An integrated in silico-in vitro approach for identifying therapeutic targets against osteoarthritis

Lesage, Raphaëlle; Ferrao Blanco, Mauricio N.; Narcisi, Roberto; Welting, Tim; van Osch, Gerjo J.V.M.; Geris, Liesbet

Citation (APA)
Lesage, R., Ferrao Blanco, M. N., Narcisi, R., Welting, T., van Osch, G. J. V. M., & Geris, L. (2022). An integrated in silico-in vitro approach for identifying therapeutic targets against osteoarthritis. *BMC Biology*, *20*(1), Article 253. https://doi.org/10.1186/s12915-022-01451-8

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

BMC Biology

RESEARCH ARTICLE

An integrated *in silico-in vitro* approach for identifying therapeutic targets against osteoarthritis

Raphaëlle Lesage[1,2], Mauricio N. Ferrao Blanco[3], Roberto Narcisi[3], Tim Welting[4], Gerjo J. V. M. van Osch[3,5,6] and Liesbet Geris[1,2,7*]

Abstract

Background: Without the availability of disease-modifying drugs, there is an unmet therapeutic need for osteoarthritic patients. During osteoarthritis, the homeostasis of articular chondrocytes is dysregulated and a phenotypical transition called hypertrophy occurs, leading to cartilage degeneration. Targeting this phenotypic transition has emerged as a potential therapeutic strategy. Chondrocyte phenotype maintenance and switch are controlled by an intricate network of intracellular factors, each influenced by a myriad of feedback mechanisms, making it challenging to intuitively predict treatment outcomes, while *in silico* modeling can help unravel that complexity. In this study, we aim to develop a virtual articular chondrocyte to guide experiments in order to rationalize the identification of potential drug targets via screening of combination therapies through computational modeling and simulations.

Results: We developed a signal transduction network model using knowledge-based and data-driven (machine learning) modeling technologies. The *in silico* high-throughput screening of (pairwise) perturbations operated with that network model highlighted conditions potentially affecting the hypertrophic switch. A selection of promising combinations was further tested in a murine cell line and primary human chondrocytes, which notably highlighted a previously unreported synergistic effect between the protein kinase A and the fibroblast growth factor receptor 1.

Conclusions: Here, we provide a virtual articular chondrocyte in the form of a signal transduction interactive knowledge base and of an executable computational model. Our *in silico-in vitro* strategy opens new routes for developing osteoarthritis targeting therapies by refining the early stages of drug target discovery.

Keywords: Network of signal transduction, Computational modeling, Drug targets, Osteoarthritis, Chondrocyte hypertrophy, *In vitro* validation, Regulatory network inference, Virtual cell

Graphical Abstract

Background

Osteoarthritis (OA) is a degenerative disease of the joint increasingly prevalent due to the aging population. It is a major societal burden as no disease-modifying drugs are currently available on the market [1]. 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

*Correspondence: liesbet.geris@kuleuven.be

[7] GIGA In silico Medicine, University of Liège, Liège, Belgium
Full list of author information is available at the end of the article

Lesage *et al. BMC Biology*   (2022) 20:253

Page 2 of 25

influence the shift from stable healthy cartilage towards a diseased state [2]. Regardless of the exact inducing mechanisms, during that transition, some of the chondrocytes enter a maturation process called hypertrophy [3, 4], 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 [2, 5–8]. Therefore, controlling chondrocyte phenotype to prevent hypertrophic maturation has emerged as a potential therapeutic strategy to treat OA patients [7, 9].

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 Indian hedgehog (IHH) and inflammatory signaling pathways [10]. Routes downstream of various growth factors are thought to be important in the control or disruption of chondrocyte homeostasis, such as the WNT and Bone morphogenic protein (BMP) pathways, the parathyroid hormone-related peptide (PTHrP), as well as the insulin-like growth factor (IGF)-I, fibroblast growth factors (FGF) and transforming growth factors (TGF)-B [9, 11, 12]. However, the interplay of intracellular pathways is highly intricate with extensive feedback loops, non-linear pathways, redundancy, and intertwining [11, 13, 14]. This complicates the intuitive prediction of what will happen in case of perturbation 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 Dickkopf1 (DKK1), both induced a reduction of glycosaminoglycan rich ECM in human articular chondrocytes [15]. 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 external perturbations and potential therapies requires a systemic view on the process and a holistic approach [13, 14].

We propose to unravel the complexity of these regulatory pathways and to rationalize the identification of potential drug targets via screening of (pairwise) perturbations 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 target identification in the drug discovery process. The *in silico* approach starts with collecting intracellular biological mechanisms. 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 allows to investigate questions that would lead to new insights into yet unexplored situations and interactions [16]. Such computational mechanistic approaches were already used in the past to identify influential candidates for cancer therapeutics [17], study the control cell fate decision [18], including in cartilage [19, 20] or prioritize personalized combination therapies [21].

In this study, we developed an *in silico* model of the regulatory intracellular 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 intracellular states of the articular chondrocyte model and investigated 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 investigated *in vitro* the role of a selection of predicted factors in the regulation of the phenotypic change, for both single and combinations of factors to guide future therapeutic strategy discovery.

Results

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 (Fig. 1). In this graphical influence network, or activity flow representation, the nodes represent biological components (i.e., proteins or genes, see Additional file 1 Table S1 for definitions) important for chondrocyte biology [9]. Directed edges linking two nodes represent interactions or activating/inhibitory influences between source and target proteins or between transcription factors (TFs) and target genes. Information was collected and adapted from a previously published model of growth plate chondrocytes [22, 23] as well as from additional deep literature and database curation (see the "Methods" section). The annotations, descriptions, and cross-references for each network node and its interactions can be consulted in two interactive subnetworks, a protein signaling one and a gene regulatory one. The former goes from the growth

Lesage *et al. BMC Biology*   (2022) 20:253

Page 3 of 25

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 (Fig. 1). The two subnetworks are interconnected as each biological component is represented by a gene in the gene regulatory network (GRN) and its corresponding protein in the protein signaling network. Both subnetworks, forming the full network, can be regarded as an interactive knowledge base on chondrocyte signaling pathways, which is available through the online platform Cell Collective [24]. We refer the reader to that interactive knowledge base for the literature support of our network model (see links in the "Availability of data and materials" section), while the main pathways that were included are listed below. It is noteworthy that some biological factors were not regulated in one of the subnetworks (unknown transcriptional regulators or absence of post-translational modifications) and/or did not have downstream targets, but they were included in the knowledge base anyway (possibly resulting in disconnected—but annotated—nodes in that subnetwork) since regulatory relationships were described in the other subnetwork (see Additional file 2 for more details).

Unless stated otherwise, all components of the network and of the *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 Additional file 3 Table S2 and reported in the Cell Collective interactive networks, for information.

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 [9, 25]. The influence of pro-inflammatory cytokines, such as interleukin 1beta (IL-1B) or Tumor Necrosis Factor Alpha (TNF-A), was summarized through a node labeled 'cytokines', in the network (Fig. 1). That node can signal through a single receptor (labeled 'Rinfl') that can activate

well known downstream pathways such as the phosphatidylinositol 3-kinase (PI3K)/AKT axis, the nuclear factor kappa B (NFKB) pathway and mitogen-activated protein kinases (MAPK) pathways. These MAPKs in question include extracellular signal-regulated kinase 1&2 (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 transducer GLI2, the signal transducer and activator of transcription (STAT1), the transcription factor 7 (TCF), the myocyte enhancer factor 2C (MEF2C) as well as SRY-Box Transcription Factor (SOX9), a marker of differentiated healthy chondrocytes, and runt-related transcription factor 2 (RUNX2), a hypertrophy marker. All pathways in the model are highly interconnected as shown by Fig. 1.

In total, there are 60 biological components in the network, which are listed in Additional file 3, Table S2. 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 gene regulatory networks 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.

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

(See figure on next page.)

Fig. 1 Influence map of the signaling and gene regulatory networks (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. In this image, each biological component is represented by a gene in the GRN and a protein in the signaling network, except for the ones that are not involved in one of the subnetworks (e.g., NKX3.2 only plays a role in the GRN while COL-II and COL-X do not have upstream and downstream influences in the protein signaling network). Network images were designed with Cell Designer
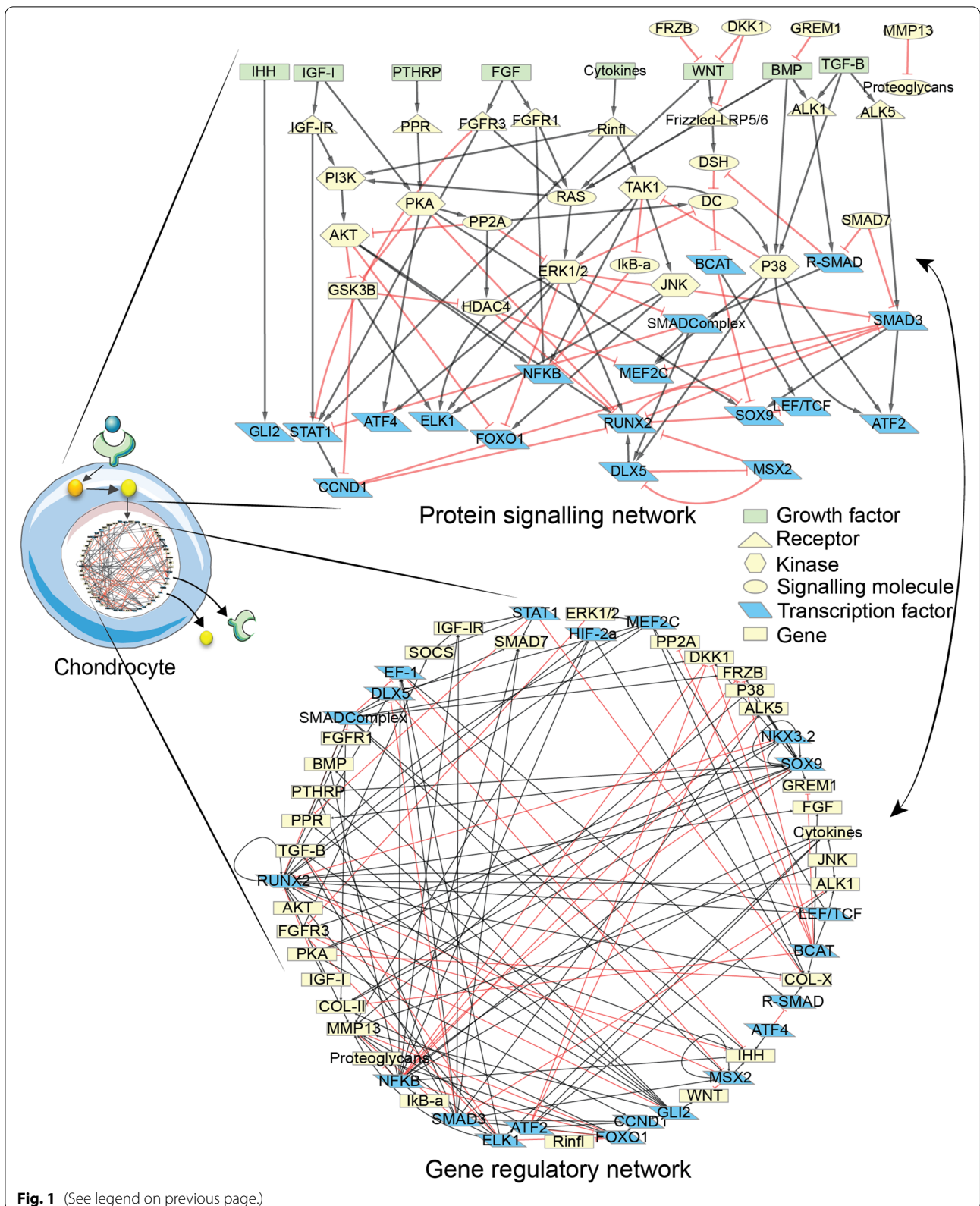
Lesage *et al. BMC Biology* (2022) 20:253

