



# Reconstructing Boolean network ensembles from single-cell data for unraveling dynamics in the aging of human hematopoietic stem cells



Julian D. Schwab<sup>a,1</sup>, Nensi Ikononi<sup>a,1</sup>, Silke D. Werle<sup>a,1</sup>, Felix M. Weidner<sup>a,1</sup>, Hartmut Geiger<sup>b</sup>, Hans A. Kestler<sup>a,\*</sup>

<sup>a</sup>Institute of Medical Systems Biology, Ulm University, Albert-Einstein-Allee 11, Ulm 89081, Germany

<sup>b</sup>Institute of Molecular Medicine, Ulm University, Albert-Einstein-Allee 11, Ulm 89081, Germany

## ARTICLE INFO

### Article history:

Received 15 May 2021

Received in revised form 20 August 2021

Accepted 12 September 2021

Available online 15 September 2021

### Keywords:

Single-cell network reconstruction

Boolean network ensembles

Ensemble dynamic analyses

HSC

Aging

NF- $\kappa$ B

Single-cell RNA sequencing

## ABSTRACT

Regulatory dependencies in molecular networks are the basis of dynamic behaviors affecting the phenotypic landscape. With the advance of high throughput technologies, the detail of omics data has arrived at the single-cell level. Nevertheless, new strategies are required to reconstruct regulatory networks based on populations of single-cell data. Here, we present a new approach to generate populations of gene regulatory networks from single-cell RNA-sequencing (scRNA-seq) data. Our approach exploits the heterogeneity of single-cell populations to generate pseudo-timepoints. This allows for the first time to uncouple network reconstruction from a direct dependency on time series measurements. The generated time series are then fed to a combined reconstruction algorithm. The latter allows a fast and efficient reconstruction of ensembles of gene regulatory networks. Since this approach does not require knowledge on time-related trajectories, it allows us to model heterogeneous processes such as aging. Applying the approach to the aging-associated NF- $\kappa$ B signaling pathway-based scRNA-seq data of human hematopoietic stem cells (HSCs), we were able to reconstruct eight ensembles, and evaluate their dynamic behavior. Moreover, we propose a strategy to evaluate the resulting attractor patterns. Interaction graph-based features and dynamic investigations of our model ensembles provide a new perspective on the heterogeneity and mechanisms related to human HSCs aging.

© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Technologies that analyze regulatory dependencies in biological high throughput data support the notion that cellular decisions and aging are based on complex molecular networks [1]. Systems biology addresses the understanding of these complex interactions by proposing a holistic view of such complex regulatory processes. Central problems in single-cell-omics analyses are the identification of the underlying regulatory networks and modeling of their dynamic behavior over time.

For modeling dynamic behaviors, multiple dynamic models have been suggested, ranging from Boolean networks [2] to probabilistic rule-based models [3], to differential equation-based models [4]. By describing regulatory interactions through binary activity levels (0/FALSE/inactive, 1/TRUE/active), Boolean network models do not require kinetic parameters, which are often not

available in biological studies [2]. Therefore, Boolean network models are very versatile and frequently applied for inferring and modeling gene regulatory networks.

Depending on the available type of data, Boolean networks can be inferred either by intense literature search [5–8] or by reconstruction from time series data [9–11]. The literature-based inference has the advantage of combining different levels of information from expression, to protein interactions, to phenotypic traits [12]. Nevertheless, the reproducibility of curated models has been questioned due to the possibility of expert-biased knowledge [13]. On the other hand, the reconstruction of regulatory networks from expression data relies on algorithms able to reconstruct regulatory dependencies based on time series. These data-driven approaches have generated more and more interest with the quick evolution of omics data and sequencing technologies. In particular single-cell RNA sequencing has given the unique possibility to explore phenotypes at the individual cellular level. While the perspective of reconstructing regulatory networks from single-cell data is appealing, these approaches have been limited so far to the reconstruction of developmental or differentiation phenotypes

\* Corresponding author.

E-mail address: [hans.kestler@uni-ulm.de](mailto:hans.kestler@uni-ulm.de) (H.A. Kestler).

<sup>1</sup> equal contribution.

[14]. This is due to the dependency of current reconstruction algorithms on time-ordering [14]. Our new approach instead, by assuming the cell populations's intrinsic heterogeneity, allows for reconstruction of single-cell-derived networks without time-ordering constraints, which we applied to the study of HSC aging.

HSCs form all blood cells in a process termed hematopoiesis [15]. The isolation and *in vitro* maintenance of HSCs are still experimentally very challenging approaches [16–19]. Therefore, results from *in silico* modeling are of particular interest in this field. To this end, single-cell reconstruction approaches to study the differentiation of HSCs have recently been proposed [20–22].

Upon aging, HSCs are increased in number, but their activity is highly heterogeneous and impaired. This impaired function of aged HSCs might influence the immune system (AAIR) and is likely to be associated with leukemia [15]. Currently, the aging of human HSCs is primarily described in phenotypic terminology, with a number of molecular hallmarks of HSC aging being characterized [23,24]. These are metabolic activation, reactive oxygen species (ROS) production, impaired autophagy, loss of polarity, cellular senescence, telomere extension, and increased DNA damage [23]. All these hallmarks are known to be connected to inflammation. Here, the NF- $\kappa$ B pathway is known to play a crucial role, but the main players and mechanisms regulating this process are still not well characterized [23]. In addition, most of the information on HSC physiology and aging comes from pooled murine data, whereas human data are still not widely explored (Table A1). Murine and human HSCs share common regulatory mechanisms, still some differences do exist [25]. Hence, the determination of underlying mechanisms and regulatory networks of aging of HSCs is critical to design approaches to attenuate the aging-related impaired function of HSCs and identify markers distinguishing between chronological and biological aging [1,26].

In the present work, we focus on developing and analyzing a network reconstruction pipeline taking advantage of the emerging single-cell sequencing techniques. The latter was applied to investigate the NF- $\kappa$ B pathway involved in human HSC aging. The network reconstruction is based on data from the recently published work from Ratliff and colleagues [27], analyzing HSCs from eight healthy human beings of different ages separated into two groups – young and aged. Based on the heterogeneity present in single-cell data and especially HSCs, which are known to have varying levels of activation [28,29] (Fig. A1-), we were able to reconstruct specific regulatory networks for each individual. This was only possible by implementing a new concept to retrieve pseudotime series within a certain population of single-cell data. Next, we established a new hybrid pipeline to reconstruct ensembles of regulatory networks to boost the speed of reconstruction, which allowed us to reconstruct large networks. Our new reconstruction strategy, which we named “*filtered best-fit*”, combines the detection of a filtered selection of input candidates [30] which serves, then, as a preprocessing step for the exact algorithm from Lähdesmäki and colleagues (Best-Fit Extension algorithm) [31]. Interaction graph-based measures and dynamic analyses of the resulting population of networks (ensembles) representing different age states of human HSCs indicated significant differences in terms of structural connections, heterogeneity, and activity levels in the populations of networks of aged HSCs.

## 2. Materials and methods

### 2.1. Boolean network models

Boolean networks are dynamic mathematical models applied to describe biological regulatory processes. These models are defined as a set of  $n$  compounds  $X = \{x_1, x_2, \dots, x_n\}$ ,  $x_i \in \mathbb{B}$  whose regula-

tory dependencies are described in a set of Boolean transition functions  $\{f_1, \dots, f_n\}$ ,  $f_i: \mathbb{B}^n \rightarrow \mathbb{B}$ . In these functions, regulatory interactions are summarized by logical operators. The activity of each compound  $x_i$  is considered to be either active (1) or inactive (0). The state of a network at a specific point in time  $t$  is therefore determined by a vector  $\vec{x}(t) = (x_1(t), \dots, x_n(t))$  containing all assigned activities for each compound at that time point. The study of network dynamics over time can be performed based on different updating schemes – synchronous [2], asynchronous [32], and probabilistic [33] updating. Applying these updating schemes to Boolean functions creates state transitions and, thus, edges from vertex to vertex (each representing one particular state) in the state transition graph [12]. Here, we consider synchronous updating, which requires the least assumptions [12]. Under synchronous updates, the transitions from one state to the next one  $\vec{x}(t) \mapsto \vec{x}(t+1)$  are executed by updating all regulatory functions  $f_i$  at the same time. This transition is described as  $x_i(t+1) = f_i(\vec{x}(t))$ . Taking into account that each compound in the network has only two possible assigned values, the total number of states scales to  $2^n$ ,  $n$  being the total number of compounds in the network [2]. Due to this deterministic nature of the state space, under synchronous update schemes, the model will eventually enter a recurrent sequence of states called an attractor. Each network can have either single or multiple attractors, which in turn can have single (fixed points) or multiple states (cyclic attractors). Attractors depict the long-term behavior of the model, and they have been connected to biological phenotypes [34–38]. Among others, Boolean networks are frequently used to simulate biological experiments and, thus, screen, for example, for potential drug targets [39].

Boolean network models were simulated using the R-package BoolNet [40] and synchronous update strategy. For attractor search, the SAT-based search algorithm in the BoolNet package was used.

