Present	
AREG/EGFR	M1 macrophage	Fibroblast		Present	(SDY998, SDY998) (E_MTAB_8322, GSE109449)
SEMA4A/ PLXNB1+ MET	M1 macrophage	Fibroblast	26303122	Present	(E_MTAB_8322, SDY998)
FASL/FAS	M2 macrophage	Fibroblast		Absent	(SDY998, SDY998) (SDY998, GSE109449)
CCL18/ PITPNM3	M2 macrophage	Fibroblast		Absent	(SDY998, GSE109449) (SDY998, SDY998) (E_MTAB_8322, GSE109449)
EDA/ EDA2R	M2 macrophage	Fibroblast		Absent	(SDY998, GSE109449) (E_MTAB_8322, SDY998)
JAG1/ NOTCH3	M2 macrophage	Fibroblast		Present	(SDY998, SDY998)
IL10/ L10RA+ IL10RB	M2 macro-phage	Fibroblast		Present	(SDY998, GSE109449) (SDY998, SDY998)
PDGFC/ PDGFRB	M2 macrophage	Fibroblast		Present	(SDY998, SDY998) (E_MTAB_8322, GSE109449)
HBEGF/ EGFR	M2 macrophage	Fibroblast	31068444	Present	
COL4A3/ ITGA4+ ITGB1	M2 macrophage	Fibroblast		Present	(SDY998, SDY998) (E_MTAB_8322, SDY998) (SDY998, GSE109449)
SEMA7A/ ITGA4+ ITGB1	M2 macrophage	Fibroblast		Absent	(SDY998, GSE109449) (SDY998, SDY998)

VEGFA/ VEGFR	M2 macrophage	Fibroblast		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)
TGFB1/ TGFBR1	M2 macrophage	Fibroblast		Present	(SDY998, GSE109449) (E_MTAB_8322, GSE109449)
SEMA4A/ PLXNB1+ MET	M2 macrophage	Fibroblast	26303122	Present	(E_MTAB_8322, SDY998)
CD40L/ CD40	M2 macrophage	Fibroblast		Absent	(SDY998, GSE109449) (SDY998, SDY998)

5.3.2 Bidirectional interactions between RA macrophage and RA CD4+ Th1

Table 11 summarizes the bidirectional intercellular interactions between RA macrophage and RA CD4+ Th1. Each interaction is associated with diverse resources used to identify it: published literature in PubMed, the internal database of CellPhoneDB and the various pairs of omics datasets used to infer the interaction.

Table 11. Cell-cell interactions taking place between RA macrophage and RA CD4+ Th1 in the synovium.

Ligand/receptor complexes	Sending cell	Receiving cell	Source		
			Literature mining	CellPhoneDB database	Omics dataset(s)
IFNG/ IFNGR1+ IFNGR2	Th1	M1 macrophage	25329467	Present	(SDY998, E_MTAB_8322)
CXCL13/ACKR4	Th1	M1 macrophage		Absent	(SDY998, SDY998) (SDY998, E_MTAB_8322)
FASL/FAS	Th1	M1 macrophage		Absent	(SDY998, SDY998) (SDY998, E_MTAB_8322)
COL4A5/ ITGA1+ITGB1	Th1	M1 macrophage		Present	(SDY998, SDY998) (SDY998, E_MTAB_8322)
TNFSF11/ TNFRSF11A	Th1	M1 macrophage	16220542	Absent	(SDY998, SDY998)
CD40L/ ITGB1+ITGA1	Th1	M1 macrophage		Absent	(SDY998, SDY998) (SDY998, E_MTAB_8322)
SEMA4A/ PLXNB2+MET	Th1	M2 macrophage		Present	(SDY998, E_MTAB_8322)
CXCL13/ACKR4	Th1	M2 macrophage		Absent	(SDY998, SDY998) (SDY998, E_MTAB_8322)
FASL/FAS	Th1	M2 macrophage		Absent	(SDY998, SDY998) (SDY998, E_MTAB_8322)

CD40L/ITGB1+ ITGA1	Th1	M2 macrophage		Absent	(SDY998, SDY998) (SDY998, E_MTAB_8322)
COL4A5/ ITGB1+ITGA1	Th1	M2 macrophage		Present	(SDY998, E_MTAB_8322)
TGFB1/TGFBR1+ TGFBR2	Th1	M2 macrophage		Present	(SDY998, SDY998)
IL12/IL12Rb1+ IL12Rb2	M1 macrophage	Th1	26635790	Present	
IFNG/IFNGR1+ IFNGR2	M1 macrophage	Th1		Present	(SDY998, SDY998) (E_MTAB_8322, SDY998)
CCL4/CCR5	M1 macrophage	Th1		Present	(E_MTAB_8322, SDY998)
CXCL10/CXCR3	M1 macrophage	Th1		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)
IL18/IL18R1+ IL18RAP	M1 macrophage	Th1	10562301	Present	
LGALS9/TIM3	M1 macrophage	Th1		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)
ICOSLG/ICOS	M1 macrophage	Th1		Present	(E_MTAB_8322, SDY998)
HLA-DP-DQ-DR/ TCR+CD3	M1 macro-phage	Th1	30915067	Absent	(SDY998, SDY998)
HLA-DP-DQ-DR /LAG3	M1 macrophage	Th1		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)
CXCL16/CXCR6	M1 macrophage	Th1	26635790	Absent	(E_MTAB_8322, SDY998)
CD40L/ITGAM+ ITGB2	M1 macrophage	Th1		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)
SEMA4A/PLXNB1	M1 macrophage	Th1		Present	(SDY998, SDY998)
JAG1/NOTCH1	M1 macrophage	Th1		Present	(SDY998, SDY998)
CD28/CD86	M1 macrophage	Th1	29868020	Present	
LTA/TNFRSF14	M1 macrophage	Th1		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)
CCL2/CCR2	M1 macrophage	Th1	26635790	Present	
COL4A3/ ITGA4+ITGB7	M2 macrophage	Th1		Present	(SDY998, SDY998)

					(E_MTAB_8322, SDY998)
ICOSLG/ICOS	M2 macrophage	Th1		Present	(E_MTAB_8322, SDY998)
HLA-DP-DQ-DR /TCR+CD3	M2 macrophage	Th1	30915067	Absent	(SDY998, SDY998)
HLA-DP-DQ-DR /LAG3	M2 macrophage	Th1		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)
CXCL16/CXCR6	M2 macrophage	Th1	26635790	Absent	(E_MTAB_8322, SDY998)
CD40L/ITGAM+ ITGB2	M2 macrophage	Th1		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)
SEMA4A/PLXNB1	M2 macrophage	Th1		Present	(SDY998, SDY998)
JAG1/NOTCH1	M2 macrophage	Th1		Present	(SDY998, SDY998)
CD28/CD86	M2 macrophage	Th1	29868020	Present	
MIF/CD74	M2 macrophage	Th1		Absent	(SDY998, SDY998) (E_MTAB_8322, SDY998)

5.3.3 Bidirectional interactions between RA fibroblast and RA CD4+ Th1

Table 12 summarizes the bidirectional intercellular interactions between RA fibroblast and RA CD4+ Th1. Each interaction is associated with diverse sources of information that enabled its identification: published literature in PubMed, the internal database of CellPhoneDB and the various pairs of omics datasets used to infer it.

Table 12. Cell-cell interactions taking place between RA CD4+ Th1 and RA fibroblast in the synovium.

Ligand/receptor complexes	Sending cell	Receiving cell	Source		
			Literature mining	CellPhoneDB database	Omics dataset(s)
CCL5/CCR5	Th1	Fibroblast		Present	(SDY998, SDY998)
COL4A5/ ITGA5+ITGB1	Th1	Fibroblast		Absent	(SDY998, SDY998) (SDY998, GSE109449)
TGFB1/TGFBR1	Th1	Fibroblast		Present	(SDY998, SDY998)
SEMA4A/ PLXNB1+MET	Th1	Fibroblast		Present	(SDY998, GSE109449)
CD40L/CD40	Th1	Fibroblast	35844494 15077296	Absent	(SDY998, SDY998) (SDY998, GSE109449)

ICAM1/ ITGB2+ITGAL	Th1	Fibroblast	27623446	Present	
TNFSF11/ TNFRSF11A	Th1	Fibroblast		Absent	(SDY998, SDY998) (SDY998, GSE109449)
IL18/IL18R	Th1	Fibroblast		Present	(SDY998, GSE109449)
CXCL13/CXCR3	Th1	Fibroblast		Absent	(SDY998, SDY998) (SDY998, GSE109449)
MIF/CXCR4	Th1	Fibroblast		Absent	(SDY998, SDY998) (SDY998, GSE109449)
IFNG/ IFNGR1+INFGR2	Th1	Fibroblast	32047926	Present	(SDY998, SDY998)
FASL/FAS	Th1	Fibroblast		Absent	(SDY998, SDY998) (SDY998, GSE109449)
CCL5/CCR5	Fibroblast	Th1		Present	(GSE109449, SDY998)
IL7R/IL2RG	Fibroblast	Th1	35844494	Absent	(SDY998, SDY998) (GSE109449, SDY998)
TNFA/TNFRSF1B	Fibroblast	Th1		Present	(SDY998, SDY998)
GAS6/MERTK	Fibroblast	Th1		Absent	(SDY998, SDY998) (GSE109449, SDY998)
CD84/CD84	Fibroblast	Th1		Absent	(SDY998, SDY998) (GSE109449, SDY998)
CCL2/CCR2	Fibroblast	Th1		Present	(SDY998, SDY998)
MIF/CD74	Fibroblast	Th1		Absent	(SDY998, SDY998) (GSE109449, SDY998)
LTA/TNFRSF14	Fibroblast	Th1		Absent	(SDY998, SDY998) (GSE109449, SDY998)
JAG1/NOTCH1	Fibroblast	Th1		Present	(GSE109449, SDY998) (SDY998, SDY998)
SEMA4D/PLXNB1	Fibroblast	Th1		Present	(SDY998, SDY998) (GSE109449, SDY998)
VCAM1/ ITGA4+ITGB7	Fibroblast	Th1	27623446 35844494	Present	
ICOSLG/ICOS	Fibroblast	Th1		Present	(SDY998, SDY998)
CXCL10/CXCR3	Fibroblast	Th1		Absent	(SDY998, SDY998) (GSE109449, SDY998)
IL10/	Fibroblast	Th1		Present	(SDY998, SDY998)

IL10RA+IL10RB					
LGALS9/TIM3	Fibroblast	Th1		Absent	(SDY998, SDY998) (GSE109449, SDY998)

5.3.4 RA multicellular map

Figure 35 depicts the RA multicellular map in CellDesigner. A full-page landscape image of the RA multicellular map is provided in Supplementary figure 5. Due to the high complexity of the network, this image is not readable in a A4 size page and is displayed for illustrative purposes only. For a better visualization of the multicellular map, an online interactive version of it will be made available in the platform MINERVA. The link to this interactive version will be provided in the publication that we are currently preparing regarding the multicellular map construction.

The RA multicellular map includes the four cell-specific molecular interaction maps that we built in Chapter 5. Construction of the RA multicellular map, connected together via 118 intercellular interactions. The multicellular map is composed of 2232 components that interact with one another via 1461 reactions.

To easily identify the sending cell in each interaction, a color coding was applied to the displayed cell-cell interactions. Pink arrows indicate interactions where the M1 macrophage is the sending cell, blue arrows indicate interactions where the M2 macrophage is the sending cell, yellow arrows indicate interactions where the fibroblast is the sending cell, and purple arrows indicate interactions where the CD4+ Th1 is the sending cell.

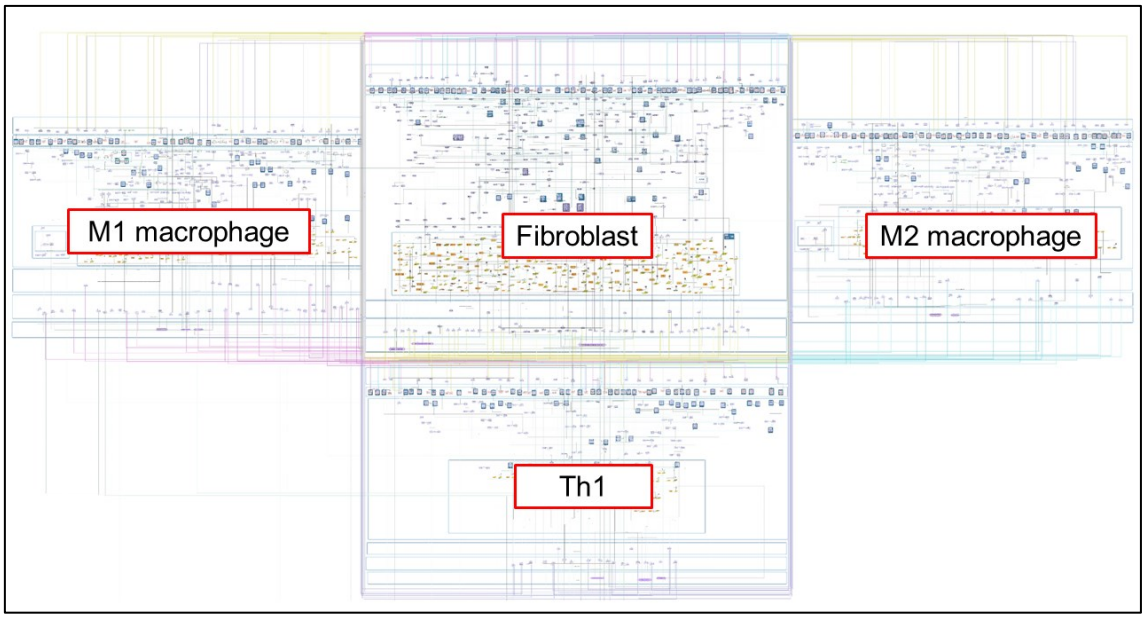


Figure 35. The RA multicellular map in CellDesigner. Red squares indicate the cell-type of each cell-specific molecular interaction map. Arrows in pink, yellow, blue and purple indicate cell-cell communication.

5.4 Discussion

The study of cell-cell interactions in the RA synovium offers an interesting and thorough picture of RA pathogenesis. Current biologic and non-biologic medicines have been shown to target some of these interactions in ways that are still unknown. In addition, safer and more specific approaches to treatment could flow from the discovery of cell-cell interaction pathways (92). Therefore, a more complete multicellular representation of the RA joint is needed.

The RA-Atlas (see Chapter 4. Construction of RA cell-specific molecular interaction maps) consists of cell-specific maps recapitulating molecular intracellular mechanisms involved in RA pathogenesis. These maps represent a great source of information but still describe a biological system where cells are in isolation from each other, cells that are in fact highly interconnected in the RA synovium. In this direction, we built the first RA multicellular map, a manually curated and enriched with extensive omics data analysis representation of both inter- and intracellular molecular mechanisms involved in RA's pathogenesis.

Going one step further, we were able to connect the RA cell-specific maps together via the addition of bidirectional intercellular interactions that were identified using CellPhoneDB internal database. The number of identified intercellular interactions varied depending on the number of ligands and receptors present in each cell-specific map. Given that the RA CD4+ Th1 map is smaller than the other

maps, with less ligands and receptors to query, knowledge about CD4+ Th1 interactions was sparser and more difficult to obtain. Moreover, interactions identified in CellPhoneDb are merely based on the expression of interacting pairs of ligands and receptors by the synovial cells of interest. These interactions are not specific to a disease or a tissue. Therefore, the expression of these ligands and receptors does not imply that they may interact together in the RA synovium. Combining CellPhoneDB queries with experimental evidence coming from published literature and/or omics data analysis can overcome this limitation.

First, we searched for cell-cell communication in published articles that were associated with experimental evidence in a cell- and disease-specific manner. We used dedicated tools like Causaly and its underlying advanced artificial intelligence to accelerate the process. The amount of information we could gather from published literature varied depending on the cell type under investigation. Indeed, studies on the interactions of fibroblasts and macrophages with their surroundings were considerably easier to find than those on CD4+ Th1 contact with other cells, as published literature usually focus on the overall T lymphocytes population and not on specific subtypes.

Second, we inferred cell-to-cell communication from transcriptomic datasets. For each bidirectional cell–cell interaction pair, we used two dedicated tools, ICELLNET and DiSiR, that we applied on disease- and cell-specific datasets. Unfortunately, macrophage, fibroblast, and Th1 cells from RA synovium were hard to find and rarely available within the same dataset. The SDY998 was the only publicly available one that included synovial monocytes, fibroblasts and a T cell population that we subdivided into subtypes, including the Th1 one. As this dataset merely contains monocytes and not macrophages, we needed to cross validate the interactions that we identified with the monocyte samples using macrophage samples. To do so, we paired samples coming from independent studies, and reduced the resulting batch effect using ComBat method.

Due to the heterogeneity of the omics datasets, the inferred cell-cell communication differs from one dataset to another, even when these datasets describe the same biological condition (RA) and cell type. Moreover, using monocyte samples rather than macrophage samples for some parts of this work, as well as combining independent datasets, even after the correction of the resulting batch effect, adds additional noise to the inferred intercellular interactions.

To ensure reliable and consistent cell- and disease specific results, we filtered the identified interactions using stringent criteria. Only interactions identified with both ICELLNET and DiSiR with statistically significant p values (FDR threshold

equal to 0.05) were kept. Also, only interactions identified with at least two different pairs of datasets, or two different approaches (literature/database mining and omics data) were kept. Regarding interactions that were not inferred from the omics datasets that we analyzed, only those coming from at least two different sources of information (published articles associated with different omics datasets for example) were kept.

Applying such stringent filtering criteria allowed us to provide consistent results that take into account the heterogeneity of omics datasets and published literature. As a counterpart, it also filters out potential intercellular interactions that may be interesting to investigate in RA disease.

The resulting RA multicellular map can be expanded via the integration of additional cell-specific molecular interaction maps. These maps can be connected to the already existing ones through the identification of cell-cell communication using the previously described methodology.

Chapter 6. Generation and calibration of large-scale Boolean models

6.1 Introduction

Molecular interaction maps are standardized, graphic representations of complex biological systems. They are useful for data visualization and comparison; however, they remain limited by their static nature when it comes to predictions and hypothesis testing. One of the primary goals of dynamical modeling is to understand the emergent features and behaviors of the biological systems described in these maps (13) through *in silico* simulations and predictions. Boolean formalism is a well-suited approach for handling such large-scale systems. The regulation of the model's components is given in a parameter-free way, without the need for kinetic parameters and precise quantitative data (14,224).

Building and analyzing Boolean models for large-scale complex biological systems remains challenging. Manually built models are usually smaller models that leave out parts of the biological system under study, while models inferred automatically from maps are far larger in size providing more accurate representations of the systems (238). These large-scale models, on the other hand, are more complex, with a substantially higher number of inputs, making it computationally challenging to identify all their possible attractors (293). Furthermore, a significantly large state space hinders the identification and interpretation of biologically consistent states, especially in a cell or disease specific context.

In this section, we present an efficient computational framework to build, analyze and validate the behaviors of large-scale Boolean models with hundreds of nodes, and a significant number of inputs. It uses molecular interaction maps as a starting point to automatically infer their corresponding executable Boolean models via the CaSQ tool (238). The generated Boolean models are analyzed in a synchronous scheme using a new version of the Bio Model Analyzer (BMA) tool (294) deployed to a high-performance computing cluster, set up for this purpose. The framework identifies all the existing attractors of the models using parallel computing. It filters them to keep only the steady states and then tests their coherence against gene expression datasets and prior knowledge. We first applied the proposed framework to the four maps of the RA-Atlas (285) specific to the RA M1 macrophage, RA M2 macrophage, RA fibroblast and RA CD4+ Th1. The resulting calibrated states of their corresponding models were then combined, and the same framework was then applied to generate and validate the first RA multicellular model using the RA multicellular molecular interaction map.

6.2 Materials and methods

In this section, we thoroughly describe the computational framework that we propose to automatically convert molecular interaction maps into their corresponding executable Boolean models, analyze their state space, and validate their behaviors using both prior knowledge and transcriptomic data. Figure 36 illustrates the proposed computational framework.

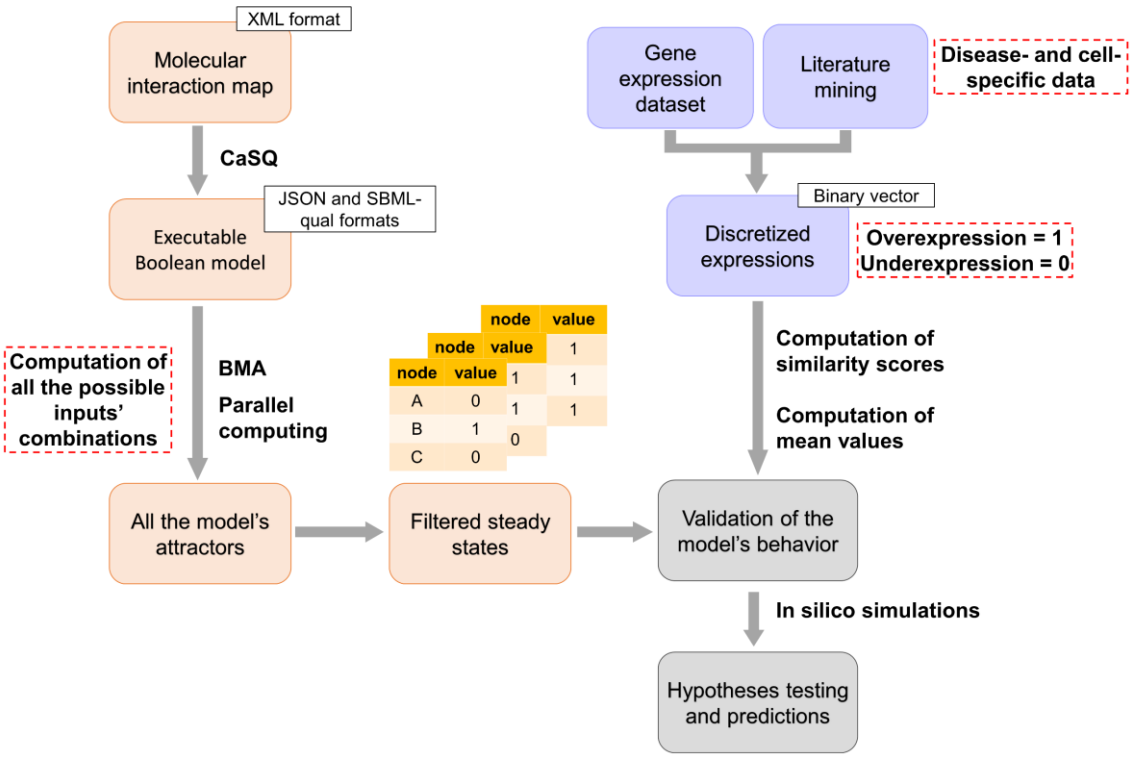


Figure 36. Schematic representation of the workflow we propose to generate and analyze large-scale Boolean models. Molecular interaction maps built in CellDesigner XML format are converted to executable Boolean models using CaSQ. Then, a new version of BMA is deployed to a high-performance computing cluster to identify all the models' attractors. These attractors are filtered to keep only the steady states. Next, the filtered steady states are validated. Differentially expressed biomolecules present in the models are identified using both literature mining and transcriptomic data analysis. The expressions of the identified biomolecules are discretized and converted to a binary vector of experimentally observed Boolean values. After that, similarity scores are computed to describe the similarity between the steady states and the experimentally observed values. The steady states with the highest score are selected, their mean vector represents the calibrated model's state.

6.2.1 Generation of Boolean models from molecular interaction maps

As a starting point for building Boolean models, we use molecular interaction maps and the tool CaSQ version 1.1.4 (238). For each node in the starting XML file, CaSQ automatically infers logical rules based on the network topology and semantics. First, the map is reduced through a pass of graph-rewriting rules. Some species in the starting map might need to be merged into a single component or into some inactive forms to avoid redundancy in the logical model. Then, logical rules are automatically inferred based on the topology of the reduced map. Each node's logical rule is defined as the (i) disjunction (OR), for all reactions producing it, of the (ii) disjunction (OR) of all catalysts of that reaction being activated and the (iii) conjunction (AND) of all products of that reaction being activated and all inhibitors being inactive.

Therefore, a target is ON if one of the reactions producing it is ON, and a reaction is ON if:

1) all its reactants are ON

2) all its inhibitors are OFF

3) One of its catalysts at least is ON

Figure 37 shows a toy example of a map in CellDesigner describing the formation of the complex AB. This reaction is catalyzed by either Activator 1 or Activator 2 and inhibited by Inhibitor 1. The logical rule of the complex AB infered using CaSQ is "(Activator 2 | Activator 1) & ! Inhibitor 1" which means that Complex AB is active if at least one of its catalysts (either Activator 1 or Activator 2) is ON and its inhibitor (Inhibitor 1) is OFF.

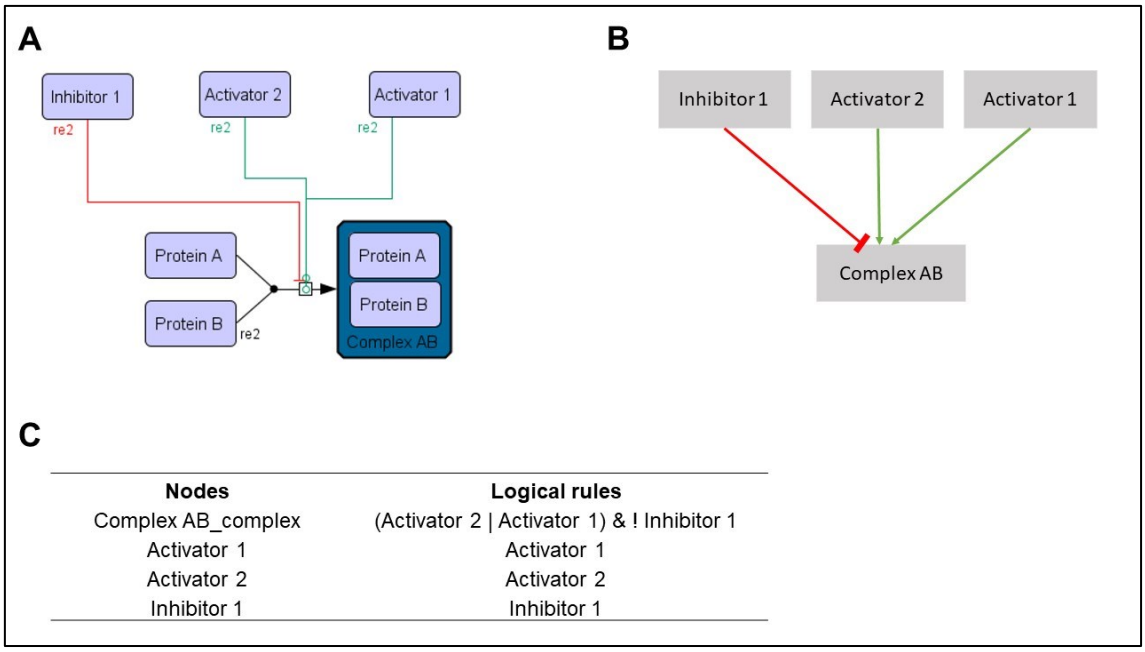


Figure 37. A) Toy XML file describing the formation of the complex AB. This reaction is catalyzed by either Activator 1 or Activator 2 (green arrows) and inhibited by Inhibitor 1 (red arrow). B) The Boolean model of the toy XML file depicted in A that we generated using CaSQ. C) The logical rules associated with each node in the Boolean model depicted in B.

The tool produces either Systems Biology Marked up Language qualitative (SBML-qual) (295) or BMA JSON executable files. During this PhD, we worked on optimizing the CaSQ export option in the latter format as we found inconsistencies between the logical rules in the SBML-qual models and the mathematical functions in the BMA JSON models, both produced by CaSQ. We identified the sources of these inconsistencies and worked in collaboration with Benjamin Hall (University College London) and Sylvain Soliman (Inria Saclay-Île-de-France) to correct them.

Using the BMA JSON format, it is also possible to produce qualitative networks where nodes can vary over a wide range of discrete values which is defined as granularity in BMA (294). Granularity defines the higher value the nodes can take in the model. Since we are using Boolean formalism, the granularity used in the following analysis is equal to one.

6.2.2 Stabilization proof using Bio Model Analyzer (BMA)

BMA is a tool for constructing and analyzing, and since recently importing, executable models of biological mechanisms (294). The user is presented with a web-based user interface, allowing for rapid and simple model construction and analysis. Whilst the GUI is the primary tool for interacting with BMA, a console

tool is also available giving access to a wide range of analysis algorithms and enabling scripting for large and complex combinatorial analyses. CaSQ is capable of generating models in the BMA JSON format, which can be used with either version of the BMA tool.

A key analysis in BMA is the proof of model stability. Using a synchronous update scheduler, a modular proof algorithm is used to show whether or not there exists a single steady state attractor, and no cycles. Briefly, this proceeds in two steps. Initially, the ranges of individual variables are reduced to the set of reachable values through examination of the input variable ranges, and the target function. If this process reduces all ranges to a singleton, stability is proven, and the global steady state attractor is returned to the user. If this fails, a constraint solver is used with the reduced variable ranges for the first step to search first for multiple fix points (bifurcation), and then cycles. Finally, if neither cycles nor multiple fix points are found, the model must be stable and a final check searches for and returns the steady state (296).

6.2.3 BMA architecture and underlying technologies

BMA is developed on the Microsoft .NET Framework and .NET standard, which tie the tools to Windows environments. The BMA web tool is hosted on Azure, and is structured as two services, one hosting the user facing client, and a second, compute service, dedicated to calculating proofs and simulations. The console tool, which provides similar functionality to the compute service, is developed for Windows and is available via this Github link: https://github.com/microsoft/BioModelAnalyzer.

To enable high-throughput model analysis and take advantage of parallelization, we needed to run the analyses on high performance computing facilities. Following Sanofi policy, these analyses had to be processed internally via Sanofi's AWS services. Since these HPCs are Linux-based only, we needed a new version of BMA that can run in Linux environments. In this direction, we worked in collaboration with Benjamin Hall and Rachel Alcraft from University College London who provided a new prototype of the BMA console tool that can run on Linux and that is based on the open-source .NET core 3.1 which can be built using the dotnet SDK. All codes are available at 10.5281/zenodo.7541023.

6.2.4 Parallel computing for the calculation of all possible attractors

Attractors depend on the external stimuli the model receives from its environment. External stimuli in Boolean models are modeled in the form of inputs. Inputs are nodes with no upstream regulation. They are not associated with any logical rule in the model; therefore, the values they take are user-defined. In the BMA console tool, the user can assign values to the input nodes with the

flag -ko that allows setting the specified nodes to be constants (zero or one). Depending on the inputs' states, the model reaches different attractors. To identify all the attractors of the model, we generate all the possible combinations of inputs' values. For each input combination, we search for the corresponding attractor. When possible, we reduce the number of input combinations by fixing the inputs associated with experimentally observed expressions. The Boolean values of these inputs are set based on the available literature and/or transcriptomic data (see Validation of the model's behavior).

The computation time is exponential. Indeed, the number of all the possible combinations of inputs' values is equal to 2^n, n being the number of inputs that vary in the model. Given the high number of inputs in the inferred large-scale Boolean models, we failed to execute the attractors search on a local Windows machine with 8 cores and 64 GBs of RAM. Therefore, we deployed BMA in a high-performance computing cluster to compensate for the lack of computational power.

Figure 38 illustrates the number of input combinations the BMA console tool can process per hour using one core and various model sizes. As the attractor search is slower on larger models, the number of processed combinations decreases proportionally with the model size. Therefore, Figure 38 can be used to estimate the computational resources required to execute the analysis depending on the model's size. We utilized the Joblib python package as well (297) to parallelize the process and considerably reduce the running time of the framework.

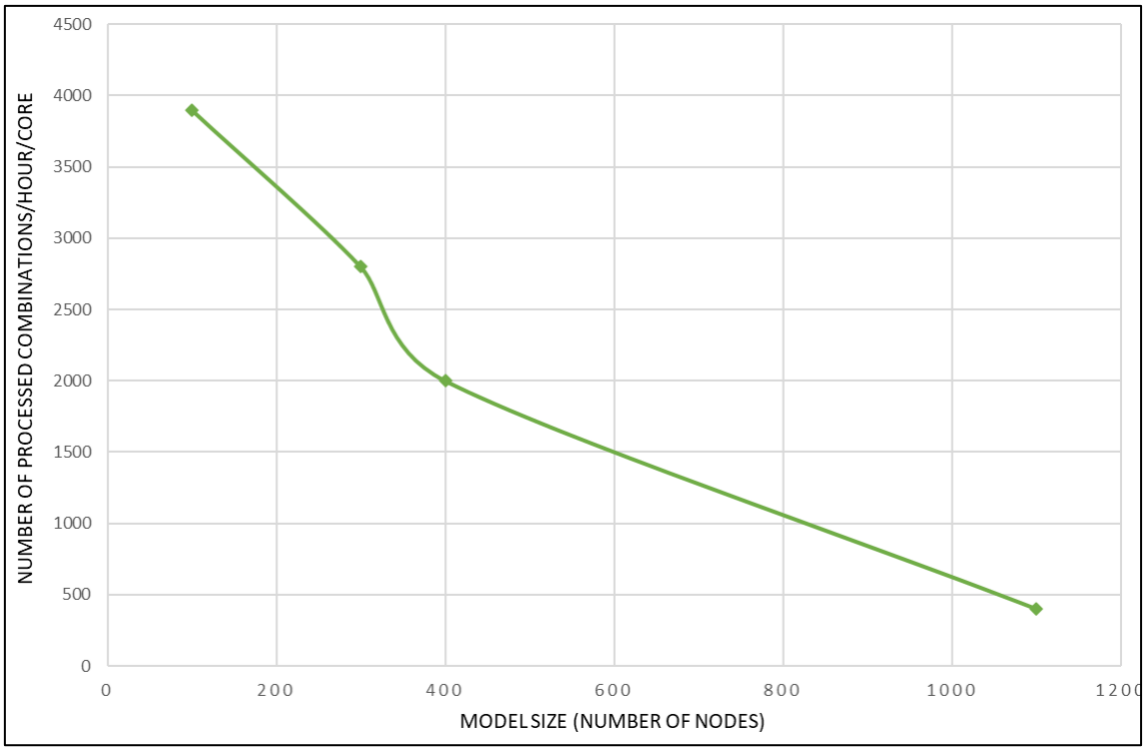


Figure 38. Plot showing the number of processed inputs' combinations by BMA per hour using a single core machine with a base frequency of 2.3 GHz.

6.2.5 Filtering the model's attractors

While cycles are linked with oscillations such as the cell cycle, steady states are characterized by stable patterns of biomolecules activities (often called phenotypes); they can therefore be compared with gene expression datasets. For this reason, we filter the model's attractors to keep only the steady states.

6.2.6 Validation of the model's behavior

The validation of the model's state is based on the signal propagation from the inputs to the internal nodes. The objective here is to select the inputs' combinations that lead to coherent states in the internal nodes of the model. To do this, we further filter the model's steady states and select the ones that are able to reproduce what is known in the literature or observed in transcriptomic datasets. First, we identify the differentially expressed biomolecules present in the model using both literature mining and differential expression analysis (DEA). Then, we discretize the expressions of the identified biomolecules to convert them to a binary vector of experimentally observed Boolean values. After that, we compare the filtered steady states to this binary vector using a similarity score. Based on this score, we select the states with the highest score and calculate their average vector. The resulting vector represents the calibrated state of the model.

6.2.6.1 Identification of differentially expressed biomolecules using literature search and gene expression data analysis

We use both low and high-throughput experimental data to identify the differentially expressed biomolecules present in the studied model. First, we conduct a thorough review of the literature regarding each node in the model. We extract information about the change of its expression level between two biological conditions. These conditions are defined based on the biological system under study. We curate the retrieved information to keep it disease- and cell-specific. Second, we integrate transcriptomic data if available. We select the dataset(s) according to the biological question we would like the model to address and perform DEA on the selected one(s). We filter the DEGs using appropriate log Fold Change (FC) and adjusted p-values thresholds. The final list of differentially expressed molecules present in the model is a combination of both literature search and DEA outcomes.

6.2.6.2 Computation of similarity scores and data discretization

To be able to compare the expressions of the identified differentially expressed components with the model's steady states, we discretize the data. Overexpressed biomolecules in the condition under study are associated with the value one while under expressed molecules are associated with the value zero. Since we are using Boolean formalism where each biomolecule can only have two possible states, biomolecules that are not differentially expressed, or associated with a low fold change are not considered (Figure 39).

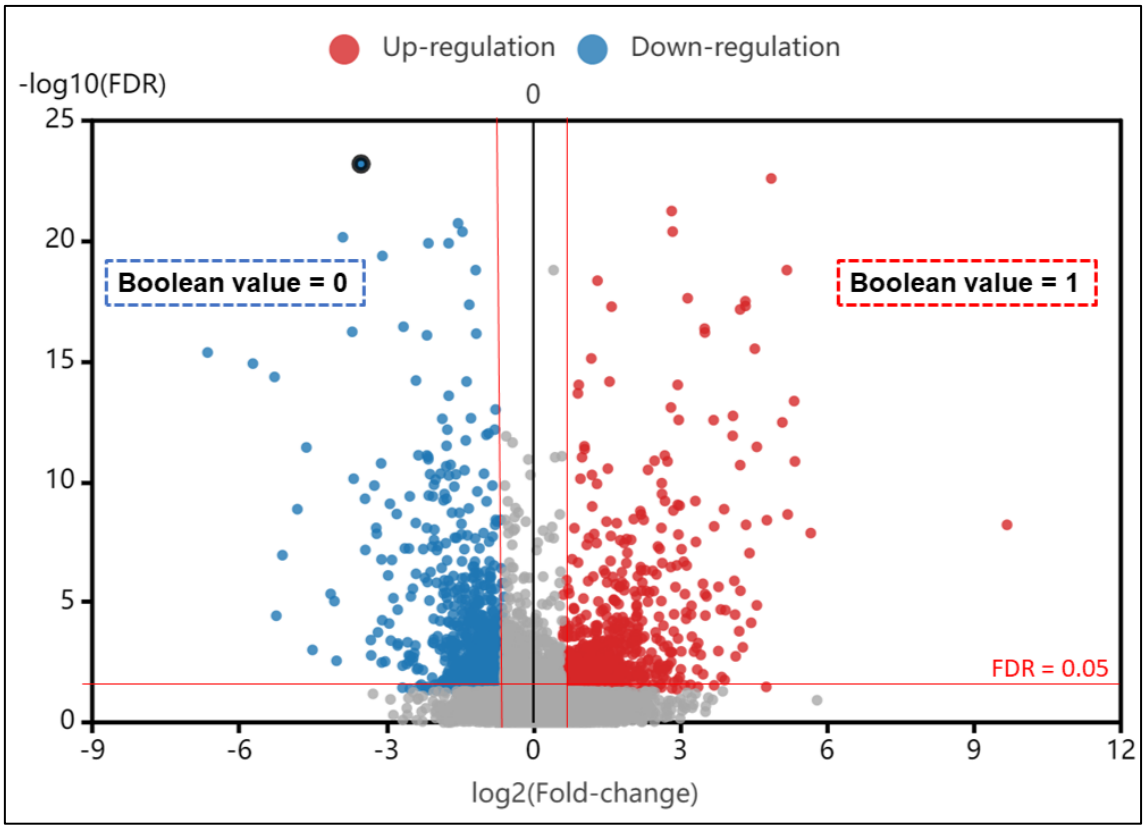


Figure 39. Example of gene expression discretization on a volcano plot showing the DEGs between RA and osteoarthritis synovial fibroblasts from GSE109449 dataset. DEGs were filtered using an FDR equal to 0.05 and a logFC equal to 0.584. Overexpressed DEGs are associated with the value one while under expressed molecules are associated with the value zero.

The resulting discretized vector of experimentally observed expressions is then used to calculate similarity scores with each steady state to describe the ability of these filtered steady states to reproduce the experimentally observed values. To do so, we calculate a similarity score S (1).

$$S = (N00 + N11)/N00 + N11 + N10 + N01 \quad (1)$$

Where,

N00 = number of nodes with a state of zero in both the steady state and the discretized vector of experimentally observed expressions

N01 and N10 = number of nodes with different states in the steady state and the discretized vector of experimentally observed expressions

N11 = number of nodes with a state of one in both the steady state and the discretized vector of experimentally observed expressions.

6.2.6.3 Selection of the steady states with the highest similarity score

We select the steady states with the highest similarity score. Then, we compute the mean value of each node over these stable states to determine the nodes that are fixed at either zero or one and those that can be found in both states. The resulting average vector represents the calibrated model's state.

6.3 Results

We illustrate in this section how we use the cell-specific molecular interaction maps of the RA-Atlas as a starting point to build cell-specific Boolean models describing RA M1 and M2 synovial macrophages, RA fibroblast and RA Th1. We also describe how we analyzed the state space, and then validated the behavior of these models using gene expression data together with prior knowledge. Once validated, the calibrated cell-specific models' states were combined, and the same analyses were performed on the RA multicellular model.

6.3.1 Updating the RA cell-specific molecular interaction maps

As a starting point of the work, we used cell-specific molecular interaction maps that are part of the RA-Atlas (285). Phenotypes are particular nodes in these maps. They describe biological states known to be active or inactive in RA. To make them more appropriate to the purpose of this work, we divided them into two categories. The first one corresponds to cell-specific phenotypes. They describe the cellular outcomes of each cell-type of interest like proliferation and apoptosis. Their names end with "M1_macrophage", "M2_macrophage", "fibroblast" or "Th1" suffixes depending on the map. The second category is not specific to a particular cell type. It corresponds to cellular signals and biological conditions in the RA joint like inflammation, angiogenesis and matrix degradation. Their names end with the ''signal'' suffix in all the updated cell-specific maps.

We updated the maps by looking for duplicates as well, removing them whenever they were found and correcting the signaling pathways accordingly. Table 13 summarizes the number of nodes and reactions in the updated RA cell-specific maps.

Table 13. Number of nodes and reactions in the RA cell-specific molecular interaction maps.

RA cell-specific maps	Number of nodes	Number of reactions
RA M1 macrophage	601	407
RA M2 macrophage	511	322
RA fibroblast	837	541
RA CD4+ Th1	319	198

6.3.2 Generation of the RA cell-specific Boolean models

We used CaSQ to convert the updated RA cell-specific maps to Boolean models using BMA JSON as export format. Table 14 shows the number of nodes, inputs, and interactions in each model.

The resulting JSON files will be made publicly available soon to allow the visualization of the mathematical functions associated with each node that CaSQ generated as well as an easy reuse of our models.

Table 14. Number of nodes, inputs, and interaction in the RA cell-specific Boolean models

Cell-specific model	Number of nodes	Number of inputs	Number of interactions
RA M1 macrophage	309	75	560
RA M2 macrophage	253	57	426
RA fibroblast	394	104	638
RA CD4+ Th1	166	38	218

In this section, we focused on the regulation of the first category of phenotypes, the cell-specific ones. As the second category of phenotypes describes the biological conditions in the RA joint and is influenced by several cell-types, its regulation will be tested and validated when calibrating the RA multicellular model.

We used the export option in CaSQ via the argument -u to identify all the nodes that are upstream of these phenotypes. Table 15 describes the cell-specific phenotypes for each model, and the number of nodes and inputs upstream these phenotypes. The nodes that are not involved in the regulation of our phenotypes of interest were not taken into consideration and the inputs regulating these nodes were fixed at one, the default value in BMA.

Table 15. Number of nodes and inputs upstream of the cell-specific phenotypes in the RA cell-specific Boolean models.

Cell-specific model	Cell-specific phenotypes	Number of nodes	Number of inputs	Number of interactions
RA M1 macrophage	Apoptosis Proliferation Osteoclastogenesis	233	64	381
RA M2 macrophage	Apoptosis Proliferation	169	39	278
RA fibroblast	Apoptosis Proliferation	275	73	446

	Migration			
RA CD4+ Th1	Apoptosis Proliferation Migration	120	28	155

6.3.3 Identification of differentially expressed molecules using literature search and transcriptomic data analysis

We extracted information from published literature regarding the differential expression of the models' components. These expressions can be at the mRNA and/or protein levels, depending on literature availability. We curated the extracted information to keep it specific to both RA disease and the relevant cell-type.

We integrated gene expression datasets as well. DEA was performed between RA and healthy synovial macrophages samples in the GSE97779 gene expression dataset. The resulting list of DEGs was combined with literature search outcomes and used to calibrate the RA M1 and M2 macrophage models. The same analyses were performed on the GSE109449 and SDY998 single cell RNA-seq datasets to calibrate the RA fibroblast model and the RA CD4+ Th1 models respectively. Details regarding datasets analyses are provided in Chapter 3. Description of data and data analysis. Assuming a linear relationship between the expression of mRNAs and the expression of their corresponding proteins, the retrieved differential expressions can be associated with nodes at mRNA or protein levels in the models. We discretized the differentially expressed molecules' expressions: molecules that were overexpressed were linked to the value 1, whereas molecules that were under expressed were linked to the value 0.

Supplementary table 1, Supplementary table 3, Supplementary table 5, and Supplementary table 7 in ANNEX A provide the lists of these differentially expressed molecules in each cell-specific model together with their observed Boolean values in the literature and/or gene expression datasets.

6.3.4 Computation of all the possible attractors of the RA cell-specific models

Given the high number of inputs in the models, we reduced the list of inputs' combinations by fixing the values of the differentially expressed ones. Based on the information displayed in Supplementary table 1, 43 out of the 64 inputs present in the RA M1 macrophage model were fixed. The total number of inputs' combinations was then equal to 2^{21}. We used the BMA tool deployed to a machine with 96 single core CPUs and 768 GBs of RAM to run the attractors

search. All the resulting attractors were steady states and were kept for further analysis. Regarding the RA M2 macrophage model, 24 inputs were fixed using the information provided in *Supplementary table 3*. The number of inputs' combinations was then equal to 2^{14}. The same machine configuration was used to run the attractor search. All the corresponding attractors were steady states that we used for the following steps.

Based on the information displayed in *Supplementary table 5*, 53 out of 73 inputs states were fixed in the RA fibroblast model. The number of inputs' combinations was then equal to 2^{20}. All the resulting attractors were steady states. Regarding the CD4+ Th1 model, 16 out of 28 inputs were fixed according to the information displayed in Supplementary table 8. The number of inputs' combinations was then equal to 2^{12}. All the corresponding attractors were steady states.

6.3.5 Validation of the cell-specific models' behaviors

First, we filtered the steady states according to the values of their cell-specific phenotypes. The biologically coherent Boolean values of these phenotypes were extracted from the literature in a disease and a cell-specific manner (Table 16).

In the RA M1 and M2 macrophage models, the Boolean state of the cell-specific phenotypes should reflect the increased M1/M2 ratio in the synovial macrophage population and the enhanced osteoclastic bone resorption in the RA joint. Indeed, RA M1 macrophages predominate in RA synovial fluid due to their excessive proliferation (Proliferation phenotype in the model should be ON) and enhanced anti-apoptosis capabilities (Apoptosis phenotype in the model should be OFF) compared to the RA M2 macrophages (Apoptosis phenotype should be ON, and Proliferation phenotype should be OFF in the model) (116). In addition, Osteoclastogenesis phenotype, which is only present in the RA M1 macrophage model, should be ON. All the steady states of the RA M1 macrophage model passed through this filtering step while only 8192 of the steady states of the RA M2 macrophage model did.

In the RA fibroblast model, the state of the cell-specific phenotypes should reflect their resistance to apoptosis and excessive proliferation in RA synovium (Apoptosis should be OFF and Proliferation should be ON in the model) (96). They also describe the fibroblast ability to spread to unaffected joints in RA patients (Migration should be ON in the model) (298). All RA fibroblast model's steady states passed through this filtering step.

