KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science
DOCTORAL SCHOOL OF BIOMEDICAL SCIENCES
Faculty of Medicine

In silico modeling of osteochondral pathophysiology and regeneration

Identification of molecular signatures and therapeutic targets with mechanistic and data-driven network models



Raphaëlle Lesage

Dissertation presented in partial fulfillment of the requirements for the interdisciplinary degree of Doctor of Engineering Science: Mechanical Engineering and of Biomedical Sciences (PhD)

Supervisors:
Prof. Dr. Liesbet Geris
Prof. Dr. Rik Lories

November 2021

KU LEUVEN

In silico modeling of osteochondral pathophysiology and regeneration

Identification of molecular signatures and therapeutic targets with mechanistic and data-driven network models.

Raphaëlle LESAGE

Supervisory Committee:

Em. Prof. Omer Van der Biest, **chair** (KU Leuven)

Prof. dr. ir. Liesbet Geris, **supervisor** (KU Leuven)

Prof. dr. Rik Lories, **supervisor** (KU Leuven)

Prof. dr. Hans Van Oosterwyck (KU Leuven)

Prof. dr. Silvia Monteagudo (KU Leuven)

Prof. dr. G.J.V.M (Gerjo) van Osch (Erasmus MC Rotterdam and University of Technology Delft, the Netherlands)

Prof. dr. Tomas Helikar (University of Nebraska-Lincoln, USA)

Prof. dr. Alexopoulos Leonidas (National Technical University of Athens, Greece)

Dissertation presented in partial fulfillment of the requirements for the interdisciplinary degree of Doctor of Engineering Science: Mechanical Engineering and of Biomedical Sciences (PhD)

November 2021

© KU Leuven – Faculty of Engineering
Celestijnenlaan 300C box 2419, B-3001 Heverlee (Belgium)
© KU Leuven– Faculty of Medicine
ON II Herestraat 49 - box 400 B-3000 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

Tous droits réservés. Aucun extrait de cet ouvrage ne peut être reproduit, ni saisi dans une banque de données, ni communiqué au public, sous quelque forme que ce soit, électronique, mécanique, par photocopie, film ou autre, sans le consentement écrit et préalable de l'éditeur.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm or any other means without written permission from the publisher.

Acknowledgement

This book is the realization of an almost 5-years PhD project, which of course has been scattered with some difficulties (including a world pandemic) but it has also been enriching on so many personal and professional aspects. Achieving this research journey would not have been possible without the help and support of many people whom I would like to thank.

First, I would like to thank my promoter Prof. Liesbet Geris for offering me the opportunity to work on this fascinating project. "Lorsqu'on nous fait confiance on est plus grand que ce qu'on est" (Erik Orsenna), so thank you Lies for always trusting me and believing in me. I also would like to thank you for the scientific and professional guidance, the positive support and research opportunities you provided me as well as for your availability for all the small and more important matters during my PhD. I also acknowledge Prof. Rik Lories who became my supervisor later in that journey, allowing me to route for one of the first official interdisciplinary PhD degree at KU Leuven, but who was also there from the beginning to provide valuable advices and scientific discussions. Thank you for always having kept the door of your office open for a chat, your availability and open-mindedness were much appreciated.

Next, I express my sincere gratitude to Prof. Hans Van Oosterwyck, Prof. Silvia Monteagudo, Prof. Gerjo van Osch, Prof. Tomas Helikar and Prof. Leonidas Alexopoulos for taking the time to read my PhD thesis and for the feedback and interesting discussions. In addition, I would like thank Prof. van Osch, Gerjo for transmitting her enthusiasm about cartilage biology and research. I am grateful for the time and enthusiasm you put in training me and the other CarBon ESRs during the bi-yearly meetings and my research stay at Erasmus MC, but also for your availability and guidance the rest of the time. You contributed to make me a better scientist.

Working in the Prometheus platform, the TE unit and with the BioMech group of ULiège has been such an enriching and valuable experience, which undoubtedly contributed to the achievement of this thesis. Not only did my amazing colleagues generate this helpful and friendly atmosphere but also the technical and administrative staff was always ready do help and bring support, a warm thank you to all of them. This environment also triggered most valuable scientific discussions and collaborations. I would like to dedicate a few words to my project partner and friend Satanik. Thank you for the fruitful collaboration and scientific exchanges from which I have learnt so much and for the fun outside of work, which I hope will

continue. In addition, I would like to thank my other CarBon buddy, Mauricio, who has been essential for the experimental part of this research. From Rotterdam to Leuven passing by Maastricht, I won't forget all the laugh and all your positive energy that made of our collaboration such a memorable experience.

In the end, what really made the difference in my PhD adventure is that this friendly atmosphere with colleagues navigated outside the office with parties, after works, volley, football, game evenings, croquets, etc... Thank you to Niki (the unique), Sébastien, Gabriella, Priyanka, Tim (my desk buddy and bioinformatics guru), Hannah (my favorite ballet teacher), Nico & Satanik (my Belgian trio), Delia (my car pooling mate), Thomas, Isaak, Lisanne, Angela, Paolo, Luis, Haftu, Naomi, Marion, Dana, Karen and Ruben (best football coach), Roger for all these moments. Without forgetting my fantastic colleagues from the Biomech team for the great science, fun activities and weekends: Morgan, Sophie (my travel acolyte for Systems Biology conferences across the world), Marc, Fernando, Ehsan, Gabriele, Laura, Mojtaba, Bernard, Sourav. In addition, to all my CarBon fellows who I used to meet every 6-months in different corners of Europe. Even though the COVID-19 tarnished our last meeting, I keep a great souvenir of our gatherings and I'm looking forward for more.

Also important for me to thanks are my most long lasting friends since I moved in Belgium and on whom I have always been able to count Fabio, Francisco, Seba, Thati, Manu, Yamid. But it's also Giorgia, Paola, Deesh and later the beach-volley crew: Luiza, Fernando, Jhon, Yeral, Julio. You all made me feel at home in Belgium, made me smile and brought the fresh air and support I needed to maintain my sanity despite the dips that any doctorate and the distance from family can bring. À mes GGs et à ma sœur de cœur, Iris, avec qui n'a rien n'a changé malgré la distance et le temps : sans que vous ne le notiez, votre présence et les fous rires partagés m'ont aussi donné l'énergie nécessaire pour réussir cette aventure.

Enfin, merci à ma famille, en particulier à mes parents et ma sœur Astrid pour leur soutien et leur amour inconditionnel. Vous m'avez toujours soutenue dans mes choix personnels et professionnels. Même si vous ne compreniez pas toujours mes recherches, vos «Alors, tu as trouvé?» me donnait la motivation pour persévérer ! C'est aussi à vous que je dois cette thèse.

Pierre-Edouard, mon pilier qui n'a jamais failli, ton amour et ton soutien quotidien ont été essentiels. Même lorsque je prenais les choses trop sérieusement tu étais toujours là pour me faire rire et relativiser. La vie n'aurait pas le même goût sans toi, je t'aime.

Abstract

Bone and cartilage tissues are both constituents of the skeletal system, which can be impaired by intrinsic and extrinsic events such as disease or traumatic injury. For bone, about 90% of defects (e.g. fractures) can heal spontaneously but few efficient regenerative solutions are available for the remainder of defects that cannot heal spontaneously. For joints, the most prevalent disease is osteoarthritis (OA) in which cartilage is progressively degraded without self-regeneration potential. No OA disease-modifying drugs are currently approved for use in patients, and limited advanced therapy medicinal products resulting from osteochondral tissue engineering (TE) strategies have actually been translated to the clinic.

Cartilage to bone transition and chondrocyte phenotypic changes (e.g. hypertrophy) are observed during both bone formation and OA. Hence, controlling and modulating this transition may be key for various therapeutic applications. Promoting bone formation from cartilage should provide potent developmental TE strategies while inhibiting cartilage to bone transition might be paramount to find potent drug targets against OA. Intricate intracellular regulatory networks tightly control that transition. In this thesis, we develop a computational framework to explore the capacity of systems biology approaches to integrate data and knowledge in order to investigate intracellular biochemical processes, recognize important patterns and identify pivotal drivers of cell phenotypic changes in OA as well as for cartilage and bone regeneration strategies.

First, we established a theoretical basis by summarizing the scientific literature on both the biomedical and the computational engineering aspects necessary for this project.

Next, a data-driven model was built to provide an assessment tool for experimental OA disease models. It revealed that assembling heterogeneous transcriptomic datasets and carrying out a meta-analysis provided informative biological insights. A newly identified OA transcriptional signature served to train machine-learning models. Validation activities showed that the model confirmed the existence of phasic changes in OA development, at molecular level. It was a proof-of concept showing the potential of the modeling tool for providing quantitative metrics to assess OA status and treatment effects in experimental disease models.

Subsequently, mechanistic models were developed to study the regulatory mechanisms underlying chondrocyte phenotypic changes during OA and identify potential therapeutic targets. *In silico* experiments were done with the virtual chondrocyte to predict key drivers, in the network, potentially blocking hypertrophy. *In vitro* experiments in cell lines and primary chondrocytes validated the anti-hypertrophic potential of some of the predicted treatments. Notably, our study suggests that concomitantly activating PKA and inhibiting FGFR1 could synergize their effects against hypertrophy.

Next, we leveraged an ensemble model of chondrocyte hypertrophy to prioritize biochemical differentiation cocktails for iPSCs-derived chondrocytes in a bone TE application. The regulatory network structure and interaction weights were optimized computationally. By comparing the *in silico* predictions with independent experimental results, we explored the capacity of the model to inform the design of cell differentiation experiments for bone formation. In line with independent experiments, we confirmed that the combinatorial application of BMP4 ligands, WNT agonist and T3 hormone stimulates chondrocytes to follow the hypertrophic path more than any of the individual treatments. The *in silico* model also corroborated the *in vivo* data about the condition with the most bone forming potential. With that, we illustrate how computational modeling can foster the shift from exploratory to confirmatory experimental trials in regenerative medicine.

Finally, we extended the previous OA model by introducing mechanotransduction pathways in relation with inflammatory pathways. This computational framework shows predictive power for non-linear effects of mechanical and inflammatory inputs on cartilage extracellular matrix catabolism and anabolism and is being tested as an intracellular module for a multiscale mechanical model of the knee joint.

In conclusion, this PhD thesis illustrates how systems biology and computational modeling and simulation can assist the understanding of osteochondral pathologies and help the early process of drug target discovery as well as the design of experiments for cell differentiation in TE.

(Ducth version)

Bot- en kraakbeenweefsels zijn beide bestanddelen van het skelet, dat aangetast kan worden door intrinsieke en extrinsieke gebeurtenissen zoals ziekte of traumatisch letsel. In bot genezen ongeveer 90% van de defecten (bv. breuken) spontaan, maar voor diegene die niet spontaan kunnen genezen, zijn er weinig efficiënte regeneratieve oplossingen beschikbaar. Osteoartritis (OA) is de meest voorkomende aandoening in gewrichten, waarbij het kraakbeen progressief wordt afgebroken zonder zelfregeneratiepotentieel. Er zijn momenteel geen geneesmiddelen voor de behandeling van OA goedgekeurd voor gebruik bij patiënten, en slechts een beperkt aantal geavanceerde medicinale therapiën die het resultaat zijn van osteochondrale tissue engineering (TE) strategieën werden daadwerkelijk naar de kliniek vertaald.

De transitie van kraakbeen naar bot en de fenotypische veranderingen van chondrocyten (bv. hypertrofie) worden waargenomen tijdens zowel botvorming als OA. Daarom kan het controleren en moduleren van deze transitie de sleutel zijn tot verschillende therapeutische toepassingen. Het bevorderen van botvorming vanuit kraakbeen kan leiden tot effectieve TE strategieën, terwijl het remmen van de overgang van kraakbeen naar bot van cruciaal belang kan zijn om potente geneesmiddelen te vinden tegen OA. Complexe intracellulaire regulatorische netwerken controleren nauwgezet deze overgang. In dit proefschrift ontwikkelen we een computationeel kader voor het bestuderen van de capaciteit van systeembiologische benaderingen om data en kennis te integreren voor het onderzoek van intracellulaire biochemische processen, voor identificatie van belangrijke patronen en cruciale drijfveren van celfenotypische veranderingen in OA, alsook voor kraakbeen- en botregeneratiestrategieën.

Eerst werd een theoretische basis gelegd door een samenvatting te geven van de wetenschappelijke literatuur over zowel de biomedische als de computationele aspecten die noodzakelijk zijn voor dit project.

Vervolgens werd een data-gedreven model gebouwd om een beoordelingsinstrument te bieden voor experimentele OA-ziektemodellen. Het bleek dat het assembleren van heterogene transcriptomische datasets en het uitvoeren van een meta-analyse informatieve biologische inzichten opleverden. Een nieuw geïdentificeerde OA transcriptionele signatuur diende om machine-learning modellen te trainen. Validatie activiteiten toonden aan dat het model het bestaan van fasische veranderingen in de ontwikkeling van OA op moleculair niveau bevestigde. Het was een proof-of-concept dat het potentieel van het modelleringsinstrumenten

aantoonde voor het leveren van kwantitatieve meetgegevens om de OA-status en de behandelingseffecten in experimentele ziektemodellen te beoordelen.

Vervolgens werden mechanistische modellen ontwikkeld om de regulerende mechanismen te bestuderen die aan de grondslag liggen van de fenotypische veranderingen van chondrocyten tijdens OA en om potentiële therapeutische doelen te identificeren. *In silico* experimenten werden uitgevoerd met de virtuele chondrocyt om de belangrijkste drijvers, in het netwerk, te voorspellen die mogelijk hypertrofie kunnen blokkeren. *In vitro* experimenten in cellijnen en primaire chondrocyten valideerden het anti-hypertrofe potentieel van sommige van de voorspelde behandelingen. In het bijzonder suggereert onze studie dat het gelijktijdig activeren van PKA en het remmen van FGFR1 synergistisch hun effecten tegen hypertrofie zouden kunnen versterken.

Vervolgens hebben we gebruik gemaakt van een ensemble model van chondrocythypertrofie om biochemische differentiatiecocktails te prioriteren voor iPSC-afgeleide chondrocyten in een bot TE toepassing. De structuur van het regulatorisch netwerk en de gewichten van de interacties werden computationeel geoptimaliseerd. Door de *in silico* voorspellingen te vergelijken met onafhankelijke experimentele resultaten onderzochten we de capaciteit van het model om celdifferentiatie-experimenten voor botvorming te ontwerpen. In lijn met onafhankelijke experimenten bevestigden we dat de combinatorische toepassing van BMP4 liganden, WNT-agonist en T3-hormoon chondrocyten meer stimuleert om het hypertrofische pad te volgen dan elk van de individuele behandelingen. Het *in silico* model bevestigde ook de *in vivo* gegevens over de conditie met het meeste botvormende potentieel. Hiermee illustreren we hoe computermodellering de verschuiving kan bevorderen van verkennende naar bevestigende experimentele proeven in de regeneratieve geneeskunde.

Tenslotte werd het vorige OA-model uitgebreid met mechanotransductiepathways in relatie tot ontstekingspathways. Dit computationeel kader toont het voorspellend vermogen voor niet-lineaire effecten van mechanische en inflammatoire inputs op het katabolisme en anabolisme van de extracellulaire matrix in kraakbeen en wordt momenteel getest als een intracellulaire module voor een multischaal mechanisch model van het kniegewricht.

Kortom, deze doctoraatsthesis illustreert hoe systeembiologie en computationele modellering en simulatie kunnen bijdragen tot het begrijpen van osteochondrale pathologieën en het vroege proces van drug target discovery alsook tot het ontwerp van differentiatie-experimenten in TE.

List of abbreviations

ADAMTS	Disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif
AKT	Thymoma viral proto-oncogene 1
ALK1	Activin A receptor like type 1 (also TBRI)
ALK5	Transforming growth factor, beta receptor I (also TBRII)
ALP	Alkaline phosphatase
APC	Axin and adenomatous polyposis coli
ATF	Activating transcription factor
BMP	Bone Morphogenic proteins
CCND1	Cyclin D1
COL-II	Type II collagen
COL-X	Type X collagen
CREB	cAMP response element-binding protein
DAMPS	Damage-associated molecular patterns
DC	Destruction complex
DKK1	Dickkopf 1 homolog
DKK1	Dickkopf 1 homolog
DLX5	Distal-less homeobox 5
DMOAD	Disease modifying OA drug
DNA	Deoxyribonucleic acid
DSH	Dishevelled protein
ELK1	ETS transcription factor ELK1
ERK	Extracellular signal-regulated kinase
ETS1	ETS proto-oncogene 1, transcription factor
FAK	Focal adhesion kinase
FAK	Focal adhesion kinase
FE	Finite element
FGF	Fibroblast growth factors
FGFR	Fibroblast growth factors receptor
FOXO	Forkhead Box O
FRZB	Frizzled-related protein 1
FRZB	Frizzled-related protein 1
GLI2	GLI family zinc finger 2
GRN	Gene regulatory network
GSK3B	Glycogen synthase kinase 3 beta
GTP	Guanosine triphosphate
HDAC4	Histone deacetylases 4
HIF2A	Hypoxia inducible factor 2, alpha subunit
HP	Hypertrophic
HPC	High power computing
IGFBP	Insulin-like growth factor binding protein.
IGF-I	Insulin-like growth factor 1

IGF-IR	Insulin-like frowth factor 1 receptor
IHH	Indian hedgehog
IL-1B	Interleukin-1B
iNOS	Inducible nitrite oxide synthase
JNK	c-Jun N-terminal kinase
LEF	Lymphoid enhancer binding factor
LPS	Lipopolysaccharide
LRP5/6	Single-pass low-density lipoprotein receptor-related protein 5 or 6
LRP5/6	LDL receptor related protein 5
MAPK	mitogen-activated protein kinase
MEF2C	Myocyte enhancer factor 2C
ML	Machine learning
MLC	Myosin light chains
MMPs	Matrix metalloproteinases
mRNA	Messenger ribonucleic acid
MSX2	Msh homeobox 2
NFKB	Nuclear factor kappa B subunit
NIS	Network influence score
NKX-3.2	NK3 homeobox 2
NO	Nitric oxide
no-HP	Non hypertrophic
OA	Osteoarthritis
PI3K	Phosphatidylinositol 3-kinase
PKA	Cyclic adenosine monophosphate dependent protein kinase
PP2A	Protein phosphatase 2A
PPI	protein protein interaction
PPR	Parathyroid hormone receptors
PTHrP	Parathyroid hormone-related protein
ROCK	Rho-associated coiled-coil containing protein kinases
RT-qPCR	Rela time quantitative polyerase chain reaction
RUNX2	Runt-related transcription factor 2
SOCS	Suppressors of cytokine signaling
SOX9	Sex determining region Y (SRY)-box 9
STAT	Signal transducer and activator of transcription
TAK1	Transforming growth factor Beta-Activated Kinase 1
TAZ	Tafazzin
TCF7	Transcription factor 7
TE	Tissue engineering
TF	Transcription factor
TGF	Transforming growth factor
TNF-A	Tumor necrosis factor alpha
VEGF	Vascular endothelial growth factor
VSC	Vlaams Supercomputer Centrum
WNT	Wingless-type MMTV integration site family
YAP	Yes-associated protein

Content

Acknowledgement iii

Abstract v

List of abbreviations ix

Content xi

List of figures xiv

List of tables xvii

Chapter 1. General introduction 1

1.1 Bone and cartilage: structure and function in development and homeostasis. 4

1.1.1 Cartilage and bone work in concert: endochondral ossification 4

1.1.2 Hyaline articular cartilage 6

1.1.3 Bone and bone regeneration 7

1.2 Osteoarthritis (OA) pathology and therapeutic strategies 8

1.3 Non-healing fractures and bone tissue engineering (TE) 10

1.4 Signaling pathways controlling chondrocyte phenotype & cartilage homeostasis 12

1.4.1 Common chondrocyte signaling pathways 12

1.4.2 Inflammatory pathways 21

1.4.3 Mechanobiology and mechanotransduction in chondrocytes 22

1.5 Computational modeling and systems biology 23

1.5.1 Introduction on *in silico* models and concepts 23

1.5.2 Computational tools and methods for regulatory network modeling. 25

1.5.3 Modeling cell fate decision, chondrocyte (mechano-)biology and cartilage degradation and repair. 30

Chapter 2. Hypotheses objectives and methodology 37

2.1 General hypotheses and aim 38

2.2 Specific objectives 38

2.3 Experimental and numerical material and methods 41

2.3.1 Experimental methods 41

2.3.2 Computational and numerical methods 43

Chapter 3. Development of a machine learning model to assess OA signature in mouse disease models 45

3.1 Introduction 46

3.2 Results 47

3.3 Discussion 59

3.4 Material and Method 64

Chapter 4. Integrated *in silico-in vitro* approach for the identification of therapeutic drug targets for osteoarthritis 69

4.1 Introduction 70

4.2 Results 71

4.3 Discussion 87

4.4 Materials and Methods 92

Chapter 5. Regulatory network modeling applied to iPSCs derived chondrocyte differentiation for bone TE 101

5.1 Introduction 102

5.2 Description of the mathematical framework and optimization method 103

5.3 Establishing the numerical baseline situation: emerging chondrocyte phenotypes 109

5.4 Application to treatment optimization for iPSc-derived chondrocyte differentiation: comparing simulations with experimental results 112

5.5 Discussion 119

Chapter 6. Interplay of mechanoregulatory and inflammatory pathways: towards a multiscale view of articular cartilage homeostasis and degeneration...123

6.1 Introduction 124

6.2 Biochemical mechanisms and assumptions for the intracellular module 125

6.3 In silico model and its relation to the experimental and biological setup 136

6.4 In silico experiments to investigate the effect of perturbations related to meniscectomy *138*
6.5 Simulation results *140*
6.6 Discussion *145*

Chapter 7. Discussion and conclusion **149**

7.1 Project summary *150*
7.2 Main conclusions and contributions *152*
7.3 Perspectives and future research *159*

References **163**

Supporting information **211**

Scientific acknowledgement **249**

Curriculum vitae **251**

List of figures

Figure 1.1: Human knee joint 2

Figure 1.2: Stages and transcription factors involved in endochondral bone formation during mouse development 5

Figure 1.3: Histology of a normal articular cartilage in the knee joint and end-stage osteoarthritic joint.. 7

Figure 1.4: Summary of the WNT and TGF superfamilly pathways in chondrocytes and their relation to the chondrocyte hypertrophic phenotype in OA 13

Figure 1.5: Summary of the IHH/PTHrP signaling pathways in chondrocytes and their relation to the hypertrophic phenotype in OA 19

Figure 1.6: Overview of the MAP kinases pathways in chondrocytes 20

Figure 1.7: Comparison of approaches for regulatory network modeling 26

Figure 1.8: Schematic overview of the systems biology process and of the state space model. *In silico* modeling and wet lab experiments are part of the same iterative cycle in systems biology 28

Figure 1.9: Graphical definition of systems biology 31

Figure 2.1: Flow chart of the thesis content 40

Figure 3.1: Study Design 48

Figure 3.2: Illustration of the micro array preprocessing pipeline for array integration as a single dataset 49

Figure 3.3: Workflow for the model construction, validation and new sample query 50

Figure 3.4: GRN signatures and features importance 52

Figure 3.5: Model testing and performance estimation 54

Figure 3.6: Class predictions and transcription factor influence for independent query samples 58

Figure 4.1: Influence map of the signaling and gene regulatory networks (GRN) implemented in the model 73

Figure 4.2: Micro array integration for GRN inference 76

Figure 4.3: Predicted chondrocyte profiles and canalization for the three emerging states during Monte Carlo analyses 78

Figure 4.4: Study of the virtual chondrocyte state transition and *in silico* screening of target perturbations 82

Figure 4.5: *In vitro* validation of *in silico* predictions on chondrocyte phenotype changes 84

Figure 4.6: *In-silico* vs. *in-vitro* dose response effect of PKA activation with FGFR1 inhibition 86

Figure 5.1: Workflow used to assemble the ensemble models 106

Figure 5.2: Graphical representation of the qualitative activity profiles occurring during hypertrophy. 107

Figure 5.3: Illustrative examples of the scoring system for an expected None-High profile.. 109

Figure 5.4: Workflow used to simulate the ensemble dynamics..... 110

Figure 5.5: The unconstrained canalization for the 30 parameter sets..... 111

Figure 5.6: Effects of the experimental conditions on the *in silico* canalization for the ensemble model 115

Figure 5.7: Hypertrophic differentiation of cartilaginous organoids 117

Figure 5.8: Orthotopic implantation of human derived cartilaginous aggregates. Hypertrophically differentiated cartilaginous organoids were implanted in critical size long bone defects..... 118

Figure 6.1: Snaphshot of the mechanotransduction and pro-inflammatory signaling and their crosstalk in the *in silico* chondrocyte 131

Figure 6.2: schematic representation of the mechano-biochemical coupling in chondrocytes, as conceptualized and modeled *in silico*. Mechanical stimuli and changes in ECM mechanical properties affect the direct mechanical environment of chondrocytes, which is sensed by sensors such as integrins and TRPV4 calcium channels..... 135

Figure 6.3: Location of the mechanobiological intracellular layer within the multiscale workflow. 136

Figure 6.4: Computed basal profiles for the superficial zone and deep zone. Global activity of indicative variables or pathways in each stable state, for the SZ (left) and the DZ (right) zone. 141

Figure 6.5: Effect of meniscectomy, addition pro-inflammatory cytokines and their combination on SZ and DZ chondrocyte phenotype..... 142

List of tables

Table 1.1: Applying engineering principles, through Systems Biology approaches, to aid process design in developmental tissue engineering. 33

Table 2.1: Targets and associated small molecules or growth factors for the validation of *in silico* predictions 42

Table 3.1: Summary of top 15 genes in the OA GRN that are associated with relevant mutant phenotypes 53

Table 5.1: Growth factor activity profiles 106

Table 5.2: Formulas for computing the fitness function penalty scores 108

Table 6.1: Normalized mechanical input values 139

Chapter 1. General introduction

Partially based on:

Computational modeling and reverse engineering to reveal dominant regulatory interactions controlling osteochondral differentiation: potential for regenerative medicine. ***R. Lesage****, J. Kerkhofs, L. G. Geris*. *Front. Bioeng. Biotechnol.* **6**, 165 (2018),

Scientific and regulatory evaluation of mechanistic *in silico* drug and disease models in drug development: Building model credibility. *F. T. Musuamba, I. Skottheim Rusten,* ***R. Lesage****, G. Russo, R. Bursi, L. Emili, G. Wangorsch, E. Manolis, K. E. Karlsson, A. Kulesza, E. Courcelles, J. P. Boissel, C. F. Rousseau, E. M. Voisin, R. Alessandrello, N. Curado, E. Dall'ara, B. Rodriguez, F. Pappalardo, L. Geris*. *CPT Pharmacometrics Syst. Pharmacol.* (2021), doi:10.1002/psp4.12669

and

Turning Nature's own processes into design strategies for living bone implant biomanufacturing: a decade of Developmental Engineering. *I. Papantoniou, G. Nilsson Hall, N. Loverdou,* ***R. Lesage****, T. Herpelinck, L. Mendes, L. Geris*, r. *Adv. Drug Deliv. Rev.* **169** (2021), pp. 22–39.

The skeletal system is at the core of the body structural support and mobility. Its maintenance is important for human well-being and daily quality of life. This system includes 206 bones and 360 joints, covered with cartilage (**Figure 1.1**). Despite the importance of this system, various intrinsic and extrinsic events, such as diseases or traumatic injuries, may lead to impairment. For bone, about 90% of defects (e.g fractures) can heal spontaneously but few efficient regenerative solutions are available for the remainder of difficult to heal defects. For joints and articular cartilage, osteoarthritis (OA) is the most prevalent pathology (*1*), which has a large socioeconomic impact since the treatment-related cost is estimated around 1.5-2% of the gross domestic product in western countries (*2*). However, the exact mechanisms involved in OA are not fully understood, yet. As a result, therapeutic strategies to target the core of the disease are still lacking and too few advances from the last decades have been really successful in this domain (*3*).

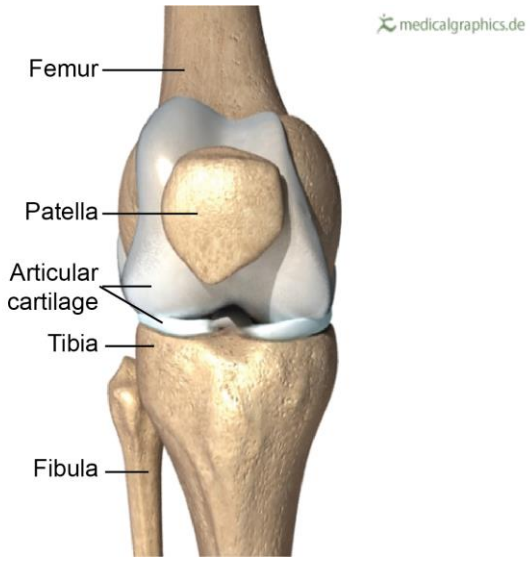


Figure 1.1: Human knee joint. Articular cartilage is a smooth tissue covering the end of the long bones (e.g. femur and tibia for the knee joint) and allowing load support and smooth joint movement without frictions. Figure adapted from www.MedicalGraphics.de, licensed under a creative commons attribution CC BY-ND 3.0.

Regenerative medicine consists in bringing back a tissue or an organ to its normal structure and function by replacing it, either via stimulation of natural repair or by producing new tissue in the lab. This second approach is called tissue engineering (TE). Normally, bones naturally form and regenerate via a cartilage intermediate through a process in which the chondrocyte (i.e. cartilage cell) enters a maturation cascade leading to chondrocyte hypertrophy and replacement by bone. Therefore, promoting chondrocyte maturation and bone formation is at the core of current regenerative strategies for bone fracture healing (*4, 5*). On the other hand, whist adult articular cartilage is a stable tissue in normal circumstances, the implication of chondrocyte

hypertrophy has been observed in pathology and is now considered as a contributing factor of OA too (*6*, *7*). Therefore, targeting and preventing the chondrocyte maturation process has emerged as a potential therapeutic strategy to treat osteoarthritis (*8*, *9*).

Controlling chondrocyte hypertrophy therapeutically requires a systemic understanding of the intracellular mechanisms involved since regulatory pathways and genetic networks govern cell states and cell fate decision. Nevertheless, cellular pathways interplay with each other in complex ways that involve extensive intertwining, feedback loops, non-linear routes and some redundancy. This makes it difficult to predict intuitively the exact role of one or another pathway or biological entity. When considering these biological systems as networks, they appear to share important structural features with other systems in engineering science. For instance, they share their modularity, robustness (*10*) and existence of recurring circuit elements such as feedback loops or mutual inhibitions (*11*, *12*). Therefore, formal mathematical representation may be key to investigate cellular systems numerically and unravel that complexity.

The accumulation of cartilage and bone related knowledge and data is a great opportunity for computational modeling approaches to tackle the skeletal pathologies-associated burden. Data scientists have stated that, while the past ten years have been about collecting data, the next ones will be about collecting less data and getting smarter about it (*13*). The focus of this thesis is on using computational modeling approaches to integrate complex knowledge and data related to chondrocyte intracellular regulatory mechanisms and mechanosensitivity to help identify new therapeutic strategies for preventing osteoarthritis or stimulating bone repair. This was done by identifying and reconstructing typical intracellular signatures as well as through the development and use of predictive *in silico* models. In this introduction, further information regarding bone and cartilage development, homeostasis and disease will be provided, along with an overview of current therapeutic strategies. Then, biochemical and signaling regulatory mechanisms underlying chondrocyte stability and hypertrophic changes will be detailed. This will be followed by an introduction to systems biology and computational modeling applied to cartilage systems, in general, and regulatory networks, in particular.

1.1 Bone and cartilage: structure and function in development and homeostasis.

1.1.1 Cartilage and bone work in concert: endochondral ossification

Bone forms in two different ways during embryonic development. The flat bones form via intramembranous (direct) ossification while most others forms through a process called endochondral ossification (EO) (*14*) (**Figure 1.2**). EO is defined as the process of bone formation from a cartilage template. Typically, during long bones development, progenitor cells from mesenchymal origin, traditionally called the mesenchymal stem cells or MSCs, are recruited and condensed into a cartilage anlage, surrounded by a connective tissue called the perichondrium (**Figure 1.2**). Progenitor cells first differentiate into chondroblasts and subsequently into chondrocytes, after what they stop proliferating and start maturating. This maturation is associated with a 20-fold increase in cell volume (*15*), increased VEGF production, a blood vessels attractant, and metalloproteinase (MMPs) production that degrade collagen fibers in the ECM (extracellular matrix). In other words, the chondrocytes become more catabolic and hypertrophic (**Figure 1.2)**. From the molecular perspective, this hypertrophic transition is associated with a switch in genetic programs marked by an increase in collagen type X alpha 1 (*COL10A1*) and runt related transcription factor 2 (*RUNX2*) expression whilst there is a loss of expression for collagen type II alpha1 (*COL2A1*) and SRY-Box transcription factor 9, 5,6 (*SOX9, SOX5, SOX6)*), among others differentially expressed genes (**Figure 1.2**). Hypertrophic chondrocytes contribute to the mineralization of the cartilage anlage, a precursor to its replacement by bone, thereby constituting the primary ossification center (**Figure 1.2**). The formation of bone by osteoblasts is facilitated both by external recruitment of cells and trans-differentiation of hypertrophic chondrocytes. The production, proliferation and maturation of chondrocyte remains active at the extremity of long bones where tightly regulated biochemical signals direct cellular organization in columnar gradients, a structure called the growth plate (*14*). While secondary ossification centers arise at the extremities of the long bone, the growth plates ensure the longitudinal bone growth by maintaining and gradually supplying a pool of differentiating chondrocytes throughout developmental and post-natal growth (*16*) (**Figure 1.2**). During this process, the perichondrium develops into the periosteum, which is a source of progenitor cells and plays an important role during growth and repair.

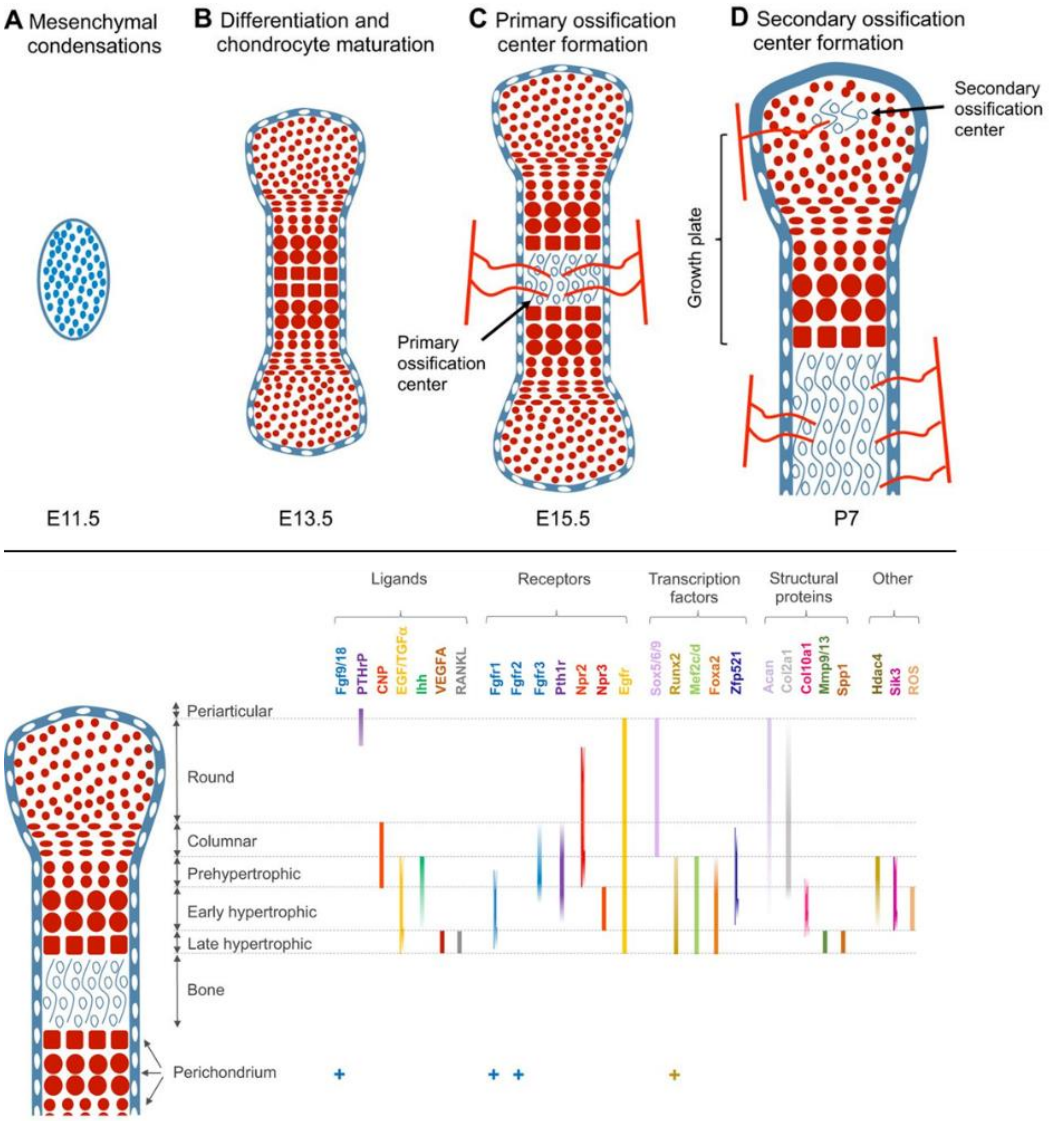


Figure 1.2: Stages and transcription factors involved in endochondral bone formation during mouse development. Representation of mesenchymal stem cells starting their condensation and differentiation into chondrocytes (**A** and **B**) . The columnar organization and succession of proliferative and hypertrophic chondrocytes push towards the formation of new bone tissue in the primary ossification center (**B** and **C**). The cartilage anlage is surrounded by a layer of perichondrium (white cells in blue layer) (**B, C** and **D**). Vascular invasion in that center contributes to the replacement of cartilage by bone (**C** and **D**). The growth plate structure remains during the postnatal growth **(D)**. Mouse embryonic stages are indicated below each stages (E11.5- P7). The mouse growth plate structure is characterized and controlled by differential gene expression profiles between the different cell phenotypes. *Runx2, Col10a1* and *Mmp13* are markers of hypertrophic chondrocytes while *Sox9* and *Acan* and *Col2a1* are rather expressed by round and proliferative chondrocytes. Modified from Kozhemyakina et al. (*16*).

During bone development, articular cartilage is formed in a separate process at the epiphyses of the bone in the synovial joint cavity. Even though some discrepancies exist between articular chondrocytes and chondrocytes from the growth plate, a certain amount of pathways that are paramount to control chondrocyte fate during development are also essential for articular cartilage (*17*). Contrary to the dynamical nature of the growth plate, articular chondrocytes

remain stable in a normal physiological environment; however, in some pathological situations and due to various environmental or genetic triggers, catabolic changes and a shift in chondrocyte phenotype have been observed, and is associated to unexpected cartilage to bone transition (*18*). These changes are believed to be reminiscent of the hypertrophic differentiation that takes place during skeletal embryonic development (*7*, *8*, *19*, *20*). The replication of the endochondral ossification process would notably explain the increased mineralization and osteophytes formation noted during osteoarthritis.

1.1.2 Hyaline articular cartilage

Hyaline (or articular) cartilage is a tough but resilient living tissue mainly found at the intersection between bones, covering the bony surface (**Figure 1.1**). Two other types of cartilage are found in the body, being the elastic cartilage, in the outer ear for instance, and the fibro cartilage, which is found in the intervertebral disc and knee meniscus. Nevertheless, most cartilage damages occur in the hyaline articular cartilage and result either from age and disease related degeneration or from traumatic injuries. This thesis focuses on the hyaline articular cartilage.

Articular cartilage presents itself in a layered structured, with the superficial zone at the joint interface, followed by the middle zone, deep zone and mineralized zone where it transitions into the subchondral bone (**Figure 1.3**). It is composed of one unique cell type, the chondrocyte. These cells are located horizontally to the joint surface in the superficial zone, and in columnar organization perpendicular to the surface in the deeper zones (**Figure 1.3**). The chondrocytes secretes and are surrounded by a rich ECM. This matrix is principally composed of water (60-85% by wet weight) and a complex and very dense network of biomolecules such as collagen fibers, proteoglycans (4-7% by wet weight) and oligomeric proteins (*21*). The main type of collagen fibers present in normal healthy cartilage are the type II collagen (15-22%), even though this so-called "type II" collagen is really a type II, IX and XI combo of fibers (*22*). Proteoglycans are protein macromolecules that are functionalized with glycosaminoglycans (GAGs), such as chondroitin sulfate (CS) in hyaline cartilage. Aggrecan is the major type of proteoglycan in articular cartilage (*23*), each aggrecan molecule consists of a 200-kDa core protein to which ~100 chains are attached (*24*). They provide to cartilage its hydrogel property, which is essential for the load-bearing and friction reduction functions of the joints (*25*, *26*). In

particular, the GAGs bear negative charges that contribute to water retention and to the formation of a viscous and elastic gel resistant to mechanical stress.

Moreover, an important particularity of cartilage tissue is the absence of vascularization, which results in a hypoxic environment for the cells; this is also one of the reasons this tissue has a limited repairing capacity.

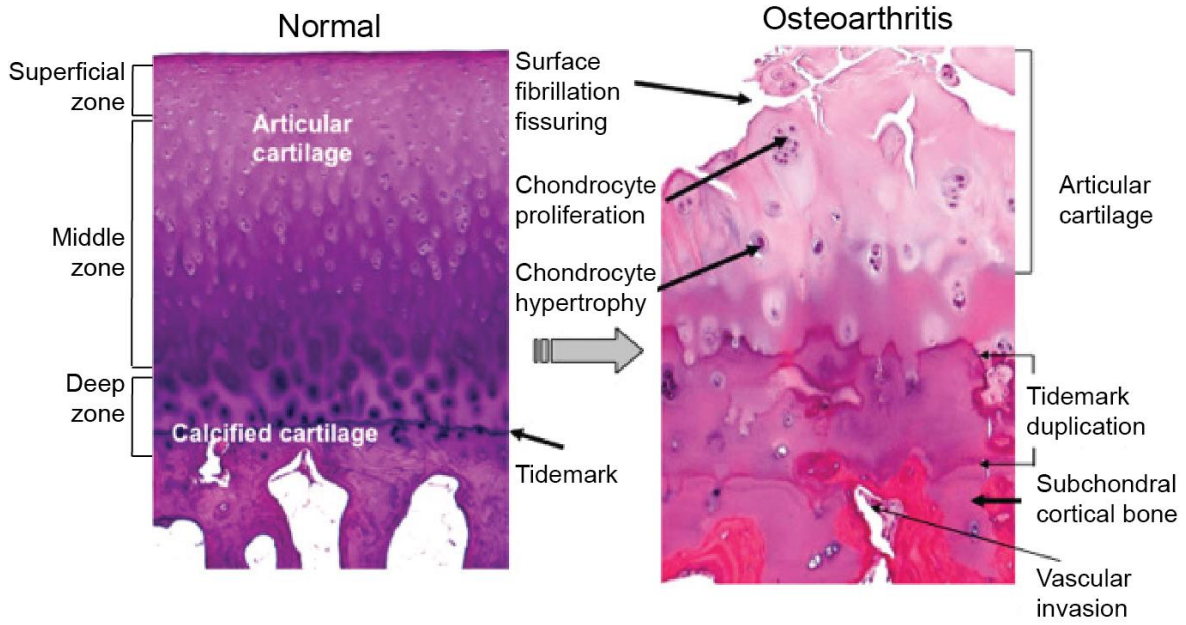


Figure 1.3: Histology of a normal articular cartilage in the knee joint and end-stage osteoarthritic joint. Cross section with hematoxylin and eosin staining. In the normal histology, chondrocytes are organized parallel to the surface in the most superficial layers while they are organized in column perpendicular to the surface in the middle and deep zone (above tidemark). Cartilage matrix is degraded in end-stage OA with the presence of hypertrophic chondrocytes. The tidemark separating the cartilage from the calcified (or mineralized) cartilage is duplicated. Figure adapted from Goldring et al. (27).

1.1.3 Bone and bone regeneration

As previously introduced, bones provide the body structure, stability and mobility capacity. In particular, long bones are composed of a periosteum enveloping the cortical hard bone in the center of which is locate the bone marrow, at the end of the bone are locate the epiphysis made of trabecular spongy bone that is covered by the smooth layer of articular cartilage. The rigidity of bone tissue makes it perfectly adapted to its function; however, it also implies a higher risk of fracture than for soft tissues in case of traumatic event. Bone fracture prevalence may be affected by various factors and it increases with age with an incidence of 50% for people over 65years.

As the bone is vascularized and alimented with stem cells, about 90% of bone defects can heal on their own (*28*). The natural healing mechanism functions in phases that involve inflammation, callus formation, endochondral ossification with attraction of blood vessels, and, bone remodeling via bone resorption by osteoclasts and formation by osteoblast.

In addition, the same markers that regulate EO during development seem to be expressed during the natural fracture healing process (*29*). It is now well accepted that similar process of bone formation, including cartilage to bone transition, are recapitulated to a certain extend and that EO is critical during bone fracture healing. This understanding has inspired regenerative treatments with the rise of developmental engineering (DE) strategies (*4*, *30*), which are further introduced in the section (1.4) Non-healing fracture and bone tissue engineering.

1.2 Osteoarthritis (OA) pathology and therapeutic strategies

With 7% of the global population and more than 33% of the people over the age of 65year affected by this condition, OA has been the most prevalent joint disease for several decades already (*31*). This prevalence increased drastically the last decade, especially in westernized countries, notably because of population ageing and increased obesity and it has greater incidence for women than men when considering the age related pathology (*1*). This is a major societal concern, not only because of the economic cost but also because this is a highly debilitating condition associated with disabling pain. Despite that situation, there is still no disease modifying OA drug (DMOA) currently available on the market for patients (*3*).

The identified causes of OA are manifold, comprising both intrinsic and extrinsic risk factors generally separated into traumatic and non-traumatic causes. Indeed, most cartilage damage results either from age and disease related degeneration, overweight or from traumatic injury (e.g. anterior cruciate ligament injury). A common ground is that alteration of the normal mechanical load on the joint plays a critically important role, in the sense that both hypo- and hyper-physiological loading are detrimental for cartilage. So much that exercise and weight loss are often part of the core treatment recommended by clinicians. Additionally, hundreds of deoxyribonucleic acid (DNA) variants have been associated with OA and these genetic risk variants would account for over 20% of the OA heritability (*32*). Nevertheless, a Lancet commission on OA recently reported that the misperception by patient and practitioners that OA is an inevitable part of ageing may be an important break for correct patient care (*31*).

From the histopathologic point of view (**Figure 1.3**), OA is a progressive disease characterized by a first silent phase followed by a stage with softening, thinning, fissuring and loss of cartilage tissue via alteration of the proteoglycan content and collagen fibers network (*33*). Remodeling also occurs at the level of the subchondral bone, which becomes thicker. At later stages, cartilage increasingly calcifies and the tidemark reduplicates at the interface between mineralizing and non-mineralizing cartilage (*33*) (**Figure 1.3**). In addition, penetration of the cartilage by blood vessels is observed in late stage OA, as well as osteophyte formation at the margin of the articular surface.

At cellular and molecular level, OA is characterized by a change in the balance between anabolic and catabolic processes leading to the overall degradation of the cartilage matrix. One of the drivers is the disruption of the normal chondrocyte resting state and the onset of inappropriate chondrocyte phenotypic changes, which lead to vascularization, activation of inflammatory pathways, matrix degradation, matrix mineralization and eventually replacement by bone. That chondrocyte maturation and hypertrophic-like phenotypic changes may occur locally and rather early in the course of OA (*19*, *34*). Following hypertophic changes, chondrocytes may direct themselves towards apoptosis or chondrocyte-to-osteoblast transdifferentiation may occur (*35*). We mainly focus on the hypertophic changes in this thesis. It is now well recognize that pro-inflammatory mediators and biomechanical factors are critically compelling factors in OA onset (*36*) and that the related inflammatory and mechanostransduction pathways are strongly intertwined. The implication of numerous signaling pathways in chondrocyte fate decision and OA is more extensively described in the section (1.4) *Cell signaling pathways controlling chondrocyte differentiation and cartilage homeostasis*, with a focus on chondrocyte hypertrophy.

Basic OA treatments focus on relieving the symptoms including with intra-articular injections, especially for pain. In extreme situations, surgery with total knee replacement may be the only way out. One of the main truly regenerative approach currently used to repair cartilage is the micro-fracture technique, which generates a marrow clot and progenitor cell invasion that eventually could lead to the production of new cartilage. Nonetheless, the newly formed cartilage often turns out to be not sustainable on the long term. A similar issue ensues from tissue engineering therapies where chondrocytes are taken from the patient, cultured in the laboratory and implanted as a cell-suspension (*37*) or combined with a biomaterial (*38*). Other therapies such as physiotherapeutic exercise or electromagnetic field applications have

provided some short-term relieve or improvement but generated no sustainable results and benefits (*39*).

Beside the aforementioned non-drug treatments, several drug-related tracks have been investigated to find efficient DMOADs. For instance, many studies investigated the use MMP inhibitors and aggrecanase inhibitors to prevent matrix degradation (*34*). Several trials have focused on hampering pro-inflammatory cytokines and downstream secondary messengers (*3*). In addition, other signaling pathways and changes in the proportion of various growth factors' membrane receptors are under scrutiny to propose therapeutic strategies preventing OA onset and progression. For example, clinical trials with recombinant fibroblast growth factor 18 (Sprifermin) gave interesting results but may require a better stratification of the target patient population (*40–42*). Nevertheless, most of those strategies have failed because of adverse effects and lack of significant efficacy, potentially due to the heterogeneity of the OA pathology (*3*, *42*). A search on clinicaltrial.gov for any drug interventions on knee osteoarthritis or cartilage returns 1082 past and current studies in the world, of which 226 are still recruiting or on-going in 2021, but this is not restricted to modifying drugs as it also includes pain relievers and bio-cement components.

1.3 Non-healing fractures and bone tissue engineering (TE)

For about 15% of bone fractures, the healing process is particularly delayed or even fails to fully bridge the fractured site (*28*); typically, clinicians consider a facture as non-healing when the bridging was not complete after up to 6-9months. Such a delayed or non-healing situation may arise due to very large (traumatic) fractures or surgically produced critical size defects, after tumor resection. The incidence increases with risk factors such as age or smoking. Bone fractures may also be due to (genetic) diseases such as diabetes or the congenital pseudarthrosis of the tibia, which leads to spontaneous tibia fractures in early childhood and requires multiple surgical intervention with an overall low success rate and a risk of amputation as the only way out (*43*, *44*).

Traditional and more recent regenerative treatment strategies for those patients involve mechanical fixators, bone graft from other part of the body or other donors, treatments with bone morphogenetic proteins (BMP) or graft with bone marrow enriched carriers, but all of them may rise donor site issues, various side effects and complications or a low efficacy. Bone tissue engineering, meaning the production of bone tissue or anlage in the lab via engineering

processes as a substitute for graft from donors, has emerged as a potent strategy in the last decades.

The generation of TE constructs generally relies on a triple helix including a cell source, a scaffold or biomaterial and a biochemical-stimulating cocktail (e.g. growth factors, bioactive molecules, etc..). The cells constitute the elementary bricks and can be either stem cells or already differentiated cells. The scaffold ensures the global structure and shape of the construct as well as certain mechanical properties, while the biochemical cocktail is critical to control the functionality, stimulation or differentiation of the cells. It is important to find the most appropriate biochemical signals for the cells, no matter the way they are provided (eg. via a functionalized scaffold or via the medium). In addition, the control of the overall culture conditions for TE constructs (i.e. temperature, medium flow or mechanical loading regimen) is also critical to achieve high quality tissue constructs. Nonetheless, despite decades of efforts put in the field of bone tissue engineering, few products successfully translated to the clinic (*4*). Some of the current strategies to overcome this issue rely on a shift from the top-down traditional TE approach towards a more bottom-up one, termed developmental engineering (DE).

Both traditional bone TE and DE aim to produce bone tissue constructs as an alternative to bone grafts. In both cases, the product would be categorized as an ATMP (advance therapy medicinal product) by regulatory agencies. However, the main difference lies in the fact that bone DE intends to do so by reproducing the natural developmental path as it is assumed that by providing the right inputs to the cells, based on known morphogen cues, the correct expected tissue structure may arise by self-organization of the cells and tissues. For bone TE this developmental path is that of EO, a well-studied process that serves as a gold standard to ensure process quality. As such, bone DE approaches focus on differentiation of progenitor cells into chondrocytes and further on towards hypertrophy, leading to the creation of a cartilage tissue intermediate that is capable of inducing EO once implanted *in vivo* (*45*, *46*). Despite successful proofs of concept (summarized in (*4*)), a major hurdle to clinical translation pertains to the scale of the tissue construct. Scaling up the size raises questions of the diffusion of biomolecules and nutrients as well as of the source and supply of the progenitor cells. On the one hand, the use of micro-organoid assemblies has emerged as a potential solution to tackle the diffusion problem (*5*). On the other hand, a cell population that is increasingly considered as a viable source is that of the induced pluripotent stem cells (iPSCs) because they may originate from the

patient own cells (autologous cells) and they have a capacity to differentiate into multiple cell types from the three basic embryonic layers, including skeletal cells (*47*, *48*).

1.4 Signaling pathways controlling chondrocyte phenotype & cartilage homeostasis

Given the focus of this thesis on chondrocyte regulatory and signaling pathways, hereafter we introduce important pathways related to the control and switch of chondrocyte phenotype. We start with common pathways followed by a focus on pro-inflammatory and mechanobiological pathways. More precisely, the description of the common pathways includes the Wingless-type MMTV integration site family (WNT), Bone morphogenic proteins (BMP), and Transforming growth factors (TGF) pathways, followed by a focus on important downstream transcriptions factors such as the Myocyte enhancer factor 2C (MEF2C), RUNX2 and SOX9. Then, the Fibroblast growth factors (FGF), Indian hedgehog (IHH), Parathyroid hormone-related protein (PTHRP), and Insulin-like Growth Factor 1 (IGF-I) pathways are introduced. The Mitogen-activated protein kinases (MAPKs) and Phosphatidylinositol 3-kinase (PI3K)/thymoma viral proto-oncogene 1 (AKT) axes are also discussed. The implications of those pathways on ECM proteins regulation is addressed too.

1.4.1 Common chondrocyte signaling pathways

WNT signaling pathway

The WNT pathway plays a major role in chondrocyte biology. First identified as an important factor during growth plate chondrocyte differentiation, the correct functioning of this pathway is also essential for articular cartilage homeostasis (**Figure 1.4**). Indeed, it plays a vital role in proliferation, stability and hypertrophy of chondrocytes as well as cartilage integrity, a role which depends on the exact WNT ligand involved (e.g. WNT3A, WNT5A, WNT5B) and the cellular context (*49–52*). Nonetheless, WNT was reported as a risk factor in osteoarthritis (*53*, *54*) and the WNT pathway was also shown to mediate cartilage matrix degradation (*55*). The typical WNT canonical pathway consists of the following steps. Upon ligand binding, a complex of the Frizzled and LRP5/6 receptors phosphorylates and activates the downstream Dishevelled (DSH) protein, which subsequently promotes the disassembly of a destruction complex (DC) (**Figure 1.4**). This complex includes several factors including the tumor suppressors Axin and adenomatous polyposis coli (APC), the glycogen synthase kinase 3 beta

GSK3B and CK1 and the protein phosphatase 2A (PP2A) (*56*, *57*). When WNT signaling is off, the DC catches cytosolic β-catenin and promotes their destruction via phosphorylation and ubiquitination. In presence of WNT ligands, the disassembly of the DC leads to B-catenin translocation to the nucleus, followed by activation of the lymphoid enhancer binding factor and the transcription factor 7 (LEF/TCF) (**Figure 1.4**). These transcriptional factors (TFs) regulate the expression of various WNT specific target genes, including *RUNX2* (*58*), a key player in chondrocyte hypertrophy. Finally, the activity of the WNT signaling is tightly regulated by crosstalks with other pathways as well as by soluble antagonists such as dickkopf 1 homolog (DKK1) and secreted Frizzled-related protein 1 (sFRP1 or FRZB), which were reported as potential brakes to chondrocyte hypertrophy and osteoarthritis (*59*, *60*) (**Figure 1.4**). In chondrocytes, a positive feedback loop downstream of the WNT pathway would decrease the transcriptional expression of the FRZB and DKK1 antagonists while increasing expression of LEF/TCF (*59*). In addition, treatment of human articular cartilage with WNT3A was reported to increase expression of Gremlin 1 (GREM1), a bone morphogenic protein (BMP) signaling antagonist. More precisely, BMP and WNT signaling upregulated the expression of each other's antagonists (*59*).

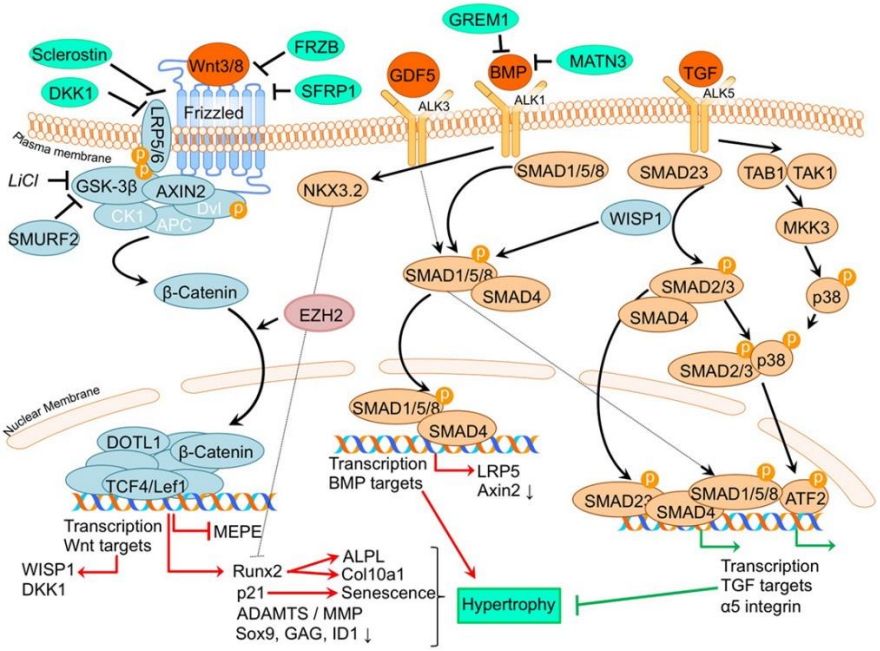


Figure 1.4: Summary of the WNT and TGF superfamilly pathways in chondrocytes and their relation to the chondrocyte hypertrophic phenotype in OA. At the external surface growth factors are indicated in orange while other secreted molecules are in green. In the cytoplasm, signaling paths downstream WNT (resp. TGF superfamily) signal are in blue (resp. in light orange). In the nucleus pro-chondondrogenic relations are represented by green arrows while pro-hypertrophic relations are in red. Figure reproduced from Ripmeester et al. (*61*).

BMP signaling

Bone morphogenic proteins (e.g. BMP2, BMP4, BMP6) are important regulators of chondrogenesis and skeletogenesis during embryonic development and of homeostasis in articular cartilage (*62*). The exact role of BMPs is not yet fully identified since they are considered to be protective of articular cartilage while they can also be involved in hypertrophic differentiation (*63–66*). Indeed, overexpression of BMP2 has been found in osteoarthritic tissues. BMPs bind and transduce signals through Activin-receptor like kinases (ALKs) (**Figure 1.4**). Upon binding a cascade of phosphorylation activates the receptor activated SMADs proteins (R-SMADs), being SMAD1/5/9 for mice or SMAD1/5/8 for humans. These R-SMADs associate with SMAD family member 4 (SMAD4) to form a complex (SMAD-complex) able to translocate to the nucleus and regulate expression of target genes (*67, 68*). Inhibitory SMADs, SMAD6 and SMAD7, are able to modulate BMP signaling via several mean: they may competitively bind with the receptors to prevent R-SMAD phosphorylation and activation, they prevent R-SMAD translocation to the nucleus and they may promote R-SMADs degradation via ubiquitination (*69, 70*). Another important non-canonical pathway that is also activated by BMPs is the p38 MAPK related pathway (*62, 71*).

Distal-less homeobox 5 (DLX5), activating transcription factor 2 (ATF2) and MEF2C are important TFs activated downstream of these BMP mediated pathways. Overall, the BMP signals have an activator effect on runx and related transcriptional programs. This happens either in a direct way by the upregulating the *RUNX2* promoter by MEF2C (*72, 73*) and DLX5 (*74–76*), or indirectly via DLX5 binding and blocking of Msh homeobox 2 (MSX2), a RUNX2 antagonist in chondrocytes (*77, 78*). ATF2 being located at the crosstalk of several pathways, its exact role most likely depends on the overall cellular state, the nature of external signals and the status of collaborating factors (*79, 80*).

TGF-B signaling

The transforming growth factors are a super family of which the TGF-B1 or TGF-B3 family members are thought to be involved in cartilage function and homeostasis (*62*). Indeed TGF-B may preserve cartilage integrity while loss of TGF-B signaling induces chondrocyte hypertrophy resulting in cartilage degradation (*70, 81, 82*). TGF-B members act through type I and type II receptors (TBRI & TBRII) at the cell surface that transduce intracellular signals via SMAD3, SMAD complex or the non-canonical-SMAD-independent signaling pathway (that is

the p38/MAPK axis). Even though TBRI and TBRII work as homo and heterodimeric complexes, each type of receptor has been associated with different and specific downstream cascades. For example, TBRII, or ALK1, more specifically induces the SMAD1/5/8 activation while TBR-I (or ALK5), rather induces SMAD3 activation (*62*) (**Figure 1.4**). Phosphorylated SMAD2/3 and SMAD1/5/8 having opposing functions regarding stable cartilage maintenance (*83, 84*), the change in the balance between ALK1/ALK5 receptor, which was observed in ageing population and during OA, could be one of the drivers of OA onset. (*70, 85–88*). Indeed, ALK5, through SMAD2/3, transduces rather pro-anabolic signals, as it is able to promote transcription factor 2 (ATF2) activity, inhibit RUNX2 activity and subsequent *COL10A1* expression (*89, 90*). As a result, loss of ALK5 leads to OA development in mice (*91*) and the TGF-β/SMAD3 signals prevent chondrocyte hypertrophy (*92*). Whereas, ALK1 transduces pro-catabolic signals with the downstream up-regulation of MMP13 expression, through SMAD1/5/9 (*86, 93*).

In addition, TGF-β non-canonically activates the p38 MAPK related pathway and the downstream ATF2 transcriptional effector. ATF2 and SMAD3 were reported to cooperate and amplify the TGF-β signal in chondrocytes (*89, 94*), which has a direct impact on the rate of chondrocyte maturation (*79*).

The MEF2C transcription factor

Expressed in (pre)hypertrophic chondrocytes, MEF2C is necessary for the normal *COL10A1* promotor activity (*95*), it also functions upstream of RUNX2 and DLX5 (*73*). Repression of MEF2C activity by histone deacetylases 4 (HDAC4) represses *RUNX2* expression (*73*, *96*), hence another mechanism of repression of hypertrophy would work via the PTHrP/HDAC4/MEF2C axis (*97*). In addition, MEF2C is regulated downstream of TGFβ/SMAD pathways and it transcriptional activity can be enhance via phosphorylation by P38 (*98*, *99*).

The RUNX2/SOX9 duo

RUN2 and SOX9 are pivotal TFs for chondrocyte proliferation, maturation and stability both during endochondral ossification and during health and disease of adult cartilage (*69, 100–102*). RUNX2 and SOX9 physically interact and have reciprocal inhibitory actions; they are antagonist factors that control opposite programs (*17, 103–105*). This negative feedback loop, together with the existence of auto-activation loops might be at the root of chondrocyte

phenotype bi-stability. On the one hand, RUNX2 (or Cbfa1) is the master TF for chondrocyte hypertrophy and deletion of RUNX2 was show to reduce the progression of osteoarthritis (*17*, *106*). On the other hand, SOX9, which acts mostly redundantly with SOX5 and SOX6 (*107*, *108*), is known to delay progression towards hypertrophy (*109*, *110*) to keep cartilage healthy and prevent osteoblastic transition (*111*). These two master TFs, RUNX2 and SOX9, integrate multiple biochemical signals and control the activity of other TFs as well as the expression of many target genes, from extracellular matrix proteins to signaling molecules. To quote few, SOX9 is able to block the WNT/B-catenin axis by binding to LEF/TCF and preventing their activation by B-catenin (*112*, *113*); SOX9 also collaborates with NK3 homeobox 2 (NKX-3.2) to up-regulate *COL2A1* and downregulate *RUNX2* expressions (*69*, *114*).

Another molecular switch between the TFs MSX2 and DLX5 happens in parallel, in skeletal cells. It is tightly controlled by the RUNX2 program and might contribute to the existence of different chondrocyte phenotypes. Indeed MSX2 and DLX5 inhibit each other by direct binding, MSX2 antagonizes RUNX2 activity and expression while DLX5 rather is an activator of RUNX2 both at the functional and the transcriptional level (*76*, *77*, *115*). Most of these functions were observed in osteogenic cells or during endochondral ossification (*116*, *117*); nevertheless, those factors are also known to be expressed in (hypertrophic) chondrocytes (*118*). Their role in articular cartilage being less known, it deserves further investigation as they were associated with OA in genome wide expression analyses (*119*).

FGF signaling

Fibroblast growth factors and related pathways are essential for skeletal development and cartilage homeostasis. Several types of ligands, such as FGF2, FGF18 or FGF1, are expressed by chondrocytes and may be grouped in several sub-families (*120*). Each ligand or sub-family of ligands preferentially binds to a different type of receptors (*120–122*). FGF receptor 3 (FGFR3) and FGF receptor 3 (FGFR1) are the major FGF receptor isoforms expressed in human articular cartilage that have opposing effects and the modification of the balance between the two, notably during ageing, might contribute to OA by promoting chondrocyte hypertrophy (*121*, *123*, *124*). Indeed, FGFR1 is upregulated in hypertrophic chondrocytes and loss of FGFR1 reduces hypertrophy (*121*, *125*). In addition, recent studies have suggested that inhibitors of FGFR1 can protect against cartilage degradation during OA and inflammation-induced catabolic effects (*123*, *124*, *126–128*). On the contrary, FGFR3 was shown to limit the progression of knee OA in adult mice (*129*), while conditional deletion of *Fgfr3* was proven to

induce OA in mice (*130*). This effect might be due to the ability to trigger downstream pathways that downregulate *Runx2*, *Ihh* and *Mmp13* expression (*129*, *131*). Currently Sprifermin, a recombinant human FGF18, which is believed to preferentially bind to FGFR3 (*132*), showed promise with positive effects such as increased cartilage thickness and little adversity, and is currently is under a phase II clinical trial (*40*, *41*).

FGF2 signals both through FGFR1 and FGFR3 and its exact role in the joint is controversial (*131*, *133*). FGF2 and FGFR1 are upregulated during OA and this is associated to a decrease in type II collagen as well as an increase in *MMP13* expression (*134*). Conversely, in normal cartilage FGF2 treatment increases FGFR3 but not FGFR1 expression and is associated with the upregulation of COL2A1, and suppressors of cytokine signaling (SOCS), through the STAT signaling (*135*). This illustrates how the transition in the ratio between the two types of receptors may define the pro-catabolic or pro-anabolic nature of FGF signaling, a transition that could be initiated by pro-inflammatory signals since IL-1B treatment increases *FGFR1* and decreases *FGFR3* expression (*134*).

Finally, the Signal transducer and activator of transcription (STAT), the MAPKs or the Nuclear Factor Kappa B Subunit (NFKB) axes are typical downstream signaling pathways that were reported to transduce FGFs signals down to the nucleus (*136*, *137*).

ECM proteins

The extracellular matrix of articular chondrocytes is principally composed of water (60-85% by wet weight). The remainder is composed of a complex and very dense network of collagen fibers (mainly type II collagen 15-22%) and proteoglycans (4-7% by wet weight) (*21*).

Many pathways and transcriptional programs converge towards the transcriptional and biosynthetic regulation of ACAN, PRG4 (Aggrecans), COL2A1 and other ECM related proteins, in chondrocytes. For instance, Aggrecan expression is upregulated by SOX9 (*138*) and FOXO (*139*) in articular cartilage, and pro-inflammatory signals were reported to reduce collagen production (*6*, *140*).

On the catabolic side, hypertrophic or pro-inflammatory cytokines treated chondrocytes produce and secrete proteases such as MMPs (MMP9, MMP13), and aggrecanases such as Disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif (ADAMTS) 4, ADAMTS5 and ADAMTS9 (*6*, *34*, *141*). Those enzymes, cleave other ECM

proteins, thereby contributing to ECM degradation, replacement of type-II collagen by type-X collagen and matrix mineralization (*142*). This process happens naturally and is necessary for bone formation during endochondral ossification; however, it is highly detrimental when happening in articular cartilage and constitute one of the hallmarks of OA. Moreover, following ECM proteins cleavage, some matrix degradation products called damage-associated molecular proteins (DAMPS) are formed (*143*). Some studies have reported the ability of these DAMPS to bind to pro-inflammatory cytokines receptors, thereby activating inflammatory like responses and contributing to the positive pro-inflammatory feedback loop (*143–145*). Nevertheless, past attempts to inhibit MMPs as a potential OA therapy showed unsatisfactory results, partly due to the difficulty of generating specific enough inhibitors (*146*).

IHH/PTHrP

The role of Indian hedgehog and parathyroid hormone-related protein during bone and cartilage development has been studied extensively over the past decades (*147–149*). IHH, more highly produced by (pre)-hypertrophic chondrocytes, up-regulates PTHrP that subsequently blocks chondrocyte maturation. This feedback loop regulates bone length and ensures the persistence of a pool of proliferative chondrocytes during growth. As a result, disruption of this axis leads to dwarfism. Their role appears to be pivotal in adult articular cartilage too (*150*).

The GLI family zinc finger 2 (GLI2) transduces IHH signal in the nucleus, mainly leading to the upregulation of pro-catabolic genes (**Figure 1.5**). *IHH* is upregulated in hypertrophic chondrocytes by RUNX2 (*150*, *151*) and would be involved in OA. As such the IHH pathway has been investigated for its therapeutic potential for OA (*152*).

PTHrP activates the parathyroid hormone receptors (PPRs) that trigger a cascade of phosphorylation in chondrocyte. A major kinase involved in that cascade is the cyclic adenosine monophosphate dependent protein kinase (PKA) (**Figure 1.5**). Pharmacological inhibition of PKA was shown to block the chondrogenic differentiation of limb bud micromass cultures, showing its importance for embryonic chondrocyte maturation (*100*, *153*). The key role of PTHrP signaling in the maintenance of cartilage homeostasis and inhibition of hypertrophy (*154*) could come from the ability of PKA to directly bind and activate SOX9 (*155*) and ATF4 (or cAMP response element-binding protein (CREB) 2). PKA activation also promotes the activity of the PP2A, a blocker of the PI3K/AKT axis and an activator of HDAC4 (*97*). Finally,

downregulation of hypertrophic genes such as *RUNX2* and *COL10A1* can be orchestrated by the PTHrP/PKA pathway (*156*, *157*).

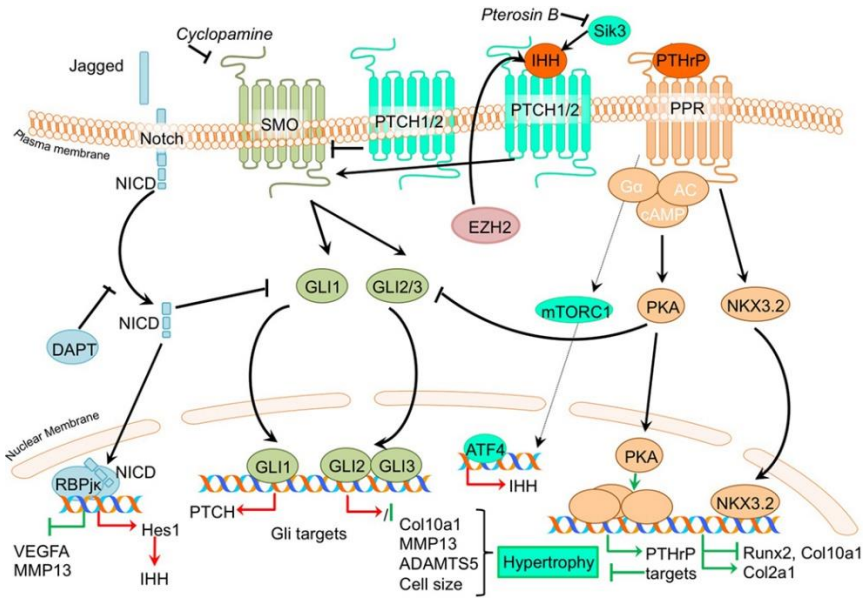


Figure 1.5: Summary of the IHH/PTHrP signaling pathways in chondrocytes and their relation to the hypertrophic phenotype in OA. GLI transcription factors are the main transducers or IHH signal while PKA, ATF4 and NKX3.2 may transduce the PTHRP signal in the nucleus. In the nucleus pro-chondondrogenic relations are represented by green arrows while pro-hypertrophic relations are in red. Figure reproduced from Ripmeester et al. (*61*).

RAS-ERK1/2 /JNK /p38 MAPK pathways

MAPK signaling is responsible for the conversion of many entering signals into cellular responses. Several cascades of phosphorylation happen in parallel for the message to propagate in chondrocytes, thereby controlling chondrocyte biology (*98*, *158*). For instance, the RAS proteins are Rho Guanosine triphosphate (GTP)ases (*159*) able to activate extracellular signal-regulated kinase1 and 2 (ERK1/2). These ERK kinases are located at the crossroads of a large number of pathways such as the PTHrP, BMP or FGF pathways, and ultimately control the activation or inhibition of downstream TFs (*55*, *160*, *161*). Beside ERK1/2, the P38 kinase and jun proto-oncogene (cJUN) terminal kinase (JNK) are the other most reported downstream enzymes of the MAPK cascades in cartilage. All together they converge on the control of TFs such as ETS proto-oncogene 1, transcription factor (ETS1), ETS transcription factor ELK1 (ELK1), ATF2 and cJUN. Notably, these signaling routes are known to induce the production of pro-catabolic enzymes (MMP13) and pro-inflammatory cytokines such as Tumor necrosis factor-alpha (TNF-A) and interleukin-1Beta (IL-1B). For that reason, they have been considered as potential therapeutic targets for osteoarthritis (*158*, *162*).

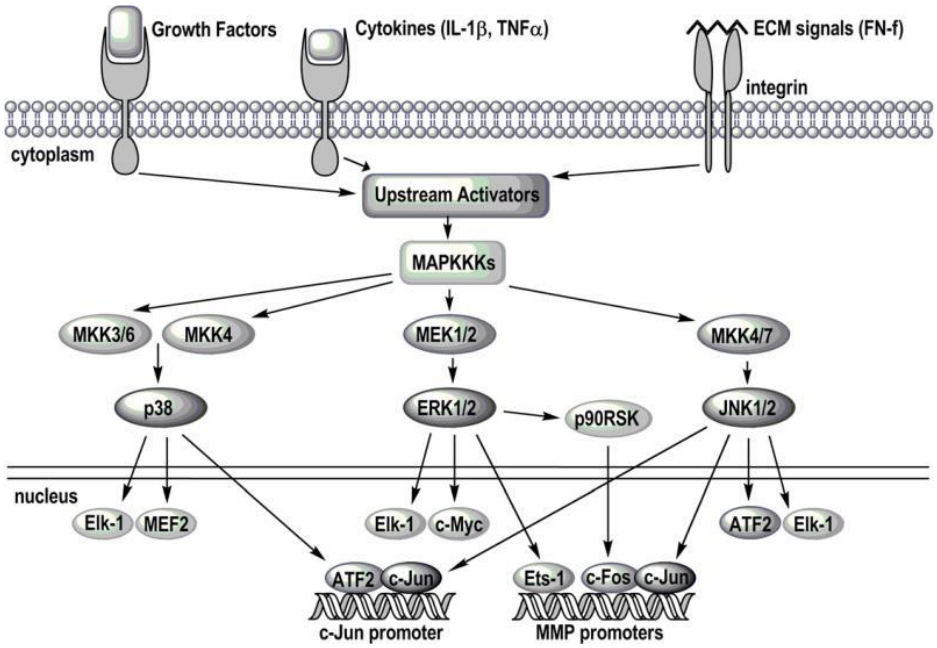


Figure 1.6: Overview of the MAP kinases pathways in chondrocytes. Various growth factors and integrin receptors may converge towards the activation of the MAPK pathway. A cascade of kinase activation starts with MAP kinase kinase kinases (MAPKKKs) phosphorylating MAP kinase kinases (e.g. MKKs, MEK), which in turn activates MAP kinases such as P38, ERK1/2 and JNK. These kinases may activate other kinases or transcription factors (e.g. ATF2, ELK1, c-JUN), translating the signal in the nucleus by binding and regulating promoters of target genes. Figure adapted from Loeser et al. (*162*)

The IGF-I / IGFBP and PI3K/AKT axes

Insulin-like Growth Factor 1 is a pro-anabolic growth factor secreted by chondrocytes that is known to promote cartilage matrix production and dysregulation of this pathway might be related to OA (*163*). It is involved in a complex cell stage dependent anabolic effect during development (*164*) and is also expressed in articular cartilage (*165*). Even though, upregulation of IGF-I was observed in OA (*166*) patients, the IGF-I activity would rather depends on the bioavailability of the ligand rather than the production rate. Indeed, some IGF binding proteins (IGFPB) were also reported to be upregulated in osteoarthritic chondrocytes by 16 to 25-fold in comparison to normal chondrocyte whereas IGF-I was only increased 3.5 fold (*166*). These binding proteins secreted in the ECM, may sequestrate IGF-I ligands and prevent the activation of downstream signals (*167*). Moreover, Olney at al. observed that TNFA and IL-1B increased IGFPB-3 production in human articular cartilage (*166*). The receptors of IGF-I, IGF-IR, show high affinity for their ligands and act as docking proteins to transduce the downstream signal thanks to a tyrosine kinase activity. It can signal through different canals including the JAK/STAT1, the PKA and the PI3K/AKT axes.

The mechanism of PI3K dependent activation of AKT lies at the crossroad of several signaling routes; therefore, its exact implication in cartilage homeostasis is controversial and may vary with the cellular context (*168*). On the one hand, AKT is able to phosphorylate and inhibit GSK3B, which otherwise can phosphorylate and promote proteasome degradation of Cyclin D1 (CCND1) and HDAC4, two blockers of RUNX2 activity. On the other hand, increase of pro-inflammatory cytokines during OA would activate that axis while inhibition of the PI3K/AKT axis was shown to decrease the inflammatory response and OA development in articular chondrocytes (*169*, *170*).

1.4.2 Inflammatory pathways

Chondrocyte express a large panel of receptors of pro-inflammatory cytokines, such as toll-like receptors (TLRs) or the receptor for advanced glycation end products (RAGE), making them sensitive to pro-inflammatory signals (*171*). Typical pro-inflammatory stimuli are triggered by IL-1β, TNF-A, lipopolysaccharide (LPS) or interferon regulatory factors and the downstream activated pathways eventually promote MMP production and cartilage matrix degradation (*141*, *172–175*). Beside protease production, pro-inflammatory cytokines trigger the activation and downregulation of several important expression programs, thus they may be responsible for the chondrocyte catabolic shift in OA (*176*, *177*). Several studies have reported that upon treatment with IL-1B, proteoglycans and GAGs are not only degraded but an overall decrease of the biosynthesis is observed in cartilage (*178*). Overall, the inflammatory process functions as an auto-activation loop since an initial pro-inflammatory stimulus results in the up-regulation of further pro-inflammatory mediators such as inducible nitrite oxide (NO) synthase (iNOS), prostaglandins, cyclooxygenase-2 (COX-2) or the IL-1B and TNF-A cytokines themselves (*179–181*).

One of the main effector pathways for inflammatory signal transduction and cartilage degradation comprises NFKB, a transcription factor, as a central actor (*182*). More precisely, pro-inflammatory cytokines binding to the cell leads to the recruitment and activation of several kinases culminating in the activation of TGFB-Activated Kinase 1 (TAK1) (*183*). In turn, TAK1 activates the nuclear factor B (IKB) kinase 2 (IKK2) that phosphorylates IKBA, the inhibitor of NFKB, thereby fixing it on proteasome degradation. When the inhibitor of NFKB is degraded, this transcription factor is free to translocate to the nucleus and activated downstream targets genes such as *MMP13*, Hypoxia inducible factor 2, alpha subunit (HIF2-

A*)*, *IHH* or *IL-1B* (*19*, *173*, *177*, *184*, *185*). Other than NFKB, the inflammatory signal transduced by TAK1 can continue through the activation of various MAPKs such as P38, JNK and ERK1/2. Inhibition of TAK1 in a post-traumatic OA rat model efficiently blocked NFKB, JNK and p38 pathways activation (*183*). Eventually, these MAPK pathways converge on the control of transcription factors including ETS1, ELK1, ATF2 and c-JUN (*162*). Overall, these inflammatory pathways converge on the up-regulation of the *MMP13* gene expression in chondrocytes (*186*).