### 2.2. Single-cell RNA-sequencing data set

Model reconstruction was performed on the publicly available single-cell RNA sequencing data from Ratliff and colleagues [27] (NCBI Gene Expression Omnibus GSE138544). This data set contains 730 samples of isolated peripheral blood long-term HSCs (LT-HSCs) (*lin<sup>-</sup>CD34<sup>+</sup>CD38<sup>-</sup>CD45RA<sup>-</sup>CD49f<sup>+</sup>*) from four young (ages 19, 21, 37, 40) and four aged (ages 61, 66, 68, 70) human individuals (two males and two females per group) as suggested by Ratliff and colleagues [27]. The dataset contains 83 to 94 single-cell measurements per individual. Sequencing was performed on a NovaSeq6000. The available dataset shows log normalized counts per million [27]. The presented dataset was selected among other available datasets of single-cell RNA sequencing datasets (Table A1). It is, to the best of our knowledge, the only dataset with multiple non-pooled individuals currently publicly available.

### 2.3. Data preparation and expression analysis

To reconstruct Boolean network ensembles of the NF- $\kappa$ B pathway, the gene symbols from the dataset were mapped to Entrez IDs using the R-package biomaRt [41]. Based on the KEGG database [42], 101 genes were extracted as belonging to the human NF- $\kappa$ B signaling pathway (hsa04064). All selected genes were then binarized using the BASCA algorithm from the R-package BiTrinA [43,44]. We used the BASC significance test to evaluate which genes were significantly binarized (FDR,  $p < 0.05$ ). In accordance with these results, we used 96 genes for the reconstruction of Boolean network ensembles. Besides Boolean network reconstruction,

we performed expression analysis to screen for clustering of expression data according to the age of the individuals. Hierarchical clustering of the samples was performed using the ComplexHeatmap (Version 2.4.3, [45]). In addition to that, we plotted the data t-Distributed Stochastic Neighbor Embedding (tSNE) using the Seurat R-package (Version 4.0.1, [46,47]). Motif search was performed using the igraph R-package (Version 1.2.6, [48]).

#### 2.4. Generation of pseudo-time series

After binarization, time series for the 96 binarizable genes were generated to proceed with the network reconstruction. Therefore, the state of each single-cell measurement is assumed to be a potential predecessor or successor of the state of each other single-cell measurement coming from the same individual. Consequently, it is possible to form a large number of tuples of predecessor and successor time steps  $X(t)$  and  $X(t + 1)$  by a combination of random single-cell measurements. Given a total amount of single-cell measurements ( $s$ ), the number of couples of predecessor and successor states is  $\binom{s}{2}$ . This is multiplied by two as each state in a given tuple can act either as a predecessor state or successor state. Hence, the total number of possible tuples is  $\binom{s}{2} \cdot 2$ .

From the binarized expression data, 1000 tuples of data points were randomly drawn and used as time points (predecessor and successor state) to generate pseudo-time series of length two for reconstruction. This means, e.g. for individual young A (19 years old) having 94 measurements, we receive 4371 couples of single-cell measurements, resulting in 8742 possible tuples. Out of them, 1000 tuples are picked to reconstruct the Boolean functions. This procedure was repeated 20 times.

#### 2.5. Reconstruction of Boolean networks from binarized time series

Monotonicity is a dominant pattern in biological functions [49]. The algorithm by Maucher et al. [30] infers regulatory dependencies from binary time series of data based on this monotonicity assumption. To infer these dependencies, the algorithm measures Pearson correlation between the different genes and the successive network states [30]. The algorithm identifies the input variables with a chosen minimum influence for each regulatory factor individually by measuring the Pearson correlation of each variable and the corresponding output value of the regulatory factor. If the correlation is above a specified threshold, the interaction is considered to be relevant. This is repeatedly done for each regulatory factor to infer the complete network dependencies [30].

While the algorithm by Maucher et al. [30] returns regulatory dependencies, the Best-Fit Extension algorithm by Lähdesmäki et al. [31] allows for the reconstruction of Boolean networks from time series of expression data. The *best-fit* approach tests all possible combinations of inputs for each Boolean function to fit the dynamics of the given time series as correctly as possible. In more detail, the algorithm screens for the subset  $X'$  of all genes  $X$ , with up to  $k$  inputs. To find Boolean functions which match the measured observations in the time series data corresponds to the consistency problem [50]. Consequently, a Boolean function that best separates true and false samples in the data is sought. In the Best-Fit Extension from Lähdesmäki and colleagues [31], this is done based on *partially defined Boolean functions*  $pdBFs(T,F)$ , with  $T, F \in \{0, 1\}^k$ . These  $pdBFs$  describe the true and false observations in the given time series of binary data. Each tuple of predecessor  $X(t)$  and successor time point  $X(t + 1)$  is added to the  $pdBF(T,F)$  as follows [31]:

$$T = \{X'(t) \in \{0, 1\}^k : X_i(t + 1) = 1\} \text{ and}$$

$$F = \{X'(t) \in \{0, 1\}^k : X_i(t + 1) = 0\}.$$

Next, the number of inconsistencies in the  $pdBF$  is measured by intersecting  $T$  and  $F$  ( $\varepsilon = |T \cap F|$ ). The algorithm then chooses the input combinations  $X'$  with the least error  $\varepsilon$ . In the final step, a Boolean function based on these inputs is created using truth tables. The truth table is filled by iterating through all examples  $s = T \cup F$  at all time steps  $j$  as follows:

$$f_i^j = \begin{cases} 0 & \text{if } s \in F \wedge f_i^{j-1} = ? \\ 1 & \text{if } s \in T \wedge f_i^{j-1} = ? \\ * & \text{else} \end{cases}$$

Here ? means undefined and \* indicates a conflict.  $f^0$  is initialized as  $f^0 = (?, \dots, ?)$ .

#### 2.6. Inference of ensemble networks

Based on the binarized pseudo-time series and the *filtered best-fit* approach, Boolean functions were reconstructed for each individual independently. For each set of single-cell data obtained from one individual, 1000 tuples of data points were randomly drawn and used as time points (predecessor and successor state) to generate pseudo-time series for reconstruction. Next, Boolean networks were reconstructed from these pseudo-time series. For each of the eight individuals, eight populations of networks (ensembles), comprising all reconstructed Boolean functions for each gene, were generated. For dynamic analysis, we sampled 100 Boolean networks from each ensemble by randomly choosing one of the potential Boolean functions suggested by the reconstruction algorithm. This procedure was repeated for 20 random picks of the 1000 tuples of pseudo-time points per individual, resulting in a total of 2000 networks per individual.

#### 2.7. Comparison of the reconstruction new pipeline with the best-fit approach

To evaluate the proposed reconstruction pipeline, we compared its computation time and performance to the original Best-Fit Extension approach. We created random networks of different sizes (20, 40, 60 to 200 in steps of 20) with scale-free topology as ground truth networks using the BoolNet R-package. Next, we created time series of different lengths from each of the random networks. For the first measurements, the number of time steps was fixed to 20 (for each of 100 networks per size 20 to 200). For the second measurements, we created a time series of  $|V| + 10$  time points (where  $|V|$  is the number of nodes in the network, see Berestovky and colleagues [9]). Based on these time series, we reconstructed Boolean networks and measured the reconstruction runtime. To assess the reconstruction quality, we compared the interactions found in the reconstructed networks to the corresponding original network and measured the sensitivity and specificity of this prediction.

#### 2.8. Interaction graph-based analyses

To assess structural changes during aging, we compared the reconstructed networks of the young and aged groups as well as the structural changes of each individual on its own. Therefore, we calculated different properties based on the suggested interactions between the genes of the reconstructed Boolean networks by measuring:

- 1). The number (#) of compounds, which are unregulated and, thus, set to a constant value (0/1) in the Boolean networks (**fixed genes**),
- 2). The number (#) of compounds that have connectivity of 0 and are, consequently, disconnected from the rest of the graph (**isolated genes**),
- 3). The mean number (#) of incoming edges across the ensemble of networks (**mean input**), and
- 4). The mean number (#) of potential regulatory functions which could be found for each compound in the reconstructed process (**mean functions**).

In addition, network motifs were also investigated across all individuals and by age groups (young vs aged). To this end, feed-forward loops and bi-fan motifs (as defined in Alon [51]) were analyzed. These motifs were selected based on their biological impact in regulating the stability of the protein-protein interaction network [52,53].

We measured the differences between the two age groups using the Wilcoxon signed-rank test. We further validated the inferred regulatory interactions across age groups as well as for each individual by comparing the existence of the found interactions in the ensembles of Boolean networks also in the STRING-DB [54]. To do so, two matrices were combined: the first considering consistency of regulatory interaction within the ensembles, the second considering the existence of a given interaction in the STRING-DB database [54]. Below, the establishment of both matrices and their combination in a trinary match matrix is described.

The ensemble adjacency matrix has an entry of 1 at the position (i,j) if the compound *i* is present in the regulatory function of the compound *j* and 0 otherwise. In the case of multiple equally accurate regulatory functions for the same compound, the union across all present interactions is taken. Furthermore, a union of interactions found across all 20 different reconstructions is taken. For analyzing each age group individually, all interactions occurring across the four corresponding individuals are taken into account.