Page 4 of 25

Fig. 1 (See legend on previous page.)

Lesage *et al. BMC Biology*   (2022) 20:253

Page 5 of 25

which we selected a subset corresponding to the expression profile of 41 genes (see the "Methods" section and full list in Additional file 4 Data S1). 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 without adding new variables (see the "Methods" section). The sub-dataset was normalized and corrected for batch effects originating from the differences in microarray platform technologies, as described in the "Methods" section. Our data indicate that the gene expression distribution was correctly normalized among the different microarrays (Fig. 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) (Fig. 2B). See the "Methods" section and Additional file 4 Data S1 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 like the ones achieved when clustering each of the individual microarray experiments separately, before batch effect removal (Additional file 5 Fig. S1). 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.
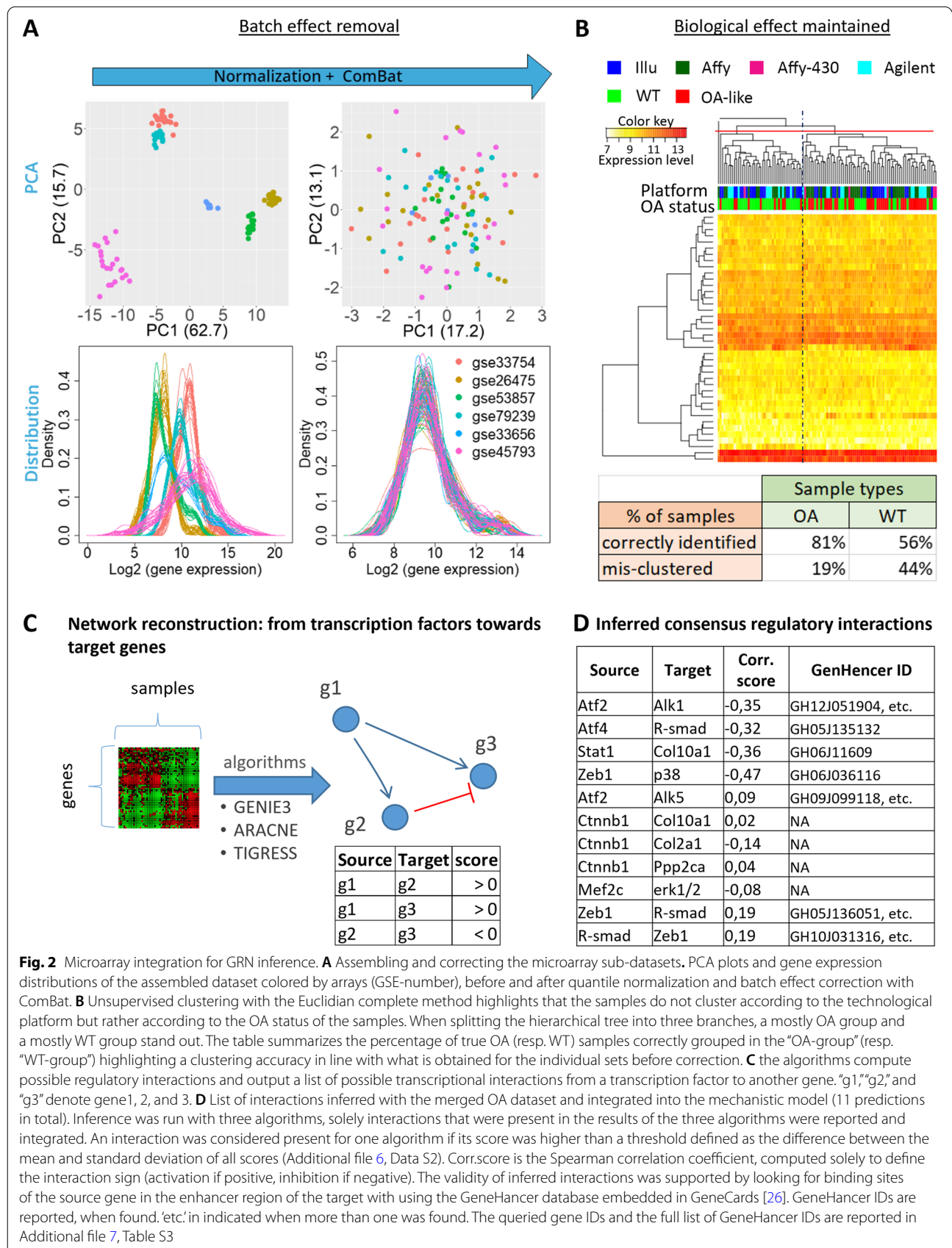
In order to complement the GRN with new data-derived hypothetical interactions, we inferred genetic interactions out of this newly assembled dataset. Directed edges between nodes were added in the GRN to account for newly inferred regulatory interactions, as represented in Fig. 2C. We inferred regulatory interactions directed from TFs towards target genes by using three different algorithms to avoid algorithm-specific bias on the inferred edges (see the "Methods" section) [27]. Only interactions predicted by all the three algorithms were implemented in the gene regulation layer of the mechanistic model. The matrix of inferred interactions is reported for each algorithm and for the consensus result in Additional file 6 Data S2. In addition, the sign of the Spearman correlation sc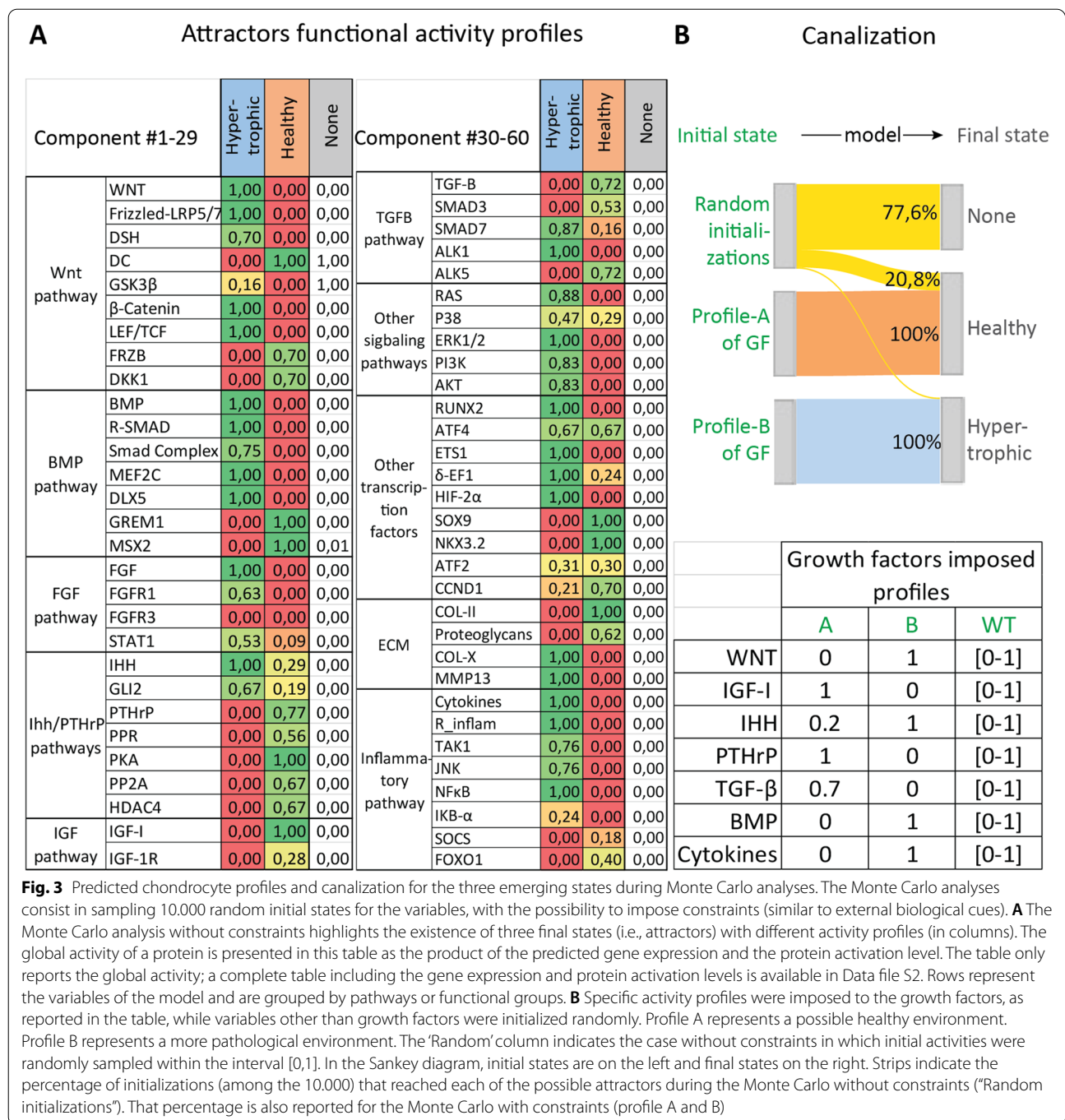ore allowed us to define whether the interactions were genetic activations or inhibitions (Fig. 2C). The inferred interactions are reported with the correlation factors in Fig. 2D. Note that the machine learning methods predicted some TF as regulator of a given target gene even for pairs that had not a high correlation score (Fig. 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.

The computational model captured two chondrocyte phenotypes and several physiologically relevant behaviors qualitatively

We translated the 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-translational level (global functional activity > 0) (see the "Methods" section, Additional file 2: supplementary computational method and Additional file 1 Table S1).

With the above-described numerical chondrocyte model, we studied the system free of fixed external cues to identify mathematically stable states that may emerge naturally. These stable states are also called attractors and they equate potentially existing cell phenotypes (see definition of attractors in Additional file 1 Table S1). They were evaluated using methodologies like those used with logical models [28]. 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 the "Methods" section for Monte Carlo analysis), we were able to explore possible model outcomes and we observed three emerging attractors that were singleton stable states. No cyclic attractors were obtained. Each of those states had a unique global activity profile for the 60 components as reported in Fig. 3A; the details of the protein activation and gene expression level for each component are available in Additional file 8 Data S3. Two of the attractor states that we found were biologically relevant, meaning they were comparable

Lesage *et al. BMC Biology*   (2022) 20:253

Page 6 of 25

Fig. 2 Microarray 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 ComBat. **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** the algorithms compute possible regulatory interactions and output a list of possible transcriptional interactions from a transcription factor to another gene. "g1," "g2," and "g3" denote gene1, 2, and 3. **D** List of interactions inferred with the merged OA dataset and integrated into the mechanistic model (11 predictions in total). 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 (Additional file 6, Data S2). Corr.score is the Spearman correlation coefficient, computed solely to define the interaction sign (activation if positive, inhibition if negative). The validity of inferred interactions was supported by looking for binding sites of the source gene in the enhancer region of the target with using the GeneHancer database embedded in GeneCards [26]. GeneHancer IDs are reported, when found. 'etc.' in indicated when more than one was found. The queried gene IDs and the full list of GeneHancer IDs are reported in Additional file 7, Table S3

Lesage *et al. BMC Biology* (2022) 20:253

Page 7 of 25

A — Attractors functional activity profiles

Component #1-29		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

Component #30-60		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 transcription 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
Inflammatory 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

B — Canalization

Initial state ——model⟶ Final state

Random initializations → None 77,6% → Healthy 20,8%
Profile-A of GF → Healthy 100%
Profile-B of GF → Hypertrophic 100%

	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]

Fig. 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 Data file S2. 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 states 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 and B)

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 states 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 (Fig. 3A and [8, 29–31]). In addition, the inflammation-associated variables had low activity. That correlates with what is known of real chondrocytes during homeostasis [32, 33]. The second state corresponded to a hypertrophic-like chondrocyte since factors such as RUNX2, COL-X, MMP13, and IHH 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

Lesage *et al. BMC Biology*   (2022) 20:253

Page 8 of 25

(functional activities nearly zero) (Fig. 3A and [4, 30, 31, 34–37]). 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) (Fig 3A and [12, 33, 38]). 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' state as it was neither healthy nor hypertrophic Fig. 3A.

The number of random initializations reaching an attractor during the Monte Carlo simulation gives a sense of the probably of reaching that state (Fig. 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 (Fig. 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 Fig. 3B, 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 (Fig. 3B).

As a first step to support the validity of our model, we established a tool enabling to test *in silico* scenarios and assess whether the model could successfully recapitulate relevant physiological behaviors. 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 [39, 40]. 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 concomitantly to activating the PTHrP-related pathway could push a hypertrophic-like chondrocyte into transitioning towards a more healthy or anabolic state (Additional file 9 Data S4). Moreover, other studies have reported that the TGF-β pathway had a protective effect against inflammation [41–44], 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 transitions towards the diseased hypertrophic state (Fig. 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 [45]. 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) [45] and impact chondrocyte maturation differentially [46]. Clinical observations reported that the ALK5/ALK1 balance decreased with age and in OA patients [47–49]. *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 $\frac{ALK1}{ALK5} > 1$ (Fig. 4 A and Additional file 10 Fig. S2). 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 (Additional file 10 Fig. S2). This is in line with what was previously modeled [41], 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 (available in GitHub [50]) allowing users, such as biologists, to easily test hypotheses by performing *in silico* experiments on the virtual chondrocyte (see interface Additional file 11 Fig. S3).

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 was the most potent condition to trigger a transition from the healthy towards the 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 (Fig. 4B). On the other hand, transition from the hypertrophic
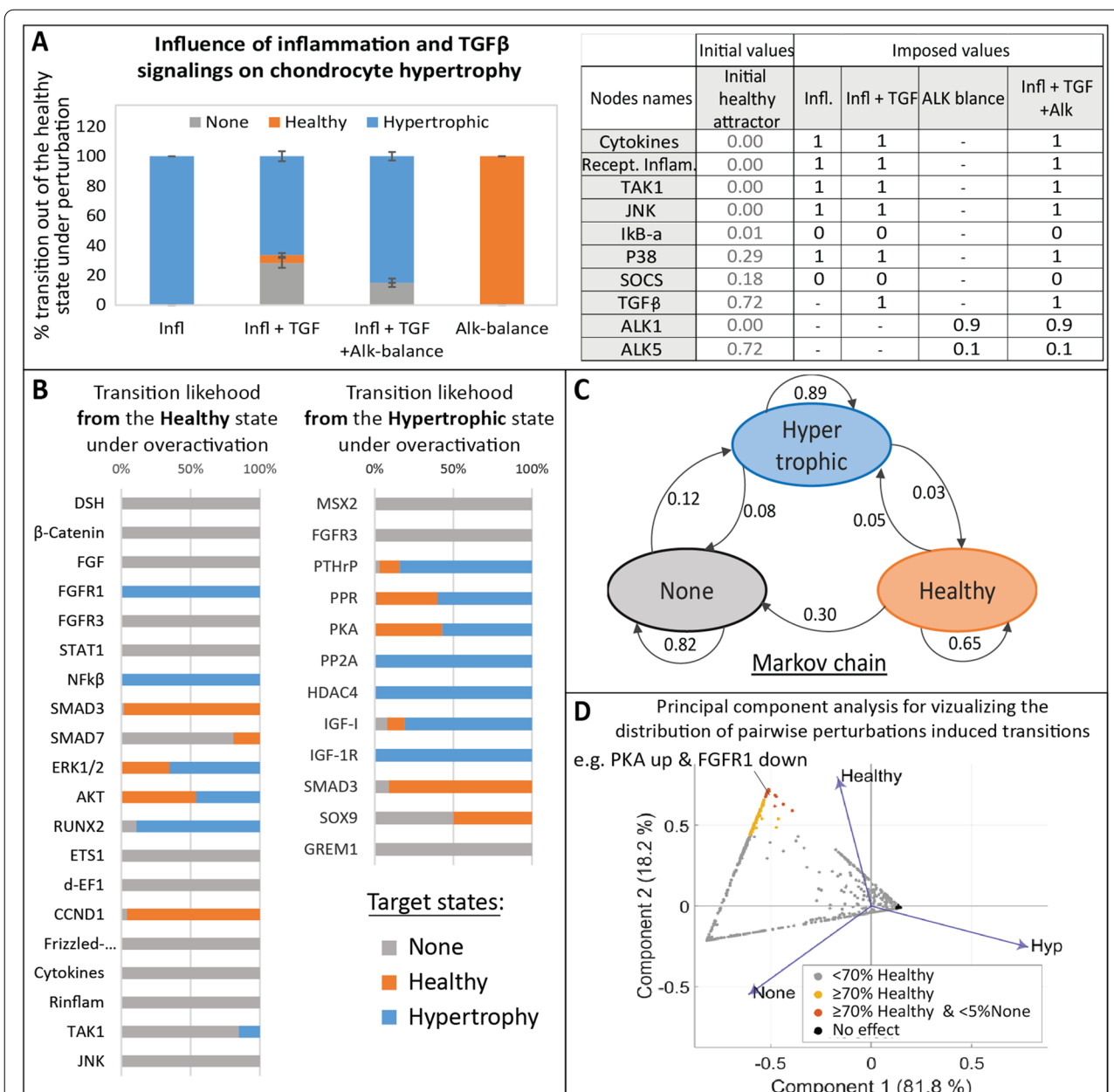
Fig. 4 Study of the virtual chondrocyte state transition and *in silico* screening of target perturbations. **A** Relation between Inflammation and TGFβ and influence on the chondrocyte state. Perturbations are applied on the healthy attractor, bar height gives the average percentage of transitions towards 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 towards a 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 3 attractors, reported as eigenvalues (blue arrows). Dot colors correspond to threshold in the percentage of transitions towards the healthy state for potential OA therapies. The details of the predicted conditions leading to 70% and up to 100% of transitions towards the healthy state are available in Additional file 9, Data S4, the ones that were selected for experimental validation are further described below.