In the RA CD4+ Th1 model, the state of the cell-specific phenotypes should describe the increased ability of this cell-type to migrate and extravasate from blood vessels to the inflamed RA joint (Migration should be ON in the model) (86). They

also show their increased apoptosis resistance (Apoptosis should be OFF in the model) and their hyperactivation and impaired proliferation in RA (Proliferation should be ON in the model) (130). 1024 steady states passed through this filtering step in the RA CD4+ Th1 model.

Table 16. Expected Boolean states of the phenotypes in the RA cell-specific Boolean models according to the literature.

Model/ Phenotypes	Proliferation	Apoptosis	Migration	Osteoclastogenesis
Fibroblast	1	0	1	/
M1 macrophage	1	0	/	1
M2 macrophage	0	1	/	/
Th1	1	0	1	/

For each cell-specific model, we calculated similarity scores between the lists of differentially expressed molecules and their matching nodes in each filtered steady state. Steady states with the highest similarity score in each model were selected and their average vectors were calculated. The resulting mean vectors represent the calibrated state of the RA cell-specific models (Supplementary table 2, Supplementary table 4, Supplementary table 6, Supplementary table 8). Table 17 shows the number of steady states with the highest similarity score for each cell-specific model. It also describes the number of nodes that were fixed at zero or one in the calibrated states of the models and the similarity score associated with each one of them.

Table 17. Description of the RA cell-specific models' calibrated states. The first row displays the number of steady states with the highest similarity scores. The second row shows similarity scores describing the ability of the calibrated models to reproduce the experimentally observed expressions. The third row indicates the

number of nodes fixed at zero or one in the calibrated states of the models.

Model	M1 macrophage	M2 macrophage	Th1	Fibroblast
Number of steady states with the highest score	384	96	128	4096
Similarity score	99%	96,5%	100%	99%
Fixed nodes/Total number of nodes	222/233	158/169	109/120	254/275

Some mismatches were identified between the models' calibrated states and the experimentally observed Boolean values (Table 18). These inconsistencies will be addressed later in the discussion section.

Table 18. Mismatches identified during the validation of the models' behaviors.

Node	Model	Boolean value in the model	Experimentally observed Boolean value
CASP7	RA M1 macrophage	0	1
BCL2L1	RA M2 macrophage	0	1
CASP3	RA M2 macrophage	1	0
MCL1	RA M2 macrophage	0	1
GAB2	RA fibroblast	1	0

6.3.6 Calibration of the RA multicellular model

We used CaSQ to convert the RA multicellular interaction map to a Boolean model using the BMA JSON as export format. It consists of 1104 nodes, including 240 inputs, and 1845 interactions (Figure 40). A full-page landscape image of the RA multicellular model displayed in BMA is provided in Supplementary figure 6. This

image is not readable in a A4 size page due to the model's complexity and is displayed for illustrative purposes only. The BMA JSON file of the corresponding model will be publicly available in the open-access publication that we are currently preparing regarding the RA multicellular model. The BMA JSON file allows for a better exploration of the model and for an easy visualization of the mathematical functions associated with each node.

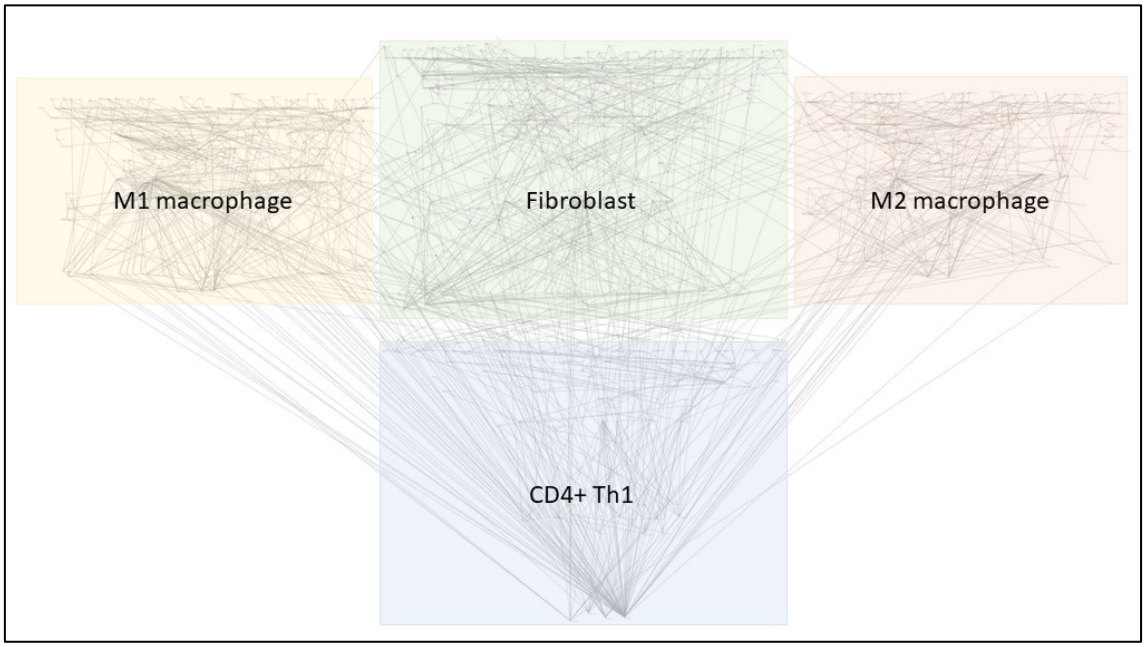


Figure 40. RA multicellular model in BMA graphical interface.

6.3.6.1 Combination of the calibrated states of the RA cell-specific models via intercellular communication

To calibrate the resulting multicellular model, first we combined the cell-specific models' calibrated states via the addition of intercellular interaction present in the multicellular model. 159 out of the 204 inputs that were already processed during the cell-specific models' calibration (64 in the RA M1 macrophage model, 39 in the RA M2 macrophage model, 73 in the RA fibroblast model, and 28 in the CD4+ Th1 model) are still inputs in the RA multicellular model. Their states were defined according to the cell-specific models' calibrated states. The remaining 45 nodes are no longer inputs in the multicellular model. The intercellular interactions that regulate them define their states in the model. Figure 41 shows an example of how cell-cell interactions can modify an input's state. In Figure 41A, CXCL13 is an input in the RA M1 macrophage model, its value is user defined. In Figure 41B, CXCL13 is no longer an input in the RA multicellular model. It participates in an intercellular interaction in which Th1 cells release CXCL13, which is then captured by the M1 macrophage cells.

Some of these 45 nodes had already been fixed in the cell-specific models. Their states remained unaltered in the multicellular model because cell-cell interactions are in accordance with the calibrated states of the cell-specific models. Some of them were not fixed in the cell-specific models. After the addition of the intercellular interactions, their states were fixed (Table 19).

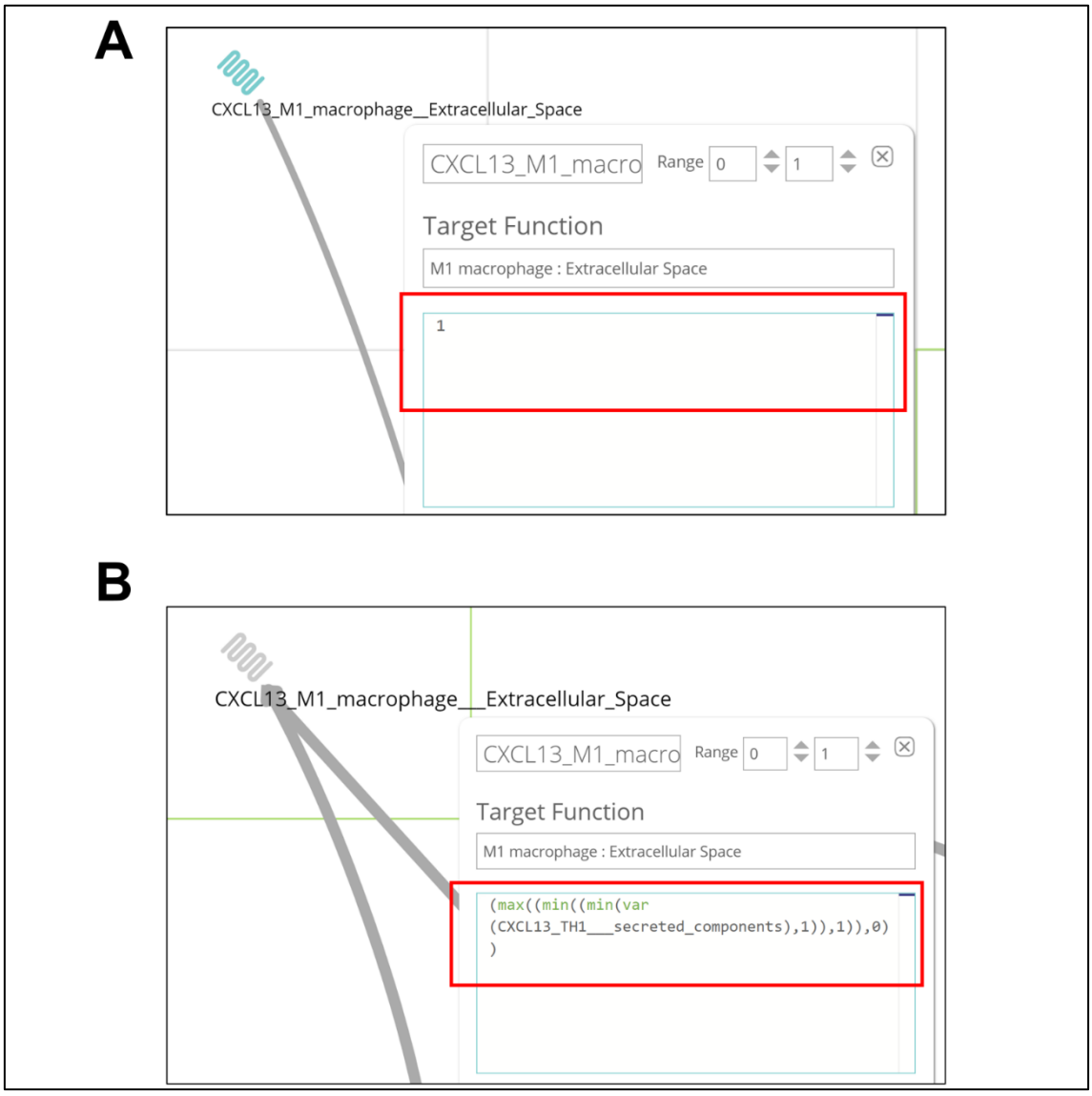


Figure 41. A) CXCL13 mathematical rule in RA M1 macrophage model before adding the intercellular interactions. B) CXCL13 mathematical rule in RA M1 macrophage model after adding the intercellular interactions.

Table 19. Number of nodes fixed in the RA cell-specific models before and after

adding cell-cell interactions.

Model	M1 macrophage	M2 macrophage	Th1	Fibroblast
Number of fixed nodes before adding intercellular interactions	222/233	158/169	109/120	254/275
Number of fixed nodes after adding intercellular interactions	227/233	162/169	120/120	256/275

6.3.6.2 Identification of differentially expressed molecules using literature search and transcriptomic data analysis

307 nodes, including 81 inputs in the multicellular model are not upstream of the cell-specific phenotypes and were not analyzed during the cell-specific models' calibration. To be able to calibrate the entirety of the multicellular model, we extracted information from published literature regarding the differential expression of these remaining biomolecules. We curated the extracted information to keep it specific to both RA disease and the relevant cell-type. We also integrated additional DEGs from the same gene expression datasets we used for the cell-specific models' calibration. *Supplementary table 9* in ANNEX A describes the additional differentially expressed biomolecules identified in both omics datasets and literature present this remaining part of the multicellular model.

6.3.6.3 Computation of all the possible attractors of the multicellular model

Given the high number of inputs in the model, we reduced the list of inputs' combinations by fixing the values of some of them. First, we made use of the cell-specific models' calibrated states. Among the 240 inputs present in the multicellular model, 141 were already fixed from the RA cell-specific models' calibration and combination. Then, we used information extracted from literature and gene expression datasets to fix additional inputs. Based on the information displayed in *Supplementary table 9*, 78 additional inputs were fixed. The total number of fixed inputs was equal to 219. The 21 remaining inputs were not associated with a Boolean value. The total number of input combinations was then equal to 2^{21}. We used the BMA tool deployed to a machine with 96 single core CPUs and 768 GBs of RAM to run the attractors search. All the resulting attractors were steady states and were kept for further analysis.

6.3.6.4 Validation of the multicellular model's behavior

We calculated the similarity score between the list of differentially expressed molecules and their matching nodes in each steady state. Steady states with the highest similarity score were selected to calculate their mean vectors. The resulting mean vector represents the calibrated state of the RA multicellular model. 32768 steady states had the highest similarity score. In their mean vector, 1076 out of 1104 nodes were fixed at zero or one, while 28 of them were not fixed (Supplementary table 10). This multicellular model's state can reproduce 98,8% of the observed Boolean values used to calibrate it. No additional mismatches were identified apart from the ones already highlighted in Table 18.

6.4 Discussion

In this chapter, we put in place a framework to calibrate large-scale Boolean models that can be either automatically inferred from molecular interaction maps using the CaSQ tool or manually built in the BMA JSON format. We analyze the models using a newly developed version of the BMA tool that can be deployed to Linux-based high-performance computing clusters to identify all the steady states of the models. Depending on the model size and complexity, the computation time varies. Given the high number of possible input combinations in large-scale Boolean models, using HPC clusters that consist of hundreds, or thousands of cores enables high throughput model analysis. It overcomes the lack of computational power and takes advantage of parallel computing to considerably reduce the running time of the analysis. We filter the resulting steady states by selecting the biologically consistent ones. The average vector of these selected steady states represents the calibrated model's state that most accurately reflects the biological condition under study. All the files and scripts needed to utilize the proposed framework are available at: https://gitlab.com/genhotel/Large_scale_computational_modelling_of_the_M1_and_M2_synovial_macrophages_in_Rheumatoid_Arthritis

We applied the proposed methodology to four RA cell-specific molecular interaction maps, namely RA M1 macrophage, RA M2 macrophage, RA fibroblast and RA CD4+ Th1, setting the path to many other disease maps to be explored.

To analyze their corresponding cell-specific models, we adjusted the framework according to the disease and cell type under study. Indeed, we divided the phenotypes of the models into cell-specific ones to describe the cellular outcomes of each cell-type of interest, and more generic ones to describe the overall condition of the RA joint. We filtered the models' steady states according to the values of their cell-specific phenotypes and kept the ones that reflect the biological conditions of the RA joint. In addition, to calibrate the models, we selected disease-

and cell-specific gene expression datasets and carefully curated the extracted information from the literature to ensure that it was specific to both RA disease and the cell-types of interest.

Mismatches between experimentally observed values and our models' states were identified. Regarding the RA M1 and M2 macrophages models, these observed values are cell and disease specific but still contain a mix of M1 and M2 phenotypes' expressions. Indeed, the identified mismatches in both models are related to biomolecules participating in the intrinsic and extrinsic apoptosis pathways. The M1 phenotype is resistant to apoptosis so pro apoptosis components (CASP7) tend to be inhibited. On the other hand, the M2 phenotype is pro-apoptotic, hence pro apoptosis components are active (CASP3) while anti apoptosis molecules are inhibited (BCL2L1, MCL1).

The calibration of the Boolean models can be performed using different datasets, publicly available or in-house. As RA is a disease of great heterogeneity, differences in certain DEGs are expected.

We successfully applied our framework to the RA multicellular model as well. To reduce the model's complexity, we first combined the cell-specific models' states via the integration of cell-cell communication. Then we applied the same protocol as for the cell-specific models. As the model's size and complexity considerably increased, we were able to demonstrate the scalability of our framework.

Since we are using Boolean formalism, the models' steady states are binary vectors. To compare them with the differentially expressed molecules, overexpressed ones were associated with the value one while under expressed ones were associated with the value zero. Biomolecules that are not differentially expressed like housekeeping genes are not considered. To overcome this limitation, qualitative models with a higher granularity could be envisioned; nevertheless, additional computational resources would be needed to cope with the exponentially higher complexity of these models.

Most of the models' nodes were fixed at zero or one in the selected steady states. Due to the lack of information regarding their expression, the remaining nodes were not fixed, meaning they could be found in both states (in different steady states). To set their Boolean values at either zero or one, we could incorporate additional layers of information via the integration of other types of omics datasets such as proteomic data. This would enrich the list of experimentally observed Boolean values and help to narrow down the number of biologically consistent steady states.

Chapter 7. Selective M1 macrophage depletion and M2 macrophage promotion in the RA synovium

7.1 Introduction

Recently, research has revealed that the innate immune system is a major player in the initiation and development of the RA pathogenesis (299). Macrophages are one of the most common innate immune cell populations in RA and their number significantly correlates with the disease severity (116). Macrophage populations are heterogeneous and can differentiate into various phenotypes in response to the local microenvironment stimuli. The M1 and M2 phenotypes represent the extremes of their activation spectrum.

Consequently, depending on their phenotype (M1 or M2), macrophages play a role in both the initiation and resolution of inflammation (87). The M1 macrophages are responsible for the overproduction of inflammatory cytokines and the release of matrix degradation enzymes, leading to cartilage destruction (299). They can also attract proinflammatory T cells and induce their hyperactivation. On the other hand, the M2 macrophages alleviate inflammation via 1) the production of anti-inflammatory cytokines, including IL-10 and TGF-β, 2) tissue homeostasis and repair (300) 3) activation of regulatory T cells functions (119). Due to their excessive activation and proliferation as well as their enhanced anti-apoptotic ability, the proportion of M1 macrophages is higher than that of M2 macrophages in RA (116). Two approaches currently exist for targeting macrophages: downregulating M1 phenotype and expanding M2 phenotype, or repolarizing M1 macrophages to M2 macrophages (301,302). In spite of this, no medicines specifically targeting macrophages are currently used in the clinic (116,300). Thus, understanding the specific approach for targeted depletion of the inflammatory macrophage while sparing other macrophage subsets and reestablishing macrophage balance might be an effective therapeutic approach in RA (303).

Although the heterogeneity of macrophages in RA has not been fully uncovered, the RA M1 and M2 macrophages models that we built and validated in the previous chapter (see Chapter 6. Generation and calibration of large-scale Boolean models) aim to cover the phenotypic diversity of macrophages through a phenotype-specific representation of their secreted cytokines/chemokines, stimulatory molecules, receptors, and transcription factors (303).

In this chapter, we utilize these validated models to investigate mono- and bi-therapies that could specifically downregulate pro-inflammatory macrophages and promote anti-inflammatory macrophages in RA synovium. This can be achieved through the induction of the Apoptosis phenotype and the inhibition of

the Proliferation phenotype in the RA M1 macrophage model, and the activation of the Proliferation phenotype and the inhibition of the Apoptosis phenotype in the RA M2 macrophage model. In this direction, we tested the models' responses to current RA therapies to evaluate the effect of these treatments on macrophages and gain a better understanding of their mechanism of action in this cell type. We also tested new drug combinations and potential therapeutic intervention points.

7.2 Methods and results

7.2.1 Single knockouts of the therapeutic targets present in the RA M1 and M2 macrophage models

To identify potential therapeutic targets (targets that have already been experimentally modulated) present in the models, we performed an exhaustive search using the Therapeutic Target Database (TTD). It is a drug database designed to provide information about the known therapeutic protein and nucleic acid targets described in the literature, the targeted disease conditions, the pathway information, and the corresponding drugs/ligands directed at each of these targets. The database currently contains 3578 targets and 38760 drugs. Targets can be divided into four categories: successful targets, clinical trial targets, preclinical trial targets and research targets (304). In order to simulate knockouts in the models, we only focused on the inhibition of the drug target. We screened the targets based on the Mode Of Action (MOA) of their associated drugs, and only kept the ones that can be targeted by at least one inhibitor (1643 targets). Then, we identified the targets that are present in the RA M1 and M2 macrophage models. Regarding the RA M1 macrophage model, 71 therapeutic targets that 988 different drugs can inhibit were identified (Supplementary table 11) while 60 targets that 1026 different drugs can inhibit were identified in the RA M2 macrophage model (Supplementary table 12).

We mimicked the effect of these targets' KOs using *in silico* knockout simulations. We used the calibrated state of both RA M1 and RA M2 macrophage models as initial conditions for the simulations (Supplementary table 2 and Supplementary table 4 respectively). Then, we compared the models' phenotype states after the target knockouts to their corresponding calibrated states. Table 20 summarizes the predicted therapeutic targets in both models.

To visualize the simulation results, we imported our models in an SBML qual format in the Cell Collective platform (232). The platform allows the simulation of loss/gain of functions of user-defined nodes. As we can see it in Figure 42 and

Figure 43, each point in the graphs provided by the platform represents the number of logical time steps in which the displayed node is active over a user-defined number of time steps called sliding window (e.g. if the sliding window is 100, and the node is active in 20 steps over the last 100 steps, the activity level displayed at that point on the graph will be 20%).

NF-κB inhibition stimulated the M1 macrophage model's death and inhibited the M1 macrophage model's growth (Figure 42). Its inhibition did not influence the M2 macrophage model's proliferation or apoptosis. Even though ERK1 inhibition did not affect the M1 macrophage model's apoptosis, it did suppress its proliferation and reduce the release of most proinflammatory cytokines (CCL2, CSF2, IFNG, IL-18, IL-1, IL-6, and TNF) in the model. It also blocked the synthesis of the angiogenic factor VEGFA in the M2 macrophage model. GSK3β inhibition, on the other hand, induced the M2 macrophage model's proliferation while suppressing its apoptosis (Figure 43).

Table 20. Single knockouts of the therapeutic targets from the TTD database that perturb the RA macrophages' phenotypes.

Target	Target type	Associated disease(s)	Drug with highest status
NF-κB	Successful target	Irritable bowel syndrome, Rheumatoid arthritis, Choreiform disorder, Lupus erythematosus, Multiple sclerosis, ...	Sulfasalazine (Approved)
ERK1	Clinical trial target	Melanoma, Pancreatic cancer, Cancer, Arteries/arterioles disorder, Mature T-cell lymphoma.	BVD-523 (Phase 2)
GSK3β	Clinical trial target	Myotonic disorder, Acute myeloid leukemia, Osteosarcoma, Fragile X chromosome, Myeloproliferative neoplasm, ...	Tideglusib (Phase 2/3)

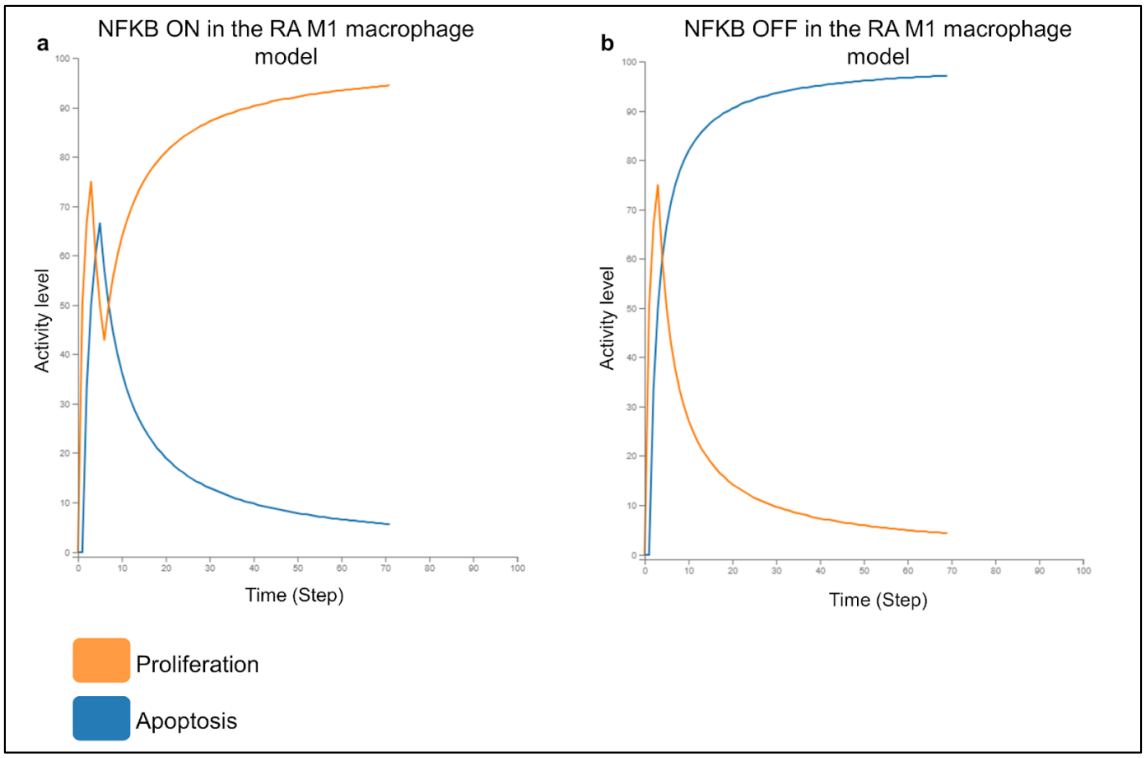


Figure 42. a) In silico simulation of the proliferation and apoptosis phenotypes in the RA M1 macrophage model before the NF-κB KO. b) In silico simulation of the proliferation and apoptosis phenotypes in the RA M1 macrophage model after the NF-κB KO.

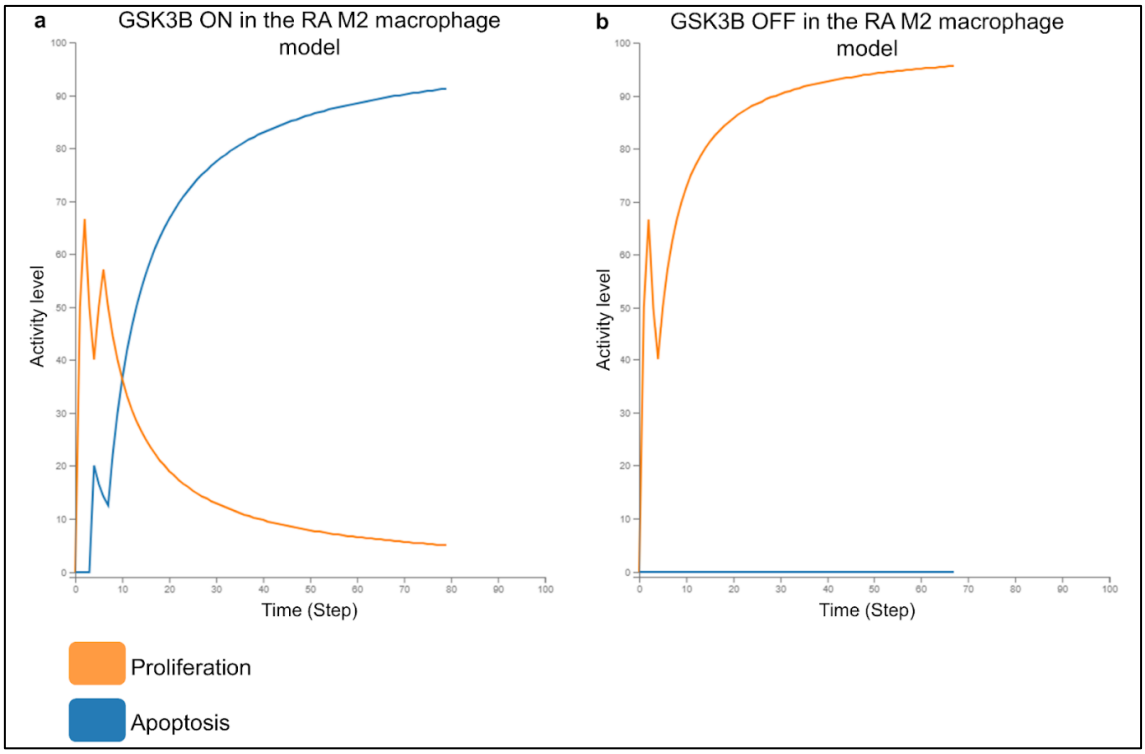


Figure 43. a) In silico simulation of the proliferation and apoptosis phenotypes in the RA M2 macrophage model before the GSK3β KO. b) In silico simulation of the proliferation and apoptosis phenotypes in the RA M2 macrophage model after the GSK3β KO.

7.2.2 Double knockouts of the therapeutic targets present in the RA M1 and M2 macrophage models

To investigate the potential synergistic effect of the previously tested therapeutic targets on the models, the targets were combined in pairs. Both RA M1 macrophage and RA M2 macrophage models were used to predict the outcome of their corresponding combined KO. We used the same initial conditions as for the mono drug testing, then we compared the perturbed states with their corresponding calibrated states.

We generated a list of all the possible unique pairs of the 71 previously identified therapeutic targets in the RA M1 macrophage model. 2485 target combinations were tested. Among these combinations, the Notch1/ERK1 pair and the JAK1/JAK2 pair were predicted as combinations having a synergistic effect on the model (Table 21). Indeed, ERK1 KO alone inhibited M1 macrophage model's proliferation. When combined with Notch1 KO, it also led to the promotion of the M1 macrophage model's apoptosis (Figure 44). JAK1 and JAK2 separate inhibitions didn't perturb the macrophages models' phenotypes either. When paired together, they suppressed the M1 macrophage model's proliferation and induced

its apoptosis. All the other drug pairs didn't provide any addidive nor synergistic effect on the models' Apoptosis and Proliferation phenotypes. Apoptosis induction and proliferation suppression were only driven by NF-κB KO in the models.

We generated a list of all the possible pairs (without repetition) of the 60 previously identified therapeutic targets in the RA M2 macrophage model. 1770 target combinations were tested. None of the drug combinations demonstrated a synergistic effect on the M2 macrophage model's phenotypes. Apoptosis suppression and proliferation activation were only driven by GSK3β knockouts in our model.

Table 21. Combinations of therapeutic targets (from TTD database) that perturb the RA macrophages' phenotypes.

Synergistic combination in the model	Targets	Target type	Associated disease(s)	Drug with highest status
JAK1/JAK2	JAK1	Successful target	Acquired hypomelanotic disorder, Atopic eczema, Crohn disease, Myeloproliferative neoplasm, Pancreatic cancer, ...	Baricitinib (Approved)
	JAK2	Successful target	Acquired hypomelanotic disorder, Atopic eczema, Myeloproliferative neoplasm, Pancreatic cancer, Rheumatoid arthritis, ...	Baricitinib (Approved)
ERK1/Notch1	ERK1	Clinical trial target	Melanoma, Pancreatic cancer, Cancer, Arteries/arterioles disorder, Mature T-cell lymphoma	BVD-523 (Phase 2)
	Notch1	Clinical trial target	Lymphoma, Mature T-cell lymphoma, Cancer	LY3039478 (Phase 1/2)

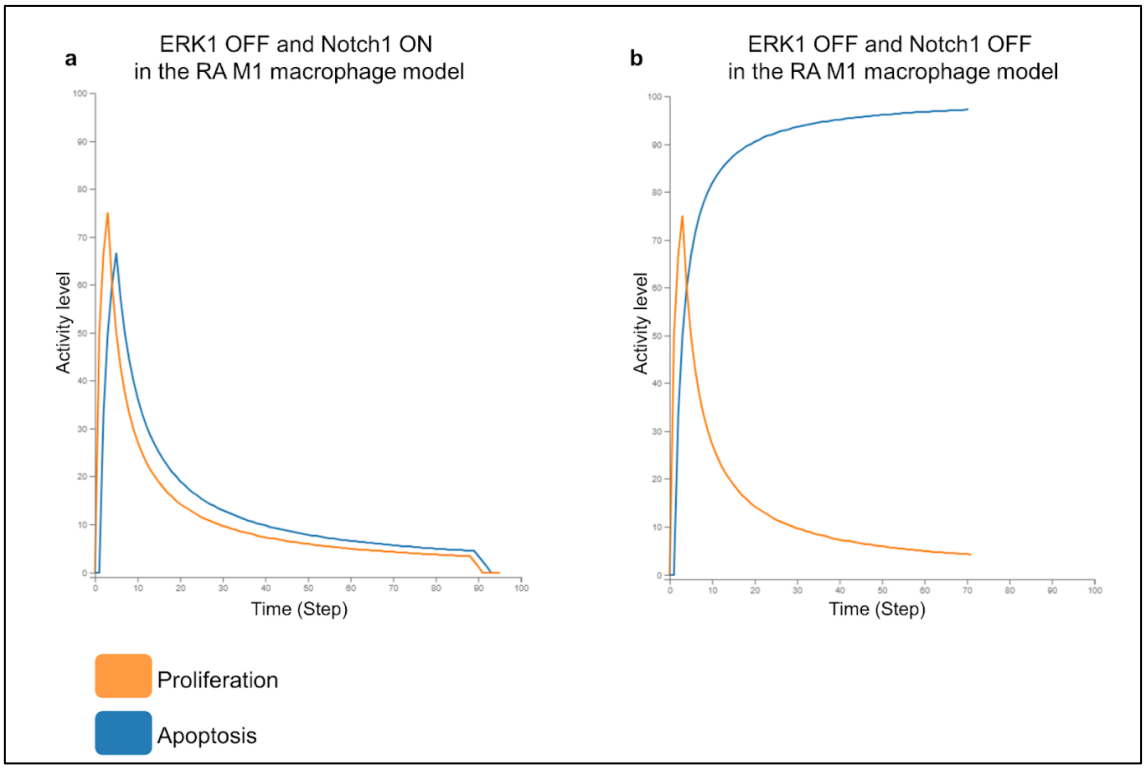


Figure 44. a) In silico simulation of the proliferation and apoptosis phenotypes in the RA M1 macrophage model after ERK1 KO. b) In silico simulation of the proliferation and apoptosis phenotypes in the RA M1 macrophage model after ERK1 and Notch1 KO.

7.2.3 Double knockouts of the receptors expressed by the RA M1 and M2 macrophages

Receptors, located on both the cell surface and within the cell, are the molecular targets through which drugs produce their beneficial effects in various disease states. They are coupled to various signal transduction systems located both within the membrane and intracellularly and can therefore regulate responses to the cellular/tissue microenvironment (305). These receptors can be targeted by antibodies in clinics. The advent of antibody therapeutics has revolutionized the pharmacological treatment of many rheumatic diseases, including RA (306). However, various factors and multiple signaling pathways are involved in such complex diseases; therefore, Monoclonal AntiBodies (MoAbs) are often associated with drug resistance, and most studies on MoAb combination therapies are still in the early stage (307). Bispecific AntiBodies (BsAbs) might treat rheumatic diseases by taking advantage of the diversity of their functions (306). Indeed, these antibodies have two binding sites directed at two different receptors, which makes their clinical therapeutic effects superior to those of MoAbs (308) with a lower rate of resistance. Approximately 70% of BsAbs in clinical trials are for the treatment of oncological indications (309). Thus, a high

unmet medical need in RA remains.

To investigate potential therapeutic targets for BsAbs in RA, we tested the effects of double KOs of the cellular receptors present in both the M1 and M2 macrophage models. Receptors were combined two by two and both RA M1 macrophage and RA M2 macrophage models were used to predict the outcome of their corresponding combined KOs. We used the same initial conditions as for the mono and dual drug testing, then we compared the perturbed states with their corresponding calibrated states.

406 double KOs were tested in the RA M1 macrophage model while 300 double KOs were tested in the RA M2 macrophage model. None of these simulations perturbed the RA macrophages models' apoptosis or proliferation phenotypes.

7.3 Discussion

The number of macrophages in inflamed synovial tissue grows rapidly during the course of RA, and their polarization plays a critical role in the physiological and pathological progression of RA (310). Thus, selectively suppressing the M1 macrophages or boosting the M2 macrophage could be a promising strategy for the treatment of RA. To investigate such complex mechanisms, we performed *in silico* simulations on the calibrated RA M1 and M2 macrophage models to investigate the effects of mono and bi-therapies on the RA macrophages' phenotypes.

NF-κB inhibition led to a selective suppression of the RA M1 macrophage model. NF-κB represents a very attractive potential therapeutic target as it is a key transcription factor of the M1 macrophages, responsible for the upregulated expression of M1 macrophage-derived cytokines in the RA synovium (311). Several studies support the concept of NF-κB inhibition for therapeutic interventions in inflammatory diseases, including RA (312–315). In RA, the in vitro inhibition of NF-κB induces apoptosis in fibroblasts and contributes to a significant downregulation of M1 markers and upregulation of M2 markers (87,316). This experimental evidence validates the model's response to NF-κB inhibition. Further investigations showed that the observed beneficial effects of NSAIDs and glucocorticoids, both used for RA treatment, are also due to NF-κB inhibition (317–319). However, their usage is limited due to severe side effects (320,321). Other NF-κB inhibitors were identified but the majority of them do not meet the standards to reach clinical development programs (322–325). Indeed, non-selective inhibition of NF-κB in all cell types has multiple detrimental effects as it is critical for maintaining homeostatic cellular pathways. Biological treatments that directly target the products of NF-κB-driven genes, such as TNF, IL-6 and IL-1 have been developed.

However, the inhibition of a single cytokine is probably not optimal as various cytokines have synergistic biological activities (326). This is further underlined by the simulations performed on our models that mimic such treatments (anti-TNF, anti-IL6, anti-IL1...) and fail to suppress the inflammatory RA macrophage population alone. Therefore, the discovery of techniques for cell-type specific NF-κB inhibition is needed to shift the benefit/risk balance (327).

Current strategies for a cell-specific uptake of substances are based on the development of immunoliposomes/particles. Whole antibodies, or ligands are attached to liposome/particle surfaces to achieve specificity for selective binding to receptors expressed on the surface of the target cells (M1 markers in this case). After receptor-mediated endocytosis, the encapsulated molecules are released into the cell and can achieve their pharmacodynamic effect via interaction with their respective intracellular targets. Another approach is based on recombinant proteins named "armed antibody" and composed of tissue specific binding domains linked to an effector domain. The armed antibody DEKAVIL is currently in the phase 1 clinical trial and showed first promising results in RA patients. DEKAVIL consists of the human antibody F8, specific for the extra-domain A of fibronectin, linked to the human anti-inflammatory cytokine IL-10. F8 exhibits a strong affinity to cells from synovial biopsies and was shown to inhibit the progression of collagen-induced arthritis (328). Despite a variety of already established strategies, there is still an unmet need for targeted treatment approaches to inhibit NF-κB.

Regarding the RA M2 macrophage model, GSK3β was predicted as a promising target leading to the promotion of M2 macrophage population in RA. GSK3β is involved in the progression of a variety of diseases including RA (329). Evidence suggests that this protein plays a central role in different signaling pathways that are relevant to macrophage function including polarization and inflammatory response (330). GSK3β inhibition in RA suppresses inflammatory responses in fibroblast-like synoviocytes and collagen-induced arthritis as well (331). Furthermore, GSK3β inhibition in allergic rhinitis inflammatory disease increases the expression of the M2 phenotypic signature markers (332). CREB1 is one of GSK3β's targets (333). When GSK3β is inhibited, it induces CREB1 gain of function which, in turn, sends an anti-inflammatory and antiapoptotic survival signal in monocytes and macrophages (334). It also increases M2 marker expression and promotes M2 phenotype in murine macrophages (335,336). These findings validate the behavior of the RA M2 macrophage model.

We also explored the synergistic effects that some therapeutic target pairs might have on the macrophages models' phenotypes. M1 macrophage model's proliferation was suppressed by ERK1 knockout alone. When paired with Notch1 KO, it also promoted M1 macrophage model's death. Published literature validates the

RA M1 macrophage model's response to ERK1 Notch1 double KO. Indeed, potential therapeutic value of co-targeting ERK1 and Notch1 have already been demonstrated in cancer but not in RA. It has been shown that targeting Notch1 enhances the efficacy of ERK1 inhibitors in cancer patients (337,338). In RA, separate ERK1 and Notch1 inhibitions reduce inflammation in mouse collagen-induced arthritis (339,340). Notch1 signaling, on the other hand, is known to regulate M1 macrophage fate through direct transcriptional regulation and indirect metabolic regulation (341).

We also identified the JAK1/JAK2 pair as a potential drug combination for the RA M1 macrophage model depletion. Baricitinib, an oral JAK inhibitor that is selective for JAK1 and JAK2, is a Food and Drug Administration (FDA) approved for treating RA (342). It prevents activation of STAT pathways and inhibits the cascade of transcription initiation of effector genes which, in turn, prevents the autoimmune and inflammatory reactions associated with RA including IFNg secretion (343). However, the way JAK inhibitors modulate macrophage's phenotypes, and whether this phenomenon explains their clinical benefit in RA is still not fully understood. A recent study showed that Baricitinib modulated the expression of membrane phenotype markers as well as the secretion of some signature cytokines in healthy macrophages (342). Another study further supports the effect of JAKs inhibition on RA macrophage's phenotypes by shifting the metabolic profile of M1 macrophage and rebalancing the metabolic reprogramming toward oxidative phosphorylation (344).

We were also interested in investigating potential therapeutic targets for BsAbs in RA, as their clinical therapeutic effects are superior to those of MoAbs with a lower rate of resistance. We tested the effects of double KOs of the macrophages' receptors represented in the models, but none of the KOs perturbed the apoptosis or proliferation phenotypes. These results underline the intricate crosstalk between the intracellular signaling pathways and the high synergistic activities of the various receptors represented in the M1 and M2 macrophage models.

Taken all together, these results further validate the behavior of our macrophage models through the identification of targets whose potential/proved therapeutical benefit in RA has already been highlighted in the literature. On the other hand, the predicted synergistic effect of the newly identified ERK1/Notch1 combination on our models still need to be validated experimentally in RA.

All files and scripts we used to run the simulations are publicly available at: https://gitlab.com/genhotel/Large_scale_computational_modelling_of_the_M1_and_M2_synovial_macrophages_in_Rheumatoid_Arthritis

Chapter 8. Investigation of new therapeutic options in RA using the RA multicellular model

8.1 Introduction

In the previous chapter, we investigated potential therapeutic options in RA by focusing on the RA macrophage population. We utilized the RA M1 and M2 macrophage models to test mono- and bi-therapies that could specifically downregulate pro-inflammatory macrophages and promote anti-inflammatory macrophages in RA synovium. To do so, we focused on the apoptosis and the proliferation phenotypes, the other macrophage-specific phenotypes were not investigated. Moreover, while the RA macrophages models provide a detailed representation of the phenotypic diversity of the macrophage population and allow us to study cell-specific processes like apoptosis and proliferation, they do not provide any information about the overall biological condition of the RA joint, like inflammation, matrix degradation and angiogenesis. Besides, these models can predict the effect of the tested drugs on the macrophage-specific phenotypes present in the models but not the cellular responses of other cell-types present in the synovium. Therefore, to investigate new therapeutic options in RA, we need to perform simulations in a multicellular environment that considers the various cell-types present in the RA synovium as well as their intercellular communication, both involved in the disease pathogenesis.

In the prior steps of this work, we described how we built the first RA multicellular map, a manually curated and enriched with extensive omics data analysis representation of both inter and intracellular molecular mechanisms involved in RA's pathogenesis. It includes four cell-types: the M1 macrophage, the M2 macrophage, the fibroblast, and the CD4+ Th1. We also demonstrated how we automatically converted the RA multicellular map to a large-scale Boolean model that we validated and calibrated using both prior knowledge and transcriptomic data.

In this chapter, the goal is to validate our calibrated multicellular model and simulation methods by identifying known therapeutic options, highighting their MoA and even identifying new potential tagets and drug combinations. As this model offers a comprehensive multicellular description of the RA synovium, we aim to better understand how the suppression and/or activation of certain cell populations and/or biological processes could regulate chronic inflammation, cartilage destruction and bone erosion in the rheumatic joint.

8.2 Methods and results

8.2.1 Single knockouts of the therapeutic targets present in the RA multicellular model

To identify potential therapeutic targets, present in the RA multicellular model, we perform an exhaustive search using the Therapeutic Target Database (TTD) (304). To simulate knockouts in the model, we only focus on the inhibition of the drug target. We screen the targets based on the Mode Of Action (MOA) of their associated drugs, and only keep the ones that can be targeted by at least one inhibitor (1643 targets), as described in Chapter 8. Investigation of new therapeutic options in RA using the RA multicellular model.

194 targets were present in the RA multicellular model (Supplementary table 13). We mimic the effect of these targets' KOs using *in silico* knockouts simulations. We use the calibrated state of the RA multicellular model as initial conditions for the simulations (Supplementary table 10). The models' phenotype states after the target knockouts are then compared to their corresponding calibrated states. Table 22 summarizes the identified therapeutic targets and their effects on the model's phenotypes while Table 23 describes the identified targets.

In the previous chapter, we demonstrated that, in our model, GSK3β inhibition induced M2 macrophage model's proliferation while suppressing its apoptosis (Figure 45). We also showed that NF-κB inhibition induced the M1 macrophage model's death and inhibited the M1 macrophage model's growth, and that ERK1 inhibition suppressed the proliferation of the inflammatory macrophage model.

The simulations we performed on the multicellular model reinforce these results by showing that GSK3β KO had no effect on the other cell-types present in the model, that NF-κB and ERK1 inhibitions suppressed the differentiation of the M1 macrophage model into osteoclasts (Figure 46) and that NF-κB KO perturbed the behavior of other cell-types in the model (Table 22).

We can see that CAV1 KO induced the apoptosis of the RA fibroblast model while suppressing its proliferation and migration (Figure 47). AKT2 KO on the other hand inhibited its proliferation and migration while preserving its ability to resist apoptosis. CREB1 inhibition led to the suppression of the RA fibroblast model's proliferation but had no effect on the migration or apoptosis phenotypes. We can also see that the migration phenotype in the fibroblast model can be suppressed via NF-κB inhibition.

Regarding the Th1 subtype model, two targets were identified, namely mTOR and TBX21. The inhibition of these targets led to the suppression of both the proliferation and migration phenotypes in the model and to the activation of the Th1 apoptosis phenotype. NF-κB KO, on the other side, inhibited both the proliferation and migration phenotypes while preserving the resistance to apoptosis in the Th1 model (Figure 48).

Potential therapeutic targets that perturb the general biological condition of the RA joint were also predicted. MIR221 inhibition led to the suppression of the angiogenesis phenotype in the multicellular model (Figure 49), and NF-κB inhibition suppressed the degradation of the extracellular matrix phenotype (Figure 50).

Table 22. Therapeutic targets from TTD database that perturb RA phenotypes in the RA multicellular model. Red arrows describe the inhibition of the phenotypes that were active in the calibrated state of the model. Green arrows describe the activation of the phenotypes that were inhibited in the calibrated state of the model. The absence of arrows means that the phenotype state remains unchanged.

Successful targets	RA multicellular model's phenotypes														
	M1 Macrophage			M2 macrophage		Fibroblast			Th1			Joint			
	Apoptosis	Proliferation	Osteoclastogenesis	Apoptosis	Proliferation	Apoptosis	Proliferation	Migration	Apoptosis	Proliferation	Migration	Inflammation	Angiogenesis	Matrix degradation	Osteoclastogenesis
AKT2							↘	↘							
CAV1						↗	↘	↘							
CREB1							↘								
GSK3β				↘	↗										
ERK1		↘	↘												
MIR221													↘		
mTOR									↗	↘	↘				
NF-κB	↗	↘	↘					↘		↘	↘			↘	
TBX21									↗	↘	↘				

Table 23. Description of the therapeutic targets perturbing the RA phenotypes in

the RA multicellular model.