Similarly, a binary matrix has been constructed from the data available in STRING-DB [54], with entries having a value of 1 if there exists an interaction and 0 otherwise. A match was assigned considering paths that are either direct or indirect via only one additional node. This analysis has been performed both considering all possible sources of evidence (including, e.g., text mining and co-expression of genes), as well as for a restriction to interactions that were either present in curated databases or that were experimentally validated.

These two binary matrices for the ensembles and the STRING-DB were then compared to obtain a matching matrix. Here, 1 indicates that there is a match between interactions present both in the Boolean network as well as in STRING-DB. If there is no interaction in the Boolean network, the match matrix has a 0 entry. An entry of -1 indicates a mismatch, meaning that the interaction in the Boolean network was not found in STRING-DB or only exists via more than one indirect node.

### 2.9. Dynamic analyses

For further analysis of the reconstructed ensembles of Boolean networks, we investigated its dynamic behavior by screening its attractor landscape. To do so, we performed an exhaustive attractor search on each of the networks in the network population (ensemble). All simulations were performed using the synchronous update strategy as previously applied in other model simulations describing complex pathway interactions [5,55–57].

First, we studied the mean number of attractors which could be found across all networks in the different ensembles and the repeated runs, and their number of states.

Second, we investigated the distribution of gene activity through the different attractor patterns. Here, the availability of ensembles of networks allows analyzing probabilities of attractor patterns. We summarized the binary states of the different attractors of each network and normalized them by the total number of attractor states. This procedure yields the probabilities of each gene to be active within the complete attractor landscape. Next, we summed up these probabilities for all network simulations within one ensemble and the repeated runs and normalized again by ensemble size and the number of repeated runs. Finally, results show an average probability of genes to be active in the long-term behavior of the networks in each ensemble over the repeated reconstruction runs.

## 3. Results and discussion

### 3.1. Fitting the heterogeneity of the aging process with a tailored systems biology approach

The aging process of HSCs is suggested to be characterized by a dynamic and heterogeneous behavior. While common effects of aging-related dysfunctions have been identified, not all HSCs in the elderly are thought to present this loss of function, which results in a likely heterogeneity of individual aged HSCs in single-cell expression data. Likely reflecting this heterogeneity, hierarchical clustering and t-SNE procedures on single HSCs from our selected dataset were not successful. In particular, analyzing the expression data of the genes in the known aging-related pathway as NF- $\kappa$ B [24,58] did not show relevant clustering between young and aged cells (Figs. A1–A4). New tools and approaches are warranted to depict this heterogeneity in the aging of HSCs.

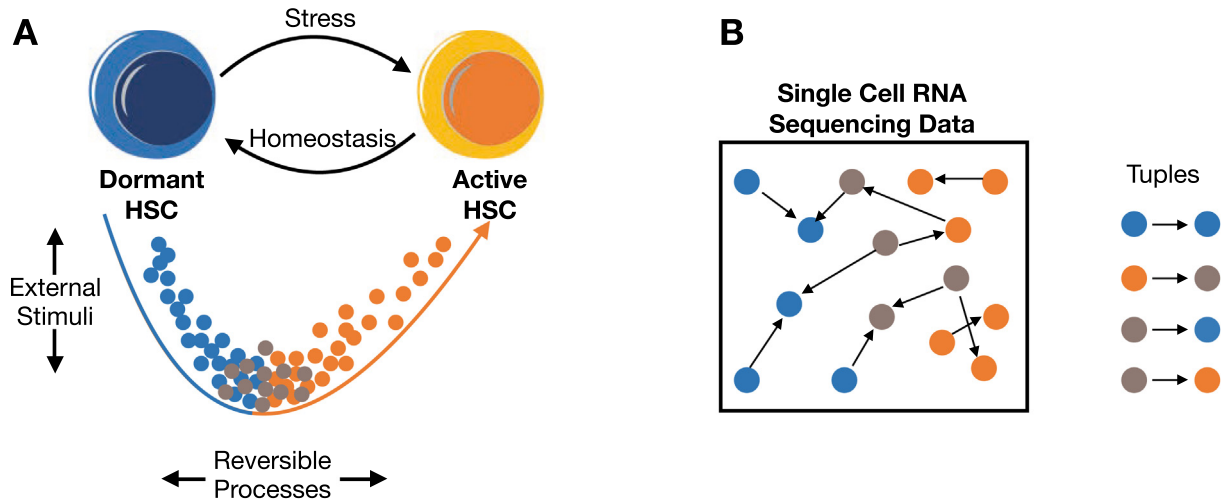
For this reason, we reconstructed populations of Boolean models from single-cell RNA sequencing data. By studying the behavior and characteristics of the ensemble of individual networks, we provide novel tools to capture the heterogeneity of the aging process of HSCs. In the following, we will present in detail the rationale of our approach, together with the major insights from the ensembles of Boolean networks.

### 3.2. Filtered reconstruction of Boolean network ensembles from time series of single-cell data

All data-based reconstruction algorithms largely depend on the number of data and especially the time points which are available for reconstruction. Typically, there is only a small number of time points, and the reconstruction potential is limited. In this approach, we take advantage of the emerging single-cell high-throughput data and propose a new strategy to increase the amount of available time series within a population of cells.

In single-cell sequencing data, though, each sample is described by a potentially large set of single-cell measurements. We group all measurements which belong to the same individual. Within these groups, we assume the state of each single-cell measurement is a potential predecessor or successor of the state of each other single-cell measurement in the same domain – each single-cell sample is treated as a pseudo-time-point and heterogeneity converted into pseudo-time-points. This assumption is based on experimental results showing that LT-HSCs can depict interchangeable activity states, implying that a population of LT-HSCs dynamically transits from one activity level to the other [28,29] (Fig. 1A). Consequently, under this assumption, we can form a large number of tuples of predecessor and successor time steps  $X(t)$  and  $X(t + 1)$  by a combination of random single-cell measurements (see Fig. 1B and methods section 2.5).





**Fig. 1.** Heterogeneity of activity levels in the LT-HSC population. (A) Biological concept. Within the LT-HSC population, stem cells can reversely transit between more dormant or activated states without changing their stemness condition (no differentiation) [28,29]. The heterogeneity of this population is influenced by environmental stimulation. (B) Generation of tuples from single-cell RNA sequencing data. Exploiting the heterogeneity of the LT-HSC population, pseudo-time-points assuming the possibility for each cell of the pool to switch to any of the other cell expression profiles. This implies that trajectories can be randomly selected within the population. Tuples were built by sets of randomly sampled trajectories from the expression data of each individual.



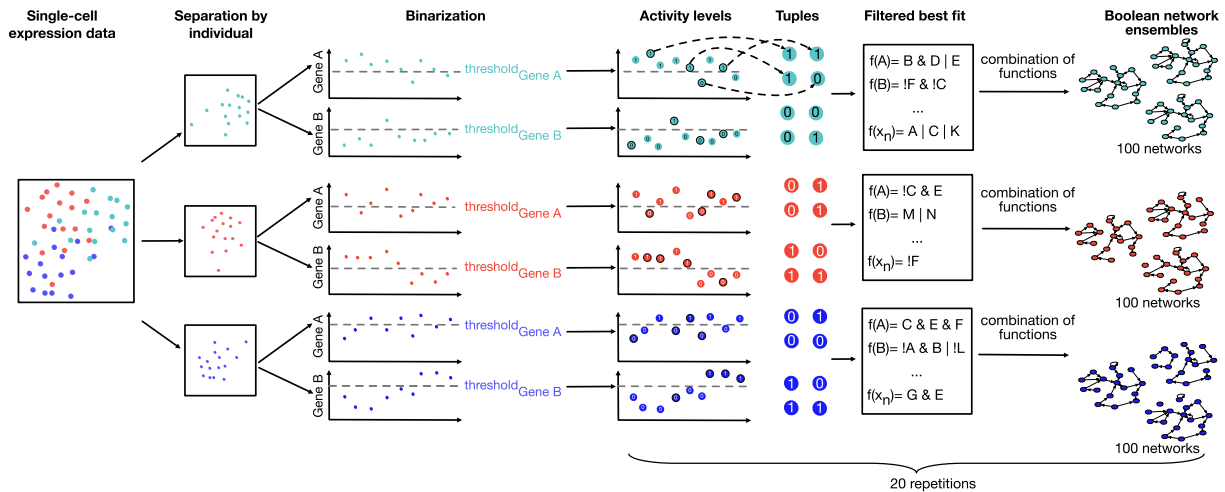
**Fig. 2.** Overview of the new Boolean network reconstruction pipeline. The Boolean networks reconstruction pipeline combines a correlation-based approach to predict regulatory inputs by Maucher et al. [30] and the Best-Fit Extension algorithm by Lähdesmäki et al. [31]. The algorithm by Maucher et al. is used as preprocessing step to reduce the number of inputs that will be tested for the succeeding est-ixtension algorithm. This combination decreases the reconstruction time without losing meaningful results.

Introducing this new concept for retrieving pseudo-time points from single-cell experiments leads to an increased number of time points compared to bulk sequencing experiments. This increase of information challenges the currently available reconstruction algorithms, impairing the efficient reconstruction of models.