Lesage *et al. BMC Biology*   (2022) 20:253

Page 10 of 25

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 (Fig. 4C) with the system's overall probability of transitioning from one state to another under random single-node perturbations. Interestingly, despite the low probability of reaching the hypertrophic state in the random canalization (Fig. 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 perturbations). The full screening amounted to $\binom{60}{2} \times 4 = 7080$ tested conditions per state. 94% of the variance was explained by the first two components in the PCA reporting the effect (percentage of transitions to destination attractors) of the combinatorial treatments on a hypertrophic-like chondrocyte (Fig. 4D). We searched for the most potent conditions to retrieve the healthy state from a hypertrophic chondrocyte (defined as conditions with more than 70% of perturbations 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 (see the complete list in Additional file 9 Data S4). 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 (Additional file 9 Data S4). Therefore, the PKA/PPR axis and SMAD3 seemed to be "enablers" that could "unblock" the system, facilitating the effect of other relevant targeting treatments, *in silico*. Moreover, some predictions among the ones triggering more than 70% of transitions towards the healthy state, in Fig. 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 transitions towards the healthy state. Additionally, inhibition of the WNT pathway while activating ALK5 as well as activation of IGF-1 while activating the destruction complex (DC) involved in the WNT pathway allowed between 70 and 90% of transitions towards the healthy state, to mention but a few (Additional file 9 Data S4). So, the *in silico* model and associated screening algorithms predicted pairwise targeting conditions with a potential role against chondrocyte hypertrophy. Those predictions can be used as an indicator to guide further validation experiments.

In vitro experiments support the role of newly predicted (combinatorial) treatments to prevent hypertrophy

We used ATDC5s, a chondrogenic murine cell line able to undergo hypertrophy as well as human OA chondrocytes, in order to *in vitro* validate the *in silico* findings obtained from the model. We measured the activity of alkaline phosphatase (ALP), an enzyme typically secreted during hypertrophy and participating to ECM

(See figure on next page.)

Fig. 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, positively linearly correlates with *Col10a1* gene expression during hypertrophic differentiation with and without Ihh treatment. Results of one representative experiment. 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 human 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 variables 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, gray 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 without significant decrease of ALP activity *in vitro*
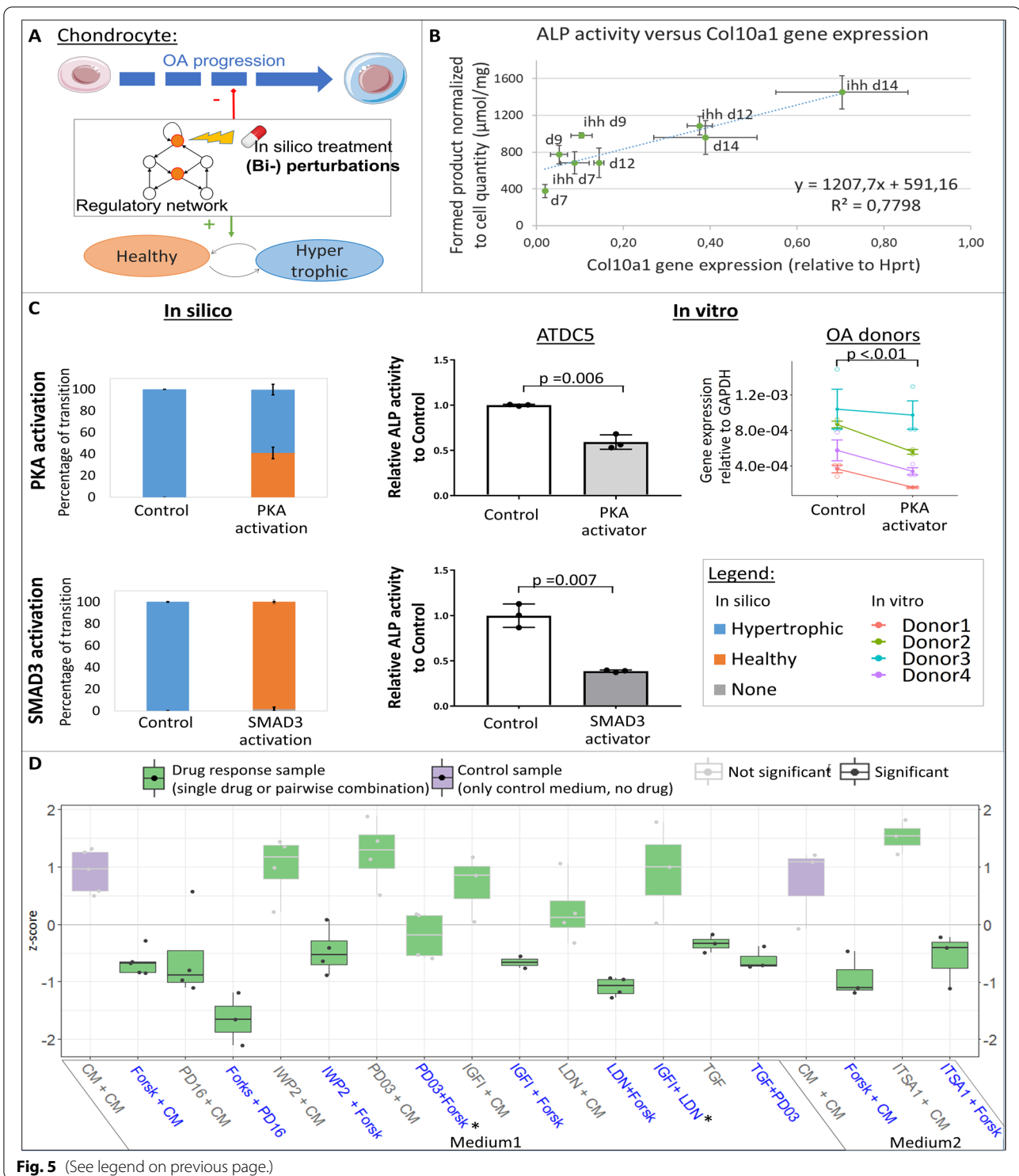
Fig. 5 (See legend on previous page.)

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 (Fig. 5B).

Hypertrophy was further increased when supplementing the differentiation medium with Ihh, as expected (Fig. 5B). Indeed, this growth factor is known for its pro-hypertrophic effect on ATDC5 [51]. These results

Lesage *et al. BMC Biology*   (2022) 20:253

Page 12 of 25

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 (Fig. 5C). Additionally, Forskolin treatment decreased *COL10A1* gene expression when applied to primary human OA chondrocytes (*p* < 0.01) (Fig. 5C). Together these results corroborated the *in silico* predictions that activation of PKA was sufficient to block hypertrophic differentiation in chondrocytes (Fig. 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. Additional file 9 file Data S4 describes the potent combinatorial conditions and their predictions that were selected for subsequent validation (PKA activation + FGFR1 inhibition, WNT inhibition + PKA activation, ERK1/2 inhibition + PKA activation, IGFIR + PKA activation, BMP inhibition + IGFI, BMP inhibition + PKA activation, ALK5 + ERK1/2 inhibition, HDAC4 activation + PKA activation). The chemical compounds used for each target are described in Additional file 12 Table S4.

The effect of those combinatorial treatments was assessed through evaluating ALP activity in ATDC5. ALP activity normalized to total DNA content is reported by means of *z*-score for all pairwise and single conditions in Fig. 5D. 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 the *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 (Fig. 5D).

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 (Fig. 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 transitions out of the hypertrophic state, showed that a minimal level of PKA activity (namely about 0.4 on a scale from 0 to 1) was required to achieve any positive effect with this combinatorial treatment (Fig. 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 (Fig. 6B). *In vitro* validation confirmed this dose effect since the overall gradient shape was very comparable between the *in silico* and *in vitro* situation (Fig. 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
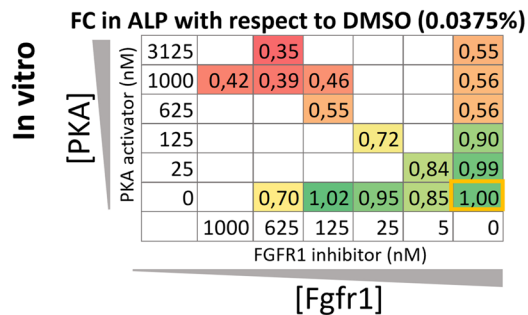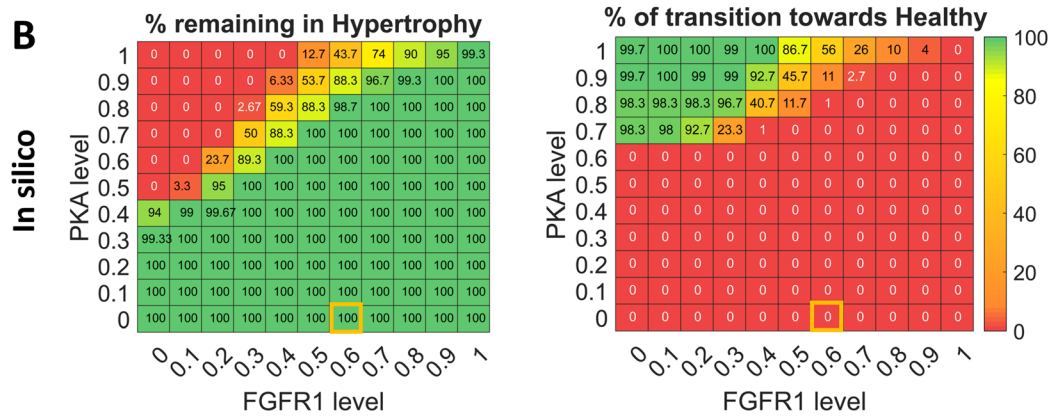
(See figure on next page.)

Fig. 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 (Forskolin, 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 transitions remaining in the hypertrophic state or transitioning towards the healthy state is reported in the upper panels, the rest of the transitions go 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

Fig. 6 (See legend on previous page.)

Lesage *et al. BMC Biology*   (2022) 20:253

Page 14 of 25

for a chondrocyte experiencing hypertrophy-inducing cues.