Successful targets	Target type	Associated disease(s)	Drug with the highest status
CAV1	Literature-reported target	/	/
AKT2	Literature-reported target	/	Akt inhibitor VIII (Investigative)
CREB1	Literature-reported target	/	/
NF-κB	Successful target	Irritable bowel syndrome, Rheumatoid arthritis, Choreiform disorder, Lupus erythematosus, Multiple sclerosis...	Sulfasalazine (Approved)
TBX21	Literature-reported target	/	/
mTOR	Successful target	Arteries/arterioles disorder, Chronic myelomonocytic leukemia, Hydrocephalus, Multiple myeloma, Renal cell carcinoma	Everolimus (Approved)
ERK1	Clinical trial target	Melanoma, Pancreatic cancer, Cancer, Arteries/arterioles disorder, Mature T-cell lymphoma	BVD-523 (Phase 2)
GSK3β	Clinical trial target	Myotonic disorder, Acute myeloid leukemia, Osteosarcoma, Fragile X chromosome, Myeloproliferative neoplasm	Tideglusib (Phase 2/3)
MIR221	Literature-reported target	/	/

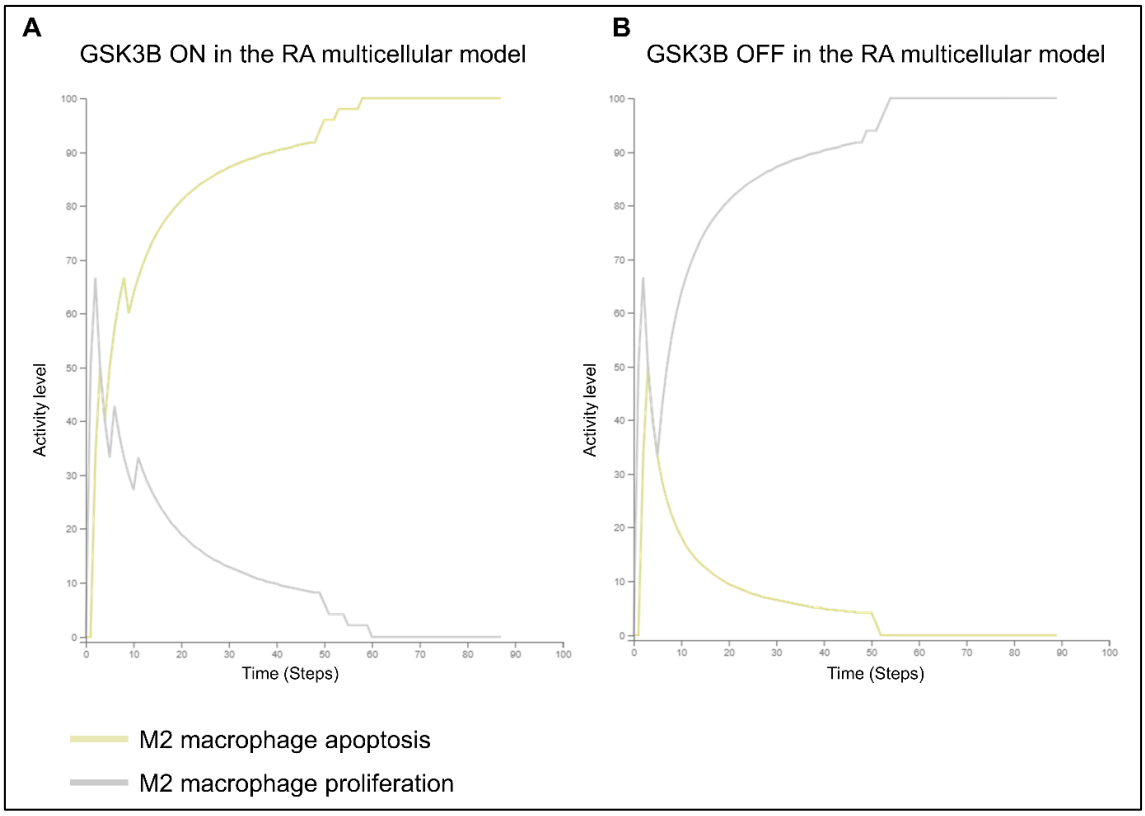


Figure 45. A) In silico simulation of the M2 macrophage cell-specific phenotypes before GSK3β KO. B) In silico simulation of the M2 macrophage cell-specific phenotypes after GSK3β KO.

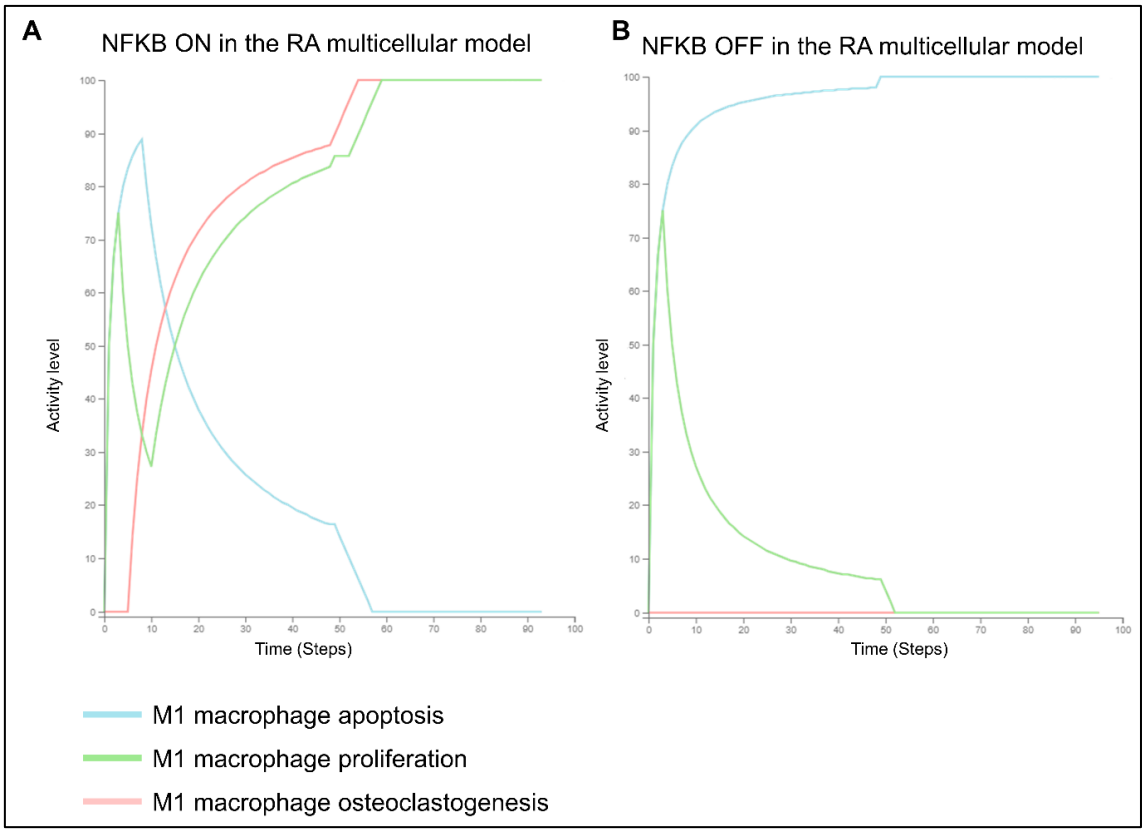


Figure 46. A) In silico simulation of the M1 macrophage cell-specific phenotypes before NF-κB KO. B) In silico simulation of the M1 macrophage cell-specific phenotypes after NF-κB KO.

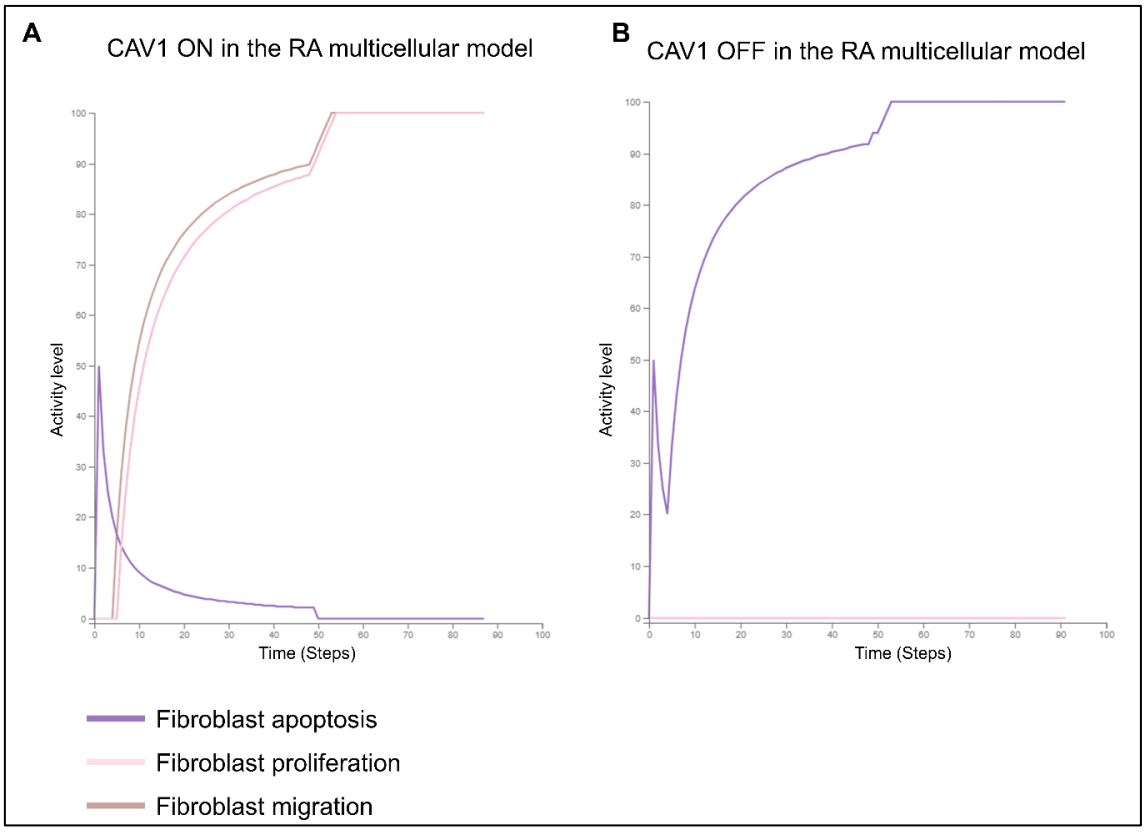


Figure 47. A) In silico simulation of the fibroblast cell-specific phenotypes before CAV1 KO. B) In silico simulation of the fibroblast cell-specific phenotypes after CAV1 KO.

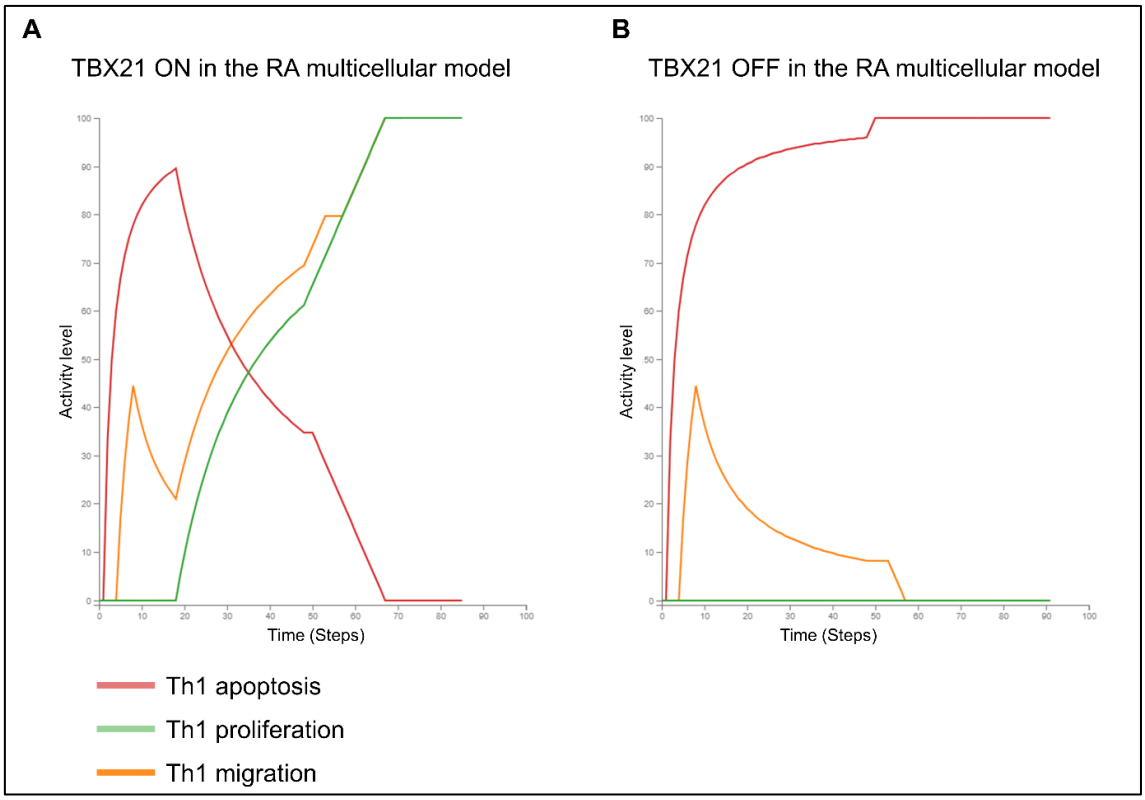


Figure 48. A) In silico simulation of the Th1 cell-specific phenotypes before TBX21 KO. B) In silico simulation of the Th1 cell-specific phenotypes after TBX21 KO.

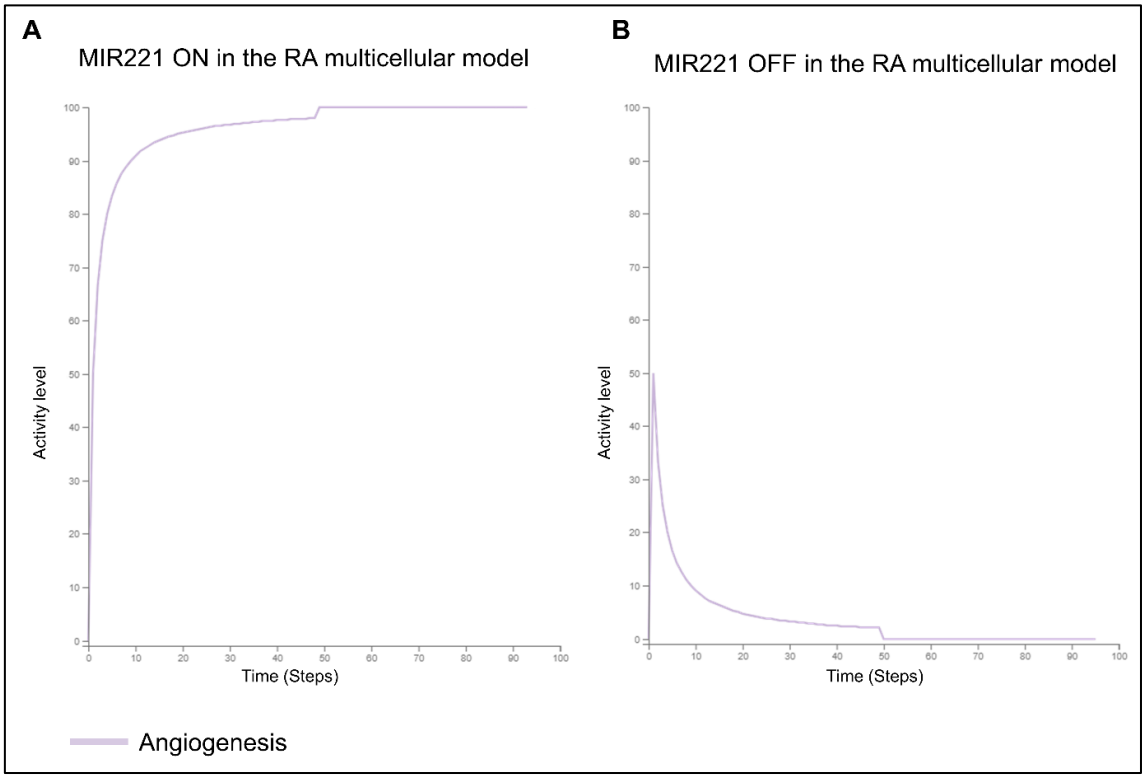


Figure 49. A) In silico simulation of the Angiogenesis phenotype before MIR221 KO.

B) In silico simulation of the Angiogenesis phenotype after MIR221 KO.

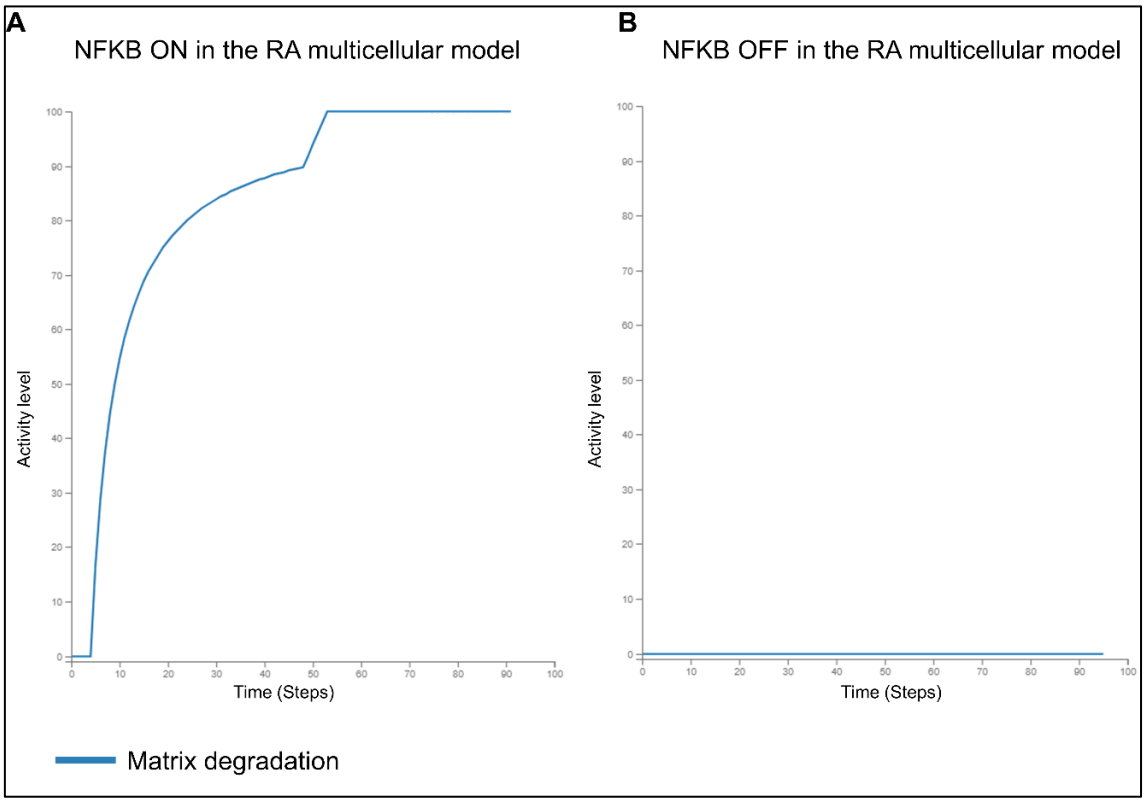


Figure 50. A) In silico simulation of the Matrix degradation phenotype before NF-κB KO. B) In silico simulation of the Matrix degradation phenotype after NF-κB KO.

8.2.2 Double knockouts of the therapeutic targets present in the RA multicellular model

To investigate the potential synergistic effect of the previously tested therapeutic targets on the model, the targets were combined in pairs. The RA multicellular model was then used to predict the outcome of these double KOs. We used the same initial conditions as for the mono drug testing, then we compared the perturbed states with their corresponding calibrated state.

We generated a list of all the possible pairs (without repetition) of the 194 previously identified therapeutic targets. 18721 drug combinations in total were tested. Among these combinations, three pairs having a synergistic effect on the model were predicted, namely ERK1/NOTCH1, JAK1/JAK2 and NF-κB/STAT3. Table 24 summarizes the identified synergistic pairs and their effects on the model's phenotypes while Table 25 describes the identified targets.

In the previous chapter, we already predicted ERK1/NOTCH1 and JAK1/JAK2 as promising combinations. We demonstrated that ERK1 KO alone inhibited the M1

macrophage model's proliferation. When combined with Notch1 KO, it also led to the promotion of the M1 macrophage model's apoptosis. We also showed that, when paired together, JAK1 and JAK2 KOs suppressed the M1 macrophage model's proliferation and induced its apoptosis.

The simulations we performed in this chapter further support these results and show a stronger impact of these double KOs at the multicellular level. Indeed, both ERK1/Notch1 and JAK1/JAK2 KOs lead to the suppression of the osteoclastogenesis phenotype in the M1 macrophage model, and JAK1/JAK2 double KO inhibits the inflammation phenotype in the model (Figure 51).

A new target pair was predicted in the model as well. The combination of NF-κB and STAT3 KOs inhibits the differentiation of the osteoclast precursor cells in the RA synovium (Osteoclastogenesis phenotype) as described in Figure 52.

Table 24. Combinations of therapeutic targets (from TTD database) that perturb RA phenotypes in the RA multicellular model. . Red arrows describe the inhibition of the phenotypes that were active in the calibrated state of the model. Green arrows describe the activation of the phenotypes that were inhibited in the calibrated state of the model. The absence of arrows means that the phenotype state remains unchanged.

Successful combination	Targeted phenotype														
	M1 macrophage			M2 macrophage		Fibroblast			Th1			Joint			
	Apoptosis	Proliferation	Osteoclastogenesis	Apoptosis	Proliferation	Apoptosis	Proliferation	Migration	Apoptosis	Proliferation	Migration	Inflammation	Angiogenesis	Matrix degradation	Osteoclastogenesis
ERK1 and NOTCH1	↗	↘	↘												
JAK1 and JAK2	↗	↘	↘									↘			
NF-κB and STAT3	↗	↘	↘					↘		↘	↘			↘	↘

Table 25. Description of the therapeutic targets perturbing the disease phenotypes

in the RA multicellular model.

Synergistic combination in the model	Targets	Target type	Associated disease(s)	Drug with highest status
JAK1/JAK2	JAK1	Successful target	Acquired hypomelanotic disorder, Atopic eczema, Crohn disease, Myeloproliferative neoplasm, Pancreatic cancer, ...	Baricitinib (Approved)
	JAK2	Successful target	Acquired hypomelanotic disorder, Atopic eczema, Myeloproliferative neoplasm, Pancreatic cancer, Rheumatoid arthritis, ...	Baricitinib (Approved)
ERK1/Notch1	ERK1	Clinical trial target	Melanoma, Pancreatic cancer, Cancer, Arteries/arterioles disorder, Mature T-cell lymphoma	BVD-523 (Phase 2)
	Notch1	Clinical trial target	Lymphoma, Mature T-cell lymphoma, Cancer	LY3039478 (Phase 1/2)
NF-κB/STAT3	NF-κB	Successful target	Irritable bowel syndrome, Rheumatoid arthritis, Choreiform disorder, Lupus erythematosus, Multiple sclerosis, ...	Sulfasalazine (Approved)
	STAT3	Successful target	Psoriasis, Brain cancer, Colorectal cancer, Liver cancer, Malignant digestive organ neoplasm.	Acitretin (Approved)

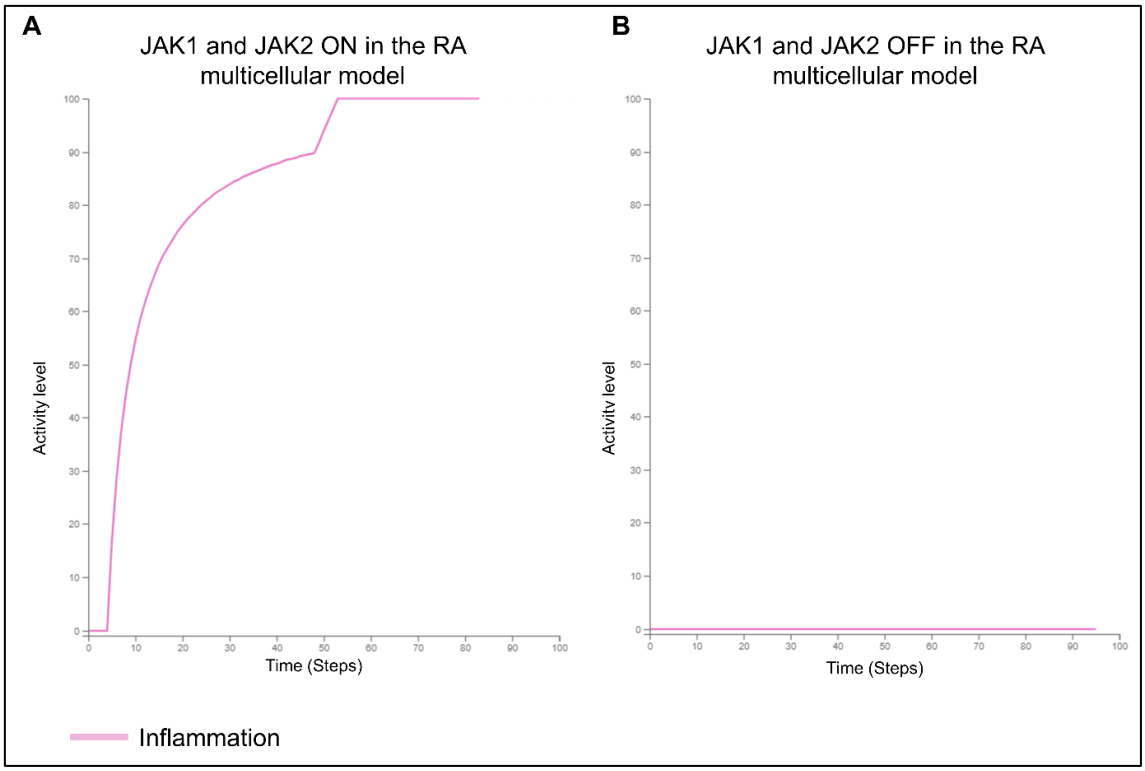


Figure 51. A) In silico simulation of the Inflammation phenotype before JAK1/JAK2 KO. B) In silico simulation of the Inflammation phenotype after JAK1/JAK2 KO.

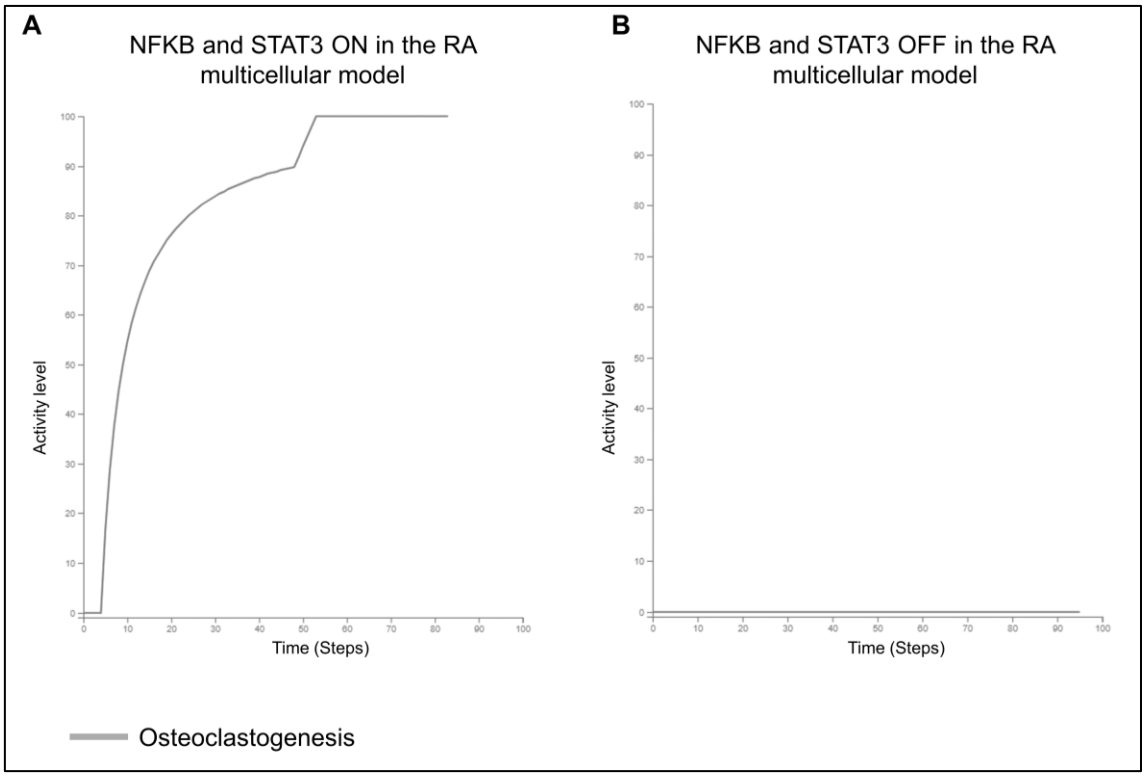


Figure 52. A) In silico simulation of the Osteoclastogenesis phenotype before NF-κB/STAT3 KO. B) In silico simulation of the Osteoclastogenesis phenotype after NF-κB /STAT3 KO.

8.3 Discussion

Various immune and non-immune cellular types, including macrophages, T cells, B cells, fibroblasts and endothelial cells are found in the joint synovium of RA patients. These cells communicate with one another through pathological cell-cell interactions that occur within a specific and defined environment known as the synovium. Therefore, to test new therapeutic options in RA and accurately predict the effects of the tested drugs, we need a multicellular environment that considers both the heterogeneity of the cellular responses to these drugs and the intricate intercellular communication taking place in the RA synovium.

To investigate such complex mechanisms, we ran *in silico* simulations on the calibrated RA multicellular model, and tested the effects of mono- and bi-therapies on the cellular behavior of the various cell types present in the model as well as the overall pathological state of the arthritic joint.

We predicted AKT2, belonging to the serine/threonine kinase family, as a potential target for inhibiting both RA fibroblast's proliferation and migration phenotypes in our model. It has been proven that the PI3K/AKT signaling pathway is

correlated with the occurrence and development of RA (345). Several studies support the concept of AKT2 inhibition for therapeutic intervention in RA. They showed that blocking AKT pathway inhibits RA progression (346). They also demonstrated that in vitro siRNA-mediated down-regulation of AKT2 significantly prevented cell proliferation and migration of human RA fibroblasts (347,348).

We also identified Caveolin 1 (CAV1) as a potential target for inducing the apoptosis phenotype and inhibiting the proliferation phenotype and the migration phenotype in the RA fibroblast model. CAV1 is one of the major structural components of caveolae and has a number of signaling functions. Regarding inflammatory disorders, CAV1 could either prevent or induce inflammation, depending on cellular context (349). In RA, in vitro silencing of CAV1 drastically reduces cell proliferation and promotes apoptosis in human RA fibroblasts (349,350). On the other hand, enforced expression of CAV1 in RA fibroblasts restores cell proliferation and attenuates apoptosis (349). CAV1 was also demonstrated to drive resistance to apoptosis in a large-scale Boolean model describing the RA fibroblasts (220).

Simulations of CREB1 KO in the model led to the inhibition of the proliferation phenotype in the RA fibroblast model while maintaining its capacity to resist apoptosis. This inhibitory effect was demonstrated in an experimental study where in vitro suppression of CREB activity downregulates aberrant synovial cell functions in patients with RA via the suppression of the RA synovial fibroblast's proliferation (351). CREB1 effect on the cellular proliferation was also highlighted through in silico simulations performed on a published large-scale Boolean model of RA fibroblasts (220). These results further validate the behavior of our model and highlight AKT2, CREB1, and CAV1 as promising targets for downregulating hyperactive fibroblasts in the rheumatic joint.

Regarding the RA macrophage models, GSK3β inhibition induced the M2 macrophage model's proliferation while suppressing its apoptosis. In contrast, NF-κB inhibition induced the M1 macrophage model's apoptosis and inhibited the M1 macrophage model's growth. ERK1 inhibition, on the other hand, had no effect on the apoptosis phenotype, but suppressed the M1 macrophage model's proliferation. These results were already described in detail in Chapter 8. Investigation of new therapeutic options in RA using the RA multicellular model. However, simulating NF-κB KO in the multicellular model further revealed that it affects other macrophage-specific phenotypes by inhibiting the osteoclastogenesis phenotype in the M1 macrophage model. It also affects other cell-types via the suppression of the migration phenotype in the RA fibroblast model, the proliferation and mi-

gration phenotypes in RA Th1 subtype, and degradation of the extracellular matrix phenotype. These results are supported by experimental evidence. Indeed, researchers have demonstrated that NF-κB mediates the transcriptional activation of several Matrix MetalloProteinases (MMPs), responsible for the destruction of the extracellular matrix and the articular cartilage in RA (352). NF-κB activation was found necessary for the induction of MMP3, MMP9, and MMP13 expression (353–355) and the inhibition of NF-κB signaling was found to block the expression of these MMPs (356). Blockade of the MMPs expression in the RA synovium leads to the suppression of RA synovial fibroblast's migration and invasion (357).

The critical importance of NF-κB in bone turnover was also highlighted experimentally. Double-knockout NF-κB1-/- NF-κB2-/- mice develop a defect in osteoclasts, showing critical functions of NF-κB proteins in the development of this cell lineage (358). Furthermore, it was shown that inhibition of NF-κB was an effective approach to inhibit osteoclast formation and bone resorptive activity, and that blocking nuclear activation of NF-κB displays anti-inflammatory and anti-osteolytic benefits (359,360).

In addition, several studies have described a role of NF-κB in Th1 differentiation. They showed that Th1 responses were significantly impaired and IFN-g production was abrogated due to diminished NF-κB activation (361,362). Another study in transgenic mice whose T cells lack the NF-κB/Rel signaling pathway shows that NF-κB plays a role in Th1-dependent hypersensitivity responses (362).

Also, we predicted two potential therapeutic targets to downregulate the RA CD4+ Th1 model, namely mTOR and TBX21. TBX21 (or T-bet) is a Th1-specific T box transcription factor that controls the expression of the hallmark Th1 cytokine, IFNg (144). Studies demonstrated that silencing of T-bet significantly suppressed IFNg and IL-17 gene expression and improved the pathogenesis of arthritis in CIA mice (363). Furthermore, CD4+ cells from T-bet-/- mice are skewed toward an anti-inflammatory Th2 differentiation via the expression of high levels of GATA-3. GATA-3 gain of function, after TBX21 depletion, was also highlighted in the KO simulations we performed on the RA multicellular model. Hence, regulation of the T-bet/GATA-3 ratio can reduce inflammatory damage to RA cells, and the mechanism may be related to the regulation of the Th1/Th2 ratio of RA cells through T-bet depletion (364).

Mechanistic Target Of Rapamycin (mTOR) has been found to play important roles in Th1 cell development as well. It was found that CD4+ T cells deficient of mTOR failed to differentiate into effector Th1 cells, and that mTOR regulates Th1 cell differentiation through STAT4 and SOCS3, which then regulate T-bet expression

(365). In RA, mTOR inhibition has shown efficacy in reducing joint inflammation in animal models of arthritis (366) and in patients with RA (367).

We also predicted potential therapeutic targets that reduce inflammation and angiogenesis in the rheumatic joint. JAK1/JAK2 double KO was already predicted in the previous chapter as a promising synergistic combination for the depletion of RA M1 macrophage model. Here we show that it also inhibits the inflammation and osteoclastogenesis phenotypes in the RA multicellular model. Baricitinib, an oral JAK inhibitor that is selective for JAK1 and JAK2, is a Food and Drug Administration (FDA) approved for treating RA (342). It prevents activation of STAT pathways and inhibits the transcription initiation of effector genes which, in turn, decreases the systemic inflammation and the progression of bone destruction associated with RA (343,368).

MIR221, on the other hand, downregulated the angiogenesis phenotype in the model. MIR221 was shown to downregulate THBS1, which acts as an anti-angiogenic factor on endothelial cells (369). Thus, inhibiting MIR221 would restore THBS1 expression, which, in RA, was found to help restore tissue homeostasis during resolution of inflammation (370).

A new combination was also predicted in the model. NF-κB and STAT3 double KO inhibited the differentiation of osteoclast precursor cells in the RA multicellular model. Whether normal or pathological, osteoclastogenesis strictly depends upon support from accessory cells which supply cytokines required for osteoclast differentiation (371). CSF1, RANKL, TNF-α and IL-6 have been found to induce the differentiation of osteoclasts and the bone resorption activity in RA patients (372). STAT3 and NF-κB are essential transcription factors for the expression of these osteoclastogenic factors' expression (373–376); therefore, targeting both STAT3 and NF-κB could be a promising strategy to inhibit bone erosion in RA.

The identification of therapeutic targets whose potential/proven value has already been highlighted in RA further validates our multicellular model. Taken all together, these results show that our approach provides a better understanding of the signalling pathways associated with these targets as well as their mechanisms of action in a disease and cell-specific manner. On the other hand, the potential new drug combinations that we predicted as having a synergistic effect on the RA multicellular model still need to be validated via in vitro experiments, but further demonstrates the power of executable mechanistic modelling to enable rapid screening of combination therapies.

General discussion

The main results and outputs that were expected from this work are defined in the CIFRE PhD project proposal section that describes the initial proposal made to the ANRT for this PhD project. These outputs are the following:

- Multiscale models of Rheumatoid Arthritis networks easy to maintain and upgrade.
- Simulated models for RA fully operational and easy to maintain able to simulate the pathogenesis.
- Better understanding of the disease with a full up to date knowledge base on RA.
- Increase Sanofi's skills/capabilities in the domain of disease map and Boolean based model simulation.

We were able to complete all of the required deliverables. Indeed, we started this PhD work with the creation of the first molecular interaction maps that describe the most abundant cell populations in the RA synovium: the macrophages (including the M1 and M2 phenotypes), the fibroblasts and the CD4+ T helper 1 (Th1) subtype. These maps recapitulate existing knowledge related to the intracellular interactions involved in RA pathogenesis in a cell-type specific manner. Maps are primarily based on high-quality manual curation of scientific literature and enrichment and cross-validation using expression datasets. All the components and reactions in the RA atlas are annotated using RA and human-specific studies and datasets. A thorough description of the methodology used to build the RA-Atlas is available in our open access publication (285). Moreover, the RA cell-specific molecular interaction maps are freely accessible at https://ramap.uni.lu/minerva/ and can be used as an interactive knowledge base, using the platform MINERVA, and serve as a template for overlaying multiple datasets.

Going one step further, we were able to provide a multicellular representation of the RA synovium by constructing the RA multicellular molecular interaction map. It is the largest manually curated map to date and the only one providing a multicellular representation of the RA synovium. The RA multicellular map is a state-of-the-art atlas of the rheumatic joint that describes manually curated intra- and intercellular interactions involved in the disease's pathogenesis. This map will be made available soon as an online interactive map on the platform MINERVA.

To build the RA multicellular map, we connected the RA cell-specific maps together via the addition of bidirectional intercellular interactions that were identified using both literature and databases mining in a disease- and cell-specific

manner. The number of identified interactions from the literature and databases varied depending on the amount of information we could gather on the cell type under investigation in published articles. Studies on the interactions of fibroblasts and macrophages with their surroundings for instance were considerably easier to find than those on CD4+ Th1 contact with other cells. It also varied depending on the number of ligands and receptors we had in each cell-specific map.

We also inferred cell-to-cell communication from transcriptomic datasets. For each bidirectional cell–cell interaction pair, we used two dedicated tools, ICELLNET and DiSiR, that we applied on disease- and cell-specific datasets. The heterogeneity and scarcity of data regarding the synvial cell-types under study were a limitation at this point of the work. To overcome this and to ensure reliable and consistent results, we filtered the identified interactions using stringent criteria. As a counterpart, we may also have filtered out potential interesting and new intercellular interactions to investigate in RA disease.

The RA multicellular map can be expanded with additional cell-specific maps to include a variety of other cells such as B lymphocytes, dendritic cells, plasma cells, mast cells, osteoclasts and chondrocytes, all involved in RA pathogenesis (4). These maps could be connected to the existing ones through the identification of additional intercellular interactions. Moreover, as several studies have highlighted the presence of various metabolic checkpoints in RA (377), It would be interesting to integrate relevant metabolic pathways to better understand the intricate interactions between metabolic pathways and the inflammatory and immune responses.

To understand the emergent features and behaviors of the biological systems described in our maps, we needed to integrate dynamical modeling approaches. Boolean formalism was the most suited option in our case as Boolean models can handle such large-scale systems and the regulation of these model's components is given in a parameter-free way, without the need for kinetic parameters and precise quantitative data. Boolean models can be constructed manually. However, defining the logical rules of a model by hands requires not only high amounts of experimental data, but also of time and efforts in retrieving the relevant information from published literature. Because it is impossible to cope with the very high number of components present in large-scale biological systems, researchers study smaller and smaller parts and attempt to understand the whole systems. Small models are efficient in answering very specific biological questions about a limited number of genes or proteins.

Leaving out parts of the biological system we wanted to study was not a viable option for us as our main objective in this work was about providing a multicellular system-level representation of the RA joint including the intricate cell-cell interactions in RA as well as the complex intracellular cascades and their regulation. We also needed to preserve all the information available in the maps we built in order to be able to understand the mechanism of action of the current RA therapies and to try to identify new therapeutic options that were not highlighted in the literature.

Therefore, we used the tool CaSQ to automatically infer the logical rules of our models from the topology of the maps. The generated models are far larger in size compared to the manually built ones, providing more accurate representations of the systems (238). These large-scale models, on the other hand, are more complex, with a substantially higher number of inputs, making it computationally challenging to identify all their possible attractors (293). Furthermore, even when the attractors of such big models can be identified, the identification of the states that are biologically consistent in a cell or disease specific context remains challenging.

In this direction, we developed an efficient computational framework to build, analyze and validate the behaviors of large-scale Boolean models. The framework analyzes the generated Boolean models in a synchronous scheme using a new prototype of the BMA console tool (294), set up for this purpose, deployed to HPCs and validates their behaviors using gene expression datasets and prior knowledge. We successfully applied our framework to explore the RA cell-specific maps in an integrative and combinatorial manner. We also applied our framework on the RA multicellular map to generate and calibrate the largest Boolean model to date (our model comprises more than 1000 nodes while the the average number of nodes in published Boolean models is 150), demonstrating the method's efficiency and scalability.

Our framework sets the path to explore many other state-of-art mechanistic maps that were already published for several diseases including Parkinson's (166), Alzheimer's (167), cancer (169), asthma (170), and more recently COVID19 (171).

We made a significant effort in this work to implement the FAIR principles (288). The maps are freely available and are part of the Disease Map community project (9,171), which makes their dissemination easier. Their content is compliant with SBGN PD for the representation, MIRIAM for the annotations, PMIDs and stable identifiers for the references, and HGNC symbols for the annotation of the signaling and gene regulation components. The Boolean models, on the other hand,

are available in SBML qual and JSON files, allowing for transparency, interoperability and reusability, and promoting open science.

The RA multicellular map and model that we were able to construct in this project are the largest multicellular static and dynamic representations of the RA synovium to date. Indeed, models and networks for RA have already been proposed. Large-scale networks were inferred using gene expression data (293,378,379). They were built using the influence interaction approach that relates the expression of a gene to the expression of the other genes in the cell and then infer gene-to-gene interaction. Such networks are often noisy, and rely on putative, not manually curated interactions. RA models based on differential equations were also published (380,381). They describe the dynamics of signaling pathways by translating them into rate equations. These quantitative executable models are limited in terms of size and cannot be used to describe entire cellular pathways. Already published large-scale Boolean models for RA overcome the size limitation. However, they remain smaller compared to our model, they only focus on the active role of fibroblasts in RA and do not consider cell-cell communication (220,382).

We integrated various gene expression datasets in this work. We carefully selected the relevant datasets and used disease and cell-specific ones when possible. However, macrophage, fibroblast, and Th1 samples from the synovium of RA patients were scarce and rarely available within the same datasets. Therefore, we had to utilize monocyte samples rather than macrophage samples for some parts of the work, and to combine datasets coming from independent studies. Besides, as phenotype-specific samples of the RA synovial macrophages were not available, we also had to use healthy M1 and M2 macrophage samples instead. Therefore, having access to more RA synovial samples of the cell-types of interest would improve the robustness of our approach. As RA is a disease of great heterogeneity, we would also need to implement a strategy to handle inconsistencies between these omics datasets.

To test the behavior of the RA Boolean models against gene expression datasets, we assumed a positive correlation between the expression of mRNAs and the expression of their corresponding proteins, which allowed us to connect experimentally observed expressions at the mRNA level with nodes at the protein level in the models. Published studies showed an overall positive correlation between mRNA and protein expression levels, which suggests that mRNA expression can be useful, but is certainly far from perfect, in predicting protein expression levels (383). Therefore, integrating multi-omics datasets with transcriptomics (single cell RNA-seq in particular) and proteomics layers, for instance, would help improve the proposed methodology. Furthermore, multi valued qualitative models may be envisioned to consider biomolecules that are not differentially expressed in

these omics datasets, such as housekeeping genes. However, more computing power would be required to cope with the exponentially increasing complexity of these models.

To calibrate the RA Boolean models, synchronous update was used to perform the attractor search, and stable states were tested against prior knowledge and gene expression datasets. It would be interesting to investigate more complex attractors using both synchronous and asynchronous modes. Even if it is difficult to estimate how the calibrated states of the model would change when the updating scheme is modified, identifying the models' attractors in both synchronous and asynchronous mode, and testing these attractors against prior knowledge and omics datasets could help understand the effects of these changes on the models' predictions.

Other promising approaches that use molecular interaction maps are being developed. They are based on machine and deep learning methods that utilize biological networks and experimental high-throughput data as inputs to predict, for instance, proteins interactions (384), proteins functions (385,386), Drug–target interactions (387), and disease diagnosis (388). This kind of approach could be used to further explore the RA cell-specific and multicellular maps that we created and might complement the methodology we propose.

Researchers interested in using our molecular interactions maps to study specific pathways or phenotypes in a quantitative manner can also opt for exporting subparts of the XML files we provide. The exported biochemical networks can then be used as a basis to build quantitative models and simulate their behavior using ordinary differential equations, stochastic differential equation, or Gillespie's stochastic simulation algorithm, all available via the tool COPASI (COmplex PAthway Simulator) (389).

To study the therapeutic options in RA, first, we focused on the synovial macrophages. The number of macrophages in the inflamed synovial tissue grows rapidly during the course of RA, and their polarization into a proinflammatory M1 or an anti-inflammatory M2 phenotype plays a critical role in the pathological progression of RA (310). Thus, selectively suppressing the M1 macrophages or boosting the M2 macrophages could be a promising strategy for the treatment of RA. In this direction, we simulated single and double KOs in the calibrated RA M1 and M2 macrophage models. The models predicted NF-κB, JAK1/JAK2, and ERK1/Notch1 as potential targets to selectively suppress pro-inflammatory RA macrophages. The models also predicted GSK3β as a promising target to promote anti-inflammatory RA macrophages.

We were also interested in studying the cellular behavior of the other cell-types present in synovium as well as the biological processes affecting the overall condition of the RA joint like inflammation, matrix degradation or angiogenesis. In this direction, we tested new therapeutic options in a multicellular setting by simulating single and double KOs in the calibrated RA multicellular model. We identified AKT2, CAV1 and CREB1 as promising targets to limit the invasive behavior of the RA fibroblast model. We showed that mTOR and TBX21 inhibition downregulates hyperactivity in the CD4+ Th1 model, that MIR221 inhibition reduces angiogenesis in the model, and that the inhibition of the NF-κB /STAT3 pair suppresses bone erosion phenotype in the RA multicellular model. Additional effects of NF-κB and JAK1/JAK2 inhibitions on the multicellular level were also revealed. Inhibiting NF-κB suppresses the differentiation of the M1 macrophage model into osteoclasts, decreases the hyperactivation of the fibroblast and Th1 models and regulates the degradation of the extracellular matrix phenotype. Inhibiting JAK1/JAK2, on the other hand, was found to downregulate inflammation and bone destruction phenotype in the model.

The *in silico* simulations we performed further validate our models via the identification of therapeutic targets whose potential/proven value has already been documented in RA. Our model provides a better understanding of the signalling pathways associated with these targets as well as their mechanisms of action in a disease and cell-specific manner. To enhance the number of identified therapeutic targets and combinations, we could expand our approach to include all the nodes in our models, instead of only targeting proteins in TTD database. We would perform *in silico* simulations of single and double KOs of all the models' components and retrieve the ones that are able tu perturb the cellular phenotypes. This would generate a massive amount of data to be analyzed but could help us identify new targets that have not been highlighted in RA yet.

To further validate the behavior of the calibrated models and to test the generated predictions, this work would benefit greatly from the integration of additional single cell RNA-seq data, in particular from patients before and after treatment. It would also benefit from in vitro experiments that target the proposed pathways/factors while measuring cellular processes like apoptosis, proliferation or inflammation. The synergies we predicted with our models need also to be validated experimentally using Bliss definition of drugs independence (390).

To conclude, we contribute, with this PhD work, to a better understanding of RA at the multicellular level through an efficient computational system biology approach. We provide a detailed representation of the structure and dynamics of the intricate cell-cell communication in RA as well as the complex intracellular cascades and their regulation in this disease. We also validate our approach by

identifying therapeutic targets that are already validated or currently under investigation in RA, highighting their mechanism of action in a disease- and cell-specific manner, and proposing new potential drug combinations in RA for in vitro validation as well.

Bibliography

1. Sparks JA. Rheumatoid Arthritis. Ann Intern Med. 2019 Jan 1;170(1):ITC1–16.

2. Figus FA, Piga M, Azzolin I, McConnell R, Iagnocco A. Rheumatoid arthritis: Extra-articular manifestations and comorbidities. Autoimmun Rev. 2021 Apr;20(4):102776.

3. Rheumatoid arthritis: MedlinePlus Medical Encyclopedia [Internet]. [cited 2023 Aug 26]. Available from: https://medlineplus.gov/ency/article/000431.htm

4. Tran CN, Lundy SK, Fox DA. Synovial biology and T cells in rheumatoid arthritis. Pathophysiology. 2005 Oct;12(3):183–9.

5. El-Gabalawy HD, Lipsky PE. Why do we not have a cure for rheumatoid arthritis? Arthritis Res. 2002;4 Suppl 3(Suppl 3):S297-301.

6. Lin YJ, Anzaghe M, Schülke S. Update on the Pathomechanism, Diagnosis, and Treatment Options for Rheumatoid Arthritis. Cells. 2020 Apr 3;9(4):880.

7. Bécède M, Alasti F, Gessl I, Haupt L, Kerschbaumer A, Landesmann U, et al. Risk profiling for a refractory course of rheumatoid arthritis. Semin Arthritis Rheum. 2019 Oct;49(2):211–7.

8. Zupanic A, Bernstein HC, Heiland I. Systems biology: current status and challenges. Cell Mol Life Sci. 2020 Feb;77(3):379–80.

9. Mazein A, Ostaszewski M, Kuperstein I, Watterson S, Le Novère N, Lefaudeux D, et al. Systems medicine disease maps: community-driven comprehensive representation of disease mechanisms. NPJ Syst Biol Appl. 2018;4:21.

10. Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, et al. The Systems Biology Graphical Notation. Nat Biotechnol. 2009 Aug;27(8):735–41.

11. Juty N, Le Novère N, Laibe C. Identifiers.org and MIRIAM Registry: community resources to provide persistent identification. Nucleic Acids Res. 2012 Jan;40(Database issue):D580-586.

12. Kreutz C. A New Approximation Approach for Transient Differential Equation Models. Frontiers in Physics [Internet]. 2020 [cited 2022 Dec 20];8. Available from: https://www.frontiersin.org/articles/10.3389/fphy.2020.00070

13. Albert R, Thakar J. Boolean modeling: a logic-based dynamic approach for

understanding signaling and regulatory networks and for making useful predictions. Wiley Interdiscip Rev Syst Biol Med. 2014;6(5):353–69.

14. Hall BA, Niarakis A. Data integration in logic-based models of biological mechanisms. Current Opinion in Systems Biology. 2021 Dec 1;28:100386.

15. Mori T, Akutsu T. Attractor detection and enumeration algorithms for Boolean networks. Comput Struct Biotechnol J. 2022;20:2512–20.

16. Montagud A, Béal J, Tobalina L, Traynard P, Subramanian V, Szalai B, et al. Patient-specific Boolean models of signalling networks guide personalised treatments. Elife. 2022 Feb 15;11:e72626.

17. Association Nationale Recherche Technologie [Internet]. 2021 [cited 2023 Oct 23]. Cifre. Available from: https://www.anrt.asso.fr/fr/cifre-35654

18. McInnes IB, Schett G. The Pathogenesis of Rheumatoid Arthritis. N Engl J Med. 2011 Dec 8;365(23):2205–19.

19. van Delft MAM, Huizinga TWJ. An overview of autoantibodies in rheumatoid arthritis. J Autoimmun. 2020 Jun;110:102392.

20. van der Woude D, Houwing-Duistermaat JJ, Toes REM, Huizinga TWJ, Thomson W, Worthington J, et al. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. Arthritis Rheum. 2009 Apr;60(4):916–23.

21. Messemaker TC, Huizinga TW, Kurreeman F. Immunogenetics of rheumatoid arthritis: Understanding functional implications. J Autoimmun. 2015 Nov;64:74–81.

22. Scherer HU, Häupl T, Burmester GR. The etiology of rheumatoid arthritis. J Autoimmun. 2020 Jun;110:102400.

23. Verpoort KN, van Gaalen FA, van der Helm-van Mil AHM, Schreuder GMT, Breedveld FC, Huizinga TWJ, et al. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. Arthritis Rheum. 2005 Oct;52(10):3058–62.

24. Okada Y, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, et al. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. Nat Genet. 2012 May;44(5):511–6.

25. Wei WH, Loh CY, Worthington J, Eyre S. Immunochip Analyses of Epistasis in Rheumatoid Arthritis Confirm Multiple Interactions within MHC and Suggest Novel Non-MHC Epistatic Signals. J Rheumatol. 2016 May;43(5):839–45.

26. Perdigones N, Vigo AG, Lamas JR, Martínez A, Balsa A, Pascual-Salcedo D, et al. Evidence of epistasis between TNFRSF14 and TNFRSF6B polymorphisms in patients with rheumatoid arthritis. Arthritis Rheum. 2010 Mar;62(3):705–10.

27. Kallberg H, Padyukov L, Plenge RM, Ronnelid J, Gregersen PK, van der Helm-van Mil AHM, et al. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. Am J Hum Genet. 2007 May;80(5):867–75.

28. Seddighzadeh M, Korotkova M, Källberg H, Ding B, Daha N, Kurreeman FAS, et al. Evidence for interaction between 5-hydroxytryptamine (serotonin) receptor 2A and MHC type II molecules in the development of rheumatoid arthritis. Eur J Hum Genet. 2010 Jul;18(7):821–6.

29. Glant TT, Mikecz K, Rauch TA. Epigenetics in the pathogenesis of rheumatoid arthritis. BMC Med. 2014 Feb 26;12:35.

30. Peeters JGC, Boltjes A, Scholman RC, Vervoort SJ, Coffer PJ, Mokry M, et al. Epigenetic changes in inflammatory arthritis monocytes contribute to disease and can be targeted by JAK inhibition. Rheumatology (Oxford). 2023 Aug 1;62(8):2887–97.

31. Glossop JR, Emes RD, Nixon NB, Packham JC, Fryer AA, Mattey DL, et al. Genome-wide profiling in treatment-naive early rheumatoid arthritis reveals DNA methylome changes in T and B lymphocytes. Epigenomics. 2016 Feb;8(2):209–24.

32. Xiao F, Rui K, Shi X, Wu H, Cai X, Lui KO, et al. Epigenetic regulation of B cells and its role in autoimmune pathogenesis. Cell Mol Immunol. 2022 Nov;19(11):1215–34.

33. Zhao J, Wei K, Chang C, Xu L, Jiang P, Guo S, et al. DNA Methylation of T Lymphocytes as a Therapeutic Target: Implications for Rheumatoid Arthritis Etiology. Front Immunol. 2022 Mar 3;13:863703.

34. Toussirot E, Abbas W, Khan KA, Tissot M, Jeudy A, Baud L, et al. Imbalance between HAT and HDAC activities in the PBMCs of patients with ankylosing spondylitis or rheumatoid arthritis and influence of HDAC inhibitors on TNF

alpha production. PLoS One. 2013;8(8):e70939.

35. Nemtsova MV, Zaletaev DV, Bure IV, Mikhaylenko DS, Kuznetsova EB, Alekseeva EA, et al. Epigenetic Changes in the Pathogenesis of Rheumatoid Arthritis. Front Genet. 2019;10:570.

36. Maeshima K, Stanford SM, Hammaker D, Sacchetti C, Zeng L fan, Ai R, et al. Abnormal PTPN11 enhancer methylation promotes rheumatoid arthritis fibroblast-like synoviocyte aggressiveness and joint inflammation. JCI Insight [Internet]. [cited 2021 Mar 7];1(7). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4889026/

37. Hammaker D, Whitaker JW, Maeshima K, Boyle DL, Ekwall AKH, Wang W, et al. LBH Gene Transcription Regulation by the Interplay of an Enhancer Risk Allele and DNA Methylation in Rheumatoid Arthritis. Arthritis Rheumatol. 2016 Nov;68(11):2637–45.

38. Whitaker JW, Boyle DL, Bartok B, Ball ST, Gay S, Wang W, et al. Integrative omics analysis of rheumatoid arthritis identifies non-obvious therapeutic targets. PLoS One. 2015;10(4):e0124254.

39. Huber LC, Brock M, Hemmatazad H, Giger OT, Moritz F, Trenkmann M, et al. Histone deacetylase/acetylase activity in total synovial tissue derived from rheumatoid arthritis and osteoarthritis patients. Arthritis Rheum. 2007 Apr;56(4):1087–93.

40. Trenkmann M, Brock M, Gay RE, Kolling C, Speich R, Michel BA, et al. Expression and function of EZH2 in synovial fibroblasts: epigenetic repression of the Wnt inhibitor SFRP1 in rheumatoid arthritis. Ann Rheum Dis. 2011 Aug;70(8):1482–8.

41. Karouzakis E, Trenkmann M, Gay RE, Michel BA, Gay S, Neidhart M. Epigenome analysis reveals TBX5 as a novel transcription factor involved in the activation of rheumatoid arthritis synovial fibroblasts. J Immunol. 2014 Nov 15;193(10):4945–51.

42. Karami J, Aslani S, Tahmasebi MN, Mousavi MJ, Sharafat Vaziri A, Jamshidi A, et al. Epigenetics in rheumatoid arthritis; fibroblast-like synoviocytes as an emerging paradigm in the pathogenesis of the disease. Immunol Cell Biol. 2020 Mar;98(3):171–86.

43. Spinelli FR, Pecani A, Conti F, Mancini R, Alessandri C, Valesini G. Post-translational modifications in rheumatoid arthritis and atherosclerosis: Focus

on citrullination and carbamylation. J Int Med Res. 2016 Sep;44(1 Suppl):81–4.

44. Darrah E, Andrade F. Rheumatoid arthritis and citrullination. Curr Opin Rheumatol. 2018 Jan;30(1):72–8.

45. Curran AM, Naik P, Giles JT, Darrah E. PAD enzymes in rheumatoid arthritis: pathogenic effectors and autoimmune targets. Nat Rev Rheumatol. 2020 Jun;16(6):301–15.

46. Pruijn GJM. Citrullination and Carbamylation in the Pathophysiology of Rheumatoid Arthritis. Front Immunol. 2015 Apr 27;6:192.

47. Klareskog L, Gregersen PK, Huizinga TWJ. Prevention of autoimmune rheumatic disease: state of the art and future perspectives. Ann Rheum Dis. 2010 Dec;69(12):2062–6.

48. Arnson Y, Shoenfeld Y, Amital H. Effects of tobacco smoke on immunity, inflammation and autoimmunity. J Autoimmun. 2010 May;34(3):J258-265.

49. Stolt P, Yahya A, Bengtsson C, Källberg H, Rönnelid J, Lundberg I, et al. Silica exposure among male current smokers is associated with a high risk of developing ACPA-positive rheumatoid arthritis. Ann Rheum Dis. 2010 Jun;69(6):1072–6.

50. Klockars M, Koskela RS, Järvinen E, Kolari PJ, Rossi A. Silica exposure and rheumatoid arthritis: a follow up study of granite workers 1940-81. Br Med J (Clin Res Ed). 1987 Apr 18;294(6578):997–1000.

51. Chang KH, Hsu CC, Muo CH, Hsu CY, Liu HC, Kao CH, et al. Air pollution exposure increases the risk of rheumatoid arthritis: A longitudinal and nationwide study. Environ Int. 2016 Sep;94:495–9.

52. Merlino LA, Curtis J, Mikuls TR, Cerhan JR, Criswell LA, Saag KG, et al. Vitamin D intake is inversely associated with rheumatoid arthritis: results from the Iowa Women's Health Study. Arthritis Rheum. 2004 Jan;50(1):72–7.

53. Cerhan JR, Saag KG, Merlino LA, Mikuls TR, Criswell LA. Antioxidant micronutrients and risk of rheumatoid arthritis in a cohort of older women. Am J Epidemiol. 2003 Feb 15;157(4):345–54.

54. Benito-Garcia E, Feskanich D, Hu FB, Mandl LA, Karlson EW. Protein, iron, and meat consumption and risk for rheumatoid arthritis: a prospective cohort study. Arthritis Res Ther. 2007;9(1):R16.

55. Hu Y, Costenbader KH, Gao X, Al-Daabil M, Sparks JA, Solomon DH, et al. Sugar-sweetened soda consumption and risk of developing rheumatoid arthritis in women1234. Am J Clin Nutr. 2014 Sep;100(3):959–67.

56. Sundström B, Johansson I, Rantapää-Dahlqvist S. Interaction between dietary sodium and smoking increases the risk for rheumatoid arthritis: results from a nested case-control study. Rheumatology (Oxford). 2015 Mar;54(3):487–93.

57. Gan RW, Demoruelle MK, Deane KD, Weisman MH, Buckner JH, Gregersen PK, et al. Omega-3 fatty acids are associated with a lower prevalence of autoantibodies in shared epitope-positive subjects at risk for rheumatoid arthritis. Ann Rheum Dis. 2017 Jan;76(1):147–52.

58. Jin Z, Xiang C, Cai Q, Wei X, He J. Alcohol consumption as a preventive factor for developing rheumatoid arthritis: a dose-response meta-analysis of prospective studies. Ann Rheum Dis. 2014 Nov;73(11):1962–7.

59. Pikwer M, Bergström U, Nilsson JÅ, Jacobsson L, Turesson C. Early menopause is an independent predictor of rheumatoid arthritis. Ann Rheum Dis. 2012 Mar;71(3):378–81.

60. Merlino LA, Cerhan JR, Criswell LA, Mikuls TR, Saag KG. Estrogen and other female reproductive risk factors are not strongly associated with the development of rheumatoid arthritis in elderly women. Semin Arthritis Rheum. 2003 Oct;33(2):72–82.

61. Jørgensen KT, Pedersen BV, Jacobsen S, Biggar RJ, Frisch M. National cohort study of reproductive risk factors for rheumatoid arthritis in Denmark: a role for hyperemesis, gestational hypertension and pre-eclampsia? Ann Rheum Dis. 2010 Feb;69(2):358–63.

62. Adab P, Jiang CQ, Rankin E, Tsang YW, Lam TH, Barlow J, et al. Breastfeeding practice, oral contraceptive use and risk of rheumatoid arthritis among Chinese women: the Guangzhou Biobank Cohort Study. Rheumatology (Oxford). 2014 May;53(5):860–6.

63. Orellana C, Saevarsdottir S, Klareskog L, Karlson EW, Alfredsson L, Bengtsson C. Postmenopausal hormone therapy and the risk of rheumatoid arthritis: results from the Swedish EIRA population-based case-control study. Eur J Epidemiol. 2015 May;30(5):449–57.

64. Spector TD, Roman E, Silman AJ. The pill, parity, and rheumatoid arthritis. Arthritis Rheum. 1990 Jun;33(6):782–9.

65. Liao KP, Alfredsson L, Karlson EW. Environmental influences on risk for rheumatoid arthritis. Curr Opin Rheumatol. 2009 May;21(3):279–83.

66. Finckh A, Gilbert B, Hodkinson B, Bae SC, Thomas R, Deane KD, et al. Global epidemiology of rheumatoid arthritis. Nat Rev Rheumatol. 2022 Oct;18(10):591–602.

67. Safiri S, Kolahi AA, Hoy D, Smith E, Bettampadi D, Mansournia MA, et al. Global, regional and national burden of rheumatoid arthritis 1990-2017: a systematic analysis of the Global Burden of Disease study 2017. Ann Rheum Dis. 2019 Nov;78(11):1463–71.

68. Almutairi K, Nossent J, Preen D, Keen H, Inderjeeth C. The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. Rheumatol Int. 2021 May;41(5):863–77.

69. Kvien TK, Uhlig T, Ødegård S, Heiberg MS. Epidemiological aspects of rheumatoid arthritis: the sex ratio. Ann N Y Acad Sci. 2006 Jun;1069:212–22.

70. Sparks JA, Chang SC, Liao KP, Lu B, Fine AR, Solomon DH, et al. Rheumatoid Arthritis and Mortality Among Women During 36 Years of Prospective Follow-Up: Results From the Nurses' Health Study. Arthritis Care Res (Hoboken). 2016 Jun;68(6):753–62.

71. van den Hoek J, Boshuizen HC, Roorda LD, Tijhuis GJ, Nurmohamed MT, van den Bos GAM, et al. Mortality in patients with rheumatoid arthritis: a 15-year prospective cohort study. Rheumatol Int. 2017;37(4):487–93.

72. Zhang Y, Lu N, Peloquin C, Dubreuil M, Neogi T, Aviña-Zubieta JA, et al. Improved survival in rheumatoid arthritis: a general population-based cohort study. Ann Rheum Dis. 2017 Feb;76(2):408–13.

73. Lacaille D, Avina-Zubieta JA, Sayre EC, Abrahamowicz M. Improvement in 5-year mortality in incident rheumatoid arthritis compared with the general population-closing the mortality gap. Ann Rheum Dis. 2017 Jun;76(6):1057–63.

74. Dadoun S, Zeboulon-Ktorza N, Combescure C, Elhai M, Rozenberg S, Gossec L, et al. Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. Joint Bone Spine. 2013 Jan;80(1):29–33.

75. Nishimura K, Sugiyama D, Kogata Y, Tsuji G, Nakazawa T, Kawano S, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. Ann Intern Med.

2007 Jun 5;146(11):797–808.

76. Ingegnoli F, Castelli R, Gualtierotti R. Rheumatoid Factors: Clinical Applications. Dis Markers. 2013;35(6):727–34.

77. Ioan-Facsinay A, Willemze A, Robinson DB, Peschken CA, Markland J, van der Woude D, et al. Marked differences in fine specificity and isotype usage of the anti-citrullinated protein antibody in health and disease. Arthritis Rheum. 2008 Oct;58(10):3000–8.

78. Kurowska W, Kuca-Warnawin EH, Radzikowska A, Maśliński W. The role of anti-citrullinated protein antibodies (ACPA) in the pathogenesis of rheumatoid arthritis. Cent Eur J Immunol. 2017;42(4):390–8.

79. Liang Z, Wang N, Shang L, Wang Y, Feng M, Liu G, et al. Evaluation of the immune feature of ACPA-negative rheumatoid arthritis and the clinical value of matrix metalloproteinase-3. Front Immunol. 2022 Jul 27;13:939265.

80. Kolarz B, Ciesla M, Rosenthal AK, Dryglewska M, Majdan M. The value of anti-CarP and anti-PAD4 as markers of rheumatoid arthritis in ACPA/RF negative rheumatoid arthritis patients. Ther Adv Musculoskelet Dis. 2021;13:1759720X21989868.

81. Kgoebane K, Ally MMTM, Duim-Beytell MC, Suleman FE. The role of imaging in rheumatoid arthritis. SA Journal of Radiology [Internet]. 2018 [cited 2023 Aug 27];22(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6837821/

82. Kay J, Upchurch KS. ACR/EULAR 2010 rheumatoid arthritis classification criteria. Rheumatology (Oxford). 2012 Dec;51 Suppl 6:vi5-9.

83. Moore TL, Dorner RW. Rheumatoid factors. Clin Biochem. 1993 Apr;26(2):75–84.

84. van der Woude D, Alemayehu WG, Verduijn W, de Vries RRP, Houwing-Duistermaat JJ, Huizinga TWJ, et al. Gene-environment interaction influences the reactivity of autoantibodies to citrullinated antigens in rheumatoid arthritis. Nat Genet. 2010 Oct;42(10):814–6; author reply 816.

85. Luban S, Li ZG. Citrullinated peptide and its relevance to rheumatoid arthritis: an update. Int J Rheum Dis. 2010 Oct;13(4):284–7.

86. Jang S, Kwon EJ, Lee JJ. Rheumatoid Arthritis: Pathogenic Roles of Diverse Immune Cells. Int J Mol Sci. 2022 Jan 14;23(2):905.

87. Cutolo M, Campitiello R, Gotelli E, Soldano S. The Role of M1/M2 Macrophage Polarization in Rheumatoid Arthritis Synovitis. Front Immunol. 2022;13:867260.

88. Komatsu N, Takayanagi H. Mechanisms of joint destruction in rheumatoid arthritis - immune cell-fibroblast-bone interactions. Nat Rev Rheumatol. 2022 Jul;18(7):415–29.

89. Iwamoto N, Kawakami A. The monocyte-to-osteoclast transition in rheumatoid arthritis: Recent findings. Front Immunol. 2022;13:998554.

90. Physiology and Pathology of Autoimmune Diseases: Role of CD4+ T cells in Rheumatoid Arthritis | IntechOpen [Internet]. [cited 2023 Aug 27]. Available from: https://www.intechopen.com/chapters/56532

91. Tu J, Huang W, Zhang W, Mei J, Zhu C. A Tale of Two Immune Cells in Rheumatoid Arthritis: The Crosstalk Between Macrophages and T Cells in the Synovium. Front Immunol. 2021;12:655477.

92. Fox DA, Gizinski A, Morgan R, Lundy SK. Cell-cell Interactions in Rheumatoid Arthritis Synovium. Rheum Dis Clin North Am. 2010 May;36(2):311–23.

93. Ib M, G S. Nature reviews. Immunology. 2007 [cited 2021 Jan 14]. Cytokines in the pathogenesis of rheumatoid arthritis. Available from: https://pubmed.ncbi.nlm.nih.gov/17525752/

94. Firestein GS. Evolving concepts of rheumatoid arthritis. Nature. 2003 May 15;423(6937):356–61.

95. Kinne RW, Bräuer R, Stuhlmüller B, Palombo-Kinne E, Burmester GR. Macrophages in rheumatoid arthritis. Arthritis Res. 2000;2(3):189–202.

96. Huber LC, Distler O, Tarner I, Gay RE, Gay S, Pap T. Synovial fibroblasts: key players in rheumatoid arthritis. Rheumatology (Oxford). 2006 Jun;45(6):669–75.

97. Schulze-Koops H, Kalden JR. The balance of Th1/Th2 cytokines in rheumatoid arthritis. Best Pract Res Clin Rheumatol. 2001 Dec;15(5):677–91.

98. Walsh MC, Choi Y. Biology of the RANKL–RANK–OPG System in Immunity, Bone, and Beyond. Front Immunol. 2014 Oct 20;5:511.

99. Ospelt C. Synovial fibroblasts in 2017. RMD Open. 2017 Oct 15;3(2):e000471.

100. Jay GD, Britt DE, Cha CJ. Lubricin is a product of megakaryocyte stimulating factor gene expression by human synovial fibroblasts. J Rheumatol. 2000 Mar;27(3):594–600.

101. Firestein GS, Nguyen K, Aupperle KR, Yeo M, Boyle DL, Zvaifler NJ. Apoptosis in rheumatoid arthritis: p53 overexpression in rheumatoid arthritis synovium. Am J Pathol. 1996 Dec;149(6):2143–51.

102. Sabeh F, Fox D, Weiss SJ. Membrane-type I matrix metalloproteinase-dependent regulation of rheumatoid arthritis synoviocyte function. J Immunol. 2010 Jun 1;184(11):6396–406.

103. Bartok B, Firestein GS. Fibroblast-like synoviocytes: key effector cells in rheumatoid arthritis. Immunol Rev. 2010 Jan;233(1):233–55.

104. Alivernini S, MacDonald L, Elmesmari A, Finlay S, Tolusso B, Gigante MR, et al. Distinct synovial tissue macrophage subsets regulate inflammation and remission in rheumatoid arthritis. Nat Med. 2020 Aug;26(8):1295–306.

105. Knab K, Chambers D, Krönke G. Synovial Macrophage and Fibroblast Heterogeneity in Joint Homeostasis and Inflammation. Front Med (Lausanne). 2022;9:862161.

106. Fuentelsaz-Romero S, Cuervo A, Estrada-Capetillo L, Celis R, García-Campos R, Ramírez J, et al. GM-CSF Expression and Macrophage Polarization in Joints of Undifferentiated Arthritis Patients Evolving to Rheumatoid Arthritis or Psoriatic Arthritis. Front Immunol. 2020;11:613975.

107. Tran CN, Davis MJ, Tesmer LA, Endres JL, Motyl CD, Smuda C, et al. Presentation of arthritogenic peptide to antigen-specific T cells by fibroblast-like synoviocytes. Arthritis Rheum. 2007 May;56(5):1497–506.

108. Yamamura Y, Gupta R, Morita Y, He X, Pai R, Endres J, et al. Effector function of resting T cells: activation of synovial fibroblasts. J Immunol. 2001 Feb 15;166(4):2270–5.

109. Tu J, Huang W, Zhang W, Mei J, Zhu C. Two Main Cellular Components in Rheumatoid Arthritis: Communication Between T Cells and Fibroblast-Like Synoviocytes in the Joint Synovium. Front Immunol. 2022;13:922111.

110. Harada S, Yamamura M, Okamoto H, Morita Y, Kawashima M, Aita T, et al. Production of interleukin-7 and interleukin-15 by fibroblast-like synoviocytes from patients with rheumatoid arthritis. Arthritis Rheum. 1999 Jul;42(7):1508–16.

111. Bustamante MF, Garcia-Carbonell R, Whisenant KD, Guma M. Fibroblast-like synoviocyte metabolism in the pathogenesis of rheumatoid arthritis. Arthritis Res Ther. 2017 May 31;19(1):110.

112. Bottini N, Firestein GS. Duality of fibroblast-like synoviocytes in RA: passive responders and imprinted aggressors. Nat Rev Rheumatol. 2013 Jan;9(1):24–33.

113. Kurowska-Stolarska M, Alivernini S. Synovial tissue macrophages: friend or foe? RMD Open. 2017;3(2):e000527.

114. Tak PP, Smeets TJ, Daha MR, Kluin PM, Meijers KA, Brand R, et al. Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. Arthritis Rheum. 1997 Feb;40(2):217–25.

115. McInnes IB, Schett G. Cytokines in the pathogenesis of rheumatoid arthritis. Nat Rev Immunol. 2007 Jun;7(6):429–42.

116. Yang X, Chang Y, Wei W. Emerging role of targeting macrophages in rheumatoid arthritis: Focus on polarization, metabolism and apoptosis. Cell Prolif. 2020 Jul;53(7):e12854.

117. Ohradanova-Repic A, Machacek C, Charvet C, Lager F, Le Roux D, Platzer R, et al. Extracellular Purine Metabolism Is the Switchboard of Immunosuppressive Macrophages and a Novel Target to Treat Diseases With Macrophage Imbalances. Front Immunol. 2018;9:852.

118. Ahmed I, Ismail N. M1 and M2 Macrophages Polarization via mTORC1 Influences Innate Immunity and Outcome of Ehrlichia Infection. J Cell Immunol. 2020;2(3):108–15.

119. Denning TL, Wang Y chong, Patel SR, Williams IR, Pulendran B. Lamina propria macrophages and dendritic cells differentially induce regulatory and interleukin 17-producing T cell responses. Nat Immunol. 2007 Oct;8(10):1086–94.

120. Nanki T, Shimaoka T, Hayashida K, Taniguchi K, Yonehara S, Miyasaka N. Pathogenic role of the CXCL16-CXCR6 pathway in rheumatoid arthritis. Arthritis Rheum. 2005 Oct;52(10):3004–14.

121. Rückert R, Brandt K, Ernst M, Marienfeld K, Csernok E, Metzler C, et al. Interleukin-15 stimulates macrophages to activate CD4+ T cells: a role in the pathogenesis of rheumatoid arthritis? Immunology. 2009 Jan;126(1):63–73.

122. Roberts CA, Dickinson AK, Taams LS. The Interplay Between Monocytes/Macrophages and CD4+ T Cell Subsets in Rheumatoid Arthritis. Front Immunol. 2015 Nov 19;6:571.

123. Tosiek MJ, Fiette L, El Daker S, Eberl G, Freitas AA. IL-15-dependent balance between Foxp3 and RORγt expression impacts inflammatory bowel disease. Nat Commun. 2016 Mar 11;7:10888.

124. Fujimoto M, Nakano M, Terabe F, Kawahata H, Ohkawara T, Han Y, et al. The influence of excessive IL-6 production in vivo on the development and function of Foxp3+ regulatory T cells. J Immunol. 2011 Jan 1;186(1):32–40.

125. Kuo D, Ding J, Cohn IS, Zhang F, Wei K, Rao DA, et al. HBEGF+ macrophages in rheumatoid arthritis induce fibroblast invasiveness. Sci Transl Med. 2019 May 8;11(491):eaau8587.

126. Ikeuchi H, Kuroiwa T, Hiramatsu N, Kaneko Y, Hiromura K, Ueki K, et al. Expression of interleukin-22 in rheumatoid arthritis: potential role as a proinflammatory cytokine. Arthritis Rheum. 2005 Apr;52(4):1037–46.

127. Aoki H, Ohnishi H, Hama K, Shinozaki S, Kita H, Yamamoto H, et al. Existence of autocrine loop between interleukin-6 and transforming growth factor-beta1 in activated rat pancreatic stellate cells. J Cell Biochem. 2006 Sep 1;99(1):221–8.

128. Udalova IA, Mantovani A, Feldmann M. Macrophage heterogeneity in the context of rheumatoid arthritis. Nat Rev Rheumatol. 2016 Aug;12(8):472–85.

129. Klarenbeek PL, de Hair MJH, Doorenspleet ME, van Schaik BDC, Esveldt REE, van de Sande MGH, et al. Inflamed target tissue provides a specific niche for highly expanded T-cell clones in early human autoimmune disease. Ann Rheum Dis. 2012 Jun;71(6):1088–93.

130. van Loosdregt J, Rossetti M, Spreafico R, Moshref M, Olmer M, Williams GW, et al. Increased autophagy in CD4+ T cells of rheumatoid arthritis patients results in T cell hyperactivation and apoptosis resistance. Eur J Immunol. 2016 Dec;46(12):2862–70.

131. Luckheeram RV, Zhou R, Verma AD, Xia B. CD4+T Cells: Differentiation and Functions. Clin Dev Immunol. 2012;2012:925135.

132. Podojil JR, Miller SD. Molecular mechanisms of T-cell receptor and costimulatory molecule ligation/blockade in autoimmune disease therapy. Immunol Rev. 2009 May;229(1):337–55.

133. Chemin K, Gerstner C, Malmström V. Effector Functions of CD4+ T Cells at the Site of Local Autoimmune Inflammation-Lessons From Rheumatoid Arthritis. Front Immunol. 2019;10:353.

134. Luckheeram RV, Zhou R, Verma AD, Xia B. $CD4^+$T cells: differentiation and functions. Clin Dev Immunol. 2012;2012:925135.

135. Li S, Yin H, Zhang K, Wang T, Yang Y, Liu X, et al. Effector T helper cell populations are elevated in the bone marrow of rheumatoid arthritis patients and correlate with disease severity. Sci Rep. 2017 Jul 6;7(1):4776.

136. Yamada H, Nakashima Y, Okazaki K, Mawatari T, Fukushi JI, Kaibara N, et al. Th1 but not Th17 cells predominate in the joints of patients with rheumatoid arthritis. Ann Rheum Dis. 2008 Sep;67(9):1299–304.

137. Luo P, Wang P, Xu J, Hou W, Xu P, Xu K, et al. Immunomodulatory role of T helper cells in rheumatoid arthritis. Bone Joint Res. 2022 Jul 20;11(7):426–38.

138. Berner B, Akça D, Jung T, Muller GA, Reuss-Borst MA. Analysis of Th1 and Th2 cytokines expressing CD4+ and CD8+ T cells in rheumatoid arthritis by flow cytometry. J Rheumatol. 2000 May;27(5):1128–35.

139. Wehr P, Purvis H, Law SC, Thomas R. Dendritic cells, T cells and their interaction in rheumatoid arthritis. Clin Exp Immunol. 2019 Apr;196(1):12–27.

140. Li P, Schwarz EM, O'Keefe RJ, Ma L, Boyce BF, Xing L. RANK signaling is not required for TNFalpha-mediated increase in CD11(hi) osteoclast precursors but is essential for mature osteoclast formation in TNFalpha-mediated inflammatory arthritis. J Bone Miner Res. 2004 Feb;19(2):207–13.

141. Sallusto F, Lenig D, Mackay CR, Lanzavecchia A. Flexible programs of chemokine receptor expression on human polarized T helper 1 and 2 lymphocytes. J Exp Med. 1998 Mar 16;187(6):875–83.

142. James EA, Rieck M, Pieper J, Gebe JA, Yue BB, Tatum M, et al. Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. Arthritis Rheumatol. 2014 Jul;66(7):1712–22.

143. Patel DD, Zachariah JP, Whichard LP. CXCR3 and CCR5 ligands in rheumatoid arthritis synovium. Clin Immunol. 2001 Jan;98(1):39–45.

144. Szabo SJ, Kim ST, Costa GL, Zhang X, Fathman CG, Glimcher LH. A novel transcription factor, T-bet, directs Th1 lineage commitment. Cell. 2000 Mar

17;100(6):655–69.

145. Yang QB, He YL, Zhang QB, Mi QS, Zhou JG. Downregulation of Transcription Factor T-Bet as a Protective Strategy in Monosodium Urate-Induced Gouty Inflammation. Front Immunol. 2019;10:1199.

146. Zimmermann J, Kühl AA, Weber M, Grün JR, Löffler J, Haftmann C, et al. T-bet expression by Th cells promotes type 1 inflammation but is dispensable for colitis. Mucosal Immunol. 2016 Nov;9(6):1487–99.

147. Wang J, Fathman JW, Lugo-Villarino G, Scimone L, von Andrian U, Dorfman DM, et al. Transcription factor T-bet regulates inflammatory arthritis through its function in dendritic cells. J Clin Invest. 2006 Feb;116(2):414–21.

148. Mosser DM, Edwards JP. Exploring the full spectrum of macrophage activation. Nat Rev Immunol. 2008 Dec;8(12):958–69.

149. Yamada H, Haraguchi A, Sakuraba K, Okazaki K, Fukushi JI, Mizu-Uchi H, et al. Th1 is the predominant helper T cell subset that produces GM-CSF in the joint of rheumatoid arthritis. RMD Open. 2017;3(1):e000487.

150. Sebbag M, Parry SL, Brennan FM, Feldmann M. Cytokine stimulation of T lymphocytes regulates their capacity to induce monocyte production of tumor necrosis factor-alpha, but not interleukin-10: possible relevance to pathophysiology of rheumatoid arthritis. Eur J Immunol. 1997 Mar;27(3):624–32.

151. Kotake S, Nanke Y, Mogi M, Kawamoto M, Furuya T, Yago T, et al. IFN-gamma-producing human T cells directly induce osteoclastogenesis from human monocytes via the expression of RANKL. Eur J Immunol. 2005 Nov;35(11):3353–63.

152. Petrasca A, Phelan JJ, Ansboro S, Veale DJ, Fearon U, Fletcher JM. Targeting bioenergetics prevents CD4 T cell-mediated activation of synovial fibroblasts in rheumatoid arthritis. Rheumatology (Oxford). 2020 Oct 1;59(10):2816–28.

153. Mori M, Hashimoto M, Matsuo T, Fujii T, Furu M, Ito H, et al. Cell-contact-dependent activation of CD4+ T cells by adhesion molecules on synovial fibroblasts. Mod Rheumatol. 2017 May;27(3):448–56.

154. NSAIDs (nonsteroidal anti-inflammatory drugs) [Internet]. [cited 2023 Aug 27]. Available from: https://rheumatology.org/patients/nsaids-nonsteroidal-anti-inflammatory-drugs

155. Hua C, Buttgereit F, Combe B. Glucocorticoids in rheumatoid arthritis: current status and future studies. RMD Open. 2020 Jan;6(1):e000536.

156. Benjamin O, Goyal A, Lappin SL. Disease-Modifying Antirheumatic Drugs (DMARD). In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 [cited 2023 Aug 27]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK507863/

157. Aletaha D, Smolen JS. Diagnosis and Management of Rheumatoid Arthritis: A Review. JAMA. 2018 Oct 2;320(13):1360–72.

158. Qiu Q, Huang J, Lin Y, Shu X, Fan H, Tu Z, et al. Polymorphisms and pharmacogenomics for the toxicity of methotrexate monotherapy in patients with rheumatoid arthritis: A systematic review and meta-analysis. Medicine (Baltimore). 2017 Mar;96(11):e6337.

159. Johnson KJ, Sanchez HN, Schoenbrunner N. Defining response to TNF-inhibitors in rheumatoid arthritis: the negative impact of anti-TNF cycling and the need for a personalized medicine approach to identify primary non-responders. Clin Rheumatol. 2019 Nov;38(11):2967–76.

160. Rubbert-Roth A, Szabó MZ, Kedves M, Nagy G, Atzeni F, Sarzi-Puttini P. Failure of anti-TNF treatment in patients with rheumatoid arthritis: The pros and cons of the early use of alternative biological agents. Autoimmun Rev. 2019 Dec;18(12):102398.

161. Huang J, Fu X, Chen X, Li Z, Huang Y, Liang C. Promising Therapeutic Targets for Treatment of Rheumatoid Arthritis. Front Immunol. 2021;12:686155.

162. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004 Feb;5(2):101–13.

163. Kitano H. Computational systems biology. Nature. 2002 Nov;420(6912):206–10.

164. Paul D, Komarova NL. Multi-scale network targeting: A holistic systems-biology approach to cancer treatment. Prog Biophys Mol Biol. 2021 Oct;165:72–9.

165. Diaz LPM, Stumpf MPH. Gaining confidence in inferred networks. Sci Rep. 2022 Feb 14;12(1):2394.

166. Fujita KA, Ostaszewski M, Matsuoka Y, Ghosh S, Glaab E, Trefois C, et al. Integrating Pathways of Parkinson's Disease in a Molecular Interaction Map.

Mol Neurobiol. 2014 Feb 1;49(1):88–102.

167. Mizuno S, Iijima R, Ogishima S, Kikuchi M, Matsuoka Y, Ghosh S, et al. AlzPathway: a comprehensive map of signaling pathways of Alzheimer's disease. BMC Syst Biol. 2012 May 30;6:52.

168. Wu G, Zhu L, Dent JE, Nardini C. A Comprehensive Molecular Interaction Map for Rheumatoid Arthritis. PLoS One. 2010 Apr 16;5(4):e10137.

169. Kuperstein I, Bonnet E, Nguyen HA, Cohen D, Viara E, Grieco L, et al. Atlas of Cancer Signalling Network: a systems biology resource for integrative analysis of cancer data with Google Maps. Oncogenesis. 2015 Jul 20;4(7):e160.

170. Mazein A, Knowles RG, Adcock I, Chung KF, Wheelock CE, Maitland-van der Zee AH, et al. AsthmaMap: An expert-driven computational representation of disease mechanisms. Clin Exp Allergy. 2018 Aug;48(8):916–8.

171. Ostaszewski M, Niarakis A, Mazein A, Kuperstein I, Phair R, Orta-Resendiz A, et al. COVID19 Disease Map, a computational knowledge repository of virus-host interaction mechanisms. Mol Syst Biol. 2021 Oct;17(10):e10387.

172. Ostaszewski M, Gebel S, Kuperstein I, Mazein A, Zinovyev A, Dogrusoz U, et al. Community-driven roadmap for integrated disease maps. Brief Bioinform. 2019 Mar 25;20(2):659–70.

173. Moodie S, Le Novère N, Demir E, Mi H, Villéger A. Systems Biology Graphical Notation: Process Description language Level 1 Version 1.3. J Integr Bioinform. 2015 Sep 4;12(2):263.

174. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021 Feb 17;12(1):1088.

175. Rao VS, Srinivas K, Sujini GN, Kumar GNS. Protein-protein interaction detection: methods and analysis. Int J Proteomics. 2014;2014:147648.

176. Jakobsson JET, Spjuth O, Lagerström MC. scConnect: a method for exploratory analysis of cell-cell communication based on single-cell RNA-sequencing data. Bioinformatics. 2021 Oct 25;37(20):3501–8.

177. Zhang Y, Liu T, Wang J, Zou B, Li L, Yao L, et al. Cellinker: a platform of ligand-receptor interactions for intercellular communication analysis. Bioinformatics. 2021 Jan 20;btab036.

178. Noël F, Massenet-Regad L, Carmi-Levy I, Cappuccio A, Grandclaudon M, Trichot C, et al. Dissection of intercellular communication using the transcriptome-based framework ICELLNET. Nat Commun. 2021 Feb 17;12(1):1089.

179. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc. 2020 Apr;15(4):1484–506.

180. Vahid MR, Kurlovs AH, Andreani T, Augé F, Olfati-Saber R, de Rinaldis E, et al. DiSiR: fast and robust method to identify ligand-receptor interactions at subunit level from single-cell RNA-sequencing data. NAR Genom Bioinform. 2023 Mar;5(1):lqad030.

181. Dimitrov D, Türei D, Garrido-Rodriguez M, Burmedi PL, Nagai JS, Boys C, et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. Nat Commun. 2022 Jun 9;13(1):3224.

182. Brodland GW. How computational models can help unlock biological systems. Seminars in Cell & Developmental Biology. 2015 Dec 1;47–48:62–73.

183. How computational models can help unlock biological systems - PubMed [Internet]. [cited 2023 Aug 28]. Available from: https://pubmed.ncbi.nlm.nih.gov/26165820/

184. Samaga R, Klamt S. Modeling approaches for qualitative and semi-quantitative analysis of cellular signaling networks. Cell Communication and Signaling. 2013 Jun 26;11(1):43.

185. A multidimensional ODE-based model of Alzheimer's disease progression | Scientific Reports [Internet]. [cited 2023 Oct 13]. Available from: https://www.nature.com/articles/s41598-023-29383-5

186. Komarova NL, Wodarz D. ODE models for oncolytic virus dynamics. J Theor Biol. 2010 Apr 21;263(4):530–43.

187. Sharpe S, Dobrovolny HM. Predicting the effectiveness of chemotherapy using stochastic ODE models of tumor growth. Communications in Nonlinear Science and Numerical Simulation. 2021 Oct 1;101:105883.

188. Koziol JA, Falls TJ, Schnitzer JE. Different ODE models of tumor growth can deliver similar results. BMC Cancer. 2020 Mar 17;20(1):226.

189. Saqlain M, Alam M, Rönnegård L, Westin J. Investigating Stochastic Differential Equations Modelling for Levodopa Infusion in Patients with Parkinson's Disease. Eur J Drug Metab Pharmacokinet. 2020 Feb;45(1):41–9.

190. Jit M, Henderson B, Stevens M, Seymour RM. TNF-alpha neutralization in cytokine-driven diseases: a mathematical model to account for therapeutic success in rheumatoid arthritis but therapeutic failure in systemic inflammatory response syndrome. Rheumatology (Oxford). 2005 Mar;44(3):323–31.

191. Baker M, Denman-Johnson S, Brook BS, Gaywood I, Owen MR. Mathematical modelling of cytokine-mediated inflammation in rheumatoid arthritis. Math Med Biol. 2013 Dec;30(4):311–37.

192. Odisharia K, Odisharia V, Tsereteli P, Janikashvili N. On the Mathematical Model of Drug Treatment of Rheumatoid Arthritis. In: Jaiani G, Natroshvili D, editors. Mathematics, Informatics, and Their Applications in Natural Sciences and Engineering. Cham: Springer International Publishing; 2019. p. 161–8. (Springer Proceedings in Mathematics & Statistics).

193. Rao R, DuBois D, Almon R, Jusko WJ, Androulakis IP. Mathematical modeling of the circadian dynamics of the neuroendocrine-immune network in experimentally induced arthritis. Am J Physiol Endocrinol Metab. 2016 Aug 1;311(2):E310-324.

194. Moise N, Friedman A. Rheumatoid arthritis - a mathematical model. J Theor Biol. 2019 Jan 14;461:17–33.

195. Scholz S, Mittendorf T. Modeling rheumatoid arthritis using different techniques - a review of model construction and results. Health Econ Rev. 2014 Dec;4(1):18.

196. Roberts-Thomson PJ, Jones ME, Walker JG, Macfarlane JG, Smith MD, Ahern MJ. Stochastic processes in the causation of rheumatic disease. J Rheumatol. 2002 Dec;29(12):2628–34.

197. Joint symmetry in early and late rheumatoid and psoriatic arthritis: comparison with a mathematical model - PubMed [Internet]. [cited 2023 Oct 13]. Available from: https://pubmed.ncbi.nlm.nih.gov/10765932/

198. Agent-Based Modeling and Simulation in Mathematics and Biology Education | SpringerLink [Internet]. [cited 2023 Oct 13]. Available from: https://link.springer.com/article/10.1007/s11538-020-00778-z

199. Bailey AM, Thorne BC, Peirce SM. Multi-cell agent-based simulation of the microvasculature to study the dynamics of circulating inflammatory cell trafficking. Ann Biomed Eng. 2007 Jun;35(6):916–36.

200. Examining the controllability of sepsis using genetic algorithms on an agent-based model of systemic inflammation | PLOS Computational Biology [Internet]. [cited 2023 Oct 13]. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005876

201. Powathil GG, Swat M, Chaplain MAJ. Systems oncology: towards patient-specific treatment regimes informed by multiscale mathematical modelling. Semin Cancer Biol. 2015 Feb;30:13–20.

202. Srikrishnan V, Keller K. Small increases in agent-based model complexity can result in large increases in required calibration data. Environmental Modelling & Software. 2021 Apr 1;138:104978.

203. Llacay B, Peffer G. Categorical surrogation of agent-based models: A comparative study of machine learning classifiers. Expert Systems. n/a(n/a):e13342.

204. Saez-Rodriguez J, Simeoni L, Lindquist JA, Hemenway R, Bommhardt U, Arndt B, et al. A logical model provides insights into T cell receptor signaling. PLoS Comput Biol. 2007 Aug;3(8):e163.

205. Schlatter R, Philippi N, Wangorsch G, Pick R, Sawodny O, Borner C, et al. Integration of Boolean models exemplified on hepatocyte signal transduction. Brief Bioinform. 2012 May;13(3):365–76.

206. Samaga R, Saez-Rodriguez J, Alexopoulos LG, Sorger PK, Klamt S. The logic of EGFR/ErbB signaling: theoretical properties and analysis of high-throughput data. PLoS Comput Biol. 2009 Aug;5(8):e1000438.

207. Hardy S, Robillard PN. Modeling and simulation of molecular biology systems using petri nets: modeling goals of various approaches. J Bioinform Comput Biol. 2004 Dec;2(4):595–613.

208. Klamt S, Saez-Rodriguez J, Lindquist JA, Simeoni L, Gilles ED. A methodology for the structural and functional analysis of signaling and regulatory networks. BMC Bioinformatics. 2006 Feb 7;7:56.

209. Wang RS, Saadatpour A, Albert R. Boolean modeling in systems biology: an overview of methodology and applications. Phys Biol. 2012 Oct;9(5):055001.

210. Morris MK, Saez-Rodriguez J, Sorger PK, Lauffenburger DA. Logic-based models for the analysis of cell signaling networks. Biochemistry. 2010 Apr 20;49(15):3216–24.

211. Reddy VN, Mavrovouniotis ML, Liebman MN. Petri net representations in metabolic pathways. Proc Int Conf Intell Syst Mol Biol. 1993;1:328–36.

212. Fisher J, Henzinger TA. Executable cell biology. Nat Biotechnol. 2007 Nov;25(11):1239–49.

213. Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. J Theor Biol. 1969 Mar;22(3):437–67.

214. Kauffman SA, Kauffman SA. The Origins of Order: Self-Organization and Selection in Evolution. Oxford, New York: Oxford University Press; 1993. 728 p.

215. Thomas R. Boolean formalization of genetic control circuits. J Theor Biol. 1973 Dec;42(3):563–85.

216. Howell R, Clarke MA, Reuschl AK, Chen T, Abbott-Imboden S, Singer M, et al. Executable network of SARS-CoV-2-host interaction predicts drug combination treatments. NPJ Digit Med. 2022 Feb 14;5(1):18.

217. Béal J, Pantolini L, Noël V, Barillot E, Calzone L. Personalized logical models to investigate cancer response to BRAF treatments in melanomas and colorectal cancers. PLOS Computational Biology. 2021 Jan 28;17(1):e1007900.

218. Drug Target Optimization in Chronic Myeloid Leukemia Using Innovative Computational Platform | Scientific Reports [Internet]. [cited 2023 Oct 13]. Available from: https://www.nature.com/articles/srep08190

219. Dutta P, Ma L, Ali Y, Sloot PMA, Zheng J. Boolean network modeling of β-cell apoptosis and insulin resistance in type 2 diabetes mellitus. BMC Syst Biol. 2019 Apr 5;13(Suppl 2):36.

220. Singh V, Naldi A, Soliman S, Niarakis A. A large-scale Boolean model of the rheumatoid arthritis fibroblast-like synoviocytes predicts drug synergies in the arthritic joint. npj Syst Biol Appl. 2023 Jul 15;9(1):1–13.

221. Miagoux Q, Singh V, de Mézquita D, Chaudru V, Elati M, Petit-Teixeira E, et al. Inference of an Integrative, Executable Network for Rheumatoid Arthritis Combining Data-Driven Machine Learning Approaches and a State-of-the-

Art Mechanistic Disease Map. J Pers Med. 2021 Aug 12;11(8):785.

222. Wynn ML, Consul N, Merajver SD, Schnell S. Logic-based models in systems biology: a predictive and parameter-free network analysis method. Integr Biol (Camb). 2012 Nov;4(11):10.1039/c2ib20193c.

223. Bordon J, Moskon M, Zimic N, Mraz M. Fuzzy Logic as a Computational Tool for Quantitative Modelling of Biological Systems with Uncertain Kinetic Data. IEEE/ACM Trans Comput Biol Bioinform. 2015;12(5):1199–205.

224. Hemedan AA. Boolean modelling as a logic-based dynamic approach in systems medicine. Computational and Structural Biotechnology Journal. 2022;12.

225. Hemedan AA, Schneider R, Ostaszewski M. Applications of Boolean modeling to study the dynamics of a complex disease and therapeutics responses. Frontiers in Bioinformatics [Internet]. 2023 [cited 2023 Oct 13];3. Available from: https://www.frontiersin.org/articles/10.3389/fbinf.2023.1189723

226. Li F, Long T, Lu Y, Ouyang Q, Tang C. The yeast cell-cycle network is robustly designed. Proc Natl Acad Sci U S A. 2004 Apr 6;101(14):4781–6.

227. Schwab JD, Kühlwein SD, Ikonomi N, Kühl M, Kestler HA. Concepts in Boolean network modeling: What do they all mean? Comput Struct Biotechnol J. 2020;18:571–82.

228. Garg A, Di Cara A, Xenarios I, Mendoza L, De Micheli G. Synchronous versus asynchronous modeling of gene regulatory networks. Bioinformatics. 2008 Sep 1;24(17):1917–25.

229. Hinkelmann F, Brandon M, Guang B, McNeill R, Blekherman G, Veliz-Cuba A, et al. ADAM: Analysis of Discrete Models of Biological Systems Using Computer Algebra. BMC Bioinformatics. 2011 Jul 20;12(1):295.

230. Müssel C, Hopfensitz M, Kestler HA. BoolNet--an R package for generation, reconstruction and analysis of Boolean networks. Bioinformatics. 2010 May 15;26(10):1378–80.

231. Albert I, Thakar J, Li S, Zhang R, Albert R. Boolean network simulations for life scientists. Source Code for Biology and Medicine. 2008 Nov 14;3(1):16.

232. Helikar T, Kowal B, Rogers J. A Cell Simulator Platform: The Cell Collective. Clin Pharmacol Ther. 2013 May;93(5):393–5.

233. Klamt S, Saez-Rodriguez J, Gilles ED. Structural and functional analysis of cellular networks with CellNetAnalyzer. BMC Systems Biology. 2007 Jan 8;1(1):2.

234. Naldi A, Berenguier D, Fauré A, Lopez F, Thieffry D, Chaouiya C. Logical modelling of regulatory networks with GINsim 2.3. Biosystems. 2009 Aug 1;97(2):134–9.

235. Hall BA, Fisher J. Constructing and Analyzing Computational Models of Cell Signaling with BioModelAnalyzer. Curr Protoc Bioinformatics. 2020 Mar;69(1):e95.

236. Kelemen A, Abraham A, Chen Y, editors. Computational Intelligence in Bioinformatics [Internet]. Berlin, Heidelberg: Springer; 2008 [cited 2023 Aug 28]. (Kacprzyk J, editor. Studies in Computational Intelligence; vol. 94). Available from: http://link.springer.com/10.1007/978-3-540-76803-6

237. Hopfensitz M, Müssel C, Maucher M, Kestler HA. Attractors in Boolean networks: a tutorial. Comput Stat. 2013 Feb 1;28(1):19–36.

238. Aghamiri SS, Singh V, Naldi A, Helikar T, Soliman S, Niarakis A. Automated inference of Boolean models from molecular interaction maps using CaSQ. Bioinformatics. 2020 Aug 15;36(16):4473–82.

239. Naldi A. BioLQM: A Java Toolkit for the Manipulation and Conversion of Logical Qualitative Models of Biological Networks. Front Physiol. 2018 Nov 19;9:1605.

240. Cimatti A, Clarke E, Giunchiglia E, Giunchiglia F, Pistore M, Roveri M, et al. NuSMV 2: An OpenSource Tool for Symbolic Model Checking. In: Brinksma E, Larsen KG, editors. Computer Aided Verification. Berlin, Heidelberg: Springer; 2002. p. 359–64. (Lecture Notes in Computer Science).

241. Monteiro PT, Chaouiya C. Efficient Verification for Logical Models of Regulatory Networks. In: Rocha MP, Luscombe N, Fdez-Riverola F, Rodríguez JMC, editors. 6th International Conference on Practical Applications of Computational Biology & Bioinformatics. Berlin, Heidelberg: Springer; 2012. p. 259–67. (Advances in Intelligent and Soft Computing).

242. Bernot G, Comet JP, Richard A, Guespin J. Application of formal methods to biological regulatory networks: extending Thomas' asynchronous logical approach with temporal logic. J Theor Biol. 2004 Aug 7;229(3):339–47.

243. Benque D, Bourton S, Cockerton C, Cook B, Ishtiaq S, Piterman N, et al. Bma:

Visual Tool for Modeling and Analyzing Biological Networks. :6.

244. Veliz-Cuba A. Reduction of Boolean network models. J Theor Biol. 2011 Nov 21;289:167–72.

245. Mizera A, Pang J, Qu H, Yuan Q. Taming Asynchrony for Attractor Detection in Large Boolean Networks. IEEE/ACM Trans Comput Biol Bioinform. 2019;16(1):31–42.

246. The Gene Expression Omnibus Database - PubMed [Internet]. [cited 2023 Aug 28]. Available from: https://pubmed.ncbi.nlm.nih.gov/27008011/

247. Kang K, Park SH, Chen J, Qiao Y, Giannopoulou E, Berg K, et al. Interferon-γ Represses M2 Gene Expression in Human Macrophages by Disassembling Enhancers Bound by the Transcription Factor MAF. Immunity. 2017 Aug 15;47(2):235-250.e4.

248. preprocessCore: A collection of pre-processing functions version 1.52.1 from Bioconductor [Internet]. [cited 2022 Nov 29]. Available from: https://rdrr.io/bioc/preprocessCore/

249. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics. 2005 Aug 15;21(16):3439–40.

250. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47.

251. Carvalho K, Rebboah E, Jansen C, Williams K, Dowey A, McGill C, et al. Uncovering the Gene Regulatory Networks Underlying Macrophage Polarization Through Comparative Analysis of Bulk and Single-Cell Data [Internet]. bioRxiv; 2021 [cited 2023 Aug 28]. p. 2021.01.20.427499. Available from: https://www.biorxiv.org/content/10.1101/2021.01.20.427499v1

252. Le T, Phan T, Pham M, Tran D, Lam L, Nguyen T, et al. BBrowser: Making single-cell data easily accessible [Internet]. bioRxiv; 2020 [cited 2023 Aug 28]. p. 2020.12.11.414136. Available from: https://www.biorxiv.org/content/10.1101/2020.12.11.414136v1

253. Vuong H, Truong T, Phan T, Pham S. Venice: A New Algorithm for Finding Marker Genes in Single-Cell Transcriptomic Data [Internet]. bioRxiv; 2020 [cited 2023 Aug 28]. p. 2020.11.16.384479. Available from: https://www.biorxiv.org/content/10.1101/2020.11.16.384479v1

254. Mizoguchi F, Slowikowski K, Wei K, Marshall JL, Rao DA, Chang SK, et al. Functionally distinct disease-associated fibroblast subsets in rheumatoid arthritis. Nat Commun. 2018 Feb 23;9(1):789.

255. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. Sci Data. 2018 Feb 27;5(1):180015.

256. Zhang F, Wei K, Slowikowski K, Fonseka CY, Rao DA, Kelly S, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat Immunol. 2019 Jul;20(7):928–42.

257. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021 Jun 24;184(13):3573-3587.e29.

258. Monaco G, Lee B, Xu W, Mustafah S, Hwang YY, Carré C, et al. RNA-Seq Signatures Normalized by mRNA Abundance Allow Absolute Deconvolution of Human Immune Cell Types. Cell Rep. 2019 Feb 5;26(6):1627-1640.e7.

259. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005 Oct 25;102(43):15545–50.

260. Zhang W, Ferguson J, Ng SM, Hui K, Goh G, Lin A, et al. Effector CD4+ T cell expression signatures and immune-mediated disease associated genes. PLoS One. 2012;7(6):e38510.

261. Shakya K, Ruskin HJ, Kerr G, Crane M, Becker J. Comparison of microarray preprocessing methods. Adv Exp Med Biol. 2010;680:139–47.

262. Höllbacher B, Duhen T, Motley S, Klicznik MM, Gratz IK, Campbell DJ. Transcriptomic Profiling of Human Effector and Regulatory T Cell Subsets Identifies Predictive Population Signatures. Immunohorizons. 2020 Oct 9;4(10):585–96.

263. Parkinson H, Kapushesky M, Shojatalab M, Abeygunawardena N, Coulson R, Farne A, et al. ArrayExpress—a public database of microarray experiments and gene expression profiles. Nucleic Acids Res. 2007 Jan;35(Database issue):D747–50.

264. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto

Encyclopedia of Genes and Genomes. Nucleic Acids Res. 1999 Jan 1;27(1):29–34.

265. Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. Nucleic Acids Res. 2022 Jan 7;50(D1):D687–92.

266. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. Nucleic Acids Res. 2021 Jan 8;49(D1):D613–21.

267. Mi H, Thomas P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. Methods Mol Biol. 2009;563:123–40.

268. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014 Feb 15;30(4):523–30.

269. Biopharma Intelligence & Insights Services - Clarivate [Internet]. [cited 2023 Aug 28]. Available from: https://clarivate.com/products/biopharma/

270. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The Human Cell Atlas. Elife. 2017 Dec 5;6:e27041.

271. International HapMap Consortium. The International HapMap Project. Nature. 2003 Dec 18;426(6968):789–96.

272. Funahashi A, Morohashi M, Matsuoka Y, Jouraku A, Kitano H. CellDesigner: A Graphical Biological Network Editor and Workbench Interfacing Simulator. In: Choi S, editor. Introduction to Systems Biology [Internet]. Totowa, NJ: Humana Press; 2007 [cited 2023 Aug 28]. p. 422–34. Available from: http://link.springer.com/10.1007/978-1-59745-531-2_21

273. Gawron P, Ostaszewski M, Satagopam V, Gebel S, Mazein A, Kuzma M, et al. MINERVA—a platform for visualization and curation of molecular interaction networks. npj Syst Biol Appl. 2016 Sep 22;2(1):1–6.

274. Kitano H, Funahashi A, Matsuoka Y, Oda K. Using process diagrams for the graphical representation of biological networks. Nat Biotechnol. 2005 Aug;23(8):961–6.

275. Tweedie S, Braschi B, Gray K, Jones TEM, Seal RL, Yates B, et al. Genenames.org: the HGNC and VGNC resources in 2021. Nucleic Acids Res. 2021 Jan 8;49(D1):D939–46.

276. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics. 2003 Mar 1;19(4):524–31.

277. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015 Dec 23;1(6):417–25.

278. Oda K, Kimura T, Matsuoka Y, Funahashi A, Muramatsu M, Kitano H. Molecular Interaction Map of a Macrophage. AfCS Research Reports. 2004 Jan 1;2.

279. Wentker P, Eberhardt M, Dreyer FS, Bertrams W, Cantone M, Griss K, et al. An Interactive Macrophage Signal Transduction Map Facilitates Comparative Analyses of High-Throughput Data. J Immunol. 2017 Mar 1;198(5):2191–201.

280. 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 Nov 1;35(21):4496–8.

281. Patra BG, Roberts K, Wu H. A content-based dataset recommendation system for researchers-a case study on Gene Expression Omnibus (GEO) repository. Database (Oxford). 2020 Jan 1;2020:1.

282. Singh V, Kalliolias GD, Ostaszewski M, Veyssiere M, Pilalis E, Gawron P, et al. RA-map: building a state-of-the-art interactive knowledge base for rheumatoid arthritis. Database (Oxford). 2020 Jan 1;2020:baaa017.

283. Serhan CN, Gupta SK, Perretti M, Godson C, Brennan E, Li Y, et al. The Atlas of Inflammation Resolution (AIR). Mol Aspects Med. 2020 Aug;74:100894.

284. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002 Jan 1;30(1):207–10.

285. Zerrouk N, Aghakhani S, Singh V, Augé F, Niarakis A. A Mechanistic Cellular Atlas of the Rheumatic Joint. 2022;2:14.

286. Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res. 2008 Jan;36(Database issue):D901-906.

287. Gaulton A, Hersey A, Nowotka M, Bento AP, Chambers J, Mendez D, et al.

The ChEMBL database in 2017. Nucleic Acids Res. 2017 Jan 4;45(D1):D945–54.

288. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar 15;3:160018.

289. About Causaly [Internet]. [cited 2023 Aug 28]. Available from: https://www.causaly.com/about-causaly

290. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007 Jan;8(1):118–27.

291. Risso D, Perraudeau F, Gribkova S, Dudoit S, Vert JP. A general and flexible method for signal extraction from single-cell RNA-seq data. Nat Commun. 2018 Jan 18;9(1):284.

292. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012 Mar 15;28(6):882–3.

293. Sciacca E, Surace AEA, Alaimo S, Pulvirenti A, Rivellese F, Goldmann K, et al. Network analysis of synovial RNA sequencing identifies gene-gene interactions predictive of response in rheumatoid arthritis. Arthritis Research & Therapy. 2022 Jul 11;24(1):166.

294. Hall BA, Fisher J. Constructing and Analyzing Computational Models of Cell Signaling with BioModelAnalyzer. Current Protocols in Bioinformatics. 2020;69(1):e95.

295. Chaouiya C, Bérenguier D, Keating SM, Naldi A, van Iersel MP, Rodriguez N, et al. SBML qualitative models: a model representation format and infrastructure to foster interactions between qualitative modelling formalisms and tools. BMC Syst Biol. 2013 Dec 10;7:135.

296. Cook B, Fisher J, Krepska E, Piterman N. Proving Stabilization of Biological Systems. :16.

297. Joblib: running Python functions as pipeline jobs — joblib 1.3.0.dev0 documentation [Internet]. [cited 2022 Nov 29]. Available from: https://joblib.readthedocs.io/en/latest/

298. Lefèvre S, Knedla A, Tennie C, Kampmann A, Wunrau C, Dinser R, et al.

Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. Nature medicine. 2009 Dec;15(12):1414.

299. Edilova MI, Akram A, Abdul-Sater AA. Innate immunity drives pathogenesis of rheumatoid arthritis. Biomed J. 2021 Apr;44(2):172–82.

300. Boutet MA, Courties G, Nerviani A, Le Goff B, Apparailly F, Pitzalis C, et al. Novel insights into macrophage diversity in rheumatoid arthritis synovium. Autoimmun Rev. 2021 Mar;20(3):102758.

301. He W, Kapate N, Shields CW, Mitragotri S. Drug delivery to macrophages: A review of targeting drugs and drug carriers to macrophages for inflammatory diseases. Adv Drug Deliv Rev. 2020;165–166:15–40.

302. Hu G, Guo M, Xu J, Wu F, Fan J, Huang Q, et al. Nanoparticles Targeting Macrophages as Potential Clinical Therapeutic Agents Against Cancer and Inflammation. Front Immunol. 2019;10:1998.

303. Li J, Hsu HC, Mountz JD. Managing macrophages in rheumatoid arthritis by reform or removal. Curr Rheumatol Rep. 2012 Oct;14(5):445–54.

304. Zhou Y, Zhang Y, Lian X, Li F, Wang C, Zhu F, et al. Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents. Nucleic Acids Res. 2022 Jan 7;50(D1):D1398–407.

305. Williams M, Raddatz R. Receptors as drug targets. Curr Protoc Pharmacol. 2006 Apr;Chapter 1:Unit 1.1.

306. Schmid AS, Neri D. Advances in antibody engineering for rheumatic diseases. Nat Rev Rheumatol. 2019 Apr;15(4):197–207.

307. Henricks LM, Schellens JHM, Huitema ADR, Beijnen JH. The use of combinations of monoclonal antibodies in clinical oncology. Cancer Treat Rev. 2015 Dec;41(10):859–67.

308. Ma J, Mo Y, Tang M, Shen J, Qi Y, Zhao W, et al. Bispecific Antibodies: From Research to Clinical Application. Frontiers in Immunology [Internet]. 2021 [cited 2023 Aug 29];12. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2021.626616

309. Sedykh SE, Prinz VV, Buneva VN, Nevinsky GA. Bispecific antibodies: design, therapy, perspectives. Drug Des Devel Ther. 2018;12:195–208.

310. Zhou X, Huang D, Wang R, Wu M, Zhu L, Peng W, et al. Targeted therapy of

rheumatoid arthritis via macrophage repolarization. Drug Deliv. 2021 Dec;28(1):2447–59.

311. Handel ML, McMorrow LB, Gravallese EM. Nuclear factor-kappa B in rheumatoid synovium. Localization of p50 and p65. Arthritis Rheum. 1995 Dec;38(12):1762–70.

312. Prescott JA, Cook SJ. Targeting IKKβ in Cancer: Challenges and Opportunities for the Therapeutic Utilisation of IKKβ Inhibitors. Cells. 2018 Aug 23;7(9):115.

313. Herrington FD, Carmody RJ, Goodyear CS. Modulation of NF-κB Signaling as a Therapeutic Target in Autoimmunity. J Biomol Screen. 2016 Mar;21(3):223–42.

314. D'Acquisto F, May MJ, Ghosh S. Inhibition of nuclear factor kappa B (NF-B): an emerging theme in anti-inflammatory therapies. Mol Interv. 2002 Feb;2(1):22–35.

315. Park MH, Hong JT. Roles of NF-κB in Cancer and Inflammatory Diseases and Their Therapeutic Approaches. Cells. 2016 Mar 29;5(2):15.

316. Xia ZB, Meng FR, Fang YX, Wu X, Zhang CW, Liu Y, et al. Inhibition of NF-κB signaling pathway induces apoptosis and suppresses proliferation and angiogenesis of human fibroblast-like synovial cells in rheumatoid arthritis. Medicine (Baltimore). 2018 Jun;97(23):e10920.

317. Yin MJ, Yamamoto Y, Gaynor RB. The anti-inflammatory agents aspirin and salicylate inhibit the activity of I(kappa)B kinase-beta. Nature. 1998 Nov 5;396(6706):77–80.

318. Kopp E, Ghosh S. Inhibition of NF-kappa B by sodium salicylate and aspirin. Science. 1994 Aug 12;265(5174):956–9.

319. Takada Y, Bhardwaj A, Potdar P, Aggarwal BB. Nonsteroidal anti-inflammatory agents differ in their ability to suppress NF-kappaB activation, inhibition of expression of cyclooxygenase-2 and cyclin D1, and abrogation of tumor cell proliferation. Oncogene. 2004 Dec 9;23(57):9247–58.

320. Kavanaugh A, Wells AF. Benefits and risks of low-dose glucocorticoid treatment in the patient with rheumatoid arthritis. Rheumatology (Oxford). 2014 Oct;53(10):1742–51.

321. Crofford LJ. Use of NSAIDs in treating patients with arthritis. Arthritis Res Ther. 2013;15 Suppl 3(Suppl 3):S2.

322. D'Acquisto F, Ialenti A, Ianaro A, Di Vaio R, Carnuccio R. Local administration of transcription factor decoy oligonucleotides to nuclear factor-kappaB prevents carrageenin-induced inflammation in rat hind paw. Gene Ther. 2000 Oct;7(20):1731–7.

323. Neubert K, Meister S, Moser K, Weisel F, Maseda D, Amann K, et al. The proteasome inhibitor bortezomib depletes plasma cells and protects mice with lupus-like disease from nephritis. Nat Med. 2008 Jul;14(7):748–55.

324. McIntyre KW, Shuster DJ, Gillooly KM, Dambach DM, Pattoli MA, Lu P, et al. A highly selective inhibitor of I kappa B kinase, BMS-345541, blocks both joint inflammation and destruction in collagen-induced arthritis in mice. Arthritis Rheum. 2003 Sep;48(9):2652–9.

325. May MJ, D'Acquisto F, Madge LA, Glöckner J, Pober JS, Ghosh S. Selective inhibition of NF-kappaB activation by a peptide that blocks the interaction of NEMO with the IkappaB kinase complex. Science. 2000 Sep 1;289(5484):1550–4.

326. Simmonds RE, Foxwell BM. Signalling, inflammation and arthritis: NF-kappaB and its relevance to arthritis and inflammation. Rheumatology (Oxford). 2008 May;47(5):584–90.

327. Sehnert B, Burkhardt H, Dübel S, Voll RE. Cell-Type Targeted NF-kappaB Inhibition for the Treatment of Inflammatory Diseases. Cells. 2020 Jul 6;9(7):1627.

328. Galeazzi M, Bazzichi L, Sebastiani GD, Neri D, Garcia E, Ravenni N, et al. A phase IB clinical trial with Dekavil (F8-IL10), an immunoregulatory "armed antibody" for the treatment of rheumatoid arthritis, used in combination wilh methotrexate. Isr Med Assoc J. 2014 Oct;16(10):666.

329. Arioka M, Takahashi-Yanaga F. Glycogen synthase kinase-3 inhibitor as a multi-targeting anti-rheumatoid drug. Biochemical Pharmacology. 2019;165:207–13.

330. Patel S, Werstuck GH. Macrophage Function and the Role of GSK3. Int J Mol Sci. 2021 Feb 23;22(4):2206.

331. Kwon YJ, Yoon CH, Lee SW, Park YB, Lee SK, Park MC. Inhibition of glycogen synthase kinase-3β suppresses inflammatory responses in rheumatoid arthritis fibroblast-like synoviocytes and collagen-induced arthritis. Joint Bone Spine. 2014 May;81(3):240–6.

332. Peng LY, Li BB, Deng KB, Wang WG. MicroRNA-214-3p facilitates M2 macrophage polarization by targeting GSK3B. Kaohsiung J Med Sci. 2022 Apr;38(4):347–56.

333. Wen AY, Sakamoto KM, Miller LS. The role of the transcription factor CREB in immune function. J Immunol. 2010 Dec 1;185(11):6413–9.

334. Shabestari RM, Safa M, Alikarami F, Banan M, Kazemi A. CREB knockdown inhibits growth and induces apoptosis in human pre-B acute lymphoblastic leukemia cells through inhibition of prosurvival signals. Biomedicine & Pharmacotherapy. 2017;87:274–9.

335. Luan B, Yoon YS, Le Lay J, Kaestner KH, Hedrick S, Montminy M. CREB pathway links PGE2 signaling with macrophage polarization. Proc Natl Acad Sci U S A. 2015 Dec 22;112(51):15642–7.

336. Polumuri S, Perkins DJ, Vogel SN. cAMP levels regulate macrophage alternative activation marker expression. Innate Immun. 2021 Feb;27(2):133–42.

337. Krepler C, Xiao M, Samanta M, Vultur A, Chen HY, Brafford P, et al. Targeting Notch enhances the efficacy of ERK inhibitors in BRAF-V600E melanoma. Oncotarget. 2016 Nov 1;7(44):71211–22.

338. Porcelli L, Mazzotta A, Garofoli M, Di Fonte R, Guida G, Guida M, et al. Active notch protects MAPK activated melanoma cell lines from MEK inhibitor cobimetinib. Biomed Pharmacother. 2021 Jan;133:111006.

339. Choi BY, Choi Y, Park JS, Kang LJ, Baek SH, Park JS, et al. Inhibition of Notch1 induces population and suppressive activity of regulatory T cell in inflammatory arthritis. Theranostics. 2018;8(17):4795–804.

340. Ohori M. ERK inhibitors as a potential new therapy for rheumatoid arthritis. Drug News Perspect. 2008 Jun;21(5):245–50.

341. Keewan E, Naser SA. The Role of Notch Signaling in Macrophages during Inflammation and Infection: Implication in Rheumatoid Arthritis? Cells. 2020 Jan 2;9(1):111.

342. Magnol M, Rauwel B, Sayegh S, Diallo K, Baron M, Lacassagne E, et al. AB0040 JAK INHIBITORS – BARICITINIB AND TOFACITINIB – MODULATE THE IN VITRO INFLAMMATORY AND ALTERNATIVE POLARIZATIONS OF MACROPHAGES. Annals of the Rheumatic Diseases. 2019;78(Suppl 2):1486–7.

343. McInnes IB, Byers NL, Higgs RE, Lee J, Macias WL, Na S, et al. Comparison of baricitinib, upadacitinib, and tofacitinib mediated regulation of cytokine signaling in human leukocyte subpopulations. Arthritis Res Ther. 2019 Aug 2;21(1):183.

344. Palasiewicz K, Umar S, Romay B, Zomorrodi RK, Shahrara S. Tofacitinib therapy intercepts macrophage metabolic reprogramming instigated by SARS-CoV-2 Spike protein. Eur J Immunol. 2021 Sep;51(9):2330–40.

345. Ding Q, Hu W, Wang R, Yang Q, Zhu M, Li M, et al. Signaling pathways in rheumatoid arthritis: implications for targeted therapy. Sig Transduct Target Ther. 2023 Feb 17;8(1):1–24.

346. Xing R, Jin Y, Sun L, Yang L, Li C, Li Z, et al. Interleukin-21 induces migration and invasion of fibroblast-like synoviocytes from patients with rheumatoid arthritis. Clin Exp Immunol. 2016 May;184(2):147–58.

347. Xu X, Chen H, Zhang Q, Xu J, Shi Q, Wang M. MiR-650 inhibits proliferation, migration and invasion of rheumatoid arthritis synovial fibroblasts by targeting AKT2. Biomed Pharmacother. 2017 Apr;88:535–41.

348. Yu FY, Xie CQ, Jiang CL, Sun JT, Feng HC, Li C, et al. MiR-92a inhibits fibroblast-like synoviocyte proliferation and migration in rheumatoid arthritis by targeting AKT2. J Biosci. 2018 Dec;43(5):911–9.

349. Trzybulska D, Olewicz-Gawlik A, Sikora J, Frydrychowicz M, Kolecka-Bednarczyk A, Kaczmarek M, et al. The effect of caveolin-1 knockdown on interleukin-1β-induced chemokine (C-C motif) ligand 2 expression in synovial fluid-derived fibroblast-like synoviocytes from patients with rheumatoid arthritis. Adv Clin Exp Med. 2018 Nov;27(11):1491–7.

350. Li S, Jin Z, Lu X. MicroRNA-192 suppresses cell proliferation and induces apoptosis in human rheumatoid arthritis fibroblast-like synoviocytes by downregulating caveolin 1. Mol Cell Biochem. 2017 Aug;432(1–2):123–30.

351. Takeba Y, Suzuki N, Wakisaka S, Takeno M, Kaneko A, Asai T, et al. Involvement of cAMP responsive element binding protein (CREB) in the synovial cell hyperfunction in patients with rheumatoid arthritis. Clin Exp Rheumatol. 2000;18(1):47–55.

352. Vincenti MP, Coon CI, Brinckerhoff CE. Nuclear factor kappaB/p50 activates an element in the distal matrix metalloproteinase 1 promoter in interleukin-1beta-stimulated synovial fibroblasts. Arthritis Rheum. 1998

Nov;41(11):1987–94.

353. Bond M, Fabunmi RP, Baker AH, Newby AC. Synergistic upregulation of metalloproteinase-9 by growth factors and inflammatory cytokines: an absolute requirement for transcription factor NF-kappa B. FEBS Lett. 1998 Sep 11;435(1):29–34.

354. Bond M, Baker AH, Newby AC. Nuclear factor kappaB activity is essential for matrix metalloproteinase-1 and -3 upregulation in rabbit dermal fibroblasts. Biochem Biophys Res Commun. 1999 Oct 22;264(2):561–7.

355. Bondeson J, Brennan F, Foxwell B, Feldmann M. Effective adenoviral transfer of IkappaBalpha into human fibroblasts and chondrosarcoma cells reveals that the induction of matrix metalloproteinases and proinflammatory cytokines is nuclear factor-kappaB dependent. J Rheumatol. 2000 Sep;27(9):2078–89.

356. Nejatbakhsh Samimi L, Farhadi E, Tahmasebi MN, Jamshidi A, Sharafat Vaziri A, Mahmoudi M. NF-κB signaling in rheumatoid arthritis with focus on fibroblast-like synoviocytes. Autoimmunity Highlights. 2020 Aug 8;11(1):11.

357. Xue M, McKelvey K, Shen K, Minhas N, March L, Park SY, et al. Endogenous MMP-9 and not MMP-2 promotes rheumatoid synovial fibroblast survival, inflammation and cartilage degradation. Rheumatology (Oxford). 2014 Dec;53(12):2270–9.

358. Iotsova V, Caamaño J, Loy J, Yang Y, Lewin A, Bravo R. Osteopetrosis in mice lacking NF-kappaB1 and NF-kappaB2. Nat Med. 1997 Nov;3(11):1285–9.

359. Lin T h., Pajarinen J, Lu L, Nabeshima A, Cordova LA, Yao Z, et al. NF-κB as a Therapeutic Target in Inflammatory-Associated Bone Diseases. Adv Protein Chem Struct Biol. 2017;107:117–54.

360. Clohisy JC, Yamanaka Y, Faccio R, Abu-Amer Y. Inhibition of IKK activation, through sequestering NEMO, blocks PMMA-induced osteoclastogenesis and calvarial inflammatory osteolysis. J Orthop Res. 2006 Jul;24(7):1358–65.

361. Hilliard BA, Mason N, Xu L, Sun J, Lamhamedi-Cherradi SE, Liou HC, et al. Critical roles of c-Rel in autoimmune inflammation and helper T cell differentiation. J Clin Invest. 2002 Sep;110(6):843–50.

362. Aronica MA, Mora AL, Mitchell DB, Finn PW, Johnson JE, Sheller JR, et al. Preferential role for NF-kappa B/Rel signaling in the type 1 but not type 2 T cell-dependent immune response in vivo. J Immunol. 1999 Nov

1;163(9):5116–24.

363. Xue Y, Yang Y, Su Z, Barnie PA, Zheng D, Zhang Y, et al. Local delivery of T-bet shRNA reduces inflammation in collagen II-induced arthritis via downregulation of IFN-γ and IL-17. Molecular Medicine Reports. 2014 Mar 1;9(3):899–903.

364. Usui T, Preiss JC, Kanno Y, Yao ZJ, Bream JH, O'Shea JJ, et al. T-bet regulates Th1 responses through essential effects on GATA-3 function rather than on IFNG gene acetylation and transcription. J Exp Med. 2006 Mar 20;203(3):755–66.

365. Delgoffe GM, Kole TP, Zheng Y, Zarek PE, Matthews KL, Xiao B, et al. The mTOR kinase differentially regulates effector and regulatory T cell lineage commitment. Immunity. 2009 Jun 19;30(6):832–44.

366. Cejka D, Hayer S, Niederreiter B, Sieghart W, Fuereder T, Zwerina J, et al. Mammalian target of rapamycin signaling is crucial for joint destruction in experimental arthritis and is activated in osteoclasts from patients with rheumatoid arthritis. Arthritis Rheum. 2010 Aug;62(8):2294–302.

367. Bruyn G a. W, Tate G, Caeiro F, Maldonado-Cocco J, Westhovens R, Tannenbaum H, et al. Everolimus in patients with rheumatoid arthritis receiving concomitant methotrexate: a 3-month, double-blind, randomised, placebo-controlled, parallel-group, proof-of-concept study. Ann Rheum Dis. 2008 Aug;67(8):1090–5.

368. Harrington R, Al Nokhatha SA, Conway R. JAK Inhibitors in Rheumatoid Arthritis: An Evidence-Based Review on the Emerging Clinical Data. J Inflamm Res. 2020 Sep 14;13:519–31.

369. Dawson DW, Pearce SF, Zhong R, Silverstein RL, Frazier WA, Bouck NP. CD36 mediates the In vitro inhibitory effects of thrombospondin-1 on endothelial cells. J Cell Biol. 1997 Aug 11;138(3):707–17.

370. McMorrow JP, Crean D, Gogarty M, Smyth A, Connolly M, Cummins E, et al. Tumor necrosis factor inhibition modulates thrombospondin-1 expression in human inflammatory joint disease through altered NR4A2 activity. Am J Pathol. 2013 Oct;183(4):1243–57.

371. O'Brien CA. Control of RANKL gene expression. Bone. 2010 Apr;46(4):911–9.

372. Yokota K. Osteoclast differentiation in rheumatoid arthritis. Immunol Med. 2023 Jun 13;1–6.

373. O'Brien CA, Gubrij I, Lin SC, Saylors RL, Manolagas SC. STAT3 activation in stromal/osteoblastic cells is required for induction of the receptor activator of NF-kappaB ligand and stimulation of osteoclastogenesis by gp130-utilizing cytokines or interleukin-1 but not 1,25-dihydroxyvitamin D3 or parathyroid hormone. J Biol Chem. 1999 Jul 2;274(27):19301–8.

374. Brasier AR. The nuclear factor-κB–interleukin-6 signalling pathway mediating vascular inflammation. Cardiovasc Res. 2010 May 1;86(2):211–8.

375. Rajavashisth TB, Yamada H, Mishra NK. Transcriptional activation of the macrophage-colony stimulating factor gene by minimally modified LDL. Involvement of nuclear factor-kappa B. Arterioscler Thromb Vasc Biol. 1995 Oct;15(10):1591–8.

376. Liu T, Zhang L, Joo D, Sun SC. NF-κB signaling in inflammation. Sig Transduct Target Ther. 2017 Jul 14;2(1):1–9.

377. Pucino V, Certo M, Varricchi G, Marone G, Ursini F, Rossi FW, et al. Metabolic Checkpoints in Rheumatoid Arthritis. Front Physiol. 2020 Apr 17;11:347.

378. Zerrouk N, Miagoux Q, Dispot A, Elati M, Niarakis A. Identification of putative master regulators in rheumatoid arthritis synovial fibroblasts using gene expression data and network inference. Scientific Reports. 2020 Oct 1;10(1):16236.

379. Balasundaram A, Udhaya Kumar S, George Priya Doss C. A computational model revealing the immune-related hub genes and key pathways involved in rheumatoid arthritis (RA). Adv Protein Chem Struct Biol. 2022;129:247–73.

380. Moise N, Friedman A. Rheumatoid arthritis - a mathematical model. Journal of Theoretical Biology. 2019 Jan 14;461:17–33.

381. Friedman A, Lam KY. Analysis of a mathematical model of rheumatoid arthritis. J Math Biol. 2020 May 1;80(6):1857–83.

382. Aghakhani S, Soliman S, Niarakis A. Metabolic reprogramming in Rheumatoid Arthritis Synovial Fibroblasts: A hybrid modeling approach. PLOS COMPUTATIONAL BIOLOGY.

383. Guo Y, Xiao P, Lei S, Deng F, Xiao GG, Liu Y, et al. How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. Acta Biochim Biophys Sin (Shanghai). 2008 May;40(5):426–36.

384. Liu L, Zhu X, Ma Y, Piao H, Yang Y, Hao X, et al. Combining sequence and

network information to enhance protein–protein interaction prediction. BMC Bioinformatics. 2020 Dec 16;21(16):537.

385. Zitnik M, Leskovec J. Predicting multicellular function through multi-layer tissue networks. Bioinformatics. 2017 Jul 15;33(14):i190–8.

386. Yue Y, Ye C, Peng PY, Zhai HX, Ahmad I, Xia C, et al. A deep learning framework for identifying essential proteins based on multiple biological information. BMC Bioinformatics. 2022 Aug 4;23(1):318.

387. Zeng X, Zhu S, Lu W, Liu Z, Huang J, Zhou Y, et al. Target identification among known drugs by deep learning from heterogeneous networks. Chem Sci. 2020 Feb 19;11(7):1775–97.

388. Matsubara T, Ochiai T, Hayashida M, Akutsu T, Nacher JC. Convolutional neural network approach to lung cancer classification integrating protein interaction network and gene expression profiles. J Bioinform Comput Biol. 2019 Jun;17(3):1940007.

389. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, et al. COPASI—a COmplex PAthway SImulator. Bioinformatics. 2006 Dec 15;22(24):3067–74.

390. Demidenko E, Miller TW. Statistical determination of synergy based on Bliss definition of drugs independence. PLoS One. 2019 Nov 25;14(11):e0224137.

ANNEX A

Supplementary Data

Supplementary table 1. *The list of differentially expressed genes present in the RA M1 macrophage model that we identified using literature search and omics data analysis. . The first column contains the DEGs HGNC names. The second and fifth columns contain their corresponding Boolean values (after data discretization) observed in GSE97779 dataset and literature respectively.*

DEG	Boolean value in GSE97779	Adjusted p_value	logFC	Boolean value in literature	Reference
C5A				1	29220376
PRKCD	1	0,0008	0,86		
BAD	0	0,02492	-0,42		
STAT2				1	27626941
IRF9	1	0,00987	0,65	1	27626941
INHBA				1	13130463
SMAD4	0	0,012	-0,46		
CASP1	1	0,03031	0,71		
SIRT1	0	0,0284	-0,61	0	25799392
CCL21				1	21225692
PRKCQ	1	0,0037	1,53		
INHBB				1	26359667
TRAF6	0	0,00612	-0,73		
SMAD7	1	7,88E-05	2,17		
INPP5A	1	0,0091	0,9		
DUSP1	1	5,59E-07	4,17		
PRKG1	1	0,00311	2,17		
ACVR2A	0	0,01746	-0,79		
ACVR2B	0	0,00045	-2,85		
BCL2	1	0,01417	1,67		
BCL2L1				1	28118944
BCL2L11	1	0,0833	1,78		
BCL3	1	0,04434	0,42		
BIRC2	1	0,00425	0,594		
C5AR1				1	29220376
CASP3	0	0,00542	-0,86		
CASP7	1	0,00655	0,96		
CCL2				1	33330982

CCR2	1	7,69E-07	5,85		
CFLAR	1	0,0029	0,8	1	12228167
CSF2RA	1	0,00287	1,11	1	24936585
CSF2RB	1	0,00099	0,93	1	24936585
CXCL1				1	7561066
FAS	1	0,0005	2,45		
FOS	1	2,99E-08	3,93	1	27626941
GNA13	1	0,00017	1,62		
GNAI3	1	0,00098	0,72		
GNB1	1	0,00507	0,47		
HLA-B	1	0,02221	0,42		
HRAS	1	0,00388	0,94		
IFNA1	1	0,00831	2,43		
IFNB				1	15878901
IFNGR1	1	0,01789	1,35	1	25708927
IFNGR2	1	0,00415	0,67	1	25708927
IL11				1	29327326
IL12RB1	1	0,00343	1,29		
IL18				1	10562301
IL18R1	1	0,00057	3,64		
IL1RAP	1	0,00018	2,73		
IL23				1	25799392
IL6	1	0,02477	1,56		
IL6ST	1	0,01567	0,56		
IL8				1	10491366
IRF7	1	1,89E-05	2,44	1	22614743
JAK1	1	0,00433	1		
JAK2	1	0,0011	2,3		
LILRB1	1	0,0015	0,96		
MAP2K1	1	0,0001	1,12		
MAP2K2	1	0,00845	0,64		
MAP2K3	1	0,00258	1,03		
MAP2K6	1	0,02686	1,78		
MAPK14	0	0,0243	-0,61		
MAPK3				1	17907188
MAPK8	1	0,00374	0,71		
MAPKAPK2	1	0,00399	0,78		
MCL1	1	0,00118	1,39	1	17009247
MDM2	0	0,01706	-1,06		

MYC	1	0,00027	1,66		
NFAT5	1	0,01494	0,97		
NFKB1	1	0,00235	0,89	1	8630106
NFKBIA	1	0,00021	0,98		
NFKBIE	1	0,01037	0,79		
NLRP3	1	8,50E-06	4,44		
OPN3	0	0,00517	-0,87		
PPIA	1	0,01626	1,05		
PTK2	1	0,00073	1,87		
PTPN6	1	2,86E-05	1		
RAC1	1	0,03366	0,44		
RAF1	1	0,01113	1		
RBPJ	1	3,82E-05	2,5		
RELA	1	0,01152	0,66		
SMAD2	0	0,01636	-0,45		
SOS1	1	0,01592	1,32		
STAT1	1	0,00057	2,22	1	22614743
STAT3	1	0,00043	0,89		
STAT4				1	10779770
STAT5B	1	0,00037	2,34		
SYK	0	0,03668	-0,38		
TLR1	1	1,21E-06	1,65		
TLR2	1	8,13E-07	2,2	1	15146415
TLR4	0	0,01061	-0,67		
TLR5	0	0,00032	-1,25		
TLR8	1	0,04753	0,68		
TLR9				1	26759164
TNF	1	7,84E-05	2,5	1	2109776
TNFRSF11	1	0,01287	1,79		
TNFRSF1A				1	9189061
TRADD	1	0,03339	0,47		
TRAF3	0	0,00022	-1,14		
TRAM1	0	0,01224	-0,4		
XIAP	1	0,00215	1,12	1	19171073
PPP4C	1	0,00131	0,68		
bigly-can_simple_mole-cule				1	19772831

DNA_simple_molecule				1	19772831
dsRNA_simple_molecule				1	19772831

Supplementary table 2. *List of nodes upstream the phenotypes of interest in the RA M1 macrophage model associated with their mean values over the fixpoints having the highest similarity score.*

Nodes	Mean values
CSF2RA_CSF2RB_complex	1
RELA_NFKB1_NFKBIE_complex	1
TLR5	0
IRF9	1
STAT2	1
IKK1_phosphorylated	0,5
col4a4	0,66666667
PTPN6	1
PRKG1	1
ACVR2A_ACVR2B_complex	0
ACVR2A_ACVR2B_INHBA_complex	0
ACVR2A_ACVR2B_INHBB_complex	0
AP_1	1
AP_1_phosphorylated	1
apoptosis_M1_macrophage_phenotype	0
ASC	1
ASK1	1
BAD	0
BCL2_M1_macrophage__Mitochondria_membrane	1
BCL2_M1_macrophage__Mitochondria_membrane_active	0
Bcl2_rna	1
BCL2L1_M1_macrophage__Mitochondria	1
BCL2L1_M1_macrophage__Mitochondria_active	0
BCL3_rna	1
biglycan_simple_molecule	1
Bim	1
c_FOS_M1_macrophage___Cytoplasm	1
c_FOS_M1_macrophage___Cytoplasm_active	1
c_FOS_M1_macrophage___nucleus	1
c_JUN	1

c_JUN_phosphorylated_M1_macrophage___Cytoplasm	1
c_JUN_phosphorylated_M1_macrophage___nucleus	1
c_Myc_rna	1
c5a	1
C5a_C5aR1_complex	1
Casp1	1
CASP3	0
CASP7	0
CASP8	1
CCL2_CCR2_complex	1
CCL2_M1_macrophage___Extracellular_Space	1
CCL2_M1_macrophage___Secreted_components	1
CCl21	1
CCL21_CCR7_complex	1
CD40LG_ITGB1_ITGA1_complex	0,5
cFLIP	1
cIAP1	1
col4a5	0,66666667
CRKL_phosphorylated	1
CSF2_M1_macrophage___Extracellular_Space	1
CSF2_M1_macrophage___Secreted_components	1
CSF2RA_CSF2RB_CSF2_complex	1
CXCL1_CXCR1_complex	1
CXCL1_M1_macrophage___Extracellular_Space	1
CXCL1_M1_macrophage___Secreted_components	1
CXCR1	1
CXCR1_IL8_complex	1
CypA	1
DAXX	1
DNA_simple_molecule	1
dsRNA_simple_molecule	1
DUSP1	1
ECSIT	0,5
ERK1_phosphorylated_M1_macrophage___Cytoplasm	1
ERK1_phosphorylated_M1_macrophage___nucleus	1
FADD	1
FAS	1
FASL_FAS_complex	1
FASL_M1_macrophage___Extracellular_Space	1

