Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Systems Biology

How to address cellular heterogeneity by distribution biology

Niko Komin[1] and Alexander Skupin[1,2]

Abstract

Cellular heterogeneity is an immanent property of biological systems that covers very different aspects of life ranging from genetic diversity to cell-to-cell variability driven by stochastic molecular interactions, and noise induced cell differentiation. Here, we review recent developments in characterizing cellular heterogeneity by distributions and argue that understanding multicellular life requires the analysis of heterogeneity dynamics at single cell resolution by integrative approaches that combine methods from non-equilibrium statistical physics, information theory and omics biology.

Addresses

[1] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belval, Luxembourg

[2] University California San Diego, La Jolla, USA

Corresponding author: Skupin, Alexander (alexander.skupin@uni.lu)

Current Opinion in Systems Biology 2017, **3**:154–160

This review comes from a themed issue on **Mathematical modelling (2017)**

Edited by Martin Falcke

For a complete overview see the Issue and the Editorial

Available online 18 May 2017

http://dx.doi.org/10.1016/j.coisb.2017.05.010

Keywords

Cellular heterogeneity, Distribution biology, Cell fate dynamics, Multicellular organization, Multiscale modeling.

Introduction

Life is heterogeneous — at nearly all biological scales and levels [1]. Maybe the most obvious heterogeneity can be observed when looking at the different species that evolution has created. But heterogeneity is much further spread in living systems where individuals within one species exhibit unique properties and even cells of the same cell type within the same (multicellular) organism can possess a wide range of divergent characteristics.

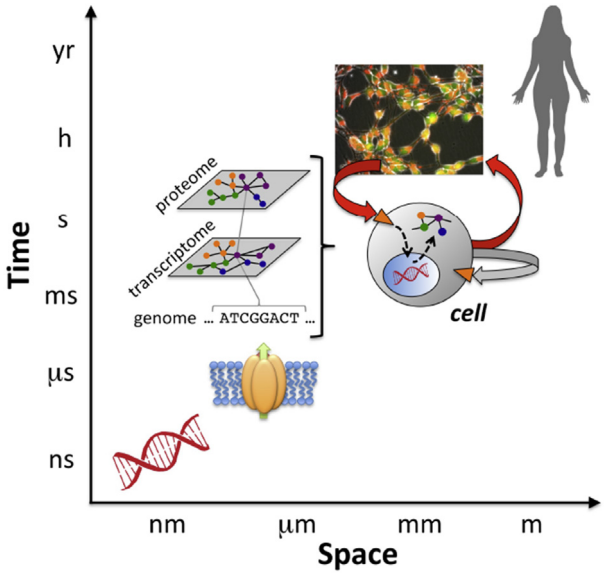
An essential function of heterogeneity is to ensure robustness of a biological system in fluctuating environments [2**]. Thus, the spread of phenotypic traits within a population of a species allows for a broader niche in which the population as a whole can survive and adapt to different conditions including competition for resources. This central mechanism of life has led to the general perception that "Nothing in Biology Makes Sense Except in the Light of Evolution" as stated by Theodosius Dobzhansky in his landmark essay in 1973 [3]. While Dobzhansky was mainly focusing on the diversity of different organisms and criticizing creationism, his conclusion may be also instructive to address currently open questions on the organization of multicellularity.

Besides the fundamental question how multicellularity arose, a central question is how an organism can give rise to all desired different cell types originating from the same genome in an coordinated manner [4,5]. Considering the concept of evolution as a general underlying mechanism of life with its two major components, mutation and the interplay of induction and selection, we may scale Dobzhansky's concept down to the level of cell populations. Thereby, mutations can be generalized to immanent heterogeneity and the interplay between induction and selection represents intra- and intercellular signaling. Hence, understanding (multicellular) life across its different scales relies on investigating *cellular nanoevolution* — that is the dynamics of cellular heterogeneity.

The origin of cellular heterogeneity (even within clonal populations) is the multiscale organization of life as depicted in Figure 1. On the smallest relevant scales, the stochastic nature of molecular interactions induces individual transcription profiles that are subsequently translated into heterogeneous cellular phenotypes. These rather randomly induced phenotypes are subsequently instructed and selected on the level of the cell population by intercellular signaling or cell—cell interactions leading to the coordinated generation of tissues, organs and organisms. This underlying noise-driven cellular heterogeneity is the mechanism to balance cellular robustness and adaptability [6].

Until recently, cellular heterogeneity was only accessible on rather low dimensional readouts such as specific protein abundance by antibody staining or since more recently by flux-cytometry analysis. These limited investigations hindered a more systematic approach to dissect underlying mechanisms. But with the recent developments of several single cell analysis methods [7–12], we are now able to characterize cellular heterogeneity in great detail. Despite these advancements, a systematic approach how to interpret and use the resulting high-dimensional data for identifying

Figure 1



Multiscale organization of life. Considering the different scales of multicellular cell fate from noisy gene expression to intercellular signaling enables adaptation and reliable morphogenesis. In this *cellular nano-evolution* process, the stochastic molecular dynamics is regulated on the level of the cell and cell population by signaling to generate ordered tissues and organisms.

biological principles of multicellular life is still lacking. Currently, methods from statistical physics are intensively discussed to be exploited but the underlying non-equilibrium dynamics and biological complexity of intra- and intercellular interactions make a direct application difficult [6,13*,14].

While previous reviews summarized potential functions of cellular heterogeneity from an experimental perspective [2**,15], we will review here more recent developments and how mathematical modeling and analysis can be used to reveal general functions and guiding principles from the central descriptor of heterogeneity – the (phenotype) trait distribution.

Characterization of cellular heterogeneity

Heterogeneity of cells can have different properties (Figure 2) and lead to multiple benefits for the population [2**,15,16]: it ensures robustness of the biological system to fluctuating environments and allows the population to adapt to a wider variety of environments (bet hedging); binary decisions on a cellular scale can yield a fractional or dose-dependent response; rare individuals can coordinate population behavior by emitting local signals; subpopulations can be primed for multiple cell fates.

Direct phenotypic heterogeneity

Cellular heterogeneity is observed relatively easily when it directly affects phenotypic traits such as cell survival times and cell division times [17,18] or size distributions (e.g. axon diameter variability [19]). The main driver of heterogeneity can be an underlying genetic variability as exemplified in Figure 2A by the biomass production of yeast strains. The observed trait variability within the two strains YO490 and YO512 is much smaller compared to the distribution of the progenies originating from a cross of these two parental strains. The non-trivial effects of the genetic recombination and the genome–environment interaction lead to a wide range of biomass production for the progenies that is not simply between the parental strains but exhibits more extreme traits.

But cellular heterogeneity does also occur independent of genetic diversity such as in clonal populations due to the multiscale organization depicted in Figure 1. Here, the stochastic nature of molecular interