Conceptual Challenges of the Systemic Approach in Understanding Cell Differentiation

Andras Paldi [1]✉

Email paldi@genethon.fr

AQ1

[1] Ecole Pratique des Hautes Etudes, PSL Research University, UMRS_951, INSERM, Univ-Evry, Genethon, 1 rue de I'Internationale, Evry, 91002 France

Abstract

The cells of a multicellular organism are derived from a single zygote and genetically identical. Yet, they are phenotypically very different. This difference is the result of a process commonly called cell differentiation. How the phenotypic diversity emerges during ontogenesis or regeneration is a central and intensely studied but still unresolved issue in biology. Cell biology is facing conceptual challenges that are frequently confused with methodological difficulties. How to define a cell type? What stability or change means in the context of cell differentiation and how to deal with the ubiquitous molecular variations seen in the living cells? What are the driving forces of the change? We propose to reframe the problem of cell differentiation in a systemic way by incorporating different theoretical approaches. The new conceptual framework is able to capture the insights made at different levels of cellular organization and considered previously as contradictory. It also provides a formal strategy for further experimental studies.

Key words

Conceptual framework
Cell differentiation
Stochastic fluctuations
Metabolism
Epigenetic mechanisms
Chromatin

1. Introduction

Biology is an empirical science; nothing makes sense in the eyes of a biologist unless it is derived from experimental observations. A typical research project in biology usually follows a naive inductive logic and the role of the underlying theory is usually underestimated. Concepts are usually taken for granted and rarely questioned directly. As a consequence, biology has a tendency to see methodological or technical problems even when the difficulty is conceptual. History of science taught us that paradigm shifts and breakthroughs in a field usually require a theoretical re-foundation or conceptual reframing of the major issues. Concepts in biology, as well as in any other scientific discipline, must be revised periodically, upgraded, or replaced if necessary. The last years witnessed emerging discussions on some of the fundamental concepts in cell biology such as "cell identity" or "cell fate." These discussions were made necessary, among others, by the rapid evolution of techniques with single-cell resolution. Perhaps

this is the reason why very often they are “hidden” behind methodological issues. In addition to their high resolution, these methods are very efficient in collecting astronomic amounts of morphological, physiological, or molecular data. Contrary to the expectation of many, accumulation of data alone did not provide us with the understanding of why and how cells become phenotypically different (i.e., differentiate) and what are the driving forces of this process. There is a growing feeling that the difficulties to address cell differentiation result from the inappropriate conceptual framing of the problem. Where these difficulties come from? What are the possible ways to resolve them? The present reflection aims to show that it is possible to define a new theoretical approach and to rethink cell differentiation on the basis of our present knowledge.

2. The Quest for Classification

Traditionally, cell differentiation is defined as the phenotypic transformation of the cells from one type into another. Therefore, the first challenge is the definition of the concept of “cell type.” Textbooks and reviews usually claim that there are about 200 different cell types in the human organism. They are classified according to their morphological similarity, tissue location, function or patterns of gene expression, etc. This is easy to do when very different cells are compared; there is no difficulty to assign neurons, lymphocytes, or epithelial cells for example to different categories. But how to classify closely related cells such as those separated only by a few cell divisions? Such cells usually resemble to each other, yet they may display a broad spectrum of gene expression levels, physiological or morphological traits, etc. How to decide if they belong to the same or to a different category? The following quote taken from a recent paper illustrates the difficulties: “… should these subcategories be declared distinct cell types? What differences, be they functional, regulatory, or morphological, are sufficient to define an organism’s cellular taxonomy?” [1]. The same problem is sometimes formulated as an issue of “cell identity” and cell differentiation as a process of change of identity [2]. Others are debating whether various cell types such as stem cells represent a state or entity [3, 4]. “Identity” or “entity” are concepts borrowed from philosophical ontology. The debates on them are as old as our systematic thinking about the world and can be tracked back to Plato and Aristotle. The cell biological re-formulation of the fundamental ontological question is: do cell types exist as independent entities? If so, what are the essential distinguishing features of cell types? Unfortunately, the issue of cell identity is usually treated in a simple intuitive way. Although experimental biology is not expected to provide a solution for fundamental philosophical issues, but understanding the origin and true significance of the concepts directly imported from philosophy would stimulate their constructive conceptual framing of the cell type issue.