FASL_M1_macrophage___Secreted_components	1
FOXO_M1_macrophage___Cytoplasm	1
FOXO_M1_macrophage___nucleus	1
gal	1
GAL_GALR2_complex	1
gamma_secretase_complex_complex	1
GNA12_GNA13_complex	1
GNAI3	1
GNB_GNG_GNAI3_complex	1
HLA_B_LILRB1_complex	1
HRAS	1
IFNa_M1_macrophage___Extracellular_Space	1
IFNa_M1_macrophage___Secreted_components	1
IFNAR1_IFNAR2_complex	1
IFNAR1_IFNAR2_IFNa_complex	1
IFNAR1_IFNAR2_IFNb_complex	1
IFNb_M1_macrophage___Extracellular_Space	1
IFNb_M1_macrophage___Secreted_components	1
IFNE	0,5
IFNE_IFNAR1_IFNAR2_complex	0,5
IFNg_M1_macrophage___Extracellular_Space	1
IFNg_M1_macrophage___Secreted_components	1
IFNGR1_IFNGR2_complex	1
IFNGR1_IFNGR2_IFNg_complex	1
IKK_complex	1
IKK1_IKK2_complex	0,5
IKK2_phosphorylated	1
IKKE_TBK1_complex	0,5
IKKE_TBK1_TRAF3_complex	0
IL1_IL1R_complex	1
IL11_IL11Ra_IL6ST_complex	1
IL12_M1_macrophage___Extracellular_Space	1
IL12_M1_macrophage___Secreted_components	1
IL12RB_complex	1
IL12RB_IL12_complex	1
IL18_IL18R1_complex	1
IL18_M1_macrophage___Cytoplasm	1
IL18_M1_macrophage___Extracellular_Space	1
IL18_M1_macrophage___Secreted_components	1

IL1B_M1_macrophage___Cytoplasm	1
IL1B_M1_macrophage___Extracellular_Space	1
IL1B_M1_macrophage___Secreted_components	1
AKT1	0
IL1R_complex	1
IL23_M1_macrophage___Extracellular_Space	1
IL23_M1_macrophage___Secreted_components	1
IL23R_IL12RB1_complex	1
IL23R_IL12RB1_IL23_complex	1
IL6_IL6R_IL6ST_complex	1
IL6_M1_macrophage___Extracellular_Space	1
IL6_M1_macrophage___Secreted_components	1
IL8_M1_macrophage___Extracellular_Space	1
IL8_M1_macrophage___Secreted_components	1
INHBA	1
INHBB	1
INPP5A	1
IRAK1	1
IRAK1_IRAK4_complex	1
IRAK4_phosphorylated	1
IRF3_phosphorylated	1
IRF7	1
ITGB1_ITGA1_col4a_complex	1
ITGB1_ITGA1_complex	1
JAK1	1
JAK1_JAK2_complex	1
JAK1_TYK2_complex	1
JAK2	1
JAK2_TYK2_complex	1
JNK1_phosphorylated_M1_macrophage___Cytoplasm	1
JNK1_phosphorylated_M1_macrophage___nucleus	1
Mcl1_rna	1
MDM2_phosphorylated	0
MEK1_phosphorylated	1
MEK2_phosphorylated	1
MEKK1	0,5
MK2_phosphorylated	1
MKK3_phosphorylated	1
MKK4_phosphorylated	1

MKK6_phosphorylated	1
MKK7_phosphorylated	1
MYD88	1
MYD88_TIRAP_TOLLIP_complex	1
NCID	1
NFAT5	1
NFKB1_TPL2_complex	1
NFKBIA_RELA_NFKB1_complex	1
NICD	1
NICD_CSL_SKIP_MAML1_ep300_complex	1
NIK	1
NLRP3	1
NLRP3_INFLAMMASOME_complex	1
notch1_JAG1_complex	1
OPN	0
osteoclastogenesis_M1_macrophage_phenotype	1
p15_rna	0
p21_rna	1
p300_SP1_complex	0
p38_MAP_KINASE_phosphorylated_M1_macrophage___Cyto-plasm	0
p38_MAP_KINASE_phosphorylated_M1_macrophage___nu-cleus	0
p53_phosphorylated	1
PI3K	0
PIK3AP1_phosphorylated	0
PP4	1
Prkcd	1
PRKCQ	1
proliferation_survival_M1_macrophage_phenotype	1
PTK2	1
Rac1	1
RAF1	1
RELA_NFKB1_complex_M1_macrophage___Cytoplasm	1
RELA_NFKB1_complex_M1_macrophage___nucleus	1
RHOA	0
SHP2_GRB2_complex	1
Sirt1	0
SMAD2_phosphorylated	0

SMAD2_SARA_complex	0
SMAD2_SMAD4_complex	0
SMAD4	0
SMAD7	1
SOS1	1
Src	1
STAT1	1
STAT1_STAT1_complex	1
STAT1_STAT2_IRF9_complex	1
STAT3	1
STAT3_STAT3_complex	1
STAT4	1
STAT4_STAT4_complex	1
STAT5_CRKL_complex	1
STAT5_phosphorylated	1
SYK	0
TAB1	1
TAB2_phosphorylated	1
TAK1	1
TLR1_TLR2_biglycan_complex	1
TLR1_TLR2_complex	1
TLR2_TLR6_biglycan_complex	1
TLR2_TLR6_complex	1
TLR3_dsRNA_complex	1
TLR4_Md2_CD14_fibrinogen_complex	0
TLR7_TLR8_ssRNA_complex	1
TLR9_DNA_complex	1
TNF_M1_macrophage___Extracellular_Space	1
TNF_M1_macrophage___Secreted_components	1
TNF_TNFRSF1A_complex	1
TNFA_rna	1
TNFSF11	1
TNFSF11_TNFRSF11_complex	1
TPL2	1
TRADD	1
TRADD_TRAF2_RIP1_complex	1
TRAF2_RIP1_TRADD_TAK1_TAB1_TAB2_complex	1
TRAF2_TRAF6_complex	1
TRAF3	0

TRAF6	0
TRAF6_ECSIT_MEKK1_TAB1_TAB2_TAK1_complex	0
TRAF6_TAB1_TAB2_TAK1_complex	0
TRAF6_ubiquitinated	0
TRAM1	0
TRAM1_TRIF_complex	0
TRIF	1
TSG6	1
UEV1A_UBC13_complex	0,5
XIAP	1

Supplementary table 3. *The list of differentially expressed genes present in the RA M2 macrophage model that we identified using literature search and omics data analysis. The first column contains the DEGs HGNC names. The second and fifth columns contain their corresponding Boolean values (after data discretization) observed in GSE97779 dataset and literature respectively.*

DEG	Boolean value in GSE97779	Adjusted p_value	logFC	Boolean value in literature	Reference
VEGFB	0	9,64E-06	-1,46		
PRLR	0	0,01306	-1,45		
MDM2	0	0,01706	-1,06		
CASP3	0	0,00542	-0,86		
TRAF6	0	0,00612	-0,73		
CREB1	0	0,03541	-0,66		
NRP2	0	0,03197	-0,65		
MAPK14	0	0,0243	-0,61		
SIRT1	0	0,0284	-0,61	0	25799392
SYK	0	0,015846	-0,53		
SMAD4	0	0,012	-0,46		
SMAD2	0	0,01636	-0,45		
BAD	0	0,02492	-0,42		
BCL3	1	0,04434	0,42		
HLA-B	1	0,02221	0,42		
CEBPB	1	0,00957	0,43		
SHC1	1	0,03891	0,44		
STAT6	1	0,02689	0,44		
BAX	1	0,02724	0,45	1	12634940
GNB1	1	0,00507	0,47		
IL6ST	1	0,01567	0,56		

CYLD	1	0,01375	0,61		
FCGR2A	1	0,02327	0,63	1	17521421
MAP2K2	1	0,00845	0,64		
RELA	1	0,01152	0,66		
CASP1	1	0,03031	0,71		
GNAI3	1	0,00098	0,72		
TGFB1	1	0,00044	0,78		
NFKBIE	1	0,01037	0,79		
CFLAR	1	0,0029	0,8	1	12228167
RXRA	1	0,0023	0,84		
IL4R	1	0,00047	0,86	1	7492352
PRKCD	1	0,0008	0,86		
NFKB1	1	0,00235	0,89	1	8630106
STAT3	1	0,00043	0,89		
HRAS	1	0,00388	0,94		
CASP7	1	0,00655	0,96		
LILRB1	1	0,0015	0,96		
NFAT5	1	0,01494	0,97		
NFKBIA	1	0,00021	0,98		
JAK1	1	0,00433	1		
PTPN6	1	2,86E-05	1		
RAF1	1	0,01113	1		
SH2D1A	1	0,04778	1,01		
IL17RA	1	0,0044	1,02	1	19265168
MAP2K3	1	0,00258	1,03		
EFNB1	1	0,0094	1,09		
MAP2K1	1	0,0001	1,12		
PLCG2	1	0,00027	1,12		
XIAP	1	0,00215	1,12	1	19171073
HCK	1	5,23E-05	1,13	1	17963503
IL12RB1	1	0,00343	1,29		
SOS1	1	0,01592	1,32		
HBEGF	1	0,04278	1,34	1	31068444
MCL1	1	0,00118	1,39	1	17009247
LIFR	1	0,01414	1,53		
PRKCQ	1	0,0037	1,53		
MYC	1	0,00027	1,66		
BCL2	1	0,01417	1,67		
BCL2L11	1	0,0833	1,78		

MAP2K6	1	0,02686	1,78		
PTK2	1	0,00073	1,87		
VEGFA	1	0,00429	1,9		
SMAD7	1	7,88E-05	2,17		
JAK2	1	0,0011	2,3		
FAS	1	0,0005	2,45		
FCGR3A	1	2,49E-05	2,77	1	22235253
FCGR1A	1	0,00011	3,32	1	17521421
DUSP1	1	5,59E-07	4,17		
NLRP3	1	8,50E-06	4,44		
KLF4	1	1,77E-08	4,66	1	29997611
BCL2L1				1	28118944
C5A				1	29220376
C5AR1				1	29220376
CCL21				1	21225692
CSF1				1	27036883
CSF1R				1	27036883
IL10				1	20001767
IL11				1	29327326
IL17F				1	19265168
IL17RC				1	19265168
IL23				1	25799392
IL34				1	22039170
IL4				1	7492352
IL8				1	10491366
MAPK3				1	17907188
TGFBR1				1	7614737
TGFBR2				1	7614737

Supplementary table 4. *List of nodes upstream the phenotypes of interest in the RA M2 macrophage model associated with their mean values over the fixpoints having the highest similarity score.*

Nodes	Mean values
AKT1	0
AKT1_phosphorylated	0
apoptosis_M2_macrophage_phenotype	1
ASC	0,5
ASK1	1
BAD	0

BAX	1
BCL2_M2_macrophage_mitochondrion_membrane	1
BCL2_M2_macrophage_mitochondrion_membrane_active	0
Bcl2_rna	1
BCL2L1	0
BCL3_rna	1
Bim	1
C_EBPb_phosphorylated	1
c_Myc_rna	1
c5a	1
C5a_C5aR1_complex	1
Casp1	1
CASP3	1
CASP7	1
CASP8	1
CASP9	0
CCl21	1
CCL21_CCR7_complex	1
CD32a	1
CD40LG_ITGB1_ITGA1_complex	0,5
cFLIP	1
cMyc	1
cMyc_phosphorylated	1
col4a4	0,66666667
col4a5	0,66666667
CREB1_phosphorylated	0
CSF1_M2_macrophage__extracellular_space	1
CSF1_M2_macrophage__secreted_components	1
CSF1R	1
CSF1R_CSF1_complex	1
CSFR1R_IL34_complex	1
CXCR1_IL8_complex	1
CYLD	1
DAG_simple_molecule	1
DAXX	1
DUSP1	1
EFNB1_EPHB1_complex	1
ERK1_phosphorylated_M2_macrophage__cytoplasm	1
ERK1_phosphorylated_M2_macrophage_nucleus	1

FADD	1
FAS	1
FASL_FAS_complex	1
FASL_M2_macrophage__extracellular_space	1
FASL_M2_macrophage__secreted_components	1
FOXO1	1
GAB2_phosphorylated	1
GAS6	0,5
GAS6_MERTK_complex	0,5
GNAI3	1
GNB_GNG_complex	1
GNB_GNG_GNAI3_complex	1
Grb2	1
GSK3B	1
HBGEF	1
HCK	1
HLA_B_LILRB1_complex	1
HRAS	1
IKK_complex	1
IKK1_IKK2_complex	0,5
IKK1_phosphorylated	0,5
IKK2_phosphorylated	1
IL10_M2_macrophage__extracellular_space	1
IL10_M2_macrophage__secreted_components	1
IL10R1_IL10R2_complex	1
IL10R1_IL10R2_IL10_complex	1
IL11Ra_IL6ST_IL11_complex	1
IL17Ra_IL17Rc_IL17F_complex	1
IL23_M2_macrophage__extracellular_space	1
IL23_M2_macrophage__secreted_components	1
IL23R_IL12RB1_complex	1
IL23R_IL12RB1_IL23_complex	1
IL34	1
IL4	1
IL4_IL4Ra_complex	1
IL4R	1
IL8_M2_macrophage__extracellular_space	1
IL8_M2_macrophage__secreted_components	1
immune_complex_CD16a_complex	1

immune_complex_CD32a_complex	1
immune_complex_CD32b_complex	1
immune_complex_CD64_complex	1
immune_complex_complex	1
ITGB1_ITGA1_col4a_complex	1
ITGB1_ITGA1_complex	1
JAK1	1
JAK1_TYK2_complex	1
JAK2	1
klf4	1
LIFR_IL6ST_CTF1_complex	1
mcl1	0
Mcl1_rna	0
MDM2_phosphorylated	0
MEK1_phosphorylated	1
MEK2_phosphorylated	1
MKK3_phosphorylated	1
MKK6_phosphorylated	1
MSK1_phosphorylated	0
NFAT5_phosphorylated	1
NFKB1_TPL2_complex	1
NFKBIA_RELA_NFKB1_complex	1
NLRP3	1
NLRP3_INFLAMMASOME_complex	0,5
p15_rna	1
p38_MAP_KINASE_phosphorylated_M2_macrophage__cyto-plasm	0
p38_MAP_KINASE_phosphorylated_M2_macrophage_nu-cleus	0
p53_phosphorylated_M2_macrophage__cytoplasm	1
p53_phosphorylated_M2_macrophage_nucleus	1
PI3K	0
PIK3AP1_phosphorylated	0
PIP2_simple_molecule	1
PLCG2	1
PRKCD	1
Prkcd	1
PRKCQ	1
PRL	0

PRL_PRLR_complex	0
proliferation_survival_M2_macrophage_phenotype	0
PTK2	1
PTPN6	1
RAF1	1
RBL1_E2F4_DP1_complex	0
RELA_NFKB1_complex_M2_macrophage__cytoplasm	1
RELA_NFKB1_complex_M2_macrophage_nucleus	1
RELA_NFKB1_NFKBIE_complex	1
RXRa_NUR77_complex	1
SH2D1A	1
Shc_phosphorylated	1
SHIP1	1
Sirt1	0
SMAD2_phosphorylated	0
SMAD2_SARA_complex	0
SMAD2_SMAD4_complex	0
SMAD4	0
SMAD7	1
SOS1	1
Src	1
STAT3	1
STAT3_STAT3_complex	1
STAT6	1
STAT6_STAT6_complex	1
SYK	0
Syk_phosphorylated	1
TAK1_phosphorylated	1
TGFB1_M2_macrophage__extracellular_space	1
TGFB1_M2_macrophage__secreted_components	1
TGFBR1_TGFBR2_complex	1
TGFBR1_TGFBR2_TGFB1_complex	1
TPL2	1
TRAF3IP2_phosphorylated	1
TRAF6_ubiquitinated	0
VEGFa_M2_macrophage__extracellular_space	1
VEGFa_M2_macrophage__secreted_components	1
Vegfa_rna	1
Vegfb_M2_macrophage__extracellular_space	0

Vegfb_M2_macrophage__secreted_components	0
Vegfc_M2_macrophage__extracellular_space	0
Vegfc_M2_macrophage__secreted_components	0
vegfc_vegfr3_complex	0
VegfR1	0,5
VegfR1_Vegfa_complex	0,5
VegfR1_vegfb_complex	0
VegfR2_vegfc_nrp2_complex	0
XIAP	1

Supplementary table 5. *The list of differentially expressed genes present in the RA fibroblast model that we identified using literature search and omics data analysis. The first column contains the DEGs HGNC names. The second and fifth columns contain their corresponding Boolean values (after data discretization) observed in GSE109449 dataset and literature respectively.*

DEG	Boolean value in GSE109449	Adjusted p_value	logFC	Boolean value in literature	Reference
ADCY8	0	0.002	-3,02		
AREG				1	18439312
ARHGEF2				1	17515956
BAX				0	23421940
BCL2				1	17515956
BCL2L1				1	28118944
BCL2L11				0	28118944
BID	0	0.004	-1,47		
CASP3	0	4.524e-05	-0,72		
CCL18				1	11745396
CCL2	1	0.001	2,19	1	33330982
CCL5				1	9756723
CCR2				1	11673556
CCR5				1	11673556
CD14				1	18452992
CD40				1	10799861
CFLAR				1	15593196
Cgas				1	26819496
COMP	1	5.036e-18	4,33	1	19652761
CSF2				1	31142839
CSF2RA				1	24936585
CSF2RB				1	24936585

CSK				1	17515956
CXCL10				1	28148302
CXCL13				1	27102921
CXCL8				1	11178128
CXCR2				1	20036936
CXCR3				1	21811993
CXCR4				1	32782501
CYCS	0	0.004	-0,687		
FN1				1	36225320
EDA				1	7748223
EGF				1	18439312
EGFR				1	18439312
GAB2	0	0.0004	-1,35		
FAS				1	11169523
FASLG	1	0.022	0,14	1	11169523
FGF1	1	0.004	1,76	1	16893535
FGFR4	1	0.039	0,84	1	8651984
IGF1R	0	0.003	-1,17		
FOS				1	7747113
FOXO1				0	24812285
FZD5				1	11315916
NRP1	0	0.001	-0,89		
HBEGF				1	31068444
ICAM1				1	17568789
IFNB				1	15878901
IFNG				1	31061532
IGF1				0	11934980
RHOA	0	0.004	-0,59		
IL10				1	20001767
IL10RA				1	27626941
IL10RB				1	27626941
IL17				1	23858337
IL18				1	10562301
IL18R				1	17530707
IL1A				1	30272996
IL1B				1	30272996
DUSP7	1	0.003	0,85		
IL6				1	32718086
IRAK4				1	18452992

IRF5				1	26315890
EPHB2	1	0.013	1,75		
GAB1	1	0.0003	1,55		
IL1R1	1	0.01	1,16		
JUN				1	7747113
LY96				1	26352601
MAP2K3				1	14695331
MAP2K4				1	13130464
MAP2K6				1	14695331
MAP2K7				1	13130464
MAP3K1				1	22736089
ITGA4	1	0.0002	2,12		
MCL1				1	16339575
MDM2				1	26655743
MIF				1	12011381
MIR146A				0	29844864
MIR192				0	28321538
MIR650				0	28129626
MMP1				1	12379519
MMP13				1	11040455
MMP3				1	12379519
NFAT5				1	21717420
NFKB1				1	12010604
ITGA5	1	0.019	1,35		
P53				1	11169523
PDGFC	1	0.042	0,18	1	26976956
PDGFRB	1	3.556e-08	1,86	1	26976956
PITPNM3				1	23728190
MAP4K4	1	1.941e-06	1,86		
PTPN11				1	23335101
PLXNB1	1	0.0003	1,45		
PTPRC	1	0.012	1,36		
RANKL				1	10693864
RAC1	1	0.035	0,61		
RIPK1				1	32116107
SEMA4A				1	26303122
SEMA7A				1	28109308
SOCS3				1	30854695
SRC	1	0.003	1,78		

SST				0	9844773
SSTR				0	9844773
STAT1	1	0.006	1,79	1	25630235
STAT3				1	30477351
TAB2				1	18452992
TBK1				1	18452992
TCF19	1	0.030	1,76		
TGFB1				1	11966774
TGFBR1				1	17594488
TIRAP				1	18452992
TLR2				1	32256787
TLR4				1	32256787
TLR5				1	22661088
TNF				1	34588517
TNFR				1	16951485
TRADD				1	16951485
TRAF2				1	16951485
TRAF6				1	18452992
TXK	1	0.010	0,79		
VAV3	1	0.038	1,41		
VEGFA	1	0.002	1,55	1	18439312
WNT5A				1	11315916
YY1				1	26821827

Supplementary table 6. *List of nodes upstream the phenotypes of interest in the RA fibroblast model associated with their mean values over the fixpoints having the highest similarity score.*

Nodes	Mean values
ADCY8	0
AREG	1
ARHGEF2	1
BAX_Fibroblast___Mitochondrion	0
BCL2	1
BCL2L1_rna	1
BCL2L11_Fibroblast___Cytoplasm	0
BID	0
CASP3	0
CCL18	1
CCL2_Fibroblast___secreted_components	1

CCL5_Fibroblast___secreted_components	1
CCR2_CCL2_complex	1
CCL5_CCR5_complex	1
LBP_CD14_complex	1
CFLAR_Fibroblast___Cytoplasm	1
cGAS	1
COMP	1
CRKL_phosphorylated	1
CSF2_Fibroblast___Extracellular_Space	1
CSF2RA_CSF2RB_complex	1
CSK	1
CXCL10_Fibroblast___Extracellular_Space	1
CXCL13	1
CXCL8_Fibroblast___Extracellular_Space	1
CXCR2_CXCL8_complex	1
CXCL10_CXCR3_complex	1
MIF_CXCR4_complex	1
CYCS	0
FN1_Fibroblast___secreted_components	0
EDA	1
EGF	1
EGFR	1
GAB2_phosphorylated	1
FAS_FASL_complex	1
FGF1	1
IGF1_IGF1R_complex	0
FOS	1
FOXO1	0
WNT_FRIZZLED_complex	1
FN1_ITGAV_complex	0
VEGFR_NRP1_complex	0
HBEGF	1
IFNb1	1
IFNG_IFNGR1_R2_complex	1
IGF1	0
RHOA	0
IL10RA_IL10RB_IL10_complex	1
IL17A_Fibroblast___secreted_components	1
IL18_Fibroblast___secreted_components	1

IL18_IL18R_complex	1
IL1A	1
IL1B_Fibroblast___secreted_components	1
DUSP7	1
IL6_Fibroblast___secreted_components	1
IRAK4_Fibroblast___Cytoplasm	1
IRF5_Fibroblast___Cytoplasm	1
EPHB2	1
GAB1	1
IL1R1	1
JUN_phosphorylated	1
LY96	1
MAP2K3_phosphorylated	1
MAP2K4_phosphorylated	1
MAP2K6_phosphorylated	1
MAP2K7_phosphorylated	1
MAP3K1_phosphorylated	1
ITGA4_ITGB1_complex	1
MCL1	1
MDM2	1
MIF	1
MIR146A_rna	0
MIR192_rna	0
MIR650_rna	0
MMP1	1
MMP13	1
MMP3	1
NFAT5_phosphorylated	1
NFKB_complex	1
ITGA5_ITGB1_col4a5_complex	1
TP53_phosphorylated	1
PDGFC	1
PDGFC_PDGFRB_complex	1
PITPNM3_CCL18_complex	1
MAP4K4_phosphorylated	1
PTPN11_phosphorylated	1
PLXNB1_MET_sema4a_complex	1
NECTIN3_PTPRC_NECTIN1_complex	1
RANK_RANKL_complex	1

RAC1_2	1
RIPK1	1
ITGA4_ITGB1_sema7a_complex	1
SOCS3	1
SRC_phosphorylated	1
sst	0
sst_sstr_complex	0
STAT1_Fibroblast___Cytoplasm	1
STAT3_Fibroblast___Cytoplasm	1
TAB1_TAB2_complex	1
TBK1_IKBKE_complex	1
TCF_LEF	1
TGFb1	1
TGFB1_TGFBR1_complex	1
TIRAP	1
LY96_TLR2_4_complex	1
TLR5	1
TNF_Fibroblast___secreted_components	1
TNF_TNFRSF1A_B_complex	1
TRADD	1
TRAF2	1
TRAF6	1
TXK	1
VAV1_2_3_phosphorylated	1
ADAMTS9_rna	0,5
AKT2	1
AKT2_phosphorylated	1
AMAP1	1
APAF1	0,5
apoptosis_fibroblast_phenotype	0
Apoptosome_complex	0
AREG_EGFR_complex	1
BAD	0
BAX_Fibroblast___Mitochondrion_active	0
BCL2L11_Fibroblast___Cytoplasm_active	0
BRAF_phosphorylated	0
CALCINEURIN	1
CASP8	0
CASP9	0,5

CAV1_rna	1
CCL2_Fibroblast___Extracellular_Space	1
CCL5_Fibroblast___Extracellular_Space	1
CD40LG_CD40_complex	1
Cell_chemotaxis_migration_fibroblast_phenotype	1
CFLAR_Fibroblast___Cytoplasm_active	1
col4a3	0,5
CREB1_phosphorylated	1
CSF2_Fibroblast___secreted_components	1
CSF2RA_CSF2RB_CSF2_complex	1
CTNNB_CK1A_AXIN_GSK3B_APC_complex	0
CTNNB1	0
CXCL10_Fibroblast___secreted_components	1
CXCL13_CXCR3_complex	1
CXCL8_Fibroblast___secreted_components	1
CXCR3	1
DAXX	1
DOCK2	0,5
DOCK2_CRKL_complex	0,5
DVL1_phosphorylated	1
EDA_EDA2R_complex	1
EGF_EGFR_complex	1
FADD	1
FASLG	1
FGF1_FGFR4_complex	1
FN1_Fibroblast___Extracellular_Space	0
FOXO3	0
GNAI3	1
GRB2	1
HBEGF_EGFR_complex	1
HOMODIMER_space_STAT1	1
ICAM1_ITGB2_ITGAL_complex	1
IFNa1	1
IFNa1_B1	1
IFNA1_B1_IFNAR1_R2_complex	1
IFNB1_rna	1
IKBA_NFKB1_RELA_complex	1
IKBKB_phosphorylated	1
IKK_complex	1

IKK1_IKK2_complex	0,25
IKK1_phosphorylated	0,5
IKK2_phosphorylated	0,5
IL17A_Fibroblast___Extracellular_Space	1
IL17A_IL17RA_complex	1
IL18_Fibroblast___Extracellular_Space	1
IL1A_IL1R1_complex	1
IL1B_Fibroblast___Extracellular_Space	1
IL1B_IL1R1_complex	1
IL6_Fibroblast___Extracellular_Space	1
IL6_IL6R_complex	1
IL6_IL6ST_complex	1
IRAK1_Fibroblast___Cytoplasm	1
IRAK1_Fibroblast___Cytoplasm_active	1
IRAK1_IRAK4_complex	1
IRAK4_Fibroblast___Cytoplasm_active	1
IRF3_phosphorylated	1
IRF5_Fibroblast___nucleus	1
IRF9	1
ISGF3_complex	1
ITGA4_ITGB1_col4a3_complex	0,5
JAK1	0
JAK2	0
JAK3	0
LCK_phosphorylated	0,5
LTBP1	0
MAP2K1_phosphorylated	1
MAP3K14_phosphorylated	1
MAP3K2_3_4	0,5
MAP3K5_phosphorylated	1
MAP3K7_phosphorylated	1
MAP3K8_phosphorylated	1
MAPK1_empty	0
MAPK1_phosphorylated	0
MAPK14_phosphorylated	1
MAPK3_complex	0,5
MAPK3_phosphorylated	1
MAPK8_phosphorylated	0,5
MAPK9_phosphorylated	1

MAPKAPK2_phosphorylated	1
MDM2_phosphorylated	1
MIR10a_rna	0
MIR338_5P_rna	0,5
MIR451A_rna	0
MYD88	1
NFKB_N_complex	1
NFKB1_MAP3K8_complex	1
ngef	1
NRAS	0
p38MAPK_empty	1
p38MAPK_phosphorylated	0
PI3_4_5_P__3_simple_molecule	0,5
PI4_5_P__2_simple_molecule	1
PIK3R5_phosphorylated	1
PP2A	0
PRKACA	0
PRKCQ	0,5
PRKG1	1
proliferation_survival_fibroblast_phenotype	1
PTEN	0,5
PTK2B_phosphorylated	1
PTPN6	0,5
pyk2_phosphorylated	1
RAB5A	1
RAF1	0,5
rasa1	1
RELA_NFKB1_NFKBIE_complex	1
RIPK1_TRAF6_complex	1
RPS6KB1_phosphorylated	1
SARM1	0
sema7a	1
SH2D1A	0,5
SHC2_phosphorylated	1
SHP2_GRB2_complex	1
SOS1	1
STAT1_Fibroblast___nucleus	1
STAT1_STAT2_complex	1
STAT2	1

STAT3_Fibroblast___nucleus	1
TAB1_TAB2_TAK1_complex	1
TAB1_TAB2_TRAF6_complex	1
tBID	0
TGFB1_Fibroblast___Extracellular_Space	1
TGFB1_Fibroblast___Extracellular_Space_active	1
TICAM1	1
TICAM1_TICAM2_complex	1
TICAM2	1
TIRAP_MYD88_complex	1
TNF_Fibroblast___Extracellular_Space	1
TNFRSF10A_rna	1
TNFRSF10B_rna	1
TNFSF11_Fibroblast___Extracellular_Space	1
TNFSF11_Fibroblast___secreted_components	1
TP73_phosphorylated	0
TRAF2_TRAF5_complex	1
TRAF3	1
TRAF3_TRAF6_complex	1
TRAF3IP2_phosphorylated	1
TRAF6_phosphorylated	1
TYK2	1
VEGFa_Fibroblast___Extracellular_Space	1
VEGFa_Fibroblast___secreted_components	1
VEGFA_rna	1
VEGFA_VEGFR_NRP1_complex	0
WNT5A	1
YWHAQ	0
YY1	1
ZC3H12A	0

Supplementary table 7. *The list of differentially expressed genes present in the RA CD4+ Th1 model that we identified using literature search and omics data analysis. The first column contains the DEGs HGNC names. The second and fifth columns contain their corresponding Boolean values (after data discretization) observed in SDY998 dataset and literature respectively.*

DEG	Boolean value in SDY998	Adjusted p_value	logFC	Boolean value in literature	Reference
F2R	1	3,47E-13	3,77		

CCL2				1	1522232
CCR5	1	1,19E-16	3,81	1	10323208
CCL5	1	2,48E-11	6,45		
CCL4	1	1,12E-13	5,97		
CXCL10				1	28148302
CXCR3				1	27190305
CXCL16				1	16200580
EOMES	1	1,59E-44	5,67		
FGR	1	1,59E-18	6,4		
LAG3	1	1,46E-11	4,8		
HLA-DRB5	1	0,0003	5,52		
HLA-DQA1	1	4,19E-14	5,27		
HLA-DQB2	1	1,53E-08	4,94		
HLA-DQA2	1	4,44E-16	5,2		
HLA-DRB1	1	5,97E-09	3,97		
HLA-DPB1	1	6,87E-08	3,59		
IFNG	1	2,98E-15	5,06	1	8961900
IFNGR1				1	25708927
IFNGR2				1	25708927
IL12				1	32264938
IL12RB1				1	32264938
IL12RB2				1	32264938
IL18R1				1	14994387
IL18				1	14532149
IL9R				1	19723899
IL2R				1	32264938
IL9				1	26078482
TIM3	1	1,04E-19	7,35		
CD74	1	0,014	2,89		
IL27				1	17015723
IL27R				1	32518420
STAT1				1	33779079
STAT4				1	17804842
TBX21				1	3333523
TNFSF4	1	1,54E-09	5,96	1	11069062
VCAM1	1	2,97E-15	7,01		

Supplementary table 8. *List of nodes upstream the phenotypes of interest in the RA CD4+ Th1 model associated with their mean values over the fixpoints having*

the highest similarity score.

Nodes	Mean values
AGT	1
AGT_F2R_complex	1
AKT1	1
apoptosis_TH1_phenotype	0
ARHGEF2	1
BCL2L11	0
CCL2	1
CCL2_CCR2_complex	1
CCL4_5_CCR5_complex	1
CCL4_TH1___extracellular_space	1
CCL4_TH1___secreted_components	1
CCL5_TH1___extracellular_space	1
CCL5_TH1___secreted_components	1
CCR5	1
CD28_CD86_complex	0,5
CD32b_igG_complex	0
CD84_CD84_complex	0,5
Cell_chemotaxis_migration_TH1_phenotype	1
COL4A3_ITGA4_ITGB7_complex	0,5
CREB1_phosphorylated	1
CXCL10_CXCR3_complex	1
CXCL10_TH1___extracellular_space	1
CXCL10_TH1___secreted_components	1
CXCL16	1
CXCL16_CXCR6_complex	1
EOMES	1
FGR	1
GAS6	0,5
gas6_mertk_complex	0,5
GNA12_GNA13_complex	1
GNAI_GNB_GNG_complex	1
GNAI_TH1___cytoplasm	1
GNAI_TH1___cytoplasm_active	1
GNB_GNG_complex	1
GRB2	1
HLA_DP_DQ_DR_LAG3_complex	1
HLA_DP_DQ_DR_TCR_CD3_complex	1

icoslg_icos_complex	0,5
IFNG_IFNGR1_IFNGR2_complex	1
IFNg_TH1___extracellular_space	1
IFNg_TH1___secreted_components	1
IFNGR1_IFNGR2_complex	1
igG	0
IKK_complex	1
IL10_IL10RA_IL10RB_complex	1
IL12_TH1___extracellular_space	1
IL12_TH1___secreted_components	1
IL12Rb1_IL12Rb2_complex	1
IL12Rb1_IL12Rb2_IL12_complex	1
IL18_IL18R1_IL18RAP_complex	1
IL18_TH1___extracellular_space	1
IL18_TH1___secreted_components	1
IL18R1_IL18RAP_complex	1
IL27	1
IL27_IL27RA_complex	1
IL7_IL7R_IL2RG_complex	0,5
IL9R_IL2RG_IL9_complex	1
IRAK1_IRAK4_complex	1
JAK1_JAK2_complex	1
JAK1_JAK2_TYK2_complex	1
JAK1_JAK3_complex	1
JAK1_TYK2_complex	1
JAK2_TYK2_complex	1
KRAS	1
LAT_phosphorylated	0
LGALS9	1
LGALS9_TIM3_complex	0
LTA_TH1___extracellular_space	0
LTA_TH1___secreted_components	0
LTA_TNFRSF14_complex	0
lyn	1
MAP2K1_phosphorylated	1
MAPK1_phosphorylated	1
MIF	1
MIF_CD74_complex	1
mtor	1

MYD88	1
NFKB1_RELA_complex_TH1___cytoplasm	1
NFKB1_RELA_complex_TH1___nucleus	1
NFKB1_RELA_NFKBIA_complex	1
NIK	1
PDCD1	1
PI3K_phosphorylated	1
PLCG1	1
PLXNB1	1
PRKCQ	1
proliferation_survival_TH1_phenotype	1
PTK2	1
Rac1	0
RAF1	1
RHOA	1
RPS6KA4	1
RUNX3	1
SEMA4_PLXNB1_complex	0,75
SEMA4A_TH1___extracellular_space	0,5
SEMA4A_TH1___secreted_components	0,5
SEMA4D	1
SH2D1A	0,5
SHC1_phosphorylated	1
SHIP1	0
SOS1	1
Src	1
STAT1	1
STAT1_STAT1_complex	1
STAT4	1
STAT4_STAT4_complex	1
tbx21	1
tbx21_phosphorylated	1
TNF	1
TNF_TNFRSF1B_complex	1
TNFA	1
TNFSF4	1
TNFSF4_TNFRSF4_complex	1
TRAF1_TRAF2_TRAF3_complex	1
TRAF2_TRAF5_complex	0

TRAF2_TRAF5_TRAF6_complex	1
TRAF6_phosphorylated	1
VAV1	0
VCAM1_ITGA4_ITGB7_complex	1
ZAP70	0

Supplementary table 9. *The list of additional differentially expressed biomolecules present in the multicellular model that we identified using literature search and omics data analysis. The first column contains the DEGs HGNC names. The second and fifth columns contain their corresponding Boolean values (after data discretization) observed in the cell-specific datasets and literature respectively.*

DEG	Boolean value in gene expression datasets	Adjusted p_value	logFC	Boolean value in literature	References
ACKR3	1	0,0001	5,13	1	33453247
AREG	1	1,00E-07	4,48		
BAX				1	12634940
CCL20				1	12695561
CCL3				1	15878203
CCL5	1	0,01	2,19		
CD40	1	0,002	2,29		
CD80	1	0,027	1,84		
CFLAR	1	0,002	0,8	1	12228167
COX2				1	32530555
CSF1				1	27036883
CXCL10	1	0,0004	6,02	1	32211348
CXCL11	1	0,0005	4,98		
CXCL12				1	12574387
CXCL13	1	0,0001	8,76	1	33292827
CXCL16				1	32211348
CXCL9	1	0,00004	5,13		
GNB1	1	0,005	0,67		
HBGEF				1	31068444
HES1	1	0,001	2,12		
HIF1A				1	11465705
HLA-DPB2	1	0,001	2,82		
HMGB1				1	35250993
ICAM1	1	0,044	0,68		

IL10				1	20001767
IL15	1	0,0001	2,69		
IL4				1	7492352
IRAK3	1	0,0001	1,51		
IRF5				1	26315890
ITCH	0	0,044	-0,6		
JAG1				1	28256007
LGALS9	1	0,008	0,73	1	32771893
LTA				1	29541795
MMP14				1	18567920
MMP3	1	0,014	3,55		
MMP9				1	18567920
NFKB1	1	0,002	0,89	1	8630106
NFKBIA	1	0,0002	0,98		
NFKBIE	1	0,01	0,79		
NOS2				1	9236674
RXRA	1	0,002	0,84		
SEMA4A				1	33605067
SOCS3	1	3,80E-07	5,11		
TNFAIP3	1	0,00005	2,36		
TRAF1	1	0,002	1,37		
TRAF6	0	0,006	-0,73		
VCAN	1	1,14E-06	5,2		
VEGFA	1	0,004	1,9		
ACVR2B	0	0,0004	-2,85		
CCL18				1	17350968
CCL20				1	12695561
COL4A3	1	0,008	2,8		
COX2				1	32530555
CXCL13	1	0,0001	8,76	1	33292827
CXCL16				1	32211348
FCGR2A	1	0,023	0,63	1	17521421
FN1	1	0,00007	4,74		
FZD5	1	0,001	1,83		
GAL				1	15616823
GNA13	1	0,00017	1,62		
HLA-DPB2	1	0,001	2,82		
IFNA	1	0,008	2,43		
IFNB				1	15878901

IL12				1	9756640
IL15	1	0,0001	2,69		
IL18				1	10562301
IL1B	1	0,004	2,17	1	15150426
IRF7	1	0,000018	2,44	1	22614743
ITCH	0	0,044	-0,6		
JAG1				1	28256007
MAP2K4				1	13130464
MAP2K7	1	0,048	0,72	1	13130464
MAP3K1	0	0,009	-0,72		
MAPK8	1	0,003	0,71		
MAPKAPK2	1	0,04	0,54		
MIF	1	0,00007	1,36		
MMP14				1	18567920
MMP3	1	0,014	3,55		
MMP9				1	18567920
NFKB1	1	0,002	0,89	1	8630106
NFKBIA	1	0,0002	0,98		
NFKBIE	1	0,01	0,79		
NLK	1	0,003	0,73		
OPN3	0	0,005	-0,87		
PDGFC				1	16508943
PLXNB2	1	0,00001	1,48		
PPRARG				1	22829690
PRKG1	1	0,003	2,17		
RAC1	1	0,033	0,74		
SEMA4D	1	0,0004	1,86		
SEMA7A				1	28109308
SMAD4	0	0,001	-0,81		
TCF7L2	1	0,019	1,97		
TNFAIP3	1	0,00005	2,36		
TRAF1	1	0,002	1,37		
ADAMTS4				1	22324945
ARRB2				1	21855149
BAK1				0	18177509
BMP6				1	14558086
BSG	1	0,002	0,59		
CCL21				1	21225692
CCL5				1	9756723

CCNB				1	16518573
CD40				1	10799861
CDKN1C	0	0,003	-1,63		
COL4A5	1	0,046	0,94		
CPNE4	1	0,003	3,09		
CSF1				1	27036883
CXCL1				1	7521808
CXCL2				1	33010041
CXCL3				1	29191223
CXCL9				1	15004722
DKK1				1	26785768
GAL				1	35023445
GAL				1	35023445
GAS6				1	34539648
HES1	1	0,003	0,64		
HLA-B				1	23259760
HLA-DRB1				1	
ICAM3	1	0,035	0,73		
IKK1				1	11160335
IL10				1	20001767
IL11				1	29327326
IL26				1	23055831
IL32				1	19248119
IL33				1	21431944
IL34				1	22264405
INHBB				1	26359667
IRF1				1	22401175
JAG1				1	32499639
JUNB	1	0,016	0,8	1	12905466
LGALS9				1	18050192
MAP3K7				1	17559674
MIR124A				0	21339227
MIR155				1	18383392
MIR203A				1	21279994
MIR221				1	25891943
MIR346				1	21611196
MIR34A				0	22161761
MMP9				1	3680518
NFKB1				1	12010604

NOTCH3				1	16307184
PDE4B	1	0,000975925	1,25		
PDIA3	1	0,0014184	0,82		
PLA2G2A	0	0,000480237	-1,67	1	21068383
PRL				1	27616146
RIPK3	0	0,028700766	-3,09		
RUNX1	0	0,000528477	-1,25		
SEMA3A	1	3,24E-12	1,76		
SEMA4D				1	25707877
STAT3				1	30477351
TGFB3	0	0,024324917	-1,62		
THBS1	0	4,40E-06	-0,73		
VCAM1	1	7,00E-05	1,71		
VEGFC	1	2,26E-08	1,18	1	11824968
CCL3	1	3,64E-18	6,85	1	32211348
CD28				1	12823856
EP300				1	31178673
GATA3				0	15981085
ICOS				1	11983910
IRF7				1	34432649
LAG3	1	1,46E-11	4,8		
NFAT				1	30538703
RBPJ				1	26604133
TCR				1	32115259
VCAM1	1	2,97E-15	7,01		

Supplementary table 10. *Calibrated state of the RA multicellular model.*

Nodes	Mean values
FOS	1
CDKN1A	0
CSF1R	1
CSK	1
CTF1	1
CYCS	0
CYLD	1
DUSP7	1
Dvl1	1
EDN1	1
EFNB1	1

EP300	1
FOXO3	0
GAB1	1
Grb2	1
HBEGF	1
IL1R1	1
IL4R	1
jag1	1
KRAS	1
LY96	1
map3k5	1
MCL1	1
mcl1	0
TIRAP	1
NRAS	0
PDE4B	1
PLCG1	1
PLXNB1	1
POU2F1	1
PRKCD	1
RPS6KA4	1
SOCS3	1
STAT6	1
TRAF2	1
vegfa	1
YWHAQ	0
ACKR3_CXCL12_complex	1
ACVR2A_ACVR2B_complex	0
ACVR2A_ACVR2B_INHBA_complex	0
ACVR2A_ACVR2B_INHBB_complex	0
ADAMTS4	1
ADAMTS9_rna	0
ADCY8	0
AGT	1
AGT_F2R_complex	1
AKT1_M1_macrophage___Cytoplasm	0
AKT1_M2_macrophage_nucleus	0
AKT1_phosphorylated	0
AKT1_TH1___cytoplasm	1

AKT2	1
AKT2_phosphorylated	1
AMAP1	1
angiogenesis_signal_phenotype	1
AP_1	1
AP_1_phosphorylated	1
APAF1	1
apoptosis_fibroblast_phenotype	0
apoptosis_M1_macrophage_phenotype	0
apoptosis_M2_macrophage_phenotype	1
apoptosis_TH1_phenotype	0
Apoptosome_complex	0
AREG_EGFR_complex	1
AREG_Fibroblast___Extracellular_Space	1
AREG_M1_macrophage___Secreted_components	1
ARHGEF12	1
ARHGEF2_Fibroblast___Cytoplasm	1
ARHGEF2_TH1___cytoplasm	1
ARI_ARII_bmp6_complex	0
ARRB2_Fibroblast___Cytoplasm	1
ARRB2_M1_macrophage___Cytoplasm	1
ARRB2_M2_macrophage__cytoplasm	1
ASC_M1_macrophage___Cytoplasm	1
ASC_M2_macrophage__cytoplasm	1
ASK1_M1_macrophage___Cytoplasm	1
ASK1_M2_macrophage__cytoplasm	1
BAD_Fibroblast___Mitochondrion	0
BAD_M1_macrophage__Mitochondria_membrane	0
BAD_M2_macrophage_mitochondrion_membrane	0
BAK1	0
BAX_Fibroblast___Mitochondrion	0
BAX_Fibroblast___Mitochondrion_active	0
BAX_M1_macrophage__Mitochondria_membrane	1
BAX_M2_macrophage_mitochondrion_membrane	1
BCL2_Fibroblast___Mitochondrion	1
BCL2_M1_macrophage__Mitochondria_membrane	1
BCL2_M1_macrophage__Mitochondria_membrane_active	0
BCL2_M2_macrophage_mitochondrion_membrane	1

BCL2_M2_macrophage_mitochondrion_membrane_active	0
Bcl2_rna_M1_macrophage___nucleus	1
Bcl2_rna_M2_macrophage_nucleus	1
BCL2A1_rna	1
BCL2L1_M1_macrophage__Mitochondria	1
BCL2L1_M1_macrophage__Mitochondria_active	0
BCL2L1_M2_macrophage__mitochondria	0
BCL2L1_rna	1
BCL2L11_Fibroblast___Cytoplasm	0
BCL2L11_Fibroblast___Cytoplasm_active	0
BCL2L11_TH1___cytoplasm	0
BCL3_rna_M1_macrophage___nucleus	1
BCL3_rna_M2_macrophage_nucleus	1
BID	0
biglycan_simple_molecule	1
Bim_M1_macrophage__Mitochondria	1
Bim_M2_macrophage__mitochondria	1
bmp6	1
BRAF_phosphorylated	0
BSG	1
BTK_rna	1
BTRC_rna	1
C_EBPb_phosphorylated	1
c_FOS_M1_macrophage___Cytoplasm	1
c_FOS_M1_macrophage___Cytoplasm_active	1
c_FOS_M1_macrophage___nucleus	1
c_JUN	1
c_JUN_phosphorylated_M1_macrophage___Cytoplasm	1
c_JUN_phosphorylated_M1_macrophage___nucleus	1
c_Myc_rna_M1_macrophage___nucleus	1
c_Myc_rna_M2_macrophage_nucleus	1
C5a_C5aR1_complex_M1_macrophage__Cytoplasmic_membrane_up	1
C5a_C5aR1_complex_M2_macrophage__cytoplasmic_membrane_up	1
c5a_M1_macrophage___Extracellular_Space	1
c5a_M2_macrophage__extracellular_space	1
CALCINEURIN	1

