Appendix

Reconstruction Boolean network ensembles form single-cell data for unravelling dynamics in the aging of human hematopoietic stem cells

Julian D. Schwab^{a,#}, Nensi Ikonomi^a, Silke D. Werle^{a,#}, Felix M. Weidner^{a,#}, Harmut Geiger^b, Hans A. Kestler^{a,*}

a Institute of Medical Systems Biology, Ulm University, Albert-Einstein-Allee 11, 89081 10 Ulm, Germany b Departement of Molecular Medicine, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany

equal contribution

*corresponding author

15

5

Tables

Table A.1 Available single-cell RNA sequencing data for HSC aging. Summary of available datasets including information on reference, species, number of samples, and details on the sample composition.

numoci of samples, and details on the sample composition.				
Author	Accession number	Species	#Samples	Comment
Kowalczyk et al. 2015 [1]	GSE59114	Mus musculus	2128	-Two mouse strains -All mice are pooled -LT-HSC, ST-HSCs, MPPs and from young and old mice
Grover et al. 2016 [2]	GSE70657	Mus musculus	135	-2 ages LT- -only HSCs
Kirschner et al. 2017 [3]	GSE87687	<i>HPC7</i> cells from mus musculus	12	-TPO treatment -4 points in times of treatment

Preprint submitted to Journal of Computational and Structural Biotechnology

Table A.2 Complete inactive genes in all ensemble networks. The color legend indicates if the literature findings match the expected behavior of a quiescent LT-HSC. Green indicates that the activity is coherent with quiescent HSC, while for pink genes, the activity is contrasting. For blue highlighted 25 genes, we did not find any information in the context of HSCs. In total, genes that have an activity level $\leq 1\%$ are considered to be inactive.

Table A.3 Expected hallmarks of HSCs aging matched by the attractor landscape. Mainly described hallmarks of HSC aging are reported and their 30 matched attractor activity. Green indicates that the hallmark is captured by the attractor landscape. Pink indicates that the hallmark is not captured by the attractor pattern.

Figures

Figure A.1: Clustering of log-normalized gene expression data of $NF-\kappa B$ pathway by individual (eight individuals in total: young $A = 19$ years old (y.o), $B = 40$ y.o., $C = 21$ 40 y.o., $D = 37$ y.o., aged A = 66 y.o., B = 70 y.o., C = 61 y.o., D = 68 y.o.). Aged Samples are colored in different shades of grey and young in shades of green.

Figure A.2: Clustering log-normalized and z-transformed gene expression data of NF-45 $\,\kappa$ B pathway by individual (eight individuals in total: young A = 19 years old (y.o), B = 40 y.o., C = 21 y.o., D = 37 y.o., aged A = 66 y.o., B = 70 y.o., C = 61 y.o., D = 68 y.o.). Aged samples are colored in different shades of grey and young in shades of green.

Figure A.3: tSNE plot of gene expression data of NF- κ B genes (according to KEGG-DB) using log-normalized expression data. Data was centered and z-transformed before applying tSNE, all using the R-package Seurat. Each dot shows one sample. Aged samples are colored in red, young samples in blue.

Figure A.4: tSNE plot of gene expression data of NF- κ B genes (according to KEGG-DB) using log-normalized expression data. Data was centered and z-transformed before applying tSNE, all using the R-package Seurat. Each dot shows one sample. Samples 60 are colored according to the corresponding individual (eight individuals in total: young A = 19 years old (y.o), B = 40 y.o., C = 21 y.o., D = 37 y.o., aged A = 66 y.o., B = 70 y.o., $C = 61$ y.o., $D = 68$ y.o.).

Figure A.5: Runtime of network reconstruction algorithms using *best-fit algorithm* with and without *inferInput* by Maucher et al. [68] as preprocessing step (here named filtered best-fit). Runtime is measured in seconds (y-axis). The x-axis shows the varying network size. 20 time series tuples were measured for each of 100 random networks of 70 each size between 20 and 200 in steps of 20. Simulation was performed for both on non-noisy (A) and noisy (5% noise, B) time-series data.

75 inputs in each regulatory function for each reconstructed network with the corresponding original network. Figure A shows the results for the reconstruction from data without noise, Figure B for reconstruction from data with 5% noise. The sensitivity distribution is measured over the 100 random networks of each size from 20 to 200 (xaxis), respectively. All reconstructions are based on time- series with 20 time points.

80 Results are measured for the different reconstruction approaches *best-fit* (red) and *filtered best-fit* with *inferInput* by Maucher et al. [68] as preprocessing step (blue).