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 recapitulated physiologically relevant observations and predicted 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 of chondrocyte phenotype and anti-hypertrophic drug targets. The *in silico* target perturbation screening and the experimental validation of our findings show the potential of *in silico* experiments to guide *in vitro* experiments for target discovery. With further validation, these predictions might form the basis for successful OA treatment. In particular, our study points towards a possible 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 [52]. Therefore, there is a 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 (e.g., ODE-based models) and 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 of 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 using 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 [53]. 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 [54, 55]. We limited the inference to our genes of interest to gain knowledge on regulatory interactions without adding new nodes in the network, which would add more unknowns. Nevertheless, an alternative strategy could be to infer interactions between all genes from the dataset and integrate interactions involving our genes of interest, new nodes could be added if relevant feedback loops or non-linear paths were observed with some genes of interest.

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 [17]. 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 formalism 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 as the size of a model and quantity of

unknow parameters should be subject to a tradeoff and is a function of the model purpose [56]. Many biostatistics and machine learning methods that make drug efficacy-related predictions solely based on omics data still lack predictability and interpretability [57]. 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. Before carrying out new validation experiments, the relevance of our model to answer osteoarthritis-related questions was investigated 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 [41] 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 [58]. 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. 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 has suggested that restricting ourselves to initializing 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.

This study also suggests that chondrocyte phenotypes are not so sensitive to small, single-factor perturbations. Even though the resistance of the current *in silico* model

to single factor perturbations might, in part, be due the omission of some parts of the real world system's complexity. This could be subject to further experimental validation in chondrocytes. This trait of robustness to small environmental variations is often considered as a fundamental and ubiquitous trait of biological systems [59]. This trait allows systems to function in noisy environments [60] and systems biologists have theorized that disease may establish its own robustness, in some cases [61]. In line with that, the Markov chain in our computational model indicates that, for articular chondrocytes, the probability to change phenotype 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, the *in silico* pairwise perturbation screening confirms that targeting at least two components concomitantly increases the chance of unlocking 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 [10], but once the hypertrophic transition has been initiated, it is rather difficult to escape that fate.

Another important outcome is the approach to validate certain 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. Future research should focus on the further validation of this result, for instance through an *in vitro* dose curve relationships study and *in vivo* testing. Activation of the PThrP pathway, to which PKA belongs, has already been reported to be rather anti-hypertrophic for growth plate chondrocytes [62]. Similarly, genetic inhibition of the Fgfr1 gene in mouse knee cartilage has been shown to attenuate the degeneration of articular cartilage in mice [63]. However, to our knowledge, this combination has never been investigated nor reported before for its synergistic potential against hypertrophy, cartilage degradation or OA.

Some predictions (IGF-1 combined with BMP inhibition, and inhibition of ERK1/2 combined with PKA activation) could not be validated experimentally. This could be explained by several factors, such as the limitations of the *in vitro* model (ALP measurement in ATDC5), the

omission of important regulatory mechanisms, or inaccurate assumptions in the *in silico* model, possibly pointing at necessary corrections in the regulatory network. Importantly, here we propose a strategy in which *in silico* predictions are used in an exploratory discovery phase. Hence, reducing the number of false positives is less important than increasing the true positives and reducing the false negatives - which hampers the discovery process.

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 number 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 [64]. Nevertheless, it is important to note that this mouse cell line is cultured in isolation from a physiological environment. Therefore, we propose to use it within an *in silico-in vitro* pipeline that would act like a funnel with sequential filters 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. The next step could be to test the filtered conditions *in vivo* to evaluate their effect on the disease progression.

We acknowledge that our study has some limitations. A 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 number of interactions we would integrate but it alleviated each method's weakness and reinforced the strengths of the predictions, as previously proved [27]. 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 [22]. 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. Moreover, 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. Finally, the mechanisms integrated in the network model include

a certain number of assumptions and regulatory relationships that represent the current state of the knowledge; however, those assumptions and regulations could still be updated and refined as knowledge grow.

Conclusions

This study is a proof-of-concept to showcase how an *in silico-in vitro* integrated approach can suggest single and combinatorial target perturbations affecting the hypertrophic transition and help to prioritize experiments for therapeutic target discovery in OA research. In that sense, our model offers the possibility to make hypotheses 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 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. Our study highlights targets - such the concomitant activation of PKA and inhibition of FGFR1 or BMP - that deserve additional investigation. With further validation, these conditions might form the basis of a successful OA treatment. 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 the 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 [65, 66], 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. The resulting network models and the effect of target perturbations could be compared across the different patient type-specific models.

In conclusion, this work provides a virtual articular chondrocyte in the form of a signal transduction interactive knowledge base and of an executable computational

Lesage *et al. BMC Biology*   (2022) 20:253

Page 17 of 25

model. The demonstrated *in silico-in vitro* strategy opens new routes for studying OA and discovering targeting therapies by refining the early stages of drug discovery.

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 [23] 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 [24]. See the links for the chondrocyte knowledge base in the "Availability of data and materials" section (separated in two subnetworks). 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 Data S1. The datasets were merged applying an in-house developed pipeline adapted from previously published methods [67]. 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 as listed in Additional file 4: Data S1. This allowed finding new regulatory relationships without adding new variables to limit the model's size. 41 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) × 109 samples (or observations)). This dataset was quantile normalized and the batch effects were removed through the ComBat algorithm based on Bayes methods [68].

Inference was performed for regulatory interactions within the aforementioned subset of genes. This was achieved by employing three different algorithms, being ARACNE, TIGRESS, and GENIE3 [69–71]; 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. The validity of the inferred interactions was corroborated by searching the GeneCards database [26]. When one or several binding sites for the source gene (transcription factor) in the enhancer region of the target was found with GeneCards, one of the GeneHancer ID was reported (Fig. 2) and the full list with the exact gene IDs that were queried is in Additional file 7: Table S3.

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/) [72].

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 [73–75]. This additive formalism resembles the Boolean threshold networks [76] and was previously implemented with priority classes as described by Kerkhofs and Geris [22]. In this formalism, the biological components were represented by variables evolving over pseudo-time steps. The model is semi-quantitative since the variables could take on a continuous activity value between 0 and 1. The evolution of all 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 Additional file 1 Table S1), in some exceptional cases (particular mechanisms) a product of several regulators was used (see Additional file 2, Supplementary computational method). Biological influences could happen at

Lesage *et al. BMC Biology*   (2022) 20:253

Page 18 of 25

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 Additional file 3 Table S2 for the definition of fast and slow reactions and variable). A formal description of the mathematical system underlying the model as well as the full list of equations is available in Additional file 2.

The asymptotic solutions were evaluated with a Monte Carlo simulation procedure, like methods employed for logical models [28, 75]. When running a simulation (also see the 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 definition in Additional file 1 Table S1) 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 Additional file 13 Fig. S4 and Additional file 14 S5 for graphical explanations about the algorithm and simulation scheme on a reduced illustrative example network. A stable state (definition in Additional file 1 Table S1) 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 input 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

A Monte Carlo canalization estimated the nature of the attractors and their reachability, given the regulatory network provided in the equations. In short, all variables were initialized 10.000 times with random values in the interval [0,1] simulations converged towards several attractors (only singleton attractors were observed). We considered that two simulations reached the same unique stable state when the absolute difference between both final states was less than a tolerance set to 10^{-2} for all the 60 components since such a difference. The number of initializations reaching each final state was computed and reported in terms of percentage of initial states. This number gives a sense of the probability of reaching the

state for the unperturbed system, i.e., without constraints [28, 75]. The number of initializations (10.000) was considered sufficient to estimate the state space of the system as higher numbers had little influence on the canalization results (see sensitivity analysis Additional file 15 Fig. S6). 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 meant to represent a normal healthy or a disease environment (profile A and B of Fig. 3B, respectively). The networkD3 package from R was used to produce the Sankey diagram for the visualization of the canalization results.

In silico target perturbations and selection

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 number 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 and going further than 1000 time-steps would not induce further changes in the result. 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). A representative example of the pseudo-time evolution of representative variables (RUNX2, SOX9, MMPs, Collagen, etc.) after an input perturbation (PKA activation + FGFR1 inhibition) is provided in Additional file 16 Fig. S7 to illustrate the choice of the perturbation duration.

The different *in silico* scenarios or treatment experiments that were tested amounted to perturbing one or several variables, 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 condition 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 transitions towards each of the possible target states. Standard deviation in the percentage of transitions 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. The independent simulations were run in parallel using high-performance computing infrastructure of the KU Leuven (Vlaams Supercomputer Centrum). The selection of potent conditions against hypertrophy that could be validated experimentally was done in two main steps. First, we automatically selected combinatorial conditions for which at least 70% of the perturbations triggered a transition from the hypertrophic state towards the healthy one. Among them, we focused on those conditions with more than 70% of transitions towards the healthy state but less than 5% towards the 'None' state as a first discriminatory factor. Conditions were classified by ranges of percentages of transitions towards the healthy state (100-90%, 89-80%, 79-70%), see Additional file 9 Data S4. Second, among those potent conditions, some were further selected for their druggability and their ease to be tested in a simple *in vitro* system. The selection criteria also involved additional elements such as the readily availability of the necessary small molecules to target the predicted component in the expected way, the availability of literature to define appropriate concentrations, and the variety of combinatorial conditions tested. 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, or for conditions that were too similar (e.g., PKA activation+ ERK1/2 inhibition *vs*. PPR activation + ERK1/2 inhibition, PPR being directly upstream of PKA in the signaling network) only the one of the two was selected.

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/cm^2 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., Fig. 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 days 0, 7, 9, 12, and 14 during differentiation (TRIzol reagent; Thermo Fisher Scientific), in addition to the ALP activity assay and DNA quantification.

To assess the effect of small molecule and growth factor 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-1 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 the 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 (Fig. 5), and with 0.0375% DMSO for the Forskolin/PD161570 synergy study and dose screening (Fig. 6).

