Current Opinion in Systems Biology

Mechanistic models of blood cell fate decisions in the era of single-cell data

--Manuscript Draft--

Manuscript Number:	COISB-D-21-00031R1
Full Title:	Mechanistic models of blood cell fate decisions in the era of single-cell data
Article Type:	Special Session "Theoretical approaches to analyze single-cell data" (2021)
Short Title:	Models of blood fate decisions for single-cells
Keywords:	mechanistic model; blood cell decision making; single cell data
Corresponding Author:	Carsten Marr Helmholtz Zentrum München Deutsches Forschungszentrum für Umwelt und Gesundheit: Helmholtz Zentrum Munchen Deutsches Forschungszentrum fur Gesundheit und Umwelt GERMANY
Corresponding Author's Institution:	Helmholtz Zentrum München Deutsches Forschungszentrum für Umwelt und Gesundheit: Helmholtz Zentrum Munchen Deutsches Forschungszentrum fur Gesundheit und Umwelt
Corresponding Author E-Mail:	carsten.marr@helmholtz-muenchen.de
First Author:	Carsten Marr
Order of Authors:	Carsten Marr
	Ingmar Glauche
Abstract:	Billions of functionally distinct blood cells emerge from a pool of hematopoietic stem cells in our bodies every day. This progressive differentiation process is hierarchically structured and remarkably robust. We provide an introductory review to mathematical approaches addressing the functional aspects of how lineage choice is potentially implemented on a molecular level. Emerging from studies on the mutual repression of key transcription factors we illustrate how those simple concepts have been challenged in recent years and subsequently extended. Especially the analysis of omics data on the single cell level with computational tools provide descriptive insights on a yet unknown level, while their embedding into a consistent mechanistic and mathematical framework is still incomplete.
Author Comments:	We would like to suggest Bertie Göttgens and Sui Huang as reviewers

Mechanistic models of blood cell fate decisions in the era of single-cell data

Ingmar Glauche[1,*] & Carsten Marr[2,*]

[1]Institute for Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Dresden, Germany.

[2]Institute of Computational Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.

*correspondence to ingmar.glauche@tu-dresden.de and carsten.marr@helmholtz-muenchen.de

Abstract

Billions of functionally distinct blood cells emerge from a pool of hematopoietic stem cells in our bodies every day. This progressive differentiation process is hierarchically structured and remarkably robust. We provide an introductory review to mathematical approaches addressing the functional aspects of how lineage choice is potentially implemented on a molecular level. Emerging from studies on the mutual repression of key transcription factors we illustrate how those simple concepts have been challenged in recent years and subsequently extended. Especially the analysis of omics data on the single cell level with computational tools provide descriptive insights on a yet unknown level, while their embedding into a consistent mechanistic and mathematical framework is still incomplete.

The production of blood cells as a dynamic equilibrium

The hematopoietic system is a primary example to illustrate how a population of resident tissue stem cells continuously contributes to the replenishment of a spectrum of mature cell types. While the dynamics and the particular contribution to the different blood lineages are tightly regulated processes, the overall theme, namely the differentiation of a pool of “undecided”, multipotent stem cells towards committed cell types, is remarkably robust. The accessibility, but also the clinical relevance in the context of stem cell transplantation, make hematopoiesis probably the best-studied example of cellular decision making in mammals.

On a more abstract level, the differentiation of hematopoietic stem cells into mature blood cells is the realization of a general principle that also governs the development of multicellular organisms, consisting of cells with distinct phenotypes and functions. During development, these different cell types arise from a common, totipotent cell (usually a zygote) in a sequence of decision steps. It is well established that a wealth of spatial, physical and temporal factors influence this process while the repertoire of developmental options is encoded in the genomic sequence. However, it is still one of the fundamental questions in biology, how this genotypic basis translates into a robust phenotypic decision process with adequate abundance, timing, and spatial orientation.

In the 1950s, Conrad Waddington [1] suggested a visualization of this decision making process by comparing it to a marble rolling down a slope with emerging valleys and ridges and thus projecting the myriad of molecular, morphological, and functional features onto a three dimensional landscape. With the rise of molecular biology, this picture provided an interpretation of the valleys as cell types with distinct biochemical configurations, making Waddington’s landscape a landmark visualization for stem cell differentiation. Already in the 1960s and 70s, Stuart Kauffman developed the formal concept that cell lineages can be viewed as attractors in gene regulatory networks and that cell differentiation represents a transition between those dynamic equilibria [2]. However, understanding the nature of the branching point that initially separates two or more distinct lineages is still an ongoing struggle.