There exist a number of natural breaks to the aforementioned inflammatory loop. For instance, the genetic expression of IKBA, the inhibitor of NFKB, is up-regulated by NFKB itself, constituting a typical feedback control mechanism. In addition, p38-mediated phosphorylation of TAB1 inactivates TAK (*177*, *187*). Moreover, healthy chondrocytes produce some suppressors of cytokine signaling (e.g. SOCS1, SOCS3), which are soluble inhibitors of pro-inflammatory cytokine receptors (*182*). To quote some pathways' complex interplay, SOCS induction has been observed downstream of the IGF-I/JAK/STAT axis, while these soluble inhibitors would also negatively regulate the JAK/STAT pathway in cartilage and during osteoarthritis (*188*). Finally, the Forkhead Box O (FOXO) transcription factor plays a critical role in articular cartilage homeostasis during ageing, excess mechanical loading but also against inflammatory cues. FOXO1 and FOXO3 expression and activation are dysregulated in OA and ageing human cartilage (*189*). In addition, the overexpression of *FoxO1* in mice OA chondrocytes was shown to reduce inflammatory mediators and cartilage-degrading enzymes (*139*). In contrast, evidence suggests that IL-1B and TNF-A indirectly suppress *Foxo1* expression (*189*, *190*). In addition FOXO is located downstream the MAPK cascades in chondrocytes (*191*).

1.4.3 Mechanobiology and mechanotransduction in chondrocytes

Mechanosensing and mechanotransduction describe the capacity of cells to sense and translate mechanical cues from their direct environment into intracellular biochemical messages. Chondrocytes are highly mechanosensitive and mechanical forces are paramount for cartilage homeostasis. However, there exists an optimal range of loading intensities and frequencies since hyper-physiological load is detrimental for cartilage and accentuates matrix degradation while hypo-physiological load might be suboptimal for the chondrocyte to produce physiological ECM (*178*). Mechanical cues are integrated through mechanosensors such as integrin (*192*), or

ion channels such as TRPV4 calcium channels (*193*). Those sensors translate physical cues into biochemical signals inside the cells. The internal machinery that is recruited downstream of the integrins is diverse and includes focal adhesion complexes with the focal adhesion kinases (FAK). The signal is then transduced through pathways involving RhoA GTPases and Rho-associated coiled-coil containing protein kinases 1 & 2 (ROCK1/2) as well as the actin cytoskeleton. Several transcription factors have been identified as potential downstream effectors of the mechanical signal. For instance, YAP/TAZ and B-catenin are recognized as mechanosensitive TFs (*194*, *195*). Overall, mechanoregulatory pathways control the expression of many genes including constitutive ECM proteins (proteoglycans, collagen) and ECM degrading enzymes (e.g. MMPs, ADMATS).

1.5 Computational modeling and systems biology

This section is partly based on Computational modeling and reverse engineering to reveal dominant regulatory interactions controlling osteochondral differentiation: potential for regenerative medicine (*196*), Scientific and regulatory evaluation of mechanistic *in silico* drug and disease models in drug development: building model credibility (2) and Turning Nature's own processes into design strategies for living bone implant biomanufacturing: a decade of Developmental Engineering (*4*). After providing a general introduction on *in silico* models and concepts, we focus on the different technologies for intracellular networks modeling in general, with the distinction between mechanistic and data-driven technologies. Then we expand on specific applications related to skeletal disease and regeneration. In particular, we discuss modeling applications for cell differentiation and tissue engineering, about models of chondrocyte and cartilage degradation and repair, as well as mechanobiological models coupling mechanics with intracellular regulatory networks

1.5.1 Introduction on *in silico* models and concepts

Systems biology is the science of analyzing holistically biological entities (i.e. considering individual entities and their interrelations as part as the whole, possibly at different scales of organization) thanks to mathematical and computational models and methods. Computational or *in silico* models are abstract and simplified representations of biological systems, composed of a set of rules describing the system's behavior and of algorithms. They are bio-mathematical models that are implemented and studied computationally. According to the mathematical nature of the rules, the behavior of the system can be studied over time and/or space and in a

quantitative or qualitative manner. The term *in silico* refers to silicon, the main component of computer chips, and hence the computational nature of the model, similar to the nomenclature of its *in vitro* and *in vivo* counterparts (*197*). Thus, it is a generic term referring to a broad scope of computational bio-models such as physics-based models, quantitative systems pharmacology, system medicine, virtual physiological human models, etc.

Computer modeling and simulations offer the possibility of testing hypotheses, acting on individual parts of the system, measuring effects and performing experiments that would otherwise not be possible in *in vivo* or *in vitro* systems. In that sense, it comes in complement and refinement of the experimental approaches. A wide spectrum of techniques exists ranging from the so called 'white box' models to the 'black box' models, passing by grey-scales. White box models denote completely comprehensive and mechanistic models based on known and observed physicochemical rules and mechanisms of actions (**Figure 1.7**). In contrast, black box models are built to match the observation content of data but bring poor mechanistic understanding of the underlying causation, making it more challenging to interpret. By extension, grey box models lay in between the fully mechanistic (or bottom-up) and fully data-driven (or top-down) models.

Even if an *in silico* model is meant to serve for research and is not going to be used in a regulatory evaluation or industrial setting, minimal requirements should be followed to ensure the model credibility. The general modeling workflow to develop a fully validated and credible computational model typically includes the following steps. It starts with data and/or information and knowledge acquisition and pre-processing. The next step is the model definition (e.g. writing the equations) and meticulous annotation, potentially followed by a calibration or optimization step (*198*). Then the equations, the codes and algorithms should be verified for correctness (*199, 200*). Generally, simulations and all *in silico* predictions and analyses arrive after that verification step. Finally, the model's results must be validated using independent data that were not used when building the model (*198–200*). The validation of an *in silico* model to answer questions in a specific context may be achieved in two ways, by showing its ability to predict or reproduce effects that were previously reported in the literature and/or by confirming newly made predictions with new, original experiments (*198*). Normally, this general workflow may apply for any modeling method employed (e.g. mechanistic or data driven) but, of course, each modeling approach comes with its particular specificities.

We will first discuss about tools and methods for intracellular network modeling in 1.5.2, followed by a more general overview of models in skeletal disease and regeneration in 1.5.3.

1.5.2 Computational tools and methods for regulatory network modeling.

A regulatory networks often represent biochemical regulatory pathways (e.g. protein signaling, gene regulation, metabolic pathways, etc.) by means of (1) nodes or vertices for the biological entities and (2) edges linking different nodes for the interrelations (physical interactions, reactions, influences)

Regulatory netwoerk modeling is a specific field of systems biology since it allows to gain a systemic and consistent view of an intracellular system and of data, which are sometimes spread, disparate or seemingly contradictory. With the substantially increased amount of generated omics data, it has become clear that studying a couple of genes or some signaling mechanisms in isolation is not enough to describe a tissue status or a biological process. Network models regard an ensemble of biochemical reactions happening in a cell or a tissue as whole, in which all factors are interconnected and might influence the rest of the system (*201*). In that sense, they are holistic approaches perfectly suited to study the complexity of biological subcellular processes. Indeed, the interrelations are sometimes so intricate that the human intuition is limited in its capacity to deal with that complexity and a mathematical representation may be the only way to predict what will happen in case of external perturbations. Simply building a literature-based static network as a knowledge base, prior to any mathematical analysis, can already improve the systemic understanding of a biological process and spotlight lacks or discrepancies in information (*202*).

While biologists usually build regulatory pathway maps as a static representation of the knowledge they get from experiments, the goal of systems biology is to turn these maps into dynamic models (*203*). Determining the most appropriate *in silico* technology amongst all of the available mathematical modeling methods is a challenging task, this is why different kind of categorizations have been proposed (**Figure 1.7**), based on either the mathematical methodology or on the biological issues (*204*). As for any types of models, regulatory network models can be conceptualized in a more data centric or a more mechanisms centric way, those alternative approaches are introduced hereafter.

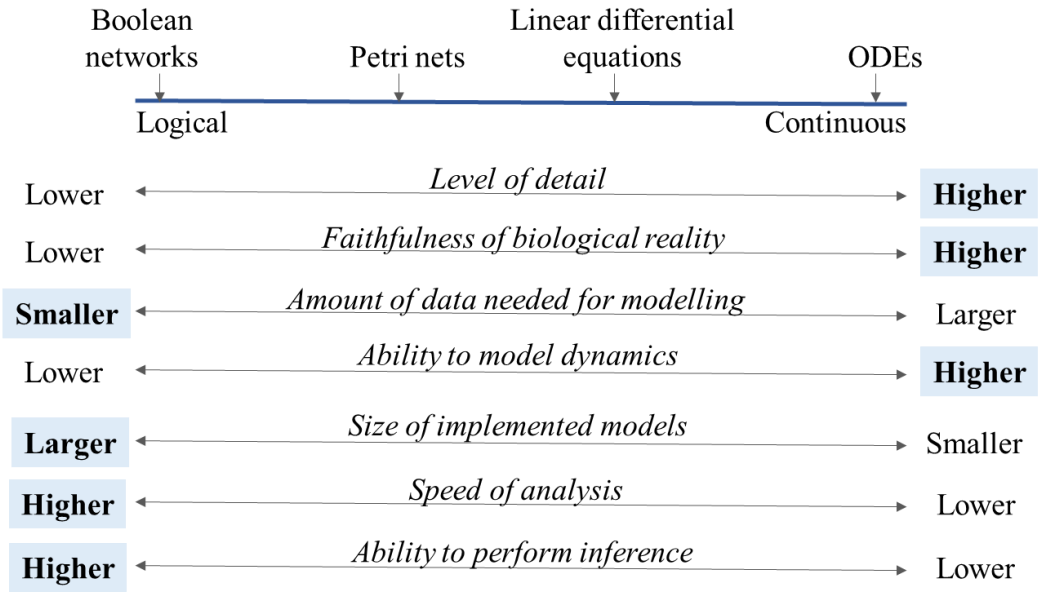


Figure 1.7: Comparison of approaches for regulatory network modeling. Different types of models are listed on an imaginary scale in which more logical (qualitative) are represented on the left and more continuous models (quantitative) on the right. This order also follow a gradient from the smaller (on the left) to the higher (on the right) level of details contained in the models. Different characteristics of these models are represented below the scale. Boolean can integrate a lot of biological details and become quite large since the amount of data needed and the computational complexity remains limited but it only provides qualitative conclusions. On the contrary, models based on ordinary differential equations (ODEs) are more quantitative and may provide dynamics that can be more easily related to continuous biological outcomes and the mechanisms are represented with details more representative of biological reality. However, the amount of information represented in those continuous model is often limited (small size) because the calibration requires a large amount of data. Petri nets can provide finer details than purely Boolean networks but the analysis is still qualitative. Adapted from Karlebach et al. (*205*)

Knowledge-based and mechanistic modeling of regulatory networks

The type of mechanistic methods employed to develop a network model will affect the level of details that can be included, the amount of data needed to establish the model, its level of interpretation and its computational complexity (**Figure 1.7**). Here we separate quantitative continuous methods from the qualitative and discrete ones.

Quantitative continuous methods for network modeling

A common way to study a system's behavior in a quantitative way is by using differential equations. In systems biology, the ordinary differential equations (ODE) describe changes of entities (e.g. protein concentration) over time whereas the partial differential equations (PDE) describe changes in time and space. In order to capture the level of biological entities over time (e.g. concentration changes of molecules) and describe the interaction between variables, differential equations rely on the law of mass action and its derivations such as the Hill equation. When ODEs are used to represent biochemical reactions and a network of protein interactions,

parameters typically are dissociation constants, kinetic rate constants, reaction order, etc (*201*). Such quantitative formalizations have the advantage of providing quantitative outputs that can be more directly related to and compared with experimental data. This type of models depends on numerical parameters such as kinetic constants and initial conditions, which are often unknown and difficult to measure, and which quickly increase in number when adding new variables to the mathematical system (**Figure 1.7**). Sensitivity analyses generally help to assess how much simulation results are affected when one or several of the parameters vary, dispensing an estimate of the model's robustness with respect to parameter values.

Qualitative and logical methods for network modeling

In qualitative and logical models, influences and dependencies are represented rather than specific mechanisms (*206*). In those models, variable values are the result of logical relationships (AND, OR, NOT gates) with other variables (**Figure 1.7**). Logical modeling was first applied to model gene regulatory networks (GRNs) by Stuart Kauffman (*207*, *208*). In Kauffman's approach, variables are evaluated using logical combinations of other variables and each variable can take only a discrete number of values (0 or 1 for Boolean models). René Thomas, another pioneer in logical models for systems biology, further refined the logical formalism by introducing multivalued variables and logical parameters equivalent to the kinetic parameters from differential equations to study cell differentiation (*209*).This enabled him to introduce asynchronous updating of the system which was later extended thanks to temporal logic (*210*). Indeed, logical systems can be updated in a synchronous way (all variables are updated at the same time during computation) or asynchronous way (variables are updated one after the other, potentially in random order). The ensemble of successive states through which a system passes during a simulation is called a trajectory and it can vary according to the initial state and the chosen updating strategy. Tracing all the trajectories across the ensemble of states allows to build a state-transition graph (*211*). With logical models, systems biologists usually study the attractors of the system, meaning that they try to identify in which stable state the system tends to settle (**Figure 1.8**). A stable state can be a fixed point or a discrete number of states over which the system oscillates. A system can settle in different attractors according to the initial condition from where it started. The ensemble of initial states leading the system to settle in the same attractor (stable state), is called a basin of attraction (*212*) (**Figure 1.8**). The study of the nature of the stable states and of their basin of attraction can give insight about cell states and state reachability, with application in cell fate decision and differentiation. Indeed, the size of a basin of attraction, meaning the number of initial states leading to a stable state,

may give an estimate of how likely it is for a cell to reach this state (*213*). An increasing amount of computational tools and software is being developed to help researchers study regulatory networks with this type of discrete mathematical formalism (*214–216*), including the establishment or optimization of network directly from experimental data (*217*) (see next section).

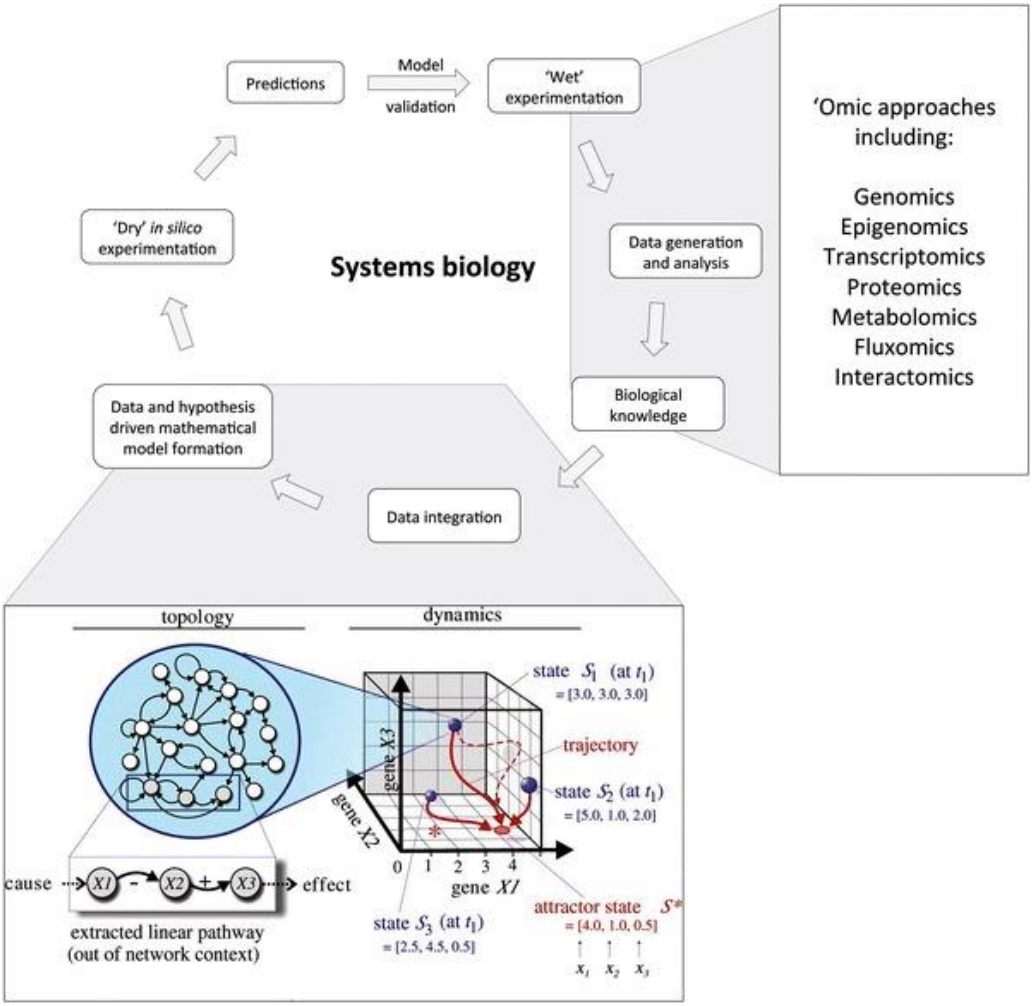


Figure 1.8: Schematic overview of the systems biology process and of the state space model. *In silico* modeling and wet lab experiments are part of the same iterative cycle in systems biology. Data generated with experiments lead to new knowledge that can be integrated into mechanistic (network) models. Numerical simulations also called *in silico* experiments allows the generation of new predictions or hypothesis that can be subsequently tested experimentally. The Inset image illustrates how the traditional pathway delineation of biochemical networks (traditional biochemistry approach) can be expanded and mapped into a multi-dimensional dynamic state space model. Each state status of the network (topology and activation level of entities) exist at a particular point (or state S) of this space, and can evolve towards a particular expression/activity profile (or stable state or attractor state) spontaneously or when relevant stimulus (e.g. a stimuli for stem cell differentiation along a lineage) is present. Figure adapted from McNamara (*218*).

Other interesting ways to analyze network dynamics in a discrete way are with probabilistic Boolean networks, which may express biological uncertainty stochastically since the same

initial state may lead to many trajectories, or with Petri Net (*205*) (**Figure 1.7**), which are associated with an easy state transition graphic representation making this formalism a good common ground for biologists and modelers.

Data-driven modeling, automatic network reconstruction and hybrid approaches

The methods described above actually resolve what we call a normal (bottom-up or forward) problem. It means that knowing all the mechanisms and parameters constituting a model, one aims to predict observations. Hence, it implies having a large background knowledge of the problem, including mechanisms and parameters. However, quite often in systems biology, such parameters are not easily available, only partial information can be obtained from the literature to build the models requiring fastidious manual literature curation. In contrast to the forward problem, the inverse (top-down or data-driven) problem aims to infer the rules and parameters describing the system given actual observations – in other words, deriving the model directly from the experimental data. The difficulty then is to find information that leads to the observed outcome given only some partially observed data at the boundaries of the domain.

In order to solve an inverse problem - meaning determining the causes knowing the effects (in contrast to identifying the effects knowing the causes for the forward problem) different methods exist depending on the question under scrutiny. Generally, inverse problems deal with an optimization problem since one needs to minimize a functional error between real values and synthetic values (*219*).

With the huge progress in the field of molecular biology, modelers now dispose of large amounts of quantitative data (Figure 1.8). Therefore, if we transfer the concept of solving inverse problems to systems biology, it is possible to imagine a way of inferring regulatory network topologies using measurements of biological activities such as messenger ribonucleic acid (mRNA) levels, protein levels, phosphorylation states, etc (*220*). Indeed, in regulatory networks, the parameters actually describe the edges linking the different components of the network. Thus, inferring the parameters of the model implies identifying the absence or presence, as well as the nature and intensity of the edges connecting the components (*11*). Following the categorization in four parts as found in (*211*, *221*) the mathematical inference methods can be listed as the statistical methods, probabilistic methods, information theoretic methods, and methods based on ordinary differential equations. Moreover, beside the network

topology, the mathematical model itself may be directly inferred or optimized in a data-driven way, be it for continuous equations or for logical Boolean gates inference (*11*, *222*, *223*).

Data-driven modeling raises the question of the identifiability of the model's parameters i.e. the ability to find an unambiguous set of parameters since several possible sets of parameters can generally fit the same data (*11*). In such case, the ensemble of parameter sets may be considered for subsequent simulations and variation across the ensemble systematically reported, this is called ensemble modeling.

Of course, both forward and inverse approaches may be relevant according to the type of questions the model is supposed to answer and the approach should be chosen wisely in a problem-centric way. An alternative can be to mix both of them as was done for instance in (*224*). In that study, the authors developed a curated knowledge-derived topology linking external stimuli to signaling pathways and a second part being data-driven, linking signaling proteins to cellular response. That way, they produced an extended network that was shown to be more performant to predict qualitative cellular responses than a totally data-derived model. We refer the reader to section 1.5.3 for other examples of hybrid regulatory network models of chondrocyte. So, data-driven techniques may complement knowledge-derived approaches but the opposite is also true as some techniques aim at inferring regulatory network from experimental data whilst incorporating curated knowledge as a prior (*225*). Overall this type of hybrid approaches allow making causal conclusion and gain comprehensive understanding based on large datasets, which might be key for recent integration efforts of multi-omics large datasets (*226*).

1.5.3 Modeling cell fate decision, chondrocyte (mechano-)biology and cartilage degradation and repair.

The importance of systems biology in the field of cartilage and bone tissue engineering and regenerative medicine has increased over the last years (*46*, *218*, *227–230*). A variety of modeling technologies is being used, covering processes at different spatial and temporal scales (genes/protein, cell, tissue, organs, physiological systems, etc.) as represented in **Figure 1.9**. In this section we discuss different types of applications. We start describing systems biology as an enabling tool for tissue engineering. Then, we discuss models of stem cell differentiation,

models of chondrocyte and cartilage degradation/repair to finish with mechanobiological models.

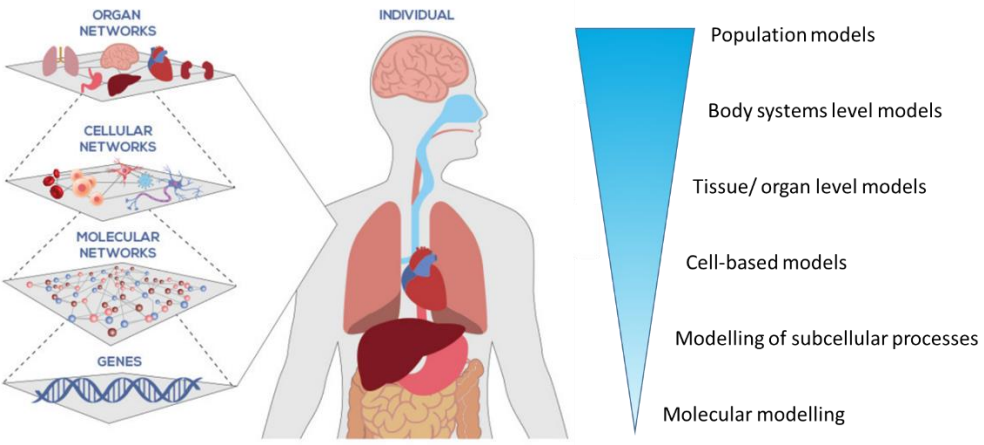


Figure 1.9: Graphical definition of systems biology. Networks may be used to organize and integrate information at different levels of organization. Modeling technologies covers all the possible layers from molecular models to full population models passing by subcellular processes and tissue or organ level models. They may also connect the different scales together. Modified from *https://isbscience.org/about/what-is-systems-biology/*)

Systems biology as enabling tools for tissue engineering

The potential of systems biology and network models in stem cell research already shows how it can help for the development and translation of tissue engineering strategies, especially when considering developmental engineering principles as guidelines for the TE construct manufacturing process. Indeed, as any process, it requires precise design and control, while with *in silico* modeling we apply well-established engineering principles from other industries and disciplines to the context of biological processes. Lenas et al. already formulated this need for applying engineering principles to DE a decade ago, emphasizing the importance of modularity, robustness, self-organization and controllability of biological processes (*46*, *231*) (**Table 1.1**). The underlying rationale is that so-called macroscopic modules (e.g. phenotype at cell level, histological structures at tissue level), critical to the acquisition of proper tissue engineered constructs, result from the reachability and robustness of underlying gene/protein regulatory modules. **Table 1.1** summarizes the matches between DE concepts, general engineering principles and systems biology methods with concrete examples. Those concepts and examples are further described below.

On the one hand, reachability of a module can be evaluated by considering the module as a mathematical system's stable state, comparable to a stable phenotype as formerly theorized (*30*,

207, *232*, *233*) (**Figure 1.8**). Inputting initial conditions to the system amounts to placing cells in a molecularly defined environment, activating transduction signals and eventually pushing the cells to acquire a new characteristic protein/metabolic/genetic profile, *in silico*. This is valuable, as it would enable the identification of a strict biochemical 'priming', and by extension the identification of critical treatment for reaching the module of interest (**Table 1.1**). In addition, a feature of importance for DE is the definition of an optimized minimal number of molecular targets in the regulatory network (e.g. genes, proteins, transcription factors) to be modulated in order to control for instance (stem) cell differentiation. When little dynamic information is available, network science and control theory provide relevant tools to determine key control nodes from the structure/topology of the network alone (i.e. the nodes and the ensemble of edges between the nodes). This was nicely theorized and applied to biological systems by Barabási et al. (*234–236*). Kobayashi et al. recently proved that simultaneous manipulation of 5 molecules among a network of 92 was sufficient to reproduce the cell specification of six different tissues in the ascidian embryo, using network science and experimental validation (*237*) (**Table 1.1**). Another challenge that DE attempts to tackle is the morphogenetic and self-designed cellular organization due to internal driving cues, so-called emergent behavior. Here again, there is a good match with current system biology techniques that have frequently been used to study pattern formation and morphogenesis during embryonic development such as chondrogenesis in vertebrate limb (*238*). It may provide mechanistic explanations allowing hypothesis testing and identification of the critical mechanisms that developmental tissue engineers have to ensure in order to replicate the wanted tissue structures. Finally, engineering cells or tissues requires quantitative metrics to evaluate how much the engineered cells resemble their native counterpart, in other words, a metric to assess and improve the quality of the process. Cahan et al. looked at the establishment of the reference target cells' GRN for such a metric (**Table 1.1**). They developed a series of machine learning tools exploiting microarrays, bulk and single-cell RNA-seq data (*239–242*) to reconstruct characteristic GRNs of various target cell types (stem cells, cardiomycoytes, neurons, etc.). These tools allow tissue engineers to input transcriptomic data from their own construct, assess quantitatively to which extent the target GRN is established and propose candidate transcription factors to be manipulated to improve the conversion (*239*).

	Developmental tissue engineering concepts			
	From gene module to phenotype, optimal conditions	**Cell organization to tissue structure/ pattern formation**	**Directed differentiation**	**Assessment of tissue constructs and intermediates**
Engineering principle	robustness, stability, modularity, reachability, directed intervention	multi-scale emerging behavior, self-designed organization	controllability, directed minimal intervention	modularity, quality control, directed intervention
Systems biology/ network science	dynamic regulatory networks, convergence to attractors and transitions	coupling subcellular signaling to multicellular behavior	scale free network topology & control theory	machine learning based GRN reconstruction and quantitative evaluation
Application example	Regulatory network of chondrocyte differentiation, convergence to attractors as emergent phenotypes and controlled transition between modules.	Spatially extended gene networks: coupling regulatory network with mechanics/ cell-cell interactions to reproduce epiboly in zebrafish during embryogenesis	Network structure based identification of 5 molecules (among 92) for which manipulation is sufficient to reproduce cell specification of six different tissues in the ascidian embryo.	CellNet: assess level of hepatocyte specific GRN establishment in induced hepatocytes and propose candidate TFs to manipulate for directing cell differentiation
Example reference	*(253)*	*(258)*	*(236)*	*(239)*

Table 1.1: Applying engineering principles, through Systems Biology approaches, to aid process design in developmental tissue engineering.

Computational models of (stem) cell differentiation

A comprehensive review about computational modeling to study embryonic stem cells differentiation, has emphasized that the analysis of the state landscape and transitions between states could give great insight into the dynamics governing cell fate decisions (*243*). The prediction of pluripotent stem cell fate decisions and regenerative medicine applications have been topics of predilection for mathematical modelers and systems biologists because the methods they use are particularly fit to predict cell lineage commitment based on internal molecular states and biological external cues (*244*). With a mechanistic approach, the work of Schittler et al. is a good example of a quantitative model predicting osteochondral cell fate decision (*245*). Indeed, they used a GRN, mathematically implemented with differential equations, for both single cell and cell population scenarios to investigate an osteochondral differentiation system. They modeled the switch mechanisms between three stable states, being the progenitor, osteogenic and chondrogenic states. On the other side of the spectrum, Bodaker et al. proposed a theoretical framework in which random Boolean networks were used to

represent cell type convergence within a multi-cellular tissue and identify conditions to induce differentiation (*246*). Interestingly, in that framework cells were governed by their own intracellular module, which could be altered by the influence of external signals from a neighboring population of cells. Systems biology is also about integrating and making sense of multi-omics data. For instance, the application of systems biology to genetic expression profiles is an efficient way to identify and demonstrate regulatory mechanisms underpinning changes in tendon and cartilage cell phenotypes (*247*).

Models of chondrocyte and cartilage degradation and repair

The use of computer modeling and simulations to study chondrocyte biology and cartilage degradation and repair is gaining traction, however the experimental validation is often still limited. To illustrate the diversity of those *in silico* approaches and their potential we mention hereafter some relevant examples. First, the concept of building a knowledge-based disease map and combining it with a data-driven network inference was recently applied to the case of the Rheumatoid Arhtritis (RA) pathology (*202*, *248*, *249*). This provided a template on which to project transcriptomic data (*248*). Eventually, that series of studies led to the identification of specific transcriptional factors as master regulators of RA thanks the comparisons between RA and OA tissues (*250*). Second, the predictive nature of ODE-based or Boolean signaling pathways models has been exploited to explain and investigate the role of oxidative changes, ages, inflammation, phosphoprotein signals and cytokines in cartilage breakdown (*93*, *251*, *252*). These studies validated previous findings and deciphered key mechanisms involving interleukin-induced production of ECM degrading enzymes. Furthermore, they highlighted underreported actors of chondrocyte catabolic changes (*252*) and offered an *in silico* interface to assess therapeutic interventions on those specific pathways (*251*). To date, few *in silico* intracellular models have focused on the growth plate chondrocyte, in particular. Nonetheless, such models may have important implication for tissue engineering and regeneration (*4*). For those reasons, Kerkhofs and coworkers implemented a series of models on the genetic switch between the chondrocyte's proliferative and hypertrophic states within the growth plate (*253*–*255*). They allowed to perform simulations with semi-quantitative outputs and to make conclusions about molecular drivers of this hypertrophic transition with implication for bone tissue engineering, as explained in the previous section. Very recently, an executable version of that model was implemented and released within the Cytoscape software interface of which the main advantage is to facilitate the exploitation of such a model by non-modelers (*256*).

Beside some examples of intracellular models, most mathematical representations of cartilage homeostasis and degradation available in the literature focus on the tissue scale and the extracellular matrix (*257, 258*). Most of the time in those models, the biochemical entities such as cytokines are solely represented via their diffusion potential and phenomenological effects on cells or tissues, without accounting for the actual cell sensing and molecular response mechanisms. In short, the level of cellular and molecular insights often remains poor in spite of their potential to provide interesting findings about the cartilage ECM production and loss under various physical cues or injuries.

Models of skeletal mechanobiology: coupling mechanics with intracellular regulatory networks

A particularity of bone and cartilage systems is the importance of the mechanical environment. The accommodation to mechanical load for a tissue and the resulting cellular response are inherently multiscale and multiphysics processes. This is why a great branch of the computational modeling field has focused on coupling mechanical tissue behavior with intracellular mechano-regulatory signals and gene regulation in various tissue types, with promising results (*259*). However, few applications targeted cartilage biology, *per se*. Generally, regulatory network models are perfectly suited to represent biochemical transduction of the mechanical signal inside the cell, as illustrated by the case of mesenchymal stem cell fate decision (*260*). In addition, representing mechanobiology at different scales or with multi-cellular models offers the possibility to study emerging collective behaviors (*261*), which is key in tissue growth and repair. Nevertheless, coupling cell regulatory networks to tissue mechanics remains a challenging task and requires further research (*262*). In skeletal cells, one of the main anchor between the extracellular space and intracellular mechanotransduction are the transmembrane integrins, therefore they constitute a strategic point to couple mechanics at the organ/tissue/cell level with intracellular regulatory networks. As matter of fact, recent modeling strategies have focused on studying integrin functions and the coupling with downstream mechanoregulatory pathways, in bone and osteoblasts (*263, 264*).

In this extensive introduction, we have presented information relative to the two branches of the present interdisciplinary thesis. It included the biological background about osteochondral tissues, related diseases and repair strategies, as well the computational modeling background that are necessary to start discussing the current thesis work. The hypotheses and aims of the thesis are provided in the following chapter.

Chapter 2. Hypotheses objectives and methodology

2.1 General hypotheses and aim

There is an increased understanding of the processes underlying the OA pathology as well as promising breakthrough for osteochondral tissue engineering. Nevertheless, no OA disease-modifying drugs are currently approved for use in patients, and limited advanced therapy medicinal products resulting from osteochondral tissue engineering strategies have actually been translated to the clinic. Since the cartilage to bone transition is a phenomenon that is observed both in the growth plate during bone development and in adult cartilage during osteoarthritis (*18*), controlling and modulating this transition may be key for different therapeutic applications. On the one hand, **promoting bone formation from cartilage is hypothesized to provide potent developmental tissue engineering strategies**. On the other hand, **inhibiting cartilage to bone transition might be paramount to find potent drug targets against osteoarthritis**. One of the major events happening at the cellular level during that transition is the **hypertrophic phenotypic change of chondrocytes**. This state transition is tightly **controlled by complex intracellular regulatory networks**.

Computational tools have shown promising capacities to integrate data and knowledge in order study intracellular biochemical processes in response to external cues, recognize important patterns and identify pivotal drivers of cell phenotypic changes. Therefore, we hypothesize that **computational models can be potent tools to investigate pathological intracellular signatures as well as cartilage to bone transition and chondrocyte response to external cues in order to identify potential therapeutic strategies for osteochondral systems, through endogenous (pharmacological) or exogenous repair (tissue engineering)**.

The aim of this thesis is then to leverage and develop different computational models to integrate complex knowledge underlying endochondral ossification with data-driven approaches and help identifying new therapeutic targets for cartilage and bone regeneration.

2.2 Specific objectives

This thesis is divided in four chapters **(Figure 2.1)**, for which the specific objectives were the following.

Chapter 3: Develop a pipeline for meta-analysis of high-throughput data in order to identify typical gene regulatory signatures in OA and evaluate mouse disease models, using machine learning methods. A computational pipeline was established to pre-process and integrate mouse

transcriptomic data from various origin in order to infer OA specific GRNs. We applied the previously published CellNet protocol (*240*, *241*) in which identified GRN signatures served as features to train a machine-learning model. This model was used to classify new samples coming from an experimental murine disease model, based on the degree of establishment of those GRN signatures.

Chapter 4: Build and validate a mechanistic regulatory network model of articular chondrocytes and identify potential drug targets against OA. This model lies at the heart of an *in silico – in vitro* strategy to search for potential drivers of chondrocyte hypertrophy. The novel articular chondrocyte model was based on an existing mechanistic regulatory network of growth plate chondrocytes (*253*, *254*), which was refined and extended for the articular chondrocyte and the OA case study. The mechanistic semi-quantitative model was further informed with regulatory interactions inferred by data-driven methods for which the meta-analysis pipeline developed in the previous chapter was leveraged. The resulting model served to investigate different scenarios relative to OA progression and treatment, by predicting cellular outcomes under a semi high-throughput screening of single and pairwise external perturbations. The main target hits were subsequently validated with dedicated *in vitro* experiments with murine and human cells.

Chapter 5: Apply and leverage an ensemble model of chondrocyte hypertrophy to prioritize biochemical differentiation cocktails for iPSCs-derived chondrocytes in a bone tissue engineering application. A knowledge-based chondrocyte regulatory network was optimized with engineering methods to fit expected chondrocyte profiles by valorizing a method initially shaped by Kerkhofs et al (*265*). The optimization strategy resulted in a non-unique solution with the identification of an ensemble of possible regulatory network topologies meeting the imposed criteria. The aim was to apply this ensemble model to a new bone TE case study. Specifically, we investigated the potential of that predictive tool to rank and prioritize biochemical differentiation cocktails for iPSCs-derived chondrocytes in a bone tissue engineering application. The experimental counterpart, executed by collaborators, dealt with iPSCs differentiation into cartilaginous aggregates comparing several biochemical treatments for their potential to promote hypertrophic differentiation and subsequent bone formation, *in vitro* and *in vivo*. We reproduced the experimental set-up numerically, predicted and compared the effect of treatments and proposed potential underlying mechanisms.

Chapter 6: Investigate the interplay between mechanoregulatory and inflammatory pathways and provide a chondrocyte intracellular module for a multiscale mechanical model of the knee joint. The articular chondrocyte model from Chapter 4 was further extended to investigate the influence of mechanotransduction and inflammatory pathways on chondrocyte phenotypes and cartilage matrix breakdown. Decades of knowledge about mechanoregulatory and inflammatory pathways and their interplay in cartilage was reviewed and the knowledge was captured in the model in the form of equations. It allowed investigating the relative role of the mechanoregulatory and inflammatory components on cartilage with the aim of understanding the tight balance between them. The ultimate goal of this model was to stand as an intracellular module bringing a real biochemical layer into a multiscale mechanical model of the human knee joint articular cartilage. The model model is being used to analyse the effect of physiological and pathological mechanical signals (e.g. post-meniscectomy) on chondrocyte fate and cartilage breakdown. Eventually it would allow numerical investigation of trial drug and ATMP therapeutic strategies.

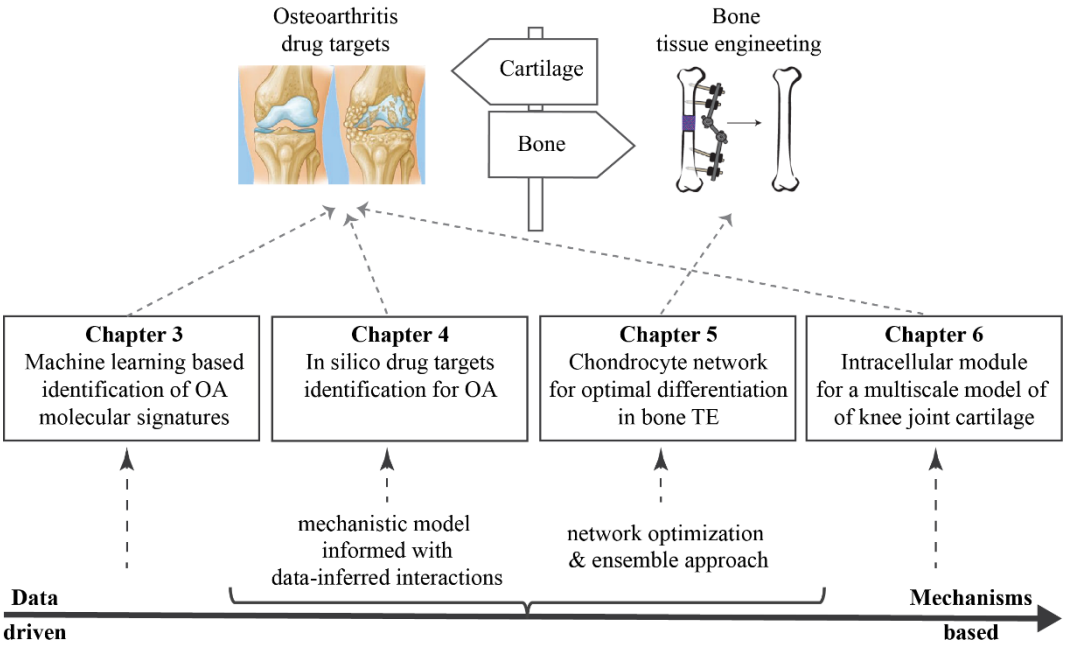


Figure 2.1: Flow chart of the thesis content. Chondrocyte hypertrophy may be modulated to either promote bone formation (e.g. for bone tissue engineering) or block adult cartilage mineralization and degradation during Osteoarthritis (OA) (e.g. for drug targets identification). The chapters of this thesis present the different models that have be developed and leveraged to tackle those issues. The mathematical models are of various types as illustrated by their positioning on the scale of data driven versus mechanistic models.

2.3 Experimental and numerical material and methods

2.3.1 Experimental methods

In vitro experiments were performed in order to validate and corroborate the *in silico* models' predictions.

Cell type and culture set-up

The **ATDC5 mouse chondroprogenitor cell line** was used to screen numerous *in vitro* treatments. They were obtained from the Riken institute BRC These cells were cultured in a thick layer and chondrogenic differentiation was induced by plating the cells at 6400 cell/cm² in proliferation medium for 24h, followed by changing the medium to differentiation medium (DM) for up to 14 days of culture (**Chapter 4**).

In addition**, primary chondrocytes** were isolated from **human articular cartilage** obtained with implicit consent as waste material from patients undergoing total knee replacement surgery. The medical ethical committee of the Erasmus MC, University Medical Center, Rotterdam approved that protocol (protocol number MEC-2004-3 22). P3-expanded chondrocytes were encapsulated in alginate beads (4 million cells/ml). Re-differentiation of chondrocytes was performed in a 24-well plate by culturing for two weeks in differentiation medium (**Chapter 4**).

The culture set-up for **iPSC-derived chondrocytes** was not part of the original work developed in this PhD thesis but results of *in vitro* and *in vivo* experiments carried out by collaborators with those cells were used in Chapter 5 as a basis for the *in silico* procedure and for comparison with the *in silico* results. Briefly, human pluripotent stem cell lines H9, 604B1, CY2 and NCRM1 were cultured and differentiated *in vitro*. Mesoderm induction preceded the chondrogenic differentiation and the resulting cartilage-like nodules were detached and cultured in suspension using a modified chondrogenic medium for 6 weeks. We refer the reader to (*266*) for the maintenance, the mesoderm induction and the chondrogenic differentiation procedures. Cartilaginous aggregates were treated to enter hypertrophy using basal hypertrophic differentiation medium (HM). Depending on the differentiation condition, HM was further supplemented with 100 ng/ml BMP4, 0.1 μM BIO (canonical Wnt agonist, Axon medchem), 1 μM 3,3,5-Triiodo-L-thyronine (T3) (Sigma) for 3 weeks. To initiate cartilage

resorption and turnover, 10 ng/ml IL-1β was added for an additional 10 days. To investigate the stability and bone inductive capacity of the cartilaginous aggregates, they were either ectopically implanted at the cervical region or orthotopically implanted in critical size long bone (tibiae) defects of NMRI nu/nu mice as described in (*267*) and (**Chapter 5**).

For all the 3 aforementioned set-ups, cells were grown in a humidified incubator at 37°C and 5% CO_2.

Small molecules and growth factors treatments

Various commercially available small activating and inhibitory molecules as well as growth factors were acquired to target specific biological molecules as listed in **Table 2**.1. They were applied to the cells in culture in order to assess the effect of the treatment on the cellular phenotype and compare the outcome with the *in silico* predicted effects (**Chapter 4**).

Target	PKA	SMAD3	FGFR1	WNT	ERK1/2	IGFI	BMP	HDAC4
Effect	+	+	-	-	-	+	-	+
Molecule name	Forskolin	Activin	PD161570	IWP2	PD0325901	IGF-I	LDN-193189	ITSA1
Cat n°	Axon 2264	338-AC-010	Axon 2098	72122	Axon 1408	791-MG-050	Axon 1509	18681

Table 2.1: Targets and associated small molecules or growth factors for the validation of *in silico* predictions. The first row indicates target to be perturbed, [+] stands for activation while [-] stands for inhibition. The name (resp. cat number) of the small molecule or growth factor employed to achieve that effect is indicated in the row called 'Molecule name' (resp. 'Cat n°')

in vitro assays

Hypertrophic differentiation of was evaluated with gene expression analysis (real-time quantitative polymerase chain reaction (RT-qPCR)) and biochemical colorimetric assay to measure medium secreted alkaline phosphatase activity (ALP). Cell quantity was assessed by DNA quantification (**Chapter 4)**.

Statistical analyses

Statistical analysis of experimental results were performed in R version 3.5.1. The employed mathematical methods are described in the respective chapters.

2.3.2 Computational and numerical methods

Most of the outputs of this thesis consisted of computational results for which the methodological aspects are described below. Overall, the mechanistic and data-driven computational research consisted in data processing and in the development of codes, algorithms and software.

Data processing and data-driven modeling:

All omics data (pre-)processing and analysis were performed with R version 3.5.1 and various packages that are cited in relevant chapters. Data-driven models and analyses were implemented with R too. Scripts from the initially published CellNet protocol were forked from the Github repository https://github.com/pcahan1/CellNet . They were modified to implement new functionalities and to train new cartilage reference ML models. GRN inference was performed using different published algorithms for which R packages, plug-ins in Cytoscape (*268*), or published Matlab scripts were available. (**Chapter 3 and 4**)

Mechanistic modeling:

Knowledge base model elaboration was taken care of by doing a rigorous literature study and assembling and the network and references in an online knowledge base via the platform CellCollective (*269*). The mechanistic models were implemented in and analyzed with the MathWorks® suite, MATLAB (2018b-2020b). A semi-quantiative framework developed by Kerkhofs et al. (*255*) was used to implement and simulate the mechanistic models. The algorithm underlying that semi-quantitative framework with priority classes is described in the flow chart **Figure S4.4**. It describes the underlying logical sequence used by the algorithm to update the model's variables during a typical simulation. A Monte Carlo analysis served to approximate the likelihood of ending up in a specific attractor (or stable state) of the system while perturbations and state transition analyses were performed to investigate model behaviors (**Chapter 4, 5 and 6**).

Whenever necessary and possible, independent large simulations were run in a parallelized manner, using the high performance computing infrastructure of the KU Leuven Vlaams Supercomputer Centrum (VSC).

An executable software was developed with MATLAB, to provide an easy interface for non-modelers to carry out *in silico* experiments with the virtual chondrocyte (**Chapter 5).** It is a standalone application, meaning it can be executed without the need for a MATLAB license.

Optimization:

Model optimization tasks were performed for the regulatory network model applied to bone tissue engineering. It consisted in the application of the genetic algorithm to optimize the parameters representing the interaction weights. This approach was reproduced from (*265*), in MATLAB. This type of optimization task typically generates families of models compatible with prior knowledge, but that are divergent in structure. We refer to those families as ensemble models (**Chapter 5**).

Multi-scale modeling

Outputs from a finite element multiscale model of the knee joint were used as input for the intracellular regulatory module described in **Chapter 6** but the simulations performed to obtain those outputs are not part of the original work reported in this thesis. They will be briefly described in the relevant chapter (**Chapter 6**).

Chapter 3. Development of a machine learning model to assess OA signature in mouse disease models

Based on

Development of a machine learning model to assess OA signature in mouse disease models.

R. Lesage, L. Geris.

Manuscript in preparation

3.1 Introduction

Many mouse osteoarthritis (OA) disease models have been established for research and drug discovery. However, truly quantitative metrics are still missing to assess how normal transcriptional programs are affected, how much the disease models resemble actual OA and how the histopathological changes are related to transcriptional dysregulation (*270*, *271*).

GRNs are composed of key transcriptional regulators, or transcription factors (TFs), and their target genes. They are usually represented by network maps in which nodes/vertices represent genes and directed edges represent regulatory interactions. The establishment of specific GRNs governs the expression program of cells and defines their state and fate. It also governs cells' response to their external environment since GRNs are rewired during differentiation or disease. Hence, specific GRN signatures tend to characterize different diseased environments. Being able to identify OA specific signatures and to recognize the status or the cellular state of a biological sample based on that GRN signature is of major interest not only for classification and diagnostic purpose but also to get quantitative metrics for quality assessment of *in vitro* or *in vivo* disease models.

Exploiting the importance of GRNs for cell fate decision, Cahan et al. (*241*) developed an online platform, CellNet, which is a computational tool measuring the extent to which target cell or tissue's GRNs are established in stem cell differentiation experiments. More precisely, CellNet was meant to classify cell types (stem cells, hepatocytes, neurons, etc.), quantify GRNs status and assess the importance of transcriptional regulators, through the use of machine learning (ML) models, being classifiers (*239–241*). ML models are a type of data-driven models that are trained on a set of data to recognize certain types of patterns, they learn from those data how to predict an outcome variable based on a pre-selection of features. Classifiers describe a subclass of ML models that predict, by looking at the selection of features, whether new observations (or query samples) belong to a class or another based what was learnt from the training set. The online CellNet platform is constituted of ML models trained on transcriptomic data from numerous types of mouse and human cells and tissues in such a way that users can query their own newly engineered samples and classify them on the basis of characteristic GRN establishment from the reference cell or tissue. To date, no skeletal tissue related data is included in the reference/training database of that tool and no study have applied it in the context of pathological changes. .

Here, we will follow the CellNet concept and train an ML model with knee joint cartilage and osteoarthritis related data. This model will allow identification of possible OA signatures, classification of new samples based on typical GRN signatures establishment and will provide a path towards systematic quantitative evaluation of established or newly developed *in vitro* or *in vivo* OA models. Given the need for a sufficiently large and informative dataset to reach the aforementioned objectives, and given the larger abundance of microarray based transcriptomic data for mouse OA models, we will start by composing a working dataset by assembling several array experiments, using a pipeline designed to integrate data from these heterogeneous sources. Next, we present our adaptation to the previously published CellNet protocol (*241*), an adaptation aimed at integrating cross-validation for estimating the model's performance. Subsequently, we will use this protocol to identify mouse OA specific signatures as well as train and evaluate a new ML model thanks to control (healthy) and OA gene expression profiles. Finally, we will validate the newly trained classifiers by querying new independent samples using the previously identified typical GRN signatures as reference.

3.2 Results

3.2.1 Data assembly and pre-processing

Machine learning approaches, whether it is for GRN reconstruction or for building classifiers, require large and informative data. Ideally, such a data would come from a multi-perturbed system, providing significant variance across conditions. However, in absence of such dataset for osteoarthritis in murine models, we assembled one by gathering together several microarray experiments containing mouse knee joint samples from genetically modified mouse spontaneously developing OA or mechanically induced OA (**see method and Table S3.1**). Indeed, it was previously shown that rich and informative data for regulatory network reconstruction could be successfully achieved by assembling naturally occurring and experimentally generated phenotypic variations of the cell type of interest (*272*). As a result, our final dataset amounted to 109 samples for training plus 26 independent samples for the final validation (see **Figure 3.1** for a graphical summary of the overall study design).

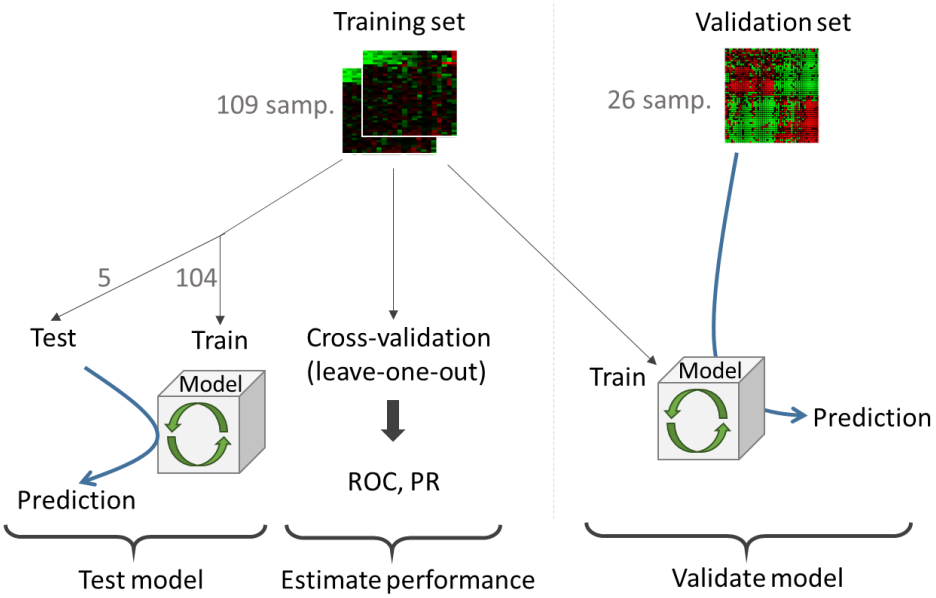


Figure 3.1: Study Design. The training set was composed of 109 samples. First, the model was trained on 104 samples and tested with the 5 remaining samples to evaluate the model and illustrate the classification system. Then formal estimation of the model's performance was obtained via a cross-validation in which all the samples were used one by one as test sample. The final model, trained with the 109 samples, was validated with 26 independent samples.

The dataset contained data on 5386 genes, a rather low amount compared to the full mouse genome, which is an inevitable result from the filtering strategy of the pre-processing pipeline combined with the integration of the different arrays. Indeed, only genes that were commonly expressed in all arrays remained in the final dataset. Among the training samples, 58 were annotated as 'WT' and 51 as 'OA' based on the sample's description in the original published metadata on GEO (see Method and manual annotations in **Table S3.2**). As a quality control, a principal component analysis showed that, after being corrected for batch effects, samples were not clustered according to their microarray of origin and had a homogeneous gene expression distribution (**Figure 3.2**). This suggested that the variance induced by the technical origin was removed and only the biological variance from the experimental conditions was maintained.

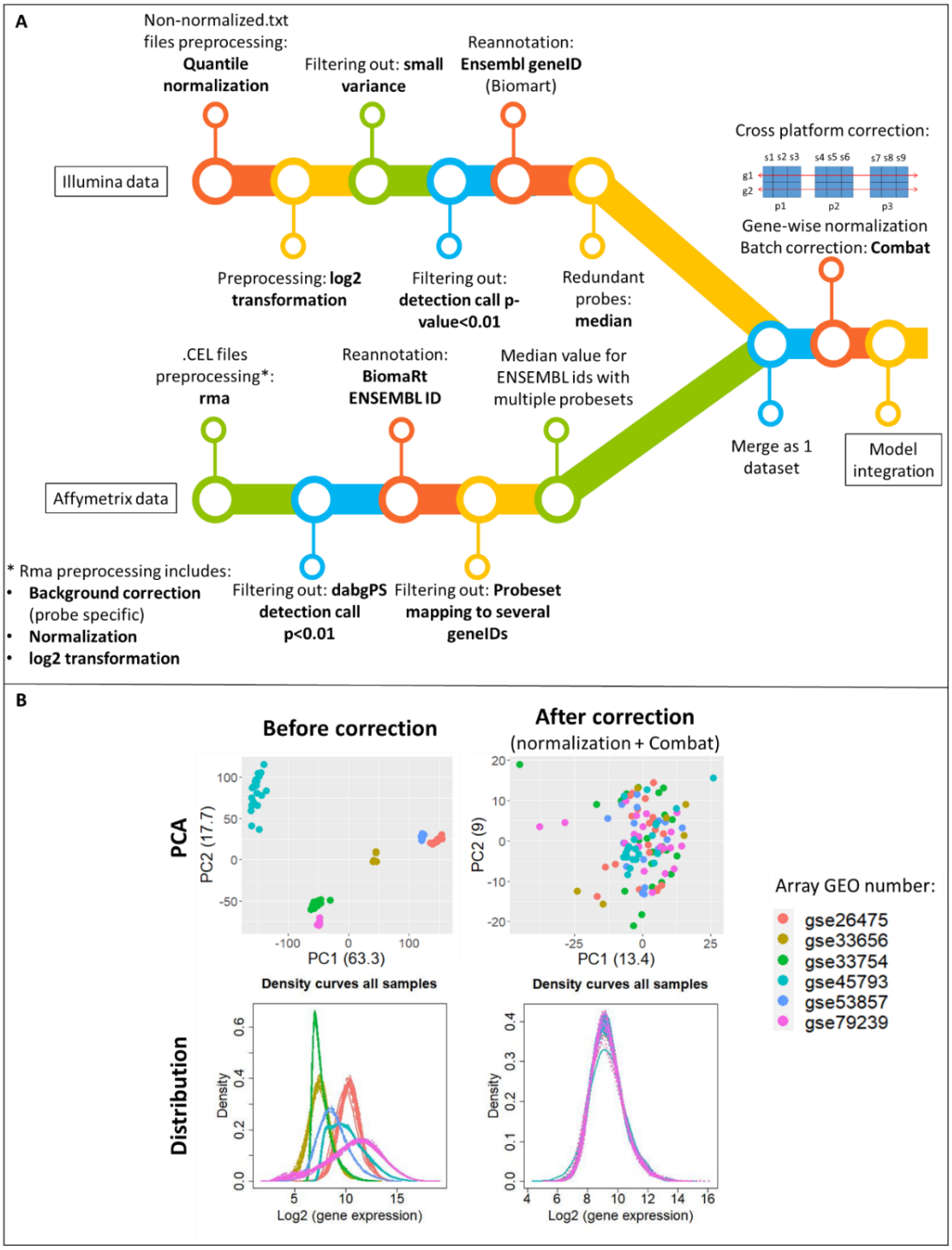


Figure 3.2: Illustration of the micro array preprocessing pipeline for array integration as a single dataset. **(A)** The pipeline includes platform specific branches, of which the serial steps are kept as much equivalent as possible between branches. The array data are then merged as a single dataset and corrected to allow comparison between genes across all arrays. In the cross platform correction representation, g_i indicates the genes, p_i the different platforms and s_i the samples. The pipeline was modified from the method of Turnbull et al. (2012) that was meant to increase the statistical power when re-analysis data from heterogeneous origins. **(B)** Principal component analysis and gene expression distribution of the merged dataset before and after batch correction with the ComBat algorithm. Samples (i.e. points in the PCA or curves in the distribution graph) are color-coded according to their array of origin, which are labelled with the GEO accession number in the legend.

3.2.2 Models construction and characteristics

The concept of CellNet relies on the identification of cell type specific signatures, which are the GRNs. The ML model consists of classifiers for each cell type that is represented in the dataset, as in the original CellNet protocol. Here, we built a binary classifier able to discriminate OA samples from all other types as well as another binary classifier able to recognize WT samples from all other types. Manual 'OA' and 'WT' annotations served as the ground truth and the genes involved in characteristic GRN signatures served as main features, or variables, to train the models or classifiers (see **Figure 3.3**, method & supporting material).

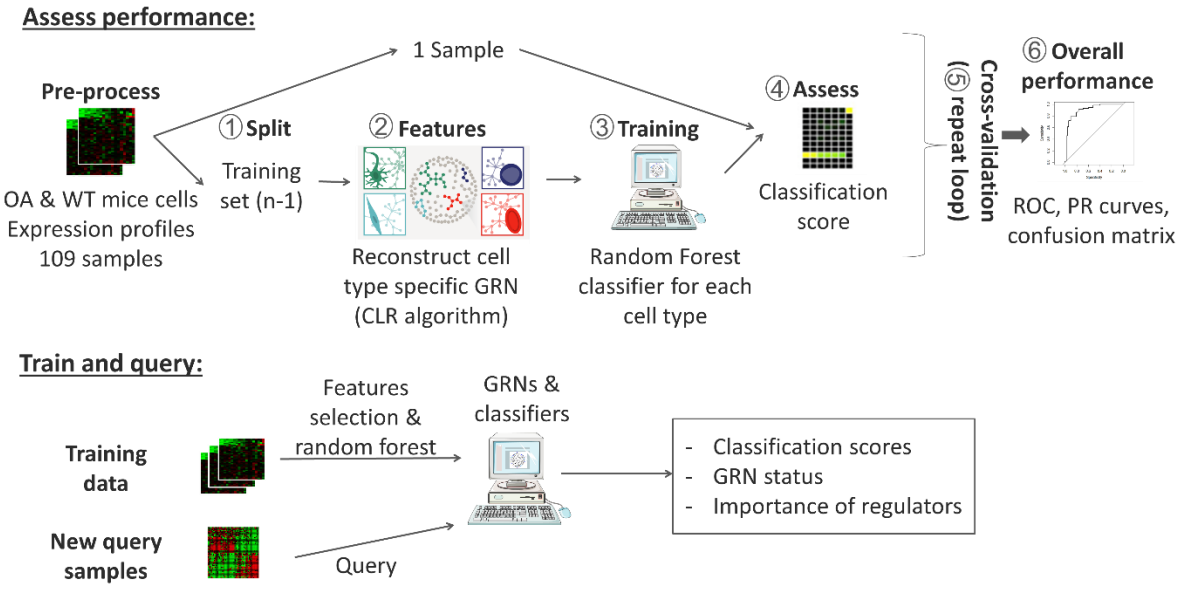


Figure 3.3: Workflow for the model construction, validation and new sample query. The upper part represents the protocol, modified from the initial Cellnet protocol, allowing for cross-validation. At each iteration a sample is kept aside from the training set (Step 1). This last one, is used to reconstruct cell type specific GRNs, which is seen as a feature selection step of the training process (Step 2). Indeed, random forest classifiers are trained to recognize each cell type based on the signatures exhibited by the genes comprised in the aforementioned GRNs (Step 3). These models are then used to score and classify the held out sample (Step 4), of which the prediction can be compared to the actual annotation or label. This process is repeated until all samples have been used as test sample once (Step 5). At the end, the overall performance is estimated (Step 6). The lower part illustrates the building and application of the final model. The training dataset is used to select features by GRN reconstruction and to build the final classifiers that are to be queried with new samples from an experiment of interest. Specific scores and metrics are computed to assess the query samples regarding their resemblance to the reference cell types and how much of the reference cell type signature is established. OA: osteoarthritis; WT: wild type; CLR: context likelihood of relatedness; ROC: receiver operator curve; PR: precision recall.

Hence, we first reconstructed OA and WT specific GRNs based on the mouse transcriptomic data. Over the 5386 genes included in the dataset, 100 genes were identified as part of a OA specific GRN by the algorithm, of which 5 transcription factors (*Elk3, Hif1a, Nfil3, Runx1* and *Hspa8*). Additionally, 33 genes, of which 5 transcription factors (*Rpa2, Sox6, Zfp101, Dbp, Thrap3*) were identified as important features to recognize WT samples (see **Figure 3.4**). The importance of each of those genes in their respective classifiers was assessed using the mean decrease GINI measure. The GINI is a measure of the node impurity in the random forest algorithm; assessing the decrease in GINI when a given gene is omitted during the training process gives an understanding of how important that gene is to classify the data correctly. For the OA classifier, the top 15 most important genes to recognize OA samples are reported in (**Figure 3.4C**), note that the ranking measure is relative and thus the absolute values should be disregarded. Among the most important genes, when searching the MGI database (the international database resource for the laboratory mouse) several ones also happened to be associated with growth/size/body, skeleton or immune system mutant phenotypes. For instance, this is the case for *Tgfbi* (Tgf, beta induced), *Timp1* (tissue inhibitor of metalloproteinase 1), *Errfi1* (ERBB receptor feedback inhibitor 1), *Ctsh* (cathepsin H), *Cd44* (CD44 antigen), or *Lbp* (LPS binding protein). **Table. 3.1** reports those genes, among the OA signature top 15 genes, that were associated with relevant phenotypes, e.g. related to bone mineralization, density, growth, inflammation, etc. Particularly, *Errf1* and *Cd44* mutants seemed to be associated with modulation or induction of arthritis and OA in mice. Since OA is a disease associated with increased pro-inflammatory pathways activity, increased mineralization and recovery of embryonic endochondral ossification processes, these results highlight the capacity of the current method to extract OA relevant signatures.

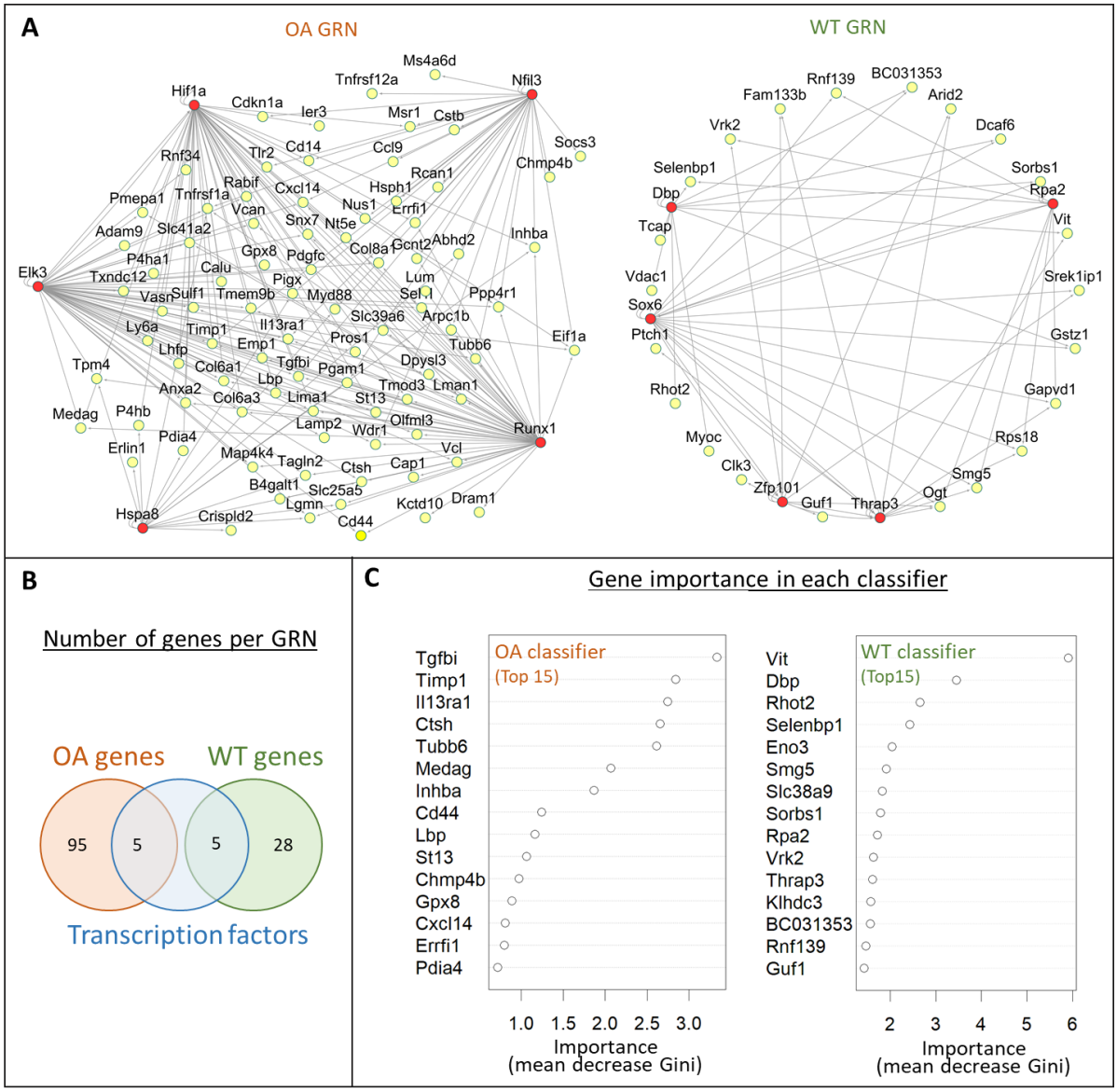


Figure 3.4: GRN signatures and features importance (A) Reconstructed GRNs for the osteoarthritis (OA) and the wild type (WT) training samples respectively. The nodes represent genes, the transcription factors (TFs) are in red and target genes in yellow. Directed edges represent regulatory interactions from TFs to target genes, with the arrow giving the direction of the interaction. **(B)** Venn diagram showing the number of genes involved in the OA (resp. WT) reconstructed GRN and highlighting the number of transcription factors among them. **(C)** Importance of the genes for each classifier is estimated as the mean decrease Gini. The Gini is measured during the random forest model building, it gives an understanding of how important is a gene within a group to classify the data correctly. The top 15 genes are ordered from the most to the least important for both classifiers.

Symbol	NCBI Gene ID	Full name	Function	Interesting phenotype annotations (associated with mutants)
Tgfbi	21810	Transforming growth factor, beta induced	Receptor binding, extracellular signaling molecule	- growth/size/body: postnatal growth retardation, decreased body weight - decrease insulin secretion
Timp1	21857	Tissue inhibitor of metalloproteinase 1	Enzyme regulator	- growth/size/body: body growth and weight phenotype - skeleton: Abnormal humerus & femur morphology
Il13ra1	16164	Interleukin 13 receptor, alpha 1	Signaling receptor activity	- homeostasis, immune system: increased circulating interleukin, increased nitric oxide homeostasis
Ctsh	13036	Cathepsin H	Protease	- skeleton: abnormal bone mineral density
Cd44	12505	CD44 antigen	Signaling receptor activity	- skeleton: decreased susceptibility to induce arthritis, short tibia, abnormal osteoclast differentiation. - immune system phenotypes
Lbp	16803	Lipopolysaccharide binding protein	Extracellular signaling, receptor binding	- skeleton: decrease bone mineral content - immune system: abnormal chemokine level, abnormal TNF secretion, abnormal inflammation
Inhba	16323	Inhibin beta-A	Receptor binding, extracellular signaling molecule. Cell differentiation	- skeleton: craniofacial, abnormal alveolar bone morphology
Cxcl14	57266	Chemokine (C-X-C motif) ligand 14	Receptor binding, extracellular signaling molecule	- growth/size/body: decrease body length
Errfi1	74155	ERBB receptor feedback inhibitor 1	Scaffolding adaptor protein, response to stimuli, cell differentiation	- skeleton: joint swelling, osteoarthritis, abnormal femur morphology, increased chondrocyte number, abnormal articular cartilage morphology, abnormal endochondral bone ossification

Table 3.1: Summary of top 15 genes in the OA GRN that are associated with relevant mutant phenotypes. The 15 most important genes for the OA classifier (based on GINI measure) for which relevant mutant phenotypes were found related to skeleton, growth/size/body or homeostasis and immune system through the MGI database (http://www.informatics.jax.org/) . For each gene, the name, NCBI ID, full name and function are reported, for information. The details of the observed phenotypes are reported in the last column.

3.2.3 Preliminary testing and assessment of the model

For the sake of testing and assessing the classifiers, intermediary models with 104 samples as input were trained in order to test and predict the class of the remaining 5 test samples. These samples were biological replicates of some of the training samples and came from mice that underwent surgical destabilization of the medial meniscus (DMM). More precisely, the test samples were from a naïve mouse without surgery and mice with DMM induced OA, at 3 and 6 days after surgery, with the respective sham operated controls.

As expected, predicting the samples' classes with the OA classifier showed increasing OA scores (indicating increasing resemblance to the OA signature in the training set) for increasing time after DMM (**Figure 3.5A**). Indeed, the lowest score was for the naïve sample without surgery and the highest one for the sample taken 7days after DMM induced OA, showing nearly

100% chance to be an actual OA sample. Interestingly, even though the OA scores for the sham-operated samples remained low, they were slightly higher than for the naïve sample (**Figure 3.5A**). This could be explained by the fact that sham-operated mice experience cellular changes comparable to some initial OA changes, such as activation of pro-inflammatory signals. Inversely, with the WT classifier, the naïve sample was indeed the one resembling most a WT-like phenotype with a score close to 1 (denoting almost 100% chance to be a WT sample) and decreasing scores were observed with increasing time after DMM.

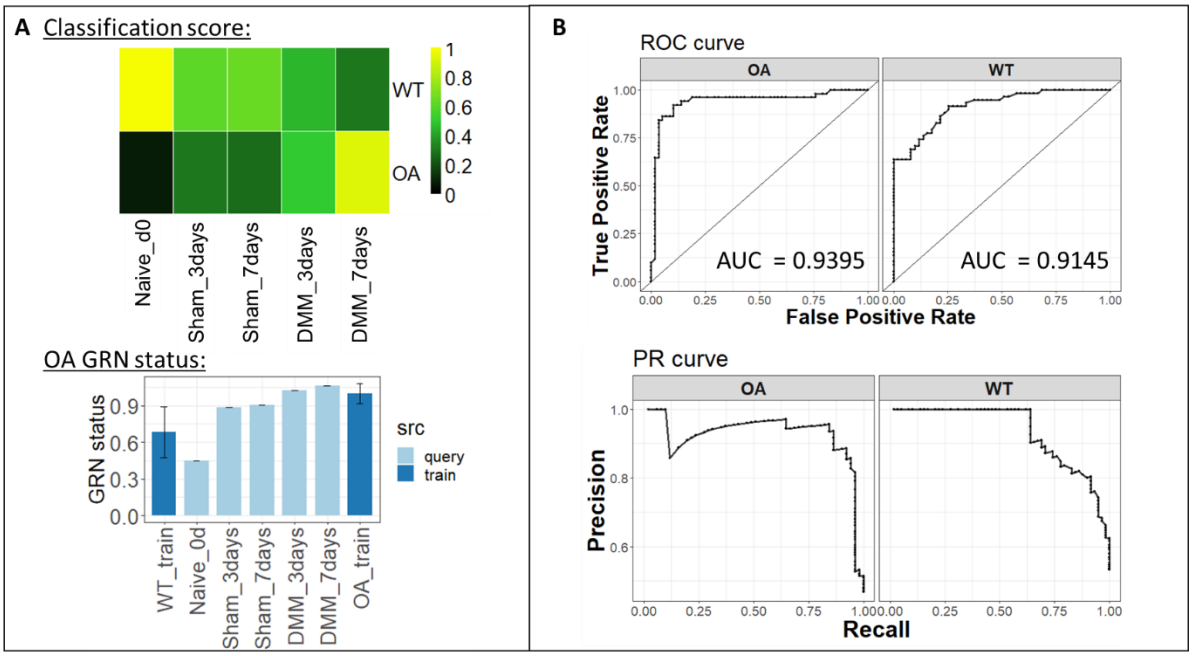


Figure 3.5: Model testing and performance estimation. (A) Receiver operating characteristic (ROC) curves with True Positive Rate (TPR) versus False Positive Rate (FPR) and Precision Recall (PR) curves with positive predictive values (i.e. precision) versus TPR (recall) for both the osteoarthritis (OA) and the wild type (WT) classifier. The ROC curves tend towards the ideal case (upper left corner in which there is no false positive and all the real positive have been classed as such). The PR curves tend towards the ideal case too (i.e. the upper right corner in which all real positives samples have been classed as positive (recall =1) and all the samples classed as positive are true positive (precision =1)). **(B)** Assessment of models behavior by predicting the score of 5 test samples taken out of the 109 samples-dataset. The heatmap reflects the prediction scores obtained for each sample when applying the OA or the WT classifier. The color indicates the classification score, being the average classification for each training cell type of the training data. Scores vary from 0 (i.e. distinct from the cell type of the training data) to 1 (i.e. indistinguishable from the cell type of the training data). The OA GRN status indicates the level of establishment of the OA specific signature for the training samples (dark blue) and the test samples (light blue). The measure of the OA-specific GRN status for a sample is derived from Cahan et al. (*240*) and is based on the mean z-score of all genes from the OA signature in the sample, weighted by their importance in the OA classifier. That GRN status is then normalized to the mean GRN status of the training OA samples. Bars plots indicate the mean ± 1s.d. There is only one sample per condition for the test samples that are queried.

Cahan et al. (*240*, *241*) established a scoring system to measure the GRN status of a sample based on comparison with a given classifier and its reference signature. We computed that score

for our query samples based on the OA classifier. In **Figure 3.5A**, the GRN score indicates the extent to which OA specific GRN is established in the training data (dark blue) and test or query samples (light blue). Both 3 days and 7 days DMM samples' GRN scores denoted that the expression levels of genes involved in the OA GRN had a profile similar to the OA samples in the training dataset. Moreover, the lower GRN status scores for the naïve samples denoted an expression profile different from the one of the OA training samples for the genes involved in the OA GRN. Surprisingly, the OA GRN score of the training samples annotated as 'WT' was rather elevated but displayed a large standard deviation (SD). This can be explained by the heterogeneity of the reference 'WT' samples in our reassembled microarray dataset. Indeed, all sham-operated samples in the training set were manually annotated as WT because no OA-specific changes were theoretically expected in these controls in the initial experiments. Nevertheless, this is not entirely accurate as proven by the relatively high OA GRN score for the 3 and 7 day sham-operated samples. Typically, this type of conditions (Sham operated) contribute to the high average and large SD of the WT training samples. First, those results support the fact that sham-operations themselves have a significant impact on mouse knee joint biology. Second, they illustrate the heterogeneity of our WT training set explaining the slightly lowered performance previously observed.

Those results highlighted important characteristics of the trained classifiers. Nevertheless, a risk of bias introduction, and therefore of overfitting, cannot be excluded given that the test samples originated from the same experiment as some of the training samples, so they were not fully independent. For that reason, a formal estimation of performance was also carried-out and is presented next.

3.2.4 Cross-validation strategy and performance

Performance assessment for a newly trained ML model is usually done by testing it against a test dataset. This test set must be kept completely separate from the model training process in order for the estimation of the model's performance to be reliable. It is common practice to split the initial dataset into a training and a test subpart according to a system of proportion (e.g. 60% training, 40% test). However, this prevents the training from exploiting the full data and may lead to information loss. Moreover, that approach might become limiting when the sample size is rather low to start with. An existing solution is to assess the model performance with a cross-validation strategy.

In the current study, we implemented a new version of the CellNet protocol integrating leave-one-out cross-validation (see method & **Figure 3.3**). As a result, there were two main differences with the previously published protocol. First, smaller dataset could be used since all samples were considered as test samples for the performance evaluation, thereby preventing the need to split and shorten our training set. For example, 109 samples populated our dataset while 16 to 20 mouse cell or tissue types with about 10 perturbations per cell/tissue type were considered in some previous CellNet implementation (*240*, *241*). Second, the GRN reconstruction and the underlying feature selection steps were considered as part of the training module and were done within the cross-validation loop, just after splitting the data into training and 1-sample test sets (see method and **Figure 3.3**). This way, we made sure that the information embedded in the test sample would not be used to extract the features upon which the classifiers were built, which is important to prevent any bias in the performance estimation.

The cross validation procedure was applied to train one classifier/model per cell type (OA and WT). The resulting models demonstrated a high performance for classification as illustrated by the shapes of the precision recall (PR), tending towards the upper right corners, and of the receiver operating characteristic (ROC) curves, tending toward the upper left corners (**Figure 3.5B**). The PR curves, indicating the tradeoff between the positive predictive values (precision) and true positive rates (recall) for different cut-off thresholds of the classification score (i.e. the score above which a sample is considered positive by the classifier), showed that the precision remained relatively high even with increasing recall values for both classifiers (**Figure 3.5B**). Therefore increasing that cut-off threshold to ensure a good precision can be achieved while maintaining a satisfying recall, which is interesting in the context where the OA classifier would be used to evaluate the quality new OA experimental systems with respect to the reference profile. Indeed, one could imagine here that a good filter or quality system should strongly prevent sub-quality experimental system to be accepted (few false positive) while avoiding the rejection of too many systems with a satisfying quality (limit false negative). The ROC curves showed that the true positive rates (sensitivity) increased faster than the false positive rates, resulting in satisfying curve shapes and an area under the curve (AUC) of 0.94 and 0.91 for the OA and the WT classifier respectively. This meant that both model were reliable to find true positive rather than false positive. Overall, the performance for the OA classifier was better than for the WT one.

3.2.5 Application to new query samples and validation

Possible applications of such a ML model would be to query new samples coming from *in vitro* or *in vivo* mouse experiments expected to mimic OA or to assess (drug) treatments expected to limit osteoarthritic changes. The computational model could provide quantitative metrics for quality assessment of the experimental OA model, suggest transcriptional regulators that would need further modulation to improve that experimental model or assess the effect of a specific treatment on the OA signature.

To illustrate some of the aforementioned applications and in order to validate our model, we predicted the class of a new sample set. These samples originated from an independent experiment that was neither used to identify features (i.e. genes in the GRN signatures) nor to train the models. The original data were produce for a time course study of gene expression in a mouse model of osteoarthritis, being the DMM (GSE41342) (*273*). It included 26 samples with naïve mice, DMM mice at 2, 4, 8 and 16weeks after surgery and the corresponding time-point sham-operated mice. As explained above, we used the full model, trained on 109 samples, for which the extracted GRN signatures and performance were reported in **Figure 3.4A & 3.5B**.

By predicting the class of the new samples we noted that naïve and sham samples were predicted with a low OA scores and high WT scores except for 2weeks after sham surgery (**Figure 3.6A**), a result that was not unexpected based on the initial study report (*273*). Indeed, in the initial paper, the 2weeks Sham samples showed a different pattern than other Sham and this condition was more alike the DMM samples. This is in line with our previous observation that Sham operated samples have some OA-like changes.

Samples after 2weeks and 4weeks of DMM surgery were predicted with the highest OA score and lowest WT scores; however, this pattern did not hold true for later samples (8 and 16weeks after DMM surgery) (Fig 3.6A). Indeed the OA scores for those samples became lower again. Even though it might seem surprising, it is again in line with the conclusion of the initial report for that dataset, since the authors observed a phasic effect. Their results supported a phasic development of OA with early matrix remodeling and transcriptional activity followed by a more quiescent period that was not maintained. At 16weeks, increased expression was observed again but not to the level of week2 and 4. They also reported a cluster analysis showing that each time point had a unique gene expression pattern supporting the hypothesis of a phasic process.