Casp1_M1_macrophage___Cytoplasm	1
Casp1_M2_macrophage__cytoplasm	1
CASP3_Fibroblast___Cytoplasm	0
CASP3_M1_macrophage___Cytoplasm	0
CASP3_M2_macrophage__cytoplasm	1
CASP7_M1_macrophage___Cytoplasm	0
CASP7_M2_macrophage__cytoplasm	1
CASP8_Fibroblast___Cytoplasm	0
CASP8_M1_macrophage___Cytoplasm	1
CASP8_M2_macrophage__cytoplasm	1
CASP9_Fibroblast___Cytoplasm	1
CASP9_M1_macrophage___Cytoplasm	0
CASP9_M2_macrophage__cytoplasm	0
CAV1_rna	1
CCL18_Fibroblast___Extracellular_Space	1
CCL18_M2_macrophage__secreted_components	1
CCL2_CCR2_complex_M1_macrophage__Cytoplas-mic_membrane_up	1
CCL2_CCR2_complex_TH1___cytoplasmic_membrane_up	1
CCL2_Fibroblast___Extracellular_Space	1
CCL2_Fibroblast___secreted_components	1
CCL2_M1_macrophage___Extracellular_Space	1
CCL2_M1_macrophage___Secreted_components	1
CCL2_TH1___extracellular_space	1
CCL20_M1_macrophage___Secreted_components	1
CCL20_M2_macrophage__secreted_components	1
CCL21	1
CCL21_CCR7_complex_M1_macrophage__Cytoplas-mic_membrane_up	1
CCL21_CCR7_complex_M2_macrophage__cytoplas-mic_membrane_up	1
CCl21_M1_macrophage___Extracellular_Space	1
CCl21_M2_macrophage__extracellular_space	1
CCL3_M1_macrophage___Secreted_components	1
CCL3_TH1___secreted_components	1
CCL4_5_CCR5_complex	1
CCL4_M1_macrophage___Secreted_components	1
CCL4_TH1___extracellular_space	1
CCL4_TH1___secreted_components	1

CCL5_ACKR2_complex	1
CCL5_CCR5_complex	1
CCL5_Fibroblast___Extracellular_Space	1
CCL5_Fibroblast___secreted_components	1
CCL5_M1_macrophage___Secreted_components	1
CCL5_TH1___extracellular_space	1
CCL5_TH1___secreted_components	1
CCNA2	1
CCNB1_CDC2_complex	1
CCND1_rna	1
CCND2	1
CCR2_CCL2_complex	1
CCR5	1
CD28	1
CD28_CD86_complex	1
CD32a	1
CD32b_igG_complex	0
CD32b_rna	1
CD40_Fibroblast__Cytoplasmic_membrane_up	1
CD40_M1_macrophage__Cytoplasmic_membrane_down	1
CD40LG_CD40_complex	1
CD40LG_ITGAM_ITGB2_complex	1
CD40LG_ITGB1_ITGA1_complex_M1_macrophage__Cyto-plasmic_membrane_up	1
CD40LG_ITGB1_ITGA1_complex_M2_macrophage__cyto-plasmic_membrane_up	1
CD40LG_M1_macrophage__Cytoplasmic_mem-brane_down	1
CD40LG_M2_macrophage__cytoplasmic_mem-brane_down	1
CD40LG_TH1___cytoplasmic_membrane_down	1
CD80	1
CD84_CD84_complex	1
CD84_Fibroblast__Cytoplasmic_membrane_down	1
CD84_M1_macrophage__Cytoplasmic_membrane_down	1
CD84_M2_macrophage__cytoplasmic_membrane_down	1
CD84_TH1___cytoplasmic_membrane_up	1
CD86_M1_macrophage__Cytoplasmic_membrane_down	1
CD86_M2_macrophage__cytoplasmic_membrane_down	1

CDC25B_C	0
Cell_chemotaxis_migration_fibroblast_phenotype	1
Cell_chemotaxis_migration_TH1_phenotype	1
CFLAR_Fibroblast___Cytoplasm	1
CFLAR_Fibroblast___Cytoplasm_active	1
cFLIP_M1_macrophage___Cytoplasm	1
cFLIP_M1_macrophage___Cytoplasm_active	1
cFLIP_M2_macrophage__cytoplasm	1
cFLIP_M2_macrophage__cytoplasm_active	1
cGAS	1
CHUK_phosphorylated	1
cIAP1	1
cMyc	1
cMyc_phosphorylated_M2_macrophage__cytoplasm	1
cMyc_phosphorylated_M2_macrophage_nucleus	1
col4a3_Fibroblast___Extracellular_Space	1
COL4A3_ITGA4_ITGB7_complex	1
col4a3_M2_macrophage__secreted_components	1
col4a3_TH1___extracellular_space	1
col4a4_Fibroblast___secreted_components	1
col4a4_M1_macrophage___Extracellular_Space	1
col4a4_M2_macrophage__extracellular_space	1
COL4A5	1
col4a5_Fibroblast___Extracellular_Space	1
col4a5_M1_macrophage___Extracellular_Space	1
col4a5_M2_macrophage__extracellular_space	1
COMP	1
COX2_rna_M1_macrophage___nucleus	1
COX2_rna_M2_macrophage_nucleus	1
CPNE3_rna	1
CREB1_phosphorylated_Fibroblast___nucleus	1
CREB1_phosphorylated_M2_macrophage_nucleus	0
CREB1_phosphorylated_TH1___nucleus	1
CRKL_phosphorylated_Fibroblast___Cytoplasm	1
CRKL_phosphorylated_M1_macrophage___Cytoplasm	1
CSF1_Fibroblast___secreted_components	1
CSF1_M2_macrophage__extracellular_space	1
CSF1_M2_macrophage__secreted_components	1
CSF1_rna	1

CSF1R_CSF1_complex	1
CSF2_Fibroblast___Extracellular_Space	1
CSF2_Fibroblast___secreted_components	1
CSF2_M1_macrophage___Extracellular_Space	1
CSF2_M1_macrophage___Secreted_components	1
CSF2RA_CSF2RB_complex_Fibroblast__Cytoplas-mic_membrane_up	1
CSF2RA_CSF2RB_complex_M1_macrophage__Cytoplas-mic_membrane_up	1
CSF2RA_CSF2RB_CSF2_complex	1
CSFR1R_IL34_complex	1
CSL	1
CTNNB_CK1A_AXIN_GSK3B_APC_complex	0
CTNNB1_Fibroblast___nucleus	0
CTNNB1_M2_macrophage_nucleus	1
CTNNB1_TCF_LEF_complex	1
CTSK_rna	1
CXCL1_CXCR1_complex	1
CXCL1_Fibroblast___secreted_components	1
CXCL1_M1_macrophage___Extracellular_Space	1
CXCL1_M1_macrophage___Secreted_components	1
CXCL10_CXCR3_complex_Fibroblast__Cytoplasmic_mem-brane_up	1
CXCL10_CXCR3_complex_TH1___cytoplasmic_mem-brane_up	1
CXCL10_Fibroblast___Extracellular_Space	1
CXCL10_Fibroblast___secreted_components	1
CXCL10_M1_macrophage___Secreted_components	1
CXCL10_TH1___extracellular_space	1
CXCL10_TH1___secreted_components	1
CXCL11	1
CXCL12	1
CXCL13_ACKR4_complex_M1_macrophage__Cytoplas-mic_membrane_up	1
CXCL13_ACKR4_complex_M2_macrophage__cytoplas-mic_membrane_up	1
CXCL13_CXCR3_complex	1
CXCL13_Fibroblast___Extracellular_Space	1
CXCL13_M1_macrophage___Extracellular_Space	1
CXCL13_M2_macrophage__extracellular_space	1

CXCL13_TH1___secreted_components	1
CXCL16_CXCR6_complex	1
CXCL16_M1_macrophage___Secreted_components	1
CXCL16_M2_macrophage__secreted_components	1
CXCL16_TH1___extracellular_space	1
CXCL2	1
CXCL3	1
CXCL8_Fibroblast___Extracellular_Space	1
CXCL8_Fibroblast___secreted_components	1
CXCL9_Fibroblast___secreted_components	1
CXCL9_M1_macrophage___Secreted_components	1
CXCR1	1
CXCR1_IL8_complex_M1_macrophage__Cytoplas-mic_membrane_up	1
CXCR1_IL8_complex_M2_macrophage__cytoplas-mic_membrane_up	1
CXCR2_CXCL8_complex	1
CXCR3	1
CypA	1
DAG_simple_molecule	1
DAXX_Fibroblast___Cytoplasm	1
DAXX_M1_macrophage___Cytoplasm	1
DAXX_M2_macrophage__cytoplasm	1
DAXX_TH1___cytoplasm	1
DKK1	1
DNA_simple_molecule	1
DOCK2	1
DOCK2_CRKL_complex	1
dsRNA_simple_molecule	1
DUSP1_M1_macrophage___Cytoplasm	1
DUSP1_M2_macrophage__cytoplasm	1
DVL1_phosphorylated	1
DYNLRB1	1
ECSIT	1
EDA_EDA2R_complex	1
EDA_Fibroblast___Extracellular_Space	1
EDA_M1_macrophage___Secreted_components	1
EDA_M2_macrophage__secreted_components	1
EFNB1_EPHB1_complex	1

efnb1_rna	1
EGF	1
EGF_EGFR_complex	1
EGFR	1
ELK1_phosphorylated	1
EOMES	1
EPHB2	1
ERK1_phosphorylated_M1_macrophage___Cytoplasm	1
ERK1_phosphorylated_M1_macrophage___nucleus	1
ERK1_phosphorylated_M2_macrophage__cytoplasm	1
ERK1_phosphorylated_M2_macrophage_nucleus	1
FADD_Fibroblast___Cytoplasm	1
FADD_M1_macrophage___Cytoplasm	1
FADD_M2_macrophage__cytoplasm	1
FAS_FASL_complex	1
FAS_M1_macrophage__Cytoplasmic_membrane_up	1
FAS_M2_macrophage__cytoplasmic_membrane_up	1
FASL_FAS_complex_M1_macrophage__Cytoplas-mic_membrane_up	1
FASL_FAS_complex_M2_macrophage__cytoplasmic_mem-brane_up	1
FASL_M1_macrophage___Extracellular_Space	1
FASL_M1_macrophage___Secreted_components	1
FASL_M2_macrophage__extracellular_space	1
FASL_M2_macrophage__secreted_components	1
FASLG_FAS_complex	1
FASLG_Fibroblast___Extracellular_Space	1
FASLG_TH1___extracellular_space	1
FASLG_TH1___secreted_components	1
FGF1	1
FGF1_FGFR4_complex	1
FGR	1
FN1_Fibroblast___Extracellular_Space	0
FN1_Fibroblast___secreted_components	0
FN1_ITGAV_complex	0
FN1_M2_macrophage__secreted_components	1
FOXO_M1_macrophage___Cytoplasm	1
FOXO_M1_macrophage___nucleus	1
FOXO1_Fibroblast___nucleus	0

FOXO1_M2_macrophage_nucleus	1
FZD1_LRP5_complex	1
FZD5_LRP5_wnt5a_complex	1
GAB2_phosphorylated_Fibroblast___Cytoplasm	1
GAB2_phosphorylated_M2_macrophage__cytoplasm	1
gal_Fibroblast___secreted_components	1
GAL_GALR2_complex_M1_macrophage__Cytoplas-mic_membrane_up	1
GAL_GALR2_complex_M2_macrophage__cytoplas-mic_membrane_up	1
gal_M1_macrophage___Extracellular_Space	1
gal_M2_macrophage__extracellular_space	1
GAL_rna	1
gamma_secretase_complex	1
gamma_secretase_complex_complex	1
GAS6_Fibroblast___secreted_components	1
GAS6_M2_macrophage__extracellular_space	1
GAS6_MERTK_complex	1
gas6_mertk_complex	1
GAS6_TH1___extracellular_space	1
gata3_rna	0
GNA12_GNA13_complex_M1_macrophage___Cytoplasm	1
GNA12_GNA13_complex_M2_macrophage__cytoplasm	1
GNA12_GNA13_complex_TH1___cytoplasm	1
GNAI_GNB_GNG_complex	1
GNAI_TH1___cytoplasm	1
GNAI_TH1___cytoplasm_active	1
GNAI3_Fibroblast___Cytoplasm	1
GNAI3_M1_macrophage___Cytoplasm	1
GNAI3_M2_macrophage__cytoplasm	1
GNB_GNG_complex_M1_macrophage___Cytoplasm	1
GNB_GNG_complex_M2_macrophage__cytoplasm	1
GNB_GNG_complex_TH1___cytoplasm	1
GNB_GNG_GNAI3_complex_M1_macrophage___Cyto-plasm	1
GNB_GNG_GNAI3_complex_M2_macrophage__cyto-plasm	1
GRB2_Fibroblast___Cytoplasm	1
GRB2_TH1___cytoplasm	1

GSK3B	0
GSK3b_APC_AXIN1_CK1A_complex	1
GSK3b_APC_AXIN1_CTNNB1_CK1A_complex	1
HBGEF_EGFR_complex	1
HBGEF_M1_macrophage___Secreted_components	1
HBGEF_M2_macrophage__secreted_components	1
HCK	1
HES1_rna	1
hes1_rna	1
HEY1_rna	1
HIF1_complex	1
Hif1a	1
Hif1b	1
HLA_B	1
HLA_B_LILRB1_complex_M1_macrophage__Cytoplasmic_membrane_up	1
HLA_B_LILRB1_complex_M2_macrophage__cytoplasmic_membrane_up	1
HLA_DP_DQ_DR_LAG3_complex	1
HLA_DP_DQ_DR_M1_macrophage__Cytoplasmic_membrane_down	1
HLA_DP_DQ_DR_M2_macrophage__cytoplasmic_membrane_down	1
HLA_DP_DQ_DR_TCR_CD3_complex	1
HLA_DRB1_rna	1
HMGB1	1
HOMODIMER_space_STAT1	1
HOMODIMER_space_STAT3	1
HRAS_M1_macrophage___Cytoplasm	1
HRAS_M2_macrophage__cytoplasm	1
HSPA5	1
ICAM1_Fibroblast__Cytoplasmic_membrane_down	1
ICAM1_ITGB2_ITGAL_complex	1
ICAM1_M1_macrophage__Cytoplasmic_membrane_down	1
ICAM1_TH1___cytoplasmic_membrane_down	1
icos	1
ICOSLG_Fibroblast__Cytoplasmic_membrane_down	1
icoslg_icos_complex	1

ICOSLG_M1_macrophage__Cytoplasmic_membrane_down	1
ICOSLG_M2_macrophage__cytoplasmic_membrane_down	1
IFNa_M1_macrophage___Extracellular_Space	1
IFNa_M1_macrophage___Secreted_components	1
IFNa_rna	1
IFNa1	1
IFNa1_B1	1
IFNA1_B1_IFNAR1_R2_complex	1
IFNAR1_IFNAR2_complex	1
IFNAR1_IFNAR2_IFNa_complex	1
IFNAR1_IFNAR2_IFNb_complex	1
IFNb_M1_macrophage___Extracellular_Space	1
IFNb_M1_macrophage___Secreted_components	1
IFNb_rna	1
IFNb1	1
IFNB1_rna	1
IFNE	1
IFNE_IFNAR1_IFNAR2_complex	1
IFNg_Fibroblast___Extracellular_Space	1
IFNG_IFNGR1_IFNGR2_complex	1
IFNG_IFNGR1_R2_complex	1
IFNg_M1_macrophage___Extracellular_Space	1
IFNg_M1_macrophage___Secreted_components	1
IFNg_rna	1
IFNg_TH1___extracellular_space	1
IFNg_TH1___secreted_components	1
IFNGR1_IFNGR2_complex_M1_macrophage__Cytoplasmic_membrane_up	1
IFNGR1_IFNGR2_complex_TH1___cytoplasmic_membrane_up	1
IFNGR1_IFNGR2_IFNg_complex	1
IFNGR1_R2_complex	1
IGF1	0
IGF1_IGF1R_complex	0
igG	0
IKBA_NFKB1_RELA_complex	1
IKBKB_phosphorylated	1

IKK_complex_Fibroblast___Cytoplasm	1
IKK_complex_M1_macrophage___Cytoplasm	1
IKK_complex_M2_macrophage__cytoplasm	1
IKK_complex_TH1___cytoplasm	1
IKK1_IKK2_complex_Fibroblast___Cytoplasm	1
IKK1_IKK2_complex_M1_macrophage___Cytoplasm	1
IKK1_IKK2_complex_M2_macrophage__cytoplasm	1
IKK1_phosphorylated_Fibroblast___Cytoplasm	1
IKK1_phosphorylated_M1_macrophage___Cytoplasm	1
IKK1_phosphorylated_M2_macrophage__cytoplasm	1
IKK2_phosphorylated_Fibroblast___Cytoplasm	1
IKK2_phosphorylated_M1_macrophage___Cytoplasm	1
IKK2_phosphorylated_M2_macrophage__cytoplasm	1
IKKE_TBK1_complex	1
IKKE_TBK1_TRAF3_complex	0
IL1_IL1R_complex	1
IL10_Fibroblast___Extracellular_Space	1
IL10_IL10RA_IL10RB_complex	1
IL10_M2_macrophage__extracellular_space	1
IL10_M2_macrophage__secreted_components	1
IL10_rna	1
IL10_TH1___extracellular_space	1
IL10R1_IL10R2_complex	1
IL10R1_IL10R2_IL10_complex	1
IL10RA_IL10RB_complex_Fibroblast__Cytoplasmic_mem-brane_up	1
IL10RA_IL10RB_complex_TH1___cytoplasmic_mem-brane_up	1
IL10RA_IL10RB_IL10_complex	1
IL11	1
IL11_IL11Ra_IL6ST_complex	1
IL11Ra_IL6ST_IL11_complex	1
IL12_M1_macrophage___Extracellular_Space	1
IL12_M1_macrophage___Secreted_components	1
IL12_rna	1
IL12_TH1___extracellular_space	1
IL12_TH1___secreted_components	1
IL12A	1
IL12B	1

IL12RB_complex	1
IL12RB_IL12_complex	1
IL12Rb1_IL12Rb2_complex	1
IL12Rb1_IL12Rb2_IL12_complex	1
IL15_M1_macrophage___Secreted_components	1
IL15_M2_macrophage__secreted_components	1
IL17A_Fibroblast___Extracellular_Space	1
IL17A_Fibroblast___secreted_components	1
IL17A_IL17RA_complex	1
IL17F	1
IL17Ra_IL17Rc_IL17F_complex	1
IL18_Fibroblast___Extracellular_Space	1
IL18_Fibroblast___secreted_components	1
IL18_IL18R_complex	1
IL18_IL18R1_complex	1
IL18_IL18R1_IL18RAP_complex	1
IL18_M1_macrophage___Cytoplasm	1
IL18_M1_macrophage___Extracellular_Space	1
IL18_M1_macrophage___Secreted_components	1
IL18_M2_macrophage__cytoplasm	1
IL18_M2_macrophage__cytoplasm_active	1
IL18_TH1___extracellular_space	1
IL18_TH1___secreted_components	1
IL18R1_IL18RAP_complex	1
IL1A	1
IL1A_IL1R1_complex	1
IL1B_Fibroblast___Extracellular_Space	1
IL1B_Fibroblast___secreted_components	1
IL1B_IL1R1_complex	1
IL1B_M1_macrophage___Cytoplasm	1
IL1B_M1_macrophage___Extracellular_Space	1
IL1B_M1_macrophage___Secreted_components	1
IL1B_M2_macrophage__cytoplasm	1
IL1RN_rna	1
IL2	1
IL23_M1_macrophage___Extracellular_Space	1
IL23_M1_macrophage___Secreted_components	1
IL23_M2_macrophage__extracellular_space	1
IL23_M2_macrophage__secreted_components	1

IL23R_IL12RB1_complex_M1_macrophage__Cytoplasmic_membrane_up	1
IL23R_IL12RB1_complex_M2_macrophage__cytoplasmic_membrane_up	1
IL23R_IL12RB1_IL23_complex_M1_macrophage__Cytoplasmic_membrane_up	1
IL23R_IL12RB1_IL23_complex_M2_macrophage__cytoplasmic_membrane_up	1
IL26	1
IL27	1
IL27_IL27RA_complex	1
IL32_rna	1
IL33_rna	1
IL34_Fibroblast___secreted_components	1
IL34_M2_macrophage__extracellular_space	1
IL4	1
IL4_IL4Ra_complex	1
IL4_rna	1
IL6_Fibroblast___Extracellular_Space	1
IL6_Fibroblast___secreted_components	1
IL6_IL6R_complex	1
IL6_IL6R_IL6ST_complex	1
IL6_IL6ST_complex	1
IL6_M1_macrophage___Extracellular_Space	1
IL6_M1_macrophage___Secreted_components	1
IL7_Fibroblast___secreted_components	1
IL7_IL7R_IL2RG_complex	1
IL7_TH1___extracellular_space	1
IL7R_IL2RG_complex	1
IL8_M1_macrophage___Extracellular_Space	1
IL8_M1_macrophage___Secreted_components	1
IL8_M2_macrophage__extracellular_space	1
IL8_M2_macrophage__secreted_components	1
IL9R_IL2RG_IL9_complex	1
immune_complex_CD16a_complex	1
immune_complex_CD32a_complex	1
immune_complex_CD32B_complex	1
immune_complex_CD64_complex	1
immune_complex_complex	1

inflammation_signal_phenotype	1
INHBA	1
INHBB_Fibroblast___secreted_components	1
INHBB_M1_macrophage___Extracellular_Space	1
INPP5A	1
IP3_simple_molecule	1
IRAK1_Fibroblast___Cytoplasm	1
IRAK1_Fibroblast___Cytoplasm_active	1
IRAK1_IRAK4_complex	1
IRAK1_IRAK4_complex_M1_macrophage___Cytoplasm	1
IRAK1_IRAK4_complex_TH1___cytoplasm	1
IRAK1_M1_macrophage___Cytoplasm	1
IRAK3	1
IRAK4_Fibroblast___Cytoplasm	1
IRAK4_Fibroblast___Cytoplasm_active	1
IRAK4_phosphorylated	1
IRF1_Fibroblast___Cytoplasm	1
IRF1_Fibroblast___nucleus	1
IRF3_phosphorylated_Fibroblast___nucleus	1
IRF3_phosphorylated_M1_macrophage___nucleus	1
IRF5_Fibroblast___Cytoplasm	1
IRF5_Fibroblast___nucleus	1
IRF5_ubiquitinated	1
IRF7_M1_macrophage___nucleus	1
IRF7_M2_macrophage_nucleus	1
IRF7_TH1___cytoplasm	1
IRF7_TH1___cytoplasm_active	1
IRF9_Fibroblast___Cytoplasm	1
IRF9_M1_macrophage___Cytoplasm	1
ISGF3_complex	1
ITCH_phosphorylated_M1_macrophage___Cytoplasm	0
ITCH_phosphorylated_M2_macrophage__cytoplasm	0
ITGA4_ITGB1_col4a3_complex	1
ITGA4_ITGB1_complex	1
ITGA4_ITGB1_sema7a_complex	1
ITGA4_ITGB7_complex	1
ITGA5_ITGB1_col4a5_complex	1
ITGA5_ITGB1_complex	1
ITGAL_ITGB2_complex	1

ITGAM_ITGB2_complex	1
ITGB1_ITGA1_col4a_complex_M1_macrophage__Cyto-plasmic_membrane_up	1
ITGB1_ITGA1_col4a_complex_M2_macrophage__cytoplas-mic_membrane_up	1
ITGB1_ITGA1_complex_M1_macrophage__Cytoplas-mic_membrane_up	1
ITGB1_ITGA1_complex_M2_macrophage__cytoplas-mic_membrane_up	1
JAG1_M1_macrophage__Cytoplasmic_membrane_down	1
JAG1_M2_macrophage__cytoplasmic_membrane_down	1
JAG1_NOTCH1_complex	1
JAG1_NOTCH3_complex	1
JAG1_rna	1
JAK1_Fibroblast___Cytoplasm	0
JAK1_JAK2_complex_M1_macrophage___Cytoplasm	1
JAK1_JAK2_complex_TH1___cytoplasm	1
JAK1_JAK2_TYK2_complex	1
JAK1_JAK3_complex	1
JAK1_M1_macrophage___Cytoplasm	1
JAK1_M2_macrophage__cytoplasm	1
JAK1_TYK2_complex_M1_macrophage___Cytoplasm	1
JAK1_TYK2_complex_M2_macrophage__cytoplasm	1
JAK1_TYK2_complex_TH1___cytoplasm	1
JAK2_Fibroblast___Cytoplasm	0
JAK2_M1_macrophage___Cytoplasm	1
JAK2_M2_macrophage__cytoplasm	1
JAK2_TYK2_complex_M1_macrophage___Cytoplasm	1
JAK2_TYK2_complex_TH1___cytoplasm	1
JAK3	0
JNK1_phosphorylated_M1_macrophage___Cytoplasm	1
JNK1_phosphorylated_M1_macrophage___nucleus	1
JNK1_phosphorylated_M2_macrophage_nucleus	1
JUN_phosphorylated	1
JUNB	1
klf4	1
LAG3	1
LAT_phosphorylated	0
LBP_CD14_complex	1

LCK_phosphorylated	0
LGALS9_Fibroblast___secreted_components	1
LGALS9_M1_macrophage___Secreted_components	1
LGALS9_TH1___extracellular_space	1
LGALS9_TIM3_complex	0
LIFR_IL6ST_CTF1_complex	1
LTA_Fibroblast___secreted_components	1
LTA_M1_macrophage___Secreted_components	1
LTA_TH1___extracellular_space	1
LTA_TH1___secreted_components	1
LTA_TNFRSF14_complex	1
LTBP1	0
LY96_TLR2_4_complex	1
lyn	1
maf_rna	0
MAML1	1
MAP2K1_phosphorylated_Fibroblast___Cytoplasm	1
MAP2K1_phosphorylated_TH1___cytoplasm	1
MAP2K3_phosphorylated	1
MAP2K4_phosphorylated	1
MAP2K6_phosphorylated	1
MAP2K7_phosphorylated	1
MAP3K1_phosphorylated	1
MAP3K14_phosphorylated	1
MAP3K2_3_4	1
MAP3K5_phosphorylated	1
MAP3K7_phosphorylated	1
MAP3K7_rna	1
MAP3K8_phosphorylated	1
MAP4K4_phosphorylated	1
MAPK1_empty	0
MAPK1_phosphorylated_Fibroblast___Cytoplasm	0
MAPK1_phosphorylated_TH1___cytoplasm	1
MAPK14_phosphorylated	1
MAPK3_complex	1
MAPK3_phosphorylated	1
MAPK8	1
MAPK8_phosphorylated	1
MAPK9_phosphorylated	1

MAPKAPK2_phosphorylated	1
Matrix_degradation_signal_phenotype	1
Mcl1_rna_M1_macrophage___nucleus	1
Mcl1_rna_M2_macrophage_nucleus	0
MDM2	1
MDM2_phosphorylated_Fibroblast___Cytoplasm	1
MDM2_phosphorylated_M1_macrophage___Cytoplasm	0
MDM2_phosphorylated_M2_macrophage__cytoplasm	0
MEK1_phosphorylated_M1_macrophage___Cytoplasm	1
MEK1_phosphorylated_M2_macrophage__cytoplasm	1
MEK2_phosphorylated_M1_macrophage___Cytoplasm	1
MEK2_phosphorylated_M2_macrophage__cytoplasm	1
MEKK1_M1_macrophage___Cytoplasm	1
MEKK1_M2_macrophage__cytoplasm	0
MIF_CD74_complex	1
MIF_CXCR4_complex	1
MIF_Fibroblast___Extracellular_Space	1
MIF_Fibroblast___secreted_components	1
MIF_M2_macrophage__secreted_components	1
MIF_TH1___extracellular_space	1
MIF_TH1___secreted_components	1
MIR10a_rna	0
MIR124A_rna	0
MIR146A_rna	0
MIR155_rna	1
MIR192_rna	0
MIR203A_rna	1
MIR221_rna	1
MIR338_5P_rna	1
MIR346_rna	1
MIR34A_rna	0
MIR451A_rna	0
MIR650_rna	0
MK2_phosphorylated_M1_macrophage___nucleus	1
MK2_phosphorylated_M2_macrophage_nucleus	1
MKK3_phosphorylated_M1_macrophage___Cytoplasm	1
MKK3_phosphorylated_M2_macrophage__cytoplasm	1
MKK4_phosphorylated_M1_macrophage___Cytoplasm	1
MKK4_phosphorylated_M2_macrophage__cytoplasm	1

MKK6_phosphorylated_M1_macrophage___Cytoplasm	1
MKK6_phosphorylated_M2_macrophage__cytoplasm	1
MKK7_phosphorylated_M1_macrophage___Cytoplasm	1
MKK7_phosphorylated_M2_macrophage__cytoplasm	1
MMP1	1
MMP13	1
MMP14_M1_macrophage___Secreted_components	1
MMP14_M2_macrophage__secreted_components	1
MMP3_Fibroblast___secreted_components	1
MMP3_M1_macrophage___Secreted_components	1
MMP3_M2_macrophage__secreted_components	1
MMP9_Fibroblast___secreted_components	1
MMP9_M1_macrophage___Secreted_components	1
MMP9_M2_macrophage__secreted_components	1
MSK1_phosphorylated	0
mtor	1
MYD88_Fibroblast___Cytoplasm	1
MYD88_M1_macrophage___Cytoplasm	1
MYD88_TH1___cytoplasm	1
MYD88_TIRAP_TOLLIP_complex	1
NCID_Fibroblast___nucleus	1
NCID_M1_macrophage___Cytoplasm	1
ncid_rbpj_snw1_complex	1
NECTIN3_PTPRC_NECTIN1_complex	1
NFAT	1
NFAT5	1
NFAT5_phosphorylated_Fibroblast___nucleus	1
NFAT5_phosphorylated_M2_macrophage_nucleus	1
NFKB_complex	1
NFKB_N_complex	1
NFKB1_Fibroblast___Cytoplasm	1
NFKB1_M1_macrophage___Cytoplasm	1
NFKB1_M2_macrophage__cytoplasm	1
NFKB1_MAP3K8_complex	1
NFKB1_RELA_complex_TH1___cytoplasm	1
NFKB1_RELA_complex_TH1___nucleus	1
NFKB1_RELA_NFKBIA_complex	1
NFKB1_TPL2_complex_M1_macrophage___Cytoplasm	1
NFKB1_TPL2_complex_M2_macrophage__cytoplasm	1

NFKB2_RELB_complex	1
NFKBIA_phosphorylated_Fibroblast___Cytoplasm	1
NFKBIA_phosphorylated_M1_macrophage___Cytoplasm	1
NFKBIA_phosphorylated_M2_macrophage__cytoplasm	1
NFKBIA_phosphorylated_TH1___cytoplasm	1
NFKBIA_RELA_NFKB1_complex_M1_macrophage___Cyto-plasm	1
NFKBIA_RELA_NFKB1_complex_M2_macrophage__cyto-plasm	1
NFKBIA_rna	1
NFKBIE_Fibroblast___Cytoplasm	1
NFKBIE_M1_macrophage___Cytoplasm	1
NFKBIE_M2_macrophage__cytoplasm	1
ngef	1
NICD_CSL_SKIP_MAML1_ep300_complex	1
NICD_EP300_SKIP_CSL_MAML1_complex	1
NICD_M1_macrophage___nucleus	1
NICD_TH1___nucleus	1
NIK_M1_macrophage___Cytoplasm	1
NIK_TH1___cytoplasm	1
NLK_phosphorylated	1
NLRP3_INFLAMMASOME_complex_M1_macro-phage___Cytoplasm	1
NLRP3_INFLAMMASOME_complex_M2_macrophage__cy-toplasm	1
NLRP3_M1_macrophage___Cytoplasm	1
NLRP3_M2_macrophage__cytoplasm	1
NOD2	0
NOS2	1
NOS2_phosphorylated	0
NOTCH1	1
notch1_JAG1_complex	1
notch3	1
OPN_M1_macrophage___Cytoplasm	0
OPN_M2_macrophage__cytoplasm	0
osteoclastogenesis_M1_macrophage_phenotype	1
osteoclastogenesis_signal_phenotype	1
p15_rna_M1_macrophage___nucleus	0
p15_rna_M2_macrophage_nucleus	1

p21_rna	1
p300_SP1_complex_M1_macrophage___nucleus	0
p300_SP1_complex_M2_macrophage_nucleus	0
p38_MAP_KINASE_phosphorylated_M1_macrophage___Cytoplasm	0
p38_MAP_KINASE_phosphorylated_M1_macrophage___nucleus	0
p38_MAP_KINASE_phosphorylated_M2_macrophage__cytoplasm	0
p38MAPK_empty	1
p38MAPK_phosphorylated	0
p53	1
p53_phosphorylated_M1_macrophage___Cytoplasm	1
p53_phosphorylated_M2_macrophage__cytoplasm	1
p53_phosphorylated_M2_macrophage_nucleus	1
PDCD1	1
PDGFC_Fibroblast___Extracellular_Space	1
PDGFC_M2_macrophage__secreted_components	1
PDGFC_PDGFRB_complex	1
PDIA3_HLA_A_B2M_complex	1
PI3_4_5_P__3_simple_molecule	0
PI3K_M1_macrophage___Cytoplasm	0
PI3K_M2_macrophage__cytoplasm	0
PI3K_phosphorylated	1
PI4_5_P__2_simple_molecule	1
PIK3AP1_phosphorylated_M1_macrophage___Cytoplasm	0
PIK3AP1_phosphorylated_M2_macrophage__cytoplasm	0
PIK3R5_phosphorylated	1
PIP2_simple_molecule	1
PITPNM3_CCL18_complex	1
PLA2G2A_phosphorylated	1
PLCG2	1
PLXNA1	1
PLXNA1_sema3A_complex	1
PLXNB1_MET_complex	1
PLXNB1_MET_sema4a_complex	1
PLXNB2_MET_complex	1
PMAIP1_Fibroblast___Cytoplasm	1
PMAIP1_Fibroblast___Cytoplasm_active	1

PP2A	0
PP4	1
PPARg_rna	1
PRKACA	1
Prkcd_M1_macrophage___Cytoplasm	1
Prkcd_M2_macrophage__cytoplasm	1
PRKCQ_Fibroblast___Cytoplasm	1
PRKCQ_M1_macrophage___Cytoplasm	1
PRKCQ_M2_macrophage__cytoplasm	1
PRKCQ_TH1___cytoplasm	1
PRKG1_Fibroblast___Cytoplasm	1
PRKG1_M1_macrophage___Cytoplasm	1
PRKG1_M2_macrophage__cytoplasm	1
PRL_Fibroblast___secreted_components	1
PRL_M2_macrophage__extracellular_space	1
PRL_PRLR_complex	1
proliferation_survival_fibroblast_phenotype	1
proliferation_survival_M1_macrophage_phenotype	1
proliferation_survival_M2_macrophage_phenotype	0
proliferation_survival_TH1_phenotype	1
PTEN	1
PTGS2_rna	1
PTK2_M1_macrophage___Cytoplasm	1
PTK2_M2_macrophage__cytoplasm	1
PTK2_TH1___cytoplasm	1
PTK2B_phosphorylated	1
PTPN11_phosphorylated	1
PTPN6_Fibroblast___Cytoplasm	1
PTPN6_M1_macrophage___Cytoplasm	1
PTPN6_M2_macrophage__cytoplasm	1
pyk2_phosphorylated	1
RAB5A	1
RAC1_2	1
Rac1_M1_macrophage___Cytoplasm	1
Rac1_M2_macrophage__cytoplasm	1
Rac1_TH1___cytoplasm	0
RAF1_Fibroblast___Cytoplasm	1
RAF1_M1_macrophage___Cytoplasm	1
RAF1_M2_macrophage__cytoplasm	1

RAF1_TH1___cytoplasm	1
RANK_RANKL_complex	1
rasa1	1
RBL1_E2F4_DP1_complex	0
rbpj	1
RDX	1
RELA_NFKB1_complex_M1_macrophage___Cytoplasm	1
RELA_NFKB1_complex_M1_macrophage___nucleus	1
RELA_NFKB1_complex_M2_macrophage__cytoplasm	1
RELA_NFKB1_complex_M2_macrophage_nucleus	1
RELA_NFKB1_NFKBIE_complex_Fibroblast___Cytoplasm	1
RELA_NFKB1_NFKBIE_complex_M1_macrophage___Cyto-plasm	1
RELA_NFKB1_NFKBIE_complex_M2_macrophage__cyto-plasm	1
RHOA_Fibroblast___Cytoplasm	0
RHOA_M1_macrophage___Cytoplasm	0
RHOA_M2_macrophage__cytoplasm	0
RHOA_TH1___cytoplasm	1
RIPK1	1
RIPK1_TRAF6_complex	1
RIPK3	0
RPS6KB1_phosphorylated	1
RUNX1	0
RUNX3	1
RXRa_NUR77_complex_M1_macrophage___Cytoplasm	1
RXRa_NUR77_complex_M2_macrophage__cytoplasm	1
SARA_M1_macrophage___Cytoplasm	0
SARA_M2_macrophage__cytoplasm	0
SARM1	0
sema3A	1
SEMA4_PLXNB1_complex	1
SEMA4_PLXNB2_MET_complex	1
SEMA4A_Fibroblast___Extracellular_Space	1
SEMA4A_M1_macrophage___Secreted_components	1
SEMA4A_M2_macrophage__extracellular_space	1
SEMA4A_M2_macrophage__secreted_components	1
SEMA4A_TH1___extracellular_space	1
SEMA4A_TH1___secreted_components	1

SEMA4D_Fibroblast___secreted_components	1
SEMA4D_M2_macrophage__extracellular_space	1
SEMA4D_TH1___extracellular_space	1
sema7a_Fibroblast___Extracellular_Space	1
sema7a_M2_macrophage__secreted_components	1
SERPINE1_rna	1
SH2D1A_Fibroblast___Cytoplasm	1
SH2D1A_M2_macrophage__cytoplasm	1
SH2D1A_TH1___cytoplasm	1
Shc_phosphorylated	1
SHC1_phosphorylated	1
SHC2_phosphorylated	1
SHIP1_M2_macrophage__cytoplasm	1
SHIP1_TH1___cytoplasm	0
SHP2_GRB2_complex	1
SHP2_GRB2_complex	1
Sirt1_M1_macrophage___nucleus	0
Sirt1_M2_macrophage__cytoplasm	0
SKIP	1
smad1_phosphorylated	0
SMAD2_phosphorylated_M1_macrophage___Cytoplasm	0
SMAD2_phosphorylated_M2_macrophage__cytoplasm	0
SMAD2_SARA_complex_M1_macrophage___Cytoplasm	0
SMAD2_SARA_complex_M2_macrophage__cytoplasm	0
SMAD2_SMAD4_complex_M1_macrophage___nucleus	0
SMAD2_SMAD4_complex_M2_macrophage_nucleus	0
SMAD4_M1_macrophage___Cytoplasm	0
SMAD4_M2_macrophage__cytoplasm	0
smad4_smad1_complex	0
SMAD7_Fibroblast___Cytoplasm	1
SMAD7_M1_macrophage___Cytoplasm	1
SMAD7_M2_macrophage__cytoplasm	1
snw1	1
SOCS3_rna	1
SOS1_Fibroblast___Cytoplasm	1
SOS1_M1_macrophage___Cytoplasm	1
SOS1_M2_macrophage__cytoplasm	1
SOS1_TH1___cytoplasm	1
Src_M1_macrophage___Cytoplasm	1

Src_M2_macrophage__cytoplasm	1
SRC_phosphorylated	1
Src_TH1___cytoplasm	1
sst	0
sst_sstr_complex	0
STAT1_Fibroblast___Cytoplasm	1
STAT1_Fibroblast___nucleus	1
STAT1_M1_macrophage___Cytoplasm	1
STAT1_STAT1_complex_M1_macrophage___nucleus	1
STAT1_STAT1_complex_TH1___nucleus	1
STAT1_STAT2_complex	1
STAT1_STAT2_IRF9_complex	1
STAT1_TH1___cytoplasm	1
STAT2_Fibroblast___Cytoplasm	1
STAT2_M1_macrophage___Cytoplasm	1
STAT3_Fibroblast___Cytoplasm	1
STAT3_Fibroblast___nucleus	1
STAT3_M1_macrophage___Cytoplasm	1
STAT3_M2_macrophage__cytoplasm	1
STAT3_STAT3_complex_M1_macrophage___nucleus	1
STAT3_STAT3_complex_M2_macrophage_nucleus	1
STAT4_M1_macrophage___Cytoplasm	1
STAT4_STAT4_complex_M1_macrophage___nucleus	1
STAT4_STAT4_complex_TH1___nucleus	1
STAT4_TH1___cytoplasm	1
STAT5_CRKL_complex	1
STAT5_phosphorylated	1
STAT6_STAT6_complex	1
SYK_M1_macrophage___Cytoplasm	0
SYK_M2_macrophage__cytoplasm	0
Syk_phosphorylated	1
TAB1	1
TAB1_TAB2_complex	1
TAB1_TAB2_TAK1_complex	1
TAB1_TAB2_TRAF6_complex	1
TAB2_phosphorylated	1
TAK1	1
TAK1_phosphorylated	1
tBID	0

TBK1	1
TBK1_IKBKE_complex	1
tbx21	1
tbx21_phosphorylated	1
TCF_LEF	1
TCF_LEF	1
TCR_CD3_complex	1
TGFb1	1
TGFB1_Fibroblast___Extracellular_Space	1
TGFB1_Fibroblast___Extracellular_Space_active	1
TGFB1_M2_macrophage__extracellular_space	1
TGFB1_M2_macrophage__secreted_components	1
TGFB1_TGFBR1_complex	1
TGFB1_TH1___secreted_components	1
tgfb3	0
TGFBR1_TGFBR2_complex	1
TGFBR1_TGFBR2_TGFB1_complex	1
THBS1_rna	0
TICAM1	1
TICAM1_TICAM2_complex	1
TICAM2	1
TIRAP_MYD88_complex	1
TLR1_TLR2_biglycan_complex	1
TLR1_TLR2_complex	1
TLR2_TLR6_biglycan_complex	1
TLR2_TLR6_complex	1
TLR3_dsRNA_complex	1
TLR4_Md2_CD14_fibrinogen_complex	0
TLR5_Fibroblast__Cytoplasmic_membrane_up	1
TLR5_M1_macrophage__Cytoplasmic_membrane_up	0
TLR7_TLR8_ssRNA_complex	1
TLR9_DNA_complex	1
TNF_Fibroblast___Extracellular_Space	1
TNF_Fibroblast___secreted_components	1
TNF_M1_macrophage___Extracellular_Space	1
TNF_M1_macrophage___Secreted_components	1
TNF_TH1___extracellular_space	1
TNF_TNFRSF1A_B_complex	1
TNF_TNFRSF1A_complex	1

TNF_TNFRSF1B_complex	1
TNFA	1
TNFA_rna	1
TNFAIP3_rna_M1_macrophage___nucleus	1
TNFAIP3_rna_M2_macrophage_nucleus	1
TNFRSF10A_rna	1
TNFRSF10B_rna	1
TNFSF11_Fibroblast___Extracellular_Space	1
TNFSF11_Fibroblast___secreted_components	1
TNFSF11_M1_macrophage___Extracellular_Space	1
TNFSF11_TH1___secreted_components	1
TNFSF11_TNFRSF11_complex	1
TNFSF4	1
TNFSF4_TNFRSF4_complex	1
TP53_phosphorylated	1
TP73_phosphorylated	0
TPL2_M1_macrophage___Cytoplasm	1
TPL2_M2_macrophage__cytoplasm	1
TRADD_Fibroblast___Cytoplasm	1
TRADD_M1_macrophage___Cytoplasm	1
TRADD_TRAF2_RIP1_complex	1
TRAF1_rna_Fibroblast___nucleus	1
TRAF1_rna_M1_macrophage___nucleus	1
TRAF1_rna_M2_macrophage_nucleus	1
TRAF1_TRAF2_TRAF3_complex	1
TRAF2_RIP1_TRADD_TAK1_TAB1_TAB2_complex	1
TRAF2_TRAF5_complex	1
TRAF2_TRAF5_complex	1
TRAF2_TRAF5_TRAF6_complex_TH1___cytoplasm	1
TRAF2_TRAF5_TRAF6_complex_TH1___cytoplasm_active	1
TRAF2_TRAF6_complex	1
TRAF3_Fibroblast___Cytoplasm	1
TRAF3_M1_macrophage___Cytoplasm	0
TRAF3_TH1___cytoplasm	1
TRAF3_TRAF6_complex	1
TRAF3IP2_phosphorylated_Fibroblast___Cytoplasm	1
TRAF3IP2_phosphorylated_M2_macrophage__cytoplasm	1
TRAF6_ECSIT_MEKK1_TAB1_TAB2_TAK1_complex	0
TRAF6_Fibroblast___Cytoplasm	1

TRAF6_IRAK1_IRAK4_complex	0
TRAF6_M1_macrophage___Cytoplasm	0
TRAF6_phosphorylated_Fibroblast___Cytoplasm	1
TRAF6_phosphorylated_TH1___cytoplasm	1
TRAF6_TAB1_TAB2_TAK1_complex	0
TRAF6_ubiquitinated_M1_macrophage___Cytoplasm	0
TRAF6_ubiquitinated_M2_macrophage__cytoplasm	0
TRAM1	0
TRAM1_TRIF_complex	0
TRIF	1
TSC2_phosphorylated	0
TSG6	1
TXK	1
TYK2	1
UEV1A_UBC13_complex	1
VAV1	0
VAV1_2_3	1
VCAM1_Fibroblast__Cytoplasmic_membrane_down	1
VCAM1_ITGA4_ITGB7_complex	1
VCAM1_TH1___cytoplasmic_membrane_down	1
vcan	1
VEGFa_Fibroblast___Extracellular_Space	1
VEGFa_Fibroblast___secreted_components	1
VEGFa_M2_macrophage__extracellular_space	1
VEGFa_M2_macrophage__secreted_components	1
VEGFA_rna	1
Vegfa_rna_M1_macrophage___nucleus	1
Vegfa_rna_M2_macrophage_nucleus	1
VEGFA_VEGFR_complex	1
Vegfb_M2_macrophage__extracellular_space	0
Vegfb_M2_macrophage__secreted_components	0
vegfc	1
Vegfc_M2_macrophage__extracellular_space	1
Vegfc_M2_macrophage__secreted_components	1
Vegfc_rna	1
vegfc_vegfr3_complex	1
VegfR1	1
VegfR1_Vegfa_complex	1
VegfR1_vegfb_complex	0

VegfR2_vegfc_complex	1
WNT_FRIZZLED_complex	1
WNT5A	1
WNT5B	1
WNT5B__rna	1
WNT5B_FZD1_LRP5_complex	1
WNT5B_gene	1
XIAP_M1_macrophage___Cytoplasm	1
XIAP_M2_macrophage__cytoplasm	1
YAP1_phosphorylated	1
YY1	1
ZAP70	0
ZC3H12A	0

Supplementary table 11. *Therapeutic drug targets in the RA M1 macrophage model.*

Therapeutic targets					
SRC	CSF2	IRAK4	XIAP	NFKB1	STAT1
SIRT1	IL23R	BCL2	PTK2	RELA	IFNG
MAPK14	SYK	C5AR1	PPIA	CXCR1	CD40LG
AKT1	RAF1	TBK1	TLR2	MAPK3	TNFSF11
PTPN6	CCR2	GALR2	TLR4	ITGB1	IL6
TLR9	BCL2L1	IRAK1	PRKCQ	MYC	INHBA
TLR8	CASP3	TYK2	IL6R	STAT3	INHBB
TLR7	CASP8	CASP1	IL6ST	RAC1	NOTCH1
MCL1	FAS	CASP7	TNF	IL1B	CSF2RB
MDM2	ACVR2B	NLRP3	DUSP1	CCL2	IL12A
JAK2	NFKBIA	PRKCD	JUN	RHOA	ITGA1
JAK1	IFNAR2	SMAD7	ACVR2A	IL18	

Supplementary table 12. *Therapeutic drug targets in the RA M2 macrophage model.*