Validating treatment effects in primary human chondrocyte 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

Lesage *et al. BMC Biology*   (2022) 20:253

Page 20 of 25

2 h 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 7500 cells/cm^2 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 trypsinized 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 real-time quantitative PCR (RT-qPCR) for ATDC5 samples

Gene expression of *Col10a1* in ATDC5 experiments was evaluated by RT-PCR. For RNA isolation, chloroform was added to the TRIzol samples (TRIzol 5: Chloroform 1), which were subsequently centrifuged for 15min at 15,000 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 15,000 rpm (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 (primer sequences available in [77]). 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 s 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.

RNA isolation and RT-qPCR for human primary OA chondrocytes

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 the 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, USA). 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. Data and primers are available in [77].

Alkaline phosphatase assay (ALP)

Enzymatic activity of secreted ALP in the supernatant medium of ATDC5 cultures was determined in a colorimetric assay as previously described [78]. 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

Lesage *et al. BMC Biology*   (2022) 20:253

Page 21 of 25

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 s, centrifuged for 1′ at 13,000 rpm (i.e. RCF= 164,380) 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 analyze *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 Forskolin and Activin treatments were compared to control in one representative experiment with 3 replicates thanks to a one-tailed unpaired $\underline{t}$-test with Welch's correction, in Fig. 5C. Effect of Forskolin treatment in OA chondrocytes from human donors was performed in triplicates for 4 donors. Comparison 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 (Fig. 5D) were performed by modifying the BraDiPluS package from the Saez Lab [79]. More precisely, the probability that a treated condition resulted in a 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 the Shapiro-Wilk test), with Benjamini–Hochberg correction for multiple testing. Three independent experiments were performed for each condition, and treatment 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 condition was considered lower than the control for p-value < 0.05 and z-score <0.

Abbreviations

ALK1: Type II TGFβ receptor ALK1; ALK5: Type I TGFβ receptor; ALP: Alkaline phosphatase; BMP: Bone morphogenic protein; COL-II: Type II collagen; COL-X: Type X collagen; DC: Destruction complex; DKK1: Dickkopf1; DMEM: Dulbecco's modified Eagle's medium; DMSO: Dimethylsulfoxyde; DNA: Deoxy-ribonucleic acid; ECM: Extracellular matrix; ERK1/2: Extracellular signal-regulated kinase 1&2; FGF: Fibroblast growth factors; FGFR1: FGF receptor 1; GRN: Gene regulatory network; HK : Housekeeping; IGF-I: Insulin-like growth factor; IHH: Indian hedgehog; IL-1B: Interleukin 1beta; JNK: c-Jun N-terminal kinase; MAPK: Mitogen-activated protein kinases; MEF2C: Myocyte enhancer factor 2C; MMP13: Matrix metalloproteinase 13; NFKB: Nuclear factor kappa B; OA: Osteoarthritis; ODE: Ordinary differential equations; PI3K: Phosphatidylinositol 3-kinase; PKA: Protein kinase A; PPR: PTHrP receptor; PTHrP: Parathyroid hormone-related peptide; RNA: Ribonucleic acid; RT-qPCR : Reverse transcriptase-quantitative polymerase chain reaction; RUNX2: Runt-related transcription factor 2; SOX9: SRY-Box Transcription Factor; STAT1: Signal transducer and activator of transcription; TAK1: TGF-β-activating kinase; TCF7: Transcription factor 7; TF: Transcription factors; TGF : Transforming growth factors; TNF-A: Tumor necrosis factor alpha.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-022-01451-8.

Additional file 1: Table S1. Definitions. Definitions of technical terms pertaining to biological signaling network modeling with a semi-quantitative additive formalism.

Additional file 2. Supplementary computational method. Equations, mathematical framework and justification of deviation from the general rule in the equations.

Additional file 3: Table S2. List of variables and mouse genes correspondence. 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 nomenclatures. 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, this is not exhaustive since some variables represent a group of factors or a family of ligands rather than a single factor.

Additional file 4: Data S1. Subset of genes for gene expression profiles and description of GEO datasets with manual binary annotations.

Additional file 5: Fig. S1. Individual dataset clustering and heatmap. 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 (see Additional file 4, Data S1.). 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 (see Additional file 4, Data S1). 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.

Additional file 6: Data S2. GRN inference and validation. (Microsoft Excel Worksheet). The GRN inference with the 3 algorithms and the consensus matrix are reported. We have set stringent rules and only interactions that were inferred by the 3 algorithms were included in the mechanistic

Lesage *et al. BMC Biology*   (2022) 20:253

Page 22 of 25

model: those interactions are all reported in Fig. 2D. The validity of the inferred interactions was investigated by searching the GeneCard database. The relevant GeneHancer IDs are reported when supportive evidence was found.

Additional file 7: Table S3. Validation of transcriptional data-inferred regulatory interactions. Transcriptional interactions were inferred from the merged OA dataset. Inference was run with three algorithms and only interactions that were present in the results of the three algorithms (GENIE3, ARACNE, TIGRESS) were considered as additions to the model. An interaction was considered present for one algorithm if it scored higher than a threshold defined as the difference between the mean and standard deviation of all scores (Additional file 6, Data S2). Inferred interactions were validated whenever possible by looking for binding sites of the source transcription factor in the enhancer region of the target gene with GeneCards. The complete list of GeneHancer IDs provided by GeneCards are reported in the last column. 'NA' indicate that no binding site was found. The exact gene names and gene IDs that were queried in GeneCards are reported in the first columns.

Additional file 8: Data S3. Attractors complete protein and gene expression profiles. (Microsoft Excel Worksheet). The predicted profiles of the 3 attractors are reported for all variables. The dataset contains the values of the fast sub-variables (i.e. protein subnetwork), slow sub-variables (i.e. gene subnetwork) and the global functional activity.

Additional file 9: Data S4. In silico screening predictions of combinatorial treatments. (Microsoft Excel Worksheet). This file reports the results of a screening of combinatorial perturbations. More particularly, it reports all conditions that lead to a transition towards a healthy chondrocyte (SOX9+) when starting from a hypertrophic-like chondrocyte (Runx2+). The first sheet reports conditions leading to such transition 100% of the time, between 99 and 90% of the time for the next sheet and so on. The TFs sheet references all the transcription factors of the model that could be retrieved in the result and that are less convenient to target experimentally (to define exclusion criteria for the experimental validation). The 'Summary' sheet, reports a list of conditions that might be interesting to test. It excludes conditions involving TFs.

Additional file 10: Fig. S2. Effect of ALKs ratio in the influence 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.

Additional file 11: Fig. S3. 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). 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 transitions 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. The application can be installed with the executable file on

the GitHub repository, [https://github.com/Rapha-L/Insilico_chondro.git]. No Matlab license is required, however, the operating system should be able to support the Matlab software. For Linux users, a Windows virtual machine may be used.

Additional file 12: Table S4. Targets and associated small molecules or growth factors for in vitro validation. The first row indicates the targets 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°').

Additional file 13: Fig. S4. 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, with a predefined tolerance interval. 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 and in a random order. For some systems (i.e. set of equations) cyclic attractors may arise, meaning that the system never reaches a fixed stable state but oscillates between several states. This situation did not occur in the current study.

Additional file 14: Fig. S5. Illustration of the algorithm for the asynchronous updating of variables with a simplified (3 nodes) example network. The network represents interactions happening in one of the subnetworks (protein= fast reactions or genetic = slow reactions). Inhibitions are represented in red and activations 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. S4.). However 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. The example illustrates that different initial states may reach the same final state but also that the same initial state (e.g. [1 0 1]) can reach different final states, depending on the order in which variables are updated, thereby introducing stochasticity in the system.

Additional file 15: Fig. S6. Sensitivity analysis: impact of the number of initialization on the canalization results during the Monte Carlo. The percentage of random initialization reaching each attractor is displayed with the Sox9 positive state (healthy) in orange, the Runx2 positive state in blue and the None in grey. Data labels indicate the absolute amount of state reaching the attractors. None of initializations reached an alternative attractor, even for higher amount of random initializations The number of initialization has no significant impact on the basal canalization and 10.000 initializations were considered sufficient to screen the state space in the current study.

Additional file 16: Fig. S7. Pseudo-time evolution of variables during simulations. The sequence of variable updating over each time steps after introducing a perturbation from the healthy state was saved as a time-series and plotted. It shows the discrete behavior of the simulation and that a steady state is reached way before the duration of the perturbation (1000 time steps) is reached in such a way that maintain the perturbation longer would not change the output of the simulation once the perturbation is released.

Lesage *et al. BMC Biology*   (2022) 20:253

Page 23 of 25

Authors' contributions

Conception: RL and LG. Design of the work: RL, MNFB, LG, GJVM, RN, and TW. Acquisition/analysis: RL and MNFB. Interpretation of data: RL, MNFB, LG, GJVM, and RN. Creation of new software: RL. Writing: RL and MNFB. Supervision and revision: LG, GJVM, RN, and TW. All authors read and approved the final manuscript.

Authors' information

@Raphaellelesag2
@BioNarci (for Roberto Narcisi)
@LiesbetGeris

Availability of data and materials

All data generated or analyzed during this study are included in this published article, its supplementary information and data files, and publicly available repositories as detailed below. The Matlab code used to simulate and analyze the model as well as the R code to analyze the experimental data and the user-friendly virtual chondrocyte application are available on the GitHub repository [https://github.com/Rapha-L/Insilico_chondro.git] [50]. The knowledge bases for the chondrocyte protein signaling (resp. gene regulatory) network are available through the Cellcollective online platform [https://cellcollective.org/#a5b66073-6769-4c88-bf6b-37ca1aa8f766] [80] (resp. [https://cellcollective.org/#474de240-8752-4c3b-aa63-23640e50bf7a] [81]). The network model is also available in the form of an SBMLqual files [82] in the GitHub repository [50], allowing modelers to easily import the topology of the networks to their favorite software. However, to reproduce the simulations results presented here, the in-house scripts and algorithms provided in [50] must be used. The datasets supporting the conclusions of this article are included within the article and its additional files. In particular, experimental validation data and relevant scripts for their analysis are available in a public repository [https://github.com/Rapha-L/Experimental_validation_for_insilicoOA] [77]. The microarray data that support the findings of this study are available in Gene Expression Omnibus database [https://www.ncbi.nlm.nih.gov/geo/]. These data were derived from the following resources available in the public domain: GSE26475 [83], GSE33754 [84], GSE79239 [85], GSE33656 [86], GSE53857 [87], GSE4579 3[88].