Stem cell decision making and the concept of bistability

Mathematical modelling has long accompanied research on molecular decision making [3–6]. Corresponding models contributed particularly to the understanding of switch-like decisions in bacteria, as they commonly occur in response to changing environmental conditions [7]. In seminal work, Timothy Gardner and colleagues took a reductionists approach to demonstrate how a synthetic genetic toggle switch built on the mutual interaction of two repressible operons establishes bistability and exemplifies the minimal requirements of a binary molecular decision making process, in which the choice of one option intrinsically represses the other [8]. This simple but elegant example combined the mathematical concept with its experimental validation in a bacterial cell line.

Around the same time, efforts intensified to experimentally investigate the role of transcription factors in initiating and maintaining decision processes during hematopoietic differentiation [9,10]. Several studies revealed an antagonism between the transcription factor PU.1, which acts as a key regulator in myeloid and lymphoid cells and is constantly upregulated in granulocytic-monocytic progenitors, and the transcription factor Gata-1, which is required for the differentiation and maturation of megakaryocytic-erythroid cells (reviewed in [10,11]). The further discovery of direct, repressive interactions between the two factors PU.1 and Gata-1 as well as the identification of an uncommitted, up-stream state, in which lineage specific transcription factors are lowly expressed, led to a series of conceptual works to explore whether bistable switches can also serve as blueprint for decision making in hematopoiesis [12]. Formalizing this concept in terms of ordinary differential equations (ODE), several models [11,13,14] predicted that the hematopoietic differentiation landscape qualitatively changes from a co-expression state of PU.1 and Gata-1 (usually referred to as priming) towards a bistable region in which the system converges to either of the two dominating states (indicating commitment, compare Figure 1). Such qualitative changes in the available system states (referred to as bifurcations) appear as a plausible conceptualization to account for progressive differentiation processes in tissue formation and maintenance, and can also be facilitated by other molecular network motifs such as feed forward loops or positive autoregulation. The concept of direct interactions between potentially counteracting transcription factors has also been applied to related phenomena in hematology [15,16], as well as to other tissues [17,18].

Extensions to the toggle switch model

The initial ODE-based models centered around the emergence of steady states through bifurcations and symmetry breaking, while sacrificing many of the molecular details. For the example of the PU.1-Gata-1 toggle switch, increasingly precise measurements revealed additional properties that had not been captured by the simple, conceptual models and fostered more sophisticated approaches. As first reported in [19], only tens of PU.1 mRNAs are present in progenitor blood cells, prompting the development of stochastic toggle switch models [20], some of them also explicitly considering transcription and translation as two processes with distinct time scales [21,22]. Bayesian networks that included epigenetic and gene expression data successfully accounted for epigenetic changes during lineage specification [23]. Beyond the regulatory motif of only two transcription factors, Boolean networks can be used to represent major interaction axes and describe on/off gene expression states [24–26]including further upstream and downstream regulators. Modeling challenges remain, in particular the inherent non-stationarity of the system due to cell cycle that induces a continuous growth of cell volume and the number of cellular

constituents, but also abrupt changes e.g. in the number of gene copies during DNA replication or during cell division.

Single cell omics profiling of lineage decisions in blood

The single cell omics revolution that we currently witness provides an unprecedentedly wide and detailed view on cellular differentiation (reviewed in [27]). After pioneer RNA sequencing of thousands of genes in thousands of single blood cells in mouse embryos [25] and adult humans [28], it is now recognized that differentiation shows many features of a continuous process instead of a sequential transition through clearly separable states [29,30]. To arrive at this conclusion and harness the large, high-dimensional single cell data, we rely on computational methods. A variety of tools has been established to allow trajectory inference and pseudotime estimation in high-dimensional state spaces (compared in [31]), recently also exploiting splicing gradients in single cells to predict future gene expression through RNA velocity [32,33]. Nowadays, multiple tools are even bundled in convenient packages [34,35]. Single cell RNA sequencing has recently also been used to link molecular profiles to stem cell functionality in transplantation assays [36] and to identify hierarchical mixed-lineage states preceding differentiation [37]. The combination of single-cell transcriptomics and lineage tracing methods can further reveal to which extent differentiation preferences are clonaly fixed or subject to external influences [38]. Beyond gene expression, chromatin accessibility via single cell ATAC sequencing in defined hematopoietic populations revealed heterogeneities in human progenitor populations and allowed analyzing regulation and expression of transcription factors [39], while CITE-seq allows to measure both protein abundance and gene expression [40]. Notably, recently launched commercial kits to simultaneously profile gene expression and chromatin state or surface markers in the same cell will allow for increasingly precise analyses of cell identity. The combination of these two data types has also been used for the assembly of regulatory networks from gene-gene correlations [41–43] and points towards a more comprehensive approach beyond mRNA abundance that also integrates information of the epigenetic level.

Tracking hematopoietic decision making with real time imaging

In contrast to the snapshot data obtained from the simultaneous analysis of thousands of cells, quantitative live cell imaging techniques with fluorescent reporters allow to continuously track the abundance of a few essential regulators over time [44]. While RNA sequencing is powerful to identify differential expression patterns along a pseudotime trajectory, image based tracking can measure gene expression in real time. Here however, the analysis is limited by the number of labelled factors and separable fluorescence channels. The resulting data catches the dynamics of potentially interacting proteins within the same cell and provides essential information about timing, coexpression and inheritance of protein expression, but also about their variability between cells. While fluorescence labeling of critical transcription factors during hematopoietic differentiation yields great potential for dynamical assessment of the interaction network [45], few works took live cell imaging data to the point at which mechanistic mathematical models could be evaluated [46]. In a recent work, single-molecule RNA imaging of PU.1, Gata-1 and Gata-2 was combined with time-lapse microscopy [47]. Stochastic modeling suggests that differentiation is preceded by a reversible transition between different co-expression states that most likely results from the intrinsic stochasticity of gene expression.

Linking single cell data with mechanistic models

The notion of a continuous differentiation process, together with the appearance of intrinsic molecular heterogeneity make the identification and understanding of molecular principles governing hematopoietic decisions a fascinating and challenging endeavor. The conceptual notion of differentiation as a transition between dynamical attractors in larger gene regulatory networks has been conceptually discussed since the late 1960s [2,48]. The single cell omics revolution largely confirms this interpretation and describes differentiation as a continuous trajectory in a high-dimensional state space. While this topological embedding is increasingly accepted, we need to further identify the "driving forces" that direct cells along those trajectories and execute decision processes at the branching points.

Several works aim to link the quasi-potential Waddington landscape with regulatory networks in more formal [49,50] or data-driven [51] manifestations, but up to now, mathematical models are rarely applied to the rapidly increasing volume of single cell data to foster a *mechanistic* understanding of the decision processes (Fig. 1). While population approaches can be fitted to the overall gene expression trajectory of a differentiation process [52], a stochastic transition model (like in [47]) captures the sequence of co-expression states preceding differentiation onset. In contrast to descriptive approaches, a general framework to evaluate mechanistic interaction models with single cell data faces a number of challenges: How can we identify the crucial regulators initiating and directing differentiation? How to infer the true differentiation time from snapshot data? And which model classes can sufficiently represent the expression and temporal heterogeneity observed in single cell data? Modern machine learning methods, such as informed neural networks [53], graph neural networks, or autoencoders [54] might help to reduce the number of genes and features to a minimum and merge different data modalities and prior knowledge.

More generally, we may need to answer the question whether the low-dimensional representation of a differentiation process as proposed by Waddington and propagated by modern embeddings truly emerges from binary decision processes that are captured with the mathematics of bifurcation theory. There is no inherent reason why gene regulatory networks should favor the cross regulation of only two factors, especially if molecular redundancies can stabilize a decision process and make it more robust to variation. The theory of gene regulatory networks is a powerful theoretical approach to explain cellular plasticity and decision making, but it needs to be extended to include other data types and to deal with variability, cellular heterogeneity, and higher-dimensional regulations. However, it remains to be shown how the phenotypic, but also molecular data obtained from advanced sequencing and imaging techniques can truly be embedded into such an extended mathematical framework describing mutual interactions of key transcription factors on the mechanistic and biochemical level, and whether our basic understanding of molecular regulations and interactions provides the right framework to account for the apparent heterogeneity and robustness of lineage specification processes.

Acknowledgements

We thank Maren Büttner, Thomas Zerjatke, Lea Schuh, and Moritz Thomas for comments on the manuscript. C.M. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 866411).