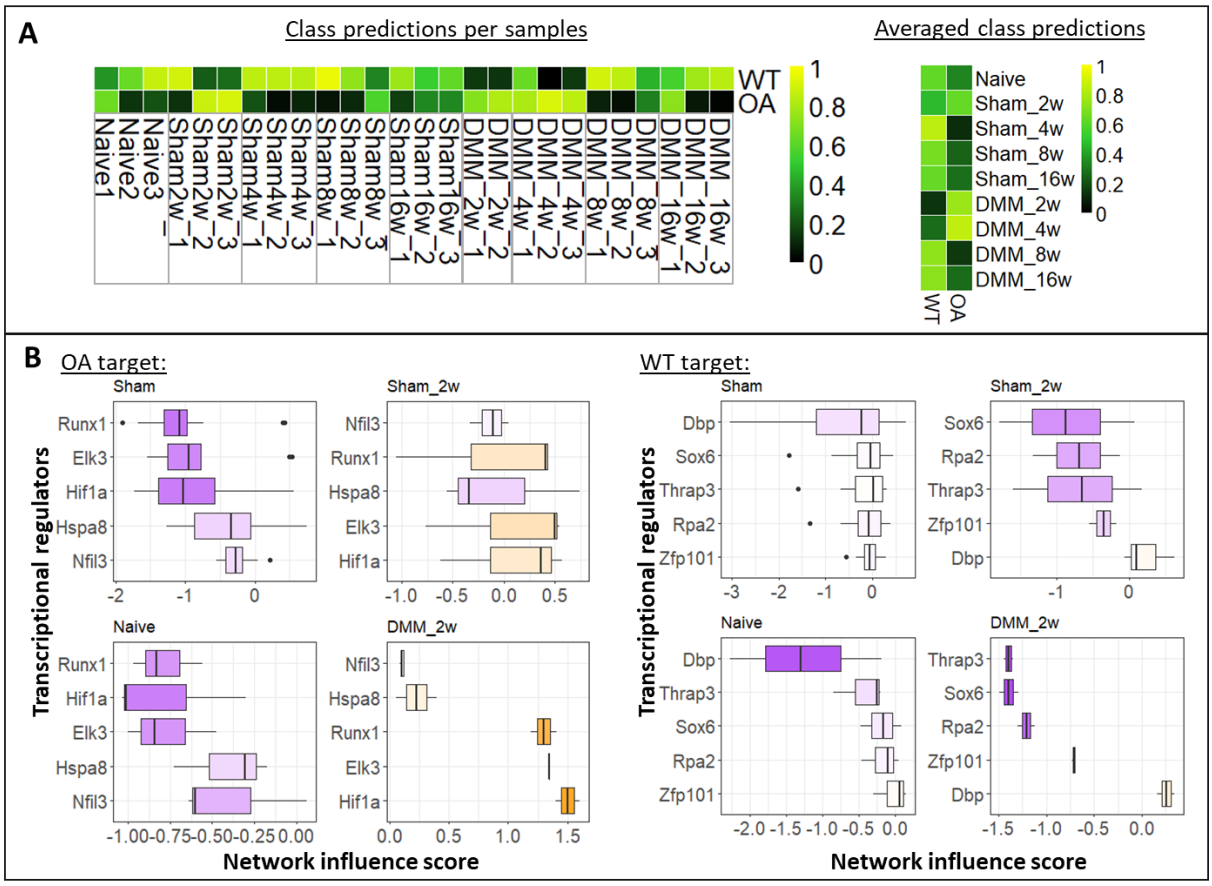


Figure 3.6: Class predictions and transcription factor influence for independent query samples. (A) Classification heatmap of samples from the knee joint of naïve and Sham operated mice as well as mice that underwent DMM at 2, 4, 8 and 16weeks after surgery. On the left side, columns represent individual samples (including biological replicates) while rows represent the cell/tissue types of the training data. On the right side, training cell types are in columns and sample conditions in rows. The color indicates the classification score of the individual samples (left) or the average classification score (across replicates) of the conditions (right) for each training cell type. Scores vary from 0 (i.e. distinct from the cell type of the training data) to 1 (i.e. indistinguishable from the cell type of the training data). **(B)** Estimation of the importance of transcriptional regulators to dysregulate gene regulatory networks (GRNs), based on the network influence scores (NIS). The colors go from purple for the most negative average NIS to the orange for the most positive ones; boxes are white for a null average NIS. 'Sham' graphs provide average NIS across all Sham samples (2,4,8 and 16 weeks), 'Naïve' graph, across all Naïve mouse samples, Sham2w across all 2weeks sham samples and DMM_2W across all 2 weeks DMM samples. NIS, on the x-axis, are computed and reported with either the osteoarthritis (OA, on the left) or the wild type (WT, on the right) cell types as target/reference. All transcriptional regulators of the corresponding classifier are listed on the y-axis. NIS<0 (resp. >0) means that the expression of the transcription factor (TF) is lower (resp. higher) in the query group as compared to the target/reference cell type. The magnitude of NIS is determined by the number and the dysregulation level of the target genes from the reconstructed GRN. The NIS measure can especially be used as way to prioritize TFs for intervention.

Finally, we computed the network influence scores (NIS) of each cell type's transcription factors, for several sample groups (see method). Initially meant to prioritize transcriptional regulators and help design of experiment for stem cell differentiation and tissue engineering (*239*, *242*), in the context of cartilage diseases the NIS could be seen as a measure of how much the cell-type specific TFs and associated regulons are dysregulated in the diseased samples as compare to the WT signature in the training set. Alternatively, it could be used to identify the

regulons that need further perturbation to reach a desired state, be it the WT state for OA-samples treated with a drug under scrutiny or the OA state when the purpose is to improve an experimental disease model or compare a new model with a reference standard model, for instance.

Following this approach with the current query samples, we observed that the transcriptional regulons of the OA signature were all downregulated in the Naive and Sham samples compared to the reference OA data, with a higher amplitude for *Runx*1 (Runt related transcription factor 1), *Hif-1a* (Hypoxia inducible factor 1, alpha subunit) and *Elk3* (ELK3, member of ETS oncogene family) (**Figure 3.6.B**, OA target). This indicated the regulons associated to those 3 TFs were the ones needing the most upregulation to reach the OA signature of the training/reference data (a NIS of 0 being the perfect match). In accordance with that observation, most WT specific regulons had a near to null NIS score in the Sham and Naïve samples, except for the *Dbp* (D site albumin promoter binding protein) regulon that showed a high standard deviation across samples (**Figure 3.6B,** WT target). In contrast, looking at DMM samples 2 weeks after surgery, highlighted that regulon related to Nfil3 and Hspa8 had small positive NIS (<0.25) so they were quite in the range of the reference OA training samples. Whereas for the *Runx2, Elk3* and *Hif1-a* regulons, the NIS was greater than 1.25, meaning they were even above the expected level. This upregulation was higher than for the corresponding sham operated samples (Sham_2w) even though 2weeks sham samples displayed a profile closer to the OA signature than the naïve and average sham operated samples, which was coherent with previous results. In the same way, the *Thrap3*, *Sox6,* and *Rpa2* regulons were the most downregulated regulons of the WT specific signature, in the 2 weeks DMM samples (**Figure 3.6B**).

3.3 Discussion

An important unfulfilled need for osteoarthritic patients is the upcoming of a disease-modifying drug, acting directly on the cause of the homeostasis disorder or blocking the drivers of the disease. The identification and/or development of such a drug requires reliable disease models and quantitative metrics to evaluate the experimental models and drugs effects on the cartilage intrinsic biology. An important read-out to evaluate how the intrinsic biology is affected by the disorder or a treatment is the establishment of specific transcriptional programs. In this study, we have applied network biology to osteoarthritis by adapting the CellNet framework, a method for the identification and evaluation of cell type specific transcriptional signatures. The

preprocessing pipeline we implemented allowed the reuse of different array platforms to generate a large perturbed transcriptomic dataset on cartilage. Thanks to this set, we have created the first skeletal cell-related CellNet classifiers, a category of cells that was missing in the public CellNet repository. The implementation of a cross-validation strategy in the algorithm demonstrated the good performance of our model to classify OA or WT samples based on the newly identified transcriptional signatures. By querying the classifiers with an independent dataset from a standard murine OA model (*274*) we validated their capacity to recognize OA over control samples. In particular, we found that the model recognized better the early-stages of OA. Overall, this provided biological insights on cartilage and OA transcriptional signatures and led to the construction of a predictive tool with promising implications for basic and clinical research, as discussed below.

Cartilage transcriptomic data, and more precisely OA related data, are largely underrepresented in public repositories when compared to other cell types or diseases (e.g. heart, breast cancer, etc…). In addition, a characteristic of omics data is the small sample size compared to the high number of observed variables. This characteristic, previously called the 'curse of dimensionality'(*275*), has been a major break to the uptake of machine learning applications in the biomedical field (*275, 276*). By reusing and assembling published datasets from various origins, we alleviated that challenge. Compiling arrays from different technical platforms have already been subject to a lot of methodological research in the past (*277, 278*), with pipelines showing good capacity for meta-analysis (*278, 279*) and with methods (e.g. batch effect removal) applicable for other types of data from more recent technologies like RNA-seq (*280*). In addition, compiling heterogeneous experiments, as we did, to increase the variance and hence the information content of the dataset has been proven to help regulatory network reconstruction, previously (*272*). Nevertheless, our final dataset remained limited in size and the introduction the cross-validation strategy into the original protocol, allowed us to make the most of this dataset. In our implementation, the GRN reconstruction was considered as a feature selection step and by placing it after the dataset splitting in the leave-one out strategy, we improved the accuracy of the performance estimator for CellNet (*281*). The high performance of the model built on that dataset demonstrated the biological relevance of the newly assembled set. This showed the power of reusing and confronting existing data, which supports the need of the ongoing efforts on the FAIR principles to ensure accessibility and reusability of data in research (*282*).

An important achievement is the reconstruction of OA and WT cartilage specific GRNs. Of course, those are not unique and do not represent the whole complexity of the real underlying mechanisms of OA. They are rather snapshots of dysregulated regulons that could be capture thanks to the disease-induced variance in our representative dataset. We found that *Hif-1a, Elk3, Nfil3, Hspa8* and *Runx1* were important transcriptional regulators of which the regulons were involved in the OA signature. Between them, they were able to modulate the expression of nearly 100 target genes, based on the data-driven network reconstruction. This identification of important factors was done with methods going beyond simple statistical differential expression, as it is frequently the case (*283*), since it was also based on information theory (e.g. CLR algorithm) (*284*, *285*). Interestingly, *Hif1a* is already known to be indispensable for chondrocyte during skeletal development (*286*). In line with our observations, Hif1a was previously reported to be expressed or even increased in OA cartilage (*287–289*). However, its exact role remains unclear since some studies report a protective role (*290*), while others suggest it would contribute to fibrosis during knee OA (*287*). Of note is that *Hif1a* is able to upregulate some targets while it downregulates others (*291*). Since the NIS score we computed is representative of the whole regulon (i.e. *Hif1a* and its target genes), analyzing more precisely the regulon's members and their expression in our transcriptional signatures might give further insights. Similarly, *Elk3, Runx1* and *Hspa8* are all factors previously reported in studies about cartilage homeostasis (*119*), osteoarthritis (*292*) or synovitis (*293*). Overall, our study suggests these signatures are relatively meaningful markers, not only because the genes were associated with biologically relevant annotations, but also because they allowed us to train performant classifiers.

The validation with an independent dataset of the GRNs-derived features and resulting classifiers confirmed their predictive power. However, the OA classifier could not fairly differentiate longer time points (8 and 16 weeks) after DMM from WT cartilage, which indicates that the GRN signatures of our reference training set are more representative of early events during OA progression. This can be explained by the fact that most OA samples in our training dataset were obtained in the first weeks (up to 6weeks) of induction of the disease, except for the 8 to 40 weeks-old mice from CBA, STR/ort lines. This specific line spontaneously develop OA in a much slower way (for example, between 12 and 20 weeks of age (*294*, *295*)), which is more representative of human rate of progression. This is an interesting feature as characterizing early OA has repeatedly been stated as an important need

for clinical research (*42*, *296*). Therefore, the current signatures are to be considered to assess early OA events and/or to be fill out with a bigger/finer panel of data.

Our strategy and the ML models we developed provide several quantitative metrics (classification score, GRN status and NIS score) to compare new cartilage samples against a reference. While CellNet was designed to optimize stem cell engineering, it is the first time this specific network biology-based concept is translated into a disease case study, so this is a first attempt to open the scope of possibilities for this approach. There are several implications of our strategy and of the ML model we developed, in the context of experimental OA models and treatment development. Indeed DMM models are often considered the most reliable and reproducible model and the most similar to human OA (*297*) but other types of experimental models (e.g. genetically induced, mechanically induced, etc..) may be considered too, depending on the research question (*296*, *297*). Those models have variable relevance to the human disease, for instance some chemically induced *in vivo* models (kaolin, collagenase) are not widely accepted as representative of human OA due to different transcriptional changes (*297*, *298*). Given, the diversity of existing *in vitro* and *in vivo* models to study OA and evaluate candidate treatments (*296*), researchers have emphasized the value and need of developing quantitative metrics, methods and assessment tools to evaluate those models, including at molecular and transcriptional level (*297*). Typical assessment tools are based on histology, gait analysis, µCT imaging, high throughput omics analysis and scoring systems (*296*, *297*), which are very important but may sometimes be subjective (*270*). On the contrary the computational strategy and the model developed here exploit further the increasing abundance of omics data. Although the current training profile would gain from further enrichment and can not yet be considered as a strong reference, this study provides a path towards systematic quantitative evaluation of established or newly developed *in vitro* or *in vivo* OA models.

Beside the evaluation of OA models, the tool we propose offers opportunities to suggest targets or monitor candidate drugs effects. First the prioritization of transcriptional regulators may suggest potential targeted intervention (to enhance an OA model or counter the disease), similarly to what was done for directed stem cell differentiation (*239*, *242*). Second, it allows for evaluation of treatment effects with regards to molecular mechanisms. For instance, we could conclude from the classification results of the validation dataset (*273*), that animals after 4 weeks of DMM surgery would be the most fit for testing candidate treatments, in that specific model. Indeed, they demonstrate the same high degree of resemblance to the OA training set as the DMM 2 weeks mice while the sham surgery-related changes have subsided contrary to the

2 weeks mice. In addition, we expect that treating DMM animals with a good candidate drug should enhance the WT transcriptional signature and lower the OA one, changes that will be quantitatively measurable and comparable between drugs thanks to this OA-CellNet tool. Future perspectives could focus on validating such a use case with a specific treatment on a DMM model, for instance.

We acknowledge that a number of parameters may affect the nature of our ML classifiers and the final output. Firstly, the OA classifier was overall more performant than the WT one, owing this to the biological heterogeneity of the WT training set. For example, the samples from sham-operated knees, which contributed to the WT training set, induced a lot of variance as some were experiencing transient changes due to the surgery. Secondly, random sample selection served to deal with an asymmetrical number of samples between different training groups in order to make sure the same amount of expression profiles represented the different groups during the training process. In our case, the WT samples outnumbered the OA ones and 7 WT samples were randomly chosen and kept out by the algorithm. Thus, the resulting GRN reconstruction and classifiers may slightly change depending on which profiles were kept out of the final model. Thirdly, during the GRN reconstruction, the CLR algorithm establishes a list of regulatory interactions between source and target genes with scores, then, the effective presence of an interaction in the final GRN is called based on a threshold system on that score. We used the default threshold value that was implemented in the CellNet R package and that led to a series of models giving remarkably good results in the context of stem cell differentiation, when validated against experimental data (*239*, *242*). Finally, the pre-processing strategy, ahead of the training steps may of course influence the output, but more precisely, it may affect the list of genes represented in the dataset. Indeed, the annotations from probeset to Esembl gene IDs and from Esembl IDs to gene names, or the labelling as transcription factors, was done by automatically fetching databases (BiomaRt, GOontology, etc..). Despite the efficiency and relative robustness of such tools and databases, maintaining one to one correspondences, taking into account the multitude of gene synonyms, sequence redundancy, knowledge update and so on, is a very challenging task. As a result, some variability in the research output may arise from those steps but the approach remains the gold standard in the current state of the art. Altogether, any modification of the GRN signature due to those parameters could affect or even refine the subsequent results. For instance, the importance of each gene in the classifiers (features' Gini importance measure) is a relative value that depends on the overall set of genes, meaning that if two genes are slightly redundant, having one of them

out from the list of features will certainly increase the importance of the other for recognizing OA (resp. WT) samples.

Finally, the binary nature of the classifiers together with the simplified training data annotations as OA or WT, might provide a too simplistic "black or white" view of osteoarthritis. Our study unveiled the potential of this tool for OA research and highlighted the need for making the distinction between different stages or types of OA and types of WT tissue (e.g. Naive vs. sham). This could be achieved with a refined dataset, therefore future improvements will be to use finer (single cell)RNA-seq training data. It will be facilitated by the recent development of the CellNet framework for bulk and single RNAseq data (*241, 299*). The existence datasets that are also evaluated histopathologically or clinically will allow sample classification according to various OA types or severity degrees (*300*) (e.g. the Mankin or OARSI (*42, 270, 271, 301*)). Recently four OA subtypes have been newly identified based on RNA-seq data, which is an important exploratory finding (*302*) but it was done with unsupervised clustering and those subtypes might not related to actual clinical phenotypes. The supervised ML approach we propose here will help relate computationally those clinical phenotypes and scores to the underlying molecular signatures, as a quantitative assessment tool for OA research.

3.4 Material and Method

Study Design

A graphical summary of the overall study design is provided in **Figure 3.1**

Microarray data assembly and pre-processing:

Several transcriptomic datasets were assembled together in order to generate a sufficiently large and informative set for the current study. To that end, six published microarray experiments related to mouse OA models were identified and an in-house pipeline for microarray platform integration was established.

The 6 microarray datasets were retrieved via their GEO accession number GSE26475, GSE33754, GSE79239, GSE33656, GSE53857 and GSE45793, for a total of 109 mouse knee joint samples (*295, 303–307*). Each dataset had a control (called wild-type (WT)) group and an OA related treatment group, with either genetically modified mice spontaneously developing

OA (e.g. CBA) or Run or DMM-induced OA. The initial annotations provided by the authors were not strictly 'WT" or 'OA' but rather experiment specific annotations such as 'DMM', 'CBA', 'Dio-/- run' or 'SHAM operated'. Therefore, the systematic binary annotation as OA-like or WT-like samples was established by hand for the current study assuming the treatment groups actually developed OA and the control ones did not, in the respective studies. The data description and annotations are reported in **supplementary material S1**.

The pipeline for data integration and preprocessing was adapted from a previously published method (*279*). The pipeline followed the same general flow no matter the array type (e.g. Affymetrix, Illumina) with some array type specific sub-steps. The general flow consisted of data pre-processing (including normalization, filtering, log2 transformation), followed by reannotation of probes with Ensembl gene IDs. Subsequently, we dealt with probes/probe sets redundancy and assembled the expression matrices as one merged dataset, followed by gene-wise normalization across experiments with batch effect removal (see pipeline workflow in **Figure 3.2.A**). The pipeline was implemented in R; Affymetrix data preprocessing was executed using the R oligo package (*308*), normalization was performed with the rma method, detection call filtering was carried out with the paCalls function using the MAS5 method (presence cutoff pval < 0.01) and probes present in at least 3 samples were kept. Equivalent Illumina data preprocessing was performed using the dedicated functions from the Matlab Bioinformatics Toolbox™. For all data, reannotation of probes into Ensembl gene IDs was carried out with the BiomaRt package (*309, 310*), Ensembl IDs redundancy was dealt with by computing the median expression across genes with the same ID. The resulting gene expression matrices were assembled by Ensembl IDs, normalized and the batch effect due to the array's origin was corrected through the application of the ComBat algorithm, which is based on the Bayes' method (*277*). The efficiency of the batch effect removal was estimated with a principal component analysis (PCA) and gene expression distribution.

An additional dataset (GSE41342) was used for validation purpose including 26 samples with naïve mice, DMM mice at 2, 4, 8 and 16weeks after surgery and the corresponding sham-operated mice (*273*). Conditions were in triplicates, except for DMM 2 weeks with 2 replicates. For the sake of comparison, this dataset was also pre-processed with the previous workflow (**Figure 3.2**), and was batch corrected together with the rest of the data to remove variation uniquely due to the originating experiment

Model training module

In order for a ML model to learn how to recognize the different classes ('OA' and 'WT'), the model has to be trained. We followed the steps of the previously published CellNet protocol, implemented in the CellNet R package (https://github.com/pcahan1/CellNet) (*241*), with some adaptations. In short, the training process was separated in two parts: (i) feature selection via cell/tissue type specific GRN reconstruction and (ii) binary classifiers building with a random forest algorithm. For the first part, transcription factors (TF) were identified in the dataset based on the following GO annotations: GO:0003700 (DNA-binding transcription factor activity), GO:0006355 (regulation of transcription) and GO:0003677 (DNA binding). Reconstructing a GRN amounted to finding regulatory interactions between TFs and target genes based on gene expression patterns across conditions and mutual information theory. This was performed via an adapted version of the context likelihood of relatedness (CLR) algorithm (*284*), with a zscore threshold of 4 applied for an interaction to be retained (default value in the CellNet R package). Then, with an enrichment analysis, cell type-enriched genes were found and, together with all their connections in the previously reconstructed GRNs, they were considered as a cell type specific GRN (*240*, *241*). For the second part, we used the cn_make_processor() and cn_makeRFs() functions from the CellNet package (based on randomForest, the R random forest algorithm (*311*)) to build binary classifiers for each cell type of the training dataset while using as features the genes from the cell type specific GRN, identified during part 1.

Cross validation pipeline

Leave-one-out cross validation was used to assess the model performance. It was implemented via specific adaptations to the CellNet R package where the training module described above was inserted in a loop in such a way that each sample would be iteratively and individually used as a test sample (**Figure 3.3**). In the cross-validation pipeline, the dataset was first split with one sample kept out (**Figure 3.3** step 1). Next, the training module, including feature selection by GRN reconstruction (**Figure 3.3** step 2) and RF classifiers building (**Figure 3.3** step 3), was applied to the training set. The resulting classifiers (or models) were used to assess and classify the holdout sample (**Figure 3.3** step 4). Subsequently, all the previous steps were repeated until all samples had been used for testing (**Figure 3.3** step 5). Finally, the class predictions and the ground truth ('OA' or 'WT' annotations) were compared and metrics such as False Positives, True Positives, Precision or Recall were computed (**Figure 3.3** step 6).

Scoring systems and metrics

The relative importance of all features (genes) for their respective classifiers was evaluated based on the mean decrease GINI measure. This measure was automatically generated during the construction of the random forest classifiers and retrieved through to the extraction function 'importance()' of the R randomForest package (*311, 312*). The class of a test or query sample (i.e. new observation), according to a newly trained model, was predicted by the CellNet 'cn_apply' function, based on the R built-in 'predict()' function.

Scoring the establishment of a cell type specific GRN was also computed with the CellNet package as described in (*240*). Briefly, the list of genes involved in the signature GRN of a reference cell type (i.e. 'OA' or 'WT') is used to compute the GRN status of a sample as the mean z-score of all those genes, weighted by their importance to the associated cell/tissue classifier. That GRN status is then normalized to the mean GRN status of the training data samples of the reference cell/tissue type.

Finally, the importance of a transcriptional regulator and its direct targets (constituting a regulon) to explain a dysregulation or a distance between a sample and a reference tissue was estimated with the network influence score (NIS). The NIS is a metric devised in (*240, 241*) and computed using to the cn_tfScore() function of the CellNet package. Three components are taken into account in this score. The first is the TF's expression change in the sample as compared to its expected expression in the reference cell type, The second is the number of predicted targets for the TF in the reference GRN. The third is the extent to which the target genes are dysregulated.

Chapter 4. Integrated *in silico-in vitro* approach for the identification of therapeutic drug targets for osteoarthritis

Based on

An integrated *in silico-in vitro* approach for the identification of drug targets against chondrocyte hypertrophy in osteoarthritis.

R. Lesage, M.N. F. Blanco, R. Narcisi, T. Welting, G.J.V.M. Van Osch, L. Geris

Manuscript submitted

BioRxive preprint : https://doi.org/10.1101/2021.09.27.461207 (*313*)

4.1 Introduction

Osteoarthritis (OA) is a degenerative disease of the joint increasingly prevalent due to the ageing population. It is a major societal burden as no disease-modifying drugs are currently available on the market (*3*). OA is characterized by cartilage damage, led by an overall increase of catabolic processes and disturbance of anabolic processes. The joint cartilage is composed of a unique cell type, the chondrocyte, which is responsible for maintaining the tissue homeostasis in an environment mainly composed of water and biomolecules such as proteoglycans and collagen fibers. Many factors, including inflammation, may influence the shift from stable healthy cartilage towards a diseased state (*314*). Regardless of the exact inducing mechanisms, during that transition, some of the chondrocytes enter a maturation process called hypertrophy (*315*, *316*), leading to extracellular matrix (ECM) degradation, mineralization and bone formation. This pathological phenomenon resembles the hypertrophic changes observed in the course of endochondral ossification during growth and development (*8*, *314*, *317–319*). Therefore, controlling the chondrocyte phenotype to prevent hypertrophic maturation has emerged as a potential therapeutic strategy to treat OA patients (*8*, *20*).

Crucial in this approach is the understanding of the process of articular chondrocyte hypertrophy for the identification of key regulators as potential drug targets. Several factors have been associated to the promotion of this phenotypic shift, such as IHH and inflammatory signaling pathways (*320*). Routes downstream of various growth factors are thought to be important in the control or disruption of chondrocyte homeostasis, such as the WNT and BMP pathways, the PTHrP, as well as the IGF-I, FGFs and TGF-B pathways (*20*, *50*, *321*). However, the interplay of intracellular pathways is highly intricate with extensive feedback loops, non-linear pathways, redundancy and intertwining (*321–323*). This complicates the intuitive prediction of what will happen in case of perturbations of a specific target. For example, it was observed that the *in vitro* activation of the WNT pathway with the WNT3A ligand and the inhibition of that same pathway with DKK1, both induced a reduction of glycosaminoglycan rich ECM in human articular chondrocytes (*324*). The fact that an activator and an inhibitor of the canonical WNT pathway both lead to the same outcome is surprising and highlights the intricacy of the underlying mechanisms. Hence, the ability to predict the effect of an external perturbation and potential therapies requires a systemic view on the process (*322*, *323*).

We propose to unravel the complexity of these regulatory pathways and to rationalize the identification of potential drug targets via screening of (combination) therapies by using a

classical engineering approach, namely that of computer modeling and simulation. Contrary to *in vitro* and *in vivo* approaches, having a systemic view of the process using an *in silico* model allows to study the system numerically, in a cost and time efficient way and with less ethical concerns. In addition, it allows to prioritize experiments, thereby refining the traditional funnel of drug targets identification in the drug discovery process. The *in silico* approach starts with acquiring a systems perspective of the intracellular biological processes. It is necessary to identify the important individual components of the system and to know how they interact and influence each other. A computational model built on this information should generate results consistent with current knowledge but also allow to ask questions that would lead to new insights into yet unexplored situations and interactions (*325*). Such computational mechanistic approaches were already used in the past to identify influential candidates for cancer therapeutics (*326*), study the control stem cell fate decision (*243*) or prioritize personalized combination therapies (*327*).

In this study, we developed an *in silico* model of the regulatory network capturing articular chondrocyte phenotypic changes during OA. This network is built by combining knowledge-based modeling and data-driven approaches to ensure the mechanistic accuracy of the network whilst taking advantage of current automatic network reconstruction technologies. We characterized the numerical intracellular states of the articular chondrocyte model and looked into consistency with known physiological behaviors, through mathematical model implementation and computer simulations. We subsequently used the model as a 'virtual chondrocyte' to perform an *in silico* high throughput screening for early predictions of the best potential therapeutic conditions to be tested in wet lab experiments. Finally, we confirmed *in vitro* the role of predicted factors in the regulation of the phenotypic change, for both single and combinations of factors.

4.2 Results

4.2.1 Highly interconnected intracellular networks regulate articular chondrocytes: building a mechanistic model.

In order to build an articular chondrocyte model, we gathered known biological mechanisms from the literature into an activity flow graph or network (**Figure 4.1).** In this graphical influence network, or activity flow representation, the nodes represent biological components (proteins or genes, see definitions **Table S4.1**) important for chondrocyte biology (*20*). Directed

edges linking two nodes represent interactions or activating/inhibitory influences between source and target proteins or between transcription factor (TFs) and target genes. Information was collected and adapted from a previously published model of growth plate chondrocytes (*254*, *255*) as well as from additional deep literature and database curation (see Method section). The annotations, descriptions and cross-references for each network node and its interactions can be consulted in two interactive networks, a protein signaling one and a gene regulatory one. The former goes from the growth factors binding their respective receptors down to the TF entering the nucleus. The latter represents a network of transcription factors regulating the expression of their target genes, coding for the corresponding proteins in the signaling network (**Figure 4.1**). The two networks are interconnected as each biological component is represented by a gene in the GRN and its corresponding protein in the signaling network. The combination of both networks can be regarded as an interactive knowledge base on chondrocyte signaling and is available through the online platform Cell Collective (*269*). We refer the reader to that interactive knowledge base for the literature support of our network model (*328*, *329*), while the main pathways that were included are listed below.

Unless stated otherwise, all components of the network and of the associated *in silico* model are referred to with names in upper cases that designate the numerical variables, which represent neither the protein nor the gene but a product of both. Therefore, neither the gene nor the protein official nomenclatures are used. A list of correspondence between the numerical variable's names and actual mouse gene names is available in **Table S4.2** and reported in the Cell Collective interactive networks, for information.

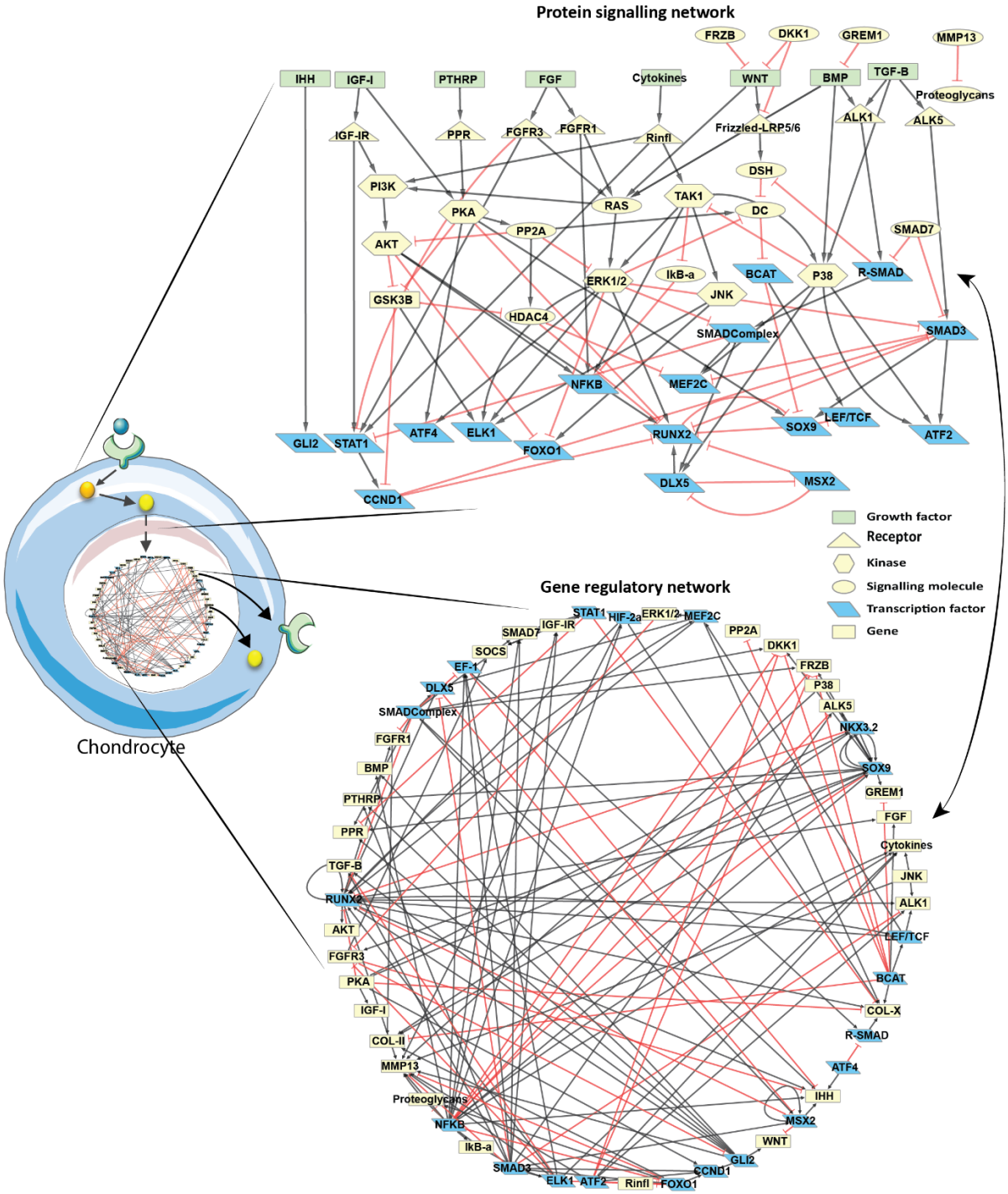


Figure 4.1: Influence map of the signaling and GRN implemented in the model. Red, T-ended arrows represent inhibitory influences and black arrows represent activating influences. On the signaling side, growth factors and pro-inflammatory cytokines are represented by green rectangular nodes, receptors are yellow triangles, kinase proteins are yellow hexagons and other signaling proteins are yellow ellipses. Transcription factors (TFs) are represented in blue both in the signaling network and the GRN. In addition, target genes are represented by yellow rectangles in the GRN and TFs might also be targets of other TFs in the GRN. Each biological component is represented by a gene in the GRN and a protein in the signaling network, except for nodes that are not involved in one of the layers (e.g. NKX3.2 only plays a role in the GRN while COL-II and COL-X do not have upstream or downstream influences in the protein signaling network)

The regulatory pathways represented in the network include the canonical and non-canonical WNT BMP pathways, the PThrP and IHH pathways, as well as the IGF-I, FGFs and TGF-B pathways since they are all reported to play a role in chondrocyte fate decisions (*20*, *62*). The influence of pro-inflammatory cytokines, such as interleukine 1Beta (IL-1B) or TNF-A, was summarized through a node labelled 'cytokines', in the network (**Figure 4.1**). That node is able to signal through a single receptor (labelled 'Rinfl') that can activate well known downstream pathways such as the PI3K/AKT axis, the NFKB pathway and MAPK pathways. These MAPKs in question include ERK1/2, c-Jun N-terminal kinase (JNK) and P38. For each of the introduced pathways we represented the downstream signaling cascades and known transcription factors as well as their target genes in the nucleus. Examples of important transcription factors that were included are the IHH signal tranducer GLI2, STAT1, TCF7, MEF2C as well as SOX9, a marker of differentiated healthy chondrocytes, and RUNX2, a hypertophy marker. All pathways in the model are strongly interconnected as shown by **Figure 4.1.**

In total, there are 60 biological components in the network, which are listed in **Table S4.2**. Each component accounts for a different biochemical entity including 8 growth factors, 8 receptors, 20 transcription factors, 4 ECM proteins and 20 signaling molecules of another type. Combining the signal transduction and the GRN into one connected network, leads to a total of 264 direct or indirect biochemical interactions. Each node has on average 7.2 direct neighbors (i.e. directly connected node) and the average shortest path (i.e. smaller number of edges) to connect two nodes is 3.01. Hence, all nodes may virtually influence others by means of a couple of intermediary components, supporting the need to represent and study that network mathematically to conclude on key controllers.

4.2.2 Learning from transcriptomic OA data by complementing the mechanistic network with data-inferred interactions

We complemented the knowledge based GRN with data-driven network inference, allowing for identification of *de novo* regulatory links, introduction of previously unstudied or undiscovered interactions and the reduction of the bias related to human literature curation. To that end, we generated an informative cross-platform merged dataset on mouse osteoarthritic cartilage. It was composed of 109 samples coming from 6 microarray experiments from which we selected a subset corresponding to the expression profile of 41 genes (see Method and **Table S3.1**). The selected genes were the ones present in or closely related to the biological factors from the

mechanistic model. The purpose was to identify potential regulatory interactions among those genes of interest only (see Method). The sub-dataset was normalized and corrected for batch effects originating from the differences in microarray platform technologies, as described in the Method section. Our data indicate that the gene expression distribution was correctly normalized among the different microarrays (**Figure 4.2A**). The principal component analysis (PCA) showed the inter-array variance was strongly reduced after merging and correction, thereby highlighting successful removal of the batch effect. As a quality control, we applied an unsupervised clustering method to the merged dataset to evaluate whether the biological information was still maintained after merging and correction. The resulting heatmap features the clustering of the 109 samples as well as their prior annotations as "OA" and "WT" (for wild-type or control samples) **(Figure 4.2B)**. See Method and **Table S3.1** for annotations definition. While splitting the clustering result in three groups, we could identify an OA-like group and a WT-like group as well as a WT-like sample clustering alone. Based on these categories, 81% of the OA tagged samples were correctly identified in the OA group whereas 56% of the WT-tagged samples were correctly located in the WT group. These accuracy levels are similar to the ones achieved when clustering each of the individual microarray experiments separately, before batch effect removal (**Figure S4.1**). We concluded that, despite the diversity of the technical platforms used in the assembled dataset, most of the variance was not due the arrays' underlying technology but rather due to the biologically meaningful information usable as input for GRN reconstruction.

In order to complement the GRN with new data-derived hypothetical interactions, we inferred genetic interactions out of this newly assembled dataset. Directed edges between nodes were added in the GRN to account for newly inferred regulatory interactions, as represented in **Figure 4.2C**. We inferred regulatory interactions directed from TFs towards target genes by using three different algorithms to avoid algorithm-specific bias on the inferred edges (see Method) (*221*). Only interactions predicted by all the three algorithms were implemented in the gene regulation layer of the mechanistic model. In addition, the sign of the Spearman correlation score allowed us to define whether the interactions were genetic activations or inhibitions (**Figure 4.2C**). The inferred interactions are reported with the correlation factors in **Figure 4.2D.** Note that the machine learning methods predicted some TFs as descriptor of a given target gene even for pairs that had not a high correlation score (**Figure 4.2D**). This showed that the ensemble of algorithms we used went beyond the simple correlation for inferring genetic interactions, relying on the concept that some form of covariation is implied by a causal

relationship. The inferred interactions were included in the GRN part of the model for all subsequent analyses presented in this study.

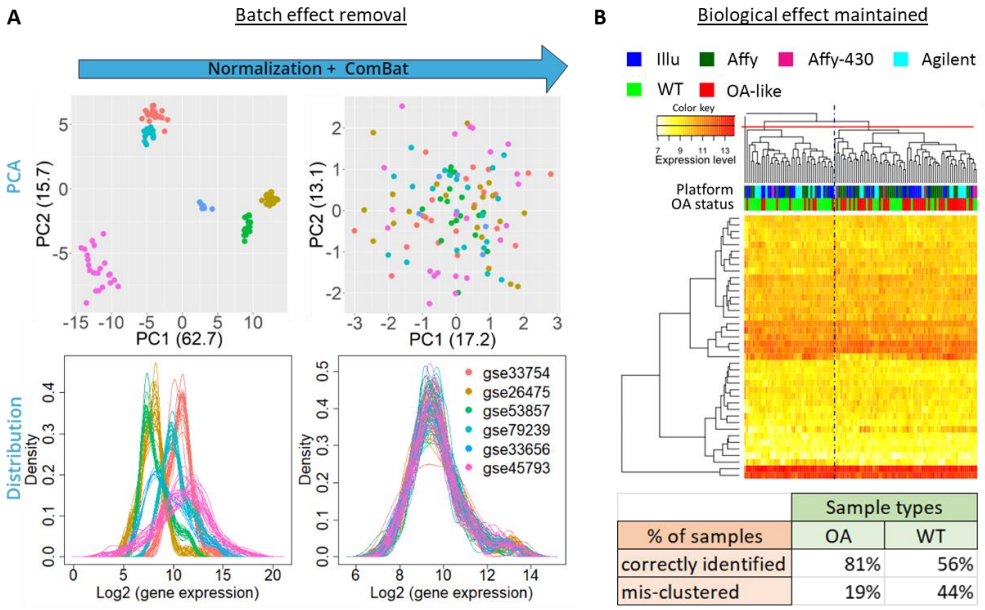


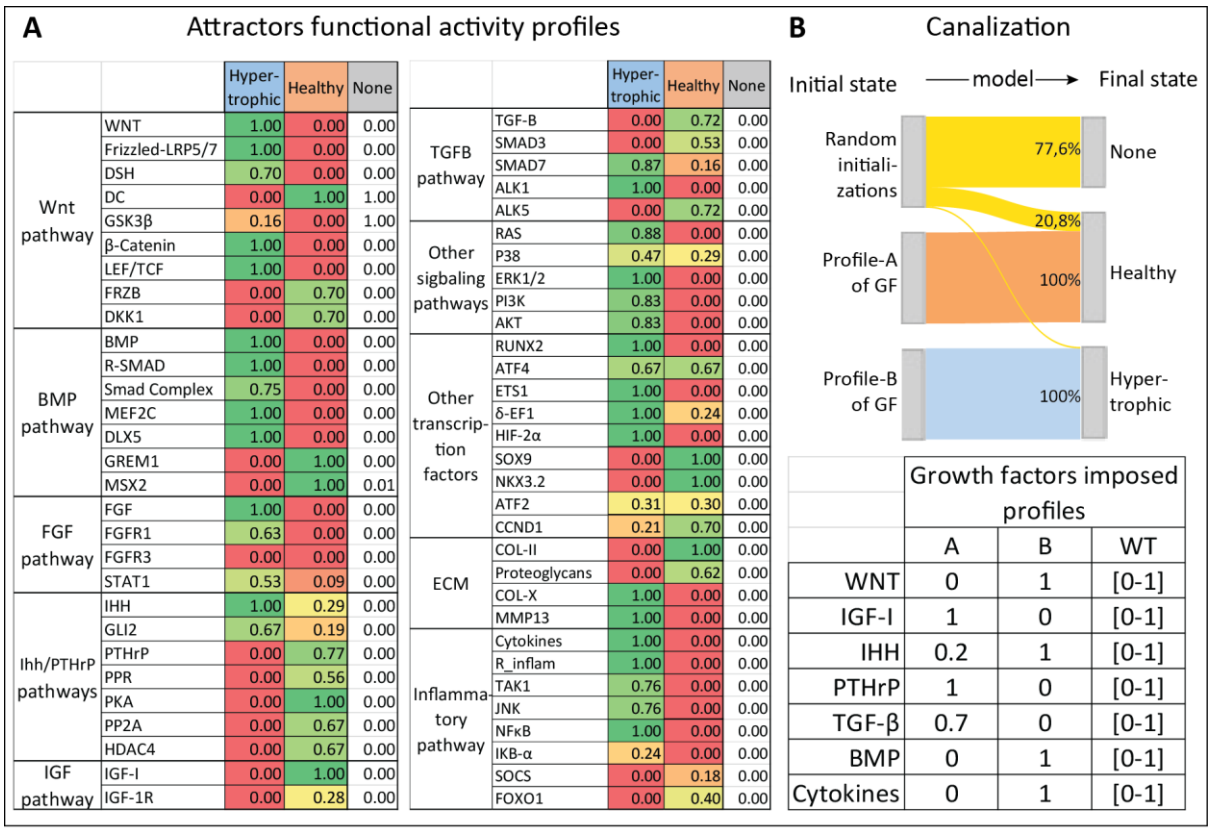
Figure 4.2: Micro array integration for GRN inference. (A) Assembling and correcting the microarray sub-datasets. PCA plots and gene expression distributions of the assembled dataset colored by arrays (GSE-number), before and after quantile normalization and batch effect correction with the Combat algorithm. **(B)** Unsupervised clustering with the Euclidian complete method highlights that the samples do not cluster according to the technological platform but rather according to the OA status of the samples. When splitting the hierarchical tree into three branches, a mostly OA group and a mostly WT group stand out. The table summarizes the percentage of true OA (resp. WT) samples correctly grouped in the 'OA-group' (resp. 'WT-group') highlighting a clustering accuracy in line with what is obtained for the individual sets before correction. **(C)** Starting from a matrix of gene expression as input, algorithms compute possible regulatory interactions and output a list of possible transcriptional regulatory interactions from a gene coding for a transcription factor to another gene. 'g1', 'g2' and 'g3' denote gene1, 2 and 3, respectively. **(D)** List of interactions inferred with the merged OA dataset and integrated into the mechanistic model. Inference was run with three algorithms, solely interactions that were present in the results of the three algorithms were reported and integrated. An interaction was considered present for one algorithm if its score was higher than a threshold defined as the difference between the mean and standard deviation of all scores. Corr.score is the spearman correlation coefficient, computed solely to define the interaction sign (activation if positive, inhibition if negative).

4.2.3 The computational model successfully recapitulated two chondrocyte phenotypes and physiologically relevant behaviors

We translated the aforementioned regulatory network into mathematical equations in order to develop an executable numerical model of the articular cartilage chondrocyte. We used a semi-quantitative additive modeling formalism with priority classes as it allows to study large networks without requiring much information on kinetics parameters. Each node takes on a continuous value in the interval [0,1], representing the global functional activity of that node, defined as the multiplication of the gene expression level and the protein activation potential. That way, a protein cannot exert its function on downstream targets unless it is both expressed at the genetic level and activated/not blocked at the post-transcriptional level (global functional activity > 0) (see definitions **Table S4.1**).

With the above described numerical chondrocyte model, we studied the system free of fixed external cues to identify all potential mathematical stable states that may emerge naturally. These stable states are also called attractors or genetic modules and they equate potentially existing cell phenotypes (see definition of attractors in **Table S4.1**). They were evaluated using methodologies similar to those used with logical models (*213*). Any initial state inputted into the system of equations is like a set of external stimuli that would trigger signal transduction inside a cell, eventually leading to a specific cell state. By randomly initializing the *in silico* model (see method for Monte Carlo analysis) we were able to explore possible model outcomes and we observed three emerging attractors. Each of these attractors had a unique global activity profile for the 60 components as reported in **Figure 4.3A;** the details of the protein activation and gene expression level for each component is available in **Table S4.3**. Two of the attractor states that we found were biologically relevant, meaning they were comparable to existing chondrocyte phenotypes identified based on known cell state biomarkers such as type II and type X collagen (COL-II and COL-X, respectively) , RUNX2 or SOX9. We identified one of the attractors as a normal healthy articular chondrocyte since markers such as SOX9, NKX.3.2 and COL-II were strongly active and expressed while RUNX2, COL-X and matrix metalloproteinase 13 (MMP13) were inactive or not expressed. In addition, the inflammation-associated nodes had low activity. That correlates with what is known of real chondrocytes during homeostasis. The second attractor corresponded to a hypertrophic-like chondrocyte since RUNX2, COL-X and MMP13 were present or active (i.e. global activities equaling to 1) while SOX9 was not and COL-II and the proteoglycans were degraded and/or not expressed

(functional activities nearly zero). In addition, the WNT and inflammation related pathways were active (e.g. WNT =1, DC=0, βcatenin = 1, Cytokines =1, NFκB = 1, TAK1= 0.76). The third state we found had nearly all protein activities close to zero, consequently, we couldn't associate it with any specific phenotype as it was more likely a trivial mathematical solution. We named it the 'None' attractor as it was neither healthy nor hypertrophic **Figure 4.3A**.



		Hyper-trophic	Healthy	None
Wnt pathway	WNT	1.00	0.00	0.00
	Frizzled-LRP5/7	1.00	0.00	0.00
	DSH	0.70	0.00	0.00
	DC	0.00	1.00	1.00
	GSK3β	0.16	0.00	1.00
	β-Catenin	1.00	0.00	0.00
	LEF/TCF	1.00	0.00	0.00
	FRZB	0.00	0.70	0.00
	DKK1	0.00	0.70	0.00
BMP pathway	BMP	1.00	0.00	0.00
	R-SMAD	1.00	0.00	0.00
	Smad Complex	0.75	0.00	0.00
	MEF2C	1.00	0.00	0.00
	DLX5	1.00	0.00	0.00
	GREM1	0.00	1.00	0.00
	MSX2	0.00	1.00	0.01
FGF pathway	FGF	1.00	0.00	0.00
	FGFR1	0.63	0.00	0.00
	FGFR3	0.00	0.00	0.00
	STAT1	0.53	0.09	0.00
Ihh/PTHrP pathways	IHH	1.00	0.29	0.00
	GLI2	0.67	0.19	0.00
	PTHrP	0.00	0.77	0.00
	PPR	0.00	0.56	0.00
	PKA	0.00	1.00	0.00
	PP2A	0.00	0.67	0.00
	HDAC4	0.00	0.67	0.00
IGF pathway	IGF-I	0.00	1.00	0.00
	IGF-1R	0.00	0.28	0.00

		Hyper-trophic	Healthy	None
TGFB pathway	TGF-B	0.00	0.72	0.00
	SMAD3	0.00	0.53	0.00
	SMAD7	0.87	0.16	0.00
	ALK1	1.00	0.00	0.00
	ALK5	0.00	0.72	0.00
Other sigbaling pathways	RAS	0.88	0.00	0.00
	P38	0.47	0.29	0.00
	ERK1/2	1.00	0.00	0.00
	PI3K	0.83	0.00	0.00
	AKT	0.83	0.00	0.00
Other transcrip-tion factors	RUNX2	1.00	0.00	0.00
	ATF4	0.67	0.67	0.00
	ETS1	1.00	0.00	0.00
	δ-EF1	1.00	0.24	0.00
	HIF-2α	1.00	0.00	0.00
	SOX9	0.00	1.00	0.00
	NKX3.2	0.00	1.00	0.00
	ATF2	0.31	0.30	0.00
	CCND1	0.21	0.70	0.00
ECM	COL-II	0.00	1.00	0.00
	Proteoglycans	0.00	0.62	0.00
	COL-X	1.00	0.00	0.00
	MMP13	1.00	0.00	0.00
Inflamma-tory pathway	Cytokines	1.00	0.00	0.00
	R_inflam	1.00	0.00	0.00
	TAK1	0.76	0.00	0.00
	JNK	0.76	0.00	0.00
	NFκB	1.00	0.00	0.00
	IKB-α	0.24	0.00	0.00
	SOCS	0.00	0.18	0.00
	FOXO1	0.00	0.40	0.00

	Growth factors imposed profiles		
	A	B	WT
WNT	0	1	[0-1]
IGF-I	1	0	[0-1]
IHH	0.2	1	[0-1]
PTHrP	1	0	[0-1]
TGF-β	0.7	0	[0-1]
BMP	0	1	[0-1]
Cytokines	0	1	[0-1]

Figure 4.3: Predicted chondrocyte profiles and canalization for the three emerging states during Monte Carlo analyses. The Monte Carlo analyses consist in sampling 10.000 random initial states for the variables, with the possibility to impose constraints (similar to external biological cues). **(A)** The Monte Carlo analysis without constraints highlights the existence of three final states (i.e. attractors) with different activity profiles (in columns). The global activity of a protein is presented in this table as the product of the predicted gene expression and the protein activation level. The table only reports the global activity; a complete table including the gene expression and protein activation levels is available in **Table S4.3**. Rows represent the variables of the model and are grouped by pathways or functional groups. **(B)** Specific activity profiles were imposed to the growth factors, as reported in the table, while variables other than growth factors were initialized randomly. Profile A represents a possible healthy environment. Profile B represents a more pathological environment. The 'Random' column indicates the case without constraints in which initial activities were randomly sampled within the interval [0,1]. In the Sankey diagram initial states are on the left and final destination states (i.e. attractors) on the right. Strips indicate the percentage of initializations (among the 10.000) that reached each of the possible attractors during the Monte Carlo without constraints ('Random initializations').That percentage is also reported for the Monte Carlo with constraints (profile A & B)

The number of random initializations reaching an attractor during the Monte Carlo simulation gives a sense of the overall attractor's size (attractors are reported in **Figure 4.3B)**. Most of the

random initial states led to the 'None' final state, reflecting the fact that most of these random initial values might very well be nonsensical from a chondrocyte biology perspective. In addition, about 21% of the initializations led to the healthy state and about 2% to the hypertrophic state (**Figure 4.3B**). These attractors are the spontaneously emerging states arising from the proposed map of biochemical interactions.

We also sampled random initial states while fixing seven growth factors at values that were physiologically relevant in either a normal healthy or a disease environment (profile A and B of Figure3B, respectively). Interestingly, when imposing the profile A to the growth factors during the Monte Carlo, the system only settled into the healthy state and when imposing profile B it settled into the hypertrophic state, thereby abolishing the 'None' state (**Figure 4.3B**)

With the aim of assessing whether the model could successfully recapitulate relevant physiological behaviors, we established a tool enabling to test experimentally observable scenarios. As a first step to validate our model, we used this tool to test specific scenarios for which the expected outcome was known from literature or hypothesized from clinical observations. For instance, inflammation in the knee is one of the symptoms of OA and has been shown to be one of the drivers of cartilage degradation, possibly by pushing chondrocytes to undergo hypertrophy and produce matrix-degrading enzymes (*330, 331*). For that reason, inflammation-related targets are subject to several investigations for potential OA therapies. Interestingly, *in silico* experiments with our model showed that blocking important transducers of inflammation such as the TGF-β-activating kinase (TAK1) or NFKB while activating the PTHrP related pathway could push a hypertrophic-like chondrocyte into transitioning towards a more healthy or anabolic state. Moreoever, other studies have reported that the TGF-β pathway had a protective effect against inflammation (*251, 332–334*), a scenario we evaluated with our model too. *In silico* mimicking the presence of inflammation in a healthy chondrocyte by forcing several inflammation related pathways of the model to be set at their highest values led to 100% of transition towards the diseased hypertrophic state (**Figure 4.4A**). This effect was partially rescued by concomitantly forcing the presence of TGF-β since 5.3% ±1.2 of the perturbations failed to exit the healthy state, thus, confirming *in silico* that TGF-β could have a protective effect against inflammation through the mechanisms present in the model. Nevertheless, the role of TGFβ in OA has been reported to be dual as this growth factor transduces signals in chondrocytes mainly via two receptors, the type I TGFβ receptor (ALK5) and the type II TGFβ receptor ALK1 (*88*). They are involved in different intra-signaling routes and depending on which receptor is activated, the downstream-activated signals would be rather

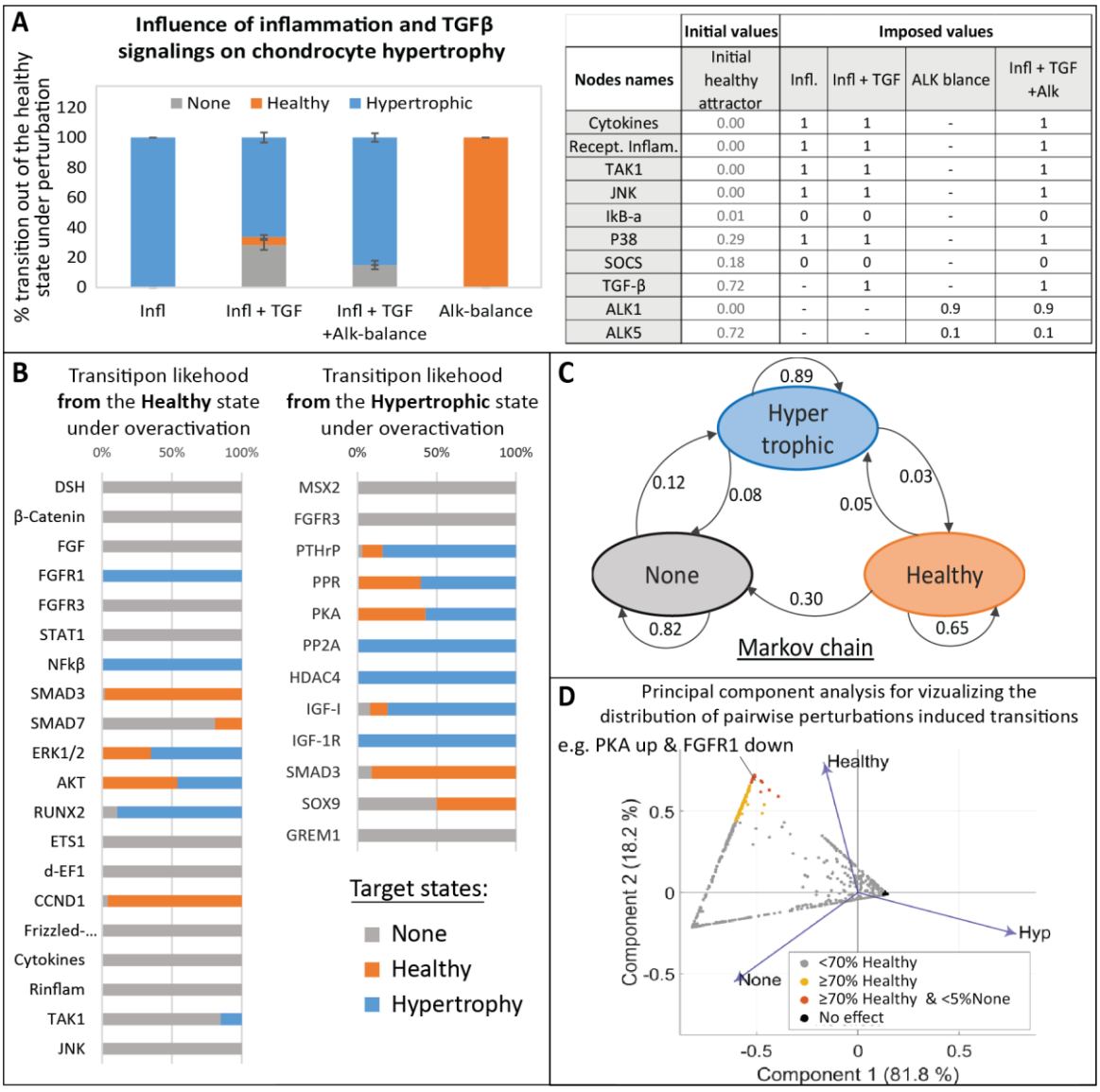
anabolic (ALK5) or catabolic (ALK1) (*88*) and impact chondrocyte maturation differentially (*335*). Clinical observations reported that the ALK5/ALK1 balance decreased with age and in OA patients (*87*, *93*, *336*). *In silico* simulations with our model showed that, roughly, the rescue by TGFβ was lost when the ratio between the receptors was forced to be $\frac{ALK1}{ALK5} > 1$ (**Figure 4.4A** and **Figure S4.2**). For higher values of ALK1, ALK1 activity could be as low as 0.86*ALK5 and still show the loss of the TGF-β protective effect (**Figure S4.2**). This is in line with what was previously modeled (*332*), demonstrating that the decrease in that balance could explain why TGF-β loses its protective effect against inflammation in OA patients.

Together, these results show the ability of the articular chondrocyte *in silico* model to behave in a physiologically relevant way and predict emerging effects qualitatively, highlighting the pro- or anti-hypertrophic nature of biological components in specific conditions. The aforementioned tool was further implemented through an executable App (https://github.com/Rapha-L/Insilico_chondro (*337*)) allowing users, such as biologists, to easily test hypotheses by performing *in silico* experiments on the virtual chondrocyte (see interface **Figure S4.3**).

4.2.4 *In silico* experiments on the modeled system predicted potential important nodes to control chondrocyte fate.

We next decided to exploit further the model by studying the effect of all possible perturbations of each component. Starting either from the healthy or the hypertrophic-like state, variables were individually activated or inhibited in a systematic manner. Over-activation of the FGF receptor 1 (FGFR1) or NFKB were the most potent conditions to trigger a transition from the healthy towardsthe hypertrophic state. To a lesser extent, activation of the variables for ERK1/2 kinases, the AKT member of the PI3K/AKT axis, the RUNX2 transcription factor or JNK kinases also promoted this diseased transition (**Figure 4.4B**). On the other hand, transition from the hypertrophic towards the healthy state was mainly triggered by forced activation of the TGFβ intracellular effector SMAD3 and partially brought about by activation of the SOX9 transcription factor, IGF-I and members of the PTHrP pathway (i.e. the protein kinase A (PKA), the PTHrP receptor (PPR) and PTHrP). This constitutes a state transition study that can be summarized in the Markov Chain representation (**Figure 4.4C**) with the system's overall probability of transitioning from one state to another under random single node perturbations. Interestingly, despite the small size of the hypertrophic attractor in the random canalization

(**Figure 4.3B**), the transition study shows that the total probability of transitioning out of that hypertrophic state under random single node perturbations is 0.11. So, in 89% of the cases, a single node perturbation will not affect this state. This means that even if the hypertrophic state is difficult to reach in the articular cartilage system, it is particularly robust to small single perturbations, and once the numerical system has reached that state, it is unlikely to escape from it, via transitioning to any other state, with a single targeting strategy. For this reason, we also systematically investigated the effect of all possible combinatorial perturbations of two constituents (or pairwise perturbation). The full screening amounted to $\binom{60}{2} \times 4 = 7080$ tested conditions per attractor. 94% of the variance was explained by the first two components in the PCA reporting the effect (percentage of transition to destination attractors) of the combinatorial treatments on a hypertrophic-like chondrocyte **(Figure 4.4D)**. We searched for the most potent conditions to retrieve the healthy state from a hypertrophic chondrocyte (by requiring more than 70% of perturbations input transitioning to the healthy state and less than 5% to the 'None' state), which could point towards potential drug therapies for OA. Based on those results, the most efficient way to transition from a hypertrophic towards a healthy chondrocyte in the *in silico* model was with the up-regulation of SMAD3 in combination with activation or inhibition of numerous other targets such as inhibition of the inflammatory mediator NFKB, inhibition of the DLX5 transcription factor or activation of SOCS, a blocker of pro-inflammatory signals transduction. Activation of PKA/PPR in combination with inhibition of various targets, such as WNT or FGFR1, also seemed to decrease hypertrophy successfully in the model with 100% of transitions towards the healthy state (data no shown). Therefore, the PKA/PPR axis and SMAD3 seemed to be 'enablers' that could 'unblock' the *in silico* system, facilitating the effect of other relevant targeting treatments. Moreover, some predictions among the ones triggering more than 70% of transition towards the healthy state, in **Figure 4.4D,** did not include the two aforementioned enablers. For example, the up-regulation of ALK5 in combination with the down-regulating ALK1, the two receptors of TGF-B in the model, gave more than 90% of transition towards the healthy state. Additionally, inhibition of the WNT pathway while activating ALK5 as well as activation of IGF-I while activating the DC involved in the WNT pathway allowed between 70 and 90% of transition towards the healthy state, to mention but a few (data not shown). So, the *in silico* model and associated screening algorithms enabled the prediction of pairwise targeting conditions with a good potential against chondrocyte hypertrophy.



Nodes names	Initial values: Initial healthy attractor	Imposed values: Infl.	Infl + TGF	ALK blance	Infl + TGF +Alk
Cytokines	0.00	1	1	-	1
Recept. Inflam.	0.00	1	1	-	1
TAK1	0.00	1	1	-	1
JNK	0.00	1	1	-	1
IkB-a	0.01	0	0	-	0
P38	0.29	1	1	-	1
SOCS	0.18	0	0	-	0
TGF-β	0.72	-	1	-	1
ALK1	0.00	-	-	0.9	0.9
ALK5	0.72	-	-	0.1	0.1

Figure 4.4: Study of the virtual chondrocyte state transition and *in silico* screening of target perturbations. (A) Relation between Inflammation and TGFβ and influence on the chondrocyte state. Perturbations are applied on the healthy attractor, bar height gives the average percentage of transition towards one of the target states, error bar denotes standard deviation. 'Infl.' refers to imposed inflammation, 'TGF' refers to TGFβ over-activation and 'Alk balance' to the modification of the ratio between TGFβ receptors (ALK1 and ALK5). Conditions were mimicked as described in the table. '-' denotes no modification of the initial value. A transition from 'healthy' to 'healthy' means no transition. **(B)** All single node perturbations triggering a state transition from the Healthy (resp. Hypertrophic) state. **(C)** Markov chain providing the overall probability of transition from one state to another, under single node perturbations. Arrows indicate transitions from an initial state towardsa target state with the associated probability. Thus, the total probability of outgoing arrows for any state is 1.0. **(D)** PCA visualizing the results of the systematic screening of all possible combinatorial perturbations on a hypertrophic-like chondrocyte. Each dot represents one of the 7080 screened conditions. Principal components are computed based on the percentage of transitions towards the three attractors, reported as eigenvalues (blue arrows). Dot colors correspond to threshold in the percentage of transition towardsthe healthy state for potential OA therapies.

4.2.5 Newly predicted (combinatorial) treatments were validated *in vitro* for their potential to prevent hypertrophy

We used ATDC5s, a chondrogenic murine cell line able to undergo hypertrophy, in order to *in vitro* validate the *in silico* findings obtained from the model. We measured the activity ALP, an enzyme typically secreted during hypertrophy and participating to ECM mineralization. It showed that there was a positive linear correlation between the level of ALP activity in the medium and the expression level of the hypertrophic gene *Col10a1* during the ATDC5 differentiation (**Figure 4.5B**). Hypertrophy was further increased when supplementing the differentiation medium with Ihh, as expected (**Figure 4.5B**). Indeed this growth factor is known for its pro-hypertrophic effect on ATDC5 (*338*). These results allowed us to evaluate ALP activity in the medium as a proxy for the level of hypertrophy, significantly increasing the throughput of the experimental set-up for testing small molecule treatments.

This semi-high throughput system was used for validating the *in silico* predictions. *In vitro* experiments evidenced that treatments with Forskolin, an activator of PKA activity or with an activator of the BMP/SMAD3 pathway, were sufficient to prevent the increase in the activity of secreted ALP and thereby were sufficient to block hypertrophy in ATDC5s (**Figure 4.5C)**. Additionally, Forskolin treatment decreased *COL10A1* gene expression when applied to primary human OA chondrocytes ($p < 0.01$) (**Figure 4.5C**). Together these results corroborated the *in silico* predictions that activation of PKA was sufficient to block hypertrophic differentiation in chondrocytes (**Figure 4.5C & D).**

Additionally, several combinatorial treatments predicted by the *in silico* screening were tested in ATDC5 and compared to the corresponding individual treatments, to assess their efficacy. In **Figure 4.5D**, ALP activities normalized to total DNA content are reported by means of z-score for all pairwise and single conditions. The ALP activity in the medium during hypertrophic differentiation was lower than the corresponding controls in 6 out of the 8 predicted combinations. The two conditions for which the strongest effect was measured were the inhibition of FGFR1 combined with the activation of PKA ('Forskolin + PD161570', $p = 0.032$) and the inhibition of BMP combined with PKA activation ('LDN-193189 + Forsk', $p = 0.014$). These two combinations seemed to generate an added effect compared to the single drugs. Contrary to *in silico* model predictions, treatment with exogenous IGF-I combined with BMP inhibition ('IGF-I + LDN-193189', $p= 0.105$) or inhibition of ERK1/2 combined with

PKA activation ('PD0325901 + Forsk', p = 0.264) did not show a significantly lower hypertrophic level, based on ALP activity (**Figure 4.5D**).

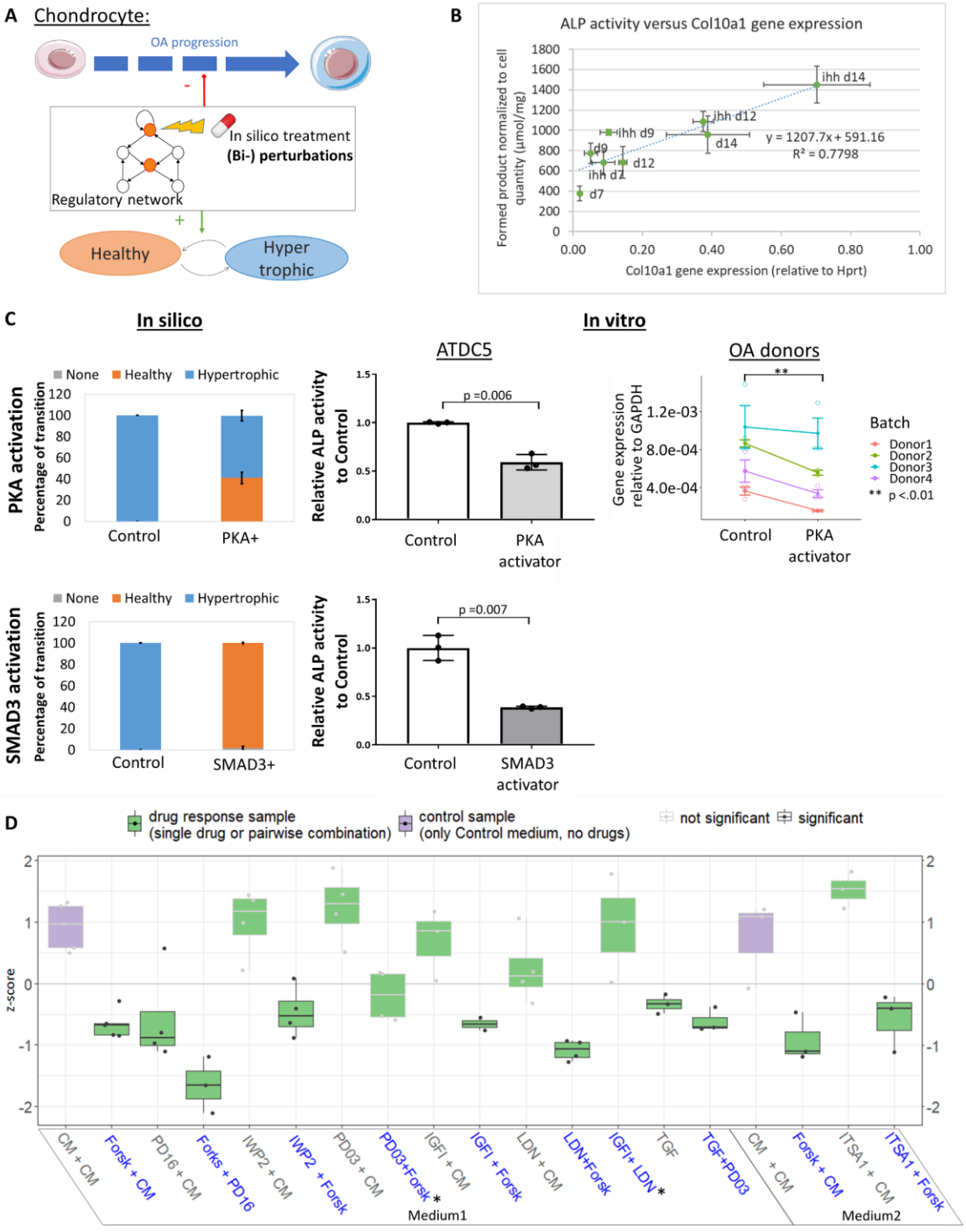


Figure 4.5: *In vitro* validation of *in silico* predictions on chondrocyte phenotype changes. (A) Concept of *in silico* identification of potential drug targets. **(B)** Secreted ALP activity, relative to DNA quantity, is positively linearly correlated to Col10a1 gene expression during hypertrophic differentiation with and without Ihh treatment. Each point is the average of 3 replicates and bars denote standard deviation. **(C)** Effect of PKA or SMAD3 activation as measured *in silico* and *in vitro* in ATDC5 (N = 3

replicates, histograms show average fold change in ALP activity relative to control and bars are standard deviations, p-values are computed with one-tailed t-test and welch's correction) and chondrocytes from OA donors (N= 4 donors with 3 replicates each, p-value is computed with one-tailed linear mixed effect model). *In silico* activation was performed by setting the nodes to their maximum value (1.0), *in vitro* PKA (resp. SMAD3) activation was performed with Forskolin 1µM (resp. Activin 100ng/ml) for 24h. **(D)** Single and combinatorial drug screening in ATDC5 with selected conditions based on *in silico* predictions. Boxplot of the series of conditions across independent replicates (z-scores of ALP activity fold change) with control conditions in purple. Conditions significantly lower than the control (combined p-value < 0.05) have dark grey borders and dots (Wilcoxon rank-sum test with BH correction and combined probabilities over independent runs). For each condition, dots are the average of biological triplicates, summary statistics are represented by a horizontal line for the median of independent experimental repetitions and a box for the interquartile range. The whiskers extend to the most extreme data point that is not >1.5 times the length of the box away from the box. Blue labels indicate potent conditions predicted by the *in silico* model, grey labeled conditions are added to the experimental set-up for information. CM stands for 'control medium', medium1 has 0.02% of DMSO and medium 2 0.035%.*indicates *in silico* predicted conditions that did not show a significant decrease of ALP activity *in vitro*.

Elaborating one of the conditions that showed the strongest response, PKA activation combined with FGFR1 inhibition, we compared the combinatorial effect with the corresponding single drug treatments. The combinatorial effect was greater than the one for either of the single drug, for both tested concentrations ratios (**Figure 4.6A**). This suggests that activating PKA (resp. inhibiting FGFR1) would potentiate or enable the effect of FGFR1 inhibition (resp. PKA activation) by blocking or unblocking key pathways and maintaining the necessary constraints on the system. Dose curve relationships need to be established to confirm that hypothesis. Screening various values of functional activities of PKA and FGFR1 with the virtual chondrocyte and looking at the percentage of transition out of the hypertrophic state, showed that a minimal level of PKA activity (namely about 0.4 on a scale from 0 to1) was required to achieve any positive effect with this combinatorial treatment (**Figure 4.6B)**. In addition, a gradient effect was observed suggesting that the lower the PKA activity, the more we needed to block FGFR1 to achieve an equivalent positive effect (**Figure 4.6B**). *In vitro* validation confirmed this dose effect since the overall gradient shape was very comparable between the *in silico* and *in vitro* situation **(Figure 4.6B)**. Comparing the diagonal (combinations of two drugs) to the single drug ranges also highlighted a likely synergistic effect in decreasing hypertrophy between the two drugs, at the tested 1:5 ratio. Taken together, these results support that targeting the regulatory mechanisms at multiple points might be necessary to maintain a physiologically healthy state for a chondrocyte experiencing hypertrophy inducing cues.

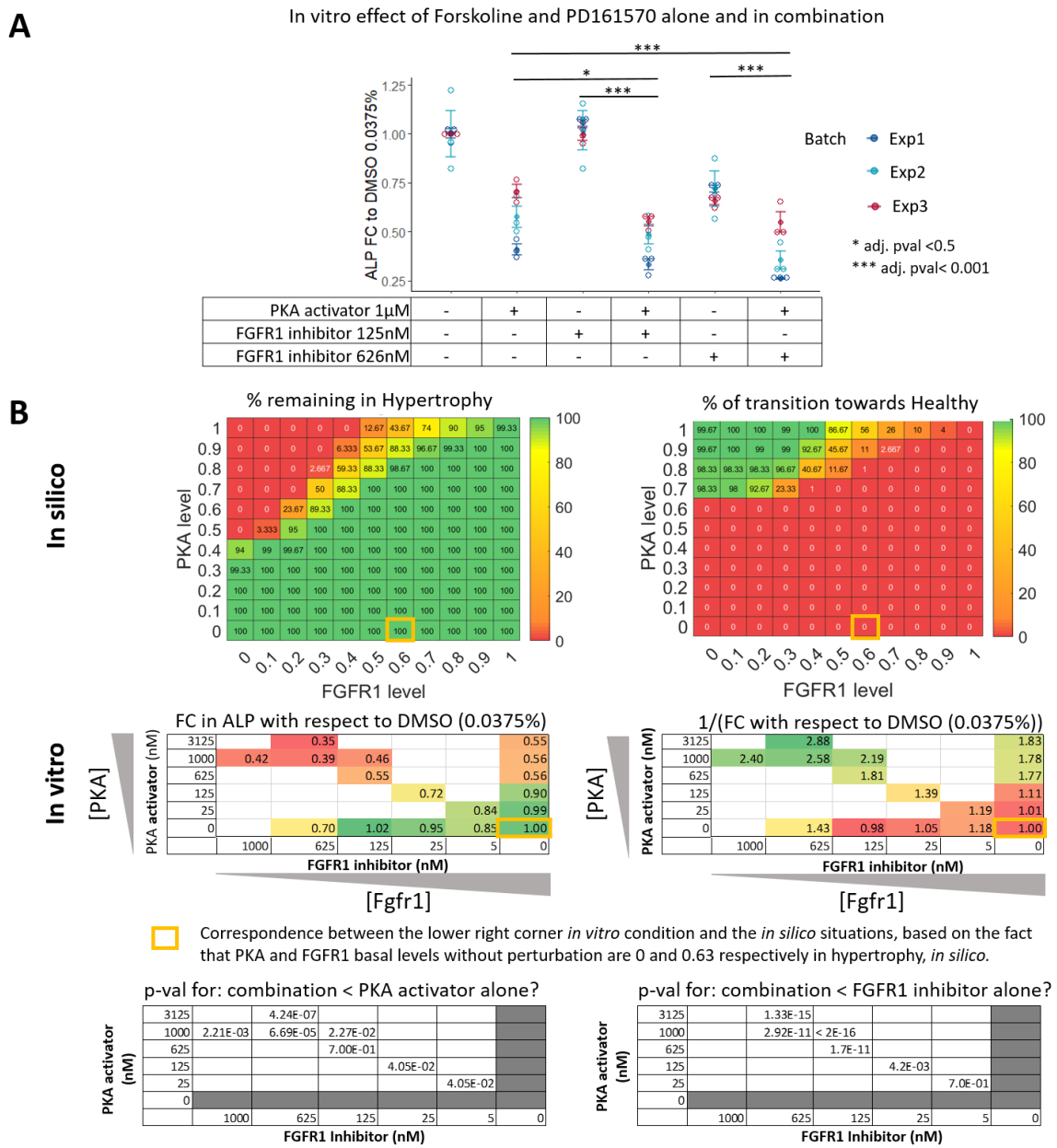


Figure 4.6: *In silico* vs. *in vitro* dose response effect of PKA activation with FGFR1 inhibition. The most potent condition from the screening is investigated further for a potential dose effect. (A) Fold change (FC) in ALP activity, with respect to control, due to PKA activator (Forskoline, 1µM) or FGFR1 inhibitor (PD161560, 125nM and 625 nM) or the combination of both. **(B)** A range of values for PKA and FGFR1 imposed activities is screened *in silico* with 0 meaning no activity and 1 being the max possible activity. The percentage of transition remaining in the hypertrophic state or transitioning towards the healthy state are reported in the upper panels, the rest of the transitions goes to the 'None' state. In the middle panels, fold change (resp. inverse of fold change) in DNA-normalized ALP activity with respect to control DMSO in ATDC5 is reported for a range of Forskolin and PD161570 concentrations. The *in vitro* situation without drugs (yellow rectangle) would correspond to the basal level of PKA and FGFR1 in *in silico* hypertrophy but there is no one to one correspondence between the *in silico* and *in vitro* ranges. All *in vitro* results represent n=9 (3 bio-replicates in 3 independent experiments), p-values are computed on log-transformed data with a linear mixed-effect model, user defined contrasts (only combination versus corresponding single doses were compared), one-sided test and adjustment for multiple comparisons with the Holm's method. The combinatorial treatment effects were greater than the ones for either of the single treatment both *in silico* and *in vitro*, for all concentrations in the gradient of dose relationships investigated.

4.3 Discussion

We report the construction of a mechanistic model of chondrocyte phenotype control in articular cartilage by combining a knowledge-based approach with a data-driven approach. We have leveraged decades of knowledge and data about chondrocyte regulatory pathways and osteoarthritis by integrating that information in a numerical predictive model. This model enables to recapitulate physiologically relevant observations and predict conditional effects resulting from intricate intracellular signaling. It offers the possibility to screen a large amount of (combinatorial) treatments and prioritize subsequent *in vitro* experiments for the identification of molecular drivers and drug targets in OA. The *in silico* targets perturbation screening and the experimental validation of our findings show the potential of *in silico* experiments to identify relevant *in vitro* experiments and refine the early drug discovery pipeline for OA. In particular, our study points towards a likely synergistic effect of PKA and FGFR1 targeting strategies to regulate chondrocyte hypotrophy. Several of our insights have implications both for the network modeling community and for cell and cartilage biology.

We have first built an interactive intracellular network as an online knowledge base and as a reference support for our numerical model. Most network-based models rely on prior mechanistic knowledge. Even current state of the art computational tools meant to reconstruct numerical models automatically from (high-throughput) data often require or offer the possibility to introduce prior knowledge (*222*). Therefore, there is high need for integrating and curating originally isolated pieces of knowledge, in a comprehensive way. However, biochemical information specifically related to cartilage and osteochondral systems are scarce in public and private pathway databases, in which cancer related cell types tend to be over-represented. The network we provide along with this study, details the prior mechanistic knowledge we have put in the model and that was predominantly chondrocyte or osteochondral cell type specific, with a focus on articular chondrocytes and osteoarthritis. In our opinion, it is valuable not only for cartilage and OA researchers but also for modelers as it can serve as a basis to derive other models to answer alternative questions.