The intensive quest for a better classification has been triggered by the rapid development of single-cell resolution techniques, hence the illusion of a technological difficulty [5]. Earlier, biochemical or molecular methods used to characterize gene expression, protein levels, or other features needed hundreds, thousands, or even more cells and were able to provide us with population averages only. Single-cell resolution techniques are able to extract similar information from a large number of individual cells. For the first time, in addition to the average we have also a reliable measure of the variability in the population. It is not surprising that the number of recognized cell types increases steadily with the resolution of these techniques. In a recent study for example, using single-cell RNA sequencing 17 different categories of CD34+ hematopoietic cells were identified on the basis of their gene expression patterns versus only two categories when a less sensitive cytometry analysis was used [6]. There are numerous similar examples [1, 7, 8]. In general, highly sensitive single-cell resolution techniques show that even very closely related cells are different to some extent with respect to their gene expression patterns. Although two different populations of cells are easy to discriminate on the basis of the average expression level of some distinctive marker genes, it is usually difficult to assign an individual cell picked up randomly to one of these defined cell types on the basis of the single-cell gene expression profile.

Although counterintuitive at a first glance, "cell type" appears as a concept that describes groups rather than individual cells. Then, how to set the limit between "irrelevant" and "important" differences between two cells? As we could learn from philosophy, there is no simple solution to this problem and perhaps the best way is to get rid definitively of these controversial concepts. The existing pragmatic solutions measure the extent of the differences without discriminating what is relevant or irrelevant. The most frequently employed strategy is based on the collection of a large number of parameters on individual cells using single-cell RT-PCR, RNA sequencing, mass cytometry, or high-throughput image analysis of the cell morphology [1, 6, 7, 9–14]. The data obtained are analyzed using multiparametric classification algorithms that group cells in categories on the basis of their phenotypic "similarity." In this context, "similarity" between the cells is calculated as a function of the distance between the cells in a multidimensional space defined by the measured parameters. The cell phenotype is represented as a location in a multidimensional parameter space. If the measured parameters are the gene expression levels, as it is frequently the case, the number of dimensions is equal to the number of genes in the genome, and their expression levels are the coordinates that determine the exact position. Statistical analysis of the distances between these positions representing the cell phenotypes gives an estimate of the probability (*p* value) that the groups identified by the classification algorithm can be obtained by chance. If the probability for this is sufficiently low, one can accept that the cells assigned to different phenotypic groups are indeed different. The advantage of these methods is that they provide an accurate measure of the phenotypic differences between the cells on the basis of clearly defined criteria (mRNA level, protein abundance, etc.) that can be used to classify the cells. This is an important methodological step, yet it does not give answer to the original question. Assigning a cell to a given cell type remains a decision of the observer, who sets the list of parameters to be considered, the threshold *p*-value, sample size, etc. This makes the classification relative, highly dependent on the experimental context, choice of the statistical methods, and, importantly, on the subjective opinion of the investigator. Clustering algorithms are incorrectly assumed to provide objective judgment on phenotypic classification and became a standard procedure for the analysis of single-cell data. Nevertheless, while subjective, the use of the chosen classification method makes the different experiments quantitatively comparable. Therefore, they open the way to testing hypotheses on the mechanisms of cell differentiation [15] without clearly defining what a cell type is.

The single-cell data confirm that every gene expression combination is not equiprobable. Some gene expression profiles and the corresponding cellular phenotypes are more frequent, hence probably more stable than others. Genes interact with each other and this can be described as a complex network where edges represent interactions between the vertices formed by the genes. The network representation of gene-gene interactions led to the proposition that frequently observed gene expression profiles or the corresponding cellular phenotypes reveal states of the gene interaction network that are close to an attractor in the multidimensional parameter space [16, 17]. These attractors emerge as a result of mutually stabilizing interactions between a set of genes making their co-expression more frequent among the possible combinations. In the attractor interpretation therefore, a cell phenotype is a state of the gene interaction network that is more or less close to an attractor and cell differentiation is a process of transition between the attractors [17, 18]. This representation makes direct reference to the now classical "epigenetic landscape" metaphor proposed by Conrad Waddington almost 70 years ago [17, 19]. The attractor concept of cell phenotype circumvents the "continuous versus discrete" dilemma of cell classification and focuses on the temporal dynamics of the phenotypic change.

3. Temporal Dynamics, Stability, and Change

Single-cell studies uncovered another important aspect of the cellular phenotypes: the expression of the genes in a cell and, consequently, the phenotypes are fluctuating continuously. As a result, even cells in a clonal population exhibit a broad distribution of various traits. Stochasticity of gene expression was

suggested and experimentally detected long time ago [20, 21], but the phenomenon gained a significant interest only after the publication of a landmark paper in 2002 [22]. Variation of gene expression is the direct consequence of the stochasticity of biochemical reactions involving molecules present in small copy-numbers in the cell. For example, in a typical eukaryotic cell there are only two copies of each gene. Transcription factors, RNA polymerase molecules, and other components of the gene expression machinery are also present in very low concentration. Under these conditions, biochemical reactions are limited by the diffusion of the molecules and occur only when the participating molecules meet by chance. In the case of gene expression involving many different partners this leads to strong fluctuations at a time scale of minutes to hours comparable with the life cycle of the cells. These fluctuations, frequently called “noise,” ubiquitous and are unavoidable because they are caused by the very nature of the biochemical reactions [23]. Therefore, stochastic fluctuations rather than stability should be considered the default state of gene expression.

This transforms radically the way we have to consider the problem of gene expression changes during differentiation. Traditionally, gene expression is supposed to be stable. Changes during differentiation are supposed to be strictly controlled, inducing regulated transition of the cell between phenotypic states. Spontaneous gene expression fluctuations have no role in the process, they are just “noise.” However, measured and characterized experimentally, we know now that the extent of the gene expression “noise” in individual cells is comparable to the variations supposed to be regulated [24, 25]. Population level measurements provide us only with average values of the gene expression levels; they show population level tendencies and hide the individual variations. Yet, not populations, but individual cells differentiate. How to reconcile then the unstable nature of almost every characteristic of individual cells with their obvious capacity to maintain a stable phenotype? How to explain that these phenotypes can change in an orderly way? Until now the phenotype and the underlying gene expression pattern were supposed to be stable and the *explanandum* was the “change.” In the new conceptual frame, stability becomes the *explanandum*. The question to be addressed now is how a naturally fluctuating living cell can be maintained in stability. This is just the opposite of the traditional deterministic view, which has dominated biology until now.

Obviously, change and stability are a pair of complementary concepts that also raise the question of “continuous versus discrete.” A slow “change” can be seen as “stability” depending on the timescale of the observer. Averaging over a sufficiently long interval of time can filter smaller fluctuations and reveal the tendency of an individual cell to conserve or change its phenotype, gene expression levels, protein abundances, etc. The key question is what is a “sufficiently long” time interval? A pragmatic approach to this question is to take the characteristic timescale of the random fluctuations as a starting point. Purely stochastic gene expression changes occur at a characteristic timescale of minutes to hours. If the timescale of the fluctuations is longer than the cell’s lifecycle, the phenotype is usually considered stable because the daughter cells remain phenotypically close to the mother cell [26]. Slow fluctuations can therefore be seen as to reflect a kind of "memory", ~~because~~ that makesthe actual phenotype state of the cell ~~remains~~ remaining close to the previous one. From this point of view, one extreme is “no change” (full stability), where the past state is identical to the present one. This is only a theoretical possibility. It is opposed to the other extreme, also theoretical, represented by random fluctuations without memory, where no prediction of the present state from the past is possible. Real cells are never fully stable, nor they are fully ergodic. Since the whole problem is further complicated by the fact that the cells divide and usually transmit their phenotype to the daughter cells, the candidate mechanisms for slowing down natural fluctuations and stabilizing cellular phenotypes are also expected to remain active during and after cell divisions.

AQ2

4. Energy for Stability

We have learned from physics that maintaining order and stability in an open system is principally a matter of energy investment. Indeed, theoretical models and experimental verification have demonstrated that the energetic costs of the noise reduction are very high and of the same order of magnitude as the cell's capacity to produce energy [27]. Consequently, the cell has no capacity to suppress molecular fluctuations such as gene expression noise and their consequences; at best it can reduce them to some extent. The putative mechanisms must be functionally dependent on and limited by the energy-producing cellular processes.

Gene expression is a "birth and death" process. Birth is a multistep process involving transcription and translation and all the steps of maturation of the mRNA-s and proteins. "Death" is also a multistep process involving the degradation of the intermediate or the final gene products. The actual level of the gene product in the cell is determined by the rate of the synthesis and degradation [28]. Simultaneous high synthesis and degradation rates can produce similar levels as low synthesis/degradation rates. Obviously, fluctuations of the gene product concentration can also be caused by the fluctuations in both the rates of synthesis and degradation. All these processes require ATP or some other form of energy-carrying substrates. Some steps are known to contribute more to the fluctuation/stabilization process than others.

Mechanisms that dissipate chemical energy generated by the metabolism to modulate gene expression fluctuations are now well known. Molecular mechanisms known as "epigenetic modifications" of the chromatin are excellent candidates for the role of the "stabilizer" of phenotype through influencing the fluctuations of the "birth" rate. Chromatin is a macromolecular structure formed by the genomic DNA associated to proteins, essentially histones. When wrapped in the chromatin, DNA is not accessible for transcription. Transcription is only possible if the chromatin dissociates from the DNA. This is a typical stability problem. Each chromatin component carries several covalent modifications, such as acetylation, methylation, phosphorylation, poly-ADP-ribosylation, etc. that determine the overall stability of the structure. The biochemical reactions that introduce or remove these modifications are catalyzed by dedicated enzymes. The reactions form a cooperative network that brings together either a stable repressive chromatin structure (heterochromatin), which makes the DNA inaccessible to the transcriptional machinery, or an open structure (euchromatin) that allows transcription. Thanks to the cooperative nature of the reactions and despite the reversibility and very short half-life of each individual modification, both the structures can stably be maintained for a long period of time. This is a dynamic, steady-state stability resulting from the equilibrium of the permanent action of the modifying and the reverse reactions and the resulting rapid dissociation-association of the corresponding chromatin proteins [29, 30]. As a result, the chromatin around a gene is either open, allowing transcription or repressed, making transcription impossible [31]. The structure is constantly adjusted depending on the dynamic equilibrium of the "on" and "off" reactions. It has been shown that the chromatin behaves as a dynamic bistable system with hysteresis [32]. The transition between the active and repressed states of a gene is switch-like. It depends on the competition of the heterochromatin- or euchromatin-generating reaction networks and on the time spent in the previous state. A heterochromatin structure formed long time ago is more difficult to reverse than a recently generated.

Whether a gene becomes silenced or accessible for transcription is *in fine* determined by the dynamic equilibrium between the processes bringing together the permissive and repressive chromatin and on the pre-existing state of the chromatin. When accessible, transcription factors can selectively bind to the DNA in a sequence-dependent way and further bias the equilibrium toward the open state usually at the sites of transcription initiation. In this way, transcription factors contribute to the stabilization of the gene expression networks as proposed by the attractor concept of cell phenotypes. However, they cannot specifically activate a repressed gene without a prior transition of the chromatin to an at least partially

open state, because the DNA is simply not accessible for binding. This is an essential point, because it means that a silenced gene can be re-activated only after or concomitantly with the change of the chromatin structure. Transcription factors alone are not sufficient to activate a gene; they can only increase the probability of the transcription initiation of the accessible genes.

Another essential distinguishing feature of the chromatin is the capacity to "record" the previous activity of the genes due to the hysteretic dynamical properties. In this way, chromatin becomes a major component of the so-called cellular memory because it can conserve its structure over mitotic and in rare cases even over meiotic divisions. This property confers the cells the capacity to differentiate in an orderly way instead of switching irregularly between the possible phenotypes.