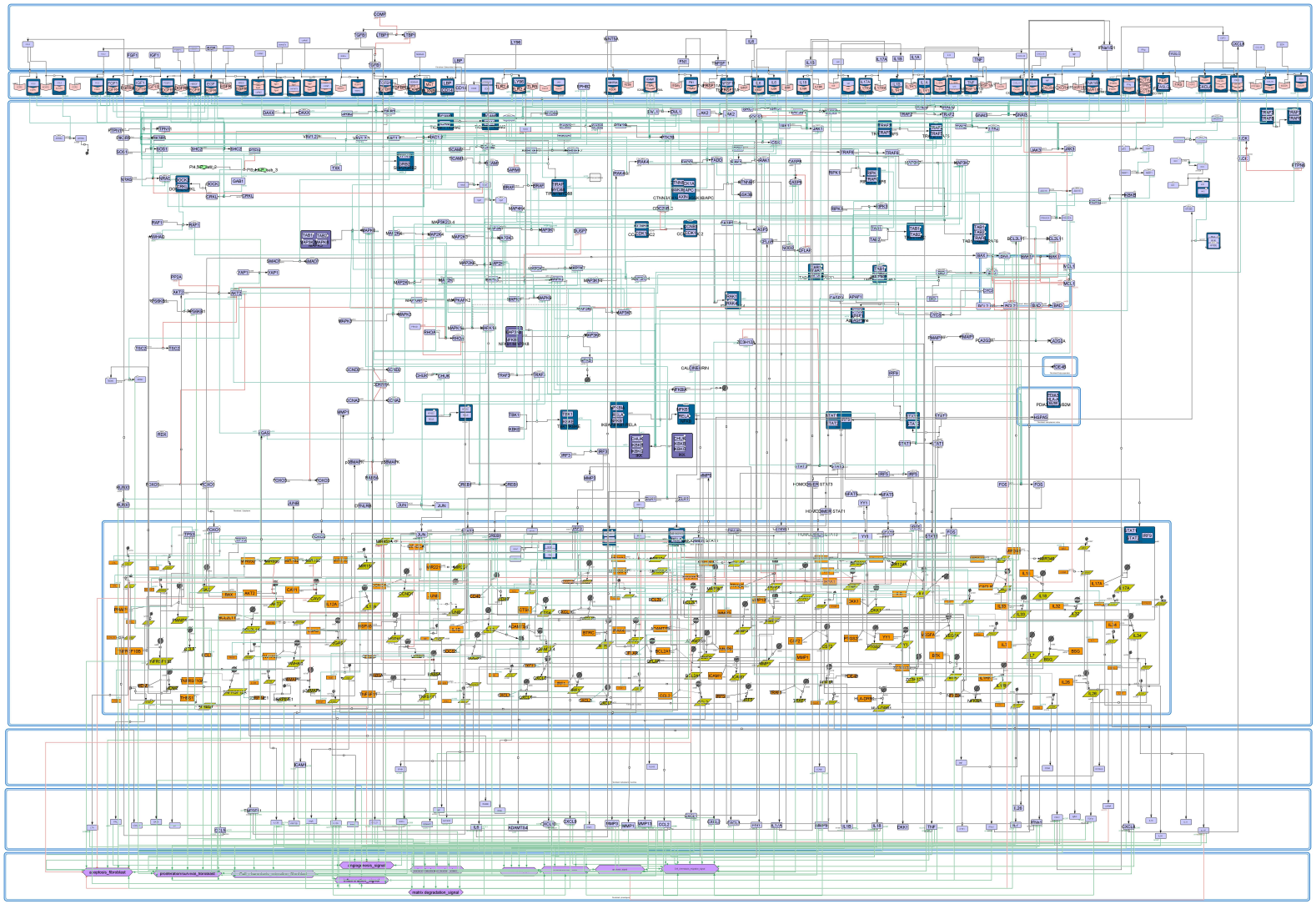
Therapeutic Targets					
SRC	VEGFA	TYK2	CASP9	CSF1	JAK2
GSK3B	IL23R	RXRA	DUSP1	LIFR	CSF1R
SIRT1	IL4R	MERTK	NFKB1	IL4	JAK1
MAPK14	SYK	TGFBR2	RELA	ITGA1	FAS
AKT1	RAF1	CASP1	CXCR1	GAS6	BCL2

PTPN6	TGFBR1	CASP7	MAPK3	VEGFC	C5AR1
HCK	BCL2L1	NLRP3	ITGB1	NFKBIA	PTK2
MCL1	CASP3	PRKCD	MYC	BAX	PRKCQ
MDM2	CASP8	XIAP	STAT3	SMAD7	IL6ST
IL17RA	IL17RC	TGFB1	CD40LG	KLF4	PRLR

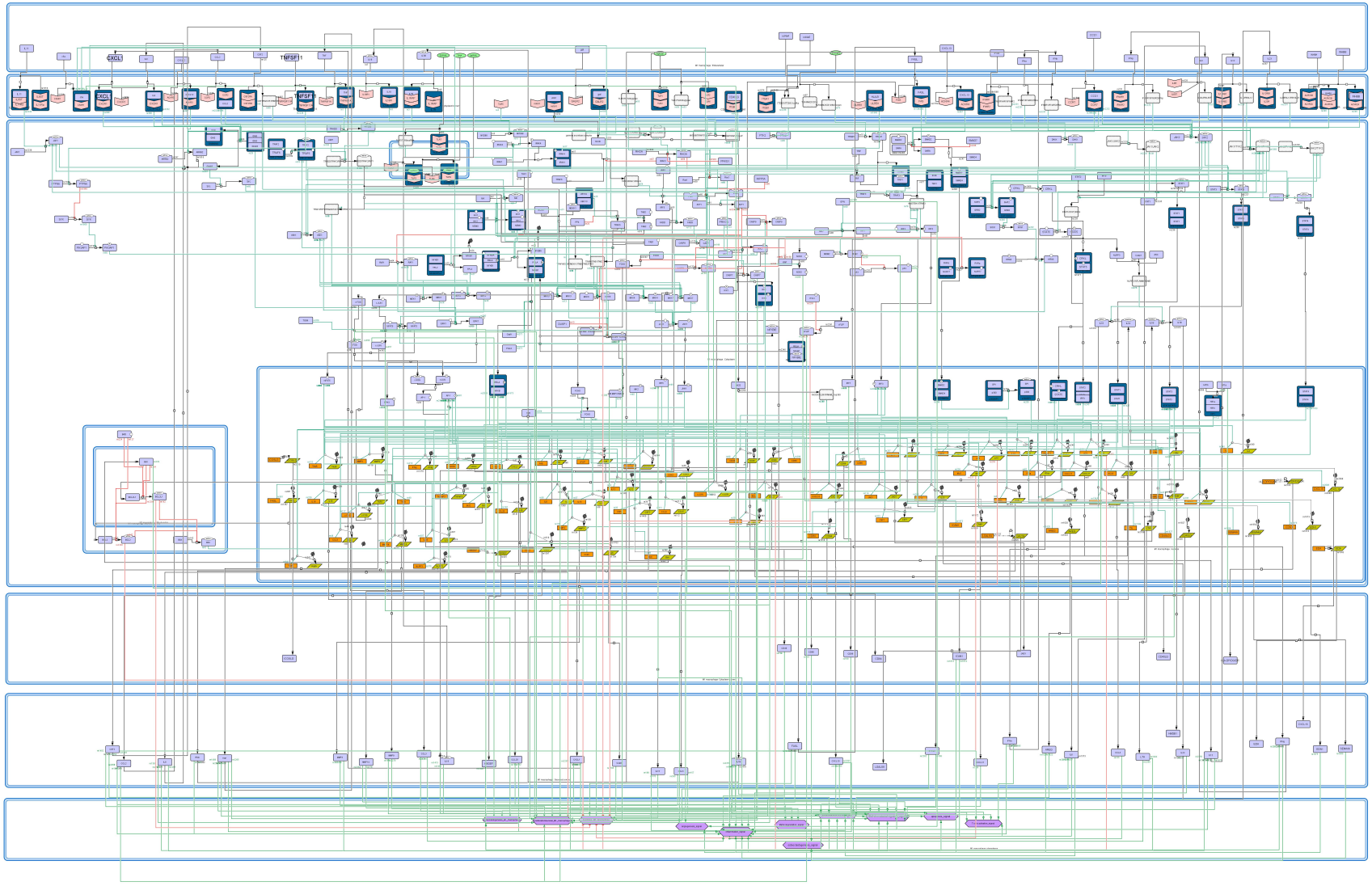
Supplementary table 13. *Therapeutic drug targets in the RA multicellular model.*

Therapeutic Targets					
PDE4B	EGFR	CTSK	BIRC2	MAPK3	KRAS
SRC	JAK1	FAS	XIAP	ITGB1	IL7R
MAP3K1	CSF2	ACVR2B	MAP4K4	HIF1A	LAG3
MAP3K2	IL15	FGFR4	RIPK1	ITGA5	CD28
IKBKB	DKK1	CHUK	PTK2	ADAMTS4	TGFB1
GSK3B	PDCD1	IRAK4	PPIA	CXCL8	LRP5
CCNB1	IL2	BCL2	TLR2	MYC	NOTCH1
SIRT1	IL9	CCR5	TLR4	PTPRC	LIFR
MAPK1	PPARG	C5AR1	MIF	STAT3	CSF2RB
RPS6KB1	IL1A	PRKACA	CDC25B	MAP2K2	CTNNB1
MAPK8	ICAM1	TBK1	FGF1	RAC1	SEMA4D
MAPK14	BTK	MAP3K5	PLA2G2A	IL1B	ICOS
LCK	VEGFA	GALR2	ITGAL	CCL2	IL1R1
AKT1	PTPN11	IRAK1	PRKCQ	RHOA	IL4
PTPN6	MTOR	LYN	IL6R	IL17A	ITGA1
PTGS2	CXCR4	TYK2	IL6ST	STAT1	TNFRSF4
TLR9	IL23R	MAPK9	TNF	IFNG	ICOSLG
TLR8	IL4R	CXCR2	CXCL10	CD40LG	GAS6
TLR7	SYK	ZAP70	SERPINE1	CXCL12	VEGFC
MMP1	NRAS	PDGFRB	TP53	TNFSF11	BMP6
MMP14	CXCR3	MAP3K8	CASP9	IL6	NFKBIA
MMP3	RAF1	MERTK	DUSP1	CCL5	BAX
MMP9	JAK3	MAPKAPK2	JUN	HSPA5	IFNAR2
MMP13	F2R	TGFBR2	NFKB1	CD86	SMAD7
HCK	TGFBR1	AKT2	NFKB2	CD80	ACVR2A
MCL1	MAP3K7	CASP1	RELA	KLF4	IL18
BRAF	CCR2	CASP7	RELB	CCL20	FN1
ITGB7	INHBA	INHBB	PRLR	CSF1	IL17RA
THBS1	CSF1R	CASP8	EPHB2	CXCR1	CREB1
MIR221	CD40	NOS2	PTK2B	ITGB2	
MET	ITGA4	NLRP3	CCND1	IL17RC	

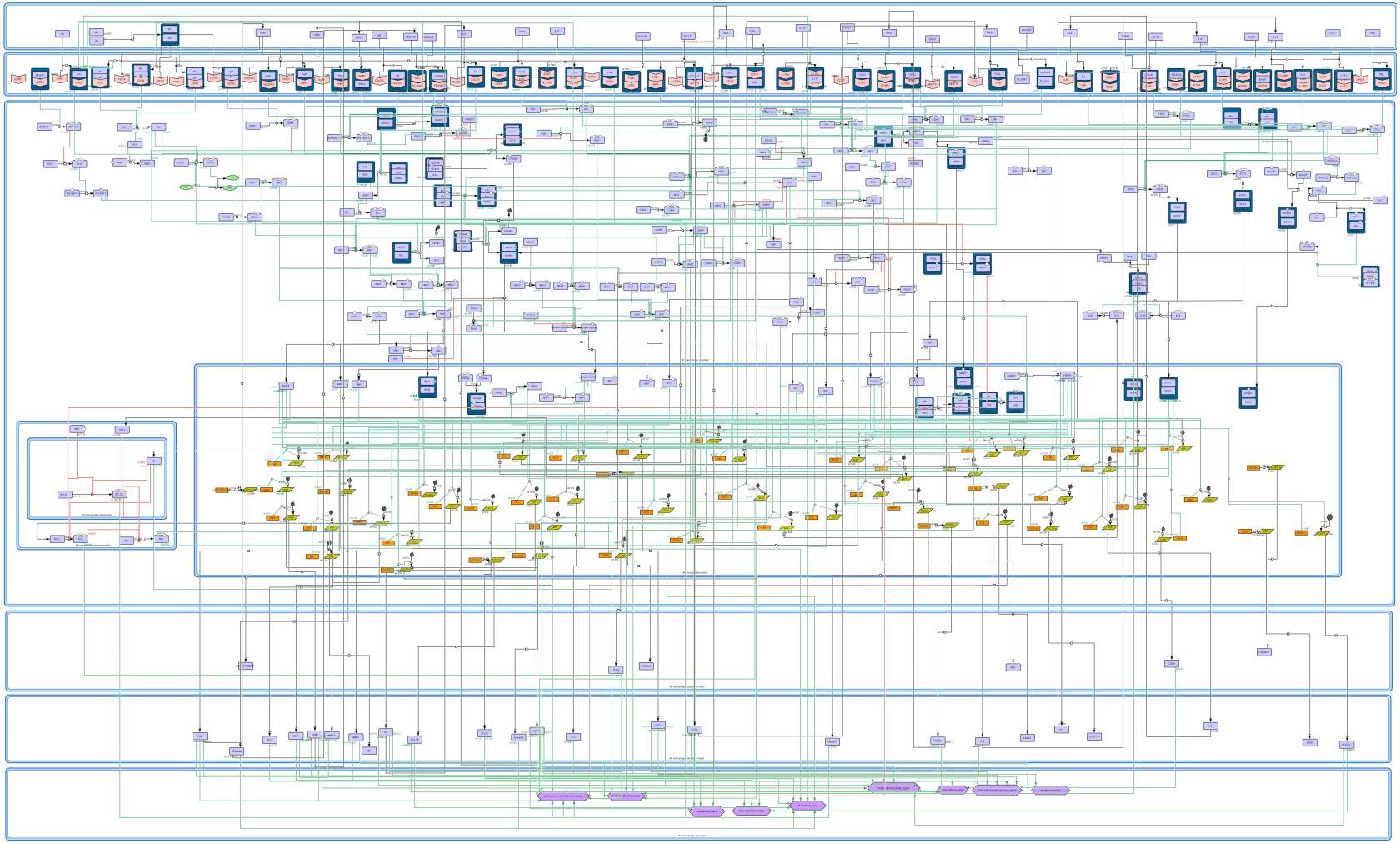
MDM2	ACKR3	IGF1R	ITGAV	CAV1	
JAK2	CASP3	PRKCD	MAP2K1	TBX21	



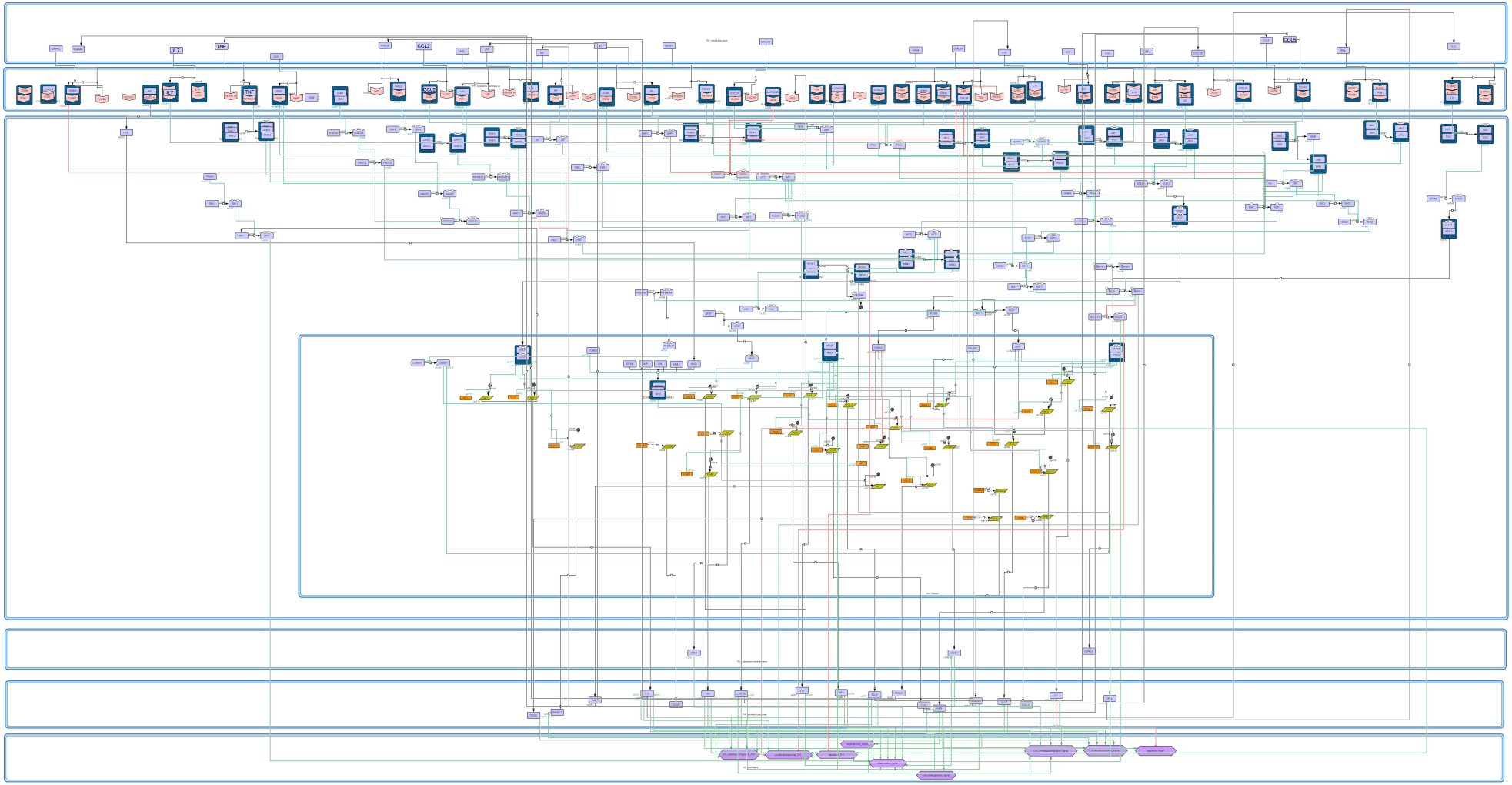
Supplementary figure 1. The RA fibroblast map in CellDesigner



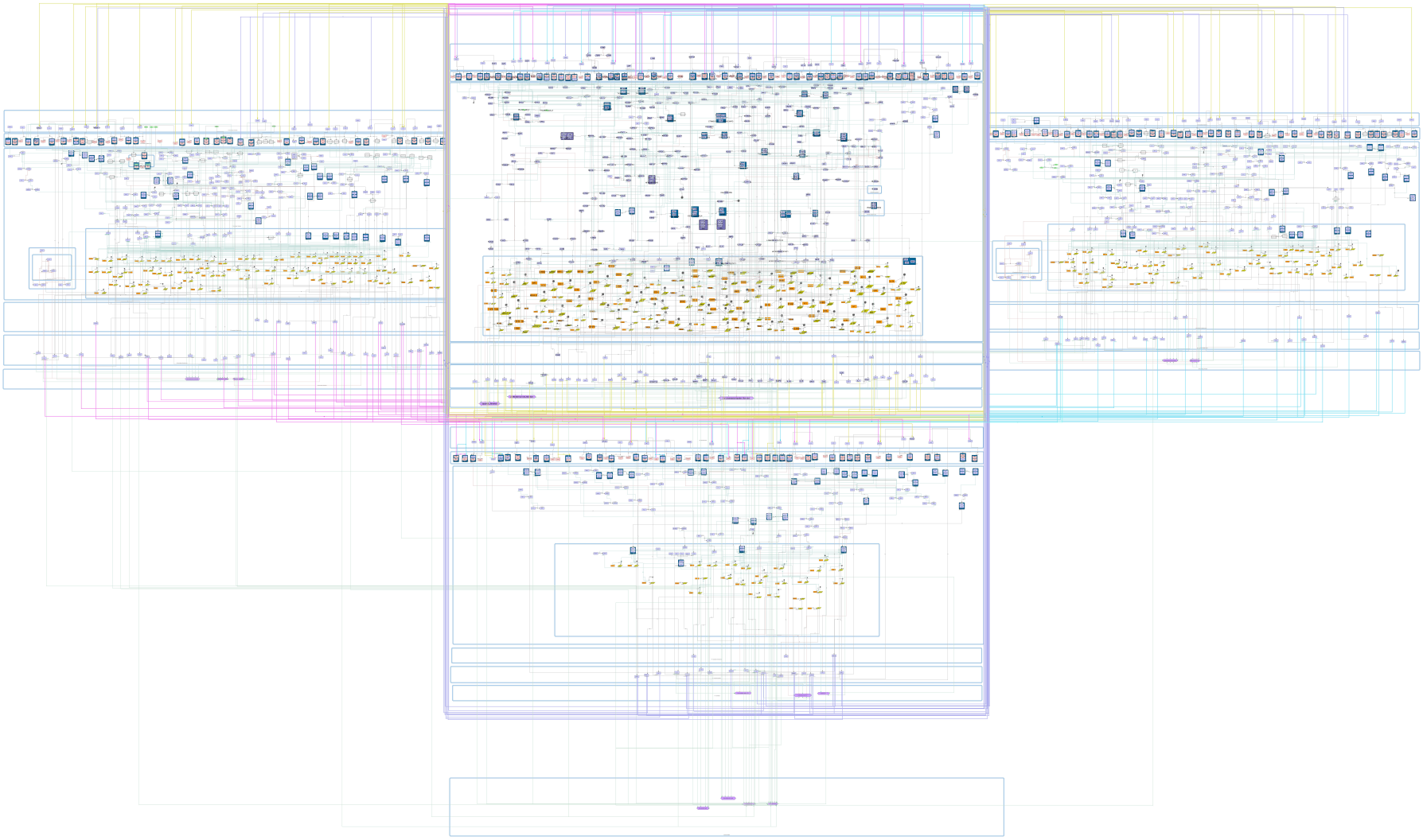
Supplementary figure 2. The RA M1 macrophage map in CellDesigner.



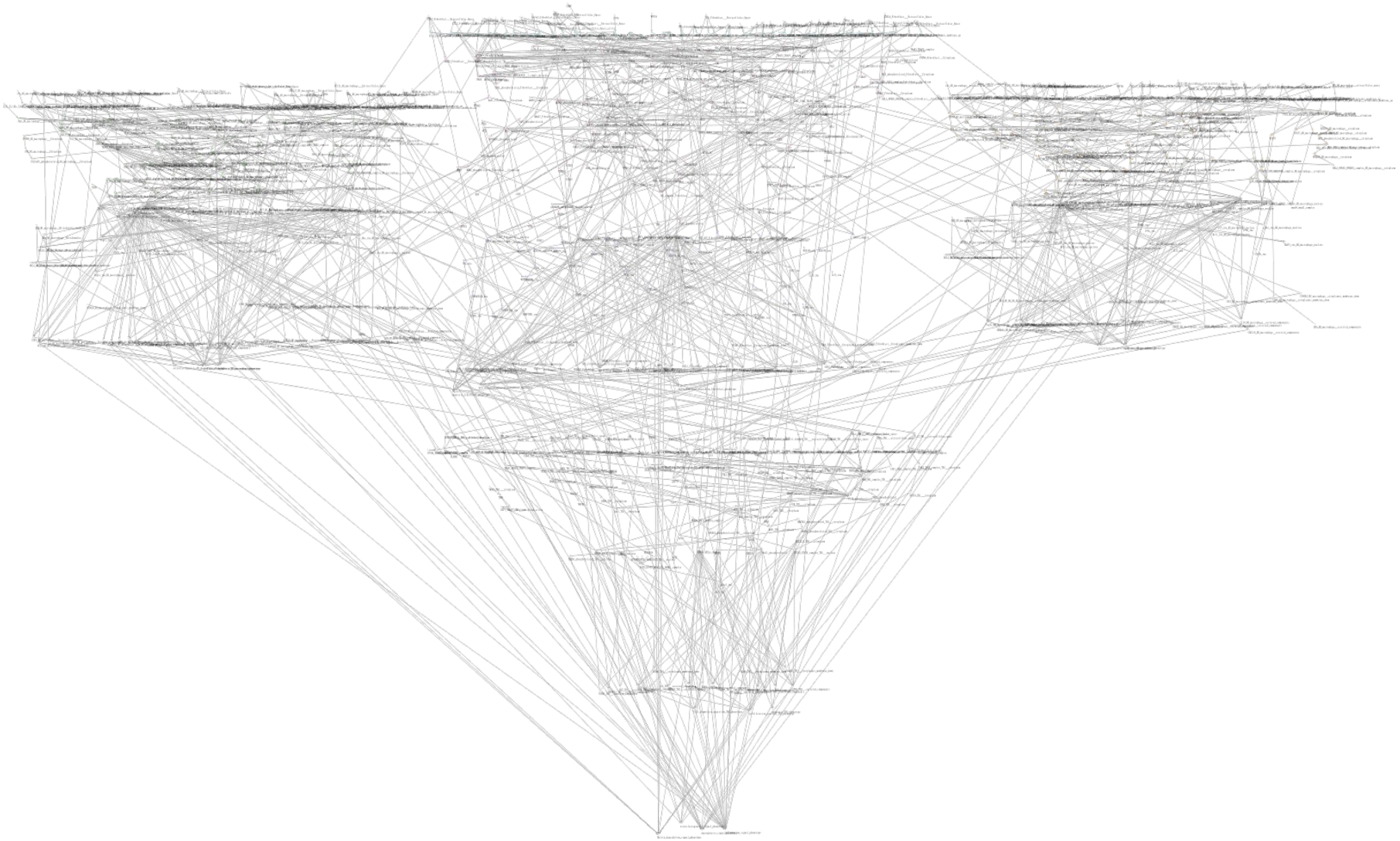
Supplementary figure 3. The RA M2 macrophage map in CellDesigner.



Supplementary figure 4. The RA CD4+ Th1 map in CellDesigner.



Supplementary figure 5. The RA multicellular map in CellDesigner.



Supplementary figure 6. The RA multicellular model in BMA graphical interface.

CIFRE PhD project proposal

INTRODUCTION:

Immune dysregulation was first implicated in the pathogenesis of rheumatoid arthritis (RA) by the discovery of anti-immunoglobulin G (IgG) antibodies known as rheumatoid factors, in the 1940s. However, concepts of how immune responses contribute to disease have evolved dramatically over the last 50 years. Autoreactivity as a pivotal step dominates the conceptual landscape, although other mechanisms, both immunologic and tissue-derived, clearly contribute to disease pathogenesis. (Firestein & McInnes, 2017).

Many cells and their cytokines play critical roles in the development of RA. (*Figure 1*) provides an overview of the main cell types implicated in the disease. The synovial compartment is infiltrated by leukocytes and the synovial fluid is inundated with pro-inflammatory mediators that are produced to induce an inflammatory cascade, which is characterized by interactions of fibroblast-like synoviocytes with the cells of the innate immune system, including monocytes, macrophages, mast cells, dendritic cells… as well as cells of adaptive immune system such as T cells and B cells. Endothelial cells contribute to the extensive angiogenesis. The fulminant stage contains hyperplastic synovium, cartilage damage, bone erosion, and systemic consequence. The destruction of the subchondral bone can eventually result in the degeneration of the articular cartilage as the result of a decrease in osteoblasts and an increase in osteoclasts and synoviocytes. (Guo et al, 2018).

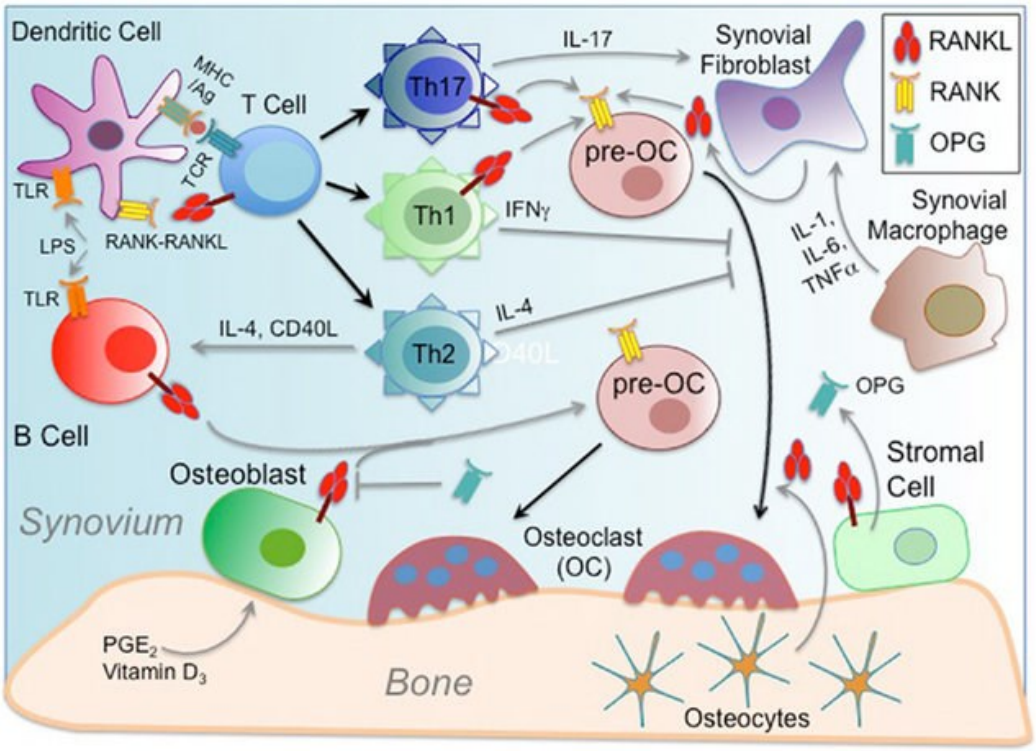


Figure 1. Cellular interplay in RA (Walsh, M. C., & Choi, Y., 2014).

OBJECTIVE:

The objective of this project is to construct a computational model able to decipher the interplay between cells of the innate and adaptive immunity in RA, that eventually leads to bone and cartilage breakdown.

To do so, we will start by mapping cell specific molecular interactions, creating separate maps for T cells, B cells, macrophages, fibroblasts, osteoblasts, osteoclasts. These maps will serve as a basis for the generation of cell specific dynamic models.

The next step is the creation of a multicellular model in order to understand how the different cells interact, contributing to the emergent behavior of the system. For this purpose, intercellular interactions must be clearly defined, mapped and described.

Appropriate choices of methodology will be made, and proper tools will be employed for every scale (creation of intra and inter cellular networks, cell specific and multicellular dynamical models). *Figure 2* shows an overview of the project. A detailed description of each step follows.

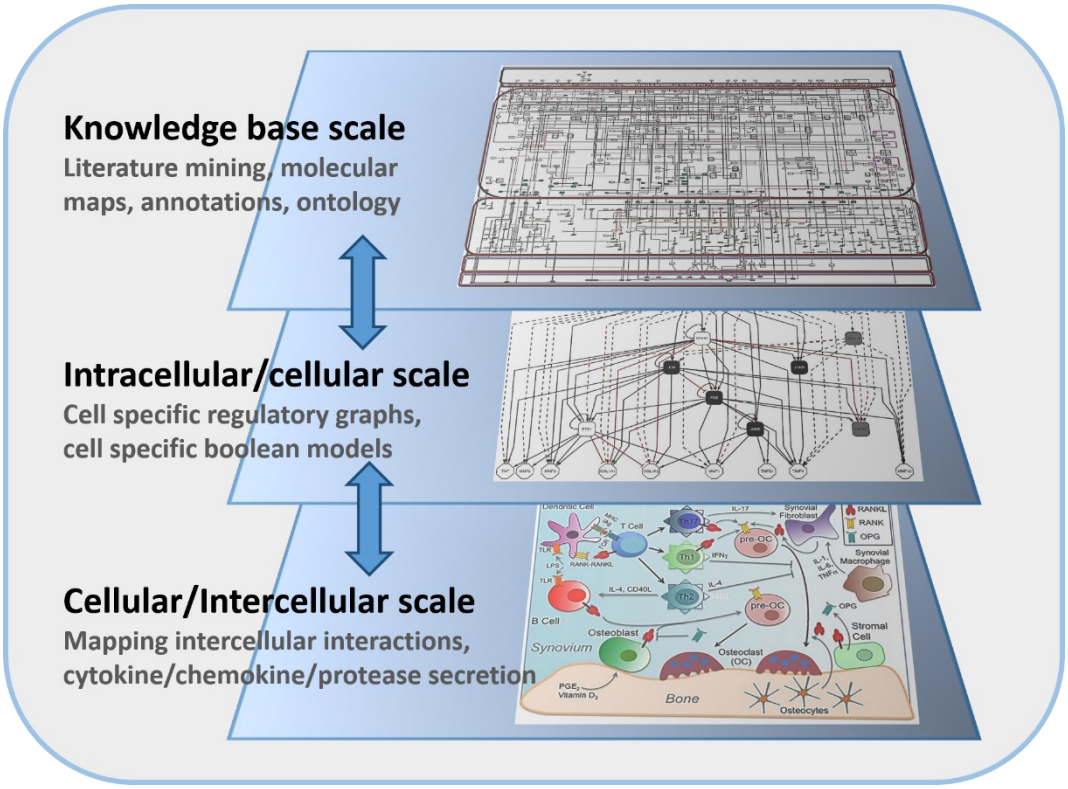


Figure 2. Multiscale model of cellular interplay in RA.

METHODOLOGY:

Construction of cell specific networks

The student will construct cell specific maps using data mining tools, appropriate data bases such as KEGG (Kanehisa et al, 2000), REACTOME (Fabregat et al, 2018), IPA (https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis), Metacore (Clarivate Analytics, https://portal.genego.com/) as well as internal data generated within Sanofi and not yet published, and the graph editor CellDesigner (Funahashi et al, 2003). He/She will use the platform MINERVA (Gawron et al, 2016) for automatic annotation and reference of the cell specific maps and exploit several MINERVA functionalities such as i) semantic zoom, ii) interface with external databases, iii) omics data overlays. The student will benefit greatly from a global, fully annotated RA specific map built in the laboratory (Singh et al, 2018, Singh et al, 2020). This map features interactions implicated in RA coming from various cell types, but due to the extensive annotations the user can opt for cell specific interactions and extract the corresponding network.

We are also going to use public datasets of expression data (microarrays, RNAseq, RNAseq single cell), data concerning metabolic pathways from MetaCyc (https://metacyc.org/) and Sanofi's proprietary datasets to enrich and expand

existing pathway resources where needed.
Moreover, for this step, the student will characterize and map the intercellular interactions that govern cell-cell communication.

Dynamical modelling of intracellular networks for cells implicated in RA
First, using the tool CaSQ (Aghamiri et al, under revision) that Dr. Niarakis develops in collaboration with INRIA, Paris-Saclay the student will create cell specific, executable Boolean models. These models will be simulated and analyzed using the modelling platform Cell Collective (https://cellcollective.org/) and the modelling software GINsim (Chaouiya et al, 2012). In this framework, a regulatory network is modeled in terms of a regulatory graph, where nodes represent regulatory components (proteins, complexes, transcription factors etc.), whereas arcs represent interactions between these components. In addition, each regulatory component is associated with a logical variable denoting its qualitative concentration or level of activity. In most cases, Boolean variables (taking the values 0 or 1) are sufficient to represent the most relevant situations.
These qualitative models will be used to decipher the input-output relationships in a cellular level (inputs: different ligands that lead to different receptor activation, output: different gene expression activation, different secreted molecules like cytokines, chemokines, MMPs etc).
The results of this analysis, i.e. the behavior of each cell according to the environmental/extracellular stimuli, along with the mapping of intercellular interactions of the previous step, will set the basis for the building of a multi-cellular model.

Dynamical modelling of intercellular networks – building a multicellular model for RA
For this step, the student will use all results of previous steps and will prioritize signature pathways for each cell type. He/She will combine them to build a logical model that will represent the intra- and intercellular relationships defined.

Network reduction
Considering that the size of the state space of a Boolean model is exponentially dependent on the node number (2^n states for n nodes), the reduction of the network is essential for computationally feasible tracking of all attractors. GINsim includes a model reduction approach (Naldi et al., 2011) that consists of the idea of allowing the user to manually select a series of components to hide. Then the software recomputes the logical rules of their targets. Provided no regulatory circuit is eliminated, the most salient dynamical properties are preserved, including all stable states.
Besides network reduction techniques, for coping with size and complexity, the

software MaBoSS (Stoll et al, 2017) (and the tools of MaBoSS environment, such as PhysiBoss) will also be employed for the computation of phenotype probabilities.

Coherence of model behavior/Simulation and model perturbations

The evaluation of the coherence of the proposed model will be tested against published data and data provided by SANOFI, together with prior knowledge. We will try to reproduce with *in silico* simulations the dynamics and temporal behavior of the various cells during joint inflammation that lead to cartilage degradation and bone erosion. The model will also be tested systematically concerning its efficiency in reproducing the effects of targeted therapy.

Predictions

Systematic testing of different initial conditions, concerning the number of each cell type, the number of secreted molecules or/and the number of simulations could further lead to predictions regarding the outcomes of specific perturbations, as single or combined effects. The ultimate goal is to gain a better understanding of the mechanism behind cartilage and bone degradation, two major debilitating symptoms of RA, and propose a strategy that could help block or even reverse these outcomes.

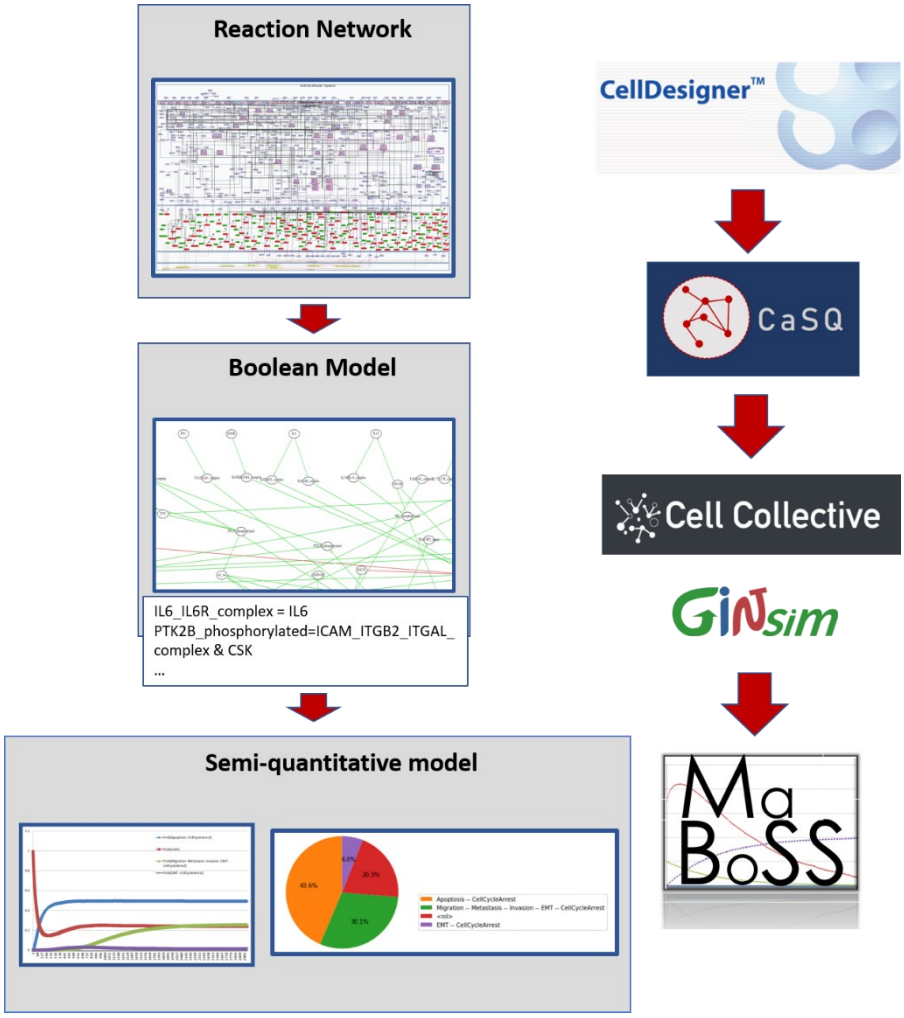


Figure 3: Schematic representation of the proposed computational framework for the building of a multicellular model for RA. The proposed tools are also highlighted in the right panel of the figure.

Novelty of the proposed project - impact

The project's novelty consists in the development of a multicellular model for the RA disease taking advantage of all the recent advances in methodologies and state of the art tools in computational systems biology. The model obtained will cover the need for versatile tools that help gain understanding of disease mechanisms using dynamic analysis and multi-omics data integration. A robust computational model can reduce the need for animal models and can also help elucidate human specific pathways that are not obligatory shared with mice counterparts. Besides, induced arthritis to mice models cannot always be considered as a representative system. In silico simulations and predictions can significantly reduce the time of experimental design and can also be used for simulation of complex scenarios that are not easily reproduced in the lab. In addition to hypothesis testing and mechanistic insights, the model can be used for predictions of novel therapeutic targets and biomarkers proposed for in vitro

validation.

REFERENCES AND PROPOSED BIBLIOGRAPHY:

1. Firestein, G. S., & McInnes, I. B. (2017). Immunopathogenesis of Rheumatoid Arthritis. Immunity, 46(2), 183-196.
2. Guo, Q., Wang, Y., Xu, D., Nossent, J., Pavlos, N. J., & Xu, J. (2018). Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies. Bone research, 6, 15. doi:10.1038/s41413-018-0016-9
3. Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research, 28(1), 27-30.
4. Antonio Fabregat, Steven Jupe, Lisa Matthews, Konstantinos Sidiropoulos, Marc Gillespie, Phani Garapati, Robin Haw, Bijay Jassal, Florian Korninger, Bruce May, Marija Milacic, Corina Duenas Roca, Karen Rothfels, Cristoffer Sevilla, Veronica Shamovsky, Solomon Shorser, Thawfeek Varusai, Guilherme Viteri, Joel Weiser, Guanming Wu, Lincoln Stein, Henning Hermjakob, Peter D'Eustachio; The Reactome Pathway Knowledgebase, Nucleic Acids Research, Volume 46, Issue D1, 4 January 2018, Pages D649–D655, https://doi.org/10.1093/nar/gkx1132
5. https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis
6. Funahashi A, Tanimura N, Morohashi M, Kitano H (2003) Cell Designer: a process diagram editor for gene-regulatory and biochemical networks. BIOSILICO 1:159–162 the proceedings of the 2nd Disease Maps Community meeting. Briefings in Bioinformatics (2018)
7. V. Singh, M. Ostaszewski, G.D Kalliolias, G. Chiocchia, R. Olaso, E. Petit-Teixeira, T. Helikar, A. Niarakis Computational Systems Biology Approach for the Study of Rheumatoid Arthritis: From a Molecular Map to a Dynamical Model. Genomics and Computational Biology Vol 4 No 1 (2018)
8. V. Singh, G.D. Kalliolias, M. Ostaszewski, M. Veyssiere, E. Pilalis, P. Gawron, A. Mazein, E. Bonnet, E. Petit-Teixeira, A. Niarakis RA-map: Building a state-of-the-art interactive knowledge base for rheumatoid arthritis, Accepted, Database
9. P. Gawron, M. Ostaszewski, V. Satagopam, S. Gebel, A. Mazein, M. Kuzma, S. Zorzan, F. McGee, B. Otjacques, R. Balling, and R. Schneider "MINERVA—a platform for visualization and curation of molecular interaction network". In : npj Systems Biology and Applications (2016), vol. 2, p. 16020, doi : 10.1038/npjsba.2016.20

10. S.S. Aghamiri, V. Singh, A. Naldi, T. Helikar, S. Soliman* and A. Niarakis*Automated inference of Boolean models from molecular interaction maps using CaSQ, Under review, Bioinformatics
11. Helikar, T. et al. The Cell Collective: Toward an open and collaborative approach to systems biology. BMC Syst. Biol. 6, 96 (2012).
12. Chaouiya C, Naldi A, Thieffry D (2012) Logical modelling of gene regulatory networks with GINsim. Methods Mol Biol 804:463–479
13. Gautier Stoll, Barthélémy Caron, Eric Viara, Aurélien Dugourd, Andrei Zinovyev, Aurélien Naldi, Guido Kroemer, Emmanuel Barillot, Laurence Calzone, MaBoSS 2.0: an environment for stochastic Boolean modeling, Bioinformatics, Volume 33, Issue 14, 15 July 2017, Pages 2226–2228,

CONTRIBUTED BY SANOFI

- Expertise in RA.
- New data to complete the existing map
- Calculation facilities
- Experiments to validate the model
- Contribution to simulations and evaluation of results (ZB)
- Testing of other logical modelling platforms (CellNetOptimizer, CellNOpt etc?) ZB
- Access to commercially available databases

CONTRIBUTED BY THE ACADEMIC PARTNER

- Expertise in RA
- Expertise in Disease map construction
- State of the art map for RA
- Expertise in Boolean model building and analysis
- Development of all simulation methods and model parametrization
- Efficient small-scale and multi-omic data integration

EXPECTED RESULTS AND OUTPUT

- Multiscale models of Rheumatoid Arthritis networks easy to maintain and upgrade.
- Simulated models for RA fully operational and easy to maintain able to simulate the pathogenesis.
- Better understanding of the disease with a full up to date knowledge base on RA.

- Increase Sanofi's skills/capabilities in the domain of disease map and Boolean based model simulation.

ANNEX B

Publications

Zerrouk N, Alcraft R, Hall BA, Augé F, Niarakis A. Large-scale computational modelling of the M1 and M2 synovial macrophages in Rheumatoid Arthritis [Internet]. bioRxiv; 2023 [cited 2023 Sep 25]. p. 2023.09.11.556838. Available from: https://www.biorxiv.org/content/10.1101/2023.09.11.556838v1 (submitted to njp systems biology in September 2023).

Zerrouk, N., Aghakhani, S., Singh, V., Augé, F. & Niarakis, A. A Mechanistic Cellular Atlas of the Rheumatic Joint. 2, 14 (2022).

Aghakhani, S., **Zerrouk, N**. & Niarakis, A. Metabolic Reprogramming of Fibroblasts as Therapeutic Target in Rheumatoid Arthritis and Cancer: Deciphering Key Mechanisms Using Computational Systems Biology Approaches. Cancers 13, (2021).

Scientific Presentations

Naouel Zerrouk, Franck Augé, Anna Niarakis. Computational modeling of the cellular interplay in Rheumatoid Arthritis. Flash presentation given at: Building immune digital twins workshop; May 5th _ June 2nd, 2023.

Naouel Zerrouk, Sahar Aghakhani, Franck Augé, Anna Niarakis. A Mechanistic Cellular Atlas of the Rheumatic Joint. Poster presented at: 21st European Conference on Computational Biology; September 12th_ September 21st, 2022.

Naouel Zerrouk, Franck Augé, Anna Niarakis. Computational modelling of the cellular interplay in Rheumatoid Arthritis. Deciphering the role of innate and adaptive immunity in cartilage destruction and bone erosion. Poster presented at: Disease Map Community Meeting 5th edition; November 12th – November 14th, 2020.

Naouel Zerrouk, Charles Bettembourg, Anna Niarakis. Computational modelling of the cellular interplay in rheumatoid arthritis. Deciphering the role of innate and adaptive immunity in cartilage destruction and bone erosion. Poster presented at: Cancer systems biology - Promises of artificial intelligence; September 28th – October 2nd, 2020.