Combining knowledge and data in a comprehensive network resulted in a systemic view of chondrocyte intracellular regulation. Many independent pieces of information have accumulated over the past decades and many databases have made curated pathways available. Most of the time in literature, these pathway descriptions stop after the (in)activation of the downstream transcription factors while the identity of the target genes downstream these

pathways is left obscure. In this study, our strategy was to complement the knowledge based network graph with automatically inferred transcriptional regulations from transcriptomic data through the use of machine learning methods. This usually requires a large multi-perturbed dataset but in absence of such dataset for chondrocytes, we reconstructed one by merging various arrays. Our strategy is supported by a previous study, which reported that equivalent informative data could be successfully achieved by assembling naturally occurring and experimentally generated phenotypic variations of a given cell type (*272*). Even though relatively few inferred gene regulatory interactions were integrated in the mechanistic model due to the stringency of our selection strategy, we limited the risk of integrating false positive predictions. This data-driven approach allowed us to take advantage of automatic network reconstruction technologies, which are becoming more and more the standard in the state of the art (*339*, *340*).

The translation of that network into mathematical rules was then further studied computationally and enabled the prediction of the overall effect of each network component on the virtual chondrocyte. Our perturbation screening approach helped to discover influencer nodes based on the mathematical dynamics, similarly to what had been previously proposed for other diseases (*326*). Important to mention is that no information was put in the model that directly made one target prevail over the others in being pro- or anti-hypertrophic. The information that fed the mathematical model was about activating or inhibitory influences of one molecule on another molecule, a complex of molecules or a pathway although some edges in the network represented indirect links for the sake of simplification. The advantage of the discrete semi-quantitative mathematical formalisms we employed lies in its ability to reproduce qualitative dynamics using only the activating or inhibitory nature of interactions and additive rules, without any prerequisite about kinetics. This characteristic makes that method perfectly suitable for a large network such as the one of this study. Many biostatistics and machine learning methods that make drug efficacy related predictions solely based on omics data still lack predictability and interpretability (*341*). In contrast, the numerical approach presented in this study has the advantage of providing mechanistic evidence supporting the predicted effects, thereby increasing the mechanistic interpretability.

Generally, the validation of an *in silico* model to answer questions in a specific context may be achieved in two ways: either by showing its ability to predict non-linear effects that were already reported in literature or by confirming predictions with new experiments. The relevance of our model to answer osteoarthritis-related questions was first confirmed by its ability to

recapitulate earlier described behavior such as the changing role of TGFβ signaling in presence of inflammatory stimuli. A recent computational study reporting a quantitative (ODE-based) time dependent model of cartilage breakdown (*332*) provided more detailed and complexed kinetics underlying the dual role of TGF-β. The fact that our model could lead to equivalent conclusions shows that the more simplifying semi-quantitative formalism that we employed could suffice to capture non-straightforward effects, thereby supporting its credibility for subsequent predictions.

The Monte Carlo analysis of the system without constraints showed that the healthy state was more easily reached than the hypertrophic one. This corroborates with the situation for normal articular cartilage in which chondrocyte hypertrophy does not spontaneously occur unless the homeostasis is disturbed (*342*). A great part of the state space was occupied by the 'None' attractor that is most likely a trivial solution towards which the system converges when an initial state or a trajectory is too far away from a feasible biological state to meet all the constraints imposed by the equations. In such a situation the trivial 'zero' solution is more easily reached. An analysis of the ensemble of initial states reaching the 'None' state, could possibly spotlight initial states that are unlikely to happen in an *in vivo* physiological environment. Conversely, the analysis of canalization under imposed growth factor profiles have suggested that restricting ourselves to the initialization of growth factor's values within biologically 'feasible' ranges can decrease prevalence of the 'None' state, although the boundaries of those feasible ranges could be further explored. Nevertheless, in the scope of finding potential therapeutic targets, we are less interested in the random canalization of the states than in their robustness to perturbations and capacity of transitioning.

An important result of this study is the highlights that chondrocyte phenotypes are not much sensitive to small, single factor perturbations. Even though, the resistance of the current *in silico* model to single factor perturbation might, in part, be due the omission of some parts of the real world system's complexity, this trait of robustness to small environmental variations is often considered as a fundamental and ubiquitous trait of biological systems (*10*). This traits allows systems to function in noisy environment (*343*). Systems biologists have theorized that disease may establish its own robustness, in some cases (*344*). In line with that, the Markov chain in our computational model indicates that, for articular chondrocytes, the probability to change phenotypes once it has been reached is rather low. Indeed, due to the very intricate interplay of molecules, it is likely that some pathways play redundant roles and that several factors should be targeted simultaneously to 'unlock' the system. Moreover, our *in silico* pairwise perturbation

screening confirms that targeting at least two proteins at the same time increase the chance to unlock the hypertrophic commitment. Overall, as its *in vivo* counterpart, the *in silico* articular chondrocyte is unlikely to display hypertrophic signs in normal conditions or even sometimes under inducing treatment (*320*), but once the hypertrophic transition has been initiated, it is rather difficult to escape that fate.

Another important outcome is the experimental validation of the model predictions on previously unreported conditions. In practice, as reverting a hypertrophic chondrocyte back to a healthy state has never been observed experimentally, we hypothesize that the *in silico* predicted conditions are, at least, more likely to block hypertrophy. The *in vitro* results that we present here support this hypothesis and open new routes for further testing the suggested combinatorial conditions. Especially, the combination of PKA activation with FGFR1 inhibition is highlighted as good candidate treatment by our integrated *in silico-in vitro* approach. Activation of the PTHrP pathway, to which PKA belongs, has already been reported to be rather anti-hypertrophic for growth plate chondrocytes (*97*). Similarly, genetic inhibition of the Fgfr1 gene in mouse knee cartilage has been shown to attenuate the degeneration of articular cartilage in mice (*124*). However, to our knowledge, this combination has never before been investigated nor reported for its synergistic potential against hypertrophy, cartilage degradation or OA.

Validating *in silico* predictions for drug target discovery experimentally is a challenging task. This is especially true when numerical high-throughput screenings, enabled by high computing power, generate a large amount of predictions. We leveraged the evaluation of secreted ALP activity in ATDC5 cell line as a semi-high throughput read-out for validating the *in silico* predictions. Indeed, it served as a convenient experimental system to assess hypertrophy modulation under many screened perturbations (*296*). Even if that cell line is of mouse origin and cultured in isolation from a physiological environment, we propose to use it within an *in silico-in vitro* pipeline that would act like a funnel for prioritization of candidate drug targets. As shown in our study the main hits can then be further evaluated in more detail using human OA chondrocytes.

We acknowledge that our study has some limitations. A first drawback is the absence of a gold standard to assess the precision of the data-driven network inference. To mitigate this, we used a consensus approach by only integrating predictions made by three different machine learning methods. This reduced the amount of interactions we would integrate but it alleviated each

method's weakness and reinforced the strengths of the predictions, as previously proved (*221*). Secondly, the type of mathematical model we employed comprises almost no numerical parameters but the main one, the saturation constant, was assigned an arbitrary value based on a previous study (*255*). This constant determines how fast a protein activity or gene expression can saturate to the maximal value depending on the amount of excess positive and negative upstream interactions. It intervenes in the weight of interactions and changing its value might slightly affect the influence of the network's constituents on the system. Finally, we did not experimentally verify the target specificity and dosing regimen for the small molecules employed for our *in vitro* validation since it was out of the scope of this study. However, those small molecules were selected based on their previously reported and well-known *in vitro* action on our targets of interest.

This study is a proof-of-concept that an *in silico* - *in vitro* integrated approach can suggest single and combinatorial target perturbations affecting the hypertrophic transition and help to prioritize therapeutic targets for OA therapy development. In that sense, our model offers the possibility to draw conclusions on the pro- or anti-hypertrophic nature of biochemical pathways and targets based on strict mathematical rules describing the intricate network connectivity. We are convinced that this type of approach can guide the process of therapy development from basic understanding to target selection early on in the drug discovery pipeline while reducing time and cost of experiments as well as the use of animal models in early stages of drug discovery. Furthermore, investigating the effect of a new target that was not present in the current model should be possible by solely informing the model on how the target interacts with and connects to the rest of the network. Typical information on the nature of the upstream activators and inhibitors of the protein's functionality, the nature of downstream proteins modulated by the target under scrutiny as well as information about its transcriptional regulators, from DNA binding assays for instance or reverse engineered from data would be needed. Ideally, this information should be as exhaustive as possible based on current state of the art knowledge, while hypothetical connections could be investigated and compared based on the simulated target's effect. Finally, as scientific research is making progress in the identification OA disease subgroups based on molecular markers and clinical phenotypes (*42, 345*), we foresee data-informed mechanistic models can become more and more patient-type specific. For instance, such knowledge-based network model could serve as a prior and be further optimized with engineering approaches, adjusting the network topology and/or interaction weights, in order to fit chondrocyte baseline profiles of typical patient subgroups.

Then, the resulting network models and the effect of targets' perturbations could be compared across the different patient type-specific models.

4.4 Materials and Methods

Network construction

The knowledge-derived networks were built by incorporating fine-grained mechanistic knowledge about signaling pathways and transcriptional regulations. A previously published model of chondrocyte differentiation in the growth plate (*254*) was used as a basis and was adapted and completed through literature curation of decades of knowledge about articular chondrocyte and osteoarthritis,. Reference sources for the experimental evidence (binding assays, clinical observations, etc..) were mostly from mouse and human origin. References used chondrocytes (or related cell lines) and predominantly involved direct protein or promoter binding information, curated pathways or observed phenomenon in cartilage during homeostasis and disease. All references and mechanism descriptions are available through the interactive networks in the online platform Cell Collective (http://www.cellcollective.org) (*269*). See references (*328, 329*) for the chondrocyte knowledge bases of the protein-signaling network and the gene regulatory network respectively. Additionally, some gene regulatory interactions were automatically identified with machine learning algorithms using gene expression multi-perturbed data as input. Such informative dataset was achieved by merging six published microarray datasets of mouse articular cartilage (GSE26475, GSE33754, GSE79239, GSE33656, GSE53857, GSE45793). Each dataset had a control or wild type group and an OA-like group, which was either using genetically modified mouse spontaneously developing OA or a DMM-induced OA mouse model. Importantly, the initial annotations provided by the authors were not necessarily 'OA' or 'WT' but rather experiment-specific annotations such as 'DMM' or 'SHAM operated'. Therefore the binary annotation as OA-like or WT-like samples was established by hand for the purpose of this study according to the original data annotations published on GEO, as reported **Table S3.1**. The datasets were merged applying an in-house developed pipeline adapted from previously published methods (*279*)**.** Briefly, the datasets were preprocessed first in a platform-specific way prior to assembly as a merged dataset. For the purpose of this study, a sub-dataset was created by restricting ourselves to the genetic profile of 74 genes of interest, being the ones present or closely related to the biological factors from the mechanistic model. Forty one genes, out of the 5470 from the full dataset, matched our list of interest, based on the Ensembl IDs, and the associated genetic

profiles constituted the final sub-dataset (41 genes (or variables) x 109 samples (or observations)). This dataset was quantile normalized and the batch effects were removed through the ComBat algorithm based on Bayes methods (*277*).

Inference was performed for regulatory interactions within the aforementioned subset of genes. This was achieved employing three different algorithms, being ARACNE, TIGRESS and GENIE3 (*346–348*); the final retained network was a consensus network of the three algorithms. Typically, an interaction between a transcription factor and a gene was kept if it was present in the three methods' results. An interaction was called present for a method if the interaction score was higher than a certain threshold for that method. This threshold was set to $m - \sigma$, where m is the average score for the given method and σ is the standard deviation of the scores.

Unless specified, all the aforementioned microarray data analytics were accomplished using the R computational environment (v.3.2.2).Topological parameter analysis of the final network was carried out with the Network Analyzer plug-in of the Cytoscape software v3.7.2 (https://cytoscape.org/) (*268*).

Implementation of the mathematical model and dynamic analysis

The information contained in the network was translated into mathematical equations through an additive formalism with 2 priority classes to distinguish between fast and slow reactions, the importance of which was repeatedly highlighted (*214*, *349*, *350*). This additive formalism resembles the Boolean threshold networks (*351*) and was previously implemented with priority classes as described by Kerkhofs and Geris (2015). In this formalism, the nodes or biological component were represented by variables for which the values evolve over pseudo-time steps. The model is semi-quantitative since the nodes or variables could take on a continuous activity values between 0 and 1. The evolution of variables (proteins or genes) was defined by the sum of the upstream activating variables and the subtraction of the upstream inhibitory variables from the network (see definitions in **Table S4.1**). Biological interactions could happen at two time scales, reflecting the priority classes: reactions related to slow biological processes such as gene expression, mRNA or protein production, were referred to as slow reactions (lower priority) and those related to fast processes such as protein activation or inhibition, were referred to as fast reactions (higher priority) (see **Table S4.1** for the definition of fast and slow

reactions & variable). A formal description of the mathematical system underlying the model as well as the full list of equations are available in **Supplementary text S4.1**.

The asymptotic solutions were evaluated with a Monte Carlo simulation procedure, similar to methods employed for logical models (*213*, *214*). When running a simulation (also see section below on Monte Carlo analysis), an initial value in the interval [0,1] was assigned to each variable. Every simulation step, the sub-variables (see **Table S4.1**) were updated asynchronously according to the rules given in the equations and following the priority classes, in such a way that fast reactions were always updated before the slow reactions. The order in which variables were updated within a priority class was random, hereby recapitulating the stochasticity inherent to any biological system. See **Figure S4.4 and Figure S4.5** for graphical explanations about the algorithm and simulation scheme on a reduced illustrative example network. A stable state (definition **Table S4.1**) was reached whenever the next iteration step did not bring any change for any of the variables up to a tolerance 10^{-2}. In other words, when initializing the system at a random point, it was considered converged when the relations detailed by the system of equations were fulfilled up to a tolerance of 10^{-2}. Thanks to the stochasticity of the model, the same initial point could lead to different types of stable states. Therefore, all computational results of this paper were computed 3 times and standard deviations were evaluated. All implementations and simulations were carried with the MathWorks® suite, MATLAB (2018b).

Monte Carlo analysis and estimation of attractors (or stable states)

The nature and size of the stable states, given the regulatory network provided in the equations, was estimated by a Monte Carlo procedure. In short, all variables were initialized 10.000 times with random values in the interval [0,1]. The amount of initializations reaching each final stable state were computed and reported in terms of percentage of initial states. This number gives a sense of the stable states' size for the unperturbed system, i.e. without constraints (*213*, *214*). Then, we performed two other Monte Carlo canalizations in which all variables were randomly initialized, except for seven growth factors that were fixed at values being physiologically relevant in either a normal healthy or a disease environment (profile A and B of **Figure3B**, respectively). The networkD3 package from R was used to produce the Sankey diagram for the visualization of the canalization results.

In silico target perturbations

By essence, the attractors are stable, meaning that variables cannot evolve anymore. These states may however be escaped by forcing, computationally, the value of one or several variables to change. Such a perturbation was imposed for a fixed amount of computational iterations, after which the system was left to evolve freely, thereby accounting for the fact that chemical treatments affect biological systems for a finite period of time. The duration of the perturbation was set to 1000 time steps as the perturbed state did not take more than 200 time-steps to be reached, on average. Imposing a perturbation on a stable state forces the system to evolve again, following rules imposed by the equations, and eventually settle down in the same initial or a new attractor (see convergence description above).

The different *in silico* scenarios or treatment experiments that were tested amounted to perturbing one or several variables/nodes, from the healthy or the diseased hypertrophic state and assessing the effect of that perturbation on the state stability. Variables were perturbed by forcing their global activity value to be 0 or 1 for inhibition or activation respectively. Imposing intermediary values between 0 and 1 was also done for some specific questions in which extreme values would be unlikely, such as varying the ratio between different membrane receptors. Each perturbed conditions was imposed starting from the relevant initial stable states (healthy or diseased) and the nature of the final state to which the perturbation led after simulation was documented. We considered that the tested perturbation triggered a state transition when the final state was different from the initial one. Given the stochastic nature of the model, the same perturbation could trigger a different outcome if simulated a second time, therefore the same perturbation was repeated 100 times and we reported the percentage of transition towards each of the possible target states. Standard deviation in the percentage of transition was assessed by repeating that experiment 3 times.

Due to the computational cost associated with the systematic screening of all possible pairwise perturbations, (for each pair there are 4 possible pairwise conditions to be tested either from the healthy state or from the hypertrophic state (**Table S4**)). the independent simulations were run in parallel using high performance computing infrastructure of the KU Leuven VSC. Afterwards, we automatically selected combinatorial conditions for which at least 70% of the perturbations triggered transitions from the hypertrophic state towards the healthy one. Among them, we focused on those conditions with more than 70% transition towards the healthy state but less than 5% towards the 'None' state as a first discriminatory factor. Conditions were

further selected for their druggability and their ease to be tested in a simple *in vitro* system. For instance, conditions involving the modulation of transcription factors were not considered for the *in vitro* validation since, no small molecule treatment could directly affect transcription factor activity.

Testing treatments in vitro with ATDC5 culture

The validation of our *in silico* predicted treatments required *in vitro* testing with small molecules. This was performed with ATDC5, a mouse chondroprogenitor cell line obtained from the Riken Biological Ressource Center. Cells were cultured in 2D in proliferation medium containing DMEM/F12 (ThermoFisher, UK), 5% Fetal Bovine Serum (Biowest, Belgium) and 1% antibiotic/antimycotic (Gibco, ThermoFisher Scientific). Chondrogenic differentiation was induced by plating the cells at 6,400 cell/cm² in proliferation medium for 24h, followed by changing the medium to differentiation medium, being proliferation medium supplemented with 10 µg/ml insulin (Sigma-Aldrich), 10 µg/ml transferrin (Sigma-Aldrich) and 30 nM sodium selenite (Sigma-Aldrich). Cells were incubated in a humid environment at 37°C, 5% CO_2 and differentiation medium was refreshed every other day for the first 10 days, and every day after the 10th day, for longer experiments (i.e. Figure 5B). Supernatant medium was taken for ALP activity assay and cells were harvested for DNA quantification (0.05% Triton-X reagent). ALP activity was reported relatively to the total DNA quantity to alleviate potential variation in cell number.

To study the correlation between *Col10a1* expression and secreted ALP activity, cells were differentiated for 14 days with or without Ihh supplement (150ng/ml, R&D Systems Europe LTD). The cells were harvested for RNA isolation on day 0, 7, 9, 12 and 14 during differentiation (TRIzol reagent; Thermo Fisher Scientific), in addition to the ALP acitivity assay and DNA quantification.

To assess the effect of small molecules and growth factors treatments on hypertrophic differentiation, cells were treated on day 8 of ATDC5 chondrogenic differentiation for readout at day 9. Cells were treated with one or a combination of the following compounds: Forskolin (1µM, Axon Medchem), Recombinant Human/Mouse/Rat Activin A Protein (100ng/ml, R&D Systems), Recombinant Mouse IGF-I/IGF-I Protein (10ng/ml, R&D Systems), Transforming Growth Factor (TGF)β1 (10ng/ml, PreproTech), PD0325901 (1µM, Axon Medchem), PD161570 (1µM, Axon Medchem), ITSA1 (50µM, Chembridge), LDN-193189 (0.5µM, Axon

Medchem), LY294002 (20µM, Axon Medchem) and IWP2 (2µM, Stem cell technology). ALP activity in treated conditions is expressed in terms of fold change with respect to the control medium with appropriate amount of DMSO, which was used as a solvent for most small molecules. Four types of control media were used throughout this study due to sparse solubilities of the compounds. The control medium was with 0.02% DMSO (Medium1) for most treatments, with 0.1% DMSO (Medium2) for the ITSA1 related conditions, without DMSO) for the ActivinA treatment (**Figure 5),** and with 0.0375% DMSO for the Forskoline/PD161570 synergy study and dose screening (**Figure 6**).

Validating treatments effects in primary human chondrocytes culture in alginate beads

Human articular cartilage was obtained with implicit consent as waste material from patients undergoing total knee replacement surgery. This protocol was approved by the medical ethical committee of the Erasmus MC, University Medical Center, Rotterdam, protocol number MEC-2004-322. To isolate chondrocytes, cartilage chips were subjected to protease (2 mg/ml, Sigma Aldrich) for 2 hours followed by overnight digestion with 1.5 mg/ml collagenase B (Roche Diagnostics, Switzerland) in Dulbecco's Modified Eagle's Medium (DMEM) high glucose supplemented with 10% fetal bovine serum. Single cell suspension was obtained by filtrating the cellular solution by a 100 µm filter. The isolated chondrocytes were expanded in monolayer at a seeding density of 7,500 cells/cm2 in DMEM high glucose supplemented with 10% fetal bovine serum, 50 µg/ml gentamicin and 1.5 µg/ml fungizone (Gibco, Grand Island, NY, USA). Upon approximately 80% confluency cells were trypsinised and reseeded at 7,500 cells/cm2. Cells were used for experiments after three passages. Redifferentiation of articular chondrocytes was performed using a 3D alginate bead culture model. For preparation of alginate beads, chondrocytes after three passages in monolayer were re-suspended in 1.2% (w/v) low viscosity alginate (Kelton LV alginate, Kelko Co, San Diego, CA, USA) in 0.9% NaCl (Sigma Aldrich) at a concentration of 4 × 106 cells/mL. Beads were made by dripping the cell-alginate suspension in 105 mM CaCl2 (Sigma Aldrich) through a 22-gauge needle. Beads were washed with 0.9% NaCl and DMEM low glucose. Beads with a size that deviated from the average after a visual inspection were not included in the experiment. Re-differentiation of chondrocytes was performed in a 24-well plate (BD Falcon) for two weeks in 100 µL/bead DMEM low glucose supplemented with 1% ITS fetal (Biosciences), 10 ng/ml transforming growth factor beta one (TGFβ1, recombinant human, R&D systems) 25 µg/mL l-ascorbic acid 2-phosphate (Sigma Aldrich), 50 µg/ml gentamicin, and 1.5 µg/mL fungizone (both Gibco). After two weeks, TGFβ1 was no longer added to the medium and cells were cultured with and

without 1µM of Forkosolin for 24h. Each experiment was performed with cells derived from 4 OA donors (3 Females, 1 Male, 65 ± 6 years), in triplicates.

RNA isolation and RT-qPCR for ATDC5 samples

Gene expression of *Col10a1* in ATDC5 expriments was evaluated by RT-qPCR. For RNA isolation, chloroform was added to the TRIzol samples (TRIzol 5: Chlororform 1), which were subsequently centrifuged for 15min at 15000 rpm (i.e. RCF = 218849) and 4°C. RNA was isolated by collecting the aqueous phase and precipitated with isopropanol (aqueous phase 1: ispropanol 1) for 30min at -80°C. After centrifugation at 15000rpm (i.e. RCF = 218849) and 4°C for 30 min, supernatant was removed and the resulting pellet was washed with 80% Ethanol. RNA pellets were dried for 10min in desiccator and dissolved in 15µl RNase free water. Finally, RNA content and purity was determined with Nanodrop. RNA was converted to cDNA with the Revert Aid H Minus First strand cDNA synthesis kit (Thermo Scientific) according to the manufacturer's protocols. Quantification of gene expression was done using Syber Select Master Mix (Applied Biosystems) adding 400nM forward and reverse oligonucleotides primers as listed in **supplementary text**. The StepOne Plus System (Applied Biosystems) was used for amplification using the following protocol: denaturation cycle at 95°C for 10min followed by 40 cycles of amplification (15 seconds 95°C and 1 min 60°C), followed by a melting curve. Expression levels were analyzed using the $2^{-\Delta Ct}$ method and normalized for the expression of the reference gene *Hprt*. This housekeeping (HK) gene was determined after verification of multiple HK genes and selecting the one that remained most constant throughout the procedure.

The specific RNA isolation protocol, RT-qPCR method and oligonucleotides used for measuring gene expression in primary human chondrocytes is reported in **supplementary text**.

ALP assay

Enzymatic activity of secreted ALP in the supernatant medium of ATDC5 cultures was determined in a colorimetric assay as previously described (*66*). Briefly, ALP activity was determined in flat-bottom 96-well plates (Sigma Aldrich, CAT M9410) containing assay buffer (1.5 M Tris-HCl, pH 9.0, 1 mM MgCl2; 7.5 mM p-nitrophenyl phosphate). The ALP activity was assessed as a function of formed nitrophenyl phosphate (pNp), the reaction colored product, which was measured by spectrophotometry at 405nM after 30min of reaction. The reaction was stopped with 1M NaOH Stop Buffer. A calibration curve containing an increasing concentration

of pNp served to determine the absolute amount of ALP-generated pNp. Sample values were normalized to total DNA amount, expressed as µmol of pNp /mg of DNA and reported as fold change with respect to the relevant control medium.

DNA quantification

After harvesting the cells with 350µl of 0.05% TritonX-100, samples were vortexed and frozen at -80°C for further processing. Samples were sonicated for 5 seconds, centrifuged for 1' at 13000rpm (i.e. RCF= 164380) and 200µl of the supernatant was harvested. Samples were diluted with a factor 1/10 with distilled water then DNA content was measured with the Qubit 3.0 fluorometer (Life Technologies). Qubit dsDNA HS (high sensitivity, 0.2 to 500ng) Assay Kit was used according to the manufacturer's protocols; a sample volume of 5µl was added to 195µl of a Qubit working solution.

Statistical analysis

In general, one-tailed statistical tests were used to analyse *in vitro* results and relate them to expected outcomes from the *in silico* model's predictions since the model predicts a directionality of the outcomes. Average effect of Forskoline and Activin treatments were compared to control in one representative experiment with 3 replicates thanks to a one-tailed unpaired t-test with welch's correction, in **Fig5.C**. Effect of Forskoline treatment in OA chondrocytes from human donors was performed in triplicates for 4 donors. Comparision of the average effect with the control is done with a linear-mixed effect model to account for donor variability. Graphical visualization and statistical analyses for the semi-high throughput small molecule screening (**Figure5D)** were performed by modifying the BraDiPluS package from the Saez Lab (https://saezlab.github.io/BraDiPluS/). More precisely, the probability that a treated condition resulted in lower ALP activity z-score than the corresponding control in the ATDC5 screening was estimated with a one-tailed Wilcoxon rank-sum test (the *in vitro* screening dataset was not normally distributed based on Shapiro-Wilk test), with Benjamini–Hochberg correction for multiple testing. Three independent experiments were performed for each conditions, and, treatments effects were assessed in triplicates, in each experiment (or run). Probabilities from the independent runs were combined with Fisher's method using the combine.test function from the survcomp R package. A treated conditions was considered lower than control for p-value < 0.05 and z-score <0.

Chapter 5. Regulatory network modeling applied to iPSCs derived chondrocyte differentiation for bone TE

This chapter is partially based on:

Human pluripotent stem cell-derived cartilaginous organoids promote scaffold-free healing of critical size long bone defects.

WL. Tam, L. Freitas Mendes, X. Chen, ***R. Lesage****, I. Van Hoven, E. Leysen, G. Kerckhofs, K. Bosmans, YC. Chai, A. Yamashita, N. Tsumaki, L. Geris, SJ. Roberts, FP. Luyten.*

Stem Cell Res Ther (2021). (*352*)

5.1 Introduction

Developmental TE strategies have emerged as potential solutions to produce bone in the lab, in case of large non-healing fractures. The advantage of such strategies lies in their intent to recapitulate the natural developmental process of endochondral ossification, which is a well-studied and understood process. Therefore, rigorous comparison between the lab-based bone formation and the normal developmental path may provide good metrics for quality assessment with the aim to foster clinical translation (*4*). Nevertheless, those strategies still require a large amount of cells as elementary building blocks. Recently, induced pluripotent stem cells (iPSCs) have been successfully engineered *in vitro* towards the formation of cell aggregates displaying bone and cartilage forming capacities *in vivo* (*47*, *48*). However, those studies used murine iPSCs and their regenerative potential has not been tested in a long-bone critical size defect. Beside the cell source, another pillar of the triple helix in tissue engineering (cells – materials – growth factors) is the use of the appropriate biochemical cues to direct cell fate decision and foster mature tissue formation. On this matter, there is still a need to find the most appropriate bone forming biochemical cues for this type of human iPSCs-based cartilaginous templates.

In an in-house parallel study, our collaborators engineered human pluripotent stem cell (hPSCs) derived callus organoids with the aim to heal critical size bone defects. In particular, the iPSC aggregates were differentiated into cartilaginous tissue intermediates and several treatments were applied to compare the treatments' potential for promoting hypertrophic differentiation *in vitro* and subsequent bone formation after implantation *in vivo*. The underlying question being: what is the best differentiation biochemical cocktail to promote the most potent hypertrophic differentiation, cartilage mineralization and subsequent bone formation? Such question may require a lot of resources as well as trials and assays to be answered. An advantage of computer modeling and simulations (CM&S) is that it allows to reproduce experimental set-ups *in silico* and simulate a large amount of experimental conditions cheaply and timely. Therefore, we aimed at using a numerical model of the mechanisms regulating chondrocyte differentiation to reproduce the cell differentiation experimental setup in a computer. The goal was then to predict and compare the effect of the treatments while gaining insights into the underlying potential mechanisms of action.

To achieve this objective, an ensemble model of chondrocyte differentiation during endochondral ossification was leveraged. Indeed, we followed a previously established strategy to optimize a chondrocyte network in order to let us reproduce the baseline situation that would

serve as a starting point for subsequent new *in silico* experiments. In this chapter, we first explain the model conceptual framework and the optimization strategy as devised in (*265*). We then reproduce the baseline situation results for that model. Subsequently, we investigate the effect of differentiation treatments such as the addition of a bone morphogenetic protein family member BMP4, a Wnt pathway agonist, the thyroid hormone Triiodothyronine (T3) and combinations thereof. Finally, we compare the *in silico* predictions with the experimental results for the aforementioned conditions from the parallel study, with the aim to rank and prioritize the differentiation treatments' capabilities.

5.2 Description of the mathematical framework and optimization method

The mathematical model employed in this application study consisted of a prior growth plate chondrocyte regulatory network model that was further optimized with a genetic algorithm by reproducing the steps devised in (*265*). More precisely, the mathematical optimization concerned the weights of all interactions in the network. Those parameters were optimized so that the model fitted reference expression profiles for chondrocytes. The underlying rationale being that all networks that produce the experimentally observed reference profiles may be considered as possible, allowing to deal with uncertainty in the network structure and tackle low confidence interactions. The reproduced ensemble model served as a basal situation for *in silico* experiments. The following subsections explain the model's conceptual framework and the computational methodology in more details.

5.2.1 Additive model with priority classes of a differentiation network in developmental chondrocyte maturation.

We started from a previously established regulatory network model describing chondrocyte maturation during endochondral ossification with well-documented mechanisms happening at the intracellular level (*253, 255*). It was that model that formed the basis of the mechanistic part of the model developed in this thesis (**Chapter 4**). We used this chondrocyte model because it is one of the only mathematical models in the current state of the art that describes so many regulatory mechanisms about chondrocyte differentiation during endochondral ossification while requiring so little kinetic information. As in **Chapter 4**, the mechanistic information included in that network model was relative to protein - protein interactions (signaling level) and transcription factor - gene interactions (gene level). We refer the reader to (*254*) for the

mechanistic content and connectome details. The backbone network was translated into a mathematical model using the additive formalism with priority classes in which biological interactions occur on two disparate time scales: the fast (signaling level) and the slow one (gene level). Therefore, each entity or node of the network could have both fast and slow upstream regulators. Thus, the global protein activity of a biological entity was the multiplication of its gene expression level by the protein activation/inhibition level. In what follows, whenever we refer to a node value, it stands for the global activity of that node/biological entity.

Determining the global activity of a node in a specific stable state (e.g. regular chondrocyte and hypertrophic chondrocyte environment) is done by solving a system of equations for both states through the application of relevant input profiles for external growth factors. Consider the adjacency matrix *A* of the network whose entries a_{ij} can be 0, 1 or -1. This matrix indicates the presence and the direction of edges in the network. If node *j* activates node *i*, then a_{ij} is 1, if node *j* inhibits node *i*, a_{ij} is -1. a_{ij} equals 0 when no interaction from node *j* to node *i* is present. The vector $L = l_{ij}$ contains a weight for each interaction in the regulatory network. z^f and z^s and $\boldsymbol{z}$ are $n \times 1$ vectors filled with the z_i^f, z_i^s *or* z_i elements respectively, which denote the variable of node *i* at fast (z_i^f), slow (z_i^s) and global level (z_i). Given *A* and *L*, the equations can be written as:

$$\begin{cases} z_1^f(t+1) = a_{11}^f l_{11}^f\, z_1(t) + a_{12}^f l_{12}^f\, z_2(t) \ldots + a_{1n}^f l_{1n}^f\, z_n(t) \\ \quad\quad\quad\

equations (**Eq. 5.1**). Importantly, this additive model was semi-quantitative since the nodes or variables could take on a continuous activity value between 0 and 1.

5.2.2 Ensemble model construction and evaluation

The system (**Eq. 5.1**) provided us with a backbone network but each interaction weight in the L vector had still to be determined. To avoid the bias introduced when using a particular configuration of interaction weights in an additive regulatory network model (as done in **Chapter 4**), we tackled the inverse problem (*265*) as shown in **Figure 5.1**. This means we started from specific constraints (chondrocyte expression profiles) that the model dynamics had to satisfy and we inferred corresponding interactions' weights (the L vector), which are considered as parameters. In the case of chondrocyte hypertrophic maturation, we imposed that the network dynamics had to be able to capture the expression profiles of both a non-hypertrophic and a hypertrophic chondrocyte. More precisely, any valid parameter set should allow the growth factor's activity input (**Table 5.1**) to give rise to the expected qualitative profiles output (**Figure 5.2**). These qualitative input-output data pairs were defined as in (*265*), in order to reproduce the basal situation (see **S5.1** and **S5.2**, for more details). As previously mentioned, given the abundance of parameters (one parameter per potential interaction in the network of 46 biological entities), the solution may be non-unique and multiple sets of parameters may match the constraints (i.e. the chondrocyte profiles) (**Figure 5.1** and **Figure 5.2**). This was addressed by considering an ensemble approach, an ensemble being a collection of models, each with a different parameter set, matching the prescribed constraints.

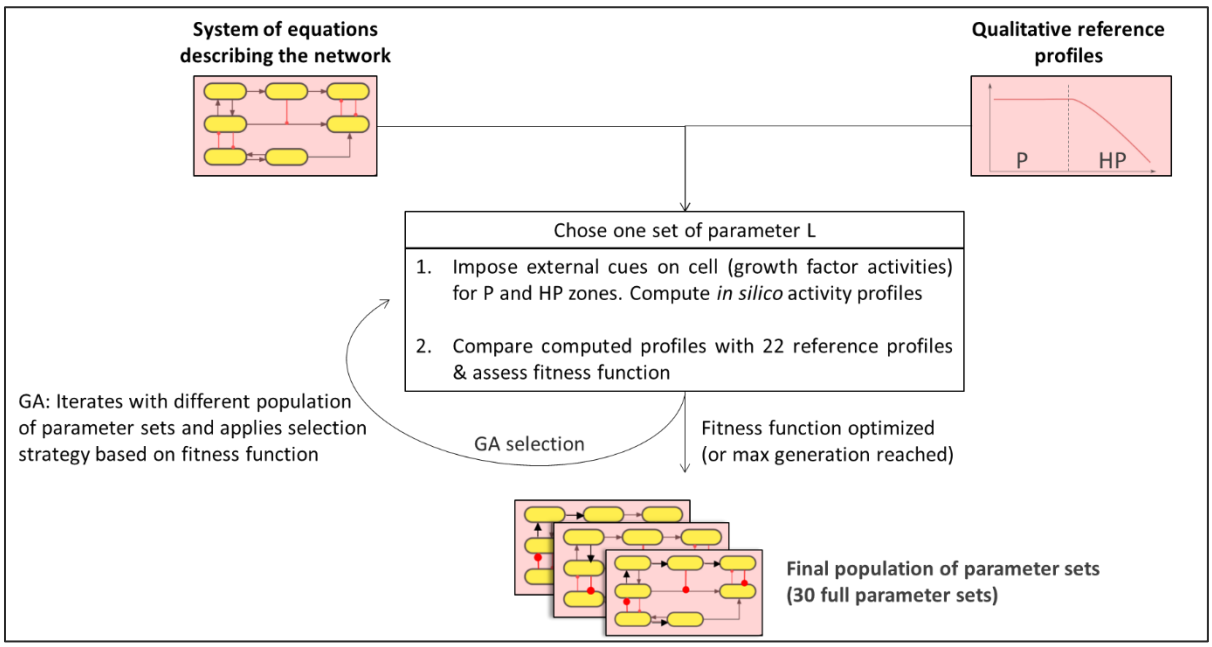


Figure 5.1: Workflow used to assemble the ensemble models. the genetic algorithm (GA) optimises a set of parameters to match associations between external cues input and resulting chondrocyte profile output. As the model was initially meant to match the profile of a chondrocyte during endochondral ossification, nHP stands for the non-hypertophic or proliferative chondrocyte and HP for the hypertrophic one. Next, the best solutions (30 parameter sets) are selected when the fitness function is optimized or when the algorithm reached a maximum number of generation. Finally, the remaining parameters are estimated based on growth factors activity in the growth plate (30 full parameter sets).

Zone:	Non-hypertrophic	Hypertrophic
Wnt	1/3	1
Ihh	1	1
PTHrP	1	0
FGF	1/3	1
TGFβ	1	1/3
BMP	1/3	1

Table 5.1: Growth factor activity profiles. These profiles are used as input for the non-hypertophic (proliferative or stable chondrocyte) and the hypertrophic state environment. The activity level is given, with 1 representing 100% activity

The reference profiles or observational data, for comparing the simulated output with the biological ground truth, were the chondrocytes' relative activity profiles of important markers in a non-hypertrophic regular chondrocyte versus a hypertrophic state. In summary, the activity profiles can be put into one of the five following qualitative categories (**Figure 5.2**):

a) **Low-high**: the node is more active during hypertrophy
b) **High-low**: the node is less active during hypertrophy
c) **None-high**: the node is exclusively active in the hypertrophic state
d) **High-none**: the node is exclusively active in the non-hypertrophic state
e) **Present**: the node is active in both states

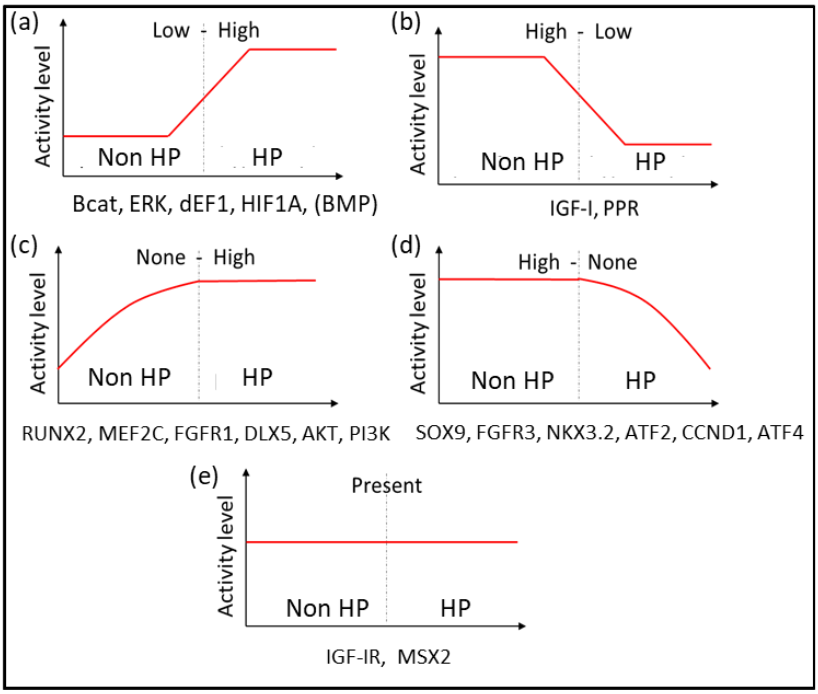


Figure 5.2: Graphical representation of the qualitative activity profiles occurring during hypertrophy. Entity or nodes are listed under the relevant graph. Each graph portrays a qualitative category: low–high, high-low, none-high, high-none and present. The profiles are derived from data of hypertrophic transition in the growth plate since the maturation observed in the current study should be similar to the one observed during embryonic endochondral ossification (Modified from (*265*)). Non-HP (resp. HP) stands for the non hyperpertrophic state (resp. hypertophic state).

To solve the inverse problem, i.e. find possible parameter sets, we employed a machine learning method called a genetic algorithm. This algorithm optimizes a fitness function based on mismatches with the pre-defined chondrocyte profiles shown in **Figure 5.2**, by performing selections on populations of parameter sets. The optimization process consisted of a series of iterations, each time evaluating a member of the population of parameter set by a fitness function and giving it a penalty score which had to be minimized. The better the fit between the computed and the reference profiles, the lower the penalty. The best parameter sets, i.e. the ones that fitted the chondrocyte expression profiles the best, were selected at each iteration (**Figure 5.3B**). This was done through a pre-defined selection strategy particular to the genetic algorithm (**Figure 5.1**). All simulations were performed using the built-in function for the genetic algorithm in MATLAB (The Mathworks): population size per generation = 100; selection based on rank, mutation and crossover; tolerance for convergence = 10^{-4}, max number of generation = 15. For efficiency, the population members (i.e. the parameter sets) were evaluated with parallel computing on the Flemish Supercomputer.

The fitness function for the genetic algorithm was derived from the qualitative activity profiles presented in **Figure 5.2**, with a penalty score system. The exact formulas to compute the penalty

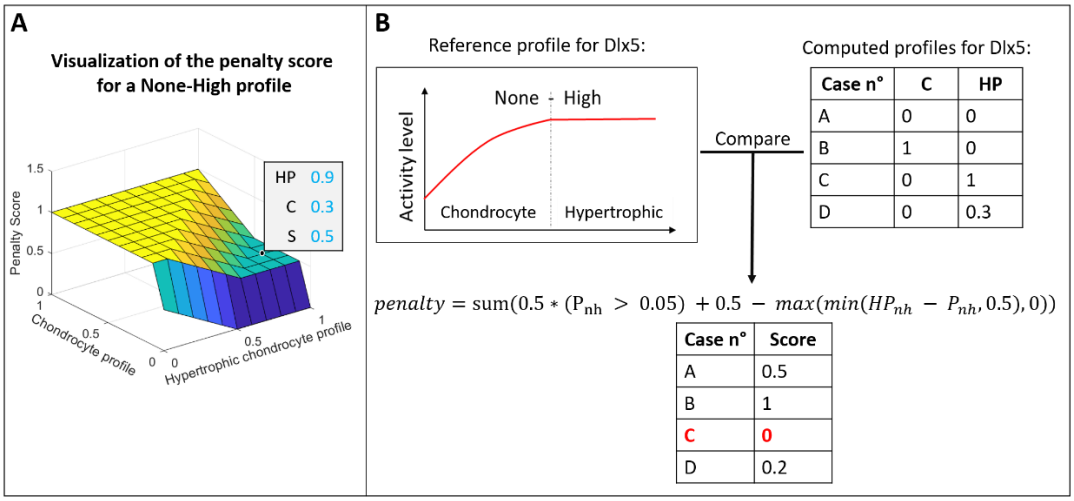
scores are reported in **Table 5.2**. The value of the final fitness function was the sum of all penalty scores. The meaning and rationale for that scoring system is explained in (*265*), summarized **Figure S5.3** and further illustrated in **Figure 5.3**. The score was solely based on the evaluation of the profiles of those nodes reported in **Figure 5.2**, the profiles of the growth factors (**Table 5.1**) were forced as input therefore the parameters for the interaction's weights upstream to the growth factors are really not optimized in that step and still need to be identified next.

Profile	penalty score
Present	$\text{sum}(1 - ((HP_p > 0.05) * 0.5 + (C_p > 0.05)*0.5))$
low-high	$\text{sum}(1 - 2*(\max(\min(HP_{lh} - C_{lh},0.5),0)))$
high-low	$\text{sum}(1 - 2*(\max(\min(C_{hl} - HP_{hl},0.5),0)))$
none-high	$\text{sum}(0.5*(C_{nh} > 0.05) + 0.5 - \max(\min(HP_{nh} - C_{nh},0.5),0))$
high-none	$\text{sum}(0.5*(HP_{hn} > 0.05) + 0.5 - \max(\min(C_{hn} - HP_{hn},0.5),0))$

Table 5.2: Formulas for computing the fitness function penalty scores. The fitness function is the sum of 5 penalty scores. HP and C contain the information about the reference profiles. More particularly, HP_i (respectively C_i) are vectors of values representing entity/nodes activity in the Hyertrophic (resp. non-hypertrophic Cartilage) state for an entity with a profile $\boldsymbol{i}$ ($\boldsymbol{i} \in [\boldsymbol{p}, \boldsymbol{lh}, \boldsymbol{hl}, \boldsymbol{nh}, \boldsymbol{hn}]$, where *p* means present, *lh* means low-high, etc..).

The case of an entity with an expected 'None-High' profile is used as an example to illustrate the scoring system (**Figure 5.3**). We consider a variable for which the expected profile should be down in a regular chondrocyte but highly expressed in a hypertrophic state, based on the reference profiles from literature. When comparing the computed profile with the reference one through the fitness function, the resulting penalty score may range from 0 (high fit between computed and expected profile) to 1 (low fit) (**Figure 5.3A**). The parameter set configuration that is retained is the one with the lowest penalty score (e.g. profile *C* in **Figure 5.3B**).

After that optimization task, the remaining parameters corresponding to regulators upstream of the growth factors were identified. It resulted in an ensemble of 30 parameter sets that made up the ensemble of models as a population of chondrocytes (**Figure 5.1**). More details about the optimization procedure and the specific options used for this analysis are provided in Supporting information **S5.4**.



Case n°	C	HP
A	0	0
B	1	0
C	0	1
D	0	0.3

Case n°	Score
A	0.5
B	1
C	**0**
D	0.2

Figure 5.3: Illustrative examples of the scoring system for an expected None-High profile. (A) The penalty score ranges from 0 (high fit between computed and expected profile) to 1 (low fit) when considering a variable for which the expected profile should be down in a regular chondrocyte but highly expressed in a hypertrophic state, based on the reference profiles from literature. The chondrocyte (resp. Hypertrophic chondrocyte) axis displays possible computed values for the variable in the regular chondrocyte (resp. Hypertrophic chondrocyte) state. The penalty score axis represent the score obtained when comparing the computed profiles with the expected None-High profile, based on the predefined fitness function. **(B)** Schematic representation of the comparison between 4 hypothetic computed profiles and the expected None-high profile for the variable Dlx5, with the resulting penalty score. Typically, the 4 profiles may be obtained with different configurations of the parameter set, the best fit is obtained with the configuration C, as highlighted in bold red in the table.

5.3 Establishing the numerical baseline situation: emerging chondrocyte phenotypes

In order to predict the effect of specific differentiation biochemical cocktails with this ensemble model, we first needed to establish a baseline situation. It consisted in computing the stable states of the model. This was done by computing the unperturbed canalization thanks to a Monte Carlo analysis, similarly to what was done in **Chapter 4** and shown schematically in **Figure 5.4**.

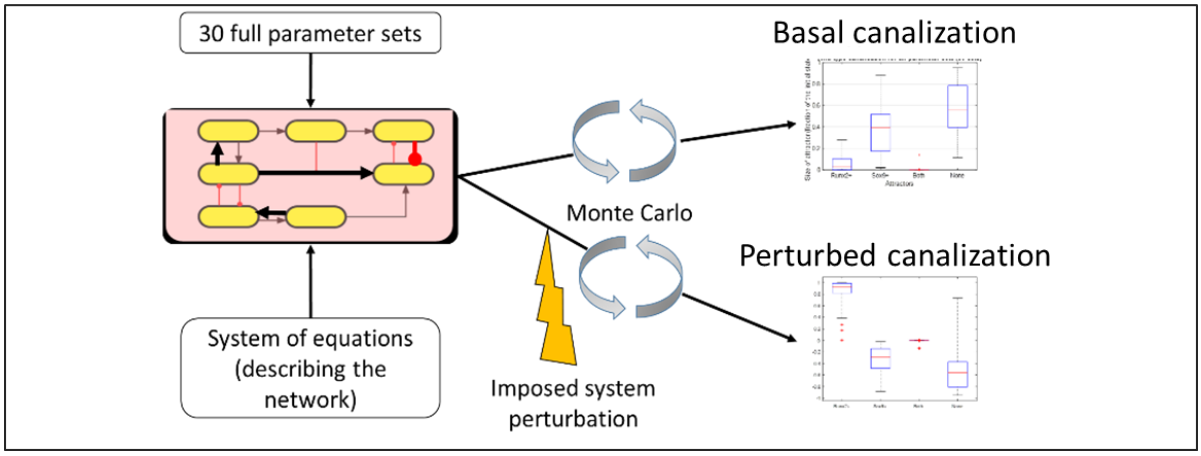


Figure 5.4: Workflow used to simulate the ensemble dynamics. The output of the optimisation is an ensemble of possible parameter sets that are used as input to the system of equations to simulate the system's dynamics. A Monte Carlo analysis is used to compute the canalization of the unconstrained and of the perturbed system. The pertubed Monte Carlo is the same as in the basal situation except that some nodes' values are imposed and fixed during the simulations, which may affect the nature and/or size of the stable states

In short, to assess the canalization of stable states emerging from the set of rules and parameters, the variables were initialised 10.000 times with random values in the interval [0,1]. The amount of initialisations reaching each stable state was computed for every model of the ensemble. This quantity gave a sense of the stability of the stable states for the unconstrained system (*353*). The same asynchronous updating strategy as in **Chapter 4** was used to run the simulations. The emerging stable states were linked with biologically meaningful states by interpreting the values of selected factors (Sox9, Col2, Col10, Runx2, etc…) in the same way as biomarkers are used in biological models to define cell states. The stable states were divided into 4 categories: Sox9-positive, Runx2-positive, Sox9- and Runx2-positive; and a state named 'none', in which all variable activities were close to zero. As previously, we considered the Sox9-positive state as a non-hypertrophic or proliferative chondrocyte and the Runx2-positive as a hypertrophic one since this is master transcription factor regulating hypertrophy. The percentage of random initialisations reaching each category after simulation in the absence of external constraints is documented in **Figure 5.5**.

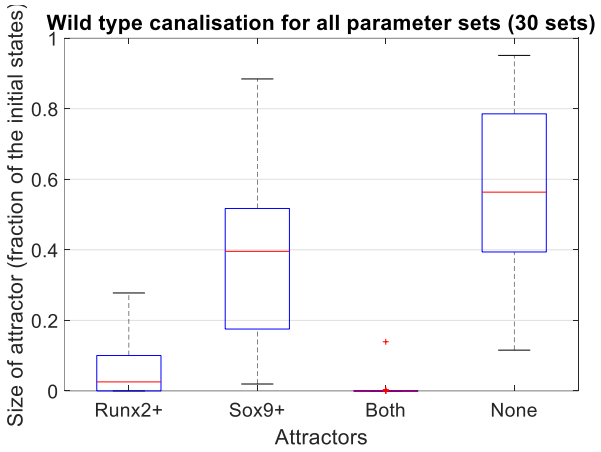


Figure 5.5: The unconstrained canalization for the 30 parameter sets. Overview of the fraction of the state space leading to the four categories of attractors, obtained with the 30 parameter sets. The y-axis indicates the fraction of the initial states (1=100%) leading to each attractor during the Monte Carlo analysis. The boxplot shows the lowest value, the first quantile, the median, the third quantile and the highest value over the ensemble.

The 'none' state had the largest basin of attraction, followed by the state with Sox9 activity. All parameter sets had a 'none' attractor (**Figure 5.5)**. The Runx2-positive state was present in 0 up to 30% of the parameter sets. In contrast, no parameter set lacked a Sox9-positive state. Only one member of the ensemble had an attractor exhibiting activity of both Sox9 and Runx2 ('both' attractor). So, members of the ensemble typically exhibit stable states that were exclusively positive for either Sox9 or Runx2.

In conclusion, these 30 parameter sets made up the ensemble of models, which could recapitulate chondrocyte-like and hypertrophic chondrocyte-like phenotypes. In that first part, we successfully recapitulated the baseline situation obtained in (*265*). In the rest of this study, the ensemble of models was regarded as several potential versions of a chondrocyte on which we simulated the addition of new external treatments since each parameter set may lead to a different outcome. By evaluating the ensemble of models we were able to estimate the robustness of results after treatment when using different parameter sets, the underlying rationale being that results with less variance through the ensemble are more likely to be robust and the corresponding treatment being effective.

5.4 Application to treatment optimization for iPSc-derived chondrocyte differentiation: comparing simulations with experimental results

5.4.1 *In vitro* and *in vivo* experimental set-up

In the parallel experimental study, chondrocytes were derived from iPSCs and were matured for the purpose of bone TE construct development (*4*). Here we summarize the experimental set-up used in that study to obtain the results that are relevant for comparison with our *in silico* experiments. Upon mesodermal induction and subsequent chondrogenic differentiation, human iPSCs condensed and formed cartilaginous nodules (see method in **Supplementary text S5.1**). Those aggregates were then treated for 3 weeks to enter hypertrophy, *in vitro*. The cartilage nodules were pushed towards hypertrophy by using basal hypertrophic differentiation medium (HM): DMEM supplemented with 1% FBS, 1% SP, 1% ITS-X, 50 μg/ml AA, 10 nM dexamethasone (Sigma), 7 mM β-Glycerolphosphate (Sigma), 50 U and 50 mg/ml Pen/Strep. Depending on the differentiation condition, HM was complemented with different growth factors or molecules. More precisely, the effect of BMP4 (100 ng/ml), the canonical Wnt pathway agonist BIO (0.1 μM), the thyroid hormone T3 (1 μM), the combination of those three (BBT3) as well as that combination supplemented for another 10 days with the pro-inflammatory cytokine IL-1B (10 ng/ml) were investigated. The hypertrophic differentiation level and bone forming potential was assessed after *in vitro* suspension culture and after *in vivo* ectopic and orthotropic implantation in a critical size long bone (tibia) defect in mice. The different experimental conditions were compared for their efficiency to promote hypertrophy and bone formation.

5.4.2 *In silico* experiments and results

In silico set-up and experiments

The ensemble model can be regarded as virtual chondrocytes with varying interaction weights in the regulatory network, as explained in the previous section. They are all possible given the data we considered for the reference profiles. All analyses were carried out considering each of the 30 possible chondrocyte networks. The results are expressed in terms of average and standard deviation over the ensemble.

We attempted to mimic *in silico* the experimental conditions in order to assess which of them or what combination of them would be the most potent to induce hypertrophy. To that end, we performed *in silico* knock-out or over-activation experiments of one or several biological factors at the same time by means of a Monte Carlo analysis and we computed the impact of those conditions on the canalization with respect to the unconstrained situation **(**see **Figure 5.4** for a schematic description). Indeed, the effect of an inhibitory molecule was mimicked by forcing the target's activity to remain 0 (inhibited). Contrarily, the effect of a growth factor or a signaling molecule was mimicked by forcing the corresponding node or target's activity to remain 1 (activated). The rest of the downstream nodes/variables were left free to evolve in reaction to those inputs. Five experimental conditions were tested *in silico* as follows*:*

(1) **BMP4**: the use of BMP4 in the medium was mimicked in the *in silico* model by over-activating the node 'BMP', which summarizes the function of BMP signaling through the SMADs but also P38, RAS/ERK and other crosstalks.

(2) **BIO**: the BIO compound (WNT agonist) blocked the activity of the DC, to which GSK3B contributes.

(3) **T3**: The action of 3,3,5-Triiodo-L-thyronine (T3) in the chondrocyte is more complex and no node directly represented T3 or its receptor in the *in silico* model. However, upon literature search we identified that T3 binds to receptors (TRα1 and TRβ1) expressed by chondrocytes and, thereby, activates the Wnt/Beta-catenin pathway (*354*, *355*). Additionally, T3 promotes FGF signaling in chondrocytes (*355*, *356*) and decreases expression of PThrP and its receptor PPR during endochondral ossification (*357*, *358*) . Consequently, in the *in silico* system the overall effect of T3 was modeled by concomitant activation of WNT and FGF (i.e. activities set to 1) and inhibition of PTHrP (i.e. activity set to 0).

(4) **combination of BMP4 + BIO + T3 (BBT3)**

(5) **BBT3 + IL-1B**: Just as T3, there was not a single node for IL-1B in this model, but IL-1B is known to trigger inflammatory pathways involving NFKB activation. Indeed, IL-1B induces phosphorylation of the inhibitor of kappa B (IκB)-α, which leads to the degradation of IKB-alpha and IKB-Beta and triggers the subsequent nuclear translocation of NFKB (p65) (*359*, *360*). As a result, activation of NFKB (i.e. activity set to 1) mimicked the addition of IL-1B, *in silico*.

In practice, to investigate the effect of these conditions on the stable states' stability, the Monte Carlo simulation was run again while imposing the experimental conditions as external

constraints (**Figure 5.4**). The number of initialisations was limited to 2.000 to reduce computational cost. Then, the effect of the imposed cues was assessed by comparing the number of initial states reaching a certain stable state during the perturbed Monte Carlo with that of the unconstrained basal situation. For instance, if 30% of the initialisations led to state *A* in the basal situation and 80% of them led to the same state in the perturbed conditions, then we concluded that the prevalence of the state *A* is enlarged by 50% of the total space in the perturbed condition compared to the basal situation.

In silico versus experimental results

In silico experiments highlight a greater efficacy with combinatorial treatments

The treatment effects during the perturbed Mont Carlo canalizations were compared to the basal situation. The effect of the five aforementioned conditions on the *in silico* system and, particularly, their impact on both the hypertrophic and non-hypertrophic stable states are depicted in **Figure 5.6A**. According to the model, the BBT3 treatment supplemented with IL-1B (condition 5) enlarged hypertrophy prevalence by 91% ± 9 of the total space, as compared to the basal situation. In addition, this treatment decreased the Sox9-positive state stability by 32% of the total space (**Figure 5.6A)**. Consequently, BBT3 combined with IL-1β was the best condition to promote hypertrophy and therefore the condition with the highest endochondral bone formation potential. From the cell population point of view, we can consider that the cells were more likely to adopt a Sox9-positive profile than a Runx2-positive one in the basal situation (**Figure 5.5**). However, under the BBT3+IL-1β condition, the same population might tend towards a Runx2 positive state more easily. BBT3 treatment alone (condition 4) enlarged hypertrophy prevalence by 88% ± 20 of the total space and decreased the Sox9-positive state stability by 33% of the total space (**Figure 5.6A)**. We concluded that under BBT3 treatment, as we model it here, the cartilaginous implants should also be likely to undergo hypertrophy but to a lesser extent than with the combination of BBT3 and IL-1B. Similarly, the combinatorial condition 4 (BBT3) seemed to be similar or slightly better than condition 3 (T3) in promoting hypertrophy. The stability of hypertrophy was more increased with BBT3 than for any of the other separate conditions, BIO or BMP4. Condition 3 (T3) is more likely to promote hypertrophy than condition 2 (BIO) and condition 1 (BMP4). Finally, condition 1 (BMP4) may be better than or the same as condition 2 (BIO) in promoting hypertrophy, given the high standard deviation for the BMP4 condition (**Figure 5.6A)**.

A

	1	2	3	4	5
	T3	BMP4	BIO	BBT3	BBT3 + IL1-β
In silico input	Wnt up FGF up PTHrP down	BMP up	DC down	[Wnt up, FGF up, PTHrP down] BMP up DC down	[Wnt up, FGF up, PTHrP down] BMP up DC down NfkB up
δ(Runx2+)	**0.79** ± 0.28	**0.48** ± 0.30	**0.10** ± 0.10	**0.88** ± 0.20	**0.91** ± 0.09
δ(Sox9+)	**-0.32** ± 0.23	**0.03** ± 0.28	**-0.22** ± 0.15	**-0.33** ± 0.24	**-0.32** ± 0.23
δ(Both)	**-0.01** ± 0.03	**0.03** ± 0.09	**0.00**± 0.06	**-0.01** ± 0.03	**-0.01** ± 0.03
δ(None)	**-0.46** ± 0.40	**-0.54** ± 0.29	**0.13** ± 0.32	**-0.54** ± 0.36	**-0.58** ± 0.25

B

	6	7	8
In silico input	DC up	BMP down	NFkB up
δ(Runx2+)	**-0.05** ± 0.07	**-0.07** ± 0.08	**0.54** ± 0.33
δ(Sox9+)	**0.12** ± 0.07	**-0.06** ± 0.09	**-0.10** ± 0.24
δ(Both)	**0.00** ± 0.03	**0.00** ± 0.03	**0.03** ± 0.09
δ(None)	**-0.07** ± 0.08	**0.13** ± 0.09	**-0.46** ± 0.34

Figure 5.6: Effects of the experimental conditions on the *in silico* canalization for the ensemble model. The row called *in silico* input indicates the network nodes that are perturbed *in silico* to mimic the experimental conditions, with 'up' indicating activation and 'down' inhibition. The lines named after δ(attractor) contain the average variation (over the ensemble) of the canalization with respect to the basal situation, in bold. That value that is also represented by the intensity of the green/red background color. Each average is associated with an error denoting the standard deviation over the ensemble of models. (A) Results for conditions 1 to 5 that reproduce the experimental set-up. (B) Extra *in silico* conditions tested either to show a differential behavior in which hypertrophy stability was not increased (DC up: activation of the β-catenin DC; BMP down: repression of the BMP signal) or to test an additional control condition that could not be assessed experimentally (NFKB up: activation of the inflammatory signal alone).

In the current study, all 5 conditions resulted in a more or less positive effect on hypertrophy promotion since we focused on the conditions of interest that were also tested *in vitro*. Therefore, one might think that any perturbation of the numerical system tended to promote hypertrophy to a lesser or greater extend. To prove this was not systematically the case, we showed that the model was also able to capture other behaviors, by applying additional conditions. The effect of the B-catenin DC activation and of BMP inhibition are reported in **Figure 5.6B** (condition 6 and 7, respectively). Both conditions showed no effect or even a slight decrease of the chance of hypertrophy compared to the baseline situation. Interestingly, the results of BMP signal activation and repression illustrate well the ambiguous effect of this signal. Indeed, even though it has some pro-hypertrophic capabilities (+48% of the state space during activation, and -7% during repression) the BMP signal is not really detrimental for the normal chondrocyte, either (condition 2 and 7 in **Figure 5.6A and B**). It corresponds to the previously reported bipotent role of BMPs in promoting chondrogenesis. Although BMP-

dependent upregulation of Runx2 has been reported, BMPs were also shown to induce Sox9, which stabilizes the articular phenotype in chondrocytes (*361*).

The experimental set-up did not investigate the effect of the pro-inflammatory pathway activation alone (+IL-1B), as the point was to test its benefit in addition to the BBT3 treatment alone. Nevertheless, such a control condition could be interesting to check in order to understand fully the individual effect of each component of the differentiation cocktail. For that reason, we also looked at the single activation of the pro-inflammatory intracellular messenger NFKB in the *in silico* model. We observed that activation of NFKB alone was indeed rather pro-hypertrophic, since the Runx2-positive state was enlarged by 54% of the state space compared to the basal situation, although that effect remained smaller than the combinatorial strategy, with a much larger standard deviation over the ensemble (**Figure 5.6.B)**. This numerical result supports the idea that the combinatorial treatment BBT3+IL-1B brings about a benefit for bone formation that is greater than the individual biochemical effects. Thereby, we showed how *in silico* experiments could complement experimental results when the experimental counterpart is challenging to execute due to technical or resource constraints.

Experimental results corroborate numerical predictions

Wet-lab experiments have shown that combination of BMP, WNT agonist and TH signaling (BBT3) induce more hypertrophic changes and cartilage mineralization *in vitro* than other separate treatments, based on the increased *Runx2/Sox9* and of the *Col10a1/Col2a1* ratios (**Figure 5.7**) as well as the composition changes of the extracellular matrix (**Figure S5.1**). This was in line with our *in silico* results showing an added benefit of the combinatorial treatment. However, despite the changes in gene expression BBT3 did not show satisfying bone formation during the 8 weeks of *in vivo* implantation (**Figure 5.8A&B**). Addition of IL-1B to the BBT3-treated aggregates progressively led to mineralized tissue, and induced bone bridging at 4 weeks of *in vivo* implantation (**Figure 5.8C &D),** as shown by the bone formation with cartilage resorption (**Figure S5.3**). This result was also supported by the numerical simulations in which BBT3 with activation of pro-inflammatory pathways (downstream of IL-1B) increased the chances of the hypertrophic changes over the BBT3 condition alone.

Together, these *in silico* and *in vitro* results support the fact that activating pro-inflammatory pathways potentiate the BBT3 treatment by adding pro-hypertrophic and thus bone-forming constraints that were less strong in the BBT3 condition alone. In addition, IL-1B pathway

activation not only potentiated the effect of BBT3 *in silico* but also significantly reduced the standard deviation in the *in silico* model results, thereby showing a reduction in the degree of variability for the chondrocyte fate in the aggregates (**Figure 5.6**).

Overall, the *in silico* predictions were well in line with the experimentally observed outcomes. As such, computer-based predictions helped identifying which combination of treatments would be the most optimal for hypertrophic differentiation and subsequent bone formation from iPSc-derived cartilaginous tissues, by evaluating their relative potential.

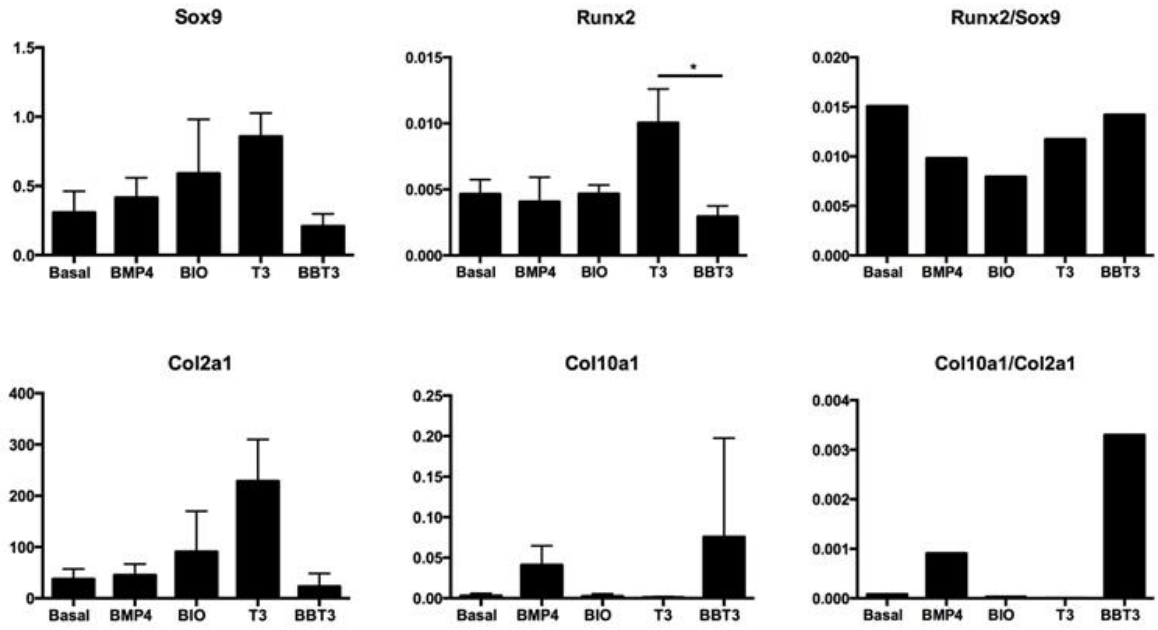


Figure 5.7: Hypertrophic differentiation of cartilaginous organoids. (A) Gene expression levels of Sox9, Runx2, Col2A1 and Col10A1. Upon BBT3 stimulation, a trend towards decrease in Sox9 and Col2A1 was observed, while Col10A1 was upregulated. The Col10A1 to Col2A1 ratio indicate that BBT3 treated samples underwent hypertrophic differentiation. Statistical significance is represented as follow: *p<0.05. (Adapted from W.L. Tam et al. 2021)

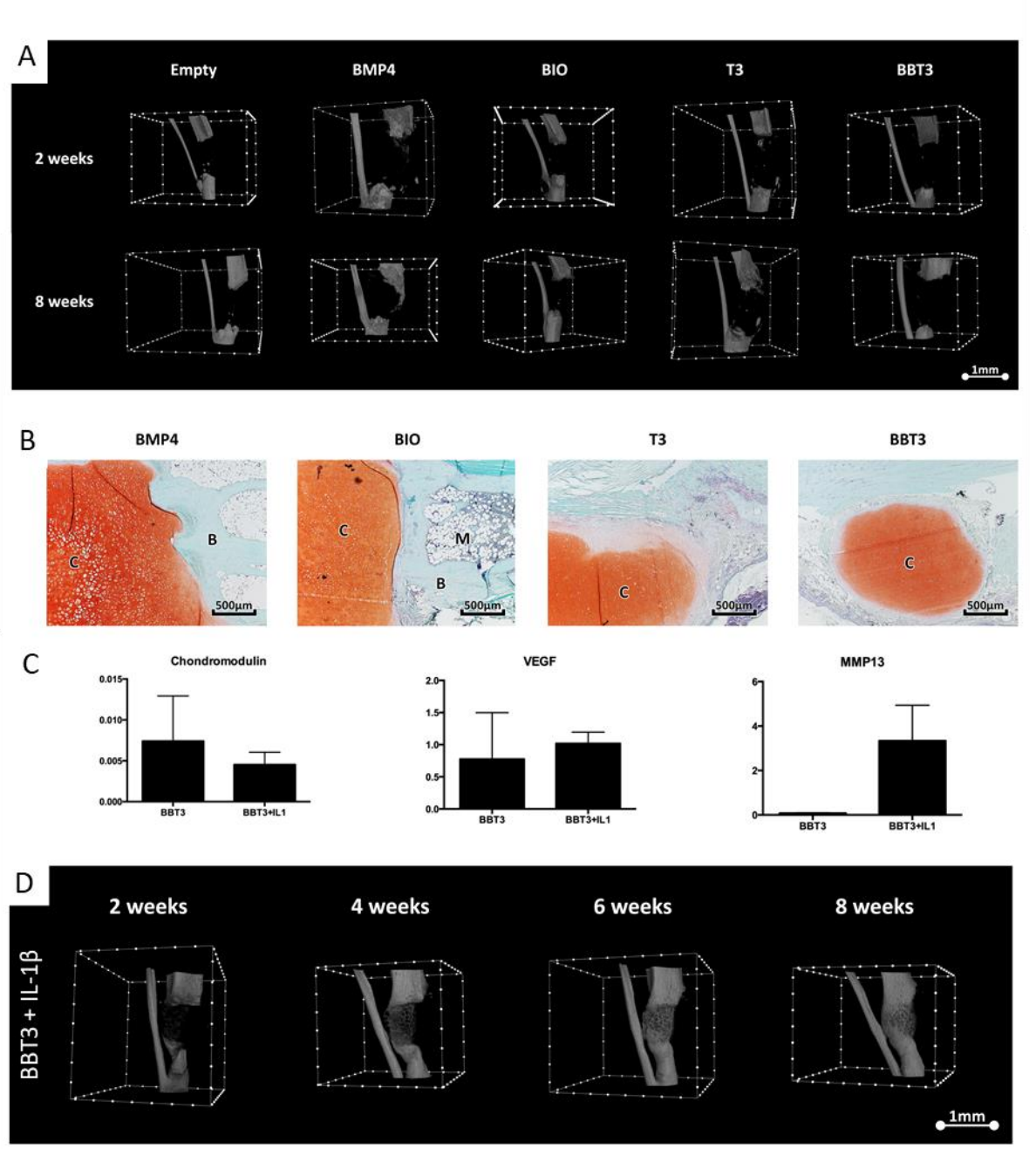


Figure 5.8: Orthotopic implantation of human derived cartilaginous aggregates. Hypertrophically differentiated cartilaginous organoids were implanted in critical size long bone defects. **(A)** Following 8 weeks of implantation, no bone bridging was detected in the BMP4, BIO, T3 and BBT3 treated samples. **(B)** Histological analysis using Safranin-O further indicated that the implanted organoids remained largely stable with no to limited bone formation at the edges of the implant (see zoom images in **Figure S5.2** with visible cartilage-bone turnover sites). The following abbreviations were used to highlight the tissue composition in the tissue sections: B: bone, C: cartilage, M: bone marrow. **(C)** Upon IL-1B treatment, a trend towards decrease of chondromodulin was detected while IL-1B upregulated VEGF and MMP13. **(D)** Following orthotopic implantation, the cartilaginous organoids underwent progressive mineralization with bone bridging being detected after 4 weeks. (Adapted from W.L.Tam et al. 2021)

5.5 Discussion

We have conducted an *in silico* study in parallel to the experimental counterpart to illustrate the capacity of network models to guide the design of experiments for cell differentiation in bone TE. iPSCs were differentiated into cartilaginous aggregates and several treatments were applied to compare the their potential for promoting hypertrophic differentiation and subsequent bone formation. Using a numerical model of the mechanisms regulating chondrocyte differentiation enabled to reproduce the experimental setup *in silico*. More precisely, we have leveraged an in-house developed ensemble network model and applied it to the optimization of bone TE differentiation protocols. We predicted and compared the effect of different treatments and successfully ranked them based on their capabilities to promote hypertrophic maturation and bone forming potential. We also investigated certain situations numerically that could not be evaluated experimentally, to complement and inform the obtained results.

The simulations confirmed the *in vitro* observed greater efficacy of the combinations of BMP4, BIO (a Wnt agonist) and T3, over the separate treatments, to induce chondrocyte maturation and hypertrophy. The *in silico* results also supported the strategy to supplement that combination with the inflammatory cytokine IL-1B to reduce the degree of variability in chondrocyte hypertrophic commitment and bone transition. A strategy that indeed proved successful *in vivo* since aggregate mineralization and bone bridging, through cartilage remodeling and bone formation was observed after 4 weeks. In addition, an extra control situation could be assessed *in silico* to mimic an IL-1B treatment alone, which confirmed that the added potential of the combinatorial strategy BBT3+ IL-1B could not be reached with IL-1B alone. Importantly, the current *in silico* study was performed in parallel to the experimental one, in a blinded way, and respective results were only revealed upon final comparison of the outcomes to ensure objectivity. This illustrated the capacity of network models and CM&S to support experimental strategies and findings, and to eventually guide experimental design when CM&S is used ahead of the experiments for informing investigators.

We showed that CM&S may support experimental results but it can also help to propose potential underlying mechanism of actions. For instance, the mathematical equations and prior knowledge network provided possible mechanistic explanations for the experimental observations. Indeed, the cumulative effect observed *in vivo* with IL- B and *in silico* with the activation of NFKB (downstream of IL-1 B) can potentially be attributed to the capacity of NFKB to bind to the promoters and upregulate the expression of *MMP13*, and *HIF2A* (*185*,

362). The pro-inflammatory signal may also tip the effect of BMP signaling towards the pro-hypertrophic path by blocking SMAD2/3 through upregulation of SMAD7 and activation of ERK1/2, two SMAD2/3-repressing proteins (*71*, *160*, *363*). This illustrates how *in silico* results may support (or challenge) the mechanistic hypothesis put forward in a study.

Network models are often build through mechanistic approaches, i.e. each edge in the network is supported by known mechanisms and observations. Nevertheless, it brings the issue of dealing with low confidence interactions and uncertainty in specific biological contexts. Using a reverse approach may indeed result in several possible networks (non-unique solutions), but that very characteristic allows to deal with the uncertainty in the network structure and tackle those low confidence interactions. In this study, we got candidate parameter sets that were consistent with expected reference profiles and from there, we generated new predictions based on the behavior of all consistent networks. Because we run our analyses on the ensemble, our conclusions account for the uncertainty of some interactions but we did not characterize or evaluate further the individual parameter sets that had been retained. Nevertheless, there exists formal verification systems to automatically test if a given network predicts certain behaviors in order to agree on the most likely possible network model, thereby allowing to conclude on the low confidence interactions. This could be subject to future investigations in the context of chondrocyte differentiation for bone TE. For instance, a tool called Reasoning Engine for Interaction Network (RE:IN) was recently developed to test whether logical expression (with temporal logic) could be satisfied by one or more concrete networks (*364*). It was applied to stem cell differentiation case studies to establish the requirement for certain interactions in myeloid progenitor differentiation (*364*, *365*) or to elucidate the dynamic regulatory network architecture for self-renewing mouse embryonic stem cells (*366*). Overall, ensemble approaches and network optimization/selection methods appear as potent tools in the context of computational (stem) cell biology and tissue engineering; either to make predictions that account for network interactions uncertainty or to discover and confirm network structures underlying cellular programs and functions.

Another challenge when reproducing an experimental set-up *in silico* is dealing with the accuracy of the correspondence between both set-ups. We made a certain amount of assumptions and simplifications while mimicking the different experimental conditions numerically, although these assumptions were based on the state of the art biological knowledge. The best illustration of this is the effect of T3 in chondrocytes, which is complex and could support alternative interpretations than the one used in this study. For instance, in our

in silico analysis T3 treatment is represented by activation of WNT and FGF signaling and repression of the PTHrP signal. So, it does not directly affect the value of the IGF-I receptor even though T3 treatment has been shown to stimulate PI3K/AKT/GSK3B signaling in some *in vivo* and *in vitro* models (*367*), which are important signal transducers of IGF-I signaling. To picture what could be the effect of IGF-I signal stimulation in the model, we can note that AKT can inactivate GSK3B, a negative regulator of the canonical Wnt/β-catenin pathway. Therefore, we can assume that the hypothetic IGF-I-mediated effect of T3 on chondrocyte hypertrophy would pass by the downstream activation of the Wnt pathway (*355*, *367*). However, in our *in silico* model, this effect cannot be retrieved since the role of GSK3B in the destruction complex is separated from its role in other crosstalks. Indeed, Doble & Woodgett, (2003) argue that GSK3B activity in the Wnt pathway appears to be insulated from regulators of GSK-3 that lie outside of the Wnt pathway (free GSK3B) (*368*, *369*). For example, some studies have shown that insulin signaling leads to inhibition of GSK3B via phosphorylation but does not cause accumulation of β-catenin. Thus, the connection between free GSK, downstream to IGF-I, and the DC (in the Wnt pathway) was not included in the *in silico* model. Because of this choice during model construction and because of other crosstalks, the global behavior of IGF-I in the ensemble model tends to be neutral for the hypertrophic-like state: single inhibition or activation of IGF-I neither increases nor decreases the hypertrophic state stability (data not shown). Still regarding the T3 treatment, another study has reported that BMP signaling is also likely to lie downstream of T3 treatment (*358*). However, the evidence is rather indirect since the authors only conclude that T3 increases *COL10A1* expression and BMP signaling would be increased prior to that stimulation of *COL10A1*. No evidence of direct interaction between the SMAD proteins (from BMP signaling) and the collagen X promotor was reported at that time (*358*). With this, we decided not to stimulate BMP signaling as part of the 'T3' condition. In any case, activation of BMP signaling alone and in combination with T3 was already part of the tested conditions of interest and did promote, to some extent, the hypertrophic state's prevalence. This shows the importance of clearly stating assumptions and making conclusions accordingly. Even though this type of uncertainty may be tackled by investigating alternative situations, we believe the simulations' credibility can be better established when the assumptions made to represent the experimental set-up are supported by literature and validated or acknowledged by experts/experimentalists from the field beforehand.

In conclusion, the application of an ensemble network model of the chondrocyte showed promising predictive power to find the best differentiation biochemical cocktail for promoting

bone formation in iPCS-based TE strategies. This encourage further efforts to improve and tailor such tool for other applications. Given the gain of resources offered by numerical simulations compared to experimental procedures and given the interpretability of mechanistic regulatory network models, we confirm with the current study that this type of tool constitutes an important asset to inform experimental design.

Chapter 6. Interplay of mechanoregulatory and inflammatory pathways: towards a multiscale view of articular cartilage homeostasis and degeneration

This chapter is based on a collaborative work with S. Mukherjee (*370*).

6.1 Introduction

Articular cartilage is a tissue whose weight bearing and load transmission functions are essential. Each physical impact on the articulation, whether it happens during normal gait or during a traumatic event, reverberates deeper in the tissue. The generated forces propagate within that tissue to the cells, the chondrocytes. Not only are these cells capable of sensing the mechanical stimulation thanks to mechanosensitive receptors, but they also react by adapting their internal machinery accordingly. While physiological mechanical cues are crucial for chondrocyte homeostasis, abnormal mechanical signals can lead to departure from the stable cartilage homeostasis with notably the rise of chondrocyte hypertrophic changes. Apart from being mechanosensitive, chondrocytes can also react to biochemical pro-inflammatory cues. As compared to rheumatoid arthritis, OA patients usually show low-grade synovitis, as such it is not considered as a typical inflammatory disease (*143*, *152*, *371*). Nevertheless, the molecular pathways that transduce signals downstream of the so-called pro-inflammatory cytokines (IL-1B, TNF-A, etc.) play an important role in the control of chondrocyte function. Multiple studies report their implication in the onset and progression of cartilage degradation during OA (*372*, *373*).