In summary, chromatin is a highly dynamic key player able to slow down the stochastic fluctuations of the transcription, essentially by its reversible repression. The way the chromatin structure is brought together by the epigenetic modification confers to it a memory function. When repressed by the heterochromatin, a gene cannot be transcribed. There is no mRNA production, hence no fluctuations. When accessible for transcription, the RNA synthesis is subject to stochastic effects resulting in a bursting production of mRNA molecules and generating stochastic fluctuations in their number. These fluctuations can be amplified or buffered by the consecutive steps of translation and degradation of the gene products and in this way they contribute to the overall fluctuations of the cellular phenotype.

5. Energy for Change

Stability of a dynamic system requires energy that compensates for the continuous stochastic fluctuations. However, changing a dynamic equilibrium into another, a gene expression profile into another is also energy dependent. Activating repressed genes and repressing active ones is achieved by the cell through changing the chromatin around these genes. The transition between the repressive and permissive configurations depends essentially on the epigenetic modifications. The dynamic nature of these modifications implies that both the maintenance of the chromatin structure and the transition between the different forms are energy-dissipating processes. Indeed, the substrates used for epigenetic modification are all small molecular intermediates of the core energy metabolic pathways (for comprehensive reviews, *see* ref. 33. I apologize for not citing original papers). For example, acetylation of the histones and many other nuclear proteins is achieved using acetyl-CoA as substrate. Acetyl-CoA is probably one of the most important hubs in the metabolic network of the cell. It is directly generated from pyruvate, the end product of the glycolysis. Acetyl-CoA is either converted into citrate in the first step of the Kebs-cycle or used as a starting point for the biosynthesis of lipids and indirectly of almost any other types of macromolecules in the cell. The levels of Acetyl-CoA fluctuate widely depending on the metabolic flux and directly influence the level of acetylation of the chromatin components in the nucleus. The same is true for all other epigenetic modifications. Methylation is dependent on S-adenosyl-methyonin, a methyl donor synthesized from methionine, an essential amino acid and ATP. Demethylation reactions use a-ketoglutarate, a key Krebs cycle intermediate, poly-ADP ribosylation is dependent on NAD+ as a substrate, phosphorylation requires ATP, etc. The direct substrate level metabolic link between energy production and chromatin structure is more than obvious. In general, the rate of enzymatic reactions is essentially dependent on the substrate concentration. The intracellular concentration of the key metabolic substrates is indeed a major determinant of the epigenetic reaction rates [34].

Energy production depends on a network of red-ox reactions. The concentration of the intermediate metabolites and final high-energy-carrying molecules, in turn, is determined by the flux and activity of the whole metabolic network, which is itself dependent on the nature and availability of electron donors and acceptors. Electron donors are essentially nutriments taken up from the cellular environment and to lesser extent the cell's own reserves. The electron transfer between them is a multistep process and involves

intermediate electron transporters (NAD, NADP, FAD). These electron transporters provide electrons to all other electron transfer reactions including biosynthesis. In the presence of a sufficient external carbon source as an electron donor and oxygen, the oxidation into H_2O and CO_2 will be dominant and the ATP production and concentration of reduced electron transporters, as NAD+ and NADP+ will be high. When oxygen is not available, glycolysis will dominate, biosynthesis will eliminate the oxidation by O_2 as a final electron acceptor. The concentration Acetyl-CoA and Krebs-cycle intermediates will be relatively high. Therefore, the nature of metabolic regimes and the transition between them can modulate the concentration of key metabolic substrates for epigenetic reactions. This, in turn, increases or decreases the rate of the corresponding epigenetic reactions and, as a corollary, modulates the frequency and amplitude of the gene expression fluctuations.