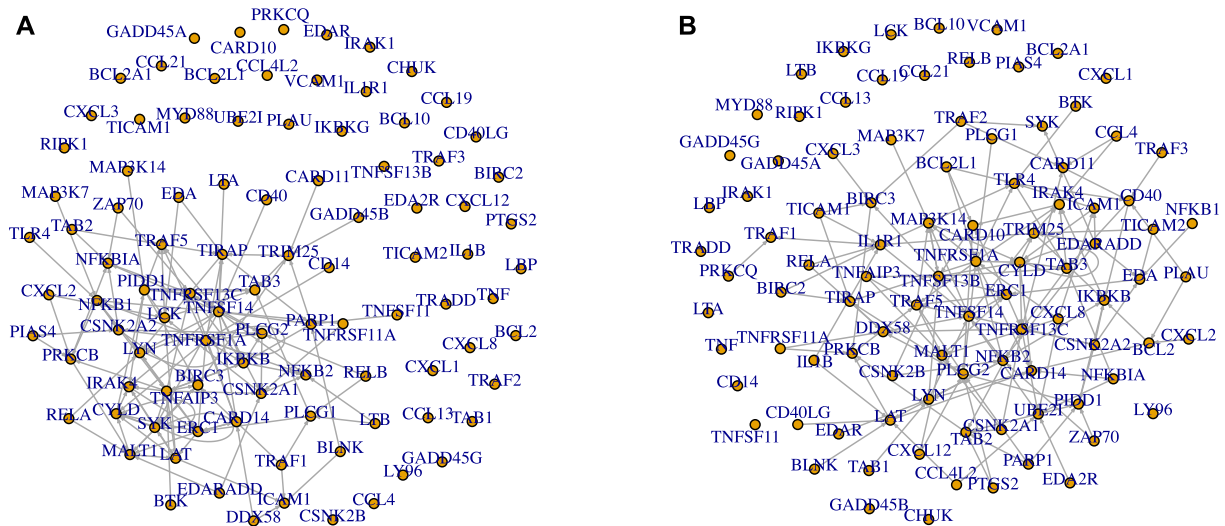
Boolean network reconstruction approaches need to be adapted to handle this data. We propose a new pipeline here that combines the search for potential regulatory inputs by Maucher et al. [30] and the Best-Fit-Extension algorithm [31] (Fig. 2). We use the approach by Maucher et al. [30] as a pre-processing step for the Best-Fit Extension by Lähdesmäki et al. [31]. In this reconstruction pipeline approach, the most feasible regulatory inputs  $X_{fi} \subseteq X$  for each regulatory input  $x_i$  are determined using this approach [30]. In the next step, the Best-Fit Extension algorithm is used to reconstruct the corresponding Boolean functions with an input combination  $X'_i \subseteq X_{fi}$  among the possible inputs  $X_{fi}$  as derived from the pre-processing step (see Fig. 3). For the following reconstruction, we set the threshold for the preprocessing to 0.03 and the

maximum number of inputs to five. Hence, instead of testing all possible input combinations in the reconstruction of Boolean networks, we only considered the ones selected by our preprocessing step. Based on the reconstruction results, we then create a population of Boolean networks (ensemble) by randomly sampling one regulatory function from the reconstructed set of potential functions for each compound of the modeled system.

In a first analysis, we compared the new pipeline to the original best-fit approach. To do so, we used randomly generated Boolean networks of different sizes, generated time series of data, and used those for the reconstruction of Boolean network ensembles. Reconstruction results were then compared to the original networks. Results showed, first, reducing the number of input combinations to test using the est-ixtension and, thus, a speed up the algorithm itself especially when considering noisy data (see Appendix Figs. A5 and A8). With an increasing number of time steps this effect becomes more prominent. Second, results also show an increased sensitivity when recapitulating the original networks



**Fig. 3.** Schematic overview of the network reconstruction approach. In our approach, we first selected all single-cell samples from each group (red, blue, turquoise – here cell samples taken from the same individual). Next, data points are randomly sampled and ordered sequentially to create a series of pseudo-time-points for each group. This time series was then used for the reconstruction of Boolean network ensembles. The approach results in a set of Boolean network ensembles – one for each selected group. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 4.** Interaction graph of one example network drawn randomly from the ensemble of one random individual from the young group of networks (A) and the aged group of networks (B).

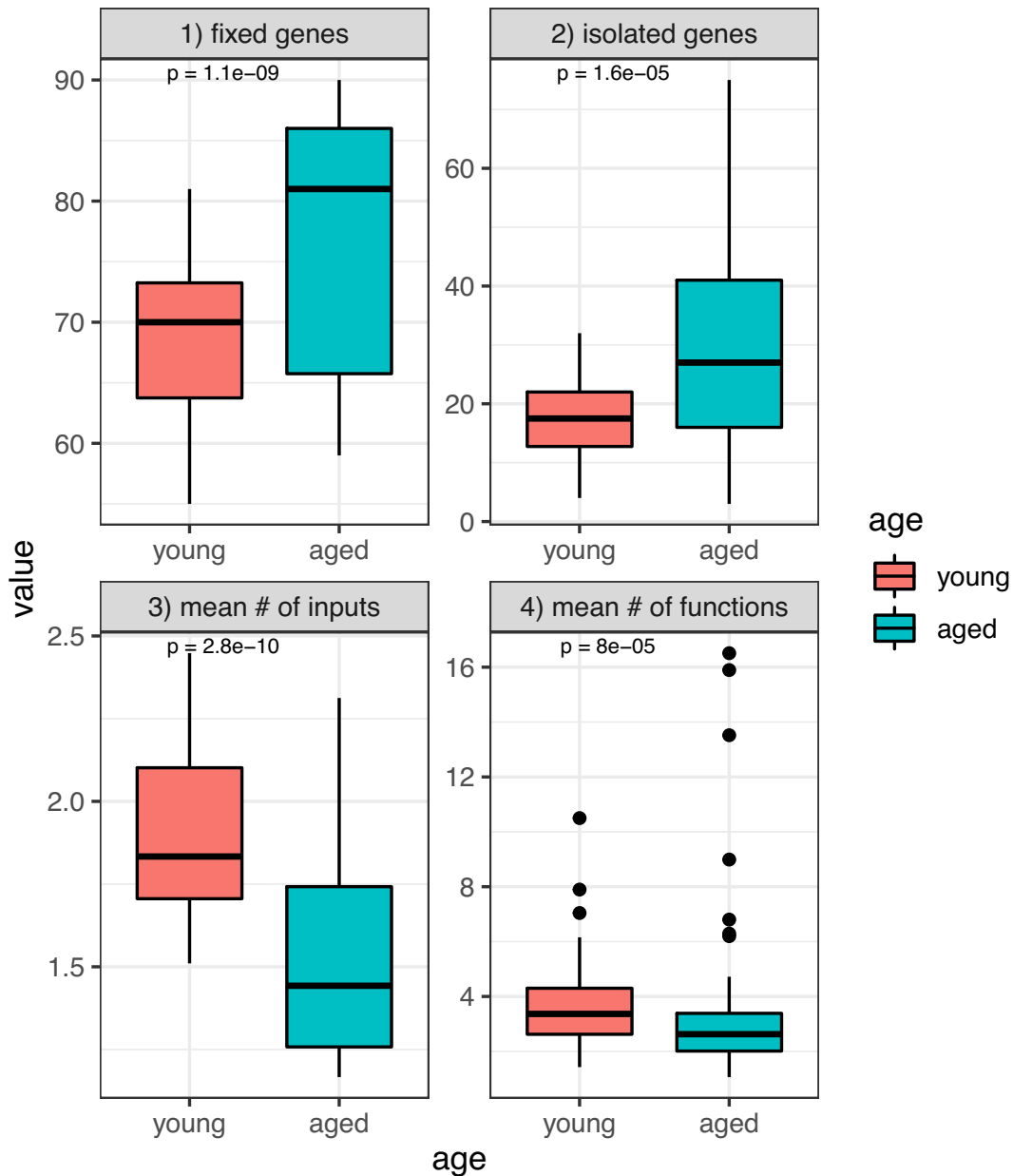
based on the noisy data (see Appendix Figs. A6 and A9) while specificity of both approaches is in similar ranges (see Appendix A7 and A10).

### 3.3. Network populations (ensembles) reconstruction and structural properties in the process of aging

Using the described approach, we reconstructed ensembles of Boolean networks modeling the NF- $\kappa$ B signaling pathway for each of the eight individuals. This revealed eight ensembles of Boolean networks – one representing each individual (exemplarily shown in Fig. 4). The reconstructed networks comprise up to 950 potential Boolean functions for one regulatory compound, depicted as a node in the network. The mean number of reconstructed Boolean functions per compound was between 1.06 and 16.50 across the different individuals and the 20 repeated runs. Reconstructed Boolean functions had between 1 and 5 regulatory inputs (Fig. 5 and Fig. A11). By comparing the reconstructed functions among young (19–40 years old) and aged (61–70 years old) individuals, we could

show an increase in mean functions ( $p = 8 \cdot 10^{-5}$ ) and regulatory inputs ( $p = 2.8 \cdot 10^{-10}$  in the young phenotype compared to the aged individuals, while fixed genes ( $p = 1.1 \cdot 10^{-9}$ ) and isolated ones ( $p = 1.6 \cdot 10^{-5}$ ) were lower in the young reconstructed networks (for a precise definition of investigated groups of nodes see methods, section 2.6). These findings hint that in aged HSCs, regulatory interactions are muted, which would be consistent with a reduced response to stimulation as described by Schwab et al. [11]. We observed that the reconstructed networks of one individual (66 years old) had similar interaction graph-based features as the networks from the young group (Appendix Figs. A11 and A14), especially compared to the youngest individuals. This might indicate a younger biological age compared to the chronological one since this individual may rely on a network wiring with similar properties to the younger ones.

For further validation of the reconstructed Boolean functions, we compared the reconstructed interactions to interactions present in the STRING-DB database [54] (see Appendix Fig. A12). Considering all interactions within this database, we were able to



**Fig. 5.** Structural properties of young and aged HSCs. We compared the static properties across the Boolean network ensembles of HSCs from all young and all aged individuals, respectively. Each panel describes one of the topological properties of the interaction graph of the underlying networks. 1) the number of fixed genes, 2) the number of isolated genes, 3) the mean number of regulatory inputs per gene per network in the ensemble, and 4) the mean number of reconstructed functions per gene per network in the ensemble. We measured the differences between the two age groups using Wilcoxon signed-rank test. Results for all measured properties show a significant difference ( $p < 0.05$ ) between the two groups.

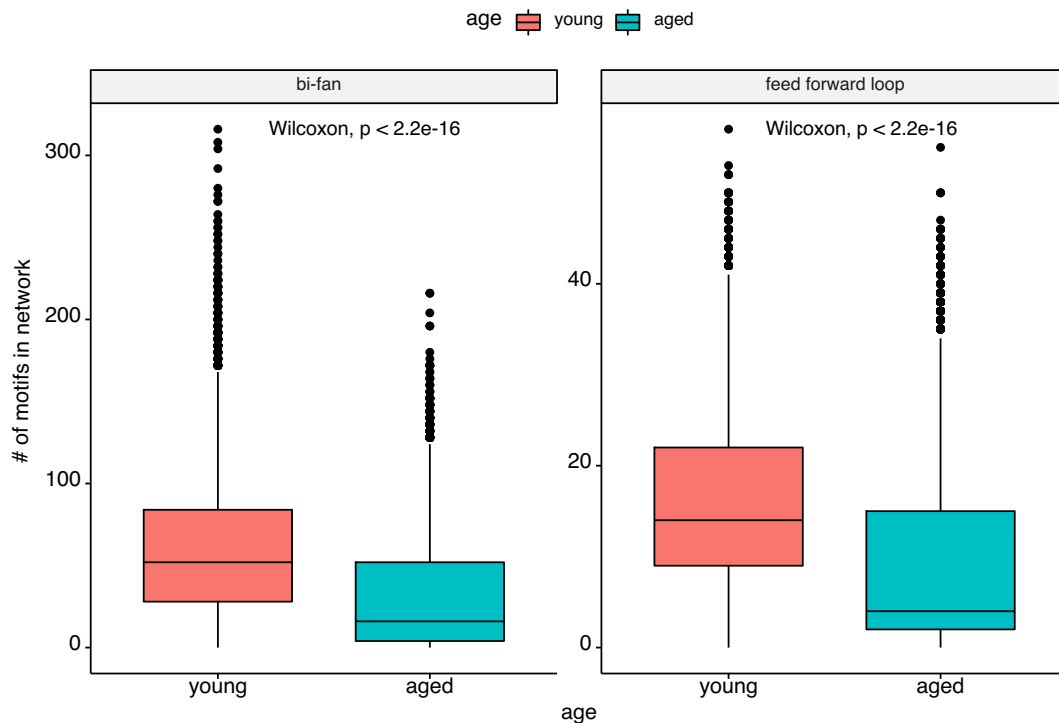
validate 94.1% of reconstructed regulations for the young phenotype and 94.0% for the aged phenotype across all individuals in the respective group. Even applying stricter measures and only considering experiments and databases (see Methods, Section 2.8), we still found 66.1% of reconstructed regulations via direct interactions or interactions with one intermediary step for the young phenotype and 65.1% for the aged phenotype. These results further validate the reconstructed interactions. Some of the reconstructed regulations were not listed as direct interactions in STRING-DB. Our reconstruction approach might thus point to so far unknown interactions, although these will require confirmation by additional experiments.

We performed screening of feed-forward loop and bi-fan motifs for each of the reconstructed networks. The number of each of

these motifs for each network was computed and then used for comparison between the two age groups (see Fig. 6) and the eight individuals (see Fig. A14). Results show that the number of both motifs is significantly increased (Wilcoxon signed-rank test,  $p < 0.05$ ) in the young group compared to the aged one.

### 3.4. Altered dynamics affecting aging

The human HSC aging process is still a widely unexplored field. In particular, still, a wide range of open questions are remaining on heterogeneity and its effects on biological aging in contrast to chronological aging. While being aware of the limited sample size of eight individuals, the results of the study might be a first attempt that justifies further exploration in this area in larger stud-



**Fig. 6.** Boxplots showing the number of bi-fan and feed-forward loop motifs measured in 2000 randomly sampled networks from each individual's ensemble (100 networks with 20 repetitions). Boxplots group the individuals by corresponding age group. We measured the differences between the two age groups using Wilcoxon signed-rank test and Bonferroni correction for multiple testing.

ies. Hence, we analyzed the dynamics of the group of Boolean networks by computing the attractor landscape. The dynamic analysis aims to get an initial understanding of general dynamics during aging in the NF- $\kappa$ B pathway. Dynamic analyses have the potential to identify new drivers of aging in humans. In the following, we will deepen our analyses on the retrieved attractor landscape focusing on gene activities attempting to match the model behaviors to literature-reported aging activities.

First, we studied the networks' general long-term behavior. To do so, we evaluated both the number of attractors retrieved and their lengths (see Fig. A13). While a general tendency towards having few attractors with only one state is observed among young and aged networks, we observe an increase in the number of attractors ( $p = 1.10^{-15}$ ) and length ( $p = 1.5 \cdot 10^{-32}$ ) for the young ones. It is likely favorable for a regulatory network to have a majority of only single state attractors since they are considered to be more stable and less likely disturbed [12,59]. This observation is consistent with a deep and stable quiescence of HSCs. In addition, Ikonomi et al. also retrieved a single state attractor describing the same HSC population in young physiological conditions [5]. The increase in the number of attractors in the network reconstructed for young HSCs might be linked to their ability to promptly respond to stimulation and which is reduced in aged HSCs [60]. Together with our results from the structural analysis (Fig. 5), our reconstructed networks in young HSCs show a larger number of connections and, therefore, hint at potential responses to sudden changes of external stimulations. Seeing these changes already in the LT-HSC population (for the NF- $\kappa$ B pathway), which are considered to be majorly quiescent, also sustains the hypothesis that major dysregulations in hematopoiesis during aging arise from differentially prompted and reactive HSCs [23,24]. Following this idea of a different intrinsic rewiring of the quiescent LT-HSCs, which prompt differential responses without *per se* causing a direct activation of the HSC, we deepened our analysis on the retrieved

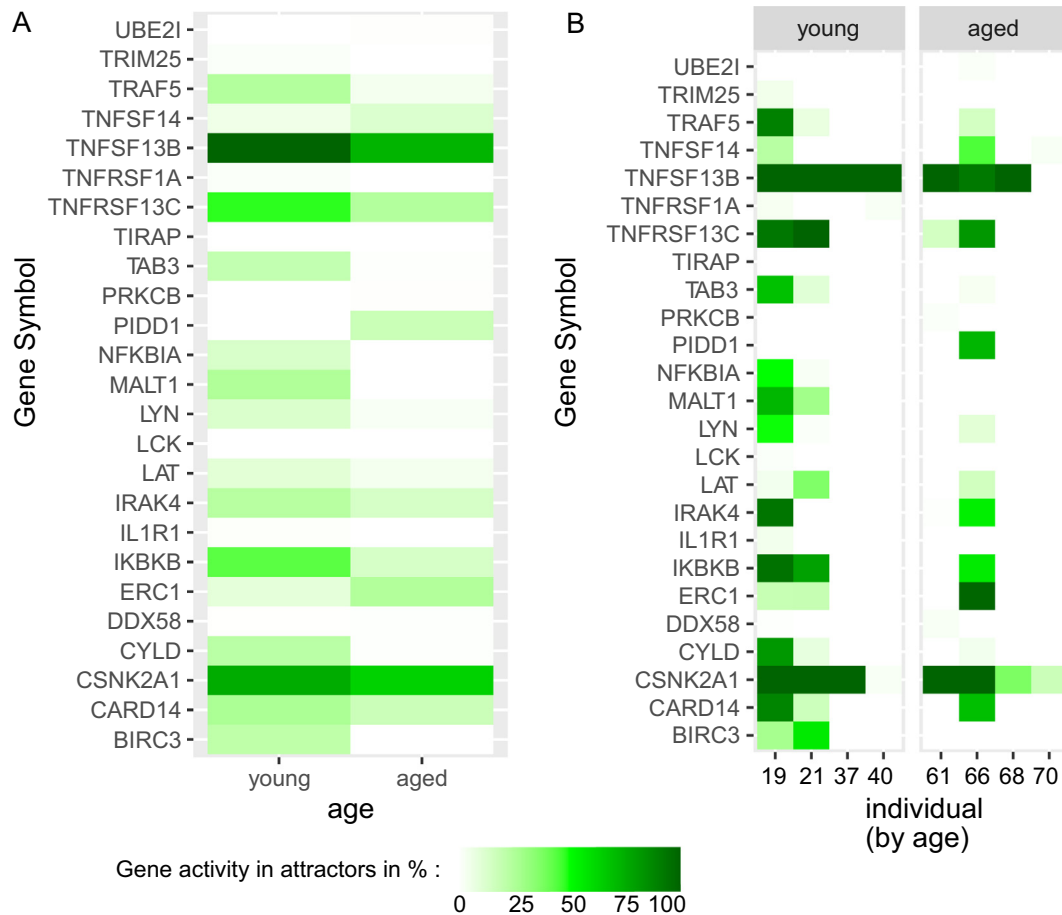
attractor landscape. To do so, we first analyzed the general activity tendencies between young and aged individuals and the inter-individual heterogeneity (Fig. 7). Second, we matched the activity of genes within the attractors to published phenotypes of young and aged HSCs (Table 1). The dynamic analysis intends to give initial insights into the general dynamics during aging in the NF- $\kappa$ B pathway and second to identify potential new drivers of aging in humans. Together with these two points, literature validation of both changing and fixed gene activities in the attractor patterns aims at a first evaluation of the attractor patterns. To do so, we matched activities in the attractor to expected changes during HSCs aging or features expected from LT-HSCs. In this sense, our literature search revealed that over the 96 genes in our attractor patterns 78.73% could be matched with expected behaviors, while 12.5% of genes have no described function in HSCs (Tables A2 and A3). Finally, only 6.25% of genes did not match behaviors expected for HSCs from literature. As a general observation, most inactive genes were either related to leukemic transformation or to blood cell differentiation (see Table A3), which are indeed not likely to influence the activity of healthy HSCs.