Mechanoregulatory and inflammatory signals are strongly related in cartilage. Indeed, there is experimental evidence for the presence of a threshold below which increased loading magnitude can decrease the cytokine-induced inflammatory response, and above which a further increase in the loading starts enhancing the inflammatory pathways (*359*). This threshold effect also resonates in the relationship between mechanical loading and cartilage degradation (*178*). Having a systemic view of the tight regulations happening between biomechanical and inflammatory signals is one of the key elements for deciphering OA mechanisms and for proposing therapeutic interventions, including mechanical adjustment and selective pharmacological targeting.

Mechanical signal transduction is inherently a multiscale process. The effect of mechanical loads on cartilage matrix biosynthesis or degradation during OA most likely originates from intracellular biochemical changes at the chondrocyte level. Indeed, matrix and collagen network damage is generally observed first around the cells in the pericellular region, and then extends progressively into the ECM (*374*). Therefore, modeling the impact of mechanical loading on local intracellular reactions and cellular metabolism (i.e. anabolism and catabolism) and linking this back to the ECM or cartilage mechanical properties might help to decipher those

mechanisms. That challenging task requires splitting the processes into scales and integrating cell regulatory networks into the mechanical models (*262*). Previous models of chondrocyte mechanostranduction or multiscale cartilage models were often linear, deterministic or phenomenological for the intracellular part (*359*, *375*) and/or lacked detailed cell regulatory networks (*257*).

Here, we aim at providing a mechanistic chondrocyte intracellular module with the perspective to integrate it into a multiscale mechanical model of the human knee joint. In particular, the purpose is to update the intracellular module with relevant mechanisms for studying the interplay between mechano-regulatory and inflammatory pathways and their influence on the chondrocyte phenotypic changes. Investigating this crosstalk is necessary for further understanding how mechanical loading influences cartilage tissue (*376*). As such, the specific objectives of this chapter are to describe the mechanistic content of the intracellular module and to study the effect of physiological and pathological mechanical signals on chondrocyte phenotype and related ECM status. This is done while taking advantage of the differential outputs from a knee joint biomechanical model, in order to investigate zonal effects (i.e. cartilage superficial zone (SZ) versus deep zone (DZ)). The pathological case study considered here is that of meniscectomy, a well-known risk factor for OA development (*377*).

This chapter starts with a description of the biological mechanisms integrated into the intracellular regulatory networks. Then, we explain the modeling strategy and the adaptations made to the algorithms (as compared to the model in Chapter 4) to reflect the experimental settings. Finally, we investigate the relation between (un-)physiological mechanical cues and inflammatory pathways as well as their impact on the SZ and DZ chondrocyte phenotypes, with the multiscale *in silico* model.

6.2 Biochemical mechanisms and assumptions for the intracellular module

This section provides a detailed description of the biochemical mechanisms about mechanotransduction, inflammatory pathways, and their crosstalk in chondrocytes that were included in the mathematical model. This extensive description constitutes both a literature study, in which mechanisms are supported by external references, and a declaration of the biological assumption and simplifications made when building the model. For other pathways (i.e. not related to mechanotransduction or inflammation), we refer the reader to the introductory **Chapter 1** and the previous articular cartilage model **(Chapter4).** Important

interactions and crosstalks related to the mechanotransduction and inflammatory pathways are reported in an activity flow diagram for illustration purposes. It represents a non-exhaustive part of the full interactome integrated in the model, the complete list of genetic and protein-protein interactions is available in supporting information **(Table S6.1).**

6.2.1 Mechanosensing and mechanotransduction

Integrin mechanoreceptor and downstream pathways

One of the main sensors by which chondrocytes sense the external loading are the transmembrane integrin proteins and focal adhesion complexes. These complexes are made of integrins and other cytoplasmic molecules, which provide kinase activity to the complexes (*378, 379*). They mediate mechanical force transmission to downstream biochemical reaction cascades. There are several integrin isoforms and subunits that differ between species (e.g. α1β1, α5β1) and that have binding sites for different ECM fibers (e.g. collagen, fibronectin), which act as external ligands (*192*). The α5β1subunint has been shown to be important for modulating the expression of MMPs in mouse chondrocytes (*380*), but too few studies distinguish downstream cascades and effects specific to each integrin subunit. Thus, the node representing integrins in the model is a generic receptor translating mechanical cues into biochemical cascades, for sake of simplification. One specific human gene coding for this type of protein could be *ITGB1*, for instance.

Generally speaking, FAKs are among of the main effectors of focal adhesion complexes for signal transduction (*381*). More actors are involved to offer refined response to various loading regimens but, for sake of simplification, our model summarizes these molecular complexes by a single node named FAK. This type of kinases function as a hub to transduce signals, including through phosphorylation of the RhoA GTPases, which are known to be involved in the regulation of cell response to mechanical cues (*382*). More particularly, these phosphorylation cascades are involved in the regulation of cell differentiation by mechanical cues in development and regenerative processes (*194*). RhoA proteins and the downstream ROCK1/2 proteins regulate the actin cytoskeleton assembly and contractility by phosphorylating myosin light chains (MLC) (*383*). This level of contractility and cytoskeleton assembly is represented by a single variable named Actin-MLC in the model. Evidence shows that RhoA and ROCK are involved in the suppression of hypertrophic chondrocyte differentiation (*384*), possibly in part through activation of the cyclin CCND1, a known suppressor of RUNX2 activity (*384*),

which is accounted for in the model. Actin-myosin phosphorylation by the RhoA/ROCK axis also takes place through LIM-domain containing protein kinase and cofilin proteins (*383*, *385*), but this redundant side pathway does not involve many crosstalk with other actors and is not represented in the model for sake of simplification.

Other downstream effectors of the RhoA/ROCK axis are multiple. First, ROCK promotes SOX9 phosphorylation, thereby promoting its transduction to the nucleus in chondrocytes (*386*). SOX9 is commonly known to directly upregulate expression of the genes producing ECM proteins such as proteoglycans and type-II collagen (*386*). In other words, it is fair to presume this axis as a mechanically induced pro-anabolic pathway, when considered in isolation of the full connectome. Second, the phosphorylation of the cytoskeleton stress fibers and MLC, which increases the stability and contractility of the cytoskeleton, regulates the transcriptional activity of various TFs (*387*). For instance, it promotes the nuclear translocation of B-catenin, an effector of the Wnt canonical pathway, thereby promoting the WNT/B-catenin axis (*194*). In the same way, it also promotes the activation of Yes-associated protein (YAP) and the transcriptional co-activator with PDZ-binding motif (TAZ). Actin contractility would also up-regulate ECM related genes such as *COL2A1* and *ACAN* (or Aggrecan, a proteoglycan), but the exact nature of the involved TF is not clear. For that reason, we introduced a dummy TF in the model representing the transcriptional activity induced by the cytoskeleton-increased contractility (more details below).

Mechanical stimulation of chondrocytes increases their sensitivity to mechanical cues, suggesting the involvement of a positive feedback loop. There is experimental evidence that the amount of cyclic loading increases integrin receptors at the membrane after mechanical stimulation (*388*). We propose a mechanism in which, upon integrin stimulation, downstream pathways promote the integrin gene up-regulation, thereby increasing the abundance of integrin receptors. In the model, we assume that this positive transcriptional activity happens through the dummy-TF introduced earlier, located downstream of the RhoA/ROCK/Actin-MLC pathway. A possible candidate for that dummy-TF could be serum response factor (SRF) because the Rho family of GTPase (RhoA, Rac1) as well as actin contractility are reported to promote the SRF transcription factor activation upon mechanical stimulation which in turn regulates cell contractility related genes (*379*, *389*, *390*). In absence of specific information, this dummy-TF or SRF is not regulated at the transcriptional level in the model.

As mentioned previously, YAP and TAZ have been identified to be mechanosensitive and to be involved in the transduction of mechanical signals (*195*). They can be activated or deactivated according to the mechanical environment, such as ECM stiffness. YAP and TAZ are considered as analogous; nevertheless, downstream effects are often reported in the literature for YAP rather than TAZ (*391*). In the current model we do not make any distinction between the two and they are represented by a single entity. YAP inactivation is induced by phosphorylation, which prevents its translocation to the nucleus and may even direct it towards proteasome degradation (*392*). Mechanically-induced YAP/TAZ activation, downstream of the RhoA/ROCK/Actin-MLC axis, would be independent from the often reported NF2/Hippo/LATS biochemical pathway (*195*). For that reason, the Hippo pathway is not represented in the model. Furthermore, YAP and TAZ are involved in the regulation of the chondrogenic phenotype in response to physical cues (*194*, *391*, *393*). For instance, YAP was reported to block hypertrophy during skeletal development (*391*) while the overall amount of YAP is reduced during OA (*392*). Mechanistically, YAP contributes to *SOX6* upregulation and reduces RUNX2 activity by direct binding, eventually preventing RUNX2 to upregulate *COL10A1* expression (*391*). Our mathematical model accounts for these mechanisms by letting YAP/TAZ increase SOX9 at genetic level, which represents the SOX trio (SOX9, 5 and 6), and decrease RUNX2 protein activation. With the lack of specific information about their transcriptional regulators in cartilage, YAP/TAZ are only modulated at the protein level and not at the transcriptional level in the present numerical model.

Finally, mTOR is a well-known mechanosensitive gene that is activated by cyclic loading in chondrocytes (*394*). This mechanically induced activation was reported to happen downstream of the PI3K/AKT/mTOR cascade in osteoblasts. This axis also being active in chondrocytes, we introduced an activation link directly from integrins to PI3K in the model. Inhibition of mTOR has been shown to reduce OA severity (*395*) but the implications of mTOR in chondrocyte biology are manifold and we only introduced a few downstream effectors in the model. We modeled mTOR capacity to phosphorylate and help the nuclear translocation of GLI2 (*396*), the nuclear effector of the IHH pathway. This interaction also accounts for the ability of mTOR to regulate the PTHrP pathway via GLI2 in chondrocytes (*396*) since GLI2 was already modeled as an upstream positive regulator of *PTHrP* expression. In addition, mTOR may enhance *RUNX2* expression through indirect activation of the transcription factors DLX5 and GLI2 (*397*), and is capable of suppressing its own inhibitor PP2A in the model (*398*, *399*), generating a positive feedback loop. The involvement of mTOR in different complexes

(i.e. mTORC1 and mTORC2) is not considered here and many mTOR-related mechanisms are disregarded in the current model.

Ion channel and calcium signaling

Beside integrins, other important mechanosensors are expressed by chondrocytes, including the the so-called chondrocyte "channelome". We focused on the cation channels entitled transient receptor potential vanilloid 4 (TRPV4), largely located at the cell membrane of articular chondrocytes (*400*). They are multimodally activated channels as they can transduce dynamic compressive loading signals and be activated by osmotic pressure variations as well as chemical signals (*193*, *401*). TRPV4 plays a major role in the physiological response to cartilage loading by mediating intracellular calcium cations (Ca^{2+}) influx (*193*). Functional mutations of the TRPV4 channel have been related to skeletal dysplasia and joint degenerative conditions such as osteoarthritis in humans (*193*, *402*). *TRPV4* expression and activation has often been related to *SOX9* increased expression as well as overexpression of ECM related proteins such as type II-collagen and proteoglycans. As such, it is believed to be involved in chondrogenic differentiation (*403*).

Importantly, Ca^{2+} signaling is an essential universal secondary messenger in articular chondrocytes and its regulation is reported to be highly complex and based on frequency of concentration oscillations (*404*, *405*). The current model being discrete in time and semi-quantitative, it does not allow to represent this tightly regulated periodic flux, which could be the subject of a dedicated time continuous model in itself. Instead, in our model, the TRPV4 entity can be activated with increased mechanical loading, based on the same forces that are sensed by integrins. It subsequently increases the activity of the node called Ca^{2+}, representing the concentration or effective level of cytoplasmic Ca^{2+}. In that sense, we can consider that the maximal value of the variable would correspond to the most optimal level and frequency of Ca^{2+} oscillations and the lower value to less physiological frequencies, since the spontaneous calcium signaling is known to be altered in disease (*406*). Actually, both intra (e.g. endoplasmic reticulum) and extra-cellular sources of Ca^{2+} exist but the current model focuses on the extracellular source of influx as supported by Lv. et al. (*404*) according to which Ca^{2+} signaling, either spontaneous or load induced, has to be initiated by the extracellular Ca^{2+} influx. CREB2 is one of the downstream transcription factors deciphering the information encoded by Ca^{2+} concentration. This connection is represented in the model by an activation link toward ATF4 (also known as CREB2), which was already present.

TRPV4 stimulation with dynamical mechanical loading or agonist treatment was reported to modulate ECM protein biosynthesis. In particular, it increases *COL2A1* expression and decreases the mRNA level of ADAMTS5, an ECM proteinase (*193*). The former observation was not directly encoded in the model but it is expected to be mediated by the activating effect of Ca^{2+} on the cytoskeleton contractility and on *SOX9* expression in the model, as mentioned above. The latter is represented by an inhibitory edge relating Ca2+ activity and ADAMTS5 gene expression, even though that relationship is likely to be indirect and to involve an intermediary TF, which is not precisely identified to our knowledge.

Interestingly, this TRPV4-dependant mechanosensitive pathway can be modulated by, and interact with, other pathways. For instance, Trompeter *et al*. have recently showed that the TRPV4 channel is regulated by the actin cytoskeleton through binding of F-actin to the Microtubule Associated Protein Binding 7 domain of TRPV4. They have also shown that TRPV4 activity is suppressed by IGF-I treatment concluding that IGF-I would modulate the activity of this channel through alteration in the actin cytoskeletal organization (*407*). Indeed, IGF-I treatment in primary chondrocytes increased cell stiffness and F-actin formation (*408*, *409*). In order to integrate this precise mechanism in the model, we made sure that concomitant activity of the IGF-I receptor and F-Actin is able to decrease the TRPV4 protein activity (i.e. channel opening) by using a multiplicative factor in the equations. That way, the Actin-MLC can inhibit TRPV4 activity only if the IGF-I receptor is activated by its ligand, and this in a dose dependent way. Moreover, this calcium signaling interacts with the Rho GTPase/Actin axis to promote F-actin stability and contractility (*410*), which is translated by an activation link between Ca^{2+} and RhoA/ROCK in our model.

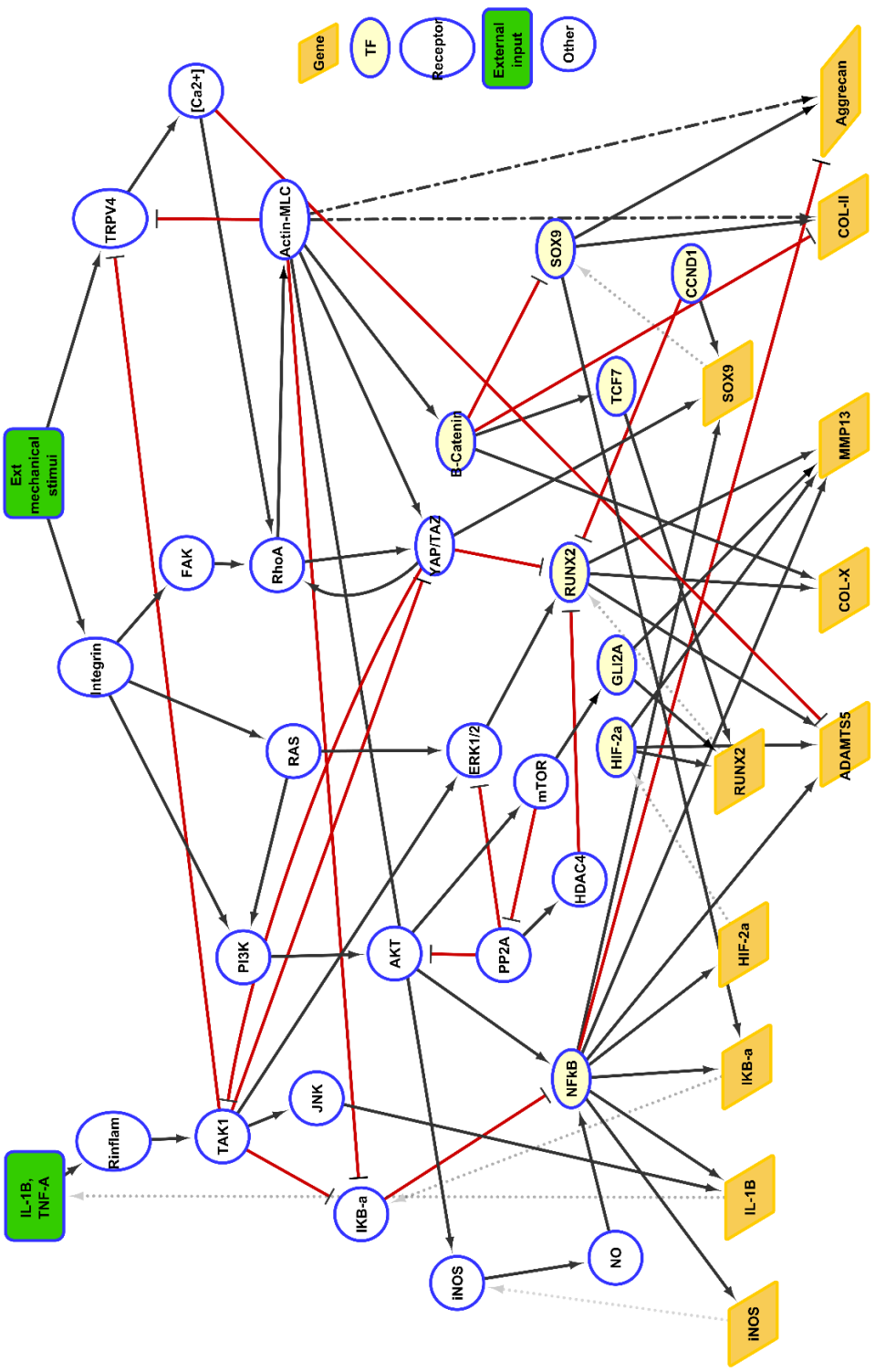


Figure 6.1: Snapshot of the mechanotransduction and pro-inflammatory signaling and their crosstalk in the *in silico* chondrocyte. This activity flow diagram is an arbitrary selection of important cross-talks in the model that are related to mechanotransduction or inflammation, for illustration. The complete network includes 74 nodes and 299 interactions. External cues (in green) are either the mechanical loading or pro-inflammatory cytokines (e.g. IL-1B or TNFA). The signal is transduced through signaling molecules ('Other') up to transcription factors ('TF' in blue circled yellow ellipse), which up- or down-regulate the expression of downstream genes (yellow parallelograms). As a result, genes increase the amount of the corresponding upstream proteins (grey dotted lines). Black arrows (resp. red T-ended lines) represent activation (resp. inhibition) interactions or influences. Influences may be direct or indirect in chondrocytes, the dashed black lines represent interactions that involve an intermediary node in the model but that is not represented in the graph.

6.2.2 Inflammatory pathways

Increasing evidence has highlighted the implications inflammatory events in OA onset and progression (*173*, *180*, *411*). Even though a relatively low synovitis grade is observed during OA, this pathology is characterized by an auto-activating inflammatory loop. Indeed, in

presence of some ECM degradation products or DAMPS, inflammatory receptors such as the Toll-like receptors trigger signal transduction resulting in pro-inflammatory cytokine production by the chondrocytes themselves (*143*, *145*, *171*, *177*). In this section, we summarize the main inflammation-related pathways that we integrated in our model and the assumptions made.

First, several pro-inflammatory mediators, for which chondrocytes express the specific receptors (e.g. Toll-like receptor), are responsible for inducing pro-inflammatory events. For the sake of simplification in the model, we represent them by a single pro-inflammatory cytokine entity (e.g. IL-1B or TNF-A) that is able to signal through a unique receptor of inflammation (R-inflam). This is an important assumption allowing us to focus on the intracellular interplay between mechanotransduction and inflammation rather than the role of specific receptors and signaling molecules.

In chondrocytes, IL-1B increases the production of proteases responsible for matrix degeneration, suppresses matrix biosynthesis and induces pro-inflammatory mediators by controlling expression of genes such as inducible NO synthase (iNOS) or matrix metalloproteinases (MMPs) (*179*, *411*). In addition, pro-inflammatory cytokines like IL-1B promote their own production, which amplifies the immune response as a positive feedback loop (*412*). This process is mediated by several signaling cascades including TAK1, NFKB and MAPKs. We refer the reader to Chapter 1 for more details and we shortly recall the main axes in the paragraphs below.

In particular, TAK1 is a critical point where signals generated by pro-inflammatory signals converge, notably to initiate the NFKB signaling cascade and pro-inflammatory gene induction (*182*, *412*). The canonical mechanism of action involves a double inhibition and the release of NFKB from its inhibitor, IKBA (*177*). Since the other upstream activators of NFKB are only able to activate it for nuclear translocation when it is free from IKBA, we devised a special term in which the sum of the upstream activators is multiplied by the term (1- [level of IKBA]) in the model's equation. In addition, there exists a negative feedback loop in which NFKB promotes the gene expression of its own inhibitor IKBA (*177*).

iNOS is an enzyme that can be induced by pro-inflammatory cytokines and that promotes the production of nitrite oxide, a pro-inflammatory mediator feeding the inflammatory loop in abnormal situations (*413*, *414*). Indeed, in cytokine-stimulated chondrocytes, induced nitric

oxide causes persistent activation of NFKB (*413*). The inflammation inducible effect is represented in the model by overexpression of *iNOS* by NFKB, since the promoter region of the *iNOS* gene contains an NFKB binding site (*413*, *415*), as well as the direct activation of NFKB by NO.

Another pathway involved in the activation of chondrocyte inflammatory response is that of Wnt signaling. Indeed, overexpression of TCF4, a WNT downstream effector, was shown to induce the expression of MMPs and the activation of NFKB in chondrocytes. TCF would directly bind to NFKB to help its nuclear translocation in articular chondrocytes (*416*). This is accounted for by a direct activation link between LEF/TCF and NFKB at the protein level.

The aforementioned auto-inflammation loop however does not function in a fully uncontrolled fashion. There exist suppressors of cytokine signaling (e.g. SOCS proteins), which are soluble inhibitors of pro-inflammatory cytokines receptors (R-inflam, in the model) (*182*). They prevent the downstream phosphorylation cascades that normally lead to NFKB activation and negatively regulate the cytokine induced-JAK/STAT pathway during osteoarthritis (*188*). SOCS proteins are upregulated by the FGF and IGF-I pathways, which illustrates some of the crosstalk happening with other signaling pathways in the model. Another way that IL-1B-R mediated signaling can be terminated is by activating some negative feedback loops. For instance, the p38 MAPK induces the phosphorylation of TAB1 inactivating TAK1, and NFKB can overexpress its own inhibitor IKBA.

Inflammatory signaling also interplays with IGF signaling in several ways. One way involves the IGF binding proteins, which bind to IGF-I and prevent downstream signaling. These proteins are known to be up-regulated during hypertrophy and OA (*166*, *167*). While IGF-I signaling may act as a break to inflammatory signaling via up-regulation of SOCS proteins, IL-1B promotes the production and release of IGFBPs (*417*), which is translated by the up-regulation of *IGFBP* expression by the nuclear effector NFKB in our model, although the real mode of action is not clear yet.

Finally, FOXO is an important anti-catabolic TF modulated and downregulated by inflammation (*190*). It is represented in the *in silico* model via its ability to promote proteoglycans and collagen production, and to downregulate *MMP13* and *ADAMTS5* (*139*).

6.2.3 Interplay between pathways

Previously we compartmented pathways into the mechanoregulatory and inflammatory pathways, though in fact they are interconnected. Madhavan *et al*. have shown that cyclic tensile strain suppresses pro-inflammatory gene induction by all the three inflammatory mediators, IL-1B, TNFA, and LPS (*412*). However, the exact relationship is complex and neither linear nor straightforward. Here, we describe some of the most direct molecular mechanisms by which mechanical loading and inflammation influence each other, and that are accounted for in the current *in silico* model.

First, signal transduction through integrin receptors such as integrin subunits α5 and β1 has been shown to up-regulate *MMP13* expression in chondrocytes, when associated with inflammatory cues. Indeed, the effect is likely mediated through induction of IL-1B (*380*, *388*). It corroborates with the finding that excessive mechanical loading causes NFKB activation and OA development (*359*). However, this observation likely depends on the status of all other connected signaling pathways as well as on the magnitude of the biomechanical signals. As to the latter, it was repeatedly reported that mechanical signals at lower or physiological magnitudes maintain healthy cartilage and are potent inhibitors of inflammation and matrix degradation (*359, 412*, *418*). The TAK1/NFKB pathway has been suggested as key player of this biomechanical signal-mediated anti-inflammatory effect (*182*). In normal physiological ranges, mechanical stimulation would block TAK1 phosphorylation, thereby preventing IKK phosphorylation and activation (*182, 412, 419*). A possible mode of action that was integrated in the model regards YAP/TAZ. Indeed, YAP and NFKB pathways reciprocally inhibit each other, possibly affecting cartilage degradation in OA (*392*). On the one hand, IL-1B and TNFA promote YAP/TAZ degradation through TAK1-mediated phosphorylation (*392*). On the other hand, YAP directly binds with TAK1, limiting NFKB activation by blocking reaction substrate accessibility (*392*).

Second, iNOS and NO are key players in the mechano-inflammatory crosstalk. Indeed, the pro-inflammatory mediator iNOs is also mechano-inducible (*420*, *421*). This is represented in the model by a direct activatory link between Actin-MLC, the most direct intracellular mechanical transducer, and iNOS at protein level. However, the exact mode of action is not known.

Third, ECM degradation products, also called DAMPs are generated when ADAMTS5 proteinases are released by chondrocytes or in case of adverse events, as observed in OA(*1*).

Those degradation products interact with membrane receptors of cytokines such as Toll-like receptors (TLRs) but also integrins, eventually increasing pro-inflammatory and catabolic products. In the model we shortcut the implication of these DAMPs and we added a link describing the indirect activation of inflammatory cytokine receptors and integrins by ADAMTS5, assuming a linear relation between ADAMTS5 and DAMPS for simplification. However, calcium signaling (TRPV4/Ca^{2+}), which is mechanosensitive, inhibits *ADAMTS5* expression in chondrocytes. This is another illustration of the multiple and complex connections between mechanotransduction and inflammation.

In the end, cells integrate all the aforementionned signals and modify their secretory profile accordingly, which in turn affects their matrix mechanical properties and cells' mechanobiological environement (*422*), eventually leading to changes in the external cues (**Figure 6.2**).

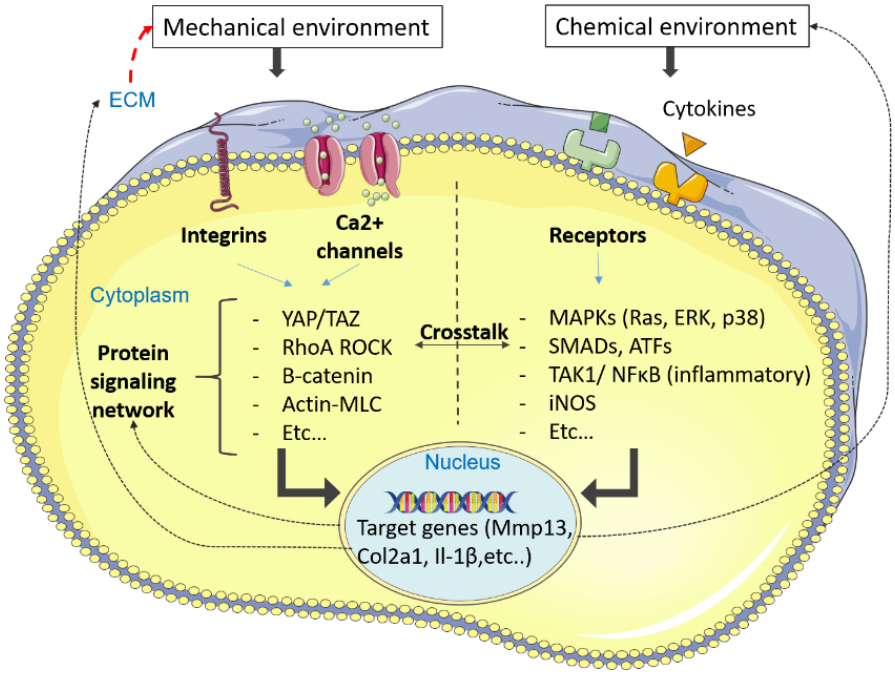


Figure 6.2: schematic representation of the mechano-biochemical coupling in chondrocytes, as conceptualized and modeled *in silico*. Mechanical stimuli and changes in ECM mechanical properties affect the direct mechanical environment of chondrocytes, which is sensed by sensors such as integrins and TRPV4 calcium channels. On the other hand, the biochemical environment (e.g. inflammation) affects chondrocytes' internal machinery via ligands binding and intracellular signals transduction. There are multiple interconnections between mechanotransduction and inflammatory pathways. Eventually, cells integrate all messages at nuclear level and modify their genetic expression profile accordingly. This generates a feedback, modifying protein cytoplasmic abundance and cellular secretory profile. Modifications in ECM proteins and degrading-enzymes production may contribute to changes in ECM composition and mechanical properties, thereby initiating a new loop. The Figure was created thanks to infographics from smart.servier.com, licensed under the Creative Commons Attribution 3.0.

6.3 *In silico* model and its relation to the experimental and biological setup

The intracellular processes, triggered by mechanical signals and sensed at cellular level, are modeled here using the gene/protein regulatory network described in **Chapter 4** and extended with the established and/or investigated mechanotransduction and inflammatory pathways described above. In order to translate (patho)physiological loading at the joint level into mechanical signals at the tissue level and further down onto the micro-scale cell level, a multiscale finite element (FE) model is required. Information about this multiscale FE workflow (**Figure 6.3)** is briefly provided below since its numerical outputs constitute the main inputs for the intracellular layer currently discussed. However, this multiscale workflow was not established in the context of the present thesis (developed by S. Mukherjee (*370*)), hence the rest of the discussion focuses on the modeling strategy of the intracellular module.

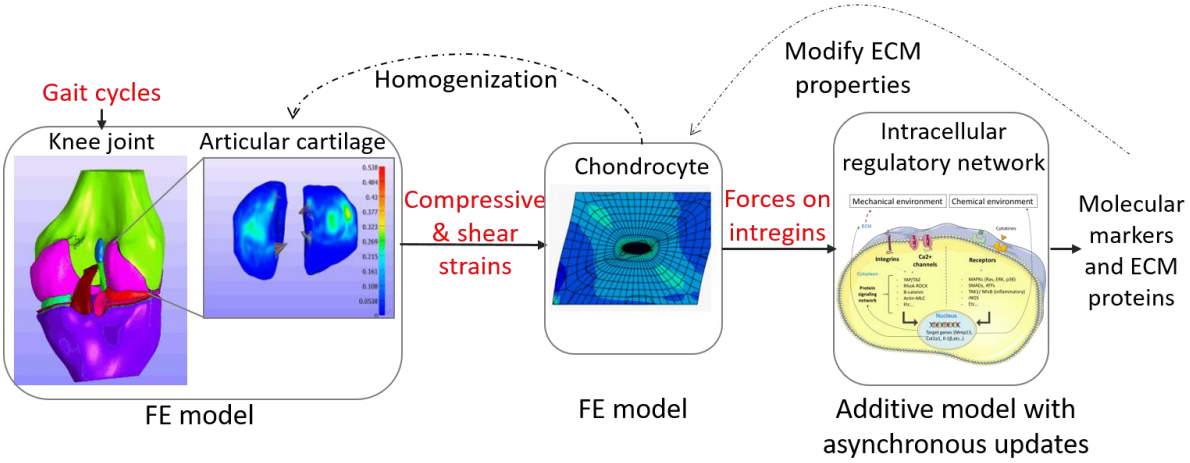


Figure 6.3: Location of the mechanobiological intracellular layer within the multiscale workflow. The first scales (tissue level: knee joint with articular cartilage and cell level: chondrocyte) are represented with finite element (FE) models. The mechanical forces resulting from the gait cycle propagate in the different scales down to the forces sensed by the integrins, modeled as springs at the chondrocyte's scale. Those scales provide input for the intracellular regulatory network model, which is represented by an additive mathematical framework and simulated with an asynchronous updating scheme. The output of the regulatory network provides insights on some molecular markers and ECM status. This may affect the ECM mechanical properties and material properties and, in turn, the forces computed in the upper scales, as represented by the hypothetical (dashed line) feedback arrow.

Two main situations were investigated numerically: the case of mechanical cues experienced by chondrocytes in a normal knee and the pathological case of a meniscectomy (i.e. the surgical removal of the meniscus). Potential differential effects were evaluated for chondrocytes located in the superficial zone of the cartilage and the ones located in the deeper zone (**Figure 1.3**). In addition, a virtual situation with application of pro-inflammatory cytokines was studied in order

to decompose *in silico* the contribution of individual factors (mechanical vs. inflammatory cues) and to better understand the previous observations on the meniscectomy case study (see *simulation results* section for more details).

FE modeling of the knee joint (tissue level): an FE model of the right knee of a 70 year old female weighing 77.1 kg (geometry obtained from the OpenKnee project (*423*)) was developed in FEBio (*424*). Hyperelastic material properties were assigned to femoral and tibial cartilage, menisci and the ligaments as obtained from literature. Joint rotations, translations and joint reaction forces for the stance phase of the gait were used as input boundary conditions for the knee joint. The forces experiences in the cartilage tissue were computed with the Abaqus software. Simulations were carried out for the healthy knee joint and a meniscectomized knee joint by removing the meniscus from the FE model (*370*).

FE modeling of the chondrocyte microenvironment (cell level): FE models of the chondrocyte and its microenvironment containing the pericellular matrix (PCM) and extracellular matrix (ECM) for the different zones of the cartilage (SZ, middle zone and DZ) were developed in Abaqus (Dassault Systèmes). The ECM and PCM were modeled as fibril reinforced poro-viscoelastic materials (*425*). The collagen fibrils in the PCM were oriented tangential to the cell, whereas the fibrils in ECM mimicked the fibril orientation in the corresponding zone of the cartilage. The cell was modeled as a biphasic material. Integrins, responsible for sensing of mechanical signals in chondrocytes, were modeled as linear springs between the cell and the PCM interface. The maximum compressive and shear strains obtained at the different zones of articular cartilage from simulations of the knee joint (tissue level) model were used as input boundary conditions for the cell-level model. Those input conditions were applied at the top of the cell level model and forces propagated through the ECM and PCM down to the springs representing integrins in a way that was determined by the material properties and fribril orientation. The forces computed on integrins were used as input for the intracellular level (*370*).

Regulatory networks for chondrocyte biochemical and mechanotransduction (intracellular): The same semi-quantitative additive formalism as in Chapter 4 and 5 was used to establish the mathematical equations. Simulations were run with priority classes for different speeds of interactions (genetic vs. protein) and with an asynchronous scheme for variable updating. It resulted in a similar type of stochastic model as presented in previous chapters.

Nevertheless, several adaptations were made in the strategy and algorithms used to run and analyse the model, as described below.

- *Mechanical input*: The intracellular module now includes a variable denoting the mechanical forces sensed by the cells. It requires an external input corresponding to the mechanical forces in the form of a continuous value in the interval [0-1], with 0 meaning no forces or mechanical stimulations, and 1 being the maximal force possibly sensed by mechanosensitive receptors (i.e. integrins and TRPV4 channels), in the system. This entry is obtained from the output of upstream layers (i.e. the FE chondrocyte environment model), with possibly differential values for the SZ and DZ chondrocytes. The integrin input for the DZ or the SZ amounts to the peak strength experienced by cells during the first gait cycle, as calculated from the FE model for the different layers. The output values from the FE model are normalized with the min-max method to fit in the interval [0,1] with the highest value being obtained in the SZ of the meniscectomized knee joint (**Table 6.1**).
- *Initial conditions for experiments:* The healthy chondrocyte profile of each zone, identified as such based on the level of standard biomarkers (e.g. SOX9, RUNX2, COL2, COL10, etc..), is used as an initial condition to apply experimental conditions.

6.4 *In silico* experiments to investigate the effect of perturbations related to meniscectomy

Two types of *in silico* experiments were executed with the regulatory network, a canalization analysis using a Monte Carlo approach to find the stable states of the updated network and a perturbation analysis to investigate the effects induced by meniscectomy. Details are provided below.

Constrained Monte Carlo analysis: The total state space was screened with a Monte Carlo analysis while imposing constraints on the mechanical loading input in order to identify eventual stable states for the intracellular model under predefined mechanical environments discussed above (cf. **Table 6.1**). Results reported are averages for 10000 random initialisations of the network (respecting the constraint on the loading input).

***In silico* state transition experiments with persistent perturbations:** Generally, chemical or medical treatments (e.g. drugs or external application of growth factors) follow bolus-like dynamics (burst release followed by a natural decay). Unlike these bolus treatments, the

changes in the mechanical or biochemical environment for cells located in a normal or meniscectomized knee joint are more likely to last and the cells must accommodate to that new environment. Hence, a major difference with the model depicted in Chapter 4 is that the external perturbations for the state transition study are now applied for an indefinite period of (computational) time steps and the new phenotypes are assessed after the system has settled in stable states accounting for the perturbed environment.

During the state transition study, the system could settle in various types of final states under the imposed environmental perturbations because of the stochastic nature of the model. Contrary to Chapter 4, the perturbations did not always trigger a transition to one of the basal stable states (i.e. one of the states arising from the Monte Carlo analysis with the physiological mechanical input constraints). Yet, multiple new types of final states could potentially be observed under the perturbed environment. Therefore, each perturbation simulation was repeated 100 times and a weighted average destination profile was computed. The weights reflected the number of occurrences of the destination, or final profiles, among the 100 repetitions. Weighted standard deviations were also evaluated.

	Healthy	**Meniscectomy**
Sup. zone	0.59	1
Deep zone	0.28	0.34

Table 6.1: Normalized mechanical input values. Input values for the deep zone and the superficial (Sup.) zone are obtained from simulations with the FE models. These correspond to the peak strain experienced by cells in the different layers during the first gait cycle. Values are normalized (min-max method) with respect to the maximum strain obtained in the superficial zone during the meniscectomy.

As readouts of the *in silico* experiments, we used several variables that can be summarized in three categories. The first corresponds to typical chondrocyte biomarkers (SOX9 and RUNX2) and ECM related proteins, such as type-II and type-X collagen, Aggrecans (a type of proteoglycans) and ECM degrading enzymes (i.e. MMP13 and ADAMTS5). They are important readouts because the ECM content and degradation level may be used as a feedback to inform the cell and tissue scales of the multiscale model. Indeed, ECM nature and content will affect its mechanical properties, which are important parameters for the FE models (**Figure 6.2** and **Figure 6.3**). The second group of variables is relevant to inform about the mechanotransduction pathways (YAP, Actin-MLC and TRPV4/Ca^{2+}), while the last category informed about inflammatory signaling (the cytokines receptors: R-inflam, as well as TAK1,

NFKB, iNOS and NO). The results for the complete set of variables (74 in total) is available in supplementary information **Figure S6.1.**

6.5 Simulation results

6.5.1 Chondrocyte location and mechanical signals define the nature of the model's stable states

The Monte Carlo analysis screening the system's state space, was run under different mechanical environments that corresponded to either the SZ or the DZ of a normal knee joint during a typical gait cycle. In both zones, two possible basal stable states were calculated from the model: a SOX9-positive (SOX9+), comparable to a non-hypertrophic chondrocyte, and a RUNX2-positive one (RUNX2+), comparable to a hypertrophic chondrocyte (**Figure 6.4**). However, we obtained slightly different canalization and profiles for the SZ and the DZ chondrocytes, as reported in **Figure 6.4**. We observed that 14% (resp. 56%) of the initial states led to the SOX9+ state and 86% (resp. 44%) to the RUNX2+ state for the SZ (resp. DZ) mechanical environment. When comparing the SZ with the DZ, the overall profiles for the normal and the hypertrophic-like states were slightly different due to the lower mechanical strains in the DZ. Especially for the non-hypertrophic state, most active entities and pathways were between 8 and 36% less abundant or active in the DZ than in the SZ, except for the IGF-I pathway as well as SOCS IKBA and FOXO that were slightly more abundant (about 20% more for IKBA).

For the same loading input (i.e. within a given zone), the integrins were more sensitive in the RUNX2+ state (integrins = 1 in the SZ, and 0.85 in the DZ) than the SOX9+ state (integrins = 0.39 in the SZ, and 0.17 in the DZ). However, the TRPV4/Ca^{2+} pathway was more activated by mechanical input in the SOX9+ state (TRPV4/Ca^{2+}= 0.28 in the SZ and 0.13 in the DZ), than in the RUNX2+ state where the activity was null.

Superficial zone

		HP	No HP
Signaling pathwaysathways	**WNT**	0.94	0.00
	BMP	0.92	0.00
	WNT antagonit	0.00	0.84
	BMP antagonist	0.00	0.97
	FGF	0.67	0.00
	IHH	1.00	0.42
	PTHRP	0.00	0.84
	IGF	0.00	0.67
	TFGB	0.07	0.84
	MAPK/ERK	0.88	0.14
TF	RUNX2	1.00	0.00
	SOX9	0.00	0.97
	NKX3.2	0.00	1.00
ECM	MMP13	1.00	0.00
	ADAMTS5	1.00	0.00
	Col-X	1.00	0.00
	Col-II	0.00	1.00
	Aggrecan	0.00	1.00
Mechanotransduction	Loading	0.59	0.59
	Integrins	1.00	0.39
	YAP/TAZ	0.00	0.67
	RhoA/Actin	1.00	0.67
	TRPV4/Ca2+	0.00	0.28
Inflammation related	SOCS	0.00	0.31
	IKB-A	0.00	0.33
	FOXO1	0.00	0.28
	Cytokines	1.00	0.12
	Inflammation	1.00	0.10
	CANALIZATION	**86%**	**14%**

Deep zone

		HP	No HP
Signaling pathwaysathways	**WNT**	0.88	0.00
	BMP	0.92	0.02
	WNT antagonit	0.00	0.74
	BMP antagonist	0.00	0.71
	FGF	0.67	0.00
	IHH	1.00	0.31
	PTHRP	0.00	0.69
	IGF	0.00	0.73
	TFGB	0.07	0.87
	MAPK/ERK	0.88	0.13
TF	RUNX2	1.00	0.00
	SOX9	0.00	0.71
	NKX3.2	0.00	0.76
ECM	MMP13	1.00	0.00
	ADAMTS5	1.00	0.00
	Col-X	1.00	0.00
	Col-II	0.00	1.00
	Aggrecan	0.00	0.95
Mechanotransduction	Loading	0.28	0.28
	Integrins	0.85	0.17
	YAP/TAZ	0.00	0.30
	RhoA/Actin	0.85	0.30
	TRPV4/Ca2+	0.00	0.13
Inflammation related	SOCS	0.00	0.34
	IKB-A	0.00	0.52
	FOXO1	0.00	0.35
	Cytokines	1.00	0.03
	Inflammation	0.94	0.04
	CANALIZATION	**44%**	**56%**

Label	Variables included in average values
WNT	(WNT, DSH, B-Cat, LLEF/TCF, Frizzled)
BMP	(BMP, R-SMADs,SMAD_complex
WNT antagonit	(DKK1, FRZB)
BMP antagonist	GREM1
FGF	(FGF, FGFR1, FGFR3)
IHH	(IGHH, GLI2)
PTHRP	(PTHrP, PKA, PPR, PP2A)
IGF	(IGFI, IGF-IR, [1-IGFBP])
TFGB	(TGF-B, SMAD3, [1-SMAD7])
MAPK/ERK	(ERK, RAS, P38, JNK)
RhoA/Actin	(Rhoa/ROCK, Actin-MLC)
TRPV4/Ca2+	(TRPV4, [Ca2+]i)
Inflammation	(R-inflam, TAK1, NFKB, iNOS, NO)

Figure 6.4: Computed basal profiles for the superficial zone and deep zone. Global activity of indicative variables or pathways in each stable state, for the SZ (left) and the DZ (right) zone. The 'Loading' variable corresponds to the imposed mechanical constraint. Bold labels indicate that the associated activity reflects the average value of several variables, as indicated in the table on the right. The first stable state is associated to a hypertrophic-like chondrocyte state (HP or RUNX2+) and the second one as a normal or non-hypertrophic one (no HP or SOX9+), based on the level of typical biological markers (RUNX2, SOX9, type II and type X collagen, Wnt pathway, etc.). The canalization indicates the percentage of initial states, among the 10000, leading to each final state during the Mont Carlo analysis.

6.5.2 *In silico* meniscectomy induces catabolic changes and inflammatory response in chondrocytes

To mimic the effect of a pathological mechanical signal (e.g. meniscectomy) on presumably healthy chondrocytes, we used the SOX9+ states from each zone as an initial condition and applied the perturbed mechanical inputs as described in **Table 6.1**. We looked at the average profile resulting from the perturbations over 100 repeated simulations. The increased mechanical input for SZ and DZ chondrocytes after meniscectomy over-activated and upregulated all mechanotransduction pathways (e.g. Actin-MLC, YAP/TAZ, TRPV4/Ca^{2+}) compared to the basal healthy situation (**Figure 6.5B**). It also led to 4 times more pro-inflammatory cytokines produced by SZ chondrocytes in meniscectomy compared to the healthy case. This small spike was associated with up-regulation/activation of pro-inflammatory mediators such as NFKB and iNOs, and NO production (**Figure 6.5C, SZ)**. As a

result, the model predicted a small induction of degrading enzymes (i.e. MMP13 and ADAMTS) in the SZ, after meniscectomy. It also predicted the initiation of ECM degradation as seen from type-II collagen and Aggrecan that went down to 0.9 and 0.8 respectively (std<0.01), in contrast to their maximum value in the SZ healthy state. Nevertheless, both SZ and DZ chondrocytes remained positive for SOX9 and negative of RUNX2 and COL-X, suggesting that the first iteration after meniscectomy induced some pro-inflammatory and catabolic changes, especially in the SZ, without triggering a full shift towards hypertrophy.

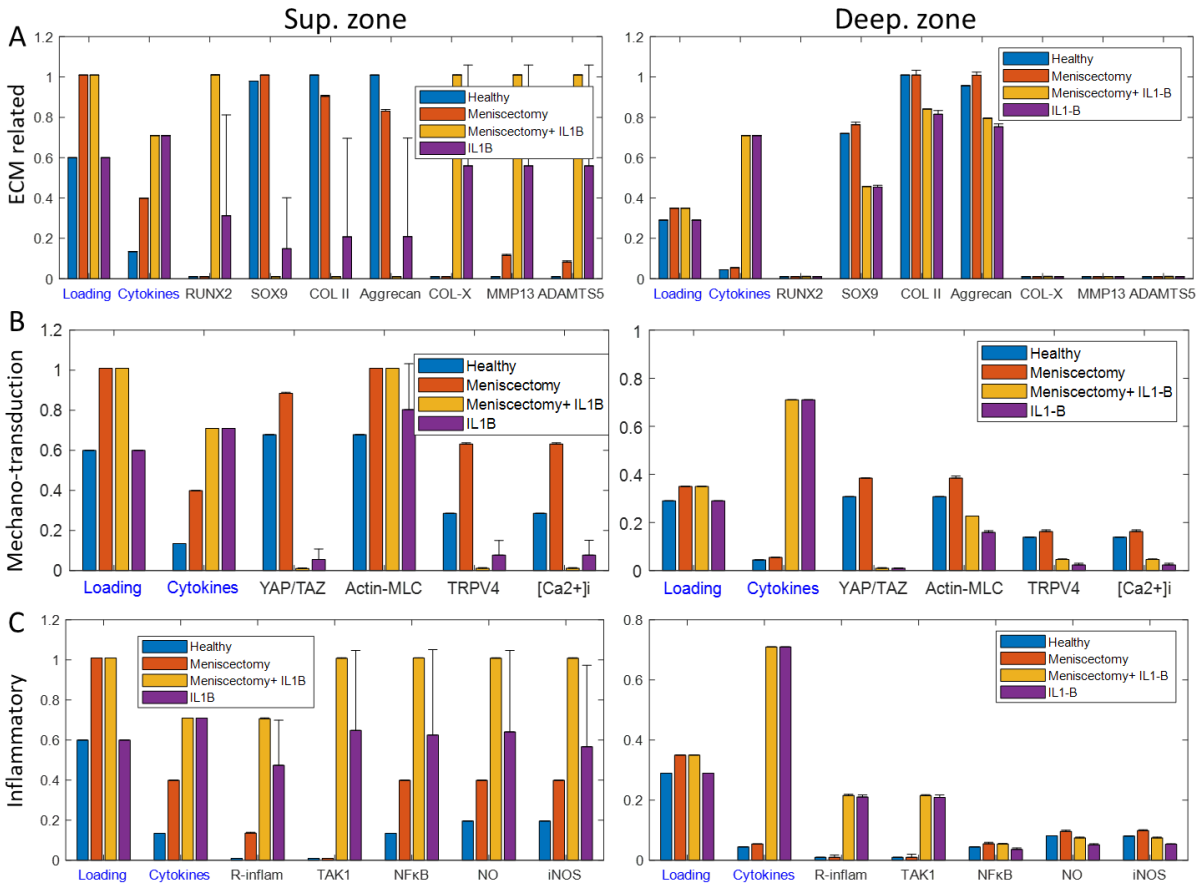


Figure 6.5: Effect of meniscectomy, addition pro-inflammatory cytokines and their combination on SZ and DZ chondrocyte phenotype. Addition of pro-inflammatory cytokines is denoted by +IL-1B, in the legend box. This 'Cytokines' input (+IL-1B) was assigned the value 0.7. The profiles of the initial SZ and DZ SOX9+ states, before perturbation, are reported under the 'Healthy' condition (blue bars), whereas other bars reflect average final profiles after external perturbation. The global activity of variables is reported for the superficial zone (left column) and the deep zone (right column); with bar height representing the average final value over 100 repeated simulations. Length of the upper error bars indicates standard deviation (std), no std is computed for the initial condition 'Healthy'. Variables labeled in blue, i.e. 'Loading' and 'Cytokines', reflect the perturbation input values due to the meniscectomy and/or the external application of pro-inflammatory cytokines. So, 'Loading' (resp. 'Cytokines') is artificially fixed in the meniscectomy and meniscectomy+IL-1B conditions (resp. IL-1B and meniscectomy+IL-1B conditions), hence std=0. Other variables are grouped into (A) master transcription factors and ECM related proteins, (B) mechanotransduction related entities, and (C) pro-inflammatory response related entities.

6.5.3 Abnormal mechanical signal amplifies phenotypic shift in case of pro-inflammatory cues build-up, in SZ chondrocytes.

The regulatory network model was built with the objective to integrate it within a multiscale framework, which implies that iterative updates should take place at the different scales to account for longer period. Based on the observed inception of the pro-inflammatory response and cytokines production during meniscectomy, we imagined a situation with a build-up of pro-inflammatory cytokines over time and repeated iterations. Therefore, we simulated a condition in which the pathological mechanical load was combined with additional cytokines (values arbitrary fixed to 0.7, i.e. about 6 times higher than in the basal situation). The experimental situation with cytokines build-up only, without changes in the forces' amplitude, was also investigated to evaluate the relative contribution of the different covariables and to better understand their interplay. The combination of pathological mechanical cues and cytokines in the SZ fully activated pro-inflammatory pathways (yellow bars in **Figure 6.5.C**) and triggered a phenotypic transition, as seen from the large increase of RUNX2, COL-X, MMP13 and ADAMTS5 and decrease of SOX9, COL-II and Aggrecan (**Figure 6.5A**, yellow bars). The sharp effect observed with the combination of both extra mechanical and inflammatory cues was not so distinct when adding pro-inflammatory cytokines only. The pro-inflammatory cytokines alone disturbed the healthy phenotype as observed from the level of the master transcription factors and ECM proteins (**Figure 6.5.A**, purple bars), but SZ cartilage still maintained some level of COL-II and Aggrecan. Most importantly, a much greater standard deviation was observed in the final states of SZ chondrocytes following perturbations with the IL-1B condition alone, compared to meniscectomy+IL-1B (**Figure 6.5. A,B &C**). The large standard deviation in the average profile was due to the variety of the final states possibly reached under that perturbation during the state transition study. Overall, those results suggested an interaction effect in which pathological mechanical cues sensitize SZ chondrocytes and make the catabolic response to pro-inflammatory cues more deterministic.

6.5.4 Pro-inflammatory signals determine the response of SZ chondrocytes to mechanical cues and the mechanotransduction paths.

Interestingly, YAP/TAZ and TRPV4/Ca^{2+} mechanotransduction pathways were shut down in the condition with meniscectomy+IL-1B, while the signal transduced through Integrins/Rho/ROCK/Actin-MLC remained highly active despite the phenotypic shift, in SZ

chondrocytes (**Figure 6.5B**, yellow bars). The mechanosensitive TRPV4 receptors and the mechano-effectors YAP/TAZ were activated under increased forces compared to the healthy situation, but they were largely down-regulated and/or inactivated when the increased forces were combined with increased pro-inflammatory cues (**Figure 6.5B, SZ**). Pro-inflammatory cues alone, already showed an ability to decrease the TRPV4 and YAP signal compared to the basal situation but not in such a deterministic way, as seen from the bigger standard deviations (**Figure 6.5B, SZ**, purple bars). Therefore, whilst increased forces alone strongly activated the YAP and TRPV4 signals, compared to the initial healthy state, those increased forces completed the decrease of those signals due to pro-inflammatory cues, in the SZ (**Figure 6.5B, SZ)**.

6.5.5 SZ and DZ chondrocytes react differentially to pathological mechanical and inflammatory signals

Most of the aforementioned effects following meniscectomy were observed in SZ and DZ chondrocytes even though the amplitudes were much lesser in the DZ. Indeed, the effect of increased mechanical loading in the DZ only displayed mild tendencies rather than bold behaviors since the forces sensed by mechanoreceptors were only increased by 6%, after the first iteration following meniscectomy. Beside the lighter effect, some differences are noticeable between the two zones. First, the increased forces alone slightly induced Aggrecan, contrary to the SZ (**Figure 6.5A**, DZ). This was certainly possible thanks to the increased mechanotransduction pathways activity together with the absence of ECM degrading enzymes production (i.e. MMP13 and ADAMTS5) (**Figure 6.5A & B, DZ)**. Although, some pro-inflammatory mediators were present but not greatly above the basal level (e.g. 0.07 vs. 0.09±0.003 for iNOS) (**Figure 6.5A & C, DZ)**.

Moreover, under this light mechanical perturbation, DZ chondrocytes appeared less sensitive to the pro-inflammatory cytokines uptake than SZ chondrocytes. Under the same amount of external pro-inflammatory cytokines (0.7) in SZ and DZ, the level of both cytokines receptors (R-inflam) and downstream effectors (TAK1) were increased up to 0.21±0.004 and 0.21±0.003 respectively in the DZ, whereas they reached a level of 0.7±0.004 and 1±0.003 respectively in the SZ (**Figure 6.5C**). In addition, the effect observed on ECM related proteins and inflammatory pathways in the meniscectomy+IL-1B condition seemed largely due to the +IL-1B covariable, in the DZ (**Figure 6.5 A & B).** Another difference regarded the mechanoregulatory pathway involving Actin-MLC, which was over-activated in both

meniscectomy and meniscectomy+IL-1B for the SZ, while it was over-activated by meniscectomy and inhibited by meniscectomy+IL-1B, in the DZ. This suggests that *in silico* chondrocytes react differently to inflammatory cues depending on the intensity of the mechanical cues. Overall, no sharp transition towards hypertrophy was observed in the DZ under the combination of mechanical cues and inflammatory cues, contrary to the SZ.

6.6 Discussion

We have developed a stochastic intracellular regulatory network module of articular chondrocytes to investigate the effect of mechanical and biochemical cues on chondrocyte phenotype. In particular, it allowed investigating the effect of pathophysiological mechanical cues as well as their interactions with biochemical cues, such as inflammation, on the chondrocyte state as well as the consequences on the surrounding ECM content. As such, it can serve as a biology-based regulatory network module within a biomechanical multiscale framework of articular cartilage. Previous models of chondrocytes with mechano-biochemical coupling mainly comprised phenomenological effects or had poor biology related mechanistic content (*257*, *359*, *375*). On the contrary, our *in silico* model enables the prediction of different pathophysiological conditions resulting from complex intracellular signaling and genetic networks. The model successfully recapitulated the loss of proteoglycans in the zone of higher mechanical stress, following meniscectomy. We also identified differential behaviors for SZ and DZ chondrocytes due to the difference in intensities of the mechanical stimuli sensed by the cells and we predicted non-linear interactions between mechanical and inflammatory cues.

As reported in the literature and in the description of the knowledge-based mechanisms at the start of this chapter, some of the mechanotransduction pathways are foreseen to be rather pro-anabolic while others could trigger both anabolic and catabolic responses depending on the status of other biochemical signals. Our results are consistent with that observation. For instance, elements of the mechanotransduction pathways such as the TRPV4 receptors, which are known to be important in healthy chondrocytes and are disrupted during OA (*402*), were indeed more active in the healthy SOX9+ than in the RUNX2+ naturally emerging states of both zones in our model. In addition, those mechanotransducers were over-activated under increased forces compared to the healthy situation, which reflected their mechanosensitivity, but, they were largely inactivated when the increased forces were combined with increased pro-catabolic inflammatory cues. On the other hand, signals transduced via the cytoskeleton (Actin-MLC) seemed more ambiguous. Our data for the SZ suggests that the mechano-response is

indeed non-linear in the current *in silico* model. The zonal difference in our model was mainly characterized by the difference of intensities of mechanical stimuli sensed by chondrocytes. While we reported an overall ECM catabolism due to increased forces in the SZ, these same increased forces in the DZ (resulting a lower mechanical stimulation compared to SZ) were not detrimental for the ECM. On the contrary, Aggrecan in the DZ was brought to its maximal amount with increased load while slightly less Aggrecan was initially available for the DZ healthy chondrocyte. This supports the experimental observation that increasing forces promote anabolism up to a certain level after which the catabolic effects overrule (*178*).

Even though the selection of the mechanisms to integrate in the model was based on literature and common knowledge in the field, they remain a non-exhaustive snapshot of real life regulatory network complexity. Nevertheless, our mechanistic modeling approach allowed taking those mechanisms in isolation from the rest of the system and testing their involvement in clinical outcomes and cartilage degeneration in a disrupted biomechanical environment. The model's prediction about the pathological mechanical cues decreasing Aggrecan corroborates with the observed proteoglycan loss following meniscectomy (*426*) or hyper-physiological load (*178, 257*). The fact that meniscectomy increases the risk of developing OA for patients (*377*) is reflected in our model with the catabolic changes following meniscectomy and the shift in chondrocyte phenotype towards hypertrophy in case of additional pro-inflammatory cytokines build-up, which are OA hallmarks. Clinical studies have also indicated that the risk of OA after meniscectomy depends on the pre-existing cartilage conditions (*427*), which is also supported by our *in silico* model since the status of other biochemical signals steered the mechanoregulatory response. Overall, the biochemical regulatory mechanisms and mechanotransduction pathways we inputted into the model allowed the recapitulation of biologically relevant behaviors, supporting the hypothesis of their implication in chondrocyte phenotype regulation and cartilage degeneration.

Despite the satisfying corroboration with experimental observations, important biological assumptions were made in order to simplify the mathematical model, leaving room for adaptations and improvement. Here we discuss a few such possibilities. An important improvement could be to consider different integrin types since various subunits have been identified and may trigger different types of signals. For example, when the cartilage matrix is damaged in OA, ECM fragments signal specifically through the α5β1 integrin to activate a pro-inflammatory and pro-catabolic response (*192, 376*). Moreover, a lot of data and studies have recently emerged regarding the involvement of the TRPV4 channels in chondrocyte osmo- and

mechano-sensing (*193, 401, 404, 428, 429*). Our model disregards important mechanisms related to these channels and the intracellular calcium as second messenger. We believe it could gain from further refinement and completion as knowledge grows (e.g. adding effect of osmotic pressure, downstream activation of NFAT nuclear factors). Especially, some studies suggest that the mechanosensitive properties of calcium channels might be dependent on interactions with integrins (*429, 430*). On the contrary, making abstraction of other pathways less directly related to the mechanoregulatory and inflammatory response (e.g. BMPs) could allow model reduction. This would help when implementing a temporal ODE-based version of the model, which is more parameter dependent. Such a time dependent model could be useful to account for the oscillating nature of the message encoded by Ca^{2+} for which the amplitude and frequency vary according to the cartilage location (DZ, SZ) and with OA severity (*406*) and that could not be represented with the current additive mathematical framework.

We acknowledge that some modeling tasks remain in order to strengthen the model credibility (in the regulatory sense of the word). First, sensitivity of the model to the input conditions and to the few parameters should be assessed. For instance, some interactions comprised a subjective weight for calibrating the basal profiles and canalizations, and the pro-inflammatory cytokines input (+IL-1B) was arbitrary set to 0.7. The latter hypothetical condition was applied with a milder amount of cytokines to analyze and understand the behavior of the modeled system, but no one-on-one relationship was established with an actually inflammatory cytokine concentration. Higher input values made the IL-1B condition overrule other cofactors, which hindered the observation of any subtle interaction effects in the SZ (see supporting information **Figure S6.1**). Finally, the predictive power of the *in silico* experiments remains to be formally validated with (*de novo*) experimental data. This would require relating *in silico* variables qualitatively and accurately to the most relevant experimental measures and should be done on a case-by-case basis. For example, the numerical 'global activity' for Aggrecan reflects the net result of Aggrecan synthesis/production and degradation by proteases. The slow (genetic) sub-variable would be more comparable to genetic expression or, by simplification, to protein synthesis, which are assumed to be correlated in the current model. The fast sub-variable (protein level) accounts for any function-modulating post-translational modifications, which are absent for Aggrecan in our model, or degrading reactions.

One of the perspectives offered by intracellular regulatory network models as the one presented here is to integrate it within a multiscale workflow. For now, the min-max normalization attributed the forces with the value 1 in the superficial cartilage with meniscectomy. It amounts

to the maximum force theoretically sensed by the mechano-sensors of the intracellular module during the gait cycle simulation. Data informing on the actual maximal forces leading to some saturation of integrins assembly and activation could help refine and calibrate that normalized input. On the other side, the output regarding ECM changes could be used as feedback to inform the upper scales about modifications of the mechanical properties (through production and degradation of ECM), allowing for a fully closed modeling loop that can capture long-term effects (weeks to years) through successive iterations. The external pro-inflammatory cytokines applied in the current study provided a glimpse of what could happen in case cytokines are allowed to build up over several iterations. Nevertheless, relating the semi-quantitative changes of the ECM to quantitative mechanical properties remains challenging and thorough calibration will be required. This will require a combination of *in vitro* set-ups mimicking a specific element or scale (e.g. (*431*)) and *in vivo* set-ups allowing to capture the overall effects (e.g. (*402*)). Eventually, this would help to assess how pathological force distribution and instability occurring in the knee can negatively affect chondrocyte molecular response and the overall joint, in return.

Overall, the intracellular module discussed in this chapter provides a mechanistic mechano-biochemical coupling and opens routes for understanding how crosstalk between mechanotransduction and signaling molecules regulates cellular phenotypic response. The ongoing integration within a multi-scale workflow will help with the optimization of new therapeutic strategies by providing suggestions about both mechanical adaptations and pharmacological targeting therapies.

Chapter 7. Discussion and conclusion

Each chapter of this thesis consists of an independent scientific study and includes its own detailed discussion. In this final chapter, we provide a general summary of the presented PhD research. Then, we formulate the main conclusions of the thesis and contributions to the scientific field. Finally, we discuss the perspectives that this PhD thesis opens for future research.

7.1 Project summary

The research carried on in this thesis lies at the interface of engineering and biomedical sciences. Despite increasing knowledge and data available on cartilage and bone biology and diseases, as described in **Chapter 1**, there are still no effective therapeutic solutions for regenerating or replacing bone, or for preventing cartilage breakdown in OA. In this thesis, *in silico* technologies focusing on bioregulatory networks are used to identify potential therapeutic strategies for both cartilage and bone disorders, through endogenous (pharmacological) or exogenous (tissue engineering) repair. We proposed an integrative strategy, combining computational and experimental approaches (**Chapter 2**) in order to integrate complex knowledge and data about chondrocyte regulatory signaling, inflammation and/or mechanotransduction into mathematical models and in order to identify key molecular signatures and phenotypic drivers. Those drivers of cell fate decision are considered potential targets for drug treatments or for improving cell differentiation. Since intracellular bioregulatory networks may be automatically reconstructed thanks to (transcript)omics data and mathematical tools, we first employed this strategy to identify typical OA GRN signatures via the development of a machine-learning (ML) model (**Chapter 3**). However, this type of approach disregards prior knowledge accumulated during decades of research in the field. Hence, we also integrated prior knowledge into mechanistic models to gain better interpretability and predictive power on non-intuitive emerging outcomes (**Chapter 4, 5, 6**). We combined those data- and knowledge-driven approaches to get the best of both worlds for two case studies. On the one hand, we built a knowledge-based articular chondrocyte model that we informed with a few data-inferred regulatory interactions to predict potential therapeutic targets in OA (**Chapter 4**). On the other hand, the topology of a knowledge-based regulatory network was optimized to fit reference data. We re-used a chondrocyte network model, previously optimized for the context of endochondral ossification, in order to predict the best differentiation biochemical cocktail for iPSc-derived chondrocytes, with the aim to enhance bone forming in TE (**Chapter 5**). Finally, the model developed in Chapter 4 was developed further in **Chapter 6** by introducing mechanoregulatory mechanisms with the aim to investigate

their interaction with inflammatory signals and to provide an intracellular bio-regulatory module for a multiscale biomechanical representation of the human knee joint.

The aim of **Chapter 3** was to perform a meta-analysis of high-throughput data in order to identify typical OA signatures and quantitatively evaluate cell states in mouse disease models, thanks to machine learning (ML) methods. The goal of this model was to provide a quantitative metric for the evaluation of *in vivo* OA models and the assessment of candidate treatments based on transcriptional signatures. The first part of that chapter describes a transcriptomic data (pre-)processing pipeline meant to integrate heterogeneous data sources in a single dataset, thereby reconstituting a multi-perturbed informative dataset for training and evaluating the model. Since establishment of GRNs governs cell states, an OA-specific GRN has been reconstructed and supervised ML models (i.e. classifiers) have been trained to classify samples based on those typical signatures, through the CellNet method. In this chapter, the performance of the ML classifiers is evaluated via cross-validation and the model's ability to recognize and score the establishment of OA is validated with a new independent dataset.

The transcriptomic dataset pre-processed in Chapter 3 is also used in **Chapter 4** to infer new transcriptional interactions with mathematical methods and to feed a mechanistic, knowledge-derived model of the articular chondrocyte. The aim of that chapter was to identify drivers of chondrocyte hypertrophy and targeting strategies against OA thanks to a mathematical mechanistic representation of the biological system. Taking a prior growth plate chondrocyte model as a basis, an extensive literature study has been carried out to extend and build the regulatory model of the articular chondrocyte. This chapter not only provides an interactive knowledge base, through CellCollective, on the multiple mechanisms pertaining to chondrocyte homeostasis and pathophysiology, but the resulting executable model also provides an *in silico* system for testing hypotheses through *in silico* experiments. Since validating numerical predictions with *de novo* wet experiments is an important step to ensure the model's credibility, different chondrocyte *in vitro* culture systems have been leveraged to validate the *in silico* findings and confirm the potential of newly identified pharmacological targeting strategies against chondrocyte hypertrophy, and therefore against OA.

A second application case study for the regulatory network modeling approach is detailed in **Chapter 5.** As mentioned above, the context is that of IPSCs-derived chondrocyte that are pushed towards hypertrophy to promote bone formation through endochondral ossification after implantation *in vivo*. The chapter explores the capacity of *in silico* regulatory network models

to inform the design of cell differentiation experiments and foster the shift from exploratory to confirmatory experimental trials. In particular, a parallel experimental study by collaborators has investigated various biochemical differentiation cocktails to optimize a protocol of bone formation via chondrocyte hypertrophy. We reproduced those experimental conditions *in silico* to rank them from the most to the least potent, and explore additional conditions. This was done by leveraging a previous model of the chondrocyte in endochondral ossification, composed of a mechanistic backbone network of which the topology and interaction weights were further optimized with growth plate expression data.

Since mechanical cues are translated into biochemical signals for cellular response, it is clear that integrating this type of biological regulatory layer into multiphysics biomechanical models offers powerful perspectives. This is investigated in **Chapter 6** where the model of Chapter 4 is expanded to account for mechanotransduction pathways and their crosstalk with inflammatory signals. The algorithms have been adapted to provide a module that not only offers the possibility to respond to external mechanical inputs and provide outputs leading to ECM property changes, but also allows exploring different scenarios in a systematic way and test mechanistic hypotheses. With that module, conclusions were drawn on the mechanisms underlying the threshold effect observed between mechanical and inflammatory cues on articular chondrocytes and cartilage breakdown. The results of that chapter as well as previous chapters led us to define some suggestions for future work and strategy for multiscale integration.

7.2 Main conclusions and contributions

Complex knowledge and data have been integrated into computer models to assist in answering biomedically relevant questions related to chondrocyte transcriptional and phenotypical changes, such as hypertrophy, during OA as well as during endochondral ossification. Several important conclusions can be drawn out of the main results, as follows.

In **chapter 3**, omics data fetched from literature about *in vivo* mouse OA models was assembled and integrated to carry out a meta-analysis. This provided new biological insights on early-OA transcriptional signatures and led to the construction of a predictive tool. Those signatures were biological meaningful as seen from the relevant mutant phenotype annotations (related to growth, skeleton, osteoarthritis or inflammation) of the top ranked biological entities included in the computationally reconstructed GRN signatures (e.g. *Tgfbi, Timp1, Cd44, Errfi*). The data-inferred signatures also suggested transcriptional regulators potentially important for the

control of transcriptional programs in the *in vivo* OA models, such as *Hif-1a, Elk3* and *Runx1*. The relevance of the newly assembled dataset and of the reconstructed OA transcriptional signatures was also evidenced by the development of a machine learning model whose training was based on those GRN signatures and whose predictive power was demonstrated on an independent dataset from literature. Indeed, the ML model scored the OA-resemblance level of the new samples in a way that was in line with the original conclusions from the statistics-based study (*273*), thereby confirming the existence of phasic changes during OA progression that was elsewhere reported through experiments and in clinics (*1*, *432*, *433*). Beside the implications for cartilage biology, this study demonstrated the power of reusing and confronting existing data, which supports the need of the ongoing efforts on the FAIR principles to ensure accessibility and reusability of data in research (*282*).

The CellNet method that we employed to build the aforementioned ML-based quantitative metrics and evaluate disease models has been repeatedly applied to various cell types in order to guide stem cell differentiation during the past years (*239*, *240*, *242*). While previously the method has successfully measured the conversion of fibroblasts to induced hepatocytes or provided hypotheses for improving the process of B-cell to macrophage conversion, it never before has been applied to skeletal cell types, to our knowledge. Similarly, it is the first time this network biology-based concept that is CellNet is translated into a disease case study, so our study is a first attempt to open the scope of possibilities for this approach. There are several clinical implications of the ML model we developed, in the context of the osteoarthritis pathology. First, given the multitude of existing *in vitro* and *in vivo* models to study OA and evaluate candidate treatments (*296*), it is valuable to develop quantitative metrics as well as a reference to evaluate and compare those models and the effect of treatments, including at molecular and transcriptional level. The strategy and the model developed in **Chapter 3** provide a path towards systematic quantitative evaluation of established or newly developed *in vitro* or *in vivo* OA models. In addition, one of the reasons candidate disease modifying drugs do not proceed to the market is the great inter-patient variability causing variable responses and disparate drug efficacies (*42*, *300*, *433*). Hence, several studies have highlighted the necessity for patient stratification and for the characterization of patient subgroups, for instance through relating histopathological changes with clear molecular signatures (*42*, *300*, *302*, *434*). Supervised machine learning strategies, such as the one we employed, offer the possibility to precisely distinguish and classify samples types based on predefined molecular features (*240*, *241*, *302*). Although, our study initially depicted a binary situation with and without OA, we

demonstrated that the current version of the model was trained to identify rather early-stage mouse OA signatures, opening perspective for further development and for introducing more granularity.

Beside the data-driven approach, the other models developed or leveraged in this thesis were at least partially knowledge-based, and therefore mechanistic. They allowed us to increase the holistic understanding of chondrocyte phenotypic changes during OA or endochondral ossification under various cues, including pro-inflammatory signals and mechanical stimuli. The step of literature curation to establish the backbone networks in itself already increases our understanding of the process, and pinpoints at gaps in the current knowledge (*200*). Indeed, 'content curation' is the act of finding, grouping, organizing and analyzing relevant content on a specific issue to help make sense of this information instead of adding more content and noise to the overload of information (*435*, *436*). Hence, formalizing that information with mathematical rules helps to structure and make sense of that information while highlighting uncertainties and contradictory information. As such, Aldridge et al. qualified such signaling pathway models as "dynamic repositories of interpretable knowledge" (*200*). The mathematical modeling strategy we used strongly relied on the mathematical framework developed by Kerkhofs et al. (*254*, *255*). Whist it does not allow for precise time-continuous and quantitative representation of the system, as do ODE-based models, this semi-quantitative additive framework proved itself very powerful to integrate large amounts of signaling pathways and mechanisms while limiting the amount of unknown parameters. Hence, it enables the holistic view of chondrocyte phenotype regulatory control and it spares the modeler from parameter identifiability issues, set aside the case of interaction weights explored in Chapter 5. The current PhD research valorizes and illustrates the potential of that additive mechanistic modeling strategy by expanding and applying the framework to new research questions (Chapter 4, 5 and 6), by curating, referencing and systematically annotating and sharing the modeled networks (Chapter 4 and 6), and by initiating formal validation with dedicated experiments or retrospective data (Chapter 4 and 5). In the end, the current articular chondrocyte network models join a series of dynamic repositories about cartilage and chondrocyte (*93*, *250–252*, *256*, *258*, *332*, *359*), which altogether extend our understanding of biochemical signalings that dictate chondrocyte fate decision.

To gain understating some intuitive reasoning can be made from the regulatory network topologies and activity flows, but solely numerical simulations formally demonstrate the actual emerging cellular states due to signal propagation and the tight balances controlling the system

(e.g. maintenance of chondrocyte homeostasis or its shift). The advantage of those mechanistic models is that they allow tracing back to what mechanisms could possibly explain the obtained results (*200*). That exercise that can be aided by additional *in silico* targeting experiments (*326*, *327*), which are often less time consuming and costly than their experimental counterparts. For example, in **Chapter 4** we described that TGF-β normally plays a protective role against inflammation in cartilage and is known to involve promiscuous ligand-receptor interactions with seemingly opposing downstream cascades. Yet, the balance between the different receptors was found to be altered with ageing and during OA (*86–88*), leading to the hypothesis that the change in the balance between ALK1 and ALK5 receptors may explain why TGF-β loses its protective role against inflammation-induced degradation. Our *in silico* simulations revealed, in line with previous (numerical) studies (*332*), that increasing the abundance of ALK1 over ALK5 was indeed detrimental to the protective role of TGF-β. Further *in silico* knock-out and activation experiments could help decipher which pathways are key to drive this observation downstream of ALKs, depending on the research question. Any type of similar scenarios about other pathways could be virtually explored in the same way. This illustrates how the numerical system we provide may assist cartilage biologists in their investigations.

Besides the increased understanding that arose from the models developed and analyzed in this PhD work, Chapter 4 provided actionable outcomes to foster drug targets identification for blocking chondrocyte hypertrophy in articular cartilage, which is an important driver of OA in some patients (*319, 437*). Other authors have shown that hypertrophic changes are associated with increased cell-volume as well as transcriptional and metabolic (catabolic) changes that negatively affect and mineralize the surrounding ECM on the long run. More particularly, the shift in activity and expression of master transcriptional regulators, including RUNX2 and SOX9, is characteristics of the hypertrophic transition. Hence the idea that pharmacologically targeting key molecular drivers could be a potent strategy to reduce hypertrophy and counter OA. Our results have highlighted the robustness of chondrocyte states to external perturbations, and suggest that targeting more than one factor is the most efficient way to control the numerous crosstalks taking place in chondrocytes. The simulations have shown that maintaining activation of the protein kinase A or SMAD3 protein alone or in combination with a second entity could be anti-hypertrophic. The model revealed that the combinatorial activation of PKA with inhibition of BMP or FGFR1 is more constraining and results in a more deterministic outcome than the respective individual treatments. In addition, the synergistic effect observed with the PKA activator and FGFR1 inhibitor was confirmed with dedicated *in vitro*

experiments. To our knowledge, this the first time an *in silico* or experimental study pinpoints this specific combination. Although, the idea to activate PKA or inhibiti FGFR1 are not counter-intuitive. Indeed, PKA is related to the known anti-hypertrophic PTHrP pathway (*97*) and would mediate the TGF-β/ALK5 signaling-induced proteoglycan 4 expression (*91*), while FGFR1 is a receptor for which inhibitory peptides were recently shown to attenuate articular cartilage degeneration in mice (*126*). Nevertheless, hundreds of *a priori* similarly meaningful combinations among the 60 biological actors represented in our model could be considered virtually. Testing them all experimentally would lead to a combinatorial explosion, and selecting them by intuition would possibly hinder *a priori* less obvious conditions. Whereas *in silico* technology allowed here, as well as in other equivalent studies (*326*), to efficiently and unbiasedly screen most, if not all, possibilities. This way the potential of intuitive and less intuitive conditions can be mathematically demonstrated. Therefore, the data of this PhD thesis contributes to refining the way therapeutic target discovery is currently approached and offers a new track of investigation for OA therapies.

A limiting factor with this type of mechanistic models is that the deep literature curation it requires remains fastidious and the choice to integrate one specific mechanisms or pathway over another can still be subjective. Therefore, the current state of the art in the field of regulatory networks modeling, including for cartilage biology, is evolving towards more automatic approaches such as text-mining, fetching of pathway databases (e.g. Reactome (*438*) , PathwayCommons (*439*), OmniPath (*440*)) and data-driven optimization of mathematical models (*196*, *248*, *327*). For instance, Melas et al. have optimized a signaling pathway model of key players in cartilage homeostasis and inflammation with phosphoproteomics data through interger linear programming (*252*). This method minimizes an objective function derived from linear constraints to limit the mismatch between predicted and experimental activation profiles. Similarly, Kerkhofs et al. have established a scoring system to fit the semi-quantitative regulatory network model of chondrocyte hypertrophy onto qualitative profiles derived from expression data (*196*, *265*). The resulting objective function served in a genetic algorithm to optimize the interaction weights, and hence the topology of the prior knowledge network. This method generates families of models (ensemble models) that are all compatible with prior knowledge, but that are divergent in structure. **Chapter 5** demonstrates the reproducibility of that method and validates the predictive power of the resulting ensemble model by applying it to a new case study related to bone tissue engineering. By reproducing *in silico* the experimental setup of Long et al. for IPSc-derived chondrocyte hypertrophy and bone formation, the model

successfully ranked the relative potential of the experimental conditions (*441*). It confirms, in line with the results of the independent experimental study, that the combinatorial application of BMP4 ligands, a WNT agonist and T3 hormone constrains chondrocytes to follow the hypertrophic path more than any of the individual treatments. In addition, the model supports the hypothesis that activation of pro-inflammatory pathways on the top of the aforementioned treatment increases further the chances of reaching the hypertrophic state, which is in line with the higher bone forming potential observed *in vivo* for that condition. Here, we exploited the variation across the ensemble as a measure of robustness for the treatment effects and as a way to prioritize the experimental conditions. Nevertheless, the model we used is optimized with generic and manually derived mouse growth plate expression profiles in absence of clearly defined transient versus hypertrophic chondrocyte multi-omics profiles. Alternatively, other authors have exploited this type of optimization of interaction weights to derive cell type- or cancer patient-specific models, starting from a common prior knowledge network and training the model with patient specific tumor data (*327*), thereby interpreting the variations in simulation outcome as inter-individual variability. This was possible thanks to the generation of dedicated high throughput multi-perturbed proteomic data through a microfluidic platform (*327*, *442*). Hence, the availability of such large-scale perturbation data for skeletal cells will unleash possibilities for *in silico* technologies to assist the design of experiments in bone tissue engineering or drug target identification for OA.

Most regulatory network models consider the system in isolation from the supra-cellular structures, although the surrounding biochemical environment can be imposed as external constraints. This may be an acceptable assumption for biochemical signals since cells are likely more affected by local, paracrine and autocrine signaling molecules. However, local mechanical cues are the results of forces propagating through the different structure and tissue layers and the cellular response is likely to affect the surrounding mechanical properties (e.g. pericellular matrix) and subsequently the nature of the cues (*422*, *443*). Taking into account those forces is important in the context of cartilage given the load-bearing function of this organ and the mechano-sensitive nature of chondrocytes. The catabolic or anabolic role that chondrocytes endorse in homeostasis and disease is dictated by the integration of mechanical cues, biochemical signals and their interplay (*173*, *178*, *359*). Consequently, increasing biomechanical modeling efforts are focusing on relating mechanical input with biochemical and cellular response (*258*, *261*, *444*), including through the integration of intracellular regulatory layers (*375*). The *in silico* data of this PhD thesis demonstrate differential responses

for chondrocytes depending on their physical location in the cartilage, and hence on their initial biochemical status and intensity of perceived forces. Our data also underpins conclusions about the influence of the inflammatory status on chondrocyte response to mechanical cues and the involvement of integrins, actin cytoskeleton, TRPV4 calcium channels as well as the TAK1/NFkB axis and nitrite oxide signaling in defining the ECM protein secretory profile and the anabolic/catabolic response to mechanical stimulation, roles that were elsewhere supported experimentally (*420*, *428*, *430*, *445*). As such, the current mechanobiological model stands as an intracellular layer for future multi-scale integration but more calibration efforts are required to accurately relate the ECM biochemical profile obtained from the model with tangible mechanical properties. With this work, we are aiming towards a model that would be the first for the knee joint to relate the tissue, cellular and intracellular scale with a high biochemical mechanisms content.