The direct dependence of the stabilizing mechanisms on the energy-producing metabolic flux also implies that the cellular environment can impact the rapidity and extent of gene expression fluctuations. In fact, the metabolic flux in the cell is dependent primarily on the external substrates as electron donors. The most efficient terminal electron acceptor O_2 is also provided by the cell's immediate environment. The oxygen concentration usually varies significantly within the tissues as a function of the physical distance to the source (blood vessels) and the local demand. In this way, the cellular microenvironment is of primary importance in determining how a cell can generate energy and impact the transcriptional fluctuations through the epigenetic modifications. Each cell is exposed by a unique microenvironment that is essentially composed by other cells. This may explain why cells in the same tissue are so different and create complementarity and interdependence between neighbors. The cells localized close to the nutrient and oxygen sources use different metabolic pathways than those cells that are located more distantly and exposed to a microenvironment composed by the resources not used by their neighbors and by their secreted metabolites. A tissue or a cell community can be considered analogous to an ecosystem and the interaction between the cells as a Darwinian selective pressure. It has been proposed that cell differentiation is a process analogous to Darwinian evolution [35, 36]. Stochastic fluctuations of gene expression in the cell generate spontaneously phenotypic fluctuations. Interactions between the cells and their microenvironment act as a selective force that can stabilize some phenotypes only. Each cell fluctuates until it can express the characteristics that allow using the available resources and maintaining a metabolic flux that produces the necessary energy in the system of interdependent individual cells placed in a given environment.

6. Conclusion

It is important to keep in mind that living cells are out-of-equilibrium open thermodynamic systems that constantly dissipate energy. The minimal energy flux required to maintain the dynamic equilibrium is a sine qua non-condition for the living state and expected to be the organizing force of the living matter [37]. This theoretical conclusion led to the proposal that the true driving force of cell differentiation is the requirement to continuously dissipate energy produced by the metabolic flux [38, 39]. Chromatin stabilizing/destabilizing epigenetic mechanisms appear as a major evolved molecular mechanism that links the environment to the fluctuations of the genome function [38]. These mechanisms transform metabolic fluctuations into gene expression fluctuations ensuring the generation of new phenotypic variants until the metabolic adaptation is achieved.

The Darwinian model of cell differentiation conceptualizes the whole process of ontogenesis using the same concepts of variation/selection as in the theory of evolution. Phenotype variations are generated by the stochastic fluctuations of the molecular processes that maintain the continuous fluctuations of gene expression levels [10, 40–42]. The necessity to maintain the permanent energy flux required for the vital cellular processes represents a strong selective pressure continuously acting on the fluctuating phenotype. Suboptimal metabolic flux acts by increasing the fluctuations; return to the steady state decreases them.

The metabolic pressure canalizes the cell phenotype through the direct substrate level link between the core energy metabolism and the chromatin modifying epigenetic mechanisms. The same epigenetic mechanisms also ensure the conservation of gene expression profiles after cell divisions.

Redefining the conceptual framework of cell differentiation by considering variation as a central player leads to a unified theory that explains the emergence of different living forms at different time scales without making the distinction between an individual as a unit of evolution and its parts as units of ontogenesis. The two processes are expressions of the same principles [43].

Acknowledgments

I thank my colleagues, Alice Moussy, Daniel Stockholm, and Guillaume Corre, for the helpful discussions and the useful comments on the manuscript.
Financial support: EPHE, Genethon, Stochagene ANR grant n° BSV6 014 02.

References

1. Trapnell C (2015) Defining cell types and states with single-cell genomics. Genome Res 25(10):1491–1498. https://doi.org/10.1101/gr.190595.115

2. Merrell AJ, Stanger BZ (2016) Adult cell plasticity in vivo: de-differentiation and transdifferentiation are back in style. Nat Rev Mol Cell Biol 17(7):413–425. https://doi.org/10.1038/nrm.2016.24

3. Blau HM, Brazelton TR, Weimann JM (2001) The evolving concept of a stem cell: entity or function? Cell 105(7):829–841

4. Zipori D (2004) The nature of stem cells: state rather than entity. Nat Rev Genet 5(11):873–878. https://doi.org/10.1038/nrg1475

5. Wagner A, Regev A, Yosef N (2016) Revealing the vectors of cellular identity with single-cell genomics. Nat Biotechnol 34(11):1145–1160. https://doi.org/10.1038/nbt.3711

6. Paul F, Arkin Y, Giladi A, Jaitin DA, Kenigsberg E, Keren-Shaul H, Winter D, Lara-Astiaso D, Gury M, Weiner A, David E, Cohen N, Lauridsen FK, Haas S, Schlitzer A, Mildner A, Ginhoux F, Jung S, Trumpp A, Porse BT, Tanay A, Amit I (2015) Transcriptional heterogeneity and lineage commitment in myeloid progenitors. Cell 163(7):1663–1677. https://doi.org/10.1016/j.cell.2015.11.013