Next, in this use-case, we investigated the change of activities by comparing gene activity of the different age groups. This analysis aims to point out possible age-related activities which need to be validated also on larger data sets once they become available.

We observed the tendency of a decrease in activity levels within the attractors ranging from young to aged phenotypes (Fig. 7). This tendency was observed in both the overall averaged young and aged groups and for each individual. Interestingly, while this decrease in activity levels is preserved during aging, still a certain heterogeneity among individuals is observed (Fig. 7).

Along with the attractor patterns, we divided the gene activity into three subgroups: 1) decreased, 2) remaining stable, and 3) increased in aging (Table 1). IRAK4 and TNFRSF13C, both down-regulated in our attractors, were shown to be decreased in aged





**Fig. 7.** Changing activity levels in the process of aging. Figures showing the average activity of each gene in the attractor landscape of the simulated Boolean network ensembles. (A) The average activity of all non-zero genes over all ensembles of the young phenotype group (left column) and the aged phenotype group (right column) is depicted respectively. (B) The gene activity in the attractors of the ensembles averaged for each individual is shown. The individuals are sorted by increasing age (see x-axis). The color of the heatmap indicates the activity of a gene, ranging from 0% (white) to 100% (dark green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**  
Activity of genes in the aging process.

Activity levels during aging young → aged	Compounds with altering in the Boolean network ensembles	Biological impact of altered activity	References
	BIRC3, IRAK4 CARD14, IKBKB, CYLD TRAF5, LAT MALT1	Loss of rescue mechanisms for apoptosis causing impaired survival Loss of quiescence Myeloid skewing	[68,69] [70-76] [77-80]
	NFKBIA LYN ERC1	Compensatory inhibition of NF-κB activation via induction of quiescence. Loss of NF-κB direct inhibition Propensity to increased inflammatory response via NF-κB activation Impaired repopulation potential	[61,62] [81] [82,83] [84-88]
	TNFRSF13C TNFSF13B CSNK2A1	Cellular polarity alteration causing reduced motility. Presence of ERC1 promotes the turnover of focal adhesions Impaired lymphoid specification Stable immune response Resistance to senescence	[89,90] [63,64] [65]
	PIDD1 TNFSF14	Control of balance between repair and apoptosis after DNA-damage Loss of quiescence due to increased cycling	[66,91] [92]

HSCs encouraging again the correctness of the ensembles retrieved dynamics. Most of the remaining genes in the first subgroup are studied in the context of loss of function of HSCs, but dysregulation of their activity has not yet been linked to HSC aging. A non-trivial example is MALT1 that was also observed to be decreased with

aging in our attractors. This reduced activity would lead to a double-edged effect. On the one side, loss of MALT1 promotes HSC quiescence [61]. On the other side, it skews hematopoiesis towards myeloid cells [61]. As MALT1 is an inhibitor of NF-κB, its loss induces the increase of the latter. Nevertheless, MALT1 loss

also has a positive effect on quiescence [61]. This might counteract the increased stress-response induced by NF- $\kappa$ B that normally leads to cell cycle entry [62].

Furthermore, few genes showed either a stable active behavior or even an increase in activity throughout the aging process (Table 1). The function of stably active genes could be linked to a stable immune response [86,87] or resistance to senescence [65].

Finally, our results show that PIDD1 is only active in one aged individual (Aged A, 66 y.o.). Interestingly, the reconstructed networks of this individual already showed more similarity to interaction graph-based features and network motifs to the youngest individuals (19 and 21 y.o.) compared to the remaining individuals of the aged group (61, 68 and 70 y.o.). A potential explanation for that might relate to a discrepancy between biological and chronological age for this individual. Notably, literature research revealed that PIDD1 activity is connected to DNA damage response [66].

Following this hypothesis, we further deepened the literature screening looking for an explanatory mechanism on the potential implication of PIDD1 in delaying age-related processes. PIDD1 is suggested to act as a switch between DNA damage response and apoptosis by regulating the activation of first damage response via NF- $\kappa$ B, and in a second moment apoptosis via Caspase-2 [66]. While PIDD1 activity has still not been investigated in HSCs, the HSC phenotype of Caspase-2-deficient cells is known [67]. Caspase-2 deficient mice develop normally but show aging-related dysfunctions. In addition, when challenged by oncogenic stimuli or stress, Caspase-2 deficient mice show enhanced tumor development [67]. Overall, this supports the idea that a reactivation of PIDD1 could contribute to a delayed aging process, leaving these initial results open for further experimental validation.

Altogether our dynamic analyses indicate a certain level of heterogeneity in the behavior of our ensembles of models. This suggests that the description of aging does not follow the idea of “one fits all” but might come with different mechanisms and activities. This idea is reflected in the distinction between biological and chronological aging [1]. Interestingly, the Boolean network analyses of this dataset showed diverging properties among the different individuals. Even though these results come from a pool of eight human individuals, we found patterns in our network analyses which match general mechanisms and hallmarks known in LT-HSC maintenance and aging (Appendix Tables A2 and A3, and Table 1). Considering the novelty of the approach, the results point to interesting findings which encourage further studies and experimental validations in this direction.

#### 4. Conclusion

Here, we present a new method to reconstruct data-driven ensembles of regulatory networks from single-cell RNA-sequencing data potentially applicable to a wide variety of sequencing data. Taking advantage of the intra-cell population heterogeneity, the approach exploits the generation of pseudo-time points of populations of the same cell type. These pseudo-time points enable the reconstruction of Boolean network ensembles. To handle such a potentially large (pseudo-) time series, we furthermore developed a new Boolean network reconstruction pipeline by adding a correlation-based screening of potential regulatory dependencies as a pre-processing step to the extension algorithm. The pre-processing step not only decreases computational time but also is more robust to noisy data.

Our use-case, the analysis of aging-related changes in NF- $\kappa$ B signaling based on human HSCs single-cell data could successfully capture the heterogeneity of the aging process. Using the network reconstruction approach enabled us to analyze structural and dynamic properties. By ordering the individuals along changes

within these properties, we could find similar tendencies to their order by age which are not evident in the expression data analyses.

#### CRedit authorship contribution statement

**Julian D. Schwab:** Conceptualization, Software, Methodology, Formal analysis, Data curation, Visualization, Writing - original draft, Writing - review & editing. **Nensi Ikonomi:** Conceptualization, Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Silke D. Werle:** Conceptualization, Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Felix M. Weidner:** Formal analysis, Software, Visualization, Writing - original draft, Writing - review & editing. **Harmut Geiger:** Formal analysis, Writing - review & editing. **Hans A. Kestler:** Project administration, Supervision, Funding acquisition, Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

HAK acknowledges funding from the German Science Foundation [Nos. DFG, 217328187 (SFB 1074) and 288342734 (GRK HEIST)] as well as funding from the German Federal Ministry of Education and Research (BMBF) TRANSCAN VI – PMTR-pNET (id 01KT1901B).

#### Data availability

The presented pipeline with the corresponding code and experimental results is available at GitHub: <https://github.com/sysbio-bioinf/FilteredReconstructionFromSCData>.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.09.012>.