Overall, a common thread in the different chapters is to leverage *in silico* technologies as tools to complement, refine and inform experimental studies meant at finding therapeutic strategies against cartilage degradation or for bone repair, through focusing on intracellular regulatory networks. That capacity was hypothesized at the beginning of this thesis and several of our results support that hypothesis. The ML model provides a quantitative metric to evaluate *in vivo* OA models through a scoring system measuring GRN signatures and comparing disease states. It is meant as tool that can be queried at will to guide experiments by suggesting intervention points to improve the disease models or evaluate the efficiency of any treatment at transcriptional level. The *in silico* models established in Chapter 4 and Chapter 6 are like a virtual articular chondrocytes to test pharmacological targeting strategies against hypertrophy and ECM destruction in OA and to evaluate the influence of both individual factors and combinatorial effects. It offers a platform for systematically screening all possible treatments with low cost and time resources as compared to experimental studies, thanks to parallel and high performance computing. Eventually, those models may come into a funnel strategy to help decision making based on mathematical evidence before proceeding to experimental confirmation. A similar concept can be applied for improving protocols for bone tissue engineering through endochondral ossification.

In conclusion, this PhD thesis established computational tools able to predict outcomes about chondrocyte hypertrophy and cartilage breakdown that would be otherwise difficult to predict or demonstrate. Furthermore, it has contributed to the increased scientific understanding of

biochemical and mechanoregulatory intracellular pathways controlling chondrocyte hypertrophy and/or ECM state, through computational and experimental analyses.

The different studies presented here open several routes for future computational and experimental analyses in order to place their outputs as strong foundations for future therapeutic strategies. Therefore, the perspectives of this thesis are further discussed in the next section.

7.3 Perspectives and future research

In this section, we make different suggestions to extend the impact of the current research and increase its relevance to applications in osteoarthritis or bone TE.

A machine learning model has been developed to quantitatively assess transcriptional programs in *in vivo* OA models and measure the impact of candidate treatments. This model and the underlying reference signatures could gain from additional application cases. For instance, additional literature derived or newly produced data about a candidate treatment applied in a DMM mouse model, preferably at an early time point after surgery, could be classified and scored with the model to confront the molecular signatures and assess the impact of the treatment. In addition, a binary picture of the disease is depicted in the current model's implementation since it was trained to recognize the presence or absence of OA with an apparent specificity for earlier stages of OA. As indicated by this result, OA evolves progressively. Hence, possible improvement for this model could be to train it with data including several levels of OA severity or types of OA, each level or type being assigned its own signature and classifier. Extending this concept to the clinical setting, human OA can be stratified with clinical phenotypes or scoring such as MANKIN or OARSI scores. An equivalent classifier model could be developed to assign each score or each clinical phenotype with its own reference transcriptional signatures and compare new samples to those references. Although the variance between for instance the MANKIN scores 2 and 3, at transcriptional level, might be lower than between any OA versus healthy patient, which might make the machine learning process less performant. Using RNA-seq or scRNA-seq data, which are generally less noisy, may be a way to alleviate this challenge.

Additionally, the CellNet platform was initially designed to evaluate stem cell differentiation level quantitatively and to improve and guide differentiation protocols. Nevertheless, no skeletal cell types are represented in current CellNet reference profiles and signatures (*241*). This opens perspective of future research to develop a quantitative quality control system for

developmental bone tissue engineering. Indeed, if the endochondral ossification process is successfully reproduced, then experimentalists should expect their experimental samples to have an increasing resemblance score for the hypertrophic chondrocyte signature and then for the one of bone cells while the score for resemblance with the initial stem cell signature should decrease during the differentiation process. Such model could relatively easily be validated with transcriptional profiles of bone TE constructs at different point of the process, such as the developmentally engineered callus organoids of Nilsson Hall et al. (*5*).

From the mechanistic perspective, the more information and knowledge about chondrocyte regulatory pathways will be available in literature and through structured public databases (e.g. KEGG pathways, OmniPath, etc..), the easier it will be to establish prior knowledge networks and the more robust will become the models of chondrocyte and cartilage. Several mechanisms could be refined and complemented in the mechanistic models of this PhD thesis via incorporation of more or different cascades and interactions, as discussed in the different thesis chapters. For example, the comprehension of the TRPV4 channel's implication in cartilage mechanotransduction has rapidly evolved in the last couple of years (*404*, *407*, *428*, *430*) and the model of **Chapter 6** could serve as a basis to confirm and explore some of the proposed mechanisms. Another example would be to expand the FGF related pathways given their seemingly important role for chondrocyte hypertrophy (through FGFR1) in the model described in **Chapter 4**. This would also be valuable given the lack of representation of the different FGF ligands with their receptor specificity in the current model, for instance FGF18 would preferably bind to FGFR3 receptors while FGF2 would bind more specifically to FGFR1, and the FGFR1/FGFR3 balance was shown to change during OA. However, clear information about the expression profiles and transcriptional regulators of the different ligands are necessary for that endeavor. Finally, integration of apoptotic related pathways and dedifferentiation pathways could be a way to account for other chondrocyte fates that are currently overlooked in the models. Our hypothesis is that doing so may help to reduce the size of the spurious 'None' attractor, present in the models of **Chapter 4** and **5**, in which settle initial states and resulting trajectories for which the mathematical rules cannot easily be met. Overall, as the information content grows, the interactive network or knowledge base we implemented in the Cell Collective platform should be updated. Indeed, it is a first step towards systematic annotations and the curation of a chondrocyte molecular map about OA. The next step could be to improve the compliance with current standard representations (e.g. SBML, SBGN etc..) developed by

the systems biology community to increase sharing and reusability perspectives (*198*, *446*, *447*), as was recently done for a rheumatoid arthritis disease map (*202*).

The knowledge-based completion of the network may turn out to be an endless task, and we believe that training the current articular chondrocyte model on relevant data (e.g. proteomic and transcriptomic profiles of healthy vs. hypertrophic chondrocytes) is one of the most interesting perspectives to improve its significance for cartilage biology. Various engineering tools are available to do so. To continue with the same mathematical additive formalism, one could head towards network optimization and ensemble modeling similarly to what was done for the growth plate model used in Chapter 5 and other studies that tackled the uncertainty carried by scarce experimental data (*448*). In that case, qualitative or even quantitative (although, scaled between 0 and 1) profiles should be derived from real experimental profiles and literature data. This model fitting task would be aided by the gene ID annotations we provided for each variable in the Cell Collective platform, allowing easier correspondence between the model's variables and actual genes or protein products represented in experimental datasets. Alternatively, the current networks could simply serve as static prior knowledge networks to derive a new type of dynamic mathematical model, for instance with logic ODEs (cf. CellNOpt package (*222*, *327*)), which maintain the advantage of limited unknown kinetic parameters while allowing a time continuous description of variables. In addition, such a continuous and time-scaled version of the mechanotransduction model from **Chapter 6** could be a good option to facilitate its calibration (i.e. relation between mechanical stimuli and molecular response) with time series data following mechanical load or with dynamical compression studies (*178*).

In this thesis, we searched for important drivers of phenotypic transition and tipping points within bioregulatory networks through mathematical implementation and computer simulations. Systems biologists have also explored alternative approaches to find important drivers based on characteristics of the network topology alone. For instance, the dynamics of a GRN in the ascidian embryo has been successfully controlled for cell fate specification of seven tissues by manipulating the activity of five molecules that were identified from the network structure alone (*237*). Those alternative approaches, do not limit themselves to looking at hub genes and node connectivity but they also borrow methodologies belonging to complex systems and structural network control theories (*235*), as briefly introduced in **Chapter 1**. A network is considered controllable if one can drive the state of the target nodes to any desired final state in finite time (*235*). Barabási et al. have proved they could identify a minimum set of driver nodes,

sufficient to control (part of) the system thanks to a several algorithms, including the "k-walk" algorithm for one 1-control input or a "greed algorithm" (graph theory) for >1-control inputs (*236*). This type of approaches showed promising results for the identification of a small number of key control nodes, within real world networks (*236*, *237*). Therefore, the topology of the networks developed in **Chapter 4, 5** and **6** could be exploited further with these alternative approaches to contribute to the objectives of the current thesis regarding the identifications of important drivers of hypertrophic changes for promoting bone formation in TE or limiting cartilage degradation in OA.

Independently of the method to find the molecular drivers, an important way to ensure the relevance of a mathematical model for therapeutic and experimental indications is to validate it, i.e. corroborate the model's outputs with actual biological observations. First, an *in silico* target screening similar to the one operated on the model of **Chapter 4** could be applied to the updated and extended mechanobiological model of **Chapter 6** to compare the impact of the originally predicted anti-hypertrophic conditions and eventually point towards other potent combinations. Some of the anti-hypertrophic combinations predicted by our model in **Chapter 4**, including the synergy between PKA activation and FGFR1 inhibition, were confirmed with *in vitro* experiments. Nevertheless, this study calls for further validation. On the one hand, additional predicted conditions from **Chapter 4** (e.g. PKA activation with BMP inhibition) or treatments newly predicted with the model of **Chapter 6** could be further tested experimentally. On the other hand, validating the beneficial impact of the aforementioned synergistic condition on chondrocyte hypertrophy, but also on OA progression, with an *in vivo* mouse disease model would reinforce the credibility of our approach - both in a regulatory sense of the word and in terms of convincing experimental researchers.

In summary, mimicking the integrated computational and experimental studies carried out in this thesis, future research in this area should have both a numerical and an experimental component. This will (continue to) establish the systems biology approach in general, and our modeling framework in particular, as a pillar for understanding cartilage biology and identifying targeted treatments for OA or bone TE.

References

1. J. Martel-Pelletier, A. J. Barr, F. M. Cicuttini, P. G. Conaghan, C. Cooper, M. B. Goldring, S. R. Goldring, G. Jones, A. J. Teichtahl, J. P. Pelletier, Osteoarthritis. *Nat. Rev. Dis. Prim.* **2** (2016), pp. 1–18.

2. M. Hiligsmann, C. Cooper, N. Arden, M. Boers, J. C. Branco, M. Luisa Brandi, O. Bruyère, F. Guillemin, M. C. Hochberg, D. J. Hunter, J. A. Kanis, T. K. Kvien, A. Laslop, J. P. Pelletier, D. Pinto, S. Reiter-Niesert, R. Rizzoli, L. C. Rovati, J. L. H. Severens, S. Silverman, Y. Tsouderos, P. Tugwell, J. Y. Reginster, Health economics in the field of osteoarthritis: An Expert's consensus paper from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). *Semin. Arthritis Rheum.* **43**, 303–313 (2013).

3. M. A. Karsdal, M. Michaelis, C. Ladel, A. S. Siebuhr, A. R. Bihlet, J. R. Andersen, H. Guehring, C. Christiansen, A. C. Bay-Jensen, V. B. Kraus, Disease-modifying treatments for osteoarthritis (DMOADs) of the knee and hip: lessons learned from failures and opportunities for the future. *Osteoarthr. Cartil.* **24**, 2013–2021 (2016).

4. I. Papantoniou, G. Nilsson Hall, N. Loverdou, R. Lesage, T. Herpelinck, L. Mendes, L. Geris, Turning Nature's own processes into design strategies for living bone implant biomanufacturing: a decade of Developmental Engineering. *Adv. Drug Deliv. Rev.* **169** (2021), pp. 22–39.

5. G. Nilsson Hall, L. F. Mendes, C. Gklava, L. Geris, F. P. Luyten, I. Papantoniou, Developmentally Engineered Callus Organoid Bioassemblies Exhibit Predictive In Vivo Long Bone Healing. *Adv. Sci.* **7**, 1902295 (2020).

6. E. Charlier, C. Deroyer, F. Ciregia, O. Malaise, S. Neuville, Z. Plener, M. Malaise, D. de Seny, Chondrocyte dedifferentiation and osteoarthritis (OA). *Biochem. Pharmacol.* **165** (2019), pp. 49–65.

7. R. Dreier, Hypertrophic differentiation of chondrocytes in osteoarthritis: The developmental aspect of degenerative joint disorders. *Arthritis Res. Ther.* **12** (2010), p.

216.

8. E. V. Tchetina, Developmental Mechanisms in Articular Cartilage Degradation in Osteoarthritis. *Arthritis*. **2011:68397**, 1–16 (2011).

9. E. G. J. Ripmeester, U. T. Timur, M. M. J. Caron, T. J. M. Welting, Recent Insights into the Contribution of the Changing Hypertrophic Chondrocyte Phenotype in the Development and Progression of Osteoarthritis. *Front. Bioeng. Biotechnol.* **6**, 18 (2018).

10. H. Kitano, Biological robustness. *Nat. Rev. Genet. 2004 511*. **5**, 826–837 (2004).

11. A. F. Villaverde, J. R. Banga, Reverse engineering and identification in systems biology: strategies, perspectives and challenges. *J. R. Soc. Interface*. **11**, 20130505–20130505 (2013).

12. U. Alon, Biological networks: The tinkerer as an engineer. *Science (80-.).* **301** (2003), pp. 1866–1867.

13. Anaconda, 8 AI Predictions for 2020: Business Leaders & Researchers Weigh Ino Title, (available at https://www.anaconda.com/blog/8-ai-ml-predictions-for-2020).

14. F. Long, D. M. Ornitz, Development of the endochondral skeleton. *Cold Spring Harb. Perspect. Biol.* **5:a008334.** (2013), doi:10.1101/cshperspect.a008334.

15. K. L. Cooper, S. Oh, Y. Sung, R. R. Dasari, M. W. Kirschner, C. J. Tabin, Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions. *Nature* (2013), doi:10.1038/nature11940.

16. E. Kozhemyakina, A. B. Lassar, E. Zelzer, A pathway to bone: Signaling molecules and transcription factors involved in chondrocyte development and maturation. *Dev.* **142**, 817–831 (2015).

17. R. Nishimura, K. Hata, E. Nakamura, T. Murakami, Y. Takahata, Transcriptional network systems in cartilage development and disease. *Histochem. Cell Biol.* **149** (2018), pp. 353–363.

18. K. A. Staines, A. S. Pollard, I. M. McGonnell, C. Farquharson, A. A. Pitsillides, Cartilage to bone transitions in health and disease. *J. Endocrinol.* **219**, R1–R12 (2013).

19. M. B. Goldring, M. Otero, D. A. Plumb, C. Dragomir, M. Favero, K. El Hachem, K. Hashimoto, H. I. Roach, E. Olivotto, R. M. Borzì, K. B. Marcu, Roles of inflammatory and anabolic cytokines in cartilage metabolism: Signals and multiple effectors converge upon MMP-13 regulation in osteoarthritis. *Eur. Cells Mater.* **21**, 202–220 (2011).

20. E. G. J. Ripmeester, U. T. Timur, M. M. J. Caron, T. J. M. Welting, Recent Insights into the Contribution of the Changing Hypertrophic Chondrocyte Phenotype in the Development and Progression of Osteoarthritis. *Front. Bioeng. Biotechnol.* **6**, 18 (2018).

21. F. Guilak, R. J. Nims, A. Dicks, C. L. Wu, I. Meulenbelt, Osteoarthritis as a disease of the cartilage pericellular matrix. *Matrix Biol.* **71**–**72** (2018), pp. 40–50.

22. D. R. Eyre, M. A. Weis, J. J. Wu, Advances in collagen cross-link analysis. *Methods*. **45**, 65–74 (2008).

23. C. Kiani, L. Chen, Y. J. Wu, A. J. Yee, B. B. Yang, Structure and function of aggrecan. *Cell Res.* **12** (2002), pp. 19–32.

24. K. Sauerland, A. H. K. Plaas, R. X. Raiss, J. Steinmeyer, The sulfation pattern of chondroitin sulfate from articular cartilage explants in response to mechanical loading. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1638**, 241–248 (2003).

25. P. Alberton, H. C. Dugonitsch, B. Hartmann, P. Li, Z. Farkas, M. M. Saller, H. Clausen-Schaumann, A. Aszodi, Aggrecan hypomorphism compromises articular cartilage biomechanical properties and is associated with increased incidence of spontaneous osteoarthritis. *Int. J. Mol. Sci.* **20** (2019), doi:10.3390/ijms20051008.

26. J. Dudhia, Aggrecan, aging and assembly in articular cartilage. *Cell. Mol. Life Sci.* **62** (2005), pp. 2241–2256.

27. M. B. Goldring, J. M. Dayer, S. R. Goldring, in *Osteoimmunology: Interactions of the Immune and Skeletal Systems: Second Edition* (Academic Press, 2016), pp. 257–269.

28. G. Victoria, B. Petrisor, B. Drew, D. Dick, in *Indian Journal of Orthopaedics* (Indian Orthopaedic Association, 2009; /pmc/articles/PMC2762251/), vol. 43, pp. 117–120.

29. A. Vortkamp, S. Pathi, G. M. Peretti, E. M. Caruso, D. J. Zaleske, C. J. Tabin, Recapitulation of signals regulating embryonic bone formation during postnatal growth

and in fracture repair. *Mech. Dev.* **71**, 65–76 (1998).

30. P. Lenas, M. J. Moos, F. P. Luyten, Developmental Engineering: A new paradigm for the design and manufacturing of cell based products. Part I: From three-dimensional cell growth to biomimetics of in vivo development. *Tissue Eng. Part B. Rev.* **15**, 381–394 (2009).

31. D. J. Hunter, L. March, M. Chew, Osteoarthritis in 2020 and beyond: a Lancet Commission. *Lancet* (2020), , doi:10.1016/S0140-6736(20)32230-3.

32. G. Aubourg, S. J. Rice, P. Bruce-Wootton, J. Loughlin, Genetics of osteoarthritis. *Osteoarthr. Cartil.* (2021), doi:10.1016/j.joca.2021.03.002.

33. I. Sulzbacher, Osteoarthritis: Histology and pathogenesis. *Wiener Medizinische Wochenschrift.* **163**, 212–219 (2013).

34. E. S. E. Mehana, A. F. Khafaga, S. S. El-Blehi, The role of matrix metalloproteinases in osteoarthritis pathogenesis: An updated review. *Life Sci.* **234** (2019), p. 116786.

35. P. Aghajanian, S. Mohan, The art of building bone: Emerging role of chondrocyte-to-osteoblast transdifferentiation in endochondral ossification. *Bone Res.* **6**, 1–9 (2018).

36. T. M. Griffin, F. Guilak, The role of mechanical loading in the onset and progression of osteoarthritis. *Exerc. Sport Sci. Rev.* **33**, 195–200 (2005).

37. European Medicine Agency (EMA), ChondroCelect | European Medicines Agency (2017), (available at https://www.ema.europa.eu/en/medicines/human/EPAR/chondrocelect).

38. M. L. Sennett, J. M. Friedman, B. S. Ashley, B. D. Stoeckl, J. M. Patel, M. Alini, M. Cucchiarini, D. Eglin, H. Madry, A. Mata, C. Semino, M. J. Stoddart, B. Johnstone, F. T. Moutos, B. T. Estes, F. Guilak, R. L. Mauck, G. R. Dodge, Long term outcomes of biomaterial -mediated repair of focal cartilage defects in a large animal model. *Eur. Cells Mater.* **41**, 40–51 (2021).

39. J. M. Hulme, V. Welch, R. de Bie, M. Judd, P. Tugwell, in *Cochrane Database of Systematic Reviews* (Cochrane Database Syst Rev, 2002; https://pubmed.ncbi.nlm.nih.gov/11869668/).

40. M. C. Hochberg, A. Guermazi, H. Guehring, A. Aydemir, S. Wax, P. Fleuranceau-Morel, A. Reinstrup Bihlet, I. Byrjalsen, J. Ragnar Andersen, F. Eckstein, Effect of Intra-Articular Sprifermin vs Placebo on Femorotibial Joint Cartilage Thickness in Patients with Osteoarthritis: The FORWARD Randomized Clinical Trial. *JAMA - J. Am. Med. Assoc.* **322**, 1360–1370 (2019).

41. L. S. Lohmander, S. Hellot, D. Dreher, E. F. W. Krantz, D. S. Kruger, A. Guermazi, F. Eckstein, Intraarticular sprifermin (recombinant human fibroblast growth factor 18) in knee osteoarthritis: A randomized, double-blind, placebo-controlled trial. *Arthritis Rheumatol.* **66**, 1820–1831 (2014).

42. F. P. Luyten, S. Bierma-Zeinstra, F. Dell'Accio, V. B. Kraus, K. Nakata, I. Sekiya, N. K. Arden, L. S. Lohmander, Toward classification criteria for early osteoarthritis of the knee. *Semin. Arthritis Rheum.* **47**, 457–463 (2018).

43. Bone4kids, (available at https://www.bone4kids.be/en).

44. A. Carlier, H. Brems, J. M. A. Ashbourn, I. Nica, E. Legius, L. Geris, Capturing the wide variety of impaired fracture healing phenotypes in Neurofibromatosis Type 1 with eight key factors: A computational study. *Sci. Rep.* **7**, 1–15 (2016).

45. E. M. Thompson, A. Matsiko, E. Farrell, D. J. Kelly, F. J. O'Brien, Recapitulating endochondral ossification: A promising route to in vivo bone regeneration. *J. Tissue Eng. Regen. Med.* **9** (2015), pp. 889–902.

46. P. Lenas, M. Moos, F. P. Luyten, Developmental engineering: a new paradigm for the design and manufacturing of cell-based products. Part II: from genes to networks: tissue engineering from the viewpoint of systems biology and network science. *Tissue Eng. Part B. Rev.* **15**, 395–422 (2009).

47. S. K. O'Connor, D. B. Katz, S. J. Oswald, L. Groneck, F. Guilak, Formation of Osteochondral Organoids from Murine Induced Pluripotent Stem Cells. *Tissue Eng. Part A* (2020), doi:10.1089/ten.tea.2020.0273.

48. M. Zhang, J. Shi, M. Xie, J. Wen, K. Niibe, X. Zhang, J. Luo, R. Yan, Z. Zhang, H. Egusa, X. Jiang, Recapitulation of cartilage/bone formation using iPSCs via biomimetic 3D rotary culture approach for developmental engineering. *Biomaterials*. **260**, 120334

(2020).

49. X. Guo, K. K. Mak, M. M. Taketo, Y. Yang, The Wnt/β-catenin pathway interacts differentially with PTHrP signaling to control chondrocyte hypertrophy and final maturation. *PLoS One*. **4**, e6067 (2009).

50. S. Monteagudo, R. J. Lories, Cushioning the cartilage: A canonical Wnt restricting matter. *Nat. Rev. Rheumatol.* **13**, 670–681 (2017).

51. T. F. Day, X. Guo, L. Garrett-Beal, Y. Yang, Wnt/β-catenin signaling in mesenchymal progenitors controls osteoblast and chondrocyte differentiation during vertebrate skeletogenesis. *Dev. Cell*. **8**, 739–750 (2005).

52. F. Xuan, F. Yano, D. Mori, R. Chijimatsu, Y. Maenohara, H. Nakamoto, Y. Mori, Y. Makii, T. Oichi, M. M. Taketo, H. Hojo, S. Ohba, U. Il Chung, S. Tanaka, T. Saito, Wnt/β-catenin signaling contributes to articular cartilage homeostasis through lubricin induction in the superficial zone. *Arthritis Res. Ther.* **21**, 247 (2019).

53. Y. Usami, A. T. Gunawardena, M. Iwamoto, M. Enomoto-Iwamoto, Wnt signaling in cartilage development and diseases: Lessons from animal studies. *Lab. Investig.* **96**, 186–196 (2016).

54. F. P. Luyten, P. Tylzanowski, R. J. Lories, Wnt signaling and osteoarthritis. *Bone*. **44** (2009), pp. 522–527.

55. S. Pérez-García, M. Carrión, R. Villanueva-Romero, T. Hermida-Gómez, M. Fernández-Moreno, M. Mellado, F. J. Blanco, Y. Juarranz, R. P. Gomariz, Wnt and RUNX2 mediate cartilage breakdown by osteoarthritis synovial fibroblast-derived ADAMTS-7 and -12. *J. Cell. Mol. Med.* **23**, 3974–3983 (2019).

56. B. W. Doble, J. R. Woodgett, GSK-3: Tricks of the trade for a multi-tasking kinase. *J. Cell Sci.* **116** (2003), pp. 1175–1186.

57. J. L. Stamos, W. I. Weis, The β-catenin destruction complex. *Cold Spring Harb. Perspect. Biol.* **5** (2013), p. a007898.

58. Y. F. Dong, D. Y. Soung, E. M. Schwarz, R. J. O'Keefe, H. Drissi, Wnt induction of chondrocyte hypertrophy through the Runx2 transcription factor. *J. Cell. Physiol.* **208**,

77–86 (2006).

59. J. C. H. Leijten, S. D. Bos, E. B. M. Landman, N. Georgi, H. Jahr, I. Meulenbelt, J. N. Post, C. A. van Blitterswijk, M. Karperien, GREM1, FRZB and DKK1 mRNA levels correlate with osteoarthritis and are regulated by osteoarthritis-associated factors. *Arthritis Res. Ther.* **15**, R126 (2013).

60. J. C. H. Leijten, J. Emons, C. Sticht, S. Van Gool, E. Decker, A. Uitterlinden, G. Rappold, A. Hofman, F. Rivadeneira, S. Scherjon, J. M. Wit, J. Van Meurs, C. A. Van Blitterswijk, M. Karperien, Gremlin 1, frizzled-related protein, and dkk-1 are key regulators of human articular cartilage homeostasis. *Arthritis Rheum.* **64**, 3302–3312 (2012).

61. E. G. J. Ripmeester, U. T. Timur, M. M. J. Caron, T. J. M. Welting, Recent insights into the contribution of the changing hypertrophic chondrocyte phenotype in the development and progression of osteoarthritis. *Front. Bioeng. Biotechnol.* **6** (2018), p. 18.

62. N. Thielen, P. van der Kraan, A. van Caam, TGFβ/BMP Signaling Pathway in Cartilage Homeostasis. *Cells*. **8**, 969 (2019).

63. P. M. van der Kraan, E. N. Blaney Davidson, W. B. van den Berg, Bone Morphogenetic Proteins and articular cartilage. To serve and protect or a wolf in sheep clothing's? *Osteoarthr. Cartil.* **18** (2010), pp. 735–741.

64. T. Nakase, T. Miyaji, T. Tomita, M. Kaneko, K. Kuriyama, A. Myoui, K. Sugamoto, T. Ochi, H. Yoshikawa, Localization of bone morphogenetic protein-2 in human osteoarthritic cartilage and osteophyte. *Osteoarthr. Cartil.* **11**, 278–284 (2003).

65. R. N. Wang, J. Green, Z. Wang, Y. Deng, M. Qiao, M. Peabody, Q. Zhang, J. Ye, Z. Yan, S. Denduluri, O. Idowu, M. Li, C. Shen, A. Hu, R. C. Haydon, R. Kang, J. Mok, M. J. Lee, H. L. Luu, L. L. Shi, Bone Morphogenetic Protein (BMP) signaling in development and human diseases. *Genes Dis.* **1** (2014), pp. 87–105.

66. M. M. J. Caron, P. J. Emans, A. Cremers, D. A. M. Surtel, M. M. E. Coolsen, L. W. van Rhijn, T. J. M. Welting, Hypertrophic differentiation during chondrogenic differentiation of progenitor cells is stimulated by BMP-2 but suppressed by BMP-7. *Osteoarthr. Cartil.* **21**, 604–613 (2013).

67. V. Dexheimer, J. Gabler, K. Bomans, T. Sims, G. Omlor, W. Richter, Differential expression of TGF-β superfamily members and role of Smad1/5/9-signalling in chondral versus endochondral chondrocyte differentiation. *Sci. Rep.* **6**, 1–14 (2016).

68. A. Nohe, E. Keating, P. Knaus, N. O. Petersen, Signal transduction of bone morphogenetic protein receptors. *Cell. Signal.* **16** (2004), pp. 291–299.

69. M. B. Goldring, K. Tsuchimochi, K. Ijiri, The control of chondrogenesis. *J. Cell. Biochem.* **97**, 33–44 (2006).

70. M. Wu, G. Chen, Y. P. Li, TGF-β and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. *Bone Res.* **4**, 1–21 (2016).

71. R. Derynck, Y. E. Zhang, Smad-dependent and Smad-independent pathways in TGF-β family signalling. *Nature*. **425** (2003), pp. 577–584.

72. E. J. Mackie, Y. A. Ahmed, L. Tatarczuch, K. S. Chen, M. Mirams, Endochondral ossification: How cartilage is converted into bone in the developing skeleton. *Int. J. Biochem. Cell Biol.* **40**, 46–62 (2008).

73. M. A. Arnold, Y. Kim, M. P. Czubryt, D. Phan, J. McAnally, X. Qi, J. M. Shelton, J. A. Richardson, R. Bassel-Duby, E. N. Olson, MEF2C Transcription Factor Controls Chondrocyte Hypertrophy and Bone Development. *Dev. Cell.* **12**, 377–389 (2007).

74. M. H. Lee, Y. J. Kim, W. J. Yoon, J. I. Kim, B. G. Kim, Y. S. Hwang, J. M. Wozney, X. Z. Chi, S. C. Bae, K. Y. Choi, J. Y. Cho, J. Y. Choi, H. M. Ryoo, Dlx5 specifically regulates Runx2 type II expression by binding to homeodomain-response elements in the Runx2 distal promoter. *J. Biol. Chem.* **280**, 35579–35587 (2005).

75. M. H. Lee, Y. J. Kim, H. J. Kim, H. D. Park, A. R. Kang, H. M. Kyung, J. H. Sung, J. M. Wozney, H. J. Kim, H. M. Ryoo, BMP-2-induced Runx2 Expression Is Mediated by Dlx5, and TGF-β1 Opposes the BMP-2-induced Osteoblast Differentiation by Suppression of Dlx5 Expression. *J. Biol. Chem.* **278**, 34387–34394 (2003).

76. H. M. Ryoo, M. H. Lee, Y. J. Kim, Critical molecular switches involved in BMP-2-induced osteogenic differentiation of mesenchymal cells. *Gene*. **366** (2006), pp. 51–57.

77. K. Shirakabe, K. Terasawa, K. Miyama, H. Shibuya, E. Nishida, Regulation of the

activity of the transcription factor Runx2 by two homeobox proteins, Msx2 and Dlx5. *Genes to Cells*. **6**, 851–856 (2001).

78. Y. lin Li, Z. sheng Xiao, Advances in Runx2 regulation and its isoforms. *Med. Hypotheses*. **68**, 169–175 (2007).

79. A. M. Ionescu, E. M. Schwarz, M. J. Zuscik, H. Drissi, J. E. Puzas, R. N. Rosier, R. J. O'Keefe, ATF-2 cooperates with Smad3 to mediate TGF-β effects on chondrocyte maturation. *Exp. Cell Res.* **288**, 198–207 (2003).

80. K. Kirsch, A. Zeke, O. Tőke, P. Sok, A. Sethi, A. Sebő, G. S. Kumar, P. Egri, Á. L. Póti, P. Gooley, W. Peti, I. Bento, A. Alexa, A. Reményi, Co-regulation of the transcription controlling ATF2 phosphoswitch by JNK and p38. *Nat. Commun.* **11**, 1–15 (2020).

81. E. Leah, Osteoarthritis: TGF-β overload at bones of cartilage degeneration. *Nat. Rev. Rheumatol.* **9** (2013), p. 382.

82. E. N. Blaney Davidson, P. M. van der Kraan, W. B. van den Berg, TGF-β and osteoarthritis. *Osteoarthr. Cartil.* **15** (2007), pp. 597–604.

83. A. van Caam, W. Madej, A. Garcia de Vinuesa, M. J. Goumans, P. ten Dijke, E. Blaney Davidson, P. van der Kraan, TGFβ1-induced SMAD2/3 and SMAD1/5 phosphorylation are both ALK5-kinase-dependent in primary chondrocytes and mediated by TAK1 kinase activity. *Arthritis Res. Ther.* **19**, 112 (2017).

84. K. W. Finnson, W. L. Parker, P. Ten Dijke, M. Thorikay, A. Philip, ALK1 opposes ALK5/Smad3 signaling and expression of extracellular matrix components in human chondrocytes. *J. Bone Miner. Res.* **23**, 896–906 (2008).

85. P. M. Van Der Kraan, M. J. Goumans, E. Blaney Davidson, P. Ten Dijke, Age-dependent alteration of TGF-β signalling in osteoarthritis. *Cell Tissue Res.* **347** (2012), pp. 257–265.

86. P. M. van der Kraan, Age-related alterations in TGF beta signaling as a causal factor of cartilage degeneration in osteoarthritis. *Biomed. Mater. Eng.* **24**, 75–80 (2014).

87. E. N. Blaney Davidson, D. F. G. Remst, E. L. Vitters, H. M. van Beuningen, A. B. Blom, M.-J. Goumans, W. B. van den Berg, P. M. van der Kraan, Increase in ALK1/ALK5

Ratio as a Cause for Elevated MMP-13 Expression in Osteoarthritis in Humans and Mice. *J. Immunol.* **182**, 7937–7945 (2009).

88. P. M. Van Der Kraan, The changing role of TGFβ in healthy, ageing and osteoarthritic joints. *Nat. Rev. Rheumatol.* **13**, 155–163 (2017).

89. T. F. Li, L. Gao, T. J. Sheu, E. R. Sampson, L. M. Flick, Y. T. Konttinen, D. Chen, E. M. Schwarz, M. J. Zuscik, J. H. Jonason, R. J. O'Keefe, Aberrant hypertrophy in Smad3-deficient murine chondrocytes is rescued by restoring transforming growth factor β-activated kinase 1/ activating transcription factor 2 signaling: A potential clinical implication for osteoarthritis. *Arthritis Rheum.* **62**, 2359–2369 (2010).

90. J. H. Jonason, G. Xiao, M. Zhang, L. Xing, D. Chen, Post-translational regulation of Runx2 in bone and cartilage. *J. Dent. Res.* **88** (2009), pp. 693–703.

91. Q. Wang, Q. Y. Tan, W. Xu, H. B. Qi, D. Chen, S. Zhou, Z. H. Ni, L. Kuang, J. Y. Guo, J. L. Huang, X. X. Wang, Z. Q. Wang, N. Su, L. Chen, B. Chen, W. L. Jiang, Y. Gao, H. G. Chen, X. L. Du, Y. L. Xie, L. Chen, Cartilage-specific deletion of Alk5 gene results in a progressive osteoarthritis-like phenotype in mice. *Osteoarthr. Cartil.* **25**, 1868–1879 (2017).

92. X. Yang, L. Chen, X. Xu, C. Li, C. Huang, C. X. Deng, TGF-β/Smad3 signals repress chondrocyte hypertrophic differentiation and are required for maintaining articular cartilage. *J. Cell Biol.* **153**, 35–46 (2001).

93. W. Hui, D. A. Young, A. D. Rowan, X. Xu, T. E. Cawston, C. J. Proctor, Oxidative changes and signalling pathways are pivotal in initiating age-related changes in articular cartilage. *Ann. Rheum. Dis.* **75**, 449–458 (2016).

94. Y. Sano, J. Harada, S. Tashiro, R. Gotoh-Mandeville, T. Maekawa, S. Ishii, ATF-2 is a common nuclear target of smad and TAK1 pathways in transforming growth factor-β signaling. *J. Biol. Chem.* **274**, 8949–8957 (1999).

95. P. Dy, W. Wang, P. Bhattaram, Q. Wang, L. Wang, R. T. Ballock, V. Lefebvre, Sox9 Directs Hypertrophic Maturation and Blocks Osteoblast Differentiation of Growth Plate Chondrocytes. *Dev. Cell.* **22**, 597–609 (2012).

96. R. B. Vega, K. Matsuda, J. Oh, A. C. Barbosa, X. Yang, E. Meadows, J. McAnally, C. Pomajzl, J. M. Shelton, J. A. Richardson, G. Karsenty, E. N. Olson, Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell*. **119**, 555–566 (2004).

97. E. Kozhemyakina, T. Cohen, T.-P. Yao, A. B. Lassar, Parathyroid Hormone-Related Peptide Represses Chondrocyte Hypertrophy through a Protein Phosphatase 2A/Histone Deacetylase 4/MEF2 Pathway. *Mol. Cell. Biol.* **29**, 5751–5762 (2009).

98. B. E. Bobick, W. M. Kulyk, Regulation of cartilage formation and maturation by mitogen-activated protein kinase signaling. *Birth Defects Res. Part C - Embryo Today Rev.* **84** (2008), pp. 131–154.

99. L. A. Stanton, S. Sabari, A. V. Sampaio, T. M. Underhill, F. Beier, p38 MAP kinase signalling is required for hypertrophic chondrocyte differentiation. *Biochem. J.* **378**, 53–62 (2004).

100. E. Kozhemyakina, A. B. Lassar, E. Zelzer, A pathway to bone: signaling molecules and transcription factors involved in chondrocyte development and maturation. *Development*. **142**, 817–831 (2015).

101. V. Y. L. Leung, B. Gao, K. K. H. Leung, I. G. Melhado, S. L. Wynn, T. Y. K. Au, N. W. F. Dung, J. Y. B. Lau, A. C. Y. Mak, D. Chan, K. S. E. Cheah, SOX9 Governs Differentiation Stage-Specific Gene Expression in Growth Plate Chondrocytes via Direct Concomitant Transactivation and Repression. *PLoS Genet.* **7**, e1002356 (2011).

102. V. Lefebvre, M. Dvir-Ginzberg, SOX9 and the many facets of its regulation in the chondrocyte lineage. *Connect. Tissue Res.* **58** (2017), pp. 2–14.

103. A. Cheng, P. G. Genever, SOX9 determines RUNX2 transactivity by directing intracellular degradation. *J. Bone Miner. Res.* **25**, 2680–2689 (2010).

104. A. Augello, C. De Bari, The regulation of differentiation in mesenchymal stem cells. *Hum. Gene Ther.* **21** (2010), pp. 1226–1238.

105. S. Koelling, J. Kruegel, M. Irmer, J. R. Path, B. Sadowski, X. Miro, N. Miosge, Migratory Chondrogenic Progenitor Cells from Repair Tissue during the Later Stages of Human Osteoarthritis. *Cell Stem Cell*. **4**, 324–335 (2009).

106. L. Liao, S. Zhang, J. Gu, T. Takarada, Y. Yoneda, J. Huang, L. Zhao, C. Do Oh, J. Li, B. Wang, M. Wang, D. Chen, Deletion of Runx2 in Articular Chondrocytes Decelerates the Progression of DMM-Induced Osteoarthritis in Adult Mice. *Sci. Rep.* **7** (2017), doi:10.1038/s41598-017-02490-w.

107. T. Ikeda, S. Kamekura, A. Mabuchi, I. Kou, S. Seki, T. Takato, K. Nakamura, H. Kawaguchi, S. Ikegawa, U. Chung, The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. *Arthritis Rheum.* **50**, 3561–3573 (2004).

108. P. Smits, P. Li, J. Mandel, Z. Zhang, J. M. Deng, R. R. Behringer, B. De Crombrugghe, V. Lefebvre, The Transcription Factors L-Sox5 and Sox6 Are Essential for Cartilage Formation. *Dev. Cell*. **1**, 277–290 (2001).

109. V. Lefebvre, in *Current Topics in Developmental Biology* (Academic Press Inc., 2019), vol. 133, pp. 171–193.

110. V. Lefebvre, M. Angelozzi, A. Haseeb, SOX9 in cartilage development and disease. *Curr. Opin. Cell Biol.* **61** (2019), pp. 39–47.

111. A. Haseeb, R. Kc, M. Angelozzi, C. de Charleroy, D. Rux, R. J. Tower, L. Yao, R. Pellegrino da Silva, M. Pacifici, L. Qin, V. Lefebvre, SOX9 keeps growth plates and articular cartilage healthy by inhibiting chondrocyte dedifferentiation/osteoblastic redifferentiation. *Proc. Natl. Acad. Sci.* **118**, e2019152118 (2021).

112. J. D. Kormish, D. Sinner, A. M. Zorn, Interactions between SOX factors and Wnt/β-catenin signaling in development and disease. *Dev. Dyn.* **239** (2010), pp. 56–68.

113. H. Akiyama, J. P. Lyons, Y. Mori-Akiyama, X. Yang, R. Zhang, Z. Zhang, J. M. Deng, M. M. Taketo, T. Nakamura, R. R. Behringer, P. D. McCrea, B. De Crombrugghe, Interactions between Sox9 and β-catenin control chondrocyte differentiation. *Genes Dev.* **18**, 1072–1087 (2004).

114. Y. Kawato, M. Hirao, K. Ebina, K. Shi, J. Hashimoto, Y. Honjo, H. Yoshikawa, A. Myoui, Nkx3.2 Promotes Primary Chondrogenic Differentiation by Upregulating Col2a1 Transcription. *PLoS One*. **7**, e34703 (2012).

115. H. Zhang, G. Hu, H. Wang, P. Sciavolino, N. Iler, M. M. Shen, C. Abate-Shen, Heterodimerization of Msx and Dlx homeoproteins results in functional antagonism. *Mol. Cell. Biol.* **17**, 2920–2932 (1997).

116. D. Ferrari, L. Sumoy, J. Gannon, H. Sun, A. M. C. Brown, W. B. Upholt, R. A. Kosher, The expression pattern of the Distal-less homeobox-containing gene Dlx-5 in the developing chick limb bud suggests its involvement in apical ectodermal ridge activity, pattern formation, and cartilage differentiation. *Mech. Dev.* **52**, 257–264 (1995).

117. T. Komori, in *Advances in Experimental Medicine and Biology* (Springer New York LLC, 2017; https://link.springer.com/chapter/10.1007/978-981-10-3233-2_6), vol. 962, pp. 83–93.

118. K. Kishimoto, R. Kitazawa, M. Kurosaka, S. Maeda, S. Kitazawa, Expression profile of genes related to osteoclastogenesis in mouse growth plate and articular cartilage. *Histochem. Cell Biol.* **125**, 593–602 (2006).

119. C. Karlsson, T. Dehne, A. Lindahl, M. Brittberg, A. Pruss, M. Sittinger, J. Ringe, Genome-wide expression profiling reveals new candidate genes associated with osteoarthritis. *Osteoarthr. Cartil.* **18**, 581–592 (2010).

120. D. M. Ornitz, N. Itoh, The fibroblast growth factor signaling pathway. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 215–266 (2015).

121. D. M. Ornitz, P. J. Marie, Fibroblast growth factor signaling in skeletal development and disease. *Genes Dev.* **29** (2015), pp. 1463–1486.

122. D. M. Ornitz, FGF signaling in the developing endochondral skeleton. *Cytokine Growth Factor Rev.* **16** (2005), pp. 205–213.

123. D. Yan, D. Chen, S. M. Cool, A. J. van Wijnen, K. Mikecz, G. Murphy, H. J. Im, Fibroblast growth factor receptor 1 is principally responsible for fibroblast growth factor 2-induced catabolic activities in human articular chondrocytes. *Arthritis Res. Ther.* **13**, R130 (2011).

124. T. Weng, L. Yi, J. Huang, F. Luo, X. Wen, X. Du, Q. Chen, C. Deng, D. Chen, L. Chen, Genetic inhibition of fibroblast growth factor receptor 1 in knee cartilage attenuates the

degeneration of articular cartilage in adult mice. *Arthritis Rheum.* **64**, 3982–3992 (2012).

125. Z. Wang, J. Huang, S. Zhou, F. Luo, Q. Tan, X. Sun, Z. Ni, H. Chen, X. Du, Y. Xie, L. Chen, Loss of Fgfr1 in chondrocytes inhibits osteoarthritis by promoting autophagic activity in temporomandibular joint. *J. Biol. Chem.* **293**, 8761–8774 (2018).

126. Q. Tan, B. Chen, Q. Wang, W. Xu, Y. Wang, Z. Lin, F. Luo, S. Huang, Y. Zhu, N. Su, M. Jin, C. Li, L. Kuang, H. Qi, Z. Ni, Z. Wang, X. Luo, W. Jiang, H. Chen, S. Chen, F. Li, B. Zhang, J. Huang, R. Zhang, K. Jin, X. Xu, C. Deng, X. Du, Y. Xie, L. Chen, A novel FGFR1-binding peptide attenuates the degeneration of articular cartilage in adult mice. *Osteoarthr. Cartil.* **26**, 1733–1743 (2018).

127. W. Xu, Y. Xie, Q. Wang, X. Wang, F. Luo, S. Zhou, Z. Wang, J. Huang, Q. Tan, M. Jin, H. Qi, J. Tang, L. Chen, X. Du, C. Zhao, G. Liang, L. Chen, A novel fibroblast growth factor receptor 1 inhibitor protects against cartilage degradation in a murine model of osteoarthritis. *Sci. Rep.* **6**, 1–12 (2016).

128. L. Yi, G. Lan, Y. Ju, X. Yin, P. Zhang, Y. Xu, T. Weng, Blockade of Fgfr1 with PD166866 Protects Cartilage from the Catabolic Effects Induced by Interleukin-1β: A Genome-Wide Expression Profiles Analysis. *Cartilage*, 194760352096820 (2020).

129. J. Tang, N. Su, S. Zhou, Y. Xie, J. Huang, X. Wen, Z. Wang, Q. Wang, W. Xu, X. Du, H. Chen, L. Chen, Fibroblast Growth Factor Receptor 3 Inhibits Osteoarthritis Progression in the Knee Joints of Adult Mice. *Arthritis Rheumatol.* **68**, 2432–2443 (2016).

130. S. Zhou, Y. Xie, W. Li, J. Huang, Z. Wang, J. Tang, W. Xu, X. Sun, Q. Tan, S. Huang, F. Luo, M. Xu, J. Wang, T. Wu, L. Chen, H. Chen, N. Su, X. Du, Y. Shen, L. Chen, Conditional Deletion of Fgfr3 in Chondrocytes leads to Osteoarthritis-like Defects in Temporomandibular Joint of Adult Mice. *Sci. Rep.* **6**, 1–13 (2016).

131. M. B. Ellman, D. Yan, K. Ahmadinia, D. Chen, H. S. An, H. J. Im, Fibroblast growth factor control of cartilage homeostasis. *J. Cell. Biochem.* **114**, 735–742 (2013).

132. Z. Liu, J. Xu, J. S. Colvin, D. M. Ornitz, Coordination of chondrogenesis and osteogenesis by fibroblast growth factor 18. *Genes Dev.* **16**, 859–869 (2002).

133. T. L. Vincent, Fibroblast growth factor 2: Good or bad guy in the joint? *Arthritis Res. Ther.* **13** (2011), p. 127.

134. H. Song, H. Du, J. Li, M. Wang, J. Wang, X. Ju, W. Mu, Effect of fibroblast growth factor 2 on degenerative endplate chondrocyte: From anabolism to catabolism. *Exp. Mol. Pathol.* **118** (2021), doi:10.1016/j.yexmp.2020.104590.

135. P. Krejci, J. Prochazkova, V. Bryja, P. Jelinkova, K. Pejchalova, A. Kozubik, L. Michels Thompson, W. R. Wilcox, Fibroblast growth factor inhibits interferon γ-STAT1 and interleukin 6-STAT3 signaling in chondrocytes. *Cell. Signal.* **21**, 151–160 (2009).

136. B. S. Yoon, R. Pogue, D. A. Ovchinnikov, I. Yoshii, Y. Mishina, R. R. Behringer, K. M. Lyons, BMPs regulate multiple aspects of growth-plate chondrogenesis through opposing actions on FGF pathways. *Development.* **133**, 4667–4678 (2006).

137. R. Raju, S. M. Palapetta, V. K. Sandhya, A. Sahu, A. Alipoor, L. Balakrishnan, J. Advani, B. George, K. R. Kini, N. P. Geetha, H. S. Prakash, T. S. K. Prasad, Y.-J. Chang, L. Chen, A. Pandey, H. Gowda, A Network Map of FGF-1/FGFR Signaling System. *J. Signal Transduct.* **2014**, 1–16 (2014).

138. M. Cucchiarini, T. Thurn, A. Weimer, D. Kohn, E. F. Terwilliger, H. Madry, Restoration of the extracellular matrix in human osteoarthritic articular cartilage by overexpression of the transcription factor SOX9. *Arthritis Rheum.* **56**, 158–167 (2007).

139. T. Matsuzaki, O. Alvarez-Garcia, S. Mokuda, K. Nagira, M. Olmer, R. Gamini, K. Miyata, Y. Akasaki, A. I. Su, H. Asahara, M. K. Lotz, FoxO transcription factors modulate autophagy and proteoglycan 4 in cartilage homeostasis and osteoarthritis. *Sci. Transl. Med.* **10**, eaan0746 (2018).

140. M. B. Goldring, J. Birkhead, L. J. Sandell, T. Kimura, S. M. Krane, Interleukin 1 suppresses expression of cartilage-specific types II and IX collagens and increases types I and III collagens in human chondrocytes. *J. Clin. Invest.* **82**, 2026–2037 (1988).

141. M. P. Vincenti, C. E. Brinckerhoff, Early response genes induced in chondrocytes stimulated with the inflammatory cytokine interleukin-1β. *Arthritis Res.* **3**, 381–388 (2001).

142. K. von der Mark, S. Frischholz, T. Aigner, F. Beier, J. Belke, S. Erdmann, H. Burkhardt, Upregulation of type x collagen expression in osteoarthritic cartilage. *Acta Orthop.* **66**, 125–129 (1995).

143. J. H. Rosenberg, V. Rai, M. F. Dilisio, D. K. Agrawal, Damage-associated molecular patterns in the pathogenesis of osteoarthritis: potentially novel therapeutic targets. *Mol. Cell. Biochem.* **434** (2017), pp. 171–179.

144. R. Liu-Bryan, Synovium and the innate inflammatory network in osteoarthritis progression topical collection on osteoarthritis. *Curr. Rheumatol. Rep.* **15**, 323 (2013).

145. M. Millerand, F. Berenbaum, C. Jacques, Danger signals and inflammaging in osteoarthritis. *Clin. Exp. Rheumatol.* **37** (2019), pp. 48–56.

146. P. Burrage, C. Brinckerhoff, Molecular Targets in Osteoarthritis: Metalloproteinases and Their Inhibitors. *Curr. Drug Targets*. **8**, 293–303 (2007).

147. R. Chijimatsu, T. Saito, Mechanisms of synovial joint and articular cartilage development. *Cell. Mol. Life Sci.* **76** (2019), pp. 3939–3952.

148. U. Il Chung, E. Schipani, A. P. McMahon, H. M. Kronenberg, Indian hedgehog couples chondrogenesis to osteogenesis in endochondral bone development. *J. Clin. Invest.* **107**, 295–304 (2001).

149. H. M. Kronenberg, Developmental regulation of the growth plate. *Nature*. **423** (2003), pp. 332–336.

150. X. Chen, C. M. Macica, A. Nasiri, A. E. Broadus, Regulation of articular chondrocyte proliferation and differentiation by Indian hedgehog and parathyroid hormone-related protein in mice. *Arthritis Rheum.* **58**, 3788–3797 (2008).

151. V. Lefebvre, P. Smits, Transcriptional control of chondrocyte fate and differentiation. *Birth Defects Res. Part C - Embryo Today Rev.* **75**, 200–212 (2005).

152. A. C. Lin, B. L. Seeto, J. M. Bartoszko, M. A. Khoury, H. Whetstone, L. Ho, C. Hsu, A. S. Ali, B. A. Alman, Modulating hedgehog signaling can attenuate the severity of osteoarthritis. *Nat. Med.* **15**, 1421–1425 (2009).

153. Y. M. Yoon, C. Do Oh, S. S. Kang, J. S. Chun, Protein kinase A regulates chondrogenesis of mesenchymal cells at the post-precartilage condensation stage via protein kinase C-α signaling. *J. Bone Miner. Res.* **15**, 2197–2205 (2000).

154. W. Zhang, J. Chen, S. Zhang, H. W. Ouyang, Inhibitory function of parathyroid hormone-related protein on chondrocyte hypertrophy: the implication for articular cartilage repair. *Arthritis Res. Ther.* **14** (2012), p. 221.

155. W. Huang, X. Zhou, V. Lefebvre, B. de Crombrugghe, Phosphorylation of SOX9 by Cyclic AMP-Dependent Protein Kinase A Enhances SOX9's Ability To Transactivate aCol2a1 Chondrocyte-Specific Enhancer. *Mol. Cell. Biol.* **20**, 4149–4158 (2000).

156. T. F. Li, Y. Dong, A. M. Ionescu, R. N. Rosier, M. J. Zuscik, E. M. Schwarz, R. J. O'Keefe, H. Drissi, Parathyroid hormone-related peptide (PTHrP) inhibits Runx2 expression through the PKA signaling pathway. *Exp. Cell Res.* **299**, 128–136 (2004).

157. S. Riemer, S. Gebhard, F. Beier, E. Pöschl, K. von der Mark, Role of c-fos in the regulation of type X collagen gene expression by PTH and PTHrP: Localization of a PTH/PTHrP-responsive region in the human COL10A1 enhancer. *J. Cell. Biochem.* **86**, 688–699 (2002).

158. L. A. Stanton, T. M. Underhill, F. Beier, MAP kinases in chondrocyte differentiation. *Dev. Biol.* **263** (2003), pp. 165–175.

159. D. Bar-Sagi, A. Hall, Ras and Rho GTPases: A family reunion. *Cell*. **103** (2000), pp. 227–238.

160. J. Massagué, J. Seoane, D. Wotton, Smad transcription factors. *Genes Dev.* **19** (2005), pp. 2783–2810.

161. T. F. Li, K. Yukata, G. Yin, T. Sheu, T. Maruyama, J. H. Jonason, W. Hsu, X. Zhang, G. Xiao, Y. T. Konttinen, D. Chen, R. J. O'Keefe, BMP-2 induces ATF4 phosphorylation in chondrocytes through a COX-2/PGE2 dependent signaling pathway. *Osteoarthr. Cartil.* **22**, 481–489 (2014).

162. R. F. Loeser, E. A. Erickson, D. L. Long, Mitogen-activated protein kinases as therapeutic targets in osteoarthritis. *Curr. Opin. Rheumatol.* **20** (2008), pp. 581–586.

163. F. Y. Wei, J. K. Lee, L. Wei, F. Qu, J. Z. Zhang, Correlation of insulin-like growth factor 1 and osteoarthritic cartilage degradation: a spontaneous osteoarthritis in guinea-pig. *Eur. Rev. Med. Pharmacol. Sci.* **21**, 4493–4500 (2017).

164. S. Ciarmatori, D. Kiepe, A. Haarmann, U. Huegel, B. Tönshoff, Signaling mechanisms leading to regulation of proliferation and differentiation of the mesenchymal chondrogenic cell line RCJ3.1C5.18 in response to IGF-1. *J. Mol. Endocrinol.* **38**, 493–508 (2007).

165. E. B. Hunziker, E. Kapfinger, J. Martin, J. Buckwalter, T. I. Morales, Insulin-like growth factor (IGF)-binding protein-3 (IGFBP-3) is closely associated with the chondrocyte nucleus in human articular cartilage. *Osteoarthr. Cartil.* **16**, 185–194 (2008).

166. R. C. Olney, K. Tsuchiya, D. M. Wilson, M. Mohtai, W. J. Maloney, D. J. Schurman, R. L. Smith, Chondrocytes from osteoarthritic cartilage have increased expression of insulin-like growth factor I (IGF-I) and IGF-binding protein-3 (IGFBP-3) and -5, but not IGF-II or IGFBP-4. *J. Clin. Endocrinol. Metab.* **81**, 1096–1103 (1996).

167. J. Martel-Pelletier, J. A. Di Battista, D. Lajeunesse, J. P. Pelletier, IGF/IGFBP axis in cartilage and bone in osteoarthritis pathogenesis. *Inflamm. Res.* **47** (1998), pp. 90–100.

168. M. A. Greene, R. F. Loeser, Function of the chondrocyte PI-3 kinase-Akt signaling pathway is stimulus dependent. *Osteoarthr. Cartil.* **23**, 949–956 (2015).

169. J. F. Xue, Z. M. Shi, J. Zou, X. L. Li, Inhibition of PI3K/AKT/mTOR signaling pathway promotes autophagy of articular chondrocytes and attenuates inflammatory response in rats with osteoarthritis. *Biomed. Pharmacother.* **89**, 1252–1261 (2017).

170. L. Xie, H. Xie, C. Chen, Z. Tao, C. Zhang, L. Cai, Inhibiting the PI3K/AKT/NF-κB signal pathway with nobiletin for attenuating the development of osteoarthritis: In vitro and in vivo studies. *Food Funct.* **10**, 2161–2175 (2019).

171. T. Sillat, G. Barreto, P. Clarijs, A. Soininen, M. Ainola, J. Pajarinen, M. Korhonen, Y. T. Konttinen, R. Sakalyte, M. Hukkanen, P. Ylinen, D. C. E. Nordström, Toll-like receptors in human chondrocytes and osteoarthritic cartilage. *Acta Orthop.* **84**, 585–592 (2013).

172. Q. Yang, W. Ding, Y. Cao, Y. Zhou, S. Ni, T. Shi, W. Fu, Interferonregulatoryfactor-8(IRF-8) regulates the expression of matrix metalloproteinase-13 (MMP-13) in chondrocytes. *Cell Stress Chaperones*. **23**, 393–398 (2018).

173. M. B. Goldring, M. Otero, Inflammation in osteoarthritis. *Curr. Opin. Rheumatol.* **23** (2011), pp. 471–478.

174. L. C. Tetlow, D. J. Adlam, D. E. Woolley, Matrix metalloproteinase and proinflammatory cytokine production by chondrocytes of human osteoarthritic cartilage: Associations with degenerative changes. *Arthritis Rheum.* **44**, 585–594 (2001).

175. S. Madhavan, M. Anghelina, D. Sjostrom, A. Dossumbekova, D. C. Guttridge, S. Agarwal, Biomechanical Signals Suppress TAK1 Activation to Inhibit NF-κB Transcriptional Activation in Fibrochondrocytes. *J. Immunol.* **179**, 6246–6254 (2007).

176. Z. Fan, S. Söder, S. Oehler, K. Fundel, T. Aigner, Activation of interleukin-1 signaling cascades in normal and osteoarthritis articular cartilage. *Am. J. Pathol.* **171**, 938–946 (2007).

177. A. Weber, P. Wasiliew, M. Kracht, Interleukin-1 (IL-1B) pathway. *Sci. Signal.* **3** (2010), p. 1.

178. Y. Li, E. H. Frank, Y. Wang, S. Chubinskaya, H. H. Huang, A. J. Grodzinsky, Moderate dynamic compression inhibits pro-catabolic response of cartilage to mechanical injury, tumor necrosis factor-α and interleukin-6, but accentuates degradation above a strain threshold. *Osteoarthr. Cartil.* **21**, 1933–1941 (2013).

179. C. Lianxu, J. Hongti, Y. Changlong, NF-κBp65-specific siRNA inhibits expression of genes of COX-2, NOS-2 and MMP-9 in rat IL-1β-induced and TNF-α-induced chondrocytes. *Osteoarthr. Cartil.* **14**, 367–376 (2006).

180. Y. Y. Chow, K. Y. Chin, The Role of Inflammation in the Pathogenesis of Osteoarthritis. *Mediators Inflamm.* **2020** (2020), , doi:10.1155/2020/8293921.

181. F. J. Blanco, M. Lotz, IL-1-induced nitric oxide inhibits chondrocyte proliferation via PGE2. *Exp. Cell Res.* **218**, 319–325 (1995).

182. Choi, Jo, Park, Kang, Park, NF-B Signaling Pathways in Osteoarthritic Cartilage

Destruction. *Cells*. **8**, 734 (2019).

183. J. Cheng, X. Hu, L. Dai, X. Zhang, B. Ren, W. Shi, Z. Liu, X. Duan, J. Zhang, X. Fu, W. Chen, Y. Ao, Inhibition of transforming growth factor β-activated kinase 1 prevents inflammation-related cartilage degradation in osteoarthritis. *Sci. Rep.* **6** (2016), doi:10.1038/srep34497.

184. K. B. Marcu, M. Otero, E. Olivotto, R. Maria Borzi, M. B. Goldring, NF-kB Signaling: Multiple Angles to Target OA. *Curr. Drug Targets*. **11**, 599–613 (2010).

185. T. Saito, A. Fukai, A. Mabuchi, T. Ikeda, F. Yano, S. Ohba, N. Nishida, T. Akune, N. Yoshimura, T. Nakagawa, K. Nakamura, K. Tokunaga, U. Il Chung, H. Kawaguchi, Transcriptional regulation of endochondral ossification by HIF-2α during skeletal growth and osteoarthritis development. *Nat. Med.* **16**, 678–686 (2010).

186. M. B. Goldring, Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis. *Ther. Adv. Musculoskelet. Dis.* **4**, 269–285 (2012).

187. H. Mendoza, D. G. Campbell, K. Burness, J. Hastie, N. Ronkina, J. H. Shim, J. S. C. Arthur, R. J. Davis, M. Gaestel, G. L. Johnson, S. Ghosh, P. Cohen, Roles for TAB1 in regulating the IL-1-dependent phosphorylation of the TAB3 regulatory subunit and activity of the TAK1 complex. *Biochem. J.* **409**, 711–722 (2008).

188. C. Malemud, Negative Regulators of JAK/STAT Signaling in Rheumatoid Arthritis and Osteoarthritis. *Int. J. Mol. Sci.* **18**, 484 (2017).

189. Y. Akasaki, A. Hasegawa, M. Saito, H. Asahara, Y. Iwamoto, M. K. Lotz, Dysregulated FOXO transcription factors in articular cartilage in aging and osteoarthritis. *Osteoarthr. Cartil.* **22**, 162–170 (2014).

190. A. M. Grabiec, C. Angiolilli, L. M. Hartkamp, L. G. M. Van Baarsen, P. P. Tak, K. A. Reedquist, JNK-dependent downregulation of FoxO1 is required to promote the survival of fibroblast-like synoviocytes in rheumatoid arthritis. *Ann. Rheum. Dis.* **74**, 1763–1771 (2015).

191. R. Wang, G. Chen, S. Zhang, R. Previn, D. Chen, Y. Jin, G. Zhou, Role of forkhead box

O transcription factors in oxidative stress-induced chondrocyte dysfunction: Possible therapeutic target for osteoarthritis? *Int. J. Mol. Sci.* **19** (2018), , doi:10.3390/ijms19123794.

192. R. F. Loeser, Integrins and chondrocyte-matrix interactions in articular cartilage. *Matrix Biol.* **39** (2014), pp. 11–16.

193. C. J. O'Conor, H. A. Leddy, H. C. Benefield, W. B. Liedtke, F. Guilak, TRPV4-mediated mechanotransduction regulates the metabolic response of chondrocytes to dynamic loading. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 1316–1321 (2014).

194. K. H. Vining, D. J. Mooney, Mechanical forces direct stem cell behaviour in development and regeneration. *Nat. Rev. Mol. Cell Biol.* **18** (2017), pp. 728–742.

195. S. Dupont, L. Morsut, M. Aragona, E. Enzo, S. Giulitti, M. Cordenonsi, F. Zanconato, J. Le Digabel, M. Forcato, S. Bicciato, N. Elvassore, S. Piccolo, Role of YAP/TAZ in mechanotransduction. *Nature*. **474**, 179–184 (2011).

196. R. Lesage, J. Kerkhofs, L. G. Geris, Computational modeling and reverse engineering to reveal dominant regulatory interactions controlling osteochondral differentiation: potential for regenerative medicine. *Front. Bioeng. Biotechnol.* **6**, 165 (2018).

197. L. Geris, Regenerative orthopaedics: In vitro, in vivo ... in silico. *Int. Orthop.* **38** (2014), pp. 1771–1778.

198. S. Motta, F. Pappalardo, Mathematical modeling of biological systems. *Brief. Bioinform.* **14**, 411–422 (2013).

199. M. Viceconti, F. Pappalardo, B. Rodriguez, M. Horner, J. Bischoff, F. Musuamba Tshinanu, In silico trials: Verification, validation and uncertainty quantification of predictive models used in the regulatory evaluation of biomedical products. *Methods* (2020), doi:10.1016/j.ymeth.2020.01.011.

200. B. B. Aldridge, J. M. Burke, D. A. Lauffenburger, P. K. Sorger, Physicochemical modelling of cell signalling pathways. *Nat. Cell Biol.* **8**, 1195–1203 (2006).

201. E. (Edda) Klipp, John Wiley & Sons., *Systems biology in practice : concepts, implementation and application* (Wiley-VCH, Weinheim, FRG, 2005;

https://books.google.com.co/books?id=_5o5Eb9t3ucC&dq=differences+between+SBML+2+AND+SBML+3&hl=es&source=gbs_navlinks_s).

202. 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. *Database*. **2020** (2020), doi:10.1093/database/baaa017.

203. S. Ghosh, Y. Matsuoka, Y. Asai, K. Y. Hsin, H. Kitano, Software for systems biology: From tools to integrated platforms. *Nat. Rev. Genet.* **12** (2011), pp. 821–832.

204. K. A. Janes, D. A. Lauffenburger, A biological approach to computational models of proteomic networks. *Curr. Opin. Chem. Biol.* **10**, 73–80 (2006).

205. G. Karlebach, R. Shamir, Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.* **9**, 770–780 (2008).

206. B. B. Aldridge, J. Saez-Rodriguez, J. L. Muhlich, P. K. Sorger, D. A. Lauffenburger, Fuzzy Logic Analysis of Kinase Pathway Crosstalk in TNF/EGF/Insulin-Induced Signaling. *PLoS Comput. Biol.* **5** (2009), doi:10.1371/journal.pcbi.1000340.

207. S. A. Kauffman, The origins of order; self organization and selection in evolution. *Int. J. Biochem.* **26**, 855 (1994).

208. L. Glass, S. A. Kauffman, The logical analysis of continuous, non-linear biochemical control networks. *J. Theor. Biol.* **39**, 103–129 (1973).

209. R. Thomas, M. Kaufman, Multistationarity, the basis of cell differentiation and memory. II. Logical analysis of regulatory networks in terms of feedback circuits. *Chaos*. **11** (2001), pp. 180–195.

210. G. Bernot, J.-P. Comet, A. Richard, J. Guespin, Application of formal methods to biological regulatory networks: extending Thomas' asynchronous logical approach with temporal logic. *J. Theor. Biol.* **229**, 339–347 (2004).

211. N. Le Novère, Quantitative and logic modelling of molecular and gene networks. *Nat. Rev. Genet.* **16**, 146–158 (2015).

212. M. Mojtahedi, A. Skupin, J. Zhou, I. G. Casta?o, R. Y. Y. Leong-Quong, H. Chang, K. Trachana, A. Giuliani, S. Huang, Cell Fate Decision as High-Dimensional Critical State Transition. *PLOS Biol.* **14**, e2000640 (2016).

213. W. Abou-Jaoudé, P. Traynard, P. T. Monteiro, J. Saez-Rodriguez, T. Helikar, D. Thieffry, C. Chaouiya, Logical Modeling and Dynamical Analysis of Cellular Networks. *Front. Genet.* **7**, 94 (2016).

214. C. Chaouiya, A. Naldi, D. Thieffry, Logical modelling of gene regulatory networks with GINsim. *Methods Mol. Biol.* **804**, 463–479 (2012).

215. M. K. Morris, J. Saez-Rodriguez, P. K. Sorger, D. A. Lauffenburger, Logic-based models for the analysis of cell signaling networks. *Biochemistry*. **49**, 3216–24 (2010).

216. A. Niarakis, T. Helikar, A practical guide to mechanistic systems modeling in biology using a logic-based approach. *Brief. Bioinform.* (2020), doi:10.1093/bib/bbaa236.

217. J. Saez-Rodriguez, L. G. Alexopoulos, J. Epperlein, R. Samaga, D. A. Lauffenburger, S. Klamt, P. K. Sorger, Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. *Mol. Syst. Biol.* **5,331** (2009), doi:10.1038/msb.2009.87.

218. L. E. Mcnamara, L.-A. Turner, K. V Burgess, Systems Biology Approaches Applied to Regenerative Medicine (2015), doi:10.1007/s40139-015-0072-4.

219. A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, 2005; http://epubs.siam.org/doi/book/10.1137/1.9780898717921).

220. L.-Z. Liu, F.-X. Wu, W.-J. Zhang, Reverse engineering of gene regulatory networks from biological data. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2**, 365–385 (2012).

221. D. Marbach, J. C. Costello, R. Küffner, N. M. Vega, R. J. Prill, D. M. Camacho, K. R. Allison, A. Aderhold, K. R. Allison, R. Bonneau, D. M. Camacho, Y. Chen, J. J. Collins, F. Cordero, J. C. Costello, M. Crane, F. Dondelinger, M. Drton, R. Esposito, R. Foygel, A. de la Fuente, J. Gertheiss, P. Geurts, A. Greenfield, M. Grzegorczyk, A.-C. Haury, B. Holmes, T. Hothorn, D. Husmeier, V. A. Huynh-Thu, A. Irrthum, M. Kellis, G.

Karlebach, R. Küffner, S. Lèbre, V. De Leo, A. Madar, S. Mani, D. Marbach, F. Mordelet, H. Ostrer, Z. Ouyang, R. Pandya, T. Petri, A. Pinna, C. S. Poultney, R. J. Prill, S. Rezny, H. J. Ruskin, Y. Saeys, R. Shamir, A. Sîrbu, M. Song, N. Soranzo, A. Statnikov, G. Stolovitzky, N. Vega, P. Vera-Licona, J.-P. Vert, A. Visconti, H. Wang, L. Wehenkel, L. Windhager, Y. Zhang, R. Zimmer, M. Kellis, J. J. Collins, G. Stolovitzky, Wisdom of crowds for robust gene network inference. *Nat. Methods*. **9**, 796–804 (2012).

222. C. Terfve, T. Cokelaer, D. Henriques, A. MacNamara, E. Goncalves, M. K. Morris, M. van Iersel, D. A. Lauffenburger, J. Saez-Rodriguez, CellNOptR: A flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC Syst. Biol.* **6**, 133 (2012).

223. S. S. Aghamiri, V. Singh, A. Naldi, T. Helikar, S. Soliman, A. Niarakis, Automated inference of Boolean models from molecular interaction maps using CaSQ. *Bioinformatics*. **36**, 4473–4482 (2020).

224. I. N. Melas, A. Mitsos, D. E. Messinis, T. S. Weiss, L. G. Alexopoulos, Combined logical and data-driven models for linking signalling pathways to cellular response. *BMC Syst. Biol.* **5**, 107 (2011).

225. R. Bonneau, D. J. Reiss, P. Shannon, M. Facciotti, L. Hood, N. S. Baliga, V. Thorsson, The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets de novo. *Genome Biol.* **7:R36** (2006), doi:10.1186/gb-2006-7-5-r36.

226. A. Dugourd, C. Kuppe, M. Sciacovelli, E. Gjerga, A. Gabor, K. B. Emdal, V. Vieira, D. B. Bekker-Jensen, J. Kranz, E. M. J. Bindels, A. S. H. Costa, A. Sousa, P. Beltrao, M. Rocha, J. V. Olsen, C. Frezza, R. Kramann, J. Saez-Rodriguez, Causal integration of multi-omics data with prior knowledge to generate mechanistic hypotheses. *Mol. Syst. Biol.* **17**, e9730 (2021).

227. B. G. Sengers, C. C. Van Donkelaar, C. W. J. Oomens, F. P. T. Baaijens, Computational Study of Culture Conditions and Nutrient Supply in Cartilage Tissue Engineering. *Biotechnol. Prog.* **21**, 1252–1261 (2008).

228. L. Geris, R. Schugart, H. Van Oosterwyck, In silico design of treatment strategies in wound healing and bone fracture healing. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*

368, 2683–2706 (2010).

229. P. Rajagopalan, S. Kasif, T. M. Murali, Systems Biology Characterization of Engineered Tissues. *Annu. Rev. Biomed. Eng.* **15**, 55–70 (2013).

230. A. Carlier, N. van Gastel, L. Geris, G. Carmeliet, H. Van Oosterwyck, Size Does Matter: An Integrative In Vivo-In Silico Approach for the Treatment of Critical Size Bone Defects. *PLoS Comput. Biol.* **10** (2014), doi:10.1371/journal.pcbi.1003888.

231. P. Lenas, F. P. Luyten, M. Doblare, E. Nicodemou-Lena, A. E. Lanzara, Modularity in developmental biology and artificial organs: A missing concept in tissue engineering. *Artif. Organs*. **35**, 656–662 (2011).

232. S. Kauffman, Homeostasis and differentiation in random genetic control networks. *Nature*. **224** (1969), pp. 177–178.

233. L. E. McNamara, L.-A. Turner, K. V Burgess, Systems Biology Approaches Applied to Regenerative Medicine. *Curr. Pathobiol. Rep.* **3**, 37–45 (2015).

234. A. L. Barabási, Scale-free networks: A decade and beyond. *Science (80-.).* **325**, 412–413 (2009).

235. Y. Y. Liu, J. J. Slotine, A. L. Barabási, Controllability of complex networks. *Nature*. **473**, 167–173 (2011).

236. J. Gao, Y. Y. Liu, R. M. D'Souza, A. L. Barabási, Target control of complex networks. *Nat. Commun.* **5**, 5415 (2014).

237. K. Kobayashi, K. Maeda, M. Tokuoka, A. Mochizuki, Y. Satou, Controlling Cell Fate Specification System by Key Genes Determined from Network Structure. *iScience*. **4**, 281–293 (2018).

238. T. Glimm, D. Headon, M. A. Kiskowski, Computational and mathematical models of chondrogenesis in vertebrate limbs. *Birth Defects Res. Part C - Embryo Today Rev.* **96** (2012), pp. 176–192.

239. S. A. Morris, P. Cahan, H. Li, A. M. Zhao, A. K. San Roman, R. A. Shivdasani, J. J. Collins, G. Q. Daley, Dissecting engineered cell types and enhancing cell fate conversion

via Cellnet. *Cell*. **158**, 889–902 (2014).

240. P. Cahan, H. Li, S. A. Morris, E. Lummertz Da Rocha, G. Q. Daley, J. J. Collins, CellNet: Network biology applied to stem cell engineering. *Cell*. **158**, 903–915 (2014).

241. A. H. Radley, R. M. Schwab, Y. Tan, J. Kim, E. K. W. Lo, P. Cahan, Assessment of engineered cells using CellNet and RNA-seq. *Nat. Protoc.* **12**, 1089–1102 (2017).

242. M. A. Kinney, L. T. Vo, J. M. Frame, J. Barragan, A. J. Conway, S. Li, K. K. Wong, J. J. Collins, P. Cahan, T. E. North, D. A. Lauffenburger, G. Q. Daley, A systems biology pipeline identifies regulatory networks for stem cell engineering. *Nat. Biotechnol.* **37**, 810–818 (2019).

243. M. Herberg, I. Roeder, Computational modelling of embryonic stem-cell fate control. *Development*. **142**, 2250–2260 (2015).

244. P. Pir, N. Le Novère, Mathematical models of pluripotent stem cells: At the dawn of predictive regenerative medicine. *Methods Mol. Biol.* **1386**, 331–350 (2016).

245. D. Schittler, J. Hasenauer, F. Allgower, S. Waldherr, Cell differentiation modeled via a coupled two-switch regulatory network. *Chaos An Interdiscip. J. Nonlinear Sci.* **20**, 045121 (2010).

246. M. Bodaker, Y. Louzoun, E. Mitrani, Mathematical Conditions for Induced Cell Differentiation and Trans-differentiation in Adult Cells. *Bull. Math. Biol.* **75**, 819–844 (2013).

247. A. J. Mueller, S. R. Tew, O. Vasieva, P. D. Clegg, E. G. Canty-Laird, A systems biology approach to defining regulatory mechanisms for cartilage and tendon cell phenotypes. *Sci. Rep.* **6**, 1–14 (2016).

248. Q. Miagoux, D. de Mezquita, V. Singh, V. Chaudru, M. Elati, E. Petit-Teixeira, A. Niarakis, *bioRxiv*, in press, doi:10.1101/2021.01.28.428679.

249. 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 Comput. Biol.* **4**, 100050 (2017).

250. N. Zerrouk, Q. Miagoux, A. Dispot, M. Elati, A. Niarakis, Identification of putative master regulators in rheumatoid arthritis synovial fibroblasts using gene expression data and network inference. *Sci. Rep.* **10**, 16236 (2020).

251. C. J. Proctor, C. Macdonald, J. M. Milner, A. D. Rowan, T. E. Cawston, A computer simulation approach to assessing therapeutic intervention points for the prevention of cytokine-induced cartilage breakdown. *Arthritis Rheumatol.* **66**, 979–989 (2014).

252. I. N. Melas, A. D. Chairakaki, E. I. Chatzopoulou, D. E. Messinis, T. Katopodi, V. Pliaka, S. Samara, A. Mitsos, Z. Dailiana, P. Kollia, L. G. Alexopoulos, Modeling of signaling pathways in chondrocytes based on phosphoproteomic and cytokine release data. *Osteoarthr. Cartil.* **22**, 509–518 (2014).

253. J. Kerkhofs, S. J. Roberts, F. P. Luyten, H. van Oosterwyck, L. Geris, Relating the chondrocyte gene network to growth plate morphology: From genes to phenotype. *PLoS One*. **7**, 1–11 (2012).

254. J. Kerkhofs, J. Leijten, J. Bolander, F. P. Luyten, J. N. Post, L. Geris, A qualitative model of the differentiation network in chondrocyte maturation: A holistic view of chondrocyte hypertrophy. *PLoS One*. **11**, 1–27 (2016).

255. J. Kerkhofs, L. Geris, A Semiquantitative Framework for Gene Regulatory Networks: Increasing the Time and Quantitative Resolution of Boolean Networks. *PLoS One*. **10**, e0130033 (2015).

256. S. Schivo, S. Khurana, K. Govindaraj, J. Scholma, J. Kerkhofs, L. Zhong, X. Huang, J. van de Pol, R. Langerak, A. J. van Wijnen, L. Geris, M. Karperien, J. N. Post, ECHO, the executable CHOndrocyte: A computational model to study articular chondrocytes in health and disease. *Cell. Signal.* **68**, 109471 (2020).

257. G. A. Orozco, P. Tanska, C. Florea, A. J. Grodzinsky, R. K. Korhonen, A novel mechanobiological model can predict how physiologically relevant dynamic loading causes proteoglycan loss in mechanically injured articular cartilage. *Sci. Rep.* **8** (2018), doi:10.1038/s41598-018-33759-3.

258. A. S. A. Eskelinen, P. Tanska, C. Florea, G. A. Orozco, P. Julkunen, A. J. Grodzinsky, R. K. Korhonen, Mechanobiological model for simulation of injured cartilage

degradation via proinflammatory cytokines and mechanical. *PLoS Comput. Biol.* **16**, e1007998 (2020).

259. J. Delile, M. Herrmann, N. Peyriéras, R. Doursat, A cell-based computational model of early embryogenesis coupling mechanical behaviour and gene regulation. *Nat. Commun.* **8**, 1–10 (2017).

260. T. Peng, L. Liu, A. L. MacLean, C. W. Wong, W. Zhao, Q. Nie, A mathematical model of mechanotransduction reveals how mechanical memory regulates mesenchymal stem cell fate decisions. *BMC Syst. Biol.* **11**, 55 (2017).

261. M. Mak, T. Kim, M. H. Zaman, R. D. Kamm, Multiscale mechanobiology: Computational models for integrating molecules to multicellular systems. *Integr. Biol. (United Kingdom)*. **7** (2015), pp. 1093–1108.

262. V. D. Sree, A. B. Tepole, Computational systems mechanobiology of growth and remodeling: Integration of tissue mechanics and cell regulatory network dynamics. *Curr. Opin. Biomed. Eng.* **15** (2020), pp. 75–80.

263. Z. Karagöz, L. Rijns, P. Y. W. Dankers, M. van Griensven, A. Carlier, Towards understanding the messengers of extracellular space: Computational models of outside-in integrin reaction networks. *Comput. Struct. Biotechnol. J.* **19** (2021), pp. 303–314.

264. A. Shuaib, D. Motan, P. Bhattacharya, A. McNabb, T. M. Skerry, D. Lacroix, Heterogeneity in The Mechanical Properties of Integrins Determines Mechanotransduction Dynamics in Bone Osteoblasts. *Sci. Rep.* **9**, 13113 (2019).

265. J. Kerkhofs, Chondrogenic Differentiation in the growth Plate: a Computational Modelling Approach. *[Doctoral thesis] KU Leuven, Univ. Liege* (2015) (available at https://orbi.uliege.be/handle/2268/186583).