7. Blakeley P, Fogarty NM, del Valle I, Wamaitha SE, Hu TX, Elder K, Snell P, Christie L, Robson P, Niakan KK (2015) Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. Development 142(18):3151–3165. https://doi.org/10.1242/dev.123547

8. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA (2015) Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12(5):453–457. https://doi.org/10.1038/nmeth.3337

9. Yin Z, Sadok A, Sailem H, McCarthy A, Xia X, Li F, Garcia MA, Evans L, Barr AR, Perrimon N, Marshall CJ, Wong ST, Bakal C (2013) A screen for morphological complexity identifies regulators of switch-like transitions between discrete cell shapes. Nat Cell Biol 15(7):860–871.

https://doi.org/10.1038/ncb2764

10. Moussy A, Cosette J, Parmentier R, da Silva C, Corre G, Richard A, Gandrillon O, Stockholm D, Paldi A (2017) Integrated time-lapse and single-cell transcription studies highlight the variable and dynamic nature of human hematopoietic cell fate commitment. PLoS Biol 15(7):e2001867. https://doi.org/10.1371/journal.pbio.2001867

11. Stockholm D, Edom-Vovard F, Coutant S, Sanatine P, Yamagata Y, Corre G, Le Guillou L, Neildez-Nguyen TM, Paldi A (2010) Bistable cell fate specification as a result of stochastic fluctuations and collective spatial cell behaviour. PLoS One 5(12):e14441. https://doi.org/10.1371/journal.pone.0014441

12. Amir el AD, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, Shenfeld DK, Krishnaswamy S, Nolan GP, Pe'er D (2013) viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat Biotechnol 31(6):545–552. https://doi.org/10.1038/nbt.2594

13. Gut G, Tadmor MD, Pe'er D, Pelkmans L, Liberali P (2015) Trajectories of cell-cycle progression from fixed cell populations. Nat Methods 12(10):951–954. https://doi.org/10.1038/nmeth.3545

14. Setty M, Tadmor MD, Reich-Zeliger S, Angel O, Salame TM, Kathail P, Choi K, Bendall S, Friedman N, Pe'er D (2016) Wishbone identifies bifurcating developmental trajectories from single-cell data. Nat Biotechnol 34(6):637–645. https://doi.org/10.1038/nbt.3569

15. Ezer D, Moignard V, Gottgens B, Adryan B (2016) Determining physical mechanisms of gene expression regulation from single cell gene expression data. PLoS Comput Biol 12(8):e1005072. https://doi.org/10.1371/journal.pcbi.1005072

16. Huang S (2009) Non-genetic heterogeneity of cells in development: more than just noise. Development 136(23):3853–3862. https://doi.org/10.1242/dev.035139

17. Huang S (2012) The molecular and mathematical basis of Waddington's epigenetic landscape: a framework for post-Darwinian biology? Bioessays 34(2):149–157. https://doi.org/10.1002/bies.201100031

18. Furusawa C, Kaneko K (2012) A dynamical-systems view of stem cell biology. Science 338(6104):215–217. https://doi.org/10.1126/science.1224311

19. Waddington CH (ed) (1957) The strategy of the genes. Allen & Unwin, Crows Nest

20. Hume DA (2000) Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. Blood 96(7):2323–2328

21. Ko MS (1991) A stochastic model for gene induction. J Theor Biol 153(2):181–194

22. Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297(5584):1183–1186. https://doi.org/10.1126/science.1070919

23. Balazsi G, van Oudenaarden A, Collins JJ (2011) Cellular decision making and biological noise: from microbes to mammals. Cell 144(6):910–925. https://doi.org/10.1016/j.cell.2011.01.030

24. Chen H, Larson DR (2016) What have single-molecule studies taught us about gene expression? Genes Dev 30(16):1796–1810. https://doi.org/10.1101/gad.281725.116

25. Larson DR, Singer RH, Zenklusen D (2009) A single molecule view of gene expression. Trends Cell Biol 19(11):630–637. https://doi.org/10.1016/j.tcb.2009.08.008