Declarations

Ethics approval and consent to participate

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.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflicts of interest.

Author details

¹Prometheus, Division of Skeletal Tissue Engineering, KU Leuven, Leuven, Belgium. ²Biomechanics Section, KU Leuven, Leuven, Belgium. ³Department of Orthopaedics and Sports Medicine, Erasmus MC, University Medical Center, Rotterdam, the Netherlands. ⁴Orthopedic Surgery Department, UMC+, Maastricht, the Netherlands. ⁵Department of Otorhinolaryngology, Erasmus MC, University Medical Center, Rotterdam, the Netherlands. ⁶Department of Biomechanical Engineering, Delft University of Technology, Delft, the Netherlands. ⁷GIGA In silico Medicine, University of Liège, Liège, Belgium.

References

1. Karsdal MA, Michaelis M, Ladel C, Siebuhr AS, Bihlet AR, Andersen JR, et al. Disease-modifying treatments for osteoarthritis (DMOADs) of the knee and hip: lessons learned from failures and opportunities for the future. Osteoarthr Cartil. 2016;24:2013–21.
2. Raman S, FitzGerald U, Murphy JM. Interplay of inflammatory mediators with epigenetics and cartilage modifications in osteoarthritis. Front Bioeng Biotechnol. 2018;6:22.
3. Ji Q, Zheng Y, Zhang G, Hu Y, Fan X, Hou Y, et al. Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. Ann Rheum Dis. 2019;78:100–10.
4. Von Der Mark K, Kirsch T, Nerlich A, Kuss A, Weseloh G, Glückert K, et al. Type x collagen synthesis in human osteoarthritic cartilage. indication of chondrocyte hypertrophy. Arthritis Rheum. 1992;35:806–11.
5. Tchetina EV, Poole AR, Zaitseva EM, Sharapova EP, Kashevarova NG, Taskina EA, et al. Differences in Mammalian Target of Rapamycin Gene Expression in the Peripheral Blood and Articular Cartilages of Osteoarthritic Patients and Disease Activity. Arthritis. 2013;2013:46148.
6. Poehling GG. Osteoarthritis: Diagnosis and medical/surgical management, third edition. Arthroscopy. 2002;18:678.
7. Tchetina EV. Developmental Mechanisms in Articular Cartilage Degradation in Osteoarthritis. Arthritis. 2011;2011(68397):1–16.
8. Van der Kraan PM, Van den Berg WB. Chondrocyte hypertrophy and osteoarthritis: Role in initiation and progression of cartilage degeneration? Osteoarthr Cartil. 2012;20:223–32.
9. Ripmeester EGJ, Timur UT, Caron MMJ, Welting TJM. Recent Insights into the Contribution of the Changing Hypertrophic Chondrocyte Phenotype in the Development and Progression of Osteoarthritis. Front Bioeng Biotechnol. 2018;6:18.
10. Ferrao Blanco MN, Bastiaansen-Jenniskens YM, Chambers MG, Pitsillides AA, Narcisi R, van Osch GJVM. Effect of Inflammatory Signaling on Human Articular Chondrocyte Hypertrophy: Potential Involvement of Tissue Repair Macrophages. Cartilage. 2021;13(2_suppl):168S–74S.
11. Cleary MA, Van Osch GJVM, Brama PA, Hellingman CA, Narcisi R. FGF, TGFβ and Wnt crosstalk: Embryonic to in vitro cartilage development from mesenchymal stem cells. J Tissue Eng Regen Med. 2015;9:332–42.
12. Monteagudo S, Lories RJ. Cushioning the cartilage: A canonical Wnt restricting matter. Nat Rev Rheumatol. 2017;13:670–81.
13. Green S, Şerban M, Scholl R, Jones N, Brigandt I, Bechtel W. Network analyses in systems biology: new strategies for dealing with biological complexity. Synthese. 2018;195:1751–77.
14. Rationalizing combination therapies. Nat Med. 2017;23:1113. https://doi.org/10.1038/nm.4426.
15. Nalesso G, Sherwood J, Bertrand J, Pap T, Ramachandran M, de Bari C, et al. WNT-3A modulates articular chondrocyte phenotype by activating both canonical and noncanonical pathways. J Cell Biol. 2011;193:551–64.
16. Kitano H. Systems biology: A brief overview. Science. 1979;2002(295):1662–4.
17. Puniya BL, Allen L, Hochfelder C, Majumder M, Helikar T. Systems Perturbation Analysis of a Large-Scale Signal Transduction Model Reveals Potentially Influential Candidates for Cancer Therapeutics. Front Bioeng Biotechnol. 2016;4:10.
18. Herberg M, Roeder I. Computational modelling of embryonic stem-cell fate control. Development. 2015;142:2250–60.
19. Schivo S, Khurana S, Govindaraj K, Scholma J, Kerkhofs J, Zhong L, et al. ECHO, the executable CHOndrocyte: A computational model to study articular chondrocytes in health and disease. Cell Signal. 2020;68:109471.

Lesage *et al. BMC Biology*   (2022) 20:253

Page 24 of 25

20. Khurana S, Schivo S, Plass JRM, Mersinis N, Scholma J, Kerkhofs J, et al. An ECHO of Cartilage: In Silico Prediction of Combinatorial Treatments to Switch Between Transient and Permanent Cartilage Phenotypes With Ex Vivo Validation. Front Bioeng Biotechnol. 2021;9:1103.

21. Eduati F, Jaaks P, Wappler J, Cramer T, Merten CA, Garnett MJ, et al. Patient-specific logic models of signaling pathways from screenings on cancer biopsies to prioritize personalized combination therapies. Mol Syst Biol. 2020;16:1–13.

22. Kerkhofs J, Geris L. A Semiquantitative Framework for Gene Regulatory Networks: Increasing the Time and Quantitative Resolution of Boolean Networks. PLoS One. 2015;10:e0130033.

23. Kerkhofs J, Leijten J, Bolander J, Luyten FP, Post JN, Geris L. A qualitative model of the differentiation network in chondrocyte maturation: A holistic view of chondrocyte hypertrophy. PLoS One. 2016;11:1–27.

24. Helikar T, Kowal B, McClenathan S, Bruckner M, Rowley T, Madrahimov A, et al. The Cell Collective: Toward an open and collaborative approach to systems biology. BMC Syst Biol. 2012;6:96.

25. Thielen N, van der Kraan P, van Caam A. TGFβ/BMP Signaling Pathway in Cartilage Homeostasis. Cells. 2019;8:969.

26. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Stein TI, et al. GeneHancer: Genome-wide integration of enhancers and target genes in GeneCards. Database. 2017;2017:1–17.

27. Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, Camacho DM, et al. Wisdom of crowds for robust gene network inference. Nat Methods. 2012;9:796–804.

28. Abou-Jaoudé W, Traynard P, Monteiro PT, Saez-Rodriguez J, Helikar T, Thieffry D, et al. Logical Modeling and Dynamical Analysis of Cellular Networks. Front Genet. 2016;7:94.

29. Nishimura R, Hata K, Nakamura E, Murakami T, Takahata Y. Transcriptional network systems in cartilage development and disease. Histochem Cell Biol. 2018;149:353–63.

30. Mitchell PG, Magna HA, Reeves LM, Lopresti-Morrow LL, Yocum SA, Rosner PJ, et al. Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage. J Clin Investig. 1996;97:761–8.

31. Martinez-Sanchez A, Dudek KA, Murphy CL. Regulation of human chondrocyte function through direct inhibition of cartilage master regulator SOX9 by microRNA-145 (miRNA-145). J Biol Chem. 2012;287:916–24.

32. Shlopov BV, Gumanovskaya ML, Hasty KA. Autocrine regulation of collagenase 3 (matrix metalloproteinase 13) during osteoarthritis. Arthritis Rheum. 2000;43:195–205.

33. Chow YY, Chin KY. The Role of Inflammation in the Pathogenesis of Osteoarthritis. Mediators Inflamm. 2020;2020:8293921.

34. Mak KK, Kronenberg HM, Chuang PT, Mackem S, Yang Y. Indian hedgehog signals independently of PTHrP to promote chondrocyte hypertrophy. Development. 2008;135:1947–56.

35. Chen X, Macica CM, Nasiri A, Broadus AE. Regulation of articular chondrocyte proliferation and differentiation by Indian hedgehog and parathyroid hormone-related protein in mice. Arthritis Rheum. 2008;58:3788–97.

36. Lin AC, Seeto BL, Bartoszko JM, Khoury MA, Whetstone H, Ho L, et al. Modulating hedgehog signaling can attenuate the severity of osteoarthritis. Nat Med. 2009;15:1421–5.

37. Kamekura S, Kawasaki Y, Hoshi K, Shimoaka T, Chikuda H, Maruyama Z, et al. Contribution of runt-related transcription factor 2 to the pathogenesis of osteoarthritis in mice after induction of knee joint instability. Arthritis Rheum. 2006;54:2462–70.

38. Leijten JCH, Bos SD, Landman EBM, Georgi N, Jahr H, Meulenbelt I, et al. GREM1, FRZB and DKK1 mRNA levels correlate with osteoarthritis and are regulated by osteoarthritis-associated factors. Arthritis Res Ther. 2013;15:R126.

39. Rowan A, Hui W, Cawston TE, Richards CD. Adenoviral gene transfer of interleukin-1 in combination with oncostatin M induces significant joint damage in a murine model. Am J Pathol. 2003;162:1975–84.

40. Saito T, Fukai A, Mabuchi A, Ikeda T, Yano F, Ohba S, et al. Transcriptional regulation of endochondral ossification by HIF-2α during skeletal growth and osteoarthritis development. Nat Med. 2010;16:678–86.

41. Hodgson D, Rowan AD, Falciani F, Proctor CJ. Systems biology reveals how altered TGFβ signalling with age reduces protection against proinflammatory stimuli. PLoS Comput Biol. 2019;15:e1006685.