#### References

- [1] Dean W, Morgan RF. In defense of the concept of biological aging measurement—current status. *Arch Gerontol Geriatr.* 1988;7(3):191–210.
- [2] Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol.* 1969;22(3):437–67.
- [3] Groß A, Kracher B, Kraus JM, Kühlwein SD, Pfister AS, Wiese S, et al. Representing dynamic biological networks with multi-scale probabilistic models. *Commun Biol.* 2019;2(1). <https://doi.org/10.1038/s42003-018-0268-3>.
- [4] Wang R-S. Ordinary differential equation (ODE), model encyclopedia of systems biology. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H New York, NY: Springer New York; 2013. p. 1606–1608.
- [5] Ikonomi N, Kühlwein SD, Schwab JD, Kestler HA. Awakening the HSC: Dynamic modeling of HSC maintenance unravels regulation of the TP53 pathway and quiescence. *Front Physiol.* 2020;11:848.
- [6] Meyer P, Maity P, Burkovski A, Schwab J, Müssel C, Singh K, et al. A model of the onset of the senescence associated secretory phenotype after DNA damage induced senescence. *PLoS Comput Biol* 2017;13(12):e1005741. <https://doi.org/10.1371/journal.pcbi.1005741>.
- [7] Siegle L, Schwab JD, Kühlwein SD, Lausser L, Tümpel S, Pfister AS, et al. A Boolean network of the crosstalk between IGF and Wnt signaling in aging satellite cells. *PLoS ONE* 2018;13(3):e0195126. <https://doi.org/10.1371/journal.pone.0195126>.
- [8] Werle SD, Schwab JD, Tatura M, Kirchoff S, Szekeley R, Diels R, et al. Unraveling the molecular tumor-promoting regulation of Cofilin-1 in pancreatic cancer. *Cancers (Basel).* 2021;13(4):725. <https://doi.org/10.3390/cancers13040725>.
- [9] Berestovsky N, Nakhleh L, Benos PV. An evaluation of methods for inferring Boolean networks from time-series data. *PLoS ONE* 2013;8(6):e66031. <https://doi.org/10.1371/journal.pone.0066031>.

- [10] Haider S, Pal R. Boolean network inference from time series data incorporating prior biological knowledge. *BMC Genomics* 2012;13(Suppl 6):S9. <https://doi.org/10.1186/1471-2164-13-S6-S9>.
- [11] Schwab J, Siegle L, Kühlwein S, Kühl M, Kestler H. Stability of signaling pathways during aging—a Boolean network approach. *Biology (Basel)*. 2017;6(4):46. <https://doi.org/10.3390/biology6040046>.
- [12] Schwab JD, Kühlwein SD, Ikonomi N, Kühl M, Kestler HA. Concepts in Boolean network modeling: What do they all mean? *Comput Struct Biotech J*. 2020;18:571–82.
- [13] Tanaka H, Fauré A, Matsuno H. Boolean modeling of mammalian cell cycle and cancer pathways. Miyazaki, Japan: International Conference on Artificial Life and Robotics; 2017. p. 507–510.
- [14] Pratapa A, Jalihal AP, Law JN, Bharadwaj A, Murali TM. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat Methods*. 2020;17(2):147–54.
- [15] Kurosawa S, Iwama A. Aging and leukemic evolution of hematopoietic stem cells under various stress conditions. *Inflamm Regen*. 2020;40:29.
- [16] Frisch BJ, Calvi LM. Hematopoietic stem cell cultures and assays. *Methods Mol Biol*. 2014;1130:315–24.
- [17] Jiang M, Chen H, Lai S, Wang R, Qiu Y, Ye F, et al. Maintenance of human haematopoietic stem and progenitor cells in vitro using a chemical cocktail. *Cell Discov*. 2018;4(1). <https://doi.org/10.1038/s41421-018-0059-5>.
- [18] Kobayashi H, Morikawa T, Okinaga A, Hamano F, Hashidate-Yoshida T, Watanuki S, et al. Environmental optimization enables maintenance of quiescent hematopoietic stem cells *in vivo*. *Cell Rep*. 2019;28(1):145–158. e9.
- [19] Luchsinger LL, Strikoudis A, Danzl NM, Bush EC, Finlayson MO, Satwani P, et al. Harnessing hematopoietic stem cell low intracellular calcium improves their maintenance *in vitro*. *Cell Stem Cell* 2019;25(2):225–240.e7.
- [20] Hamey FK, Göttgens B. Reconstructing gene regulatory networks that control hematopoietic commitment. *Methods Mol Biol*. 2019;1975:239–49.
- [21] Moignard V, Göttgens B. Transcriptional mechanisms of cell fate decisions revealed by single cell expression profiling. *BioEssays* 2014;36(4):419–26.
- [22] Moignard V, Woodhouse S, Haghverdi L, Lilly AJ, Tanaka Y, Wilkinson AC, et al. Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nat Biotechnol*. 2015;33(3):269–76.
- [23] De Haan G, Lazare SS. Aging of hematopoietic stem cells. *Blood* 2018;131:479–87.
- [24] Li X, Zeng X, Xu Y, Wang B, Zhao Y, Lai X, et al. Mechanisms and rejuvenation strategies for aged hematopoietic stem cells. *J Hematol Oncol*. 2020;13(1). <https://doi.org/10.1186/s13045-020-00864-8>.
- [25] Biechonski S, Milyavsky M. Differences between human and rodent DNA-damage response in hematopoietic stem cells: at the crossroads of self-renewal, aging and leukemogenesis. *Translational. Cancer Res* 2013;2.
- [26] Mejia-Ramirez E, Florian MC. Understanding intrinsic hematopoietic stem cell aging. *Haematologica* 2020;105(1):22–37.
- [27] Ratliff ML, Garton J, James JA, Webb CF. ARID3a expression in human hematopoietic stem cells is associated with distinct gene patterns in aged individuals. *Immun Ageing*. 2020;17:1–15.
- [28] Cabezas-Wallscheid N, Buettner F, Sommerkamp P, et al. Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. *Cell* 2017;169:807–823.e19.
- [29] Wilson A, Laurenti E, Oser G, van der Wath RC, Blanco-Bose W, Jaworski M, et al. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* 2008;135(6):1118–29.
- [30] Maucher M, Kracher B, Kühl M, Kestler HA (2011) Inferring Boolean network structure via correlation. *Bioinformatics*. 27:1529–1536.
- [31] Lähdesmäki H, Shmulevich I, Yli-Harja J. On learning gene regulatory networks under the Boolean network model. *Mach Learn*. 2003;52:147–67.
- [32] Harvey I, Bossomaier T. Time out of joint: Attractors in asynchronous random Boolean networks. In: *Husbands P, Harvey I, editors. Fourth European Conference on Artificial Life (ECAL97)*. Cambridge: MIT Press; 1997. p. 67–75.
- [33] Shmulevich I, Dougherty ER, Kim S, Zhang W. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics* 2002;18(2):261–74.
- [34] Huang S, Ingber DE. Shape-dependent control of cell growth, differentiation, and apoptosis: switching between attractors in cell regulatory networks. *Exp Cell Res*. 2000;261(1):91–103.
- [35] Huang S, Eichler G, Bar-Yam Y, Ingber DE. Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys Rev Lett*. 2005;94(12). <https://doi.org/10.1103/PhysRevLett.94.128701>.
- [36] Huang S. Reprogramming cell fates: reconciling rarity with robustness. *BioEssays* 2009;31(5):546–60.
- [37] Huang S. Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways. *Philos Trans R Soc Lond B Biol Sci*. 2011;366(1575):2247–59.
- [38] Kauffman SA. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: University Press; 1993.
- [39] Schwab JD, Kestler HA. Automatic Screening for Perturbations in Boolean Networks. *Front Physiol*. 2018;9:431.
- [40] Müssel C, Hopfensitz M, Kestler HA (2010) BoolNet—an R package for generation, reconstruction and analysis of Boolean networks. *Bioinformatics*. 26:1378–1380.
- [41] Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 2005;21(16):3439–40.
- [42] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30.
- [43] Hopfensitz M, Mussel C, Wawra C, Maucher M, Kuhl M, Neumann H, et al. Multiscale binarization of gene expression data for reconstructing Boolean networks. *IEEE/ACM Trans Comput Biol Bioinform*. 2012;9(2):487–98.
- [44] Müssel C, Schmid F, Blätte TJ, Hopfensitz M, Lausser L, Kestler HA. BiTrinA—multiscale binarization and trinarization with quality analysis. *Bioinformatics* 2016;32(3):465–8.
- [45] Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;32(18):2847–9.
- [46] Hao Y, Hao S, Andersen-Nissen E et al. (2020) Integrated analysis of multimodal single-cell data.
- [47] Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. *Cell* 2019;177(7):1888–1902. e21.
- [48] Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal. Complex Systems*:1695. 2006.
- [49] Sontag ED. Monotone and near-monotone biochemical networks. *Syst Synth Biol*. 2007;1(2):59–87.
- [50] Akutsu T, Miyano S, Kuhara S. Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. *Pac Symp Biocomput*. 1999;17–28.
- [51] Alon U. Network motifs: theory and experimental approaches. *Nat Rev Genet*. 2007;8(6):450–61.
- [52] Alberghante L, Blow JJ, Newman TJ (2014) Buffered Qualitative Stability explains the robustness and evolvability of transcriptional networks. *Elife*. 3: e02863.
- [53] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. *Science* 2002;298:824–7.
- [54] Snel P, Lehmann G, Bork P, Huynen MA. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res*. 2000;28:3442–4.
- [55] Martínez-Méndez D, Villarreal C, Mendoza L, Huerta L. An Integrative Network Modeling Approach to T CD4 Cell Activation. *Front Physiol*. 2020;11:380.
- [56] Méndez A, Mendoza L, Höfer T. A Network Model to Describe the Terminal Differentiation of B Cells. *PLoS Comput Biol*. 2016;12(1):e1004696. <https://doi.org/10.1371/journal.pcbi.1004696>.
- [57] Palma A, Jarrar AS, Tieri P, Cesareni G, Castiglione F. Gene Regulatory Network Modeling of Macrophage Differentiation Corroborates the Continuum Hypothesis of Polarization States. *Front Physiol*. 2018;9. <https://doi.org/10.3389/fphys.2018.01659>.
- [58] López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell* 2013;153(6):1194–217.
- [59] Klemm K, Bornholdt S. Stable and unstable attractors in Boolean networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2005;72(5). <https://doi.org/10.1103/PhysRevE.72.055101>.
- [60] Amoah A, Keller A, Emiri R, Hoenicka M, Liebald A, Vollmer A, et al. Aging of human hematopoietic stem cells is linked to changes in Cdc42 activity. *Haematologica* 2020. <https://doi.org/10.3324/haematol.2020.269670>.
- [61] Kidoya H, Muramatsu F, Shimamura T, Jia W, Satoh T, Hayashi Y, et al. Regnase-1-mediated post-transcriptional regulation is essential for hematopoietic stem and progenitor cell homeostasis. *Nat Commun*. 2019;10(1). <https://doi.org/10.1038/s41467-019-09028-w>.
- [62] Rozmus J, McDonald R, Fung S-Y, Del Bel KL, Roden J, Senger C, et al. Successful clinical treatment and functional immunological normalization of human MALT1 deficiency following hematopoietic stem cell transplantation. *Clin Immunol*. 2016;168:1–5.
- [63] Croft M, Duan W, Choi H, Eun S-Y, Madireddi S, Mehta A. TNF superfamily in inflammatory disease: translating basic insights. *Trends Immunol*. 2012;33(3):144–52.
- [64] Saito Y, Kitamura H, Hijikata A et al. (2010) Identification of therapeutic targets for quiescent, chemotherapy-resistant human leukemia stem cells. *Sci Transl Med*. 2:17ra9.
- [65] Ryu SW, Woo JH, Kim YH, Lee YS, Park JW, Bae YS. Downregulation of protein kinase CKII is associated with cellular senescence. *FEBS Lett*. 2006;580:988–94.
- [66] Janssens S, Tinel A, Lippens S, Tschopp J. PIDD mediates NF- $\kappa$ B activation in response to DNA damage. *Cell* 2005;123(6):1079–92.
- [67] Dawar S, Shahrin NH, Sladojevic N et al. (2016) Impaired haematopoietic stem cell differentiation and enhanced skewing towards myeloid progenitors in aged caspase-2-deficient mice. *Cell Death Dis*. 7:e2509.
- [68] Rhyasen GW, Bolanos L, Starczynowski DT. Differential IRAK signaling in hematologic malignancies[letter]. *Exp Hematol* 2013;41(12):1005–7.
- [69] Yamashita M, Passetgué E. TNF- $\alpha$  coordinates hematopoietic stem cell survival and myeloid regeneration. *Cell Stem Cell* 2019;25(3):357–372.e7.
- [70] Broxmeyer HE, Starnes T, Ramsey H et al. The IL-17 cytokine family members are inhibitors of human hematopoietic progenitor proliferation[letter]. *Blood*. 2006. 108. 2. 770.
- [71] Israel L, Mellett M. Clinical and genetic heterogeneity of CARD14 mutations in psoriatic skin disease. *Front Immunol*. 2018;9:2239.
- [72] Matsumoto R, Dainichi T, Tsuchiya S et al. (2018) Epithelial TRAF6 drives IL-17-mediated psoriatic inflammation. *JCI Insight*. 3.
- [73] Mellett M, Meier B, Mohanan D, Schairer R, Cheng P, Satoh TK, et al. CARD14 gain-of-function mutation alone is sufficient to drive IL-23/IL-17-mediated psoriasiform skin inflammation *in vivo*. *J Invest Dermatol*. 2018;138(9):2010–23.