266. A. Yamashita, M. Morioka, Y. Yahara, M. Okada, T. Kobayashi, S. Kuriyama, S. Matsuda, N. Tsumaki, Generation of scaffoldless hyaline cartilaginous tissue from human iPSCs. *Stem Cell Reports*. **4**, 404–418 (2015).

267. N. Van Gastel, S. Stegen, I. Stockmans, K. Moermans, J. Schrooten, D. Graf, F. P. Luyten, G. Carmeliet, Expansion of murine periosteal progenitor cells with fibroblast

growth factor 2 reveals an intrinsic endochondral ossification program mediated by bone morphogenetic protein 2. *Stem Cells* (2014), doi:10.1002/stem.1783.

268. P. Shannon, M. Andrew, O. Owen, B. Nitin S, W. Jonathan T, R. Daniel, A. Nada, S. Benno, I. Trey, Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13**, 2498–2504 (2003).

269. T. Helikar, B. Kowal, S. McClenathan, M. Bruckner, T. Rowley, A. Madrahimov, B. Wicks, M. Shrestha, K. Limbu, J. A. Rogers, The Cell Collective: Toward an open and collaborative approach to systems biology. *BMC Syst. Biol.* **6** (2012), doi:10.1186/1752-0509-6-96.

270. W. J. Pinamont, N. K. Yoshioka, G. M. Young, V. Karuppagounder, E. L. Carlson, A. Ahmad, R. Elbarbary, F. Kamal, Standardized histomorphometric evaluation of osteoarthritis in a surgical mouse model. *J. Vis. Exp.* **2020** (2020), doi:10.3791/60991.

271. S. S. Glasson, M. G. Chambers, W. B. Van Den Berg, C. B. Little, The OARSI histopathology initiative - recommendations for histological assessments of osteoarthritis in the mouse. *Osteoarthr. Cartil.* **18**, S17–S23 (2010).

272. K. Basso, A. A. Margolin, G. Stolovitzky, U. Klein, R. Dalla-Favera, A. Califano, Reverse engineering of regulatory networks in human B cells. *Nat. Genet.* **37**, 382–390 (2005).

273. R. F. Loeser, A. L. Olex, M. A. McNulty, C. S. Carlson, M. Callahan, C. Ferguson, J. S. Fetrow, Disease Progression and Phasic Changes in Gene Expression in a Mouse Model of Osteoarthritis. *PLoS One*. **8** (2013), doi:10.1371/journal.pone.0054633.

274. S. S. Glasson, T. J. Blanchet, E. A. Morris, The surgical destabilization of the medial meniscus (DMM) model of osteoarthritis in the 129/SvEv mouse. *Osteoarthr. Cartil.* **15**, 1061–1069 (2007).

275. N. Altman, M. Krzywinski, The curse(s) of dimensionality this-month. *Nat. Methods*. **15** (2018), pp. 399–400.

276. C. Xu, S. A. Jackson, Machine learning and complex biological data. *Genome Biol.* **20** (2019), pp. 1–4.

277. W. E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. **8**, 118–127 (2007).

278. D. R. Rhodes, J. Yu, K. Shanker, N. Deshpande, R. Varambally, D. Ghosh, T. Barrette, A. Pandey, A. M. Chinnaiyan, Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9309–9314 (2004).

279. A. K. Turnbull, R. R. Kitchen, A. A. Larionov, L. Renshaw, J. M. Dixon, A. H. Sims, Direct integration of intensity-level data from Affymetrix and Illumina microarrays improves statistical power for robust reanalysis. *BMC Med. Genomics*. **5**, 35 (2012).

280. Y. Zhang, G. Parmigiani, W. E. Johnson, ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics Bioinforma.* **2** (2020), doi:10.1093/nargab/lqaa078.

281. D. Ruppert, The Elements of Statistical Learning: Data Mining, Inference, and Prediction. *J. Am. Stat. Assoc.* **99**, 567–567 (2004).

282. M. D. Wilkinson, M. Dumontier, Ij. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da S. Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. . 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data 2016 31*. **3**, 1–9 (2016).

283. M. AM, K. AM, B. MK, B. MP, T. TN, Differential gene expression analysis reveals pathways important in early post-traumatic osteoarthritis in an equine model. *BMC Genomics*. **21** (2020), doi:10.1186/S12864-020-07228-Z.

284. J. J. Faith, B. Hayete, J. T. Thaden, I. Mogno, J. Wierzbowski, G. Cottarel, S. Kasif, J. J. Collins, T. S. Gardner, Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* **5**, 0054–0066 (2007).

285. P. E. Meyer, K. Kontos, F. Lafitte, G. Bontempi, Information-theoretic inference of large transcriptional regulatory networks. *Eurasip J. Bioinforma. Syst. Biol.* **2007** (2007), doi:10.1155/2007/79879.

286. E. Schipani, H. E. Ryan, S. Didrickson, T. Kobayashi, M. Knight, R. S. Johnson, Hypoxia in cartilage: HIF-1α is essential for chondrocyte growth arrest and survival. *Genes Dev.* **15**, 2865 (2001).

287. Z. L, Z. L, H. Z, X. R, L. X, Y. S, M. J, Z. N, M. W, D. L, W. P, Increased HIF-1 α in Knee Osteoarthritis Aggravate Synovial Fibrosis via Fibroblast-Like Synoviocyte Pyroptosis. *Oxid. Med. Cell. Longev.* **2019** (2019), doi:10.1155/2019/6326517.

288. S. Hu, C. Zhang, L. Ni, C. Huang, D. Chen, K. Shi, H. Jin, K. Zhang, Y. Li, L. Xie, M. Fang, G. Xiang, X. Wang, J. Xiao, Stabilization of HIF-1α alleviates osteoarthritis via enhancing mitophagy. *Cell Death Dis. 2020 116*. **11**, 1–16 (2020).

289. C. IB, J. SA, H. DF, P.-V. S, S. DG, Hypoxia inducible factor-1 alpha expression in human normal and osteoarthritic chondrocytes. *Osteoarthr. Cartil.* **12**, 336–345 (2004).

290. K. Okada, D. Mori, Y. Makii, H. Nakamoto, Y. Murahashi, F. Yano, S. H. Chang, Y. Taniguchi, H. Kobayashi, H. Semba, N. Takeda, W. Piao, K. Hanaoka, T. Nagano, S. Tanaka, T. Saito, Hypoxia-inducible factor-1 alpha maintains mouse articular cartilage through suppression of NF-κB signaling. *Sci. Reports 2020 101*. **10**, 1–11 (2020).

291. L. Slemc, T. Kunej, Transcription factor HIF1A: downstream targets, associated pathways, polymorphic hypoxia response element (HRE) sites, and initiative for standardization of reporting in scientific literature. *Tumor Biol.* **37**, 14851–14861 (2016).

292. W. W, L. Y, H. J, Z. S, W. Y, X. X, H. A, F. Q, Z. F, L. R, Comparative analysis of gene expression profiles of hip articular cartilage between non-traumatic necrosis and osteoarthritis. *Gene*. **591**, 43–47 (2016).

293. F. K, S. H, R. M, B. A, H. B, K. R, G. ML, F. B, S. P, P. A, B. I, Analysis of pigmented villonodular synovitis with genome-wide complementary DNA microarray and tissue array technology reveals insight into potential novel therapeutic approaches. *Arthritis Rheum.* **54**, 1009–1019 (2006).

294. B. Poulet, T. A. T. Westerhof, R. W. Hamilton, S. J. Shefelbine, A. A. Pitsillides, Spontaneous osteoarthritis in Str/ort mice is unlikely due to greater vulnerability to mechanical trauma. *Osteoarthr. Cartil.* **21**, 756–763 (2013).

295. B. Poulet, V. Ulici, T. C. Stone, M. Pead, V. Gburcik, E. Constantinou, D. B. Palmer, F. Beier, J. A. Timmons, A. A. Pitsillides, Time-series transcriptional profiling yields new perspectives on susceptibility to murine osteoarthritis. *Arthritis Rheum.* **64**, 3256–3266 (2012).

296. S. Thysen, F. P. Luyten, R. J. U. Lories, Targets, models and challenges in osteoarthritis research. *DMM Dis. Model. Mech.* **8**, 17–30 (2015).

297. H. Fang, F. Beier, Mouse models of osteoarthritis: Modelling risk factors and assessing outcomes. *Nat. Rev. Rheumatol.* **10** (2014), pp. 413–421.

298. B. RA, M. JC, W. DJ, M. DM, A. DJ, S. PM, W. SL, H. RD, Transcriptional profiling and pathway analysis of monosodium iodoacetate-induced experimental osteoarthritis in rats: relevance to human disease. *Osteoarthr. Cartil.* **15**, 1190–1198 (2007).

299. Y. Tan, P. Cahan, SingleCellNet: A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms and Across Species. *Cell Syst.* **9**, 207-213.e2 (2019).

300. D. A, A. R, S. SL, M. SS, S. M, Identification of clinical phenotypes in knee osteoarthritis: a systematic review of the literature. *BMC Musculoskelet. Disord.* **17**, 1–12 (2016).

301. D. Lajeunesse, P. Reboul, in *Bone and Osteoarthritis* (Springer London, 2007), pp. 19–39.

302. Y. C, P. Z, Z. K, L. J, S. Z, Y. X, L. H, Z. X, Y. Y, Y. D, Z. Y, X. Y, Z. ZY, H. T, L. W, O. H, Classification of four distinct osteoarthritis subtypes with a knee joint tissue transcriptome atlas. *Bone Res.* **8** (2020), doi:10.1038/S41413-020-00109-X.

303. A. Burleigh, A. Chanalaris, M. D. Gardiner, C. Driscoll, O. Boruc, J. Saklatvala, T. L. Vincent, Joint immobilization prevents murine osteoarthritis and reveals the highly mechanosensitive nature of protease expression in vivo. *Arthritis Rheum.* **64**, 2278–2288 (2012).

304. N. Bomer, F. M. F. Cornelis, Y. F. Ramos, W. Den Hollander, L. Storms, R. Van Der Breggen, N. Lakenberg, P. E. Slagboom, I. Meulenbelt, R. J. Lories, The effect of forced exercise on knee joints in Dio2-/- mice: Type II iodothyronine deiodinasedeficient mice are less prone to develop OA-like cartilage damage upon excessive mechanical stress. *Ann. Rheum. Dis.* **75**, 571–577 (2016).

305. N. Gossan, L. Zeef, J. Hensman, A. Hughes, J. F. Bateman, L. Rowley, C. B. Little, H. D. Piggins, M. Rattray, R. P. Boot-Handford, Q. J. Meng, The circadian clock in murine chondrocytes regulates genes controlling key aspects of cartilage homeostasis. *Arthritis Rheum.* **65**, 2334–2345 (2013).

306. L. Lodewyckx, F. Cailotto, S. Thysen, F. P. Luyten, R. J. Lories, Tight regulation of wingless-type signaling in the articular cartilage - subchondral bone biomechanical unit: Transcriptomics in Frzb-knockout mice. *Arthritis Res. Ther.* **14** (2012), doi:10.1186/ar3695.

307. M. D. Gardiner, T. L. Vincent, C. Driscoll, A. Burleigh, G. Bou-Gharios, J. Saklatvala, H. Nagase, A. Chanalaris, Transcriptional analysis of micro-dissected articular cartilage in post-traumatic murine osteoarthritis. *Osteoarthr. Cartil.* **23**, 616–628 (2015).

308. B. S. Carvalho, R. A. Irizarry, A framework for oligonucleotide microarray preprocessing. *Bioinformatics*. **26**, 2363–2367 (2010).

309. S. Durinck, P. T. Spellman, E. Birney, W. Huber, Mapping identifiers for the integration of genomic datasets with the R/ Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).

310. S. Durinck, Y. Moreau, A. Kasprzyk, S. Davis, B. De Moor, A. Brazma, W. Huber, BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics*. **21**, 3439–3440 (2005).

311. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).

312. A. L. and M. Wiener, Classification and Regression by randomForest. R News 2. **3** (2003), pp. 18–22.

313. R. Lesage, M. N. F. Blanco, R. Narcisi, T. Welting, G. J. V. M. van Osch, L. Geris,

bioRxiv, in press, doi:10.1101/2021.09.27.461207.

314. S. Raman, U. FitzGerald, J. M. Murphy, Interplay of Inflammatory Mediators with Epigenetics and Cartilage Modifications in Osteoarthritis. *Front. Bioeng. Biotechnol.* **6**, 22 (2018).

315. Q. Ji, Y. Zheng, G. Zhang, Y. Hu, X. Fan, Y. Hou, L. Wen, L. Li, Y. Xu, Y. Wang, F. Tang, Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. *Ann. Rheum. Dis.* **78**, 100–110 (2019).

316. K. Von Der Mark, T. Kirsch, A. Nerlich, A. Kuss, G. Weseloh, K. Glückert, H. Stöss, Type x collagen synthesis in human osteoarthritic cartilage. indication of chondrocyte hypertrophy. *Arthritis Rheum.* **35**, 806–811 (1992).

317. E. V Tchetina, A. R. Poole, E. M. Zaitseva, E. P. Sharapova, N. G. Kashevarova, E. A. Taskina, L. I. Alekseeva, L. A. Semyonova, S. I. Glukhova, A. N. Kuzin, M. A. Makarov, S. A. Makarov, Differences in Mammalian Target of Rapamycin Gene Expression in the Peripheral Blood and Articular Cartilages of Osteoarthritic Patients and Disease Activity. *Arthritis*. **2013:46148** (2013), doi:10.1155/2013/461486.

318. G. G. Poehling, Osteoarthritis: Diagnosis and medical/surgical management, third edition: Ronald W. Moskowitz, Davis S. Howell, Roy D. Altman, Joseph A. Buckwalter, and Victor M. Goldberg. Philadelphia, W. B. Saunders, 2001, 685 pp., $189.00. *Arthrosc. J. Arthrosc. Relat. Surg.* **18**, 678 (2002).

319. P. M. van der Kraan, W. B. van den Berg, Chondrocyte hypertrophy and osteoarthritis: role in initiation and progression of cartilage degeneration? *Osteoarthr. Cartil.* **20**, 223–232 (2012).

320. M. N. Ferrao Blanco, Y. M. Bastiaansen-Jenniskens, M. G. Chambers, A. A. Pitsillides, R. Narcisi, G. J. V. M. van Osch, Effect of Inflammatory Signaling on Human Articular Chondrocyte Hypertrophy: Potential Involvement of Tissue Repair Macrophages. *Cartilage* (2021), doi:10.1177/19476035211021907.

321. M. A. Cleary, G. J. V. M. van Osch, P. A. Brama, C. A. Hellingman, R. Narcisi, FGF, TGFβ and Wnt crosstalk: embryonic to in vitro cartilage development from mesenchymal stem cells. *J. Tissue Eng. Regen. Med.* **9**, 332–342 (2015).

322. S. Green, M. Şerban, R. Scholl, N. Jones, I. Brigandt, W. Bechtel, Network analyses in systems biology: new strategies for dealing with biological complexity. *Synthese*. **195**, 1751–1777 (2018).

323. Rationalizing combination therapies. *Nat. Med.* **23**, 1113 (2017).

324. G. Nalesso, J. Sherwood, J. Bertrand, T. Pap, M. Ramachandran, C. de Bari, C. Pitzalis, F. Dell'Accio, WNT-3A modulates articular chondrocyte phenotype by activating both canonical and noncanonical pathways. *J. Cell Biol.* **193**, 551–564 (2011).

325. H. Kitano, Systems biology: A brief overview. *Science (80-.).* **295**, 1662–1664 (2002).

326. B. L. Puniya, L. Allen, C. Hochfelder, M. Majumder, T. Helikar, Systems Perturbation Analysis of a Large-Scale Signal Transduction Model Reveals Potentially Influential Candidates for Cancer Therapeutics. *Front. Bioeng. Biotechnol.* **4**, 10 (2016).

327. F. Eduati, P. Jaaks, J. Wappler, T. Cramer, C. A. Merten, M. J. Garnett, J. Saez-Rodriguez, Patient-specific logic models of signaling pathways from screenings on cancer biopsies to prioritize personalized combination therapies. *Mol. Syst. Biol.* **16**, 1–13 (2020).

328. R. Lesage, L. Geris, PPI network- Interactive Modeling of Biological Networks | Cell Collective. *Cell Collect.* (2021), (available at https://cellcollective.org/#a5b66073-6769-4c88-bf6b-37ca1aa8f766).

329. R. Lesage, L. Geris, GRN - Interactive Modeling of Biological Networks | Cell Collective. *Cell Collect.* (2021), (available at https://cellcollective.org/#474de240-8752-4c3b-aa63-23640e50bf7a).

330. A. Rowan, W. Hui, T. E. Cawston, C. D. Richards, Adenoviral gene transfer of interleukin-1 in combination with oncostatin M induces significant joint damage in a murine model. *Am. J. Pathol.* **162**, 1975–1984 (2003).

331. T. Saito, A. Fukai, A. Mabuchi, T. Ikeda, F. Yano, S. Ohba, N. Nishida, T. Akune, N. Yoshimura, T. Nakagawa, K. Nakamura, K. Tokunaga, U. Il Chung, H. Kawaguchi, Transcriptional regulation of endochondral ossification by HIF-2α during skeletal growth and osteoarthritis development. *Nat. Med.* **16**, 678–686 (2010).

332. D. Hodgson, A. D. Rowan, F. Falciani, C. J. Proctor, Systems biology reveals how altered TGFβ signalling with age reduces protection against pro-inflammatory stimuli. *PLOS Comput. Biol.* **15**, e1006685 (2019).

333. W. Hui, A. D. Rowan, T. Cawston, Transforming growth factor-β1 blocks the release of collagen fragments from bovine nasal cartilage stimulated by oncostatin M in combination with IL-1α. *Cytokine*. **12**, 765–769 (2000).

334. M. M. Shull, I. Ormsby, A. B. Kier, S. Pawlowski, R. J. Diebold, M. Yin, R. Allen, C. Sidman, G. Proetzel, D. Calvin, N. Annunziata, T. Doetschman, Targeted disruption of the mouse transforming growth factor-β1 gene results in multifocal inflammatory disease. *Nature*. **359**, 693–699 (1992).

335. F. CM, S. EM, R. PR, P. JE, R. RN, O. RJ, Smad2 and 3 mediate transforming growth factor-beta1-induced inhibition of chondrocyte maturation. *Endocrinology*. **141**, 4728–4735 (2000).

336. P. van der Kraan, C. Matta, A. Mobasheri, Age-Related Alterations in Signaling Pathways in Articular Chondrocytes: Implications for the Pathogenesis and Progression of Osteoarthritis - A Mini-Review. *Gerontology*. **63**, 29–35 (2017).

337. R. Lesage, Github - Rapha-L/Insilico_chondro: Scripts & data used to run the analyses done in the publication. *Git Hub* (2021), (available at https://github.com/Rapha-L/Insilico_chondro).

338. H. Akiyama, C. Shigeno, K. I. Iyama, H. Ito, Y. Hiraki, J. Konishi, T. Nakamura, Indian hedgehog in the late-phase differentiation in mouse chondrogenic EC cells, ATDC5: Upregulation of type X collagen and osteoprotegerin ligand mRNAs. *Biochem. Biophys. Res. Commun.* **257**, 814–820 (1999).

339. A. Liu, P. Trairatphisan, E. Gjerga, A. Didangelos, J. Barratt, J. Saez-Rodriguez, From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL. *npj Syst. Biol. Appl.* **5**, 1–10 (2019).

340. D. Türei, T. Korcsmáros, J. Saez-Rodriguez, OmniPath: Guidelines and gateway for literature-curated signaling pathway resources. *Nat. Methods*. **13**, 966–967 (2016).

341. M. Yang, J. Simm, C. C. Lam, P. Zakeri, G. J. P. Van Westen, Y. Moreau, J. Saez-Rodriguez, Linking drug target and pathway activation for effective therapy using multi-Task learning. *Sci. Rep.* **8**, 1–10 (2018).

342. F. Binette, D. P. McQuaid, D. R. Haudenschild, P. C. Yaeger, J. M. McPherson, R. Tubo, Expression of a stable articular cartilage phenotype without evidence of hypertrophy by adult human articular chondrocytes in vitro. *J. Orthop. Res.* **16**, 207–216 (1998).

343. M. Kærn, T. C. Elston, W. J. Blake, J. J. Collins, Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet. 2005 66*. **6**, 451–464 (2005).

344. H. Kitano, The Theory of Biological Robustness and Its Implication in Cancer. *Ernst Schering Res. Found. Workshop*, 69–88 (2007).

345. A. Mobasheri, W. E. van Spil, E. Budd, I. Uzieliene, E. Bernotiene, A.-C. Bay-Jensen, J. Larkin, M. C. Levesque, O. Gualillo, Y. Henrotin, Molecular taxonomy of osteoarthritis for patient stratification, disease management and drug development. *Curr. Opin. Rheumatol.* **31**, 80–89 (2019).

346. V. A. Huynh-Thu, A. Irrthum, L. Wehenkel, P. Geurts, Inferring regulatory networks from expression data using tree-based methods. *PLoS One*. **5**, e12776 (2010).

347. A. A. Margolin, I. Nemenman, K. Basso, C. Wiggins, G. Stolovitzky, R. Favera, A. Califano, ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. *BMC Bioinformatics*. **7**, S7 (2006).

348. A. Haury, F. Mordelet, P. Vera-licona, J. Vert, Open Access TIGRESS : Trustful Inference of Gene REgulation using Stability Selection. *BMC Syst. Biol.* (2012), doi:10.1186/1752-0509-6-145.

349. A. Fauré, A. Naldi, C. Chaouiya, D. Thieffry, Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*. **22**, 124–131 (2006).

350. U. Alon, *An introduction to systems biology: Design principles of biological circuits* (Chapman and Hall/CRC Press, 2006).

351. J. G. T. Zañudo, M. Aldana, G. Martínez-Mekler, Boolean threshold networks: Virtues and limitations for biological modeling. *Intell. Syst. Ref. Libr.* **11**, 113–151 (2011).

352. W. L. Tam, L. Freitas Mendes, X. Chen, R. Lesage, I. Van Hoven, E. Leysen, G. Kerckhofs, K. Bosmans, Y. C. Chai, A. Yamashita, N. Tsumaki, L. Geris, S. J. Roberts, F. P. Luyten, Human pluripotent stem cell-derived cartilaginous organoids promote scaffold-free healing of critical size long bone defects. *Stem Cell Res. Ther.* **12**, 513 (2021).

353. A. Montagud, P. Traynard, L. Martignetti, E. Bonnet, E. Barillot, A. Zinovyev, L. Calzone, "Conceptual and computational framework for logical modelling of biological networks deregulated in diseases" (2017), , doi:10.1093/bib/bbx163.

354. L. Wang, Y. Y. Shao, R. T. Ballock, Thyroid hormone interacts with the Wnt/β-catenin signaling pathway in the terminal differentiation of growth plate chondrocytes. *J. Bone Miner. Res.* **22**, 1988–1995 (2007).

355. H.-Y. Kim, S. Mohan, Role and Mechanisms of Actions of Thyroid Hormone on the Skeletal Development. *Bone Res.* **1**, 146–161 (2013).

356. J. C. Barnard, A. J. Williams, B. Rabier, O. Chassande, J. Samarut, S. Cheng, J. H. D. Bassett, G. R. Williams, Thyroid Hormones Regulate Fibroblast Growth Factor Receptor Signaling during Chondrogenesis. *Endocrinology*. **146** (2005), pp. 5568–5580.

357. D. A. Stevens, R. P. Hasserjian, H. Robson, T. Siebler, S. M. Shalet, G. R. Williams, Thyroid Hormones Regulate Hypertrophic Chondrocyte Differentiation and Expression of Parathyroid Hormone-Related Peptide and Its Receptor During Endochondral Bone Formation. *J. Bone Miner. Res.* **15**, 2431–2442 (2000).

358. S. L. Adams, A. J. Cohen, L. Lassová, Integration of signaling pathways regulating chondrocyte differentiation during endochondral bone formation. *J. Cell. Physiol.* **213**, 635–641 (2007).

359. J. Nam, B. D. Aguda, B. Rath, S. Agarwal, Biomechanical Thresholds Regulate Inflammation through the NF-κB Pathway: Experiments and Modeling. *PLoS One*. **4**, e5262 (2009).

360. T. Lawrence, The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb. Perspect. Biol.* **1**, a001651 (2009).

361. Q. Pan, Y. Yu, Q. Chen, C. Li, H. Wu, Y. Wan, J. Ma, F. Sun, Sox9, a key transcription factor of bone morphogenetic protein-2-induced chondrogenesis, is activated through BMP pathway and a CCAAT box in the proximal promoter. *J. Cell. Physiol.* **217**, 228–241 (2008).

362. T. Sueishi, Y. Akasaki, N. Goto, I. Kurakazu, M. Toya, M. Kuwahara, T. Uchida, M. Hayashida, H. Tsushima, H. Bekki, M. K. Lotz, Y. Nakashima, GRK5 Inhibition Attenuates Cartilage Degradation via Decreased NF-κB Signaling. *Arthritis Rheumatol.* **72**, 620–631 (2020).

363. M. Bitzer, G. Von Gersdorff, D. Liang, A. Dominguez-Rosales, A. A. Beg, M. Rojkind, E. P. Böttinger, A mechanism of suppression of TGF-β/SMAD signaling by NF-κB/RelA. *Genes Dev.* **14**, 187–197 (2000).

364. B. Yordanov, S. J. Dunn, H. Kugler, A. Smith, G. Martello, S. Emmott, A method to identify and analyze biological programs through automated reasoning. *npj Syst. Biol. Appl.* **2**, 1–16 (2016).

365. S. Dunn, M. A. Li, E. Carbognin, A. Smith, G. Martello, A common molecular logic determines embryonic stem cell self-renewal and reprogramming. *EMBO J.* **38**, e100003 (2019).

366. S. J. Dunn, G. Martello, B. Yordanov, S. Emmott, A. G. Smith, Defining an essential transcription factor program for naïve pluripotency. *Science (80-.).* **344**, 1156–1160 (2014).

367. L. Wang, Y. Y. Shao, R. T. Ballock, Thyroid hormone-mediated growth and differentiation of growth plate chondrocytes involves IGF-1 modulation of β-catenin signaling. *J. Bone Miner. Res.* **25**, 1138–1146 (2010).

368. H. Yuan, J. Mao, L. Li, D. Wu, Suppression of glycogen synthase kinase activity is not sufficient for leukemia enhancer factor-1 activation. *J. Biol. Chem.* **274**, 30419–23 (1999).

369. V. W. Ding, R.-H. Chen, F. McCormick, Differential Regulation of Glycogen Synthase Kinase 3β by Insulin and Wnt Signaling. *J. Biol. Chem.* **275**, 32475–32481 (2000).

370. S. Mukherjee, R. Lesage, W. Wilson, L. Geris, in *14th World Congress in Computational Mechanics (WCCM) and ECCOMAS Congress 2020* (2020).

371. Z. Fan, S. Söder, S. Oehler, K. Fundel, T. Aigner, Activation of interleukin-1 signaling cascades in normal and osteoarthritis articular cartilage. *Am. J. Pathol.* **171**, 938–946 (2007).

372. M. Neidlin, S. Dimitrakopoulou, L. G. Alexopoulos, Multi-tissue network analysis for drug prioritization in knee osteoarthritis. *Sci. Rep.* **9** (2019), doi:10.1038/s41598-019-51627-6.

373. M. Kapoor, J. Martel-Pelletier, D. Lajeunesse, J. P. Pelletier, H. Fahmi, Role of proinflammatory cytokines in the pathophysiology of osteoarthritis. *Nat. Rev. Rheumatol.* **7** (2011), pp. 33–42.

374. A. P. Hollander, I. Pidoux, A. Reiner, C. Rorabeck, R. Bourne, A. R. Poole, Damage to type II collagen in aging and osteoarthritis starts at the articular surface, originates around chondrocytes, and extends into the cartilage with progressive degeneration. *J. Clin. Invest.* **96**, 2859–2869 (1995).

375. V. B. Shim, P. J. Hunter, P. Pivonka, J. W. Fernandez, A multiscale framework based on the Physiome markup languages for exploring the initiation of osteoarthritis at the bone-cartilage interface. *IEEE Trans. Biomed. Eng.* **58**, 3532–3536 (2011).

376. S. J. Millward-Sadler, D. M. Salter, Integrin-Dependent Signal Cascades in Chondrocyte Mechanotransduction. *Ann. Biomed. Eng.* **32**, 435–446 (2004).

377. J. B. Thorlund, A. Holsgaard-Larsen, M. W. Creaby, G. M. Jørgensen, N. Nissen, M. Englund, L. S. Lohmander, Changes in knee joint load indices from before to 12 months after arthroscopic partial meniscectomy: A prospective cohort study. *Osteoarthr. Cartil.* **24**, 1153–1159 (2016).

378. M. A. Wozniak, K. Modzelewska, L. Kwong, P. J. Keely, Focal adhesion regulation of cell behavior. *Biochim. Biophys. Acta - Mol. Cell Res.* **1692** (2004), pp. 103–119.

379. Z. Sun, S. S. Guo, R. Fässler, Integrin-mediated mechanotransduction. *J. Cell Biol.* **215** (2016), pp. 445–456.

380. E. C. Arner, M. D. Tortorella, Signal transduction through chondrocyte integrin receptors induces matrix metalloproteinase synthesis and synergizes with interleukin-1. *Arthritis Rheum.* **38**, 1304–1314 (1995).

381. N. Q. Balaban, U. S. Schwarz, D. Riveline, P. Goichberg, G. Tzur, I. Sabanay, D. Mahalu, S. Safran, A. Bershadsky, L. Addadi, B. Geiger, Force and focal adhesion assembly: A close relationship studied using elastic micropatterned substrates. *Nat. Cell Biol.* **3**, 466–472 (2001).

382. J. Hoon, M. Tan, C.-G. Koh, The Regulation of Cellular Responses to Mechanical Cues by Rho GTPases. *Cells*. **5**, 17 (2016).

383. M. Amano, M. Nakayama, K. Kaibuchi, Rho-kinase/ROCK: A key regulator of the cytoskeleton and cell polarity. *Cytoskeleton*. **67** (2010), pp. 545–554.

384. G. Wang, A. Woods, S. Sabari, L. Pagnotta, L. A. Stanton, F. Beier, RhoA/ROCK Signaling Suppresses Hypertrophic Chondrocyte Differentiation. *J. Biol. Chem.* **279**, 13205–13214 (2004).

385. M. Maekawa, T. Ishizaki, S. Boku, N. Watanabe, A. Fujita, A. Iwamatsu, T. Obinata, K. Ohashi, K. Mizuno, S. Narumiya, Signaling from Rho to the actin cytoskeleton through protein kinases ROCK and LIM-kinase. *Science (80-.).* **285**, 895–898 (1999).

386. D. R. Haudenschild, J. Chen, N. Pang, M. K. Lotz, D. D. D'Lima, Rho kinase-dependent activation of SOX9 in chondrocytes. *Arthritis Rheum.* **62**, 191–200 (2010).

387. N. Wang, Review of cellular mechanotransduction. *J. Phys. D. Appl. Phys.* **50** (2017), p. 233002.

388. E. Lucchinetti, M. M. Bhargava, P. A. Torzilli, The effect of mechanical load on integrin subunits α5 and β1 in chondrocytes from mature and immature cartilage explants. *Cell Tissue Res.* **315**, 385–391 (2004).

389. E. N. Olson, A. Nordheim, Linking actin dynamics and gene transcription to drive cellular motile functions. *Nat. Rev. Mol. Cell Biol.* **11** (2010), pp. 353–365.

390. C. S. Hill, J. Wynne, R. Treisman, The Rho family GTPases RhoA, Racl, and CDC42Hsregulate transcriptional activation by SRF. *Cell*. **81**, 1159–1170 (1995).

391. Y. Deng, A. Wu, P. Li, G. Li, L. Qin, H. Song, K. K. Mak, Yap1 Regulates Multiple Steps of Chondrocyte Differentiation during Skeletal Development and Bone Repair. *Cell Rep.* **14**, 2224–2237 (2016).

392. Y. Deng, J. Lu, W. Li, A. Wu, X. Zhang, W. Tong, K. K. Ho, L. Qin, H. Song, K. K. Mak, Reciprocal inhibition of YAP/TAZ and NF-κB regulates osteoarthritic cartilage degradation. *Nat. Commun.* **9**, 1–14 (2018).

393. W. Zhong, Y. Li, L. Li, W. Zhang, S. Wang, X. Zheng, YAP-mediated regulation of the chondrogenic phenotype in response to matrix elasticity. *J. Mol. Histol.* **44**, 587–595 (2013).

394. Y. Guan, X. Yang, W. Yang, C. Charbonneau, Q. Chen, Mechanical activation of mammalian target of rapamycin pathway is required for cartilage development. *FASEB J.* **28**, 4470–4481 (2014).

395. B. Pal, H. Endisha, Y. Zhang, M. Kapoor, mTOR: A Potential Therapeutic Target in Osteoarthritis? *Drugs R D.* **15** (2015), pp. 27–36.

396. B. Yan, Z. Zhang, D. Jin, C. Cai, C. Jia, W. Liu, T. Wang, S. Li, H. Zhang, B. Huang, P. Lai, H. Wang, A. Liu, C. Zeng, D. Cai, Y. Jiang, X. Bai, mTORC1 regulates PTHrP to coordinate chondrocyte growth, proliferation and differentiation. *Nat. Commun.* **7** (2016), doi:10.1038/ncomms11151.

397. Q. Dai, Z. Xu, X. Ma, N. Niu, S. Zhou, F. Xie, L. Jiang, J. Wang, W. Zou, MTOR/Raptor signaling is critical for skeletogenesis in mice through the regulation of Runx2 expression. *Cell Death Differ.* **24**, 1886–1899 (2017).

398. K. L. Posey, F. Coustry, A. C. Veerisetty, M. G. Hossain, M. J. Gambello, J. T. Hecht, Novel mTORC1 Mechanism Suggests Therapeutic Targets for COMPopathies. *Am. J. Pathol.* **189**, 132–146 (2019).

399. L. Hui, V. Rodrik, R. M. Pielak, S. Knirr, Y. Zheng, D. A. Foster, mTOR-dependent suppression of protein phosphatase 2A is critical for phospholipase D survival signals in human breast cancer cells. *J. Biol. Chem.* **280**, 35829–35835 (2005).

400. A. L. McNulty, H. A. Leddy, W. Liedtke, F. Guilak, TRPV4 as a therapeutic target for

joint diseases. *Naunyn. Schmiedebergs. Arch. Pharmacol.* **388** (2015), pp. 437–450.

401. M. N. Phan, H. A. Leddy, B. J. Votta, S. Kumar, D. S. Levy, D. B. Lipshutz, H. L. Suk, W. Liedtke, F. Guilak, Functional characterization of TRPV4 as an osmotically sensitive ion channel in porcine articular chondrocytes. *Arthritis Rheum.* **60**, 3028–3037 (2009).

402. C. J. O'Conor, S. Ramalingam, N. A. Zelenski, H. C. Benefield, I. Rigo, D. Little, C. L. Wu, D. Chen, W. Liedtke, A. L. McNulty, F. Guilak, Cartilage-Specific Knockout of the Mechanosensory Ion Channel TRPV4 Decreases Age-Related Osteoarthritis. *Sci. Rep.* **6**, 1–10 (2016).

403. S. Muramatsu, M. Wakabayashi, T. Ohno, K. Amano, R. Ooishi, T. Sugahara, S. Shiojiri, K. Tashiro, Y. Suzuki, R. Nishimura, S. Kuhara, S. Sugano, T. Yoneda, A. Matsuda, Functional gene screening system identified TRPV4 as a regulator of chondrogenic differentiation. *J. Biol. Chem.* **282**, 32158–32167 (2007).

404. M. Lv, Y. Zhou, X. Chen, L. Han, L. Wang, X. L. Lu, Calcium signaling of in situ chondrocytes in articular cartilage under compressive loading: Roles of calcium sources and cell membrane ion channels. *J. Orthop. Res.* **36**, 730–738 (2018).

405. E. Smedler, P. Uhlén, Frequency decoding of calcium oscillations. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 964–969 (2014).

406. X. Gong, W. Xie, B. Wang, L. Gu, F. Wang, X. Ren, C. Chen, L. Yang, Altered spontaneous calcium signaling of in situ chondrocytes in human osteoarthritic cartilage. *Sci. Rep.* **7**, 1–12 (2017).

407. N. Trompeter, J. Gardinier, V. DeBarros, M. Boggs, V. Gangadharan, W. Cain, L. Hurd, R. Duncan, *bioRxiv*, in press, doi:10.1101/2020.03.10.985713.

408. K. Novakofski, A. Boehm, L. Fortier, The small GTPase Rho mediates articular chondrocyte phenotype and morphology in response to interleukin-1α and insulin-like growth factor-1. *J. Orthop. Res.* **27**, 58–64 (2009).

409. N. D. Leipzig, S. V. Eleswarapu, K. A. Athanasiou, The effects of TGF-β1 and IGF-I on the biomechanics and cytoskeleton of single chondrocytes. *Osteoarthr. Cartil.* **14**, 1227–1236 (2006).

410. S. Pritchard, F. Guilak, Effects of interleukin-1 on calcium signaling and the increase of filamentous actin in isolated and in situ articular chondrocytes. *Arthritis Rheum.* **54**, 2164–2174 (2006).

411. M. P. Vincenti, C. E. Brinckerhoff, Early response genes induced in chondrocytes stimulated with the inflammatory cytokine interleukin-1β. *Arthritis Res.* **3**, 381–388 (2001).

412. S. Madhavan, M. Anghelina, D. Sjostrom, A. Dossumbekova, D. C. Guttridge, S. Agarwal, Biomechanical Signals Suppress TAK1 Activation to Inhibit NF-κB Transcriptional Activation in Fibrochondrocytes. *J. Immunol.* **179**, 6246–6254 (2007).

413. R. M. Clancy, P. F. Gomez, S. B. Abramson, Nitric oxide sustains nuclear factor kappaB activation in cytokine-stimulated chondrocytes. *Osteoarthr. Cartil.* **12**, 552–558 (2004).

414. J. N. Sharma, A. Al-Omran, S. S. Parvathy, Role of nitric oxide in inflammatory diseases. *Inflammopharmacology*. **15** (2007), pp. 252–259.

415. K. Vuolteenaho, T. Moilanen, U. Jalonen, A. Lahti, R. Nieminen, H. M. Van Beuningen, P. M. Van Der Kraan, E. Moilanen, TGFβ inhibits IL-1 -induced iNOS expression and NO production in immortalized chondrocytes. *Inflamm. Res.* **54**, 420–427 (2005).

416. B. Ma, L. Zhong, C. A. Van Blitterswijk, J. N. Post, M. Karperien, T cell factor 4 is a pro-catabolic and apoptotic factor in human articular chondrocytes by potentiating nuclear factor κB signaling. *J. Biol. Chem.* **288**, 17552–17558 (2013).

417. R. C. Olney, D. M. Wilson, M. Mohtai, P. J. Fielder, R. L. Smith, Interleukin-1 and tumor necrosis factor-α increase insulin-like growth factor-binding protein-3 (IGFBP-3) production and IGFBP-3 protease activity in human articular chondrocytes. *J. Endocrinol.* **146**, 279–286 (1995).

418. J. Deschner, B. Rath-Deschner, S. Agarwal, Regulation of matrix metalloproteinase expression by dynamic tensile strain in rat fibrochondrocytes. *Osteoarthr. Cartil.* **14**, 264–272 (2006).

419. A. Dossumbekova, M. Anghelina, S. Madhavan, L. He, N. Quan, T. Knobloch, S. Agarwal, Biomechanical signals inhibit IKK activity to attenuate NF-κB transcription

activity in inflamed chondrocytes. *Arthritis Rheum.* **56**, 3284–3296 (2007).

420. B. Fermor, J. Brice Weinberg, D. S. Pisetsky, M. A. Misukonis, A. J. Banes, F. Guilak, The effects of static and intermittent compression on nitric oxide production in articular cartilage explants. *J. Orthop. Res.* **19**, 729–737 (2001).

421. B. Fermor, J. B. Weinberg, D. S. Pisetsky, F. Guilak, The influence of oxygen tension on the induction of nitric oxide and prostaglandin E2 by mechanical stress in articular cartilage. *Osteoarthr. Cartil.* **13**, 935–941 (2005).

422. L. G. Alexopoulos, L. A. Setton, F. Guilak, The biomechanical role of the chondrocyte pericellular matrix in articular cartilage. *Acta Biomater.* **1**, 317–325 (2005).

423. A. Erdemir, Open knee: A pathway to community driven modeling and simulation in joint biomechanics. *J. Med. Devices, Trans. ASME*. **7**, 1 (2013).

424. S. A. Maas, B. J. Ellis, G. A. Ateshian, J. A. Weiss, FEBio: Finite elements for biomechanics. *J. Biomech. Eng.* **134** (2012), doi:10.1115/1.4005694.

425. P. Julkunen, W. Wilson, J. S. Jurvelin, R. K. Korhonen, Composition of the pericellular matrix modulates the deformation behaviour of chondrocytes in articular cartilage under static loading. *Med. Biol. Eng. Comput.* **47**, 1281–1290 (2009).

426. A. A. Young, S. McLennan, M. M. Smith, S. M. Smith, M. A. Cake, R. A. Read, J. Melrose, D. H. Sonnabend, C. R. Flannery, C. B. Little., Proteoglycan 4 downregulation in a sheep meniscectomy model of early osteoarthritis. *Arthritis Res. Ther.* **8** (2006), doi:10.1186/ar1898.

427. P. T. Paradowski, M. Englund, L. S. Lohmander, E. M. Roos, The effect of patient characteristics on variability in pain and function over two years in early knee osteoarthritis. *Health Qual. Life Outcomes*. **3**, 1–6 (2005).

428. S. Fu, H. Meng, S. Inamdar, B. Das, H. Gupta, W. Wang, C. L. Thompson, M. M. Knight, Activation of TRPV4 by mechanical, osmotic or pharmaceutical stimulation is anti-inflammatory blocking IL-1β mediated articular cartilage matrix destruction. *Osteoarthr. Cartil.* **29**, 89–99 (2021).

429. R. Potla, M. Hirano-Kobayashi, H. Wu, H. Chen, A. Mammoto, B. D. Matthews, D. E.

Ingber, Molecular mapping of transmembrane mechanotransduction through the β1 integrin–CD98hc–TRPV4 axis. *J. Cell Sci.* **133** (2020), doi:10.1242/jcs.248823.

430. C. Ji, C. A. McCulloch, TRPV4 integrates matrix mechanosensing with Ca2+ signaling to regulate extracellular matrix remodeling. *FEBS J.* (2020), , doi:10.1111/febs.15665.

431. P. Occhetta, A. Mainardi, E. Votta, Q. Vallmajo-Martin, M. Ehrbar, I. Martin, A. Barbero, M. Rasponi, Hyperphysiological compression of articular cartilage induces an osteoarthritic phenotype in a cartilage-on-a-chip model. *Nat. Biomed. Eng. 2019 37*. **3**, 545–557 (2019).

432. T. C. B. Pollard, S. E. Gwilym, A. J. Carr, The assessment of early osteoarthritis. *J. Bone Jt. Surg. - Ser. B.* **90** (2008), pp. 411–421.

433. A. Mahmoudian, L. S. Lohmander, H. Jafari, F. P. Luyten, Towards classification criteria for early-stage knee osteoarthritis: A population-based study to enrich for progressors. *Semin. Arthritis Rheum.* **51**, 285–291 (2021).

434. A. Mahmoudian, L. S. Lohmander, H. Jafari, F. P. Luyten, Towards classification criteria for early-stage knee osteoarthritis: A population-based study to enrich for progressors. *Semin. Arthritis Rheum.* **51**, 285–291 (2021).

435. R. Bhargava, The 5 Models Of Content Curation | Influential Marketing Blog. *Influ. Mark. Blog* (2011), (available at https://www.rohitbhargava.com/2011/03/the-5-models-of-content-curation.html).

436. S. Corbett, Curation nation: how to win in a world where consumers are creators. *Prometheus*. **29**, 184–187 (2011).

437. Y. A. Rim, Y. Nam, J. H. Ju, The role of chondrocyte hypertrophy and senescence in osteoarthritis initiation and progression. *Int. J. Mol. Sci.* **21** (2020), , doi:10.3390/ijms21072358.

438. A. Fabregat, S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie, P. Garapati, R. Haw, B. Jassal, F. Korninger, B. May, M. Milacic, C. D. Roca, K. Rothfels, C. Sevilla, V. Shamovsky, S. Shorser, T. Varusai, G. Viteri, J. Weiser, G. Wu, L. Stein, H. Hermjakob, P. D'Eustachio, The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–

D655 (2018).

439. I. Rodchenkov, O. Babur, A. Luna, B. A. Aksoy, J. V. Wong, D. Fong, M. Franz, M. C. Siper, M. Cheung, M. Wrana, H. Mistry, L. Mosier, J. Dlin, Q. Wen, C. O'Callaghan, W. Li, G. Elder, P. T. Smith, C. Dallago, E. Cerami, B. Gross, U. Dogrusoz, E. Demir, G. D. Bader, C. Sander, Pathway Commons 2019 Update: Integration, analysis and exploration of pathway data. *Nucleic Acids Res.* **48**, D489–D497 (2020).

440. D. Türei, A. Valdeolivas, L. Gul, N. Palacio-Escat, M. Klein, O. Ivanova, M. Ölbei, A. Gábor, F. Theis, D. Módos, T. Korcsmáros, J. Saez-Rodriguez, Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. *Mol. Syst. Biol.* **17**, e9923 (2021).

441. W. L. Tam, F. P. Luyten, S. J. Roberts, From skeletal development to the creation of pluripotent stem cell-derived bone-forming progenitors. *Philos. Trans. R. Soc. B Biol. Sci.* **373** (2018), doi:10.1098/RSTB.2017.0218.

442. F. Eduati, R. Utharala, D. Madhavan, U. P. Neumann, T. Longerich, T. Cramer, J. Saez-Rodriguez, C. A. Merten, A microfluidics platform for combinatorial drug screening on cancer biopsies. *Nat. Commun. 2018 91*. **9**, 1–13 (2018).

443. F. Guilak, L. G. Alexopoulos, M. L. Upton, I. Youn, J. B. Choi, L. Cao, L. A. Setton, M. A. Haider, in *Annals of the New York Academy of Sciences* (2006; http://www.ncbi.nlm.nih.gov/pubmed/16831947), vol. 1068, pp. 498–512.

444. J. Gao, J. L. Williams, E. Roan, Multiscale modeling of growth plate cartilage mechanobiology. *Biomech. Model. Mechanobiol.* **16**, 1–13 (2016).

445. M. Anghelina, D. Sjostrom, P. Perera, J. Nam, T. Knobloch, S. Agarwal, Regulation of biomechanical signals by NF-κB transcription factors in chondrocytes. *Biorheology*. **45**, 245 (2008).

446. N. Le Novère, M. Hucka, H. Mi, S. Moodie, F. Schreiber, A. Sorokin, E. Demir, K. Wegner, M. I. Aladjem, S. M. Wimalaratne, F. T. Bergman, R. Gauges, P. Ghazal, H. Kawaji, L. Li, Y. Matsuoka, A. Villéger, S. E. Boyd, L. Calzone, M. Courtot, U. Dogrusoz, T. C. Freeman, A. Funahashi, S. Ghosh, A. Jouraku, S. Kim, F. Kolpakov, A. Luna, S. Sahle, E. Schmidt, S. Watterson, G. Wu, I. Goryanin, D. B. Kell, C. Sander, H.

Sauro, J. L. Snoep, K. Kohn, H. Kitano, The Systems Biology Graphical Notation. *Nat. Biotechnol.* **27** (2009), pp. 735–741.

447. S. M. Keating, D. Waltemath, M. König, F. Zhang, A. Dräger, C. Chaouiya, F. T. Bergmann, A. Finney, C. S. Gillespie, T. Helikar, S. Hoops, R. S. Malik-Sheriff, S. L. Moodie, I. I. Moraru, C. J. Myers, A. Naldi, B. G. Olivier, S. Sahle, J. C. Schaff, L. P. Smith, M. J. Swat, D. Thieffry, L. Watanabe, D. J. Wilkinson, M. L. Blinov, K. Begley, J. R. Faeder, H. F. Gómez, T. M. Hamm, Y. Inagaki, W. Liebermeister, A. L. Lister, D. Lucio, E. Mjolsness, C. J. Proctor, K. Raman, N. Rodriguez, C. A. Shaffer, B. E. Shapiro, J. Stelling, N. Swainston, N. Tanimura, J. Wagner, M. Meier-Schellersheim, H. M. Sauro, B. Palsson, H. Bolouri, H. Kitano, A. Funahashi, H. Hermjakob, J. C. Doyle, M. Hucka, R. R. Adams, N. A. Allen, B. R. Angermann, M. Antoniotti, G. D. Bader, J. Červený, M. Courtot, C. D. Cox, P. Dalle Pezze, E. Demir, W. S. Denney, H. Dharuri, J. Dorier, D. Drasdo, A. Ebrahim, J. Eichner, J. Elf, L. Endler, C. T. Evelo, C. Flamm, R. M. Fleming, M. Fröhlich, M. Glont, E. Gonçalves, M. Golebiewski, H. Grabski, A. Gutteridge, D. Hachmeister, L. A. Harris, B. D. Heavner, R. Henkel, W. S. Hlavacek, B. Hu, D. R. Hyduke, H. Jong, N. Juty, P. D. Karp, J. R. Karr, D. B. Kell, R. Keller, I. Kiselev, S. Klamt, E. Klipp, C. Knüpfer, F. Kolpakov, F. Krause, M. Kutmon, C. Laibe, C. Lawless, L. Li, L. M. Loew, R. Machne, Y. Matsuoka, P. Mendes, H. Mi, F. Mittag, P. T. Monteiro, K. N. Natarajan, P. M. Nielsen, T. Nguyen, A. Palmisano, J. Pettit, T. Pfau, R. D. Phair, T. Radivoyevitch, J. M. Rohwer, O. A. Ruebenacker, J. Saez-Rodriguez, M. Scharm, H. Schmidt, F. Schreiber, M. Schubert, R. Schulte, S. C. Sealfon, K. Smallbone, S. Soliman, M. I. Stefan, D. P. Sullivan, K. Takahashi, B. Teusink, D. Tolnay, I. Vazirabad, A. Kamp, U. Wittig, C. Wrzodek, F. Wrzodek, I. Xenarios, A. Zhukova, J. Zucker, SBML Level 3: an extensible format for the exchange and reuse of biological models. *Mol. Syst. Biol.* **16** (2020), doi:10.15252/msb.20199110.

448. L. Kuepfer, M. Peter, U. Sauer, J. Stelling, Ensemble modeling for analysis of cell signaling dynamics. *Nat. Biotechnol.* **25**, 1001–1006 (2007).

Supporting information

Appendix of Chapter 3

Table S3.1: Description of microarray datasets and manual annotations as WT and OA based on authors' initial annotations.

GEO ID	sample names	description (according to authors)	Manual annotation for clustering
GSE26475	GSM649331_2_G275_0hrN1R_01.CEL.gz	Murine knee joint_Naive_0hr_rep1	WT
	GSM649338_G275_0hrN1R_02.CEL.gz	Murine knee joint_Naive_0hr_rep2	WT
	GSM649351_G275_0hrN1R_03.CEL.gz	Murine knee joint_Naive_0hr_rep3	WT
	GSM649473_G275_3dO1R_16.CEL.gz	Murine knee joint_Operated (DMM)_3day_rep1	OA
	GSM649474_G275_3dO1R_17.CEL.gz	Murine knee joint_Operated (DMM)_3day_rep2	OA
	GSM649475_G275_3dO1R_18.CEL.gz	Murine knee joint_Operated (DMM)_3day_rep3	OA
	GSM651065_G275_3dS1R_07.CEL.gz	Murine knee joint_Operated (Sham)_3day_rep1	WT
	GSM651066_G275_3dS1R_08.CEL.gz	Murine knee joint_Operated (Sham)_3day_rep2	WT
	GSM651067_G275_3dS1R_09.CEL.gz	Murine knee joint_Operated (Sham)_3day_rep3	WT
	GSM651068_G275_6hrO1R_13.CEL.gz	Murine knee joint_Operated (DMM)_6hr_rep1	OA
	GSM651069_G275_6hrO1R_14.CEL.gz	Murine knee joint_Operated (DMM)_6hr_rep2	OA
	GSM651070_G275_6hrO1R_15.CEL.gz	Murine knee joint_Operated (DMM)_6hr_rep3	OA
	GSM651071_G275_6hrS1R_04.CEL.gz	Murine knee joint_Operated (Sham)_6hr_rep3	WT
	GSM651072_G275_6hrS1R_05.CEL.gz	Murine knee joint_Operated (Sham)_6hr_rep2	WT
	GSM651073_G275_6hrS1R_06.CEL.gz	Murine knee joint_Operated (Sham)_6hr_rep1	WT
	GSM651074_G275_7dO1R_19.CEL.gz	Murine knee joint_Operated (DMM)_7days_rep1	OA
	GSM651075_G275_7dO1R_20.CEL.gz	Murine knee joint_Operated (DMM)_7days_rep2	OA
	GSM651076_G275_7dO1R_21.CEL.gz	Murine knee joint_Operated (DMM)_7days_rep3	OA
	GSM651095_G275_7dS1R_10.CEL.gz	Murine knee joint_Operated (Sham)_7days_rep1	WT
	GSM651096_G275_7dS1R_11.CEL.gz	Murine knee joint_Operated (Sham)_7days_rep2	WT
	GSM651105_G275_7dS1R_12.CEL.gz	Murine knee joint_Operated (Sham)_7days_rep3	WT
GSE33754	GSM834970.CEL.gz	CBA, 8 weeks old, rep1	WT
	GSM834971.CEL.gz	CBA, 8 weeks old, rep2	WT
	GSM834972.CEL.gz	CBA, 8 weeks old, rep3	WT
	GSM834973.CEL.gz	CBA, 8 weeks old, rep4	WT
	GSM834974.CEL.gz	Str/ort 8 weeks old, rep1	OA
	GSM834975.CEL.gz	Str/ort 8 weeks old, rep2	OA
	GSM834976.CEL.gz	Str/ort 8 weeks old, rep3	OA
	GSM834978.CEL.gz	CBA, 18 weeks old, rep2	WT
	GSM834979.CEL.gz	CBA 18 weeks old, rep3	WT

	GSM834980.CEL.gz	CBA 18 weeks old, rep4	WT
	GSM834981.CEL.gz	Str/ort 18 weeks old, rep1	OA
	GSM834982.CEL.gz	Str/ort 18 weeks old, rep2	OA
	GSM834983.CEL.gz	Str/ort 18 weeks old, rep3	OA
	GSM834984.CEL.gz	Str/ort 18 weeks old, rep4	OA
	GSM834985.CEL.gz	CBA, 40 weeks old, rep1	WT
	GSM834986.CEL.gz	CBA, 40 weeks old, rep2	WT
	GSM834987.CEL.gz	CBA, 40 weeks old, rep3	WT
	GSM834988.CEL.gz	CBA, 40 weeks old, rep4	WT
	GSM834989.CEL.gz	Str/ort, 40 weeks old, rep1	OA
	GSM834990.CEL.gz	Str/ort, 40 weeks old, rep2	OA
	GSM834991.CEL.gz	Str/ort, 40 weeks old, rep3	OA
	GSM834992.CEL.gz	Str/ort, 40 weeks old, rep4	OA
GSE79239	Sample_20	WT run	OA
	Sample_21	WT run	OA
	Sample_22	WT run	OA
	Sample_27	WT run	OA
	Sample_12	WT norun	WT
	Sample_14	WT norun	WT
	Sample_15	WT norun	WT
	Sample_35	Dio-/- run	WT
	Sample_16	Dio-/- norun	WT
	Sample_17	Dio-/- norun	WT
	Sample_19	Dio-/- norun	WT
	Sample_29	WT run	OA
	Sample_30	WT run	OA
	Sample_31	WT run	OA
	Sample_32	Dio-/- run	WT
	Sample_33	Dio-/- run	WT
	Sample_34	Dio-/- run	WT
	Sample_23	WT run	OA
	Sample_24	Dio-/- run	WT
	Sample_25	Dio-/- run	WT
GSE53857	'2w_ctrl_1'	'treatment: control'	WT
	'2w_ctrl_2'	'treatment: control'	WT
	'2w_ctrl_3'	'treatment: control'	WT
	'2w_dmm_1'	'treatment: destabilisation of the medial meniscus (DMM)'	OA
	'2w_dmm_2'	'treatment: destabilisation of the medial meniscus (DMM)'	OA
	'2w_dmm_3'	'treatment: destabilisation of the medial meniscus (DMM)'	OA
	'4w_ctrl_1'	'treatment: control'	WT
	'4w_ctrl_2'	'treatment: control'	WT
	'4w_ctrl_3'	'treatment: control'	WT
	'4w_dmm_1'	'treatment: destabilisation of the medial meniscus (DMM)'	OA

	'4w_dmm_2'	'treatment: destabilisation of the medial meniscus (DMM)'	OA
	'4w_dmm_3'	'treatment: destabilisation of the medial meniscus (DMM)'	OA
	'8w_ctrl_1'	'treatment: control'	WT
	'8w_ctrl_2'	'treatment: control'	WT
	'8w_ctrl_3'	'treatment: control'	WT
	'8w_dmm_1'	'treatment: destabilisation of the medial meniscus (DMM)'	OA
	'8w_dmm_2'	'treatment: destabilisation of the medial meniscus (DMM)'	OA
	'8w_dmm_3'	'treatment: destabilisation of the medial meniscus (DMM)'	OA
GSE33656	GSM832368	6-1 tibia Bl6	WT
	GSM832370	8-1 tibia Bl6	WT
	GSM832372	9-1 tibia Bl6	WT
	GSM832369	6-tibia AC FRZB KO	OA
	GSM832371	14-tibia AC FRZB KO	OA
	GSM832373	17-tibia AC FRZB KO	OA
GSE45793	SHAM_GSM1120018_1WK_REP1	Sham_1wk	WT
	DMM_GSM1120018_1WK_REP1	DMM_1wk	OA
	SHAM_GSM1120019_1WK_REP2	Sham_1wk	WT
	DMM_GSM1120019_1WK_REP2	DMM_1wk	OA
	SHAM_GSM1120020_1WK_REP3	Sham_1wk	WT
	DMM_GSM1120020_1WK_REP3	DMM_1wk	OA
	SHAM_GSM1120021_2WK_REP1	Sham_2wk	WT
	DMM_GSM1120021_2WK_REP1	DMM_2wk	OA
	SHAM_GSM1120022_2WK_REP2	Sham_2wk	WT
	DMM_GSM1120022_2WK_REP2	DMM_2wk	OA
	SHAM_GSM1120023_2WK_REP3	Sham_2wk	WT
	DMM_GSM1120023_2WK_REP3	DMM_2wk	OA
	SHAM_GSM1120024_2WK_REP4	Sham_2wk	WT
	DMM_GSM1120024_2WK_REP4	DMM_2wk	OA
	SHAM_GSM1120025_6WK_REP1	Sham_6wk	WT
	DMM_GSM1120025_6WK_REP1	DMM_6wk	OA
	SHAM_GSM1120026_6WK_REP2	Sham_6wk	WT
	DMM_GSM1120026_6WK_REP2	DMM_6wk	OA
	SHAM_GSM1120027_6WK_REP3	Sham_6wk	WT
	DMM_GSM1120027_6WK_REP3	DMM_6wk	OA
	SHAM_GSM1120028_6WK_REP4	Sham_6wk	WT
	DMM_GSM1120028_6WK_REP4	DMM_6wk	OA

Appendix of Chapter 4

Supplementary text S4.1: The full system of equations is provided in two parts: (a) the fast reactions, (b) the slow reactions, the variable index refers to the index provided in **Table S2**.

a) Fast reactions (protein signaling network)

$$z_1^f(t+1) = 1 - z_{49}(t) - z_{48}(t)$$

$$z_2^f(t+1) = z_{50}(t) - 0.3 * z_4(t)$$

$$z_3^f(t+1) = 1$$

$$z_4^f(t+1) = z_{52}(t) - 0.5 * z_{25}(t) + z_{52}(t)$$

$$z_5^f(t+1) = 1$$

$$z_6^f(t+1) = \frac{2}{3} * z_5(t)$$

$$z_7^f(t+1) = 1 - z_{36}(t)$$

$$z_8^f(t+1) = z_7(t) - z_{10}(t)$$

$$z_9^f(t+1) = z_{38}(t) + z_{22}(t) + z_{32}(t) - z_{10}(t) - z_{14}(t) - z_{26}(t) \times z_{30}(t) - z_{31}(t) - z_{43}(t)$$

$$z_{10}^f(t+1) = z_{14}(t) + z_{26}(t) - z_7(t) - z_{19}(t)$$

$$z_{11}^f(t+1) = 1$$

$$z_{12}^f(t+1) = z_{11}(t)$$

$$z_{13}^f(t+1) = 1$$

$$z_{14}^f(t+1) = \big(z_{12}(t) + z_3(t)\big) \times s_2$$

$$z_{15}^f(t+1) = z_{34}(t) + z_{19}(t) - z_{30}(t) - z_{26}(t)$$

$$z_{16}^{f}(t+1) = 1$$

$$z_{17}^{f}(t+1) = z_{16}(t)$$

$$z_{18}^{f}(t+1) = \frac{2}{3} \times \left(\frac{2}{3} \times z_{17}(t) + \frac{2}{3} \times z_{27}(t)\right) + \frac{1}{3} \times z_{42}(t) - z_{19}(t) + z_{54}(t) \times (1 - z_{17}(t))$$

$$z_{19}^{f}(t+1) = z_{4}(t) - 0.25 \times z_{22}(t)$$

$$z_{20}^{f}(t+1) = 1$$

$$z_{21}^{f}(t+1) = 1$$

$$z_{22}^{f}(t+1) = z_{41}(t) - z_{37}(t) + z_{55}(t)$$

$$z_{23}^{f}(t+1) = 1$$

$$z_{24}^{f}(t+1) = 1$$

$$z_{25}^{f}(t+1) = 1$$

$$z_{26}^{f}(t+1) = z_{53}(t) - 0.5 \times \Big(\big(z_{22}(t) + z_{31}(t) + z_{25}(t)\big) \times s_3\Big)$$

$$z_{27}^{f}(t+1) = z_{16}(t)$$

$$z_{28}^{f}(t+1) = \big(z_{34}(t) + z_{34}(t) \times z_{26}(t)\big) \times s_2$$

$$z_{29}^{f}(t+1) = \big(z_{27}(t) + z_{38}(t) + z_{55}(t) - z_{58}(t)\big) \times s_2$$

$$z_{30}^{f}(t+1) = z_{37}(t) - z_{35}(t)$$

$$z_{31}^{f}(t+1) = 1 - 0.5 \times \big(z_{35}(t) + z_{18}(t)\big)$$

$$z_{32}^{f}(t+1) = z_{34}(t) + z_{19}(t) - z_{43}(t)$$

$$z_{33}^{f}(t+1) = 1 - z_{47}(t)$$

$$z_{34}^{f}(t+1) = \left(z_{23}(t) + z_{33}(t) + z_{55}(t)\right) \times s_3$$

$$z_{35}^{f}(t+1) = 1 - z_{14}(t) - z_{38}(t)$$

$$z_{36}^{f}(t+1) = 1 + z_{37}(t) - z_{2}(t) - 0.5 \times z_{22}(t)$$

$$z_{37}^{f}(t+1) = z_{14}(t)$$

$$z_{38}^{f}(t+1) = z_{39}(t) - 0.5 \times z_{37}(t)$$

$$z_{39}^{f}(t+1) = \left(z_{41}(t) + z_{42}(t) + z_{54}(t)\right) \times s_3$$

$$z_{40}^{f}(t+1) = \left(z_{22}(t) + z_{56}(t)\right) \times s_2$$

$$z_{41}^{f}(t+1) = \left(z_{1}(t) + z_{33}(t) + z_{17}(t) + z_{27}(t)\right) \times s_4$$

$$z_{42}^{f}(t+1) = z_{3}(t)$$

$$z_{43}^{f}(t+1) = 1 - z_{32}(t)$$

$$z_{44}^{f}(t+1) = 1$$

$$z_{45}^{f}(t+1) = \left(z_{14}(t) + z_{22}(t)\right) \times s_2$$

$$z_{46}^{f}(t+1) = 1$$

$$z_{47}^{f}(t+1) = 1$$

$$z_{48}^{f}(t+1) = 1$$

$$z_{49}^{f}(t+1) = 1$$

$$z_{50}^{f}(t+1) = z_{1}(t) - z_{49}(t)$$

$$z_{51}^{f}(t+1) = 1$$

$$z_{52}^{f}(t+1) = z_{23}(t) - z_{33}(t)$$

$$z_{53}^{f}(t+1) = z_{23}(t)$$

$$z_{54}^{f}(t+1) = z_{51}(t) - z_{59}(t)$$

$$z_{55}^{f}(t+1) = z_{54}(t) - 0.5 \times z_{34}(t)$$

$$z_{56}^{f}(t+1) = z_{55}(t)$$

$$z_{57}^{f}(t+1) = 1 - z_{24}(t)$$

$$z_{58}^{f}(t+1) = 1 - z_{55}(t)$$

$$z_{59}^{f}(t+1) = 1$$

$$z_{60}^{f}(t+1) = 0.75 + z_{56}(t) - z_{38}(t) - z_{22}(t)$$

b) Slow reactions (gene regulatory network)

$$z_{1}^{s}(t+1) = 2 \times z_{6}(t) - z_{43}(t)$$

$$z_{2}^{s}(t+1) = 1$$

$$z_{3}^{s}(t+1) = z_{14}(t)$$

$$z_{4}^{s}(t+1) = 1 - z_{45}(t) + z_{44}(t)$$

$$z_{5}^{s}(t+1) = \left(z_{9}(t) + z_{19}(t) + z_{29}(t) + z_{45}(t) - z_{44}(t) - z_{17}(t)\right) \times s_{2}$$

$$z_{6}^{s}(t+1) = 1 - z_{17}(t)$$

$$z_{7}^{s}(t+1) = 1$$

$$z_{8}^{s}(t+1) = \left(1 + z_{9}(t) + z_{7}(t)\right) \times s_{3}$$

$$z_{9}^{s}(t+1) = \left(z_{46}(t) + z_{8}(t) + z_{9}(t) + z_{15}(t) + z_{32}(t) - z_{21}(t) \times z_{19}(t) - z_{43}(t) + z_{6}(t) - z_{14}(t) + z_{40}(t)\right) \times s_{4}$$

$$z_{10}^{s}(t+1) = \left(z_{14}(t) - z_{29}(t) + z_{34}(t) + z_{29}(t) + z_{10}(t) + z_{21}(t)\right) \times s_{4}$$

$$z_{11}^{s}(t+1) = \big(z_6(t) \times z_{10}(t) + z_{10}(t) + z_{26}(t)\big) \times s_3$$

$$z_{12}^{s}(t+1) = \big(z_6(t) + z_{10}(t) + z_{19} - z_{18}(t)\big) \times s_2$$

$$z_{13}^{s}(t+1) = \big(z_9(t) + z_4(t) + z_{15}(t) - z_{14}(t) - z_{11}(t) + z_{46}(t) - z_{18}(t) + z_7(t)\big) \times s_2$$

$$z_{14}^{s}(t+1) = 1$$

$$z_{15}^{s}(t+1) = \big(z_9(t) + z_{19}(t)\big) \times s_2$$

$$z_{16}^{s}(t+1) = \big(z_7(t) + z_9(t)\big) \times s_2$$

$$z_{17}^{s}(t+1) = z_{10}(t) - z_{40}(t) \times z_{22}(t)$$

$$z_{18}^{s}(t+1) = 1$$

$$z_{19}^{s}(t+1) = 1$$

$$z_{20}^{s}(t+1) = \big(z_{10}(t) + z_{27}(t) + z_{21}(t) - z_7(t) + z_{60}(t)\big) \times s_3$$

$$z_{21}^{s}(t+1) = \big(z_{10}(t) + z_{14}(t)\big) \times s_2$$

$$z_{22}^{s}(t+1) = z_{15}(t)$$

$$z_{23}^{s}(t+1) = z_6(t) - z_{40}(t) + z_{26}(t)$$

$$z_{24}^{s}(t+1) = z_9(t) + z_{29}(t) + z_{46}(t) \ + z_6(t) + z_{46}(t) \times z_9(t) \times z_{34}(t) + z_{19}(t) + z_{40}(t) - z_{26}(t) + z_{40}(t) \times z_{28}(t) \times z_{56}(t) \ - \frac{1}{5} z_{60}(t)$$

$$z_{25}^{s}(t+1) = \big(z_{18}(t) + z_{19}(t) + z_{26}(t) + 2 \times z_{29}(t)\big) \times s_5$$

$$z_{26}^{s}(t+1) = 1$$

$$z_{27}^{s}(t+1) = z_9(t) - 0.5 \times z_{19}(t)$$

$$z_{29}^{s}(t+1) = 1$$

$$z_{30}^{s}(t+1) = 1$$

$$z_{31}^{S}(t+1) = \left(z_{28}(t) + z_{6}(t) + 1.5 \times z_{14}(t) + z_{60}(t) \times z_{26}(t)\right) \times s_4$$

$$z_{32}^{S}(t+1) = z_{15}(t) + z_{19}(t) - z_{26}(t)$$

$$z_{33}^{S}(t+1) = \left(z_{6}(t) + z_{29}(t)\right) \times s_2$$

$$z_{34}^{S}(t+1) = 1 - 0.4 \times z_{44}(t)$$

$$z_{35}^{S}(t+1) = 1$$

$$z_{36}^{S}(t+1) = 1$$

$$z_{37}^{S}(t+1) = \left(1 - z_{7}(t)\right) \times s_2$$

$$z_{38}^{S}(t+1) = z_{9}(t)$$

$$z_{39}^{S}(t+1) = z_{9}(t)$$

$$z_{40}^{S}(t+1) = 1$$

$$z_{41}^{S}(t+1) = 1$$

$$z_{42}^{S}(t+1) = \left(z_{26}(t) + z_{29}(t) + z_{18}(t)\right) \times s_3$$

$$z_{43}^{S}(t+1) = \left(z_{19}(t) + z_{26}(t) + z_{43}(t) - z_{32}(t)\right) \times s_2$$

$$z_{44}^{S}(t+1) = \left(z_{26}(t) + z_{29}(t) + z_{40}(t) - z_{19}(t) + z_{4}(t)\right) \times s_3$$

$$z_{45}^{S}(t+1) = 1$$

$$z_{46}^{S}(t+1) = z_{29}(t)$$

$$z_{47}^{S}(t+1) = z_{10}(t) - z_{7}(t)$$

$$z_{48}^{S}(t+1) = 0.5 \times z_{19}(t) - z_{7}(t) + z_{10}(t) - z_{29}(t) - z_{28}(t)$$

$$z_{49}^{S}(t+1) = 0.5 \times z_{19}(t) - z_{7}(t) + z_{10}(t) - z_{29}(t) - z_{28}(t)$$

$$z_{50}^{S}(t+1) = 1$$

$$z_{51}^{S}(t+1) = z_{40}(t) \times z_{28}(t) \times z_{56}(t) + z_{29}(t)$$

$$z_{52}^{S}(t+1) = \big(0.4 + z_{18}(t) - z_{26}(t) + z_{9}(t) + z_{28}(t) \times z_{56}(t)\big) \times s_3$$

$$z_{53}^{S}(t+1) = \big(0.4 + z_{26}(t) - z_{29}(t) + z_{10}(t)\big) \times s_2$$

$$z_{54}^{S}(t+1) = 1 - z_{60}(t)$$

$$z_{55}^{S}(t+1) = 1$$

$$z_{56}^{S}(t+1) = 1$$

$$z_{57}^{S}(t+1) = \big(z_{10}(t) + z_{60}(t) - z_{29}(t)\big) \times s_3$$

$$z_{58}^{S}(t+1) = z_{29}(t)$$

$$z_{59}^{S}(t+1) = \big(z_{42}(t) + z_{17}(t)\big) \times s_3$$

$$z_{60}^{S}(t+1) = z_{26}(t) - z_{40}(t) \times z_{28}(t) - z_{29}(t)$$

The additive sum includes a weight for each term that assumes a limited amount of saturation whenever there is a majority of stimulatory interactions. *SC* is the saturation constant, set to 2/3, and p_i the number of excess positive interactions upstream of node *i*, the saturation is defined as follows:

$$s_p^i = \begin{cases} 1, if\ p_i \leq 1 \\ 2\dfrac{SC}{p_i}, if\ p_i > 1 \end{cases}$$

If the positive interactions do not outnumber the inhibitory interactions by more than 1, all weights are set to 1. If the number of excess positive interactions is higher than 1, we introduce the a saturation factor whose value will determine how fast the node will saturate.

Supplementary text S4.2

Mouse oligonucleotides used for qRT-PCR reaction with ATDC5 samples. Primers were designed by the Primer Design tool of NCBI:

Gene	Forward primer	Reverse primer
mHrpt	GAGCGTTGGGCTTACCTCAC	ATCGTAATCACGACGCTGG
mCol10a1	TCCCAGCACCAGAATCTATCTGA	TTATGCCTGTGGGCGTTTGG

RT-qPCR protocol for samples from human primary OA chondrocytes cultured in alginate beads: Alginate beads were dissolved using citrate buffer, centrifuged at 200g and the pellet was resuspended in RLT (Qiagen, Hilden, Germany) buffer containing 1% beta-mercaptoethanol for RNA isolation. mRNA isolation was performed according to manufacturer's protocol utilizing the RNeasy Column system (Qiagen, Hilden, Germany). The RNA concentration was determined using a NanoDrop spectrophotometer (Isogen Life Science, Utrecht, the Netherlands). 0.5 µg RNA was used for cDNA synthesis following the protocol of the manufacturer of the RevertAid First Strand cDNA kit (Thermo Fisher Scientific, Waltham, MA, United States). qPCR was performed on a Bio-Rad CFX96 Real-Time PCR Detection System (Bio-Rad) to assess gene expression, Collagen type 10 (*COL10A1*) and Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), which was found stable and therefore used as reference gene. Data were analyzed by the ΔΔCt method and normalized to the expression of *GAPDH* of each condition and compared to the corresponding gene expression in the control groups

Human oligonucleotides used for qRT-PCR reaction with primary chondrocytes.

Gene	Forward primer	Reverse primer	Probe
GAPDH	ATGGGGAAGGTGAAGGTCG	TAAAAGCAGCCCTGGTGACC	CGCCCAATACGACCAAATCCGTTGAC
COL10A1	CAAGGCACCATCTCCAGGAA	AAAGGGTATTTGTGGCAGCATATT	TCCAGCACGCAGAATCCATCTGA

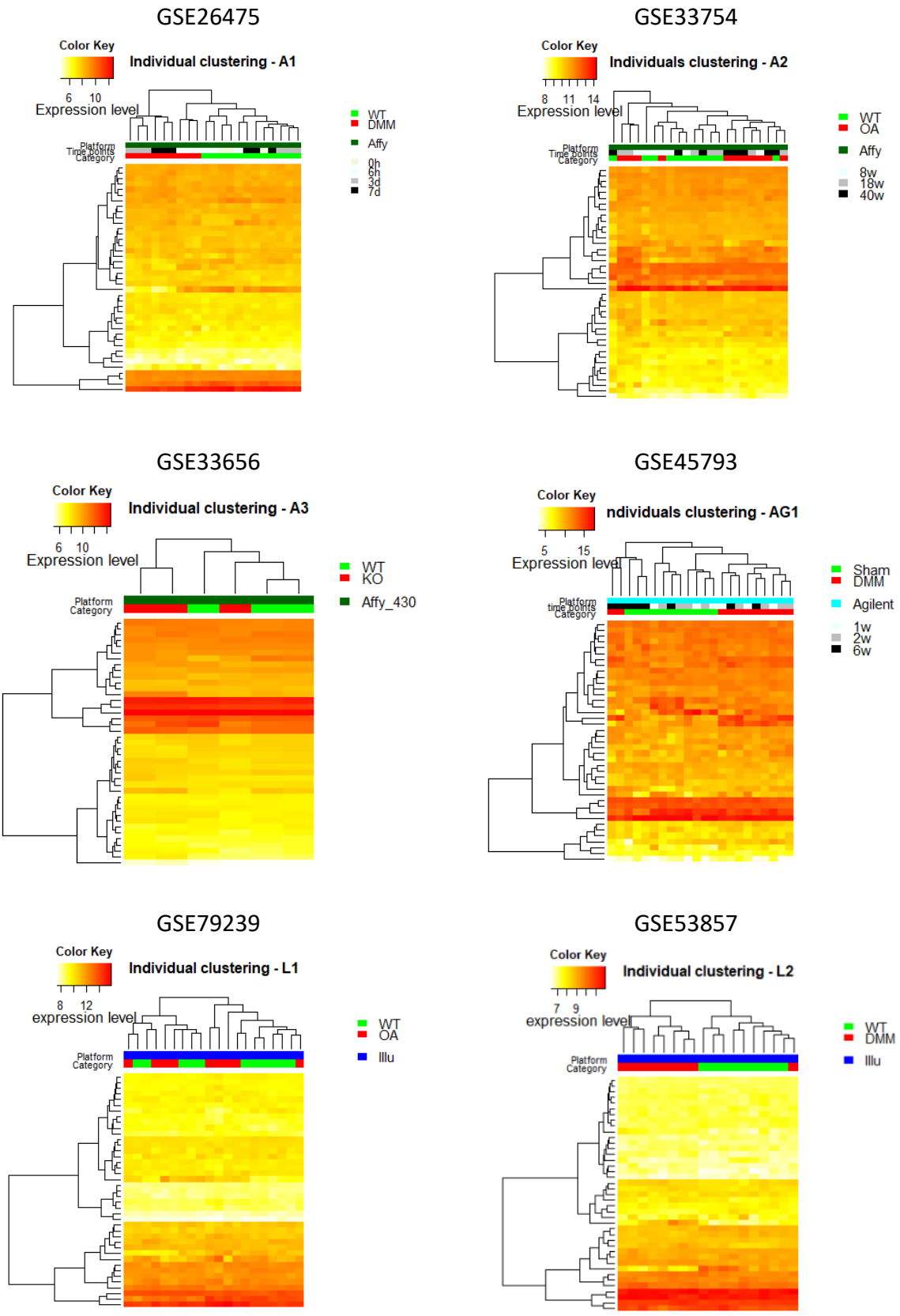


Figure. S4.1. Heatmap with unsupervised clustering on individual dataset.

The microarray sub-datasets, used in the merged dataset, were also investigated individually. They were pre-processed (i.e. processing steps before merging and correcting for batch effect) in the same way as the merged dataset but the unsupervised clustering analysis was done on each sub-dataset separately in order to compare the biological or OA related information content before and after the data merging. The heatmaps show the expression profile of the same list of genes of interest than for the merged dataset. These expression datasets were submitted to unsupervised clustering with the Euclidean distance method and the Complete aggregation method in R thanks to the heatmap3 function from the github repository https://github.com/obigriffith/biostar-tutorials/tree/master/Heatmaps. The headers indicate the GEO accession numbers of the 6 original datasets. Samples labeled as 'WT', for wild type, are in green, samples labeled as 'OA', for osteoarthritis, in red. The pre-labelling is the same as for the merged dataset **Table S3.1**. When applicable, a grey scale indicate the time points (w stands for weeks, in the legend). For some datasets (e.g. GSE26475, GSE33656, GSE53857) the OA and the WT samples are well separated in different clusters, while for other the separation is not so clear. For instance, in GSE45793 the variance due the time point 6weeks is greater than the OA induced variance, while for weeks 1 and 2 OA and WT samples are well separated due to the OA condition.