26. Corre G, Stockholm D, Arnaud O, Kaneko G, Vinuelas J, Yamagata Y, Neildez-Nguyen TM, Kupiec JJ, Beslon G, Gandrillon O, Paldi A (2014) Stochastic fluctuations and distributed control of gene expression impact cellular memory. PLoS One 9(12):e115574. https://doi.org/10.1371/journal.pone.0115574

27. Lestas I, Vinnicombe G, Paulsson J (2010) Fundamental limits on the suppression of molecular fluctuations. Nature 467(7312):174–178. https://doi.org/10.1038/nature09333

28. Schwanhausser B, Wolf J, Selbach M, Busse D (2013) Synthesis and degradation jointly determine the responsiveness of the cellular proteome. Bioessays 35(7):597–601. https://doi.org/10.1002/bies.201300017

29. Misteli T (2001) Protein dynamics: implications for nuclear architecture and gene expression. Science 291(5505):843–847

30. Phair RD, Scaffidi P, Elbi C, Vecerova J, Dey A, Ozato K, Brown DT, Hager G, Bustin M, Misteli T (2004) Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. Mol Cell Biol 24(14):6393–6402. https://doi.org/10.1128/MCB.24.14.6393-6402.2004

31. Turner BM (2012) The adjustable nucleosome: an epigenetic signaling module. Trends Genet 28(9):436–444. https://doi.org/10.1016/j.tig.2012.04.003

32. Dodd IB, Micheelsen MA, Sneppen K, Thon G (2007) Theoretical analysis of epigenetic cell memory by nucleosome modification. Cell 129(4):813–822. https://doi.org/10.1016/j.cell.2007.02.053

33. Cyr AR, Domann FE (2011) The redox basis of epigenetic modifications: from mechanisms to functional consequences. Antioxid Redox Signal 15(2):551–589. https://doi.org/10.1089/ars.2010.3492

34. Lu C, Thompson CB (2012) Metabolic regulation of epigenetics. Cell Metab 16(1):9–17. https://doi.org/10.1016/j.cmet.2012.06.001

35. Kupiec JJ (1996) A chance-selection model for cell differentiation. Cell Death Differ 3(4):385–390

36. Kupiec JJ (1997) A Darwinian theory for the origin of cellular differentiation. Mol Gen Genet 255(2):201–208

37. Lane N, Martin W (2010) The energetics of genome complexity. Nature 467(7318):929–934.

https://doi.org/10.1038/nature09486

38. Paldi A (2003) Stochastic gene expression during cell differentiation: order from disorder? Cell Mol Life Sci 60(9):1775–1778. https://doi.org/10.1007/s00018-003-23147-z

39. Paldi A (2012) What makes the cell differentiate? Prog Biophys Mol Biol 110(1):41–43. https://doi.org/10.1016/j.pbiomolbio.2012.04.003

40. Mojtahedi M, Skupin A, Zhou J, Castano IG, Leong-Quong RY, Chang H, Trachana K, Giuliani A, Huang S (2016) Cell fate decision as high-dimensional critical state transition. PLoS Biol 14(12):e2000640. https://doi.org/10.1371/journal.pbio.2000640

41. Richard A, Boullu L, Herbach U, Bonnafoux A, Morin V, Vallin E, Guillemin A, Papili Gao N, Gunawan R, Cosette J, Arnaud O, Kupiec J-J, Espinasse T, Gonin-Giraud S, Gandrillon O (2016) Single-cell-based analysis highlights a surge in cell-to-cell molecular variability preceding irreversible commitment in a differentiation process. PLoS Biol 14:e1002585. https://doi.org/10.1371/journal.pbio.1002585

42. Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, Hirche C, Lutz C, Buss EC, Nowak D, Boch T, Hofmann WK, Ho AD, Huber W, Trumpp A, Essers MA, Steinmetz LM (2017) Human haematopoietic stem cell lineage commitment is a continuous process. Nat Cell Biol 19(4):271–281. https://doi.org/10.1038/ncb3493

43. Kupiec J-J (2009) The origin of individuals. World Scientific, Hackensack, NJ