- [74] Mojsilović S, Jauković A, Santibañez JF, Bugarski D. Interleukin-17 and its implication in the regulation of differentiation and function of hematopoietic and mesenchymal stem cells. *Mediators Inflamm.* 2015;2015:1–11.
- [75] Tesio M, Tang Y, Müdder K et al. 2015. Hematopoietic stem cell quiescence and function are controlled by the CYLD-TRAF2-p38MAPK pathway. *J Exp Med.* 212. 525–538.
- [76] Zhang J, Li Li, Baldwin AS, Friedman AD, Paz-Priel I, Zambidis ET. Loss of IKK $\beta$  but not NF- $\kappa$ B p65 skews differentiation towards myeloid over erythroid commitment and increases myeloid progenitor self-renewal and functional long-term hematopoietic stem cells. *PLoS ONE* 2015;10(6):e0130441. <https://doi.org/10.1371/journal.pone.0130441>.
- [77] Chakravarti B, Abraham GN. Effect of age and oxidative stress on tyrosine phosphorylation of ZAP-70. *Mech Ageing Dev.* 2002;123(4):297–311.
- [78] Moore JC, Mulligan TS, Yordán NT, Castranova D, Pham VN, Tang Q, et al. T cell immune deficiency in zap70 mutant zebrafish. *Mol Cell Biol.* 2016;36(23):2868–76.
- [79] Nakano H, Oshima H, Chung W, et al. TRAF5, an activator of NF- $\kappa$ B and putative signal transducer for the lymphotoxin-beta receptor. *J Biol Chem.* 1996;271:14661–4.
- [80] Nakano H, Sakon S, Koseki H, Takemori T, Tada K, Matsumoto M, et al. Targeted disruption of Traf5 gene causes defects in CD40- and CD27-mediated lymphocyte activation. *Proc Natl Acad Sci U S A.* 1999;96(17):9803–8.
- [81] Ping Z, Chen Si, Hermans SJF, Kenswil KJG, Feyen J, van Dijk C, et al. Activation of NF- $\kappa$ B driven inflammatory programs in mesenchymal elements attenuates hematopoiesis in low-risk myelodysplastic syndromes. *Leukemia* 2019;33(2):536–41.
- [82] Lannutti BJ, Drachman JG. Lyn tyrosine kinase regulates thrombopoietin-induced proliferation of hematopoietic cell lines and primary megakaryocytic progenitors. *Blood* 2004;103:3736–43.
- [83] Orschell CM, Borneo J, Munugalavada V, Ma P, Sims E, Ramdas B, et al. Deficiency of Src family kinases compromises the repopulating ability of hematopoietic stem cells. *Exp Hematol.* 2008;36(5):655–66.
- [84] Astro V, Tonoli D, Chiaretti S, Badanai S, Sala K, Zerial M, et al. Liprin- $\alpha$ 1 and ERC1 control cell edge dynamics by promoting focal adhesion turnover. *Sci Rep.* 2016;6(1). <https://doi.org/10.1038/srep33653>.
- [85] Ducut Sigala JL, Bottero V, Young DB, Shevchenko A, Mercurio F, Verma IM. Activation of transcription factor NF- $\kappa$ B requires ELKS, an IkappaB kinase regulatory subunit. *Science* 2004;304:1963–7.
- [86] Gillingham AK, Bertram J, Begum F, Munro S. 2019. In vivo identification of GTPase interactors by mitochondrial relocalization and proximity biotinylation. *Elife.* 8.
- [87] Lansbergen G, Grigoriev I, Mimori-Kiyosue Y, et al. CLASPs attach microtubule plus ends to the cell cortex through a complex with LL5beta. *Dev Cell.* 2006;11:21–32.
- [88] Zamboni V, Jones R, Umbach A, Ammoni A, Passafaro M, Hirsch E, et al. Rho GTPases in intellectual disability: from genetics to therapeutic opportunities. *Int J Mol Sci.* 2018;19(6):1821. <https://doi.org/10.3390/ijms19061821>.
- [89] Rossi DJ, Bryder D, Zahn JM, Ahlenius H, Sonu R, Wagers AJ, et al. Cell intrinsic alterations underlie hematopoietic stem cell aging. *Proc Natl Acad Sci U S A.* 2005;102(26):9194–9.
- [90] Rossi D, Bryder D, Weissman I. Hematopoietic stem cell aging: mechanism and consequence. *Exp Gerontol.* 2007;42(5):385–90.
- [91] Wu Z-H, Mabb A, Miyamoto S. PIDD: a switch hitter. *Cell* 2005;123(6):980–2.
- [92] Heo S-K, Noh E-K, Gwon G-D, Kim JY, Jo J-C, Choi Y, et al. LIGHT (TNFSF14) Increases the Survival and Proliferation of Human Bone Marrow-Derived Mesenchymal Stem Cells. *PLoS ONE* 2016;11(11):e0166589. <https://doi.org/10.1371/journal.pone.0166589>.