85 Figure A.7: Boxplots showing the measured specificity when predicting the regulatory inputs in each regulatory function for each reconstructed network with the corresponding original network. Figure A shows the results for the reconstruction from data without noise, Figure B for reconstruction from data with 5% noise. The sensitivity distribution is measured over the 100 random networks of each size from 20 to 200 (x-

90 axis), respectively. All reconstructions are based on time- series with 20 time points. Results are measured for the different reconstruction approaches *best-fit* (red) and *filtered best-fit* with *inferInput* by Maucher et al. [68] as preprocessing step (blue).

Figure A.8: Runtime of network reconstruction algorithms using best-fit algorithm and filtered best-fit algorithm with *inferInput* by Maucher et al. [68] as preprocessing step. Runtime is measured in seconds (y-axis). Time-scales on the y-axis are shown in logarithmic scaling. The x-axis shows the varying network size. 10 random networks 100 of each size between $|V| = 20$ and $|V| = 200$ in steps of 20 were created. Based on these network time-series with $|V| + 10$ number of time points. Simulation was performed for

both on non-noisy (A) and noisy (B) time-series data.

- 105 Figure A.9: Boxplots showing the measured sensitivity when predicting the regulatory inputs in each regulatory function for each reconstructed network with the corresponding original network. Figure A shows the results for the reconstruction from data without noise, Figure B for reconstruction from data with 5% noise. The sensitivity distribution is measured over the 100 random networks of each size from 20 to 200 (x-
- 110 axis), respectively. Reconstructions are based on time-series with |V| + 10 time points, where IVI refers to the number of regulatory components in the network. Results are measured for the different reconstruction approaches *best-fit* (red) and *filtered best-fit* with *inferInput* by Maucher et al. [68] as preprocessing step (blue).

Figure A.10: Boxplots showing the measured specificity when predicting the regulatory inputs in each regulatory function for each reconstructed network with the corresponding original network. Figure A shows the results for the reconstruction from data without noise, Figure B for reconstruction from data with 5% noise. The sensitivity

120 distribution is measured over the 100 random networks of each size from 20 to 200 (xaxis), respectively. Reconstructions are based on time-series with $|V| + 10$ time points, where IVI refers to the number of regulatory com- ponents in the network. Results are measured for the different reconstruction approaches *best-fit* (red) and *filtered best-fit* with *inferInput* by Maucher et al. [68] as preprocessing step (blue).

Figure A.11: Static measures per individual. The four properties as described in Section 2.6 for each individual (eight individuals in total: young $A = 19$ years old. (y.o.), $B =$ 130 40 y.o., C = 21 y.o., D = 37 y.o., aged A = 66 y.o., B = 70 y.o., C = 61 y.o., D = 68 y.o.).

- Figure A.12: Comparison of interactions with String DB. (A) shows the matrix of the 135 comparison of all young network interactions to the STRING DB using experimental and database knowledge. (B) shows the matrix of the comparison of all aged network interactions to the STRING DB using experimental and database knowledge. (C) shows the matrix of the comparison of all young network interactions to the STRING DB using experimental, database knowledge, text-mining, and co-expression knowledge. 140 (D) shows the matrix of the comparison of all aged network interactions to the STRING
	- DB using experimental, database knowledge, text-mining, and co-expression knowledge. Yellow indicates matches between Boolean network and STRING DB, green mean no interaction in the Boolean network, and blue indicates a mismatch between Boolean network and STRING DB.

Figure A.13: Boxplots showing attractor properties in different ensembles. The number of attractors of each network within the ensembles was measured and grouped via age

150 (A) and individual (C). The number of attractors found for each network within the ensembles was measured and grouped via age (B) and individual (D). Attractors were simulated exhaustively for all 100 networks in the ensembles representing each individual in 20 repeated runs (1000 random pseudo time tuples per run; eight individuals in total: young A = 19 years old (y.o.), B = 40 y.o., C = 21 y.o., D = 37 y.o., 155 aged A = 66 y.o., B = 70 y.o., C = 61 y.o., D = 68 y.o.).

Figure A.14: Boxplots showing the number of bi-fan and feedforward loop motifs 160 measured in the 2000 randomly sampled networks from each individuals' ensemble (100 networks with 20 repetitions). Individuals are encoded as follows: young $A = 19$ years old (y.o.), $B = 40$ y.o., $C = 21$ y.o., $D = 37$ y.o., aged $A = 66$ y.o., $B = 70$ y.o., C $= 61$ y.o., D = 68 y.o.