42. Proctor CJ, Macdonald C, Milner JM, Rowan AD, Cawston TE. A computer simulation approach to assessing therapeutic intervention points for the prevention of cytokine-induced cartilage breakdown. Arthritis Rheum. 2014;66:979–89.

43. Hui W, Rowan AD, Cawston T. Transforming growth factor-β1 blocks the release of collagen fragments from bovine nasal cartilage stimulated by oncostatin M in combination with IL-1α. Cytokine. 2000;12:765–9.

44. Shull MM, Ormsby I, Kier AB, Pawlowski S, Diebold RJ, Yin M, et al. Targeted disruption of the mouse transforming growth factor-β1 gene results in multifocal inflammatory disease. Nature. 1992;359:693–9.

45. Van Der Kraan PM. The changing role of TGFβ in healthy, ageing and osteoarthritic joints. Nat Rev Rheumatol. 2017;13:155–63.

46. Ferguson CM, Schwarz EM, Reynolds PR, Puzas JE, Rosier RN, O'Keefe RJ. Smad2 and 3 mediate transforming growth factor-β1-induced inhibition of chondrocyte maturation. Endocrinology. 2000;141:4728–35.

47. Hui W, Young DA, Rowan AD, Xu X, Cawston TE, Proctor CJ. Oxidative changes and signalling pathways are pivotal in initiating age-related changes in articular cartilage. Ann Rheum Dis. 2016;75:449–58.

48. Van Der Kraan P, Matta C, Mobasheri A. Age-Related Alterations in Signaling Pathways in Articular Chondrocytes: Implications for the Pathogenesis and Progression of Osteoarthritis - A Mini-Review. Gerontology. 2016;63:29–35.

49. Blaney Davidson EN, Remst DFG, Vitters EL, van Beuningen HM, Blom AB, Goumans M-J, et al. Increase in ALK1/ALK5 Ratio as a Cause for Elevated MMP-13 Expression in Osteoarthritis in Humans and Mice. J Immunol. 2009;182:7937–45.

50. Lesage R. Github - Rapha-L/Insilico_chondro: Scripts & data used to run the analyses done in the publication. GitHub. 2021. https://doi.org/10.5281/zenodo.7144317.

51. Akiyama H, Shigeno C, Iyama KI, Ito H, Hiraki Y, Konishi J, et al. 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. 1999;257:814–20.

52. Terfve C, Cokelaer T, Henriques D, MacNamara A, Goncalves E, Morris MK, et al. CellNOptR: A flexible toolkit to train protein signaling networks to data using multiple logic formalisms. BMC Syst Biol. 2012;6:133.

53. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A. Reverse engineering of regulatory networks in human B cells. Nat Genet. 2005;37:382–90.

54. Liu A, Trairatphisan P, Gjerga E, Didangelos A, Barratt J, Saez-Rodriguez J. From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL. NPJ Syst Biol Appl. 2019;5:1–10.

55. Türei D, Korcsmáros T, Saez-Rodriguez J. OmniPath: Guidelines and gateway for literature-curated signaling pathway resources. Nat Methods. 2016;13:966–7.

56. Karlebach G, Shamir R. Modelling and analysis of gene regulatory networks. Nat Rev Mol Cell Biol. 2008;9:770–80.

57. Yang M, Simm J, Lam CC, Zakeri P, Van Westen GJP, Moreau Y, et al. Linking drug target and pathway activation for effective therapy using multi-Task learning. Sci Rep. 2018;8:1–10.

58. Binette F, McQuaid DP, Haudenschild DR, Yaeger PC, McPherson JM, Tubo R. Expression of a stable articular cartilage phenotype without evidence of hypertrophy by adult human articular chondrocytes in vitro. J Orthop Res. 1998;16:207–16.

59. Kitano H. Biological robustness. Nat Rev Genet. 2004;5:826–37.

60. Kærn M, Elston TC, Blake WJ, Collins JJ. Stochasticity in gene expression: From theories to phenotypes. Nat Rev Genet. 2005;6:451–64.

61. Kitano H. The Theory of Biological Robustness and Its Implication in Cancer. Ernst Schering Res Found Workshop. 2007;61:69–88.

62. Kozhemyakina E, Cohen T, Yao T-P, Lassar AB. Parathyroid Hormone-Related Peptide Represses Chondrocyte Hypertrophy through a Protein Phosphatase 2A/Histone Deacetylase 4/MEF2 Pathway. Mol Cell Biol. 2009;29:5751–62.

63. Weng T, Yi L, Huang J, Luo F, Wen X, Du X, et al. Genetic inhibition of fibroblast growth factor receptor 1 in knee cartilage attenuates the degeneration of articular cartilage in adult mice. Arthritis Rheum. 2012;64:3982–92.

64. Thysen S, Luyten FP, Lories RJU. Targets, models and challenges in osteoarthritis research. DMM Dis Models Mech. 2015;8:17–30.

65. Luyten FP, Bierma-Zeinstra S, Dell'Accio F, Kraus VB, Nakata K, Sekiya I, et al. Toward classification criteria for early osteoarthritis of the knee. Semin Arthritis Rheum. 2018;47:457–63.

Lesage *et al. BMC Biology*   (2022) 20:253

Page 25 of 25

66. Mobasheri A, Van Spil WE, Budd E, Uzieliene I, Bernotiene E, Bay-Jensen AC, et al. Molecular taxonomy of osteoarthritis for patient stratification, disease management and drug development: Biochemical markers associated with emerging clinical phenotypes and molecular endotypes. Curr Opin Rheumatol. 2019;31:80–9.

67. Turnbull AK, Kitchen RR, Larionov AA, Renshaw L, Dixon JM, Sims AH. Direct integration of intensity-level data from Affymetrix and Illumina microarrays improves statistical power for robust reanalysis. BMC Med Genomics. 2012;5:35.

68. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8:118–27.

69. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring regulatory networks from expression data using tree-based methods. PLoS One. 2010;5:e12776.

70. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera R, et al. ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. BMC Bioinformatics. 2006;7(Suppl 1):S7.

71. Haury A, Mordelet F, Vera-licona P, Vert J. Open Access TIGRESS : Trustful Inference of Gene REgulation using Stability Selection. BMC Syst Biol. 2012. https://doi.org/10.1186/1752-0509-6-145.

72. Shannon P, Andrew M, Owen O, Nitin SB, Jonathan TW, Daniel R, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. 2003;13:2498–504.

73. Fauré A, Naldi A, Chaouiya C, Thieffry D. Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. Bioinformatics. 2006;22:124–31.

74. Alon U. An introduction to systems biology: Design principles of biological circuits. Chapman and Hall/CRC Press. 2006.

75. Chaouiya C, Naldi A, Thieffry D. Logical modelling of gene regulatory networks with GINsim. Methods Mol Biol. 2012;804:463–79.

76. Zañudo JGT, Aldana M, Martínez-Mekler G. Boolean threshold networks: Virtues and limitations for biological modeling. Intell Syst Ref Libr. 2011;11:113–51.

77. Lesage R. Experimental validation data for in silico OA study. Zenodo. 2022. https://doi.org/10.5281/ZENODO.7189743.

78. Caron MMJ, Emans PJ, Cremers A, Surtel DAM, Coolsen MME, van Rhijn LW, et al. Hypertrophic differentiation during chondrogenic differentiation of progenitor cells is stimulated by BMP-2 but suppressed by BMP-7. Osteoarthr Cartil. 2013;21:604–13.

79. Eduati F, Utharala R, Madhavan D, Neumann UP, Cramer T, Seaz-Rodriguez J, et al. Data and code for "A microfluidics platform for combinatorial drug screening on cancer biopsies."; 2018. https://doi.org/10.5281/ZENODO.1248886.

80. Lesage R, Geris L. PPI network- Interactive Modeling of Biological Networks | Cell Collective. Cell Collective. 2021. https://doi.org/https://research.cellcollective.org/#a5b66073-6769-4c88-bf6b-37ca1aa8f766.

81. Lesage R, Geris L. GRN - Interactive Modeling of Biological Networks | Cell Collective. Cell Collective. 2021. https://doi.org/https://research.cellcollective.org/#474de240-8752-4c3b-aa63-23640e50bf7a.

82. Chaouiya C, Bérenguier D, Keating SM, Naldi A, van Iersel MP, Rodriguez N, et al. SBML qualitative models: A model representation format and infrastructure to foster interactions between qualitative modelling formalisms and tools. BMC Syst Biol. 2013;7:1–15.

83. Burleigh A, Chanalaris A, Gardiner MD, Driscoll C, Boruc O, Saklatvala J, et al. Joint immobilization prevents murine osteoarthritis and reveals the highly mechanosensitive nature of protease expression in vivo. Arthritis Rheum. 2012;64:2278–88.

84. Poulet B, Ulici V, Stone TC, Pead M, Gburcik V, Constantinou E, et al. Time-series transcriptional profiling yields new perspectives on susceptibility to murine osteoarthritis. Arthritis Rheum. 2012;64:3256–66.

85. Bomer N, Cornelis FMF, Ramos YF, Den Hollander W, Storms L, Van Der Breggen R, et al. 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. 2016;75:571–7.

86. Lodewyckx L, Cailotto F, Thysen S, Luyten FP, Lories RJ. Tight regulation of wingless-type signaling in the articular cartilage - subchondral bone biomechanical unit: Transcriptomics in Frzb-knockout mice. Arthritis Res Ther. 2012;14:R26

87. Gardiner MD, Vincent TL, Driscoll C, Burleigh A, Bou-Gharios G, Saklatvala J, et al. Transcriptional analysis of micro-dissected articular cartilage in post-traumatic murine osteoarthritis. Osteoarthr Cartil. 2015;23:616–28.

88. Bateman JF, Rowley L, Belluoccio D, Chan B, Bell K, Fosang AJ, et al. Transcriptomics of wild-type mice and mice lacking adamts-5 activity identifies genes involved in osteoarthritis initiation and cartilage destruction. Arthritis Rheum. 2013;65:1547–60.

Publisher's Note