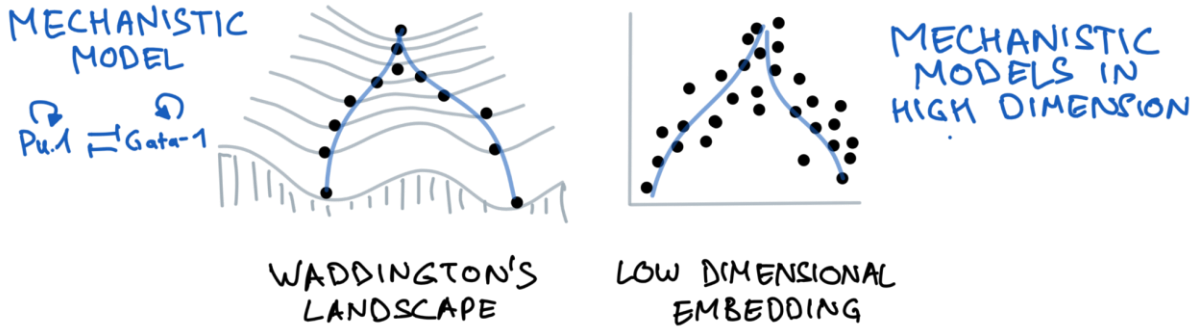

Figure 1: While the PU.1-Gata-1 toggle switch model (left) has been successfully used to describe the concept of binary fate choice from a conceptual perspective, it evidently does not capture the full complexity of the early myeloid lineage decision. Single-cell gene expression analysis provides a better understanding of differentiation trajectories (often visualised as low-dimensional embeddings), while an explicit link between causative, mechanistic mathematical models and the high dimensional data is still an ongoing challenge.

References

1. Waddington CH: *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. George Allen & Unwin, London, UK; 1957.

2. Kauffman SA: **Metabolic stability and epigenesis in randomly constructed genetic nets**. *J Theor Biol* 1969, **22**:437–467.

3. Griffith JS: **Mathematics of cellular control processes II. Positive feedback to one gene**. *Journal of Theoretical Biology* 1968, **20**:209–216.

4. Griffith JS: **Mathematics of cellular control processes. I. Negative feedback to one gene**. *J Theor Biol* 1968, **20**:202–208.

5. Tyson JJ, Othmer HG: **The dynamics of feedback control circuits in biochemical pathways**. *Progress in theoretical biology* 1978, **5**:1–62.

6. Ackers GK, Johnson AD, Shea MA: **Quantitative model for gene regulation by lambda phage repressor**. *Proc Natl Acad Sci U S A* 1982, **79**:1129–1133.

7. Mackey MC, Santillán M, Tyran-Kamińska M, Zeron ES: **Generic Deterministic Models of Prokaryotic Gene Regulation**. *Lecture Notes on Mathematical Modelling in the Life Sciences* 2016, doi:10.1007/978-3-319-45318-7_1.

8. Gardner TS, Cantor CR, Collins JJ: **Construction of a genetic toggle switch in Escherichia coli**. *Nature* 2000, **403**:339–342.

9. Sieweke MH, Graf T: **A transcription factor party during blood cell differentiation**. *Curr Opin Genet Dev* 1998, **8**:545–551.

10. Orkin SH: **Diversification of haematopoietic stem cells to specific lineages**. *Nat Rev Genet* 2000, **1**:57–64.

11. Roeder I, Glauche I: **Towards an understanding of lineage specification in hematopoietic stem cells: a mathematical model for the interaction of transcription factors GATA-1 and PU.1**. *J Theor Biol* 2006, **241**:852–865.

12. Huang S: **Chapter 14 - Multistability and Multicellularity: Cell Fates as High-Dimensional Attractors of Gene Regulatory Networks**. In *Computational Systems Biology*. Edited by Kriete A, Eils R. Academic Press; 2006:293–326.

13. Huang S, Guo Y-P, May G, Enver T: **Bifurcation dynamics in lineage-commitment in bipotent progenitor cells**. *Dev Biol* 2007, **305**:695–713.

14. Chickarmane V, Enver T, Peterson C: **Computational modeling of the hematopoietic erythroid-myeloid switch reveals insights into cooperativity, priming, and irreversibility**. *PLoS Comput Biol* 2009, **5**:e1000268.

15. Laslo P, Spooner CJ, Warmflash A, Lancki DW, Lee H-J, Sciammas R, Gantner BN, Dinner AR, Singh H: **Multilineage transcriptional priming and determination of alternate hematopoietic cell fates**. *Cell* 2006, **126**:755–766.

16. Spooner CJ, Cheng JX, Pujadas E, Laslo P, Singh H: **A Recurrent Network Involving the Transcription Factors PU.1 and Gfi1 Orchestrates Innate and Adaptive Immune Cell Fates**. *Immunity* 2009, **31**:576–586.

17. Chickarmane V, Troein C, Nuber UA, Sauro HM, Peterson C: **Transcriptional dynamics of the embryonic stem cell switch**. *PLoS Comput Biol* 2006, **2**:e123.

18. Schittler D, Hasenauer J, Allgöwer F, Waldherr S: **Cell differentiation modeled via a coupled two-switch regulatory network**. *Chaos* 2010, **20**:045121.

19. Warren L, Bryder D, Weissman IL, Quake SR: **Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR**. *Proc Natl Acad Sci U S A* 2006, **103**:17807–17812.

20. Schultz D, Onuchic JN, Wolynes PG: **Understanding stochastic simulations of the smallest genetic networks**. *J Chem Phys* 2007, **126**:245102.

21. Strasser M, Theis FJ, Marr C: **Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression**. *Biophys J* 2012, **102**:19–29.

22. Zhu R, Ribeiro AS, Salahub D, Kauffman SA: **Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models**. *J Theor Biol* 2007, **246**:725–745.

23. Schütte J, Wang H, Antoniou S, Jarratt A, Wilson NK, Riepsaame J, Calero-Nieto FJ, Moignard V, Basilico S, Kinston SJ, et al.: **An experimentally validated network of nine haematopoietic transcription factors reveals mechanisms of cell state stability**. *eLife* 2016, **5**.

24. Krumsiek J, Marr C, Schroeder T, Theis FJ: **Hierarchical differentiation of myeloid progenitors is encoded in the transcription factor network**. *PLoS One* 2011, **6**:e22649.

25. Moignard V, Woodhouse S, Haghverdi L, Lilly AJ, Tanaka Y, Wilkinson AC, Buettner F, Macaulay IC, Jawaid W, Diamanti E, et al.: **Decoding the regulatory network of early blood development from single-cell gene expression measurements**. *Nat Biotechnol* 2015, **33**:269–276.

26. Collombet S, van Oevelen C, Sardina Ortega JL, Abou-Jaoudé W, Di Stefano B, Thomas-Chollier M, Graf T, Thieffry D: **Logical modeling of lymphoid and myeloid cell specification and transdifferentiation**. *Proc Natl Acad Sci U S A* 2017, **114**:5792–5799.

** This review illustrates the potential and challenges of single-cell RNA sequencing techniques and how they can advance the understanding of lineage specification. The authors cover recent technological advances and their impact on the interpretation of a continuous hematopoietic differentiation landscape.

27. Watcham S, Kucinski I, Gottgens B: **New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing**. *Blood* 2019, **133**:blood–2018–08–835355.

28. Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, Hirche C, Lutz C, Buss EC, Nowak D, et al.: **Human haematopoietic stem cell lineage commitment is a continuous process**. *Nat Cell Biol* 2017, **19**:271–281.

29. Haas S, Trumpp A, Milsom MD: **Causes and Consequences of Hematopoietic Stem Cell Heterogeneity**. *Cell Stem Cell* 2018, **22**:627–638.

30. Laurenti E, Göttgens B: **From haematopoietic stem cells to complex differentiation landscapes**. *Nature* 2018, **553**:418–426.

** This broad and systematic overview compares 45 lineage inference methods on real and synthetic datasets for cellular ordering, topology, scalability and usability. The authors derive guidelines to help users select the best method for their research question.

31. Saelens W, Cannoodt R, Todorov H, Saeys Y: **A comparison of single-cell trajectory inference methods**. *Nat Biotechnol* 2019, doi:10.1038/s41587-019-0071-9.

32. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, et al.: **RNA velocity of single cells**. *Nature* 2018, **560**:494–498.

33. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ: **Generalizing RNA velocity to transient cell states through dynamical modeling**. *Nat Biotechnol* 2020, **38**:1408–1414.

34. Wolf FA, Angerer P, Theis FJ: **SCANPY: large-scale single-cell gene expression data analysis**. *Genome Biol* 2018, **19**:15.

35. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, Satija R: **Comprehensive Integration of Single-Cell Data**. *Cell* 2019, **177**:1888–1902.e21.

36. Dong F, Hao S, Zhang S, Zhu C, Cheng H, Yang Z, Hamey FK, Wang X, Gao A, Wang F, et al.: **Differentiation of transplanted haematopoietic stem cells tracked by single-cell transcriptomic analysis**. *Nat Cell Biol* 2020, **22**:630–639.

* Using single-cell RNA sequencing and a transcriptome-based classification of haematopoietic cell types this work tracks changes of highly purified HSCs early after transplantation. It demonstrates a rapid and robust differentiation without substantial HSC expansion during the first week post transplant.

37. Olsson A, Venkatasubramanian M, Chaudhri VK, Aronow BJ, Salomonis N, Singh H, Grimes HL: **Single-cell analysis of mixed-lineage states leading to a binary cell fate choice**. *Nature* 2016, **537**:698–702.

38. Kester L, van Oudenaarden A: **Single-Cell Transcriptomics Meets Lineage Tracing**. *Cell Stem Cell* 2018, **23**:166–179.

* Lineage tracing methods proved instrumental to delineate clonal relationships during differentiation. This review shows the potential of merging clonal analysis with single cell sequencing technologies.

39. Buenrostro JD, Corces MR, Lareau CA, Wu B, Schep AN, Aryee MJ, Majeti R, Chang HY, Greenleaf WJ: **Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation**. *Cell* 2018, **173**:1535–1548.e16.

40. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, Smibert P: **Large-scale simultaneous measurement of epitopes and transcriptomes in single cells**. *Nat Methods* 2017, **14**:865.

41. Moerman T, Aibar Santos S, Bravo González-Blas C, Simm J, Moreau Y, Aerts J, Aerts S: **GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks**. *Bioinformatics* 2019, **35**:2159–2161.

42. Kim J, Jakobsen ST, Natarajan KN, Won KJ: **Gene network reconstruction using single**

cell transcriptomic data reveals key factors for embryonic stem cell differentiation. *biorxiv* 2019, doi:10.1101/2019.12.20.884163.

43. Kamimoto K, Hoffmann CM, Morris SA: **CellOracle: Dissecting cell identity via network inference and in silico gene perturbation**. *bioRxiv* 2020, doi:10.1101/2020.02.17.947416.

* This review illustrates how noninvasive continuous quantitative single-cell imaging can complement population-based single-cell technologies. The authors focus on the potential of these methods to study continuous molecular dynamics.

44. Loeffler D, Schroeder T: **Understanding cell fate control by continuous single-cell quantification**. *Blood* 2019, **133**:1406–1414.

45. Hoppe PS, Schwarzfischer M, Loeffler D, Kokkaliaris KD, Hilsenbeck O, Moritz N, Endele M, Filipczyk A, Gambardella A, Ahmed N, et al.: **Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios**. *Nature* 2016, **535**:299–302.

46. Strasser MK, Hoppe PS, Loeffler D, Kokkaliaris KD, Schroeder T, Theis FJ, Marr C: **Lineage marker synchrony in hematopoietic genealogies refutes the PU.1/GATA1 toggle switch paradigm**. *Nat Commun* 2018, **9**:2697.

47. Wheat JC, Sella Y, Willcockson M, Skoultchi AI, Bergman A, Singer RH, Steidl U: **Single-molecule imaging of transcription dynamics in somatic stem cells**. *Nature* 2020, doi:10.1038/s41586-020-2432-4.

** This work studies hematopoietic decision making by single-molecule RNA fluorescence in situ hybridization of *PU.1, Gata1,* and *Gata2.* Using a mathematical model analysis the authors demonstrate the fundamental role of stochastic gene expression for maintaining transcriptional plasticity and stem cell robustness.

48. Kauffman SA, Member of the Santa Fe Institute and Professor of Biochemistry Stuart A Kauffman: *The Origins of Order: Self-organization and Selection in Evolution*. Oxford University Press; 1993.

49. Wang J, Zhang K, Xu L, Wang E: **Quantifying the Waddington landscape and biological paths for development and differentiation**. *Proceedings of the National Academy of Sciences* 2011, **108**:8257–8262.

50. Zhou JX, Aliyu MDS, Aurell E, Huang S: **Quasi-potential landscape in complex multi-stable systems**. *J R Soc Interface* 2012, **9**:3539–3553.

51. Schiebinger G, Shu J, Tabaka M, Cleary B, Subramanian V, Solomon A, Gould J, Liu S, Lin S, Berube P, et al.: **Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming**. *Cell* 2019, **176**:928–943.e22.

** Using single-cell sequencing data from mouse embryonic fibroblasts reprogrammed into iPSCs, the authors illustrate a computational framework that describes cells as time-dependent realizations from a probability distribution in gene-expression space while their transitions are formalized as transport processes. The study reports a broad spectrum of developmental programs initiated upon reprogramming and infers potentially relevant transcription factors.

52. Fischer DS, Fiedler AK, Kernfeld EM, Genga RMJ, Bastidas-Ponce A, Bakhti M, Lickert H, Hasenauer J, Maehr R, Theis FJ: **Inferring population dynamics from single-cell RNA-sequencing time series data**. *Nat Biotechnol* 2019, **37**:461–468.

* The PDE model in this study describes cell population flux along continuous transcriptomic trajectories that are obtained from single-cell RNA-sequencing data. This pseudodynamics framework accounts for confounding effects such as proliferation and cell death, and reconciles population dynamics with differentiation trajectories inferred from time-series single-cell data.

53. Raissi M, Perdikaris P, Karniadakis GE: **Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations**. *J Comput Phys* 2019, **378**:686–707.

54. Yang KD, Belyaeva A, Venkatachalapathy S, Damodaran K, Katcoff A, Radhakrishnan A, Shivashankar GV, Uhler C: **Multi-domain translation between single-cell imaging and sequencing data using autoencoders**. *Nat Commun* 2021, **12**:31.

Conflict of Interest

HelmholtzZentrum münchen
German Research Center for Environmental Health

Helmholtz Zentrum München Postfach 1129 85758 Neuherberg

Dr. Carsten Marr
Institute of Computational Biology
Deputy Director
Group Leader 'Quantitative
Single Cell Dynamics'
+49 89 3187 2158
carsten.marr@helmholtz-muenchen.de
http://icb.helmholtz-muenchen.de

12/03/2021

Dear Editors,

On behalf of my co-author, I state that there are not financial nor personal interests in relation to the work submitted for consideration to your Journal.

With best regards,

Carsten Marr

Carsten Marr

Helmholtz Zentrum München
German Research Center for
Environmental Health (GmbH)
Ingolstädter Landstrasse 1
85764 Neuherberg
Phone + 49 (0) 89 3187 2158
Fax + 49 (0) 89 3187 3322

info@helmholtz-muenchen.de
www.helmholtz-muenchen.de

HelmholtzZentrum münchen
German Research Center for Environmental Health

Helmholtz Zentrum München Postfach 1129 85758 Neuherberg

Dr. Carsten Marr
Institute of Computational Biology
Deputy Director
Group Leader 'Quantitative
Single Cell Dynamics'
+49 89 3187 2158
carsten.marr@helmholtz-muenchen.de
http://icb.helmholtz-muenchen.de

10/06/2021

Revision COISB-D-21-00031

Dear Editors,

We would like to thank you and the reviewer for your comments on our review manuscript.

We have addressed all the points raised in the following manner:

- We added "We provide an introductory review to mathematical approaches addressing the functional aspects of how lineage choice is potentially implemented on a molecular level." to the abstract (marked in red) to clarify the scope of the text, as suggested by you. We are of course open for suggestions if you had a different framing in mind.
- We expanded the descriptions of the references as proposed by you (also marked in red).
- As suggested, we expanded the coverage of Kauffman's early works (marked in red) and added 1 corresponding citation.
- As requested by Reviewer 1, we put the recent works on linking Waddington's landscape explicitly with fate determination into context (marked in red) and added 3 references.
- We added a Figure to illustrate the Pu.1-Gata-1 paradigm and the current gap in research.

Please find our revised version attached. We would be grateful for feedback or any further suggestions from your side, which we are happy to incorporate.

Best regards,
Ingmar & Carsten

Helmholtz Zentrum München
German Research Center for
Environmental Health (GmbH)
Ingolstädter Landstrasse 1
85764 Neuherberg
Phone + 49 (0) 89 3187 2158
Fax + 49 (0) 89 3187 3322

info@helmholtz-muenchen.de
www.helmholtz-muenchen.de