References

- ¹⁶⁵ 1. Kowalczyk MS, Tirosh I, Heckl D et al. (2015) Single-cell RNAseq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Res. 25:1860- 1872.
- 2. Grover A, Sanjuan-Pla A, Thongjuea S et al. (2016) Single-cell ¹⁷⁰ RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. Nat Commun. 7:11075.
- 3. Kirschner K, Chandra T, Kiselev V et al. (2017) Proliferation Drives Aging-Related Functional Decline in a Subpopulation of the Hematopoietic Stem Cell Compartment. Cell Rep. 19:1503- ¹⁷⁵ 1511.

renewal of hematopoietic stem/progenitor cells and induces bone marrow failure. Stem Cells. 35:777-786.

- 39. McIntyre BAS, Alev C, Tarui H, Jakt LM, Sheng G (2008) Expression profiling of circulating non-red blood cells in ²⁹⁵ embryonic blood. BMC Dev Biol. 8:21.
	- 40. Höpner SS, Raykova A, Radpour R et al. (2021) LIGHT/LTβR signaling regulates self-renewal and differentiation of hematopoietic and leukemia stem cells. Nat Commun.
- 41. Bogert NV, Furkel J, Din S et al. (2020) A novel approach to ³⁰⁰ genetic engineering of T-cell subsets by hematopoietic stem cell infection with a bicistronic lentivirus. Scientific reports.
- 42. Gallay P, Heumann D, Le Roy D, Barras C, Glauser MP (1993) Lipopolysaccharide-binding protein as a major plasma protein responsible for endotoxemic shock. Proc Natl Acad Sci U S A. ³⁰⁵ 90:9935-9938.
	- 43. Rhyasen GW, Bolanos L, Fang J et al. (2013) Targeting IRAK1 as a therapeutic approach for myelodysplastic syndrome. Cancer Cell. 24:90-104.

44. Pietras EM, Mirantes-Barbeito C, Fong S et al. (2016) Chronic ³¹⁰ interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. Nat Cell Biol. 18:607-618.

- 45. Liu YF, Zhang SY, Chen YY et al. (2018) ICAM-1 deficiency in the bone marrow niche impairs quiescence and repopulation of ³¹⁵ hematopoietic stem cells. Stem Cell Reports. 11:258-273.
	- 46. Thalheimer FB, Wingert S, De Giacomo P et al. (2014) Cytokineregulated GADD45G induces differentiation and lineage selection in hematopoietic stem cells. Stem Cell Reports. 3:34-43.

47. Gupta M, Gupta SK, Balliet AG et al. (2005) Hematopoietic cells ³²⁰ from Gadd45a- and Gadd45b-deficient mice are sensitized to genotoxic-stress-induced apoptosis. Oncogene. 24:7170-7179.

- 48. Himburg HA, Sasine J, Yan X, Kan J, Dressman H, Chute JP (2016) A Molecular Profile of the Endothelial Cell Response to Ionizing Radiation. Radiat Res. 186:141-152.
- ³²⁵ 49. Zhang N-N, Shen S-H, Jiang L-J et al. (2008) RIG-I plays a critical role in negatively regulating granulocytic proliferation. Proc Natl Acad Sci U S A. 105:10553-10558.

- 62. Li X, Zeng X, Xu Y et al. (2020) Mechanisms and rejuvenation strategies for aged hematopoietic stem cells. J Hematol Oncol. 13:31.
- ³⁷⁰ 63. Afreen S, Bohler S, Müller A et al. (2020) BCL-XL expression is essential for human erythropoiesis and engraftment of hematopoietic stem cells. Cell Death Dis. 11:8.
- 64. Orelio C, Dzierzak E (2007) Bcl-2 expression and apoptosis in the regulation of hematopoietic stem cells. Leuk Lymphoma. 48:16- ³⁷⁵ 24.
	- 65. Rozmus J, McDonald R, Fung SY et al. (2016) Successful clinical treatment and functional immunological normalization of human MALT1 deficiency following hematopoietic stem cell transplantation. Clin Immunol. 168:1-5.
- ³⁸⁰ 66. Wang Y, Sun X, Yuan S et al. (2020) Interleukin-1β inhibits normal hematopoietic expansion and promotes acute myeloid leukemia progression via the bone marrow niche. Cytotherapy. 22:127-134.
- 67. De Haan G, Lazare SS (2018) Aging of hematopoietic stem cells. ³⁸⁵ Blood. 131:479-487.

68. Maucher M, Kracher B, Kühl M, Kestler HA (2011) Inferring Boolean network structure via correlation. Bioinformatics. 27:1529-1536.