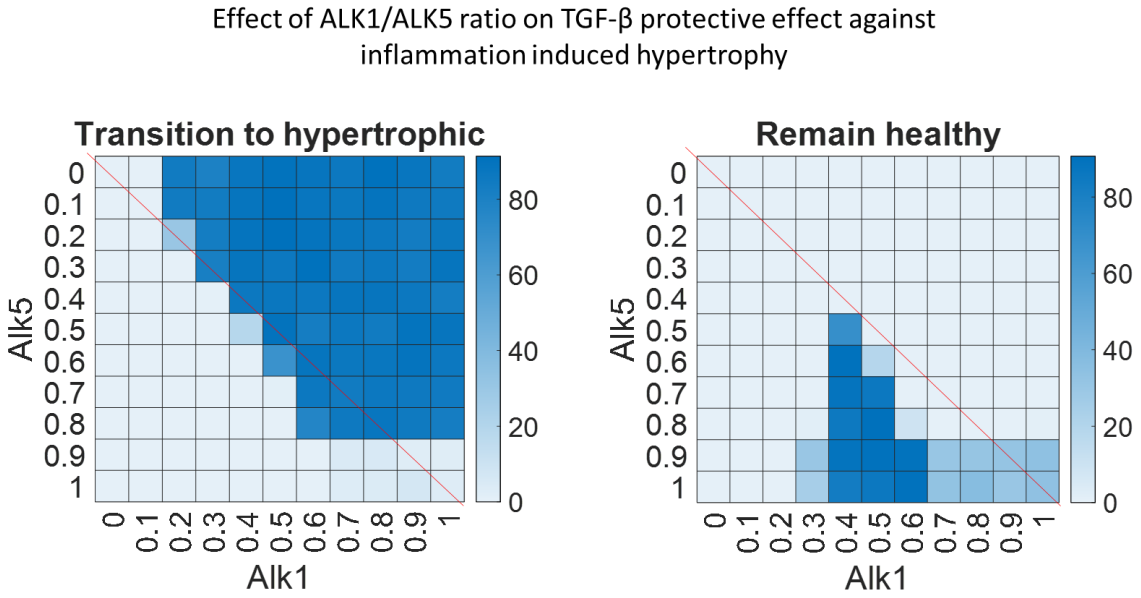


Fig. Figure S4.2. Role of the ratio between receptors ALK1 and ALK5 in the effect of inflammation and TGFβ signaling on chondrocyte hypertrophy. Percentage of perturbations remaining in the healthy state (right) or transitioning towards the hypertrophic one (left) during inflammatory pathway activation with TGF-B treatment while changing the ratio between ALK1 and ALK5. The inflammatory and TGF-B profiles that were imposed are the same as in Fig.4A except that the value imposed for ALK1 and ALK5 are varying between 0 and 1 with a 0.1 increment. ALK1=ALK5 on the red diagonals and ALK1>ALK5 in the upper right corner. Roughly, the rescue of the healthy state by TGF-B is lost when ALK1 is greater than ALK5. If ALK1 is high enough and that the difference between ALK1 and ALK5 is not greater than 10-20% then the protective effect of TGF-B is also disturbed for ALK1<ALK5

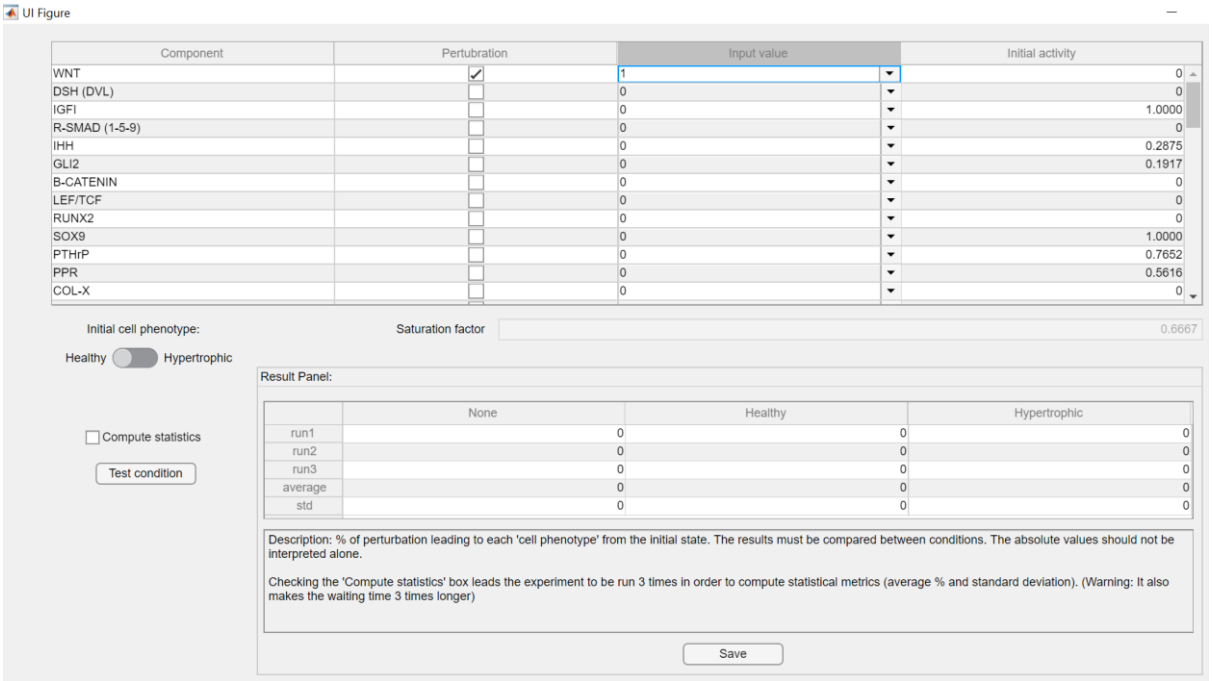


Figure. S4.3. Screenshot of the user-friendly interface for the virtual chondrocytes App. The standalone Matlab-based applications can be launched and used without Matlab license, provided that the compiler Matlab Runtime is installed (https://nl.mathworks.com/products/compiler/matlab-runtime.html (*337*)). The virtual chondrocyte initial state can be set as healthy or hypertrophic, allowing the user to test any scenarios. All the 60 components may be perturbed alone or in any sort of combination by forcing the variables to take a value in the interval [0:1], with a step of 0.1. The most left column indicate the value of the variable in the selected initial state, for information. Obviously, applying a perturbation that is equal to the initial value of the variable will not affect the system. Once the setting are done, the user can apply the experimental condition by pushing the button 'Test condition' and the percentage of transition towards each of the possible basal stable states (i.e. 'None', 'Healthy' and 'Hypertrophic') is computed. If the 'Compute statistics' box is ticked, then the experiment is repeated 3 times and the average and standard deviations are displayed (variation occurs due to the stochastic nature of the model). The results may be exported and saved in an excel file via the 'Save' button.

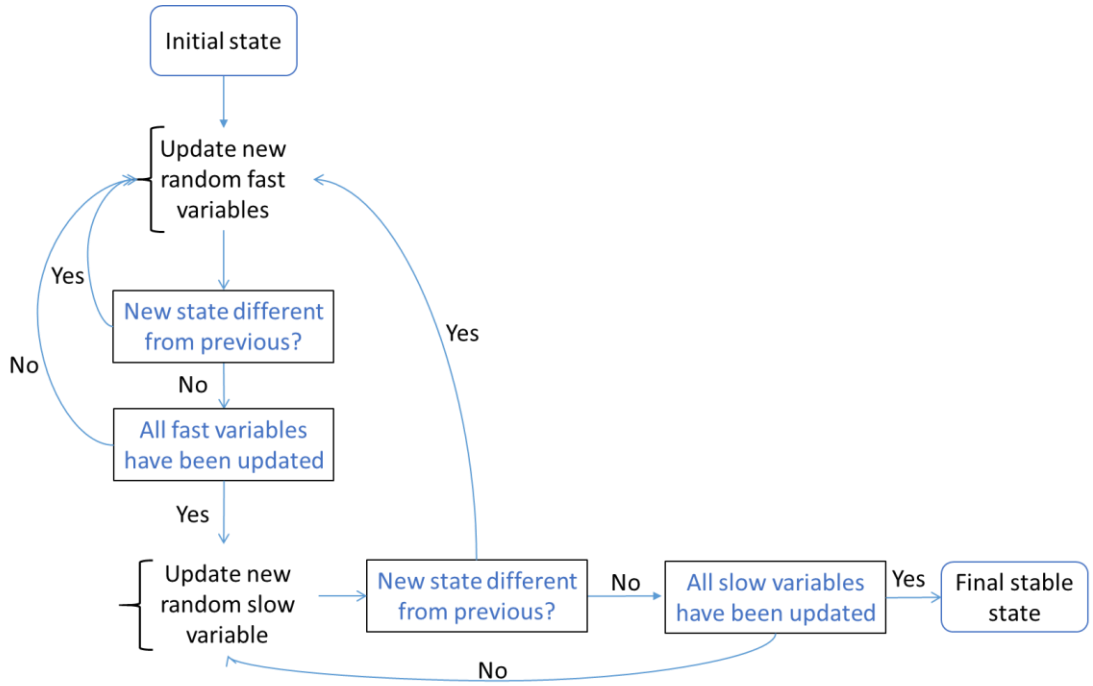


Figure S4.4. Decision tree summarizing the variable updating scheme employed in algorithm to simulate the *in silico* chondrocyte. Each biological component is represented by the gene expression level (slow variable) and the protein activity potential (fast variable). Variables are updated based on the rules stored in the model's equations. First, fast variables are updated in random order, when a pseudo-stable state is reached and that all fast variables have been updated, the next random chosen slow variable is updated. This goes on until a state that is stable both at the fast and slow level is reached. This is the final stable state. A state is considered stable if further variable updates do not bring further changes for any of the variables, more or less a predefined tolerance. The order in which variables are updated is random, thereby generating some stochasticity in the model. Within the fast (resp. slow) updating loops, variables are updated asynchronously (meaning the one after the others) according to the rules defined in the system of equations (see **Supplementary text S4.1**) and in a random order.

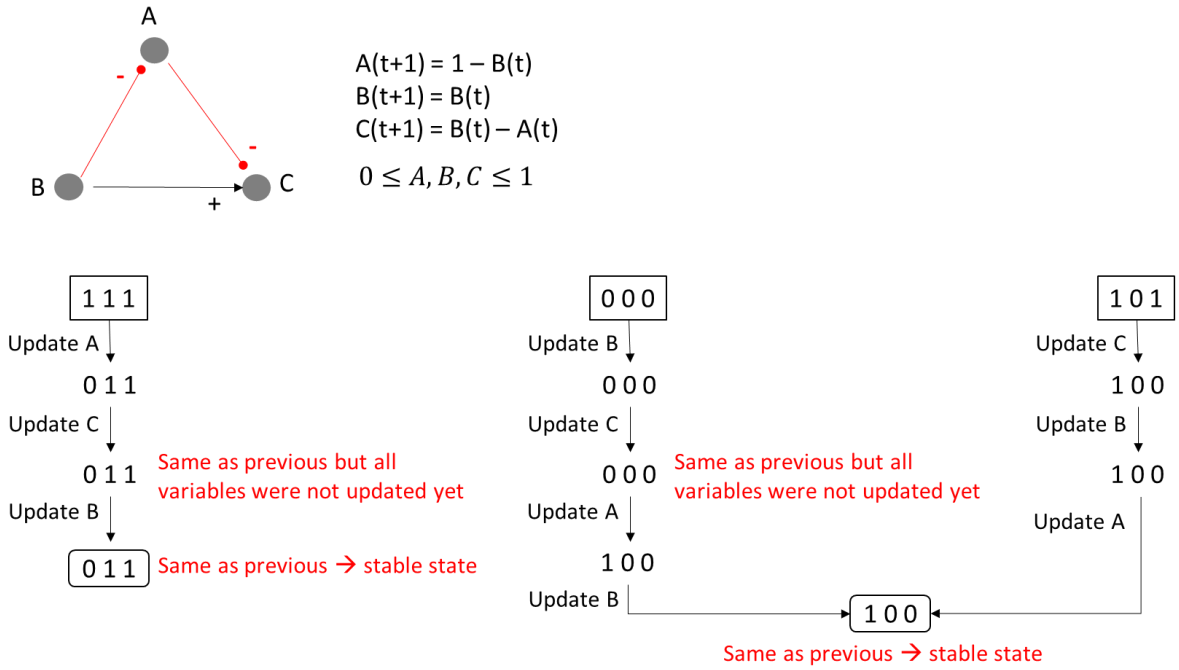


Figure S4.5. Illustration of the algorithm for the asynchronous updating of variables with a simplified example network. The network represents interactions happening in one of the layers (protein= fast reactions or genetic = slow reactions). Inhibitory influences are represented in Red and activator influences are in black. The mathematical rules corresponding to the network are displayed. If the rules result in a value lower than 0 (resp. higher than 1), the value is brought back to 0 (resp. 1). For this example, 3 different initial states are inputted and each of the three variables is updated asynchronously. The order in which variables are updated is random. The system reaches a stable state when the next update gives the same state as in the previous time step and that all variables were screened in the random ordering list. That state is a pseudo-stable state if the rules were describing fast reactions, in that case, a new slow variable can be updated (see Fig. S3.). It is a final stable state if the rules were describing slow interactions since it would mean that the system had first reached a pseudo-stable state at the fast level and would now be stable at the slow level too.

Table S4.1. Table of definitions. Modeling terms and concepts relative to the current additive regulatory network model are defined.

Definitions
Additive model: A network of protein or gene interactions is modeled with an additive method if the evolution of variables (proteins or genes) at the next time step is defined by the sum of the upstream activating variables and the subtraction of the upstream inhibitory variables of the network. For instance, in a GRN, if a transcription factor A activates a gene P while B inhibits gene P, then the evolution of the P expression level at the next time step is defined as: $P(t+1) = A(t) - B(t)$ (eq. 1)
Node or component (variable): The nodes of a regulatory network are located at the intersections of multiple interactions (edges) in the network. They represent biological components such as proteins or genes. In mathematical models, such as the additive models, components' evolutions are described with variables (e.g. A, P or B in (eq.1)).
Fast and Slow reactions & variables: All reactions related to slow biological processes such as gene expression, mRNA or protein production, were referred to as slow reactions (lower priority) and those related to fast processes such as protein activation (e.g. post transcriptional modification) or degradation, were referred to as fast reactions (higher priority). This priority order plays a role in the simulation since any network component could be regulated both at the protein and gene level, as in real life. Therefore, each network component or variable was split into a fast and a slow subpart, also called sub-variables. Let's consider that the gene P from (eq.1) produces a protein that is both activated post-transcriptionally by another protein kinase K and blocked by an inhibitory protein I. The global functional activity of the protein P is actually the multiplication of the slow by the fast subparts: $P(t+1) = [A(t) - B(t)] \times [K(t) - I(t)]$. (eq2) When simulating the system, the sub-variables are updated asynchronously following the priority classes in such a way that fast reactions are always updated before the slow reactions (*255*).
Stable state (or attractors): A stable state is an ensemble of values, one for each variable/component of the model, that meet all the rules/constraints imposed by the equations. When variables do not evolve anymore after a certain time of simulation, the system has reached a stable state. The nature of such a state depends on the system of equations and on which initial state was used. Given the ensemble of interconnected signaling pathways, it is likely that only a finite number of states can fulfil all the constraints imposed by the network structure (i.e. by the equations).

Table S4.2. Table of correspondence between the variable names in the model corresponding real world mouse gene. All mathematical variables and the corresponding node in the network have names written in upper cases and do not reflect the official human or mouse nomenclature. To relate those variables to actual genes more easily, we provide this table of correspondence. A related mouse gene name and NCBI ID is indicated for each variable. Nevertheless, it is not exhaustive since some variables represents a group of factors or a family of ligands rather than a single factor.

Variable index	Variable name	Example representative mouse gene	
		Name	NCBI ID
1	WNT	*Wnt3a*	22416
2	DSH	*Dvl1*	20423
3	IGF-I	*Igf1*	16000
4	R-SMAD	*Smad5*	17129
5	IHH	*Ihh*	16147
6	GLI2	*Gli2*	14633
7	β-Catenin	*Ctnnb1*	12387
8	LEF/TCF	*Tcf7*	21414
9	RUNX2	*Runx2 or Cbfa1*	12393
10	SOX9	*Sox9*	20682
11	PTHrP	*Pthlh*	19227
12	PPR	*Pth1r*	19228
13	COL-X	*Col10a1*	12813
14	PKA	*Prkaca*	18747
15	MEF2C	*Mef2c*	17260
16	FGF	*Fgf2*	14173
17	FGFR3	*Fgfr3*	14184
18	STAT1	*Stat1*	20846
19	Smadcomplex	*Smad4 & R-Smads*	17128
20	COL II	*Col2a1*	12824
21	NKX3.2	*Nkx3-2 or Bapx1*	12020
22	ERK1/2	*Mapk3*	26417
23	TGFβ	*Tgfb1*	21803
24	MMP13	*Mmp13*	17386
25	SMAD7	*Smad7*	17131
26	SMAD3	*Smad3*	17127
27	FGFR1	*Fgfr1*	14182
28	ATF2	*Atf2*	11909
29	NFκB	*Nfkb1*	18033
30	HDAC4	*Hdac4*	208727
31	CCND1	*Ccnd1*	12443
32	DLX5	*Dlx5*	13395
33	BMP	*Bmp2*	12156

34	P38	*Mapk14*	26416
35	GSK3β	*Gsk3b*	56637
36	DC	*Apc*	11789
37	PP2A	*Ppp2ca*	19052
38	AKT	*Akt1*	11651
39	PI3K	*Pi3kr1*	18708
40	ETS1	*Ets1*	23871
41	RAS	*Kras*	16653
42	IGF-IR	*Igf1r*	16001
43	MSX2	*Msx2*	17702
44	δEF-1	*Zeb1*	21417
45	ATF4	*Atf4 or Creb*	11911
46	HIF-2α	*Epas1*	13819
47	GREM1	*Grem1*	23892
48	DKK1	*Dkk1*	13380
49	FRZB	*Frzb*	20378
50	Frizzled-LRP5/7	*Lrp5*	16973
51	Cytokines	*Il1b*	16176
52	ALK1	*Acvrl1*	11482
53	ALK5	*Tgrb1*	21812
54	R-Infl (Receptor inflammation)	*Tlr1*	21897
55	TAK1	*Map3k7*	26409
56	JNK	*Mapk8*	26419
57	Proteoglycans	*Acan*	11595
58	IkB-a	*Nfkbi*	18035
59	SOCS	*Socs1*	12703
60	FOXO1	*Foxo1*	56458

Table S4.3 Full profiles of the three computed attractors.

	Attractor1			Attractor2			Attractor3		
	global	protein	gene	global	protein	gene	global	protein	gene
WNT	0.00	1.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00
DSH	0.00	0.00	1.00	0.00	0.00	1.00	0.70	0.70	1.00
IGFI	0.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	0.00
R-SMAD	0.00	0.00	1.00	0.00	0.00	0.57	1.00	1.00	1.00
IGG	0.00	1.00	0.00	0.29	1.00	0.29	1.00	1.00	1.00
GLI2	0.00	0.00	1.00	0.19	0.19	1.00	0.67	0.67	1.00
B-Catenin	0.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00
LEF/TCF	0.00	0.00	0.44	0.00	0.00	0.44	1.00	1.00	1.00
RUNX2	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00
SOX9	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.16
PTHrP	0.00	1.00	0.00	0.77	1.00	0.77	0.00	1.00	0.00
PPR	0.00	0.00	0.00	0.56	0.77	0.73	0.00	0.00	0.59
COL-X	0.00	1.00	0.00	0.00	1.00	0.00	1.00	1.00	1.00
PKA	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00
MEF2C	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00
FGF	0.00	1.00	0.00	0.00	1.00	0.00	1.00	1.00	1.00
FGFR3	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00
STAT1	0.00	0.00	1.00	0.09	0.09	1.00	0.53	0.53	1.00
SMAD complex	0.00	0.00	1.00	0.00	0.00	1.00	0.75	0.75	1.00
COL-II	0.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	0.00
NKX3.2	0.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	0.00
ERK1/2	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00
TGFβ	0.00	1.00	0.00	0.72	1.00	0.72	0.00	1.00	0.00
MMP13	0.00	1.00	0.00	0.00	1.00	0.00	1.00	1.00	1.00
SMAD7	0.00	1.00	0.00	0.16	1.00	0.16	0.87	1.00	0.87
SMAD3	0.00	0.00	1.00	0.53	0.53	1.00	0.00	0.00	1.00
FGFR1	0.00	0.00	0.00	0.00	0.00	0.00	0.63	1.00	0.63
ATF2	0.00	0.00	1.00	0.30	0.30	1.00	0.31	0.31	1.00
NFkB	0.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00
HDAC4	0.00	0.00	1.00	0.67	0.67	1.00	0.00	0.00	1.00
CCND1	0.00	0.50	0.00	0.70	0.95	0.73	0.21	0.66	0.33
DLX5	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00
BMP	0.00	1.00	0.00	0.00	0.00	0.13	1.00	1.00	1.00
P38	0.00	0.00	1.00	0.29	0.32	0.91	0.47	0.78	0.60
GSK	1.00	1.00	1.00	0.00	0.00	1.00	0.16	0.16	1.00
DC	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00
PP2A	0.00	0.00	0.67	0.67	1.00	0.67	0.00	0.00	0.00
AKT	0.00	0.00	0.00	0.00	0.00	0.00	0.83	0.83	1.00
PI3K	0.00	0.00	0.00	0.00	0.12	0.00	0.83	0.83	1.00
ETS1	0.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00
RAS	0.00	0.00	1.00	0.00	0.00	1.00	0.88	0.88	1.00
IGF-IR	0.00	0.00	0.00	0.28	1.00	0.28	0.00	0.00	0.68
MSX2	0.01	1.00	0.01	1.00	1.00	1.00	0.00	0.00	0.00
EF-1	0.00	1.00	0.00	0.24	1.00	0.24	1.00	1.00	1.00
ATF4	0.00	0.00	1.00	0.67	0.67	1.00	0.67	0.67	1.00
HIF-2α	0.00	1.00	0.00	0.00	1.00	0.00	1.00	1.00	1.00
GREM1	0.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	0.00
FRZB	0.00	1.00	0.00	0.70	1.00	0.70	0.00	1.00	0.00
DKK1	0.00	1.00	0.00	0.70	1.00	0.70	0.00	1.00	0.00
Frizzled-LRP5/6	0.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00
Cytokines	0.00	1.00	0.00	0.00	1.00	0.00	1.00	1.00	1.00
ALK1	0.00	0.00	0.18	0.00	0.72	0.00	1.00	1.00	1.00
ALK5	0.00	0.00	0.27	0.72	0.72	1.00	0.00	0.00	0.00
R_inflamm	0.00	0.00	1.00	0.00	0.00	0.60	1.00	1.00	1.00
TAK1	0.00	0.00	1.00	0.00	0.00	1.00	0.76	0.76	1.00
JNK	0.00	0.00	1.00	0.00	0.00	1.00	0.76	0.76	1.00
Proteoglycans	0.00	1.00	0.00	0.62	1.00	0.62	0.00	0.00	0.00
IkBa	0.00	1.00	0.00	0.00	1.00	0.00	0.24	0.24	1.00
SOCS	0.00	1.00	0.00	0.18	1.00	0.18	0.00	1.00	0.00
FOXO1	0.00	0.75	0.00	0.40	0.75	0.53	0.00	0.00	0.00

Appendix of Chapter 5

Supplementary text S5.1

Cell culture

Human pluripotent stem cell lines H9, 604B1, CY2 and NCRM1 were cultured as previously described (*266*). Briefly, hPSCs were maintained on mitomycin-c (Sigma) treated SNL feeder cells in human embryonic stem cell medium consisting of Dulbecco's modified eagle medium (DMEM): Ham's F12 nutrient mix (F12) (Sigma) supplemented with 20% knockout serum replacement, 2mM Glutamax, 1% sodium pyruvate (SP), 1% non-essential amino acids (NEAA), 0.1 mM 2-mercaptoethanol (2ME), 50 U and 50 mg/ml penicillin/streptomycin (Pen/Strep) and 10 ng/ml human basic fibroblast growth factor (hbFGF). Cells were grown in a humidified incubator at 37°C and 5% CO_2. Medium was replenished daily and hPSCs were passaged once a week on fresh feeder cells. H9 and 604B1 were kindly provided by the Professors Sampaolesi (KU Leuven, Belgium) and Tsumaki (CiRA, Japan) respectively. CY2 and NCRM1 were purchased from Rutgers University Cell and DNA Repository. All components/reagents except for the growth factors and cytokines (Peprotech) were purchased from Invitrogen, unless otherwise stated.

Chondrogenic differentiation

Prior to differentiation, hPSCs were transferred and maintained in feeder-free conditions per manufacturer's recommendations using Essential 8 and Matrigel (Becton Dickinson) coated well plates. Mesoderm induction was carried out using a two-step protocol: cells were treated with a primitive streak induction medium containing Stemdiff APEL medium (APEL; Stem Cell Technologies) supplemented with 8 μM canonical wingless related integration protein (Wnt) agonist CHIR99021 (GSK3β inhibitor, Axon Medchem), 20 ng/ml hbFGF, 50 U and 50

mg/ml Pen/Strep for 36 hours followed by another 36 hours of APEL supplemented with 1 µM retinoic acid (RA), 8 ng/ml hbFGF, 50 U and 50 mg/ml Pen/Strep. The resultant cell populations were subsequently chondrogenically differentiated as previously described (*266*). Briefly, differentiation was induced for 14 days by using a chondrogenic medium (CM) consisting of DMEM supplemented with 1% Foetal Bovine Serum (FBS, Hyclone), 1% L-Glutamine, 1% NEAA, 1% SP, 1% Insulin-Transferrin-Selenite X (ITS-X), 50 µg/ml ascorbic acid (AA, Sigma), 0.1 mM 2ME, 10 ng/ml hbFGF, 10 ng/ml TGF- β1, 10 ng/ml BMP2, 10 ng/ml GDF5 (Prospec), 50 U and 50 mg/ml Pen/Strep. Subsequently, cartilage like nodules were detached and cultured in suspension using CM without hbFGF for 6 weeks.

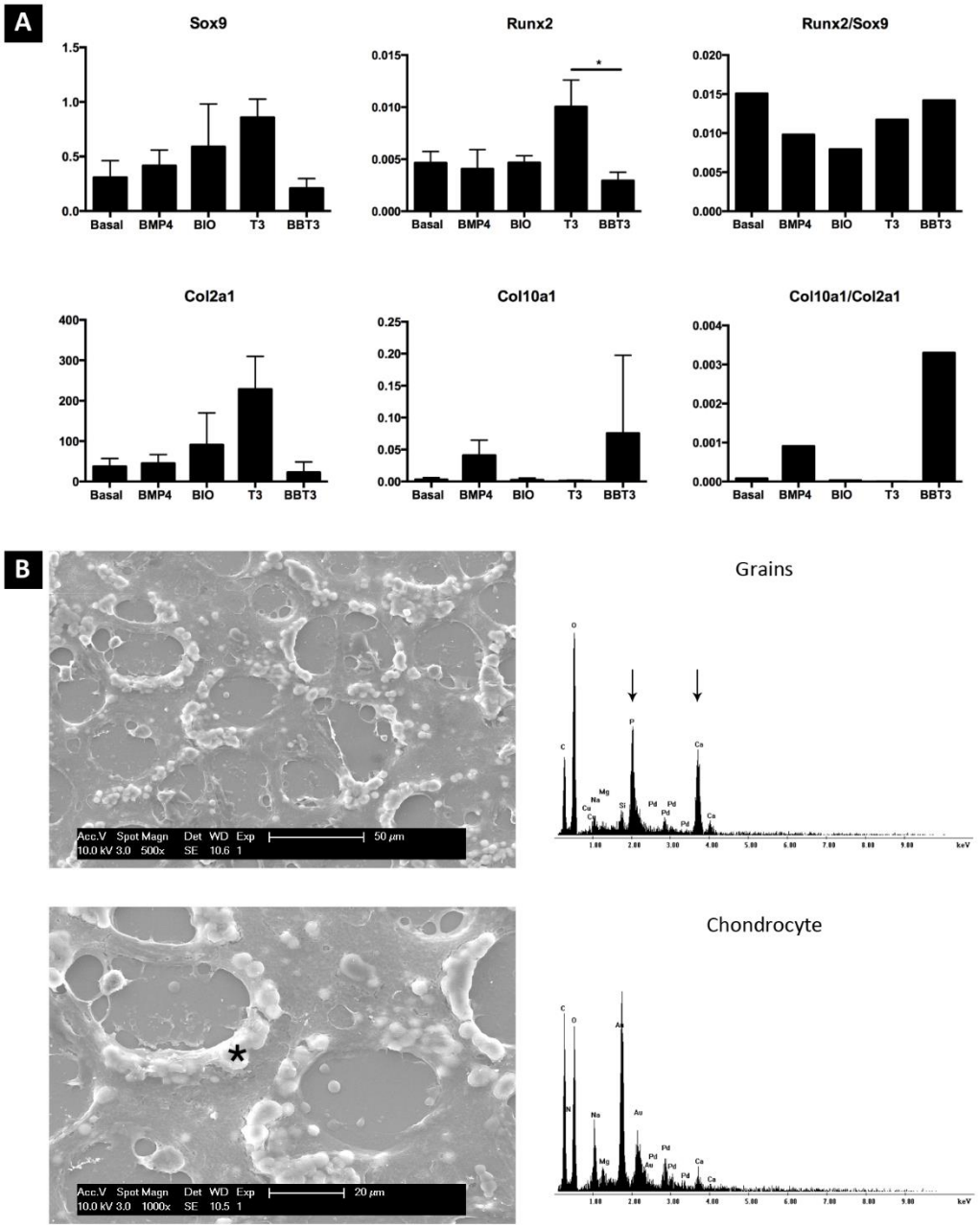


Figure S5.1. Hypertrophic differentiation of cartilaginous organoids. (A) Gene expression levels of *Sox9, Runx2, Col2A1* and *Col10A1*. Upon BBT3 stimulation, a trend towards decrease in *Sox9* and *Col2A1* was observed, while *Col10A1* was upregulated. The *Col10A1* to *Col2A1* ratio indicate that BBT3 treated samples underwent hypertrophic differentiation. Statistical significance is represented as follow: *p<0.05. **(B)** Scanning electron microscopy of the BBT3 treated samples to further investigate the composition changes of the extracellular matrix. White nodules (indicated by *) were detected surrounding the chondrocyte lacunae. Element analysis revealed the presence of calcium and phosphate within the samples (arrows). Chondrocyte embedded within the lacunaes were used as negative control. (Adapted from Tam et al. (2021) (*352*))

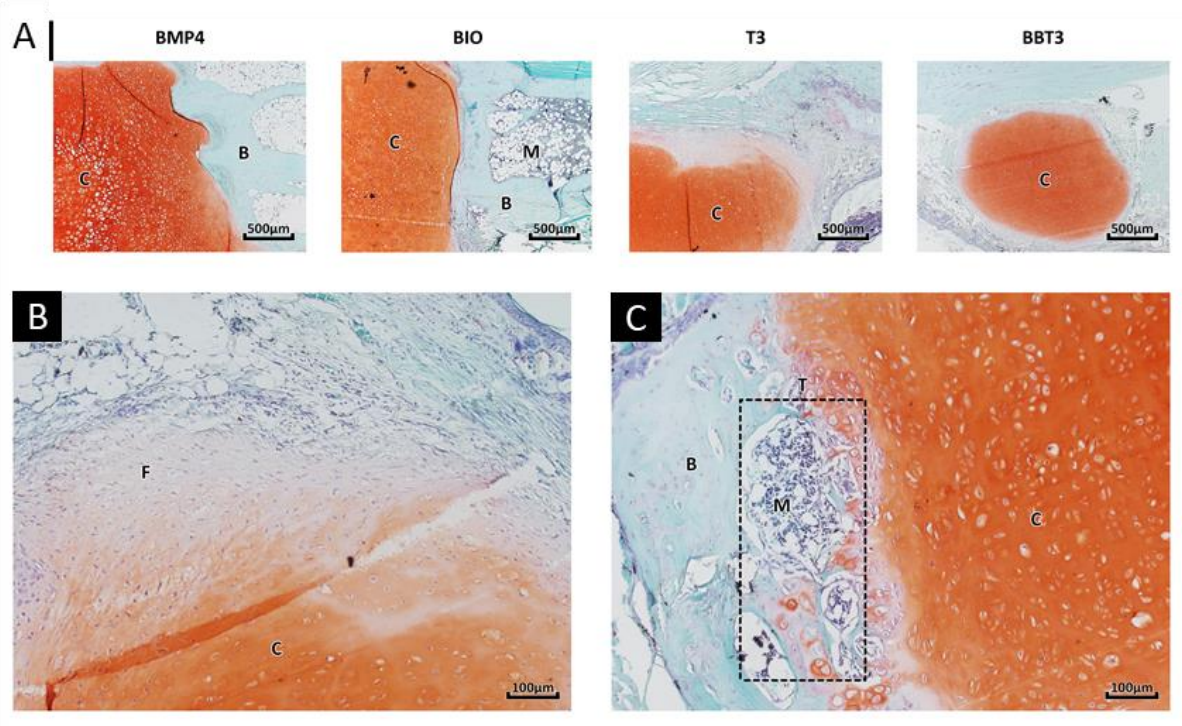


Figure S5.2. Orthotopic implantation of human derived cartilaginous aggregates. Hypertrophically differentiated cartilaginous organoids were implanted in critical size long bone defects. **(A)** Histological analysis of the BMP4, BIO, T3 and BBT3 treated samples using Safranin-O further indicated that the implanted organoids remained largely stable with no **(B)** to limited **(C)** bone formation at the edges of the implant, after 8 weeks of implantation. The following abbreviations have been used to highlight the tissue composition in the sections: B: bone, C: cartilage, F: fibrous tissue, M: bone marrow and T: cartilage-bone turnover site. (Adapted from Tam et al. (2021) (*352*)).

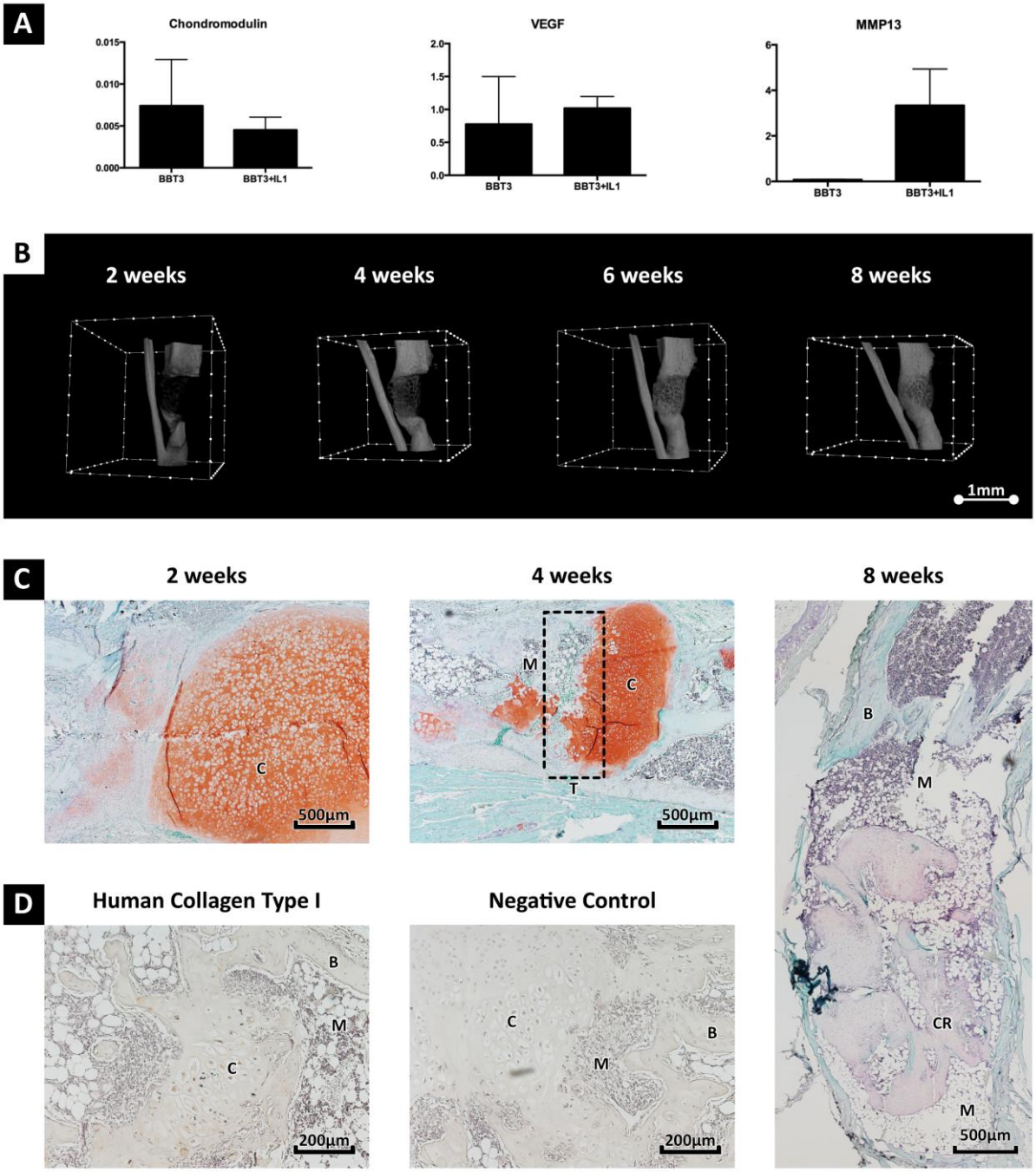


Figure S5.3: IL-1B priming allows successful fracture repair. (A) Upon IL-1B treatment a trend towards decrease of chondromodulin was detected while IL-1B upregulated VEGF and MMP13. **(B)** Following orthotopic implantation, the cartilaginous organoids underwent progressive mineralization with bone bridging being detected after 4 weeks. **(C)** Histological analysis using Safranin-O further revealed accelerated bone formation with cartilage resorption being detected at week 4. While bone union was detected after 8 weeks, it appears that some cartilaginous remnants could be detected, which indicates progressive cartilage resorption. **(D)** Immunostaining of human collagen type I further revealed that the newly formed bone was host derived, while some hypertrophic chondrocytes expressed low levels of collagen type I. (Adapted from Tam et al. 2021 (*352*))

Appendix of Chapter 6

Table S6.1: List of mechanistic interactions

The tables below provide the list of interactions/influences in the regulatory network model of **Chapter 6**. It is composed of a protein-signaling layer and a gene regulatory network. Rows represent interactions/influences. The first column gives the target protein or gene and the second gives the source protein, which are transcription factors in the GRN. The Sign of the interaction/influence indicate whether it is an activation (Sign =1) or an inhibition (Sign = -1). The three last columns are useful when importing the table in Cytoscape for the visualizing the network. All interactions are directed (TRUE). The GRN (resp. PPI) column is TRUE when the interaction belongs to the GRN layer (resp. protein-protein interaction layer).

Protein signaling layer:

Target	Source	Sign	Directed	GRN	PPI
WNT	DKK1	-1	TRUE	FALSE	TRUE
WNT	FRZB	-1	TRUE	FALSE	TRUE
DSH	FRIZZLED-LRP5/6	1	TRUE	FALSE	TRUE
DSH	R-SMADS	-1	TRUE	FALSE	TRUE
DC	DSH	-1	TRUE	FALSE	TRUE
DC	PP2A	1	TRUE	FALSE	TRUE
DC	ERK1/2	-1	TRUE	FALSE	TRUE
GSK	PKA	-1	TRUE	FALSE	TRUE
GSK	AKT	-1	TRUE	FALSE	TRUE
BCAT	DC	-1	TRUE	FALSE	TRUE
BCAT	Actin-MLC	1	TRUE	FALSE	TRUE
LEF/TCF	BCAT	1	TRUE	FALSE	TRUE
LEF/TCF	SOX9	-1	TRUE	FALSE	TRUE
BMP	GREM1	-1	TRUE	FALSE	TRUE
R-SMADS	SMAD7	-1	TRUE	FALSE	TRUE
R-SMADS	ALK1	1	TRUE	FALSE	TRUE
SMAD-Complex	R-SMADS	1	TRUE	FALSE	TRUE
SMAD-Complex	ERK1/2	-1	TRUE	FALSE	TRUE
MEF2C	P38	1	TRUE	FALSE	TRUE
MEF2C	SMAD-Complex	1	TRUE	FALSE	TRUE
MEF2C	HDAC4	-1	TRUE	FALSE	TRUE
MEF2C	SMAD3	-1	TRUE	FALSE	TRUE
DLX5	P38	1	TRUE	FALSE	TRUE
DLX5	SMAD-Complex	1	TRUE	FALSE	TRUE
DLX5	MSX2	-1	TRUE	FALSE	TRUE
DLX5	mTOR	1	TRUE	FALSE	TRUE
MSX2	DLX5	-1	TRUE	FALSE	TRUE

FGFR1	FGF	1	TRUE	FALSE	TRUE
FGFR3	FGF	1	TRUE	FALSE	TRUE
STAT1	FGFR3	1	TRUE	FALSE	TRUE
STAT1	IGF-IR	1	TRUE	FALSE	TRUE
STAT1	SMAD-Complex	-1	TRUE	FALSE	TRUE
STAT1	Rinflam	1	TRUE	FALSE	TRUE
STAT1	FGFR3	-1	TRUE	FALSE	TRUE
STAT1	SOCS	-1	TRUE	FALSE	TRUE
NFKB	FGFR1	1	TRUE	FALSE	TRUE
NFKB	AKT	1	TRUE	FALSE	TRUE
NFKB	NO	1	TRUE	FALSE	TRUE
NFKB	IKB-A	-1	TRUE	FALSE	TRUE
GLI2	IHH	1	TRUE	FALSE	TRUE
GLI2	mTOR	1	TRUE	FALSE	TRUE
PPR	PTHrP	1	TRUE	FALSE	TRUE
PKA	PPR	1	TRUE	FALSE	TRUE
PKA	IGF-I	1	TRUE	FALSE	TRUE
PP2A	PKA	1	TRUE	FALSE	TRUE
PP2A	mTOR	-1	TRUE	FALSE	TRUE
HDAC4	PP2A	1	TRUE	FALSE	TRUE
HDAC4	GSK	-1	TRUE	FALSE	TRUE
IGF-IR	IGF-I	1	TRUE	FALSE	TRUE
IGF-IR	SOCS	-1	TRUE	FALSE	TRUE
SMAD3	SMAD7	-1	TRUE	FALSE	TRUE
SMAD3	ERK1/2	-1	TRUE	FALSE	TRUE
SMAD3	CCND1	-1	TRUE	FALSE	TRUE
SMAD3	ALK5	1	TRUE	FALSE	TRUE
ATF2	P38	1	TRUE	FALSE	TRUE
ATF2	SMAD3	1	TRUE	FALSE	TRUE
ATF2	TGF-B	1	TRUE	FALSE	TRUE
ATF2	TAK1	1	TRUE	FALSE	TRUE
ATF2	JNK	1	TRUE	FALSE	TRUE
RAS	WNT	1	TRUE	FALSE	TRUE
RAS	BMP	1	TRUE	FALSE	TRUE
RAS	FGFR3	1	TRUE	FALSE	TRUE
RAS	FGFR1	1	TRUE	FALSE	TRUE
RAS	Integrins	1	TRUE	FALSE	TRUE
P38	TGF-B	1	TRUE	FALSE	TRUE
P38	BMP	1	TRUE	FALSE	TRUE
P38	TAK1	1	TRUE	FALSE	TRUE
ERK1/2	RAS	1	TRUE	FALSE	TRUE
ERK1/2	PP2A	-1	TRUE	FALSE	TRUE
ERK1/2	TAK1	1	TRUE	FALSE	TRUE
ERK1/2	PKA	1	TRUE	FALSE	TRUE
PI3K	RAS	1	TRUE	FALSE	TRUE

PI3K	IGF-IR	1	TRUE	FALSE	TRUE
PI3K	Rinflam	1	TRUE	FALSE	TRUE
PI3K	Integrins	1	TRUE	FALSE	TRUE
AKT	PI3K	1	TRUE	FALSE	TRUE
AKT	PP2A	-1	TRUE	FALSE	TRUE
RUNX2	AKT	1	TRUE	FALSE	TRUE
RUNX2	ERK1/2	1	TRUE	FALSE	TRUE
RUNX2	DLX5	1	TRUE	FALSE	TRUE
RUNX2	SOX9	-1	TRUE	FALSE	TRUE
RUNX2	PKA	-1	TRUE	FALSE	TRUE
RUNX2	SMAD3	-1	TRUE	FALSE	TRUE
RUNX2	HDAC4	-1	TRUE	FALSE	TRUE
RUNX2	CCND1	-1	TRUE	FALSE	TRUE
RUNX2	MSX2	-1	TRUE	FALSE	TRUE
RUNX2	YAP	-1	TRUE	FALSE	TRUE
ATF4	PKA	1	TRUE	FALSE	TRUE
ATF4	ERK1/2	1	TRUE	FALSE	TRUE
ATF4	[Ca2+]	-1	TRUE	FALSE	TRUE
ELK1	ERK1/2	1	TRUE	FALSE	TRUE
ELK1	ERK1/2	1	TRUE	FALSE	TRUE
ELK1	JNK	1	TRUE	FALSE	TRUE
SOX9	PKA	1	TRUE	FALSE	TRUE
SOX9	SMAD3	1	TRUE	FALSE	TRUE
SOX9	BCAT	-1	TRUE	FALSE	TRUE
SOX9	RUNX2	-1	TRUE	FALSE	TRUE
SOX9	RhoA/ROCK	1	TRUE	FALSE	TRUE
CCND1	RhoA/ROCK	1	TRUE	FALSE	TRUE
CCND1	GSK	-1	TRUE	FALSE	TRUE
CCND1	STAT1	1	TRUE	FALSE	TRUE
FRIZZLED-LRP5/6	WNT	1	TRUE	FALSE	TRUE
FRIZZLED-LRP5/6	DKK1	-1	TRUE	FALSE	TRUE
Rinflam	IL-1B	1	TRUE	FALSE	TRUE
Rinflam	SOCS	1	TRUE	FALSE	TRUE
ALK5	TGF-B	1	TRUE	FALSE	TRUE
ALK1	TGF-B	1	TRUE	FALSE	TRUE
ALK1	BMP	1	TRUE	FALSE	TRUE
TAK1	Rinflam	1	TRUE	FALSE	TRUE
TAK1	ALK1	1	TRUE	FALSE	TRUE
TAK1	YAP	-1	TRUE	FALSE	TRUE
JNK	TAK1	1	TRUE	FALSE	TRUE
Aggrecan	MMP13	-1	TRUE	FALSE	TRUE
Aggrecan	ADAMTS5	-1	TRUE	FALSE	TRUE
FOXO1	AKT	-1	TRUE	FALSE	TRUE
FOXO1	JNK	1	TRUE	FALSE	TRUE
FOXO1	ERK1/2	-1	TRUE	FALSE	TRUE

IKB-A	TAK1	-1	TRUE	FALSE	TRUE
IKB-A	LEF/TCF	-1	TRUE	FALSE	TRUE
IKB-A	Actin-MLC	-1	TRUE	FALSE	TRUE
mTOR	AKT	1	TRUE	FALSE	TRUE
COL-II	MMP13	-1	TRUE	FALSE	TRUE
Integrins	Loading	1	TRUE	FALSE	TRUE
Integrins	ADAMTS5	1	TRUE	FALSE	TRUE
FAK	Integrins	1	TRUE	FALSE	TRUE
dummyTF	Actin-MLC	1	TRUE	FALSE	TRUE
dummyTF	RhoA/ROCK	1	TRUE	FALSE	TRUE
YAP	RhoA/ROCK	1	TRUE	FALSE	TRUE
YAP	TAK1	-1	TRUE	FALSE	TRUE
YAP	Rinflam	-1	TRUE	FALSE	TRUE
RhoA/ROCK	FAK	1	TRUE	FALSE	TRUE
RhoA/ROCK	[Ca2+]	1	TRUE	FALSE	TRUE
Actin-MLC	RhoA/ROCK	1	TRUE	FALSE	TRUE
TRPV4	Loading	1	TRUE	FALSE	TRUE
TRPV4	TAK1	-1	TRUE	FALSE	TRUE
TRPV4	Actin-MLC	-1	TRUE	FALSE	TRUE
TRPV4	IGF-IR	-1	TRUE	FALSE	TRUE
[Ca2+]	TRPV4	1	TRUE	FALSE	TRUE
NO	iNOS	1	TRUE	FALSE	TRUE
iNOS	Actin-MLC	1	TRUE	FALSE	TRUE

Gene regulatory layer

Target	Source	Sign	Directed	GRN	PPI
WNT	GLI2	1	TRUE	TRUE	FALSE
WNT	MSX2	-1	TRUE	TRUE	FALSE
LEF/TCF	RUNX2	1	TRUE	TRUE	FALSE
LEF/TCF	BCAT	1	TRUE	TRUE	FALSE
BMP	GLI2	1	TRUE	TRUE	FALSE
BMP	NFKB	1	TRUE	TRUE	FALSE
MEF2C	RUNX2	1	TRUE	TRUE	FALSE
MEF2C	SMAD-Complex	1	TRUE	TRUE	FALSE
DLX5	MEF2C	1	TRUE	TRUE	FALSE
DLX5	SMAD-Complex	1	TRUE	TRUE	FALSE
DLX5	SMAD3	-1	TRUE	TRUE	FALSE
MSX2	SMAD-Complex	1	TRUE	TRUE	FALSE
MSX2	SMAD3	1	TRUE	TRUE	FALSE
MSX2	MSX2	1	TRUE	TRUE	FALSE
MSX2	DLX5	-1	TRUE	TRUE	FALSE
FGF	BCAT	1	TRUE	TRUE	FALSE
FGF	RUNX2	1	TRUE	TRUE	FALSE

FGFR1	SMAD-Complex	-1	TRUE	TRUE	FALSE
FGFR1	RUNX2	1	TRUE	TRUE	FALSE
FGFR3	NFKB	1	TRUE	TRUE	FALSE
FGFR3	ELK1	-1	TRUE	TRUE	FALSE
FGFR3	ERK1/2	-1	TRUE	TRUE	FALSE
HIF-2a	NFKB	1	TRUE	TRUE	FALSE
IHH	RUNX2	1	TRUE	TRUE	FALSE
IHH	SMAD-Complex	1	TRUE	TRUE	FALSE
IHH	NFKB	1	TRUE	TRUE	FALSE
IHH	ATF4	1	TRUE	TRUE	FALSE
IHH	EF-1	-1	TRUE	TRUE	FALSE
IHH	MSX2	1	TRUE	TRUE	FALSE
IHH	FGFR3	-1	TRUE	TRUE	FALSE
IHH	SMAD3	-1	TRUE	TRUE	FALSE
GLI2	FGFR3	-1	TRUE	TRUE	FALSE
PTHrP	SOX9	1	TRUE	TRUE	FALSE
PTHrP	GLI2	1	TRUE	TRUE	FALSE
PTHrP	SMAD3	1	TRUE	TRUE	FALSE
PPR	GLI2	1	TRUE	TRUE	FALSE
PPR	SOX9	1	TRUE	TRUE	FALSE
PPR	SMAD-Complex	1	TRUE	TRUE	FALSE
PPR	STAT1	-1	TRUE	TRUE	FALSE
IGF-I	PKA	1	TRUE	TRUE	FALSE
IGF-IR	SMAD3	1	TRUE	TRUE	FALSE
IGF-IR	NFKB	1	TRUE	TRUE	FALSE
TGF-B	GLI2	1	TRUE	TRUE	FALSE
TGF-B	ELK1	-1	TRUE	TRUE	FALSE
TGF-B	SMAD3	1	TRUE	TRUE	FALSE
SMAD7	STAT1	1	TRUE	TRUE	FALSE
SMAD7	SMAD-Complex	1	TRUE	TRUE	FALSE
SMAD7	SMAD3	1	TRUE	TRUE	FALSE
SMAD7	NFKB	1	TRUE	TRUE	FALSE
AKT	RUNX2	1	TRUE	TRUE	FALSE
RUNX2	HIF-2a	1	TRUE	TRUE	FALSE
RUNX2	LEF/TCF	1	TRUE	TRUE	FALSE
RUNX2	RUNX2	1	TRUE	TRUE	FALSE
RUNX2	MEF2C	1	TRUE	TRUE	FALSE
RUNX2	DLX5	1	TRUE	TRUE	FALSE
RUNX2	NKX3.2	-1	TRUE	TRUE	FALSE
RUNX2	SMAD-Complex	-1	TRUE	TRUE	FALSE
RUNX2	MSX2	-1	TRUE	TRUE	FALSE
RUNX2	ELK1	1	TRUE	TRUE	FALSE
RUNX2	GLI2	1	TRUE	TRUE	FALSE

EF-1	NFKB	1	TRUE	TRUE	FALSE
EF-1	ELK1	1	TRUE	TRUE	FALSE
EF-1	SMAD-Complex	-1	TRUE	TRUE	FALSE
EF-1	R-SMADS	1	TRUE	TRUE	FALSE
SOX9	PKA	1	TRUE	TRUE	FALSE
SOX9	NFKB	-1	TRUE	TRUE	FALSE
SOX9	P38	1	TRUE	TRUE	FALSE
SOX9	NFKB	1	TRUE	TRUE	FALSE
SOX9	SOX9	1	TRUE	TRUE	FALSE
SOX9	NKX3.2	1	TRUE	TRUE	FALSE
NKX3.2	SOX9	1	TRUE	TRUE	FALSE
NKX3.2	PKA	1	TRUE	TRUE	FALSE
NKX3.2	ATF4	1	TRUE	TRUE	FALSE
CCND1	ATF2	1	TRUE	TRUE	FALSE
CCND1	GLI2	1	TRUE	TRUE	FALSE
CCND1	PKA	1	TRUE	TRUE	FALSE
CCND1	ATF4	1	TRUE	TRUE	FALSE
CCND1	FOXO1	1	TRUE	TRUE	FALSE
COL-II	SOX9	1	TRUE	TRUE	FALSE
COL-II	FGFR1	1	TRUE	TRUE	FALSE
COL-II	NKX3.2	1	TRUE	TRUE	FALSE
COL-II	FOXO1	1	TRUE	TRUE	FALSE
COL-II	BCAT	-1	TRUE	TRUE	FALSE
COL-II	dummyTF	1	TRUE	TRUE	FALSE
COL-X	RUNX2	1	TRUE	TRUE	FALSE
COL-X	R-SMADS	1	TRUE	TRUE	FALSE
COL-X	MEF2C	1	TRUE	TRUE	FALSE
COL-X	PKA	-1	TRUE	TRUE	FALSE
COL-X	PTHrP	-1	TRUE	TRUE	FALSE
COL-X	HIF-2a	1	TRUE	TRUE	FALSE
COL-X	BCAT	1	TRUE	TRUE	FALSE
COL-X	STAT1	-1	TRUE	TRUE	FALSE
MMP13	RUNX2	1	TRUE	TRUE	FALSE
MMP13	NFKB	1	TRUE	TRUE	FALSE
MMP13	HIF-2a	1	TRUE	TRUE	FALSE
MMP13	GLI2	1	TRUE	TRUE	FALSE
MMP13	P38	1	TRUE	TRUE	FALSE
MMP13	ELK1	1	TRUE	TRUE	FALSE
MMP13	SMAD-Complex	1	TRUE	TRUE	FALSE
MMP13	SMAD3	-1	TRUE	TRUE	FALSE
MMP13	ELK1	1	TRUE	TRUE	FALSE
MMP13	ATF2	1	TRUE	TRUE	FALSE
MMP13	JNK	1	TRUE	TRUE	FALSE
MMP13	FOXO1	-1	TRUE	TRUE	FALSE
GREM1	SOX9	1	TRUE	TRUE	FALSE

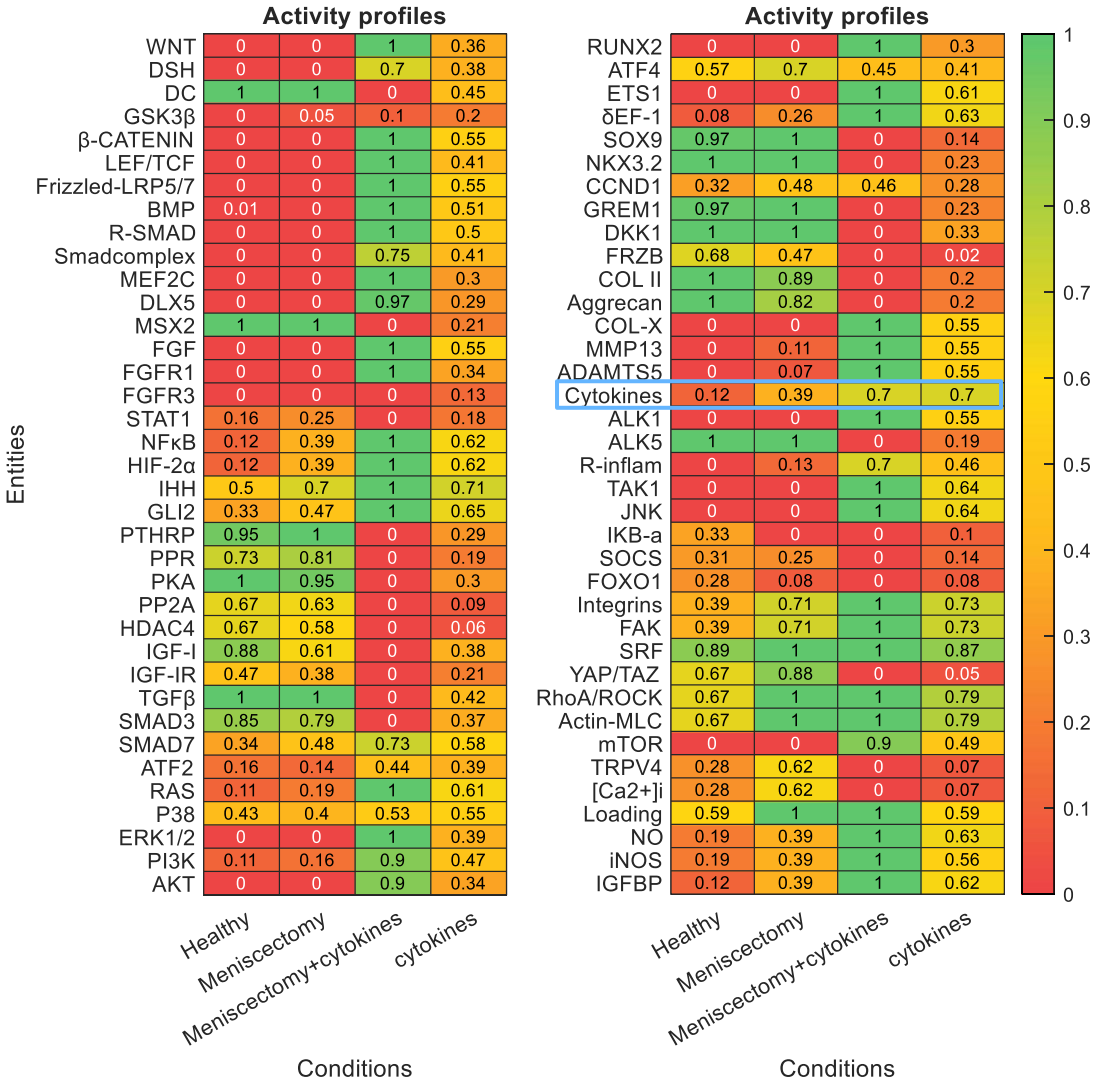
GREM1	BCAT	-1	TRUE	TRUE	FALSE
FRZB	SMAD-Complex	1	TRUE	TRUE	FALSE
FRZB	BCAT	-1	TRUE	TRUE	FALSE
FRZB	SOX9	1	TRUE	TRUE	FALSE
FRZB	ATF4	-1	TRUE	TRUE	FALSE
FRZB	NFKB	-1	TRUE	TRUE	FALSE
DKK1	SMAD-Complex	1	TRUE	TRUE	FALSE
DKK1	BCAT	-1	TRUE	TRUE	FALSE
DKK1	SOX9	1	TRUE	TRUE	FALSE
DKK1	ATF2	-1	TRUE	TRUE	FALSE
DKK1	NFKB	-1	TRUE	TRUE	FALSE
R-SMADS	ATF4	-1	TRUE	TRUE	FALSE
R-SMADS	EF-1	1	TRUE	TRUE	FALSE
PP2A	BCAT	-1	TRUE	TRUE	FALSE
ERK1/2	MEF2C	1	TRUE	TRUE	FALSE
P38	EF-1	-1	TRUE	TRUE	FALSE
PI3K	RUNX2	1	TRUE	TRUE	FALSE
ALK1	SMAD3	-1	TRUE	TRUE	FALSE
ALK1	LEF/TCF	1	TRUE	TRUE	FALSE
ALK1	ATF2	1	TRUE	TRUE	FALSE
ALK1	JNK	1	TRUE	TRUE	FALSE
ALK1	RUNX2	1	TRUE	TRUE	FALSE
ALK5	SOX9	1	TRUE	TRUE	FALSE
ALK5	SMAD3	1	TRUE	TRUE	FALSE
ALK5	NFKB	-1	TRUE	TRUE	FALSE
Aggrecan	SOX9	1	TRUE	TRUE	FALSE
Aggrecan	NFKB	-1	TRUE	TRUE	FALSE
Aggrecan	dummyTF	1	TRUE	TRUE	FALSE
Aggrecan	FOXO1	1	TRUE	TRUE	FALSE
IL-1B	NFKB	1	TRUE	TRUE	FALSE
IL-1B	ELK1	1	TRUE	TRUE	FALSE
IL-1B	ATF2	1	TRUE	TRUE	FALSE
IL-1B	JNK	1	TRUE	TRUE	FALSE
IKB-A	SOX9	1	TRUE	TRUE	FALSE
IKB-A	NFKB	1	TRUE	TRUE	FALSE
SOCS	IGF-IR	1	TRUE	TRUE	FALSE
SOCS	FGFR3	1	TRUE	TRUE	FALSE
FOXO1	ELK1	-1	TRUE	TRUE	FALSE
FOXO1	ATF2	-1	TRUE	TRUE	FALSE
FOXO1	NFKB	-1	TRUE	TRUE	FALSE
FOXO1	TGF-B	1	TRUE	TRUE	FALSE
FOXO1	TAK1	1	TRUE	TRUE	FALSE
FOXO1	P38	1	TRUE	TRUE	FALSE
Rinflam	FOXO1	-1	TRUE	TRUE	FALSE
IGFPB	RUNX2	1	TRUE	TRUE	FALSE

IGFPB	NFKB	1	TRUE	TRUE	FALSE
ADAMTS5	RUNX2	1	TRUE	TRUE	FALSE
ADAMTS5	NFKB	1	TRUE	TRUE	FALSE
ADAMTS5	HIF-2a	1	TRUE	TRUE	FALSE
ADAMTS5	FOXO1	-1	TRUE	TRUE	FALSE
ADAMTS5	[Ca2+]	-1	TRUE	TRUE	FALSE
Integrins	dummyTF	1	TRUE	TRUE	FALSE
iNOS	NFKB	1	TRUE	TRUE	FALSE
iNOS	ELK1	1	TRUE	TRUE	FALSE

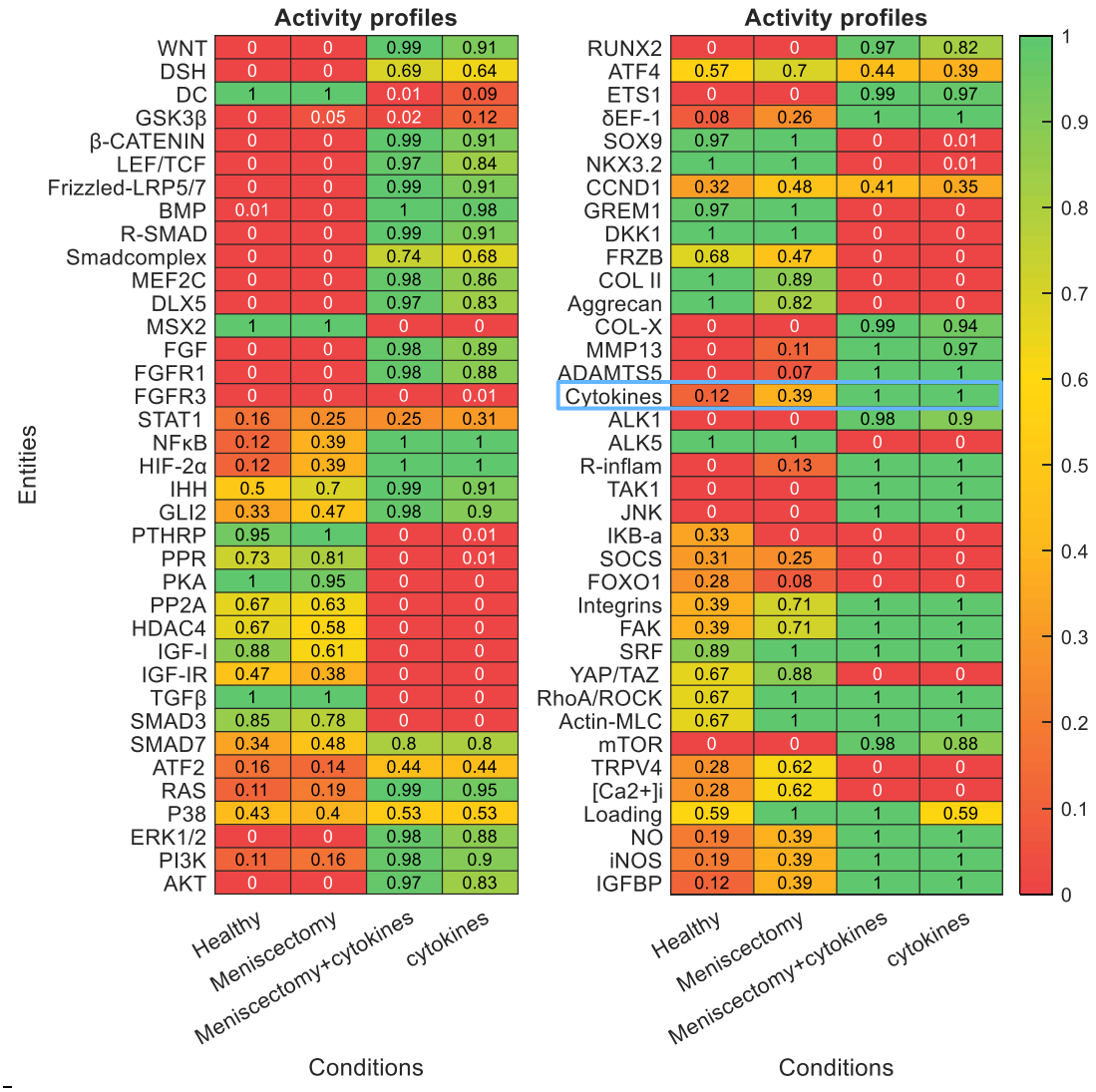
Figure S6.1: Effect of meniscectomy with a without addition of pro-inflammatory cytokines in varying amounts.

The effect of different input of pro-inflammatory cytokines is investigated: cytokines= 0.7 vs. cytokines = 1, with 1 being the maximal level (i.e. saturation). Global activities of all the 74 variables are reported for the superfical and the deep zone.

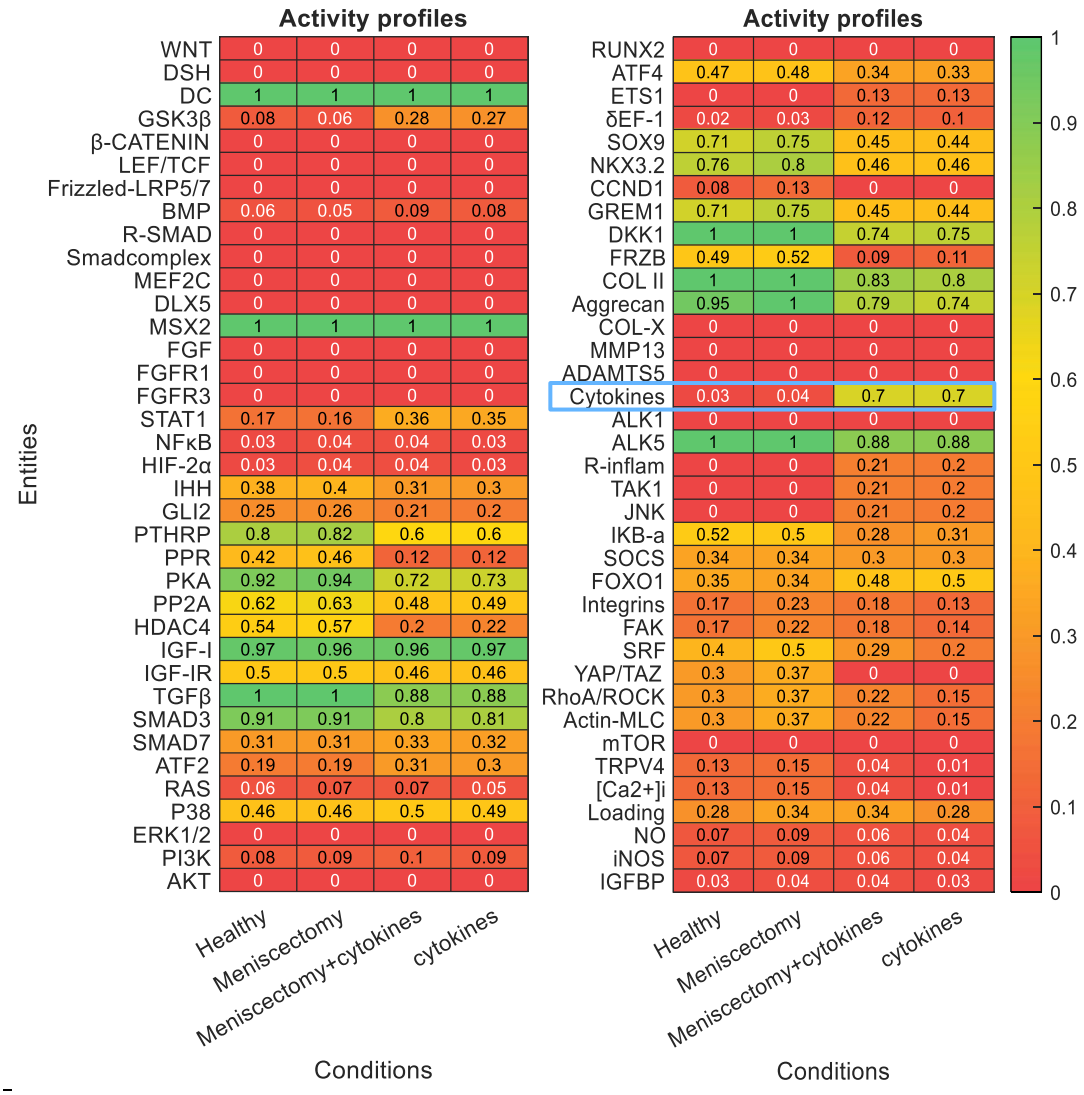
Superficial zone (pro-inflamatory cytokines input = 0.7)



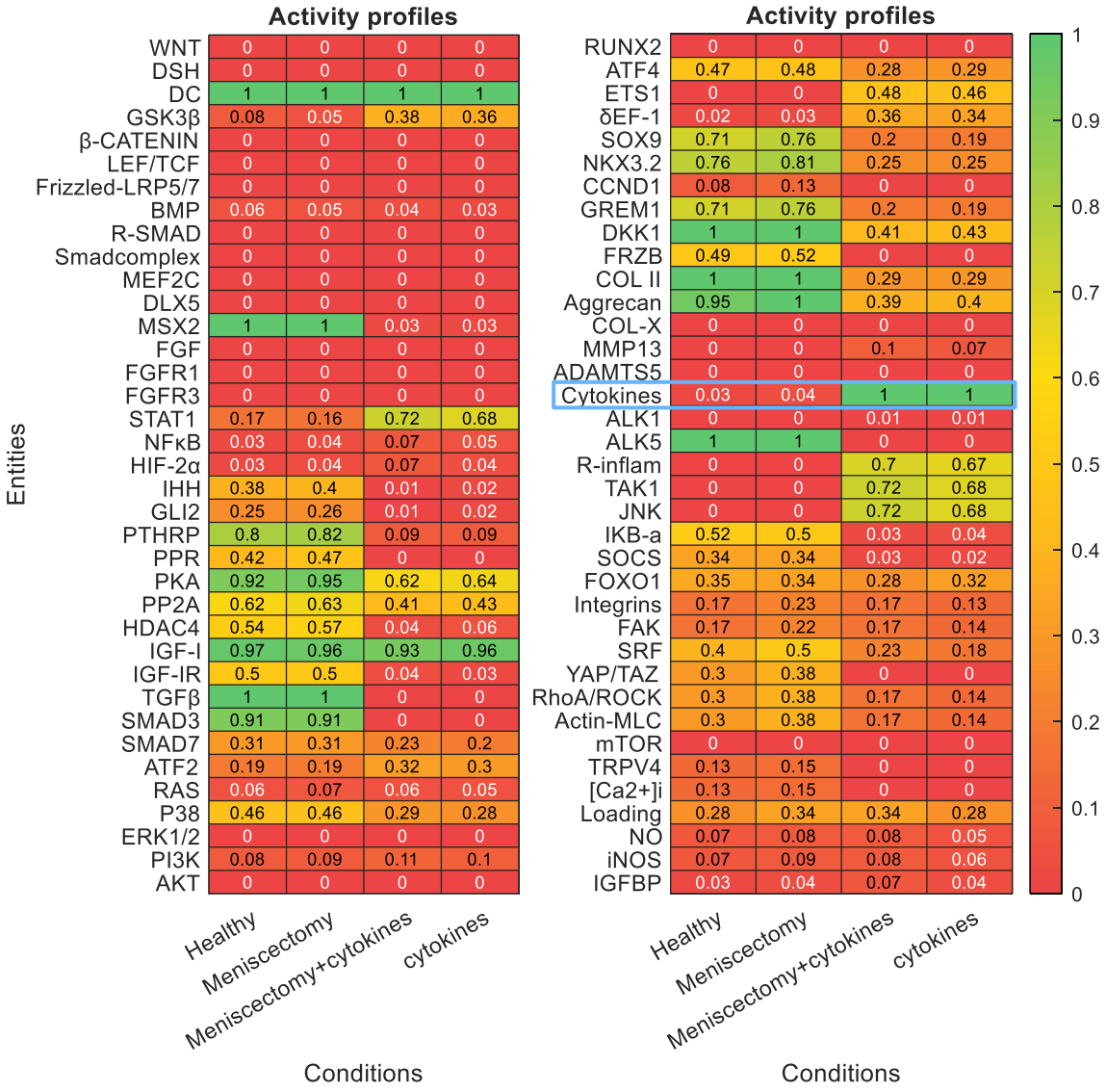
Superficial zone (pro-inflamatory cytokines input = 1)



Deep zone (pro-inflamatory cytokines input = 0.7)



Deep zone (pro-inflamatory cytokines input = 1)



Scientific acknowledgement

Research presented in this thesis was funded by the European Union's Horizon 2020 research and innovation programme under Marie Sklodowska-Curie grant agreement No 721432 and the European Research Council Consolidator Grant agreement grant 772418.

Curriculum vitae

Name:	**Raphaëlle LESAGE**
Birth:	June 25, 1993. Charenton-le-Pont, France
Address:	Avenue Marcel Thiry, 93. Box7. 1200 Woluwe Saint Lambert, Belgium
Telephone:	+32 489996931
E-mails:	raphaelle.lesage@kuleuven.be, lesage.raphaelle3@gmail.com

Education

2017 – 2021	**Interdisplinary PhD in Engineering Science and Biomedical Sciences** at KU Leuven, Belgium
2013-2016	**Bioengineering diploma - Bioinformatics and Modeling for Biology (BIMB)** Engineering school *Polytech Nice Sophia Antipolis.*
2011-2013	**Studies in preparatory classes for competitive engineering top schools - CPGE** In Biology, Chemistry, Physics and Earth Science section**.** *Janson de Sailly, Paris.*
2011	**Graduate from High School, Major Science, A-level grade** *France (77)*

Bibliography

1. (Submitted) An integrated *in silico-in vitro* approach for the identification of drug targets against chondrocyte hypertrophy in osteoarthritis. **R. Lesage**, M.F. Blanco, R.Narcisi, T.Welting, GJVM Van Osch, L.Geris. (BioRxive preprint available: https://doi.org/10.1101/2021.09.27.461207)

2. (In preparation) Development of a machine-learning model to assess osteoarthritis signature in mouse disease models. **R. Lesage**, L. Geris

3. **R. Lesage**, J. Kerkhofs, L. G. Geris, Computational modeling and reverse engineering to reveal dominant regulatory interactions controlling osteochondral differentiation: potential for regenerative medicine. *Front. Bioeng. Biotechnol.* 6, 165 (2018)

4. C. Gorgun, D. Ceresa, **R. Lesage**, F. Villa, D. Reverberi, C. Balbi, S. Santamaria, K. Cortese, P. Malatesta, L. Geris, R. Quarto, R. Tasso, Dissecting the effects of preconditioning with inflammatory cytokines and hypoxia on the angiogenic potential of mesenchymal stromal cell (MSC)-derived soluble proteins and extracellular vesicles (EVs). *Biomaterials*. 269, 120633 (2021)

5. I. Papantoniou, G. Nilsson Hall, N. Loverdou, **R. Lesage**, T. Herpelinck, L. Mendes, L. Geris, Turning Nature's own processes into design strategies for living bone implant biomanufacturing: a decade of Developmental Engineering. *Adv. Drug Deliv. Rev.* 169 (2021), pp. 22–39

6. F. T. Musuamba, R. Bursi, E. Manolis, K. Karlsson, A. Kulesza, E. Courcelles, J. Boissel, **R. Lesage**, C. Crozatier, E. M. Voisin, C. F. Rousseau, T. Marchal, R. Alessandrello, L. Geris, Verifying and Validating Quantitative Systems Pharmacology and *In silico* Models in Drug Development: Current Needs, Gaps, and Challenges. *CPT Pharmacometrics Syst. Pharmacol.* 9, 195–197 (2020)

7. F. T. Musuamba, I. Skottheim Rusten, **R. Lesage**, G. Russo, R. Bursi, L. Emili, G. Wangorsch, E. Manolis, K. E. Karlsson, A. Kulesza, E. Courcelles, J. P. Boissel, C. F. Rousseau, E. M. Voisin, R. Alessandrello, N. Curado, E. Dall'ara, B. Rodriguez, F. Pappalardo, L. Geris, Scientific and regulatory evaluation of mechanistic *in silico* drug and disease models in drug development: Building model credibility. *CPT Pharmacometrics Syst. Pharmacol.* (2021)

8. Human pluripotent stem cell derived cartilaginous organoids promote scaffold-free healing of critical size long bone defects. W.L. Tam, L. Freitas Mendes, X. Chen, **R. Lesage**, I. Van Hoven, E. Leysen, G. Kerckhofs, K. Bosmans, Y. Chin Chai, A. Yamashita, N. Tsumaki, L. Geris, S. J. Roberts and F.P. Luyten

9. Tam WL, Freitas Mendes L, Chen X, **Lesage R**, Van Hoven I, Leysen E, Kerckhofs G, Bosmans K, Chai YC, Yamashita A, Tsumaki N, Geris L, Roberts SJ, Luyten FP. Human pluripotent stem cell-derived cartilaginous organoids promote scaffold-free healing of critical size long bone defects. Stem Cell Res Ther. (2021)

10. F. M. Nielsen, S. E. Riis, J. I. Andersen, **R. Lesage**, T. Fink, C. P. Pennisi, V. Zachar, Discrete adipose-derived stem cell subpopulations may display differential functionality after *in vitro* expansion despite convergence to a common phenotype distribution. *Stem Cell Res. Ther.* **7**, 177 (2016)

Abstracts presented at international conference and symposia

- SBHD 2021 (online) conference: Development of a machine-learning model to assess osteoarthritis signature in mouse disease models. R. Lesage, L. Geris (Oral presentation)

- ESB2021 (online) conference: Interplay of mechanical and inflammatory cues in osteoarthritis: a multiscale in-silico approach. S. Mukherjee, R. Lesage, L.Geris (Oral presentation by S.Mukherjee)

- ORS2021 (online) conference: Predicting chondrocyte hypertrophy due to mechanical and inflammatory cues: a multiscale in-silico approach. S. Mukherjee, R. Lesage, L.Geris (Oral presentation by S.Mukherjee)

- ISCB 2020 (online) conference: Development of a machine-learning model to assess osteoarthritis signature in mouse disease models. R. Lesage, L.Geris (Poster)

- VPH2020 (online) conference: *In vitro* validation of potential osteoarthritis drug targets predicted by an *in silico* model of articular chondrocyte. R. Lesage, M.F. Blanco, GJVM Van Osch, R.Narcisi, T.Welting, L.Geris. (Oral presentation). ***This abstract was selected as one of the final four in the VPHi student award competition****

- EORS2019 conference: Study of osteoarthritis (OA) development and prediction of potential intervention targets by an *in silico* approach. (2019) *Orthopaedic Proceedings*. R. Lesage, M.F. Blanco, GJVM Van Osch, L.Geris. (Oral presentation + publication in conf. proceeding)

- ICSB 2019 conference: An *In silico* approach to study Osteoarthritis development and identify potential intervention targets (Poster + Oral presentation during BioNetVista workshop)

- SBHD 2019 conference: An *In silico* approach to study Osteoarthritis development and identify potential intervention targets. R.Lesage, L.Geris (Oral presentation)

- ByteMal2019: Merging data-based and mechanistic modeling to study osteoarthritis (OA) development and predict potential intervention targets (Poster + flash talk) ***This poster & associated flash talk were awarded the best poster award****

- ECCB 2018 conference: Mechanistic versus data-driven modeling for osteochondral regulatory networks. R.Lesage, J.Kekhofs, L.Geris (Poster)

- ICCS 2018 conference: Reverse Engineering Methods to Study Osteochondral Regulatory Networks. R.Lesage, J.Kekhofs, L.Geris (Poster + Oral presentation)

FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF MECHNICAL ENGINEERING
BIOMECHANICS SECTION
celestijnenlaan 300c box 2419
b-3001 heverlee

FACULTY OF MEDICINE
DEPARTMENT OF DEVELOPMENT AND REGENERATION
PROMETHEUS- THE DIVISION OF SKELETAL TISSUE ENGINEERING
SKELETAL BIOLOGY AND ENGINEERING RESEARCH CENTER
Herestraat 49. bus 813
3000 LEUVEN, België
raphaelle.lesage@kuleuven.be
www.mtm.kuleuven.be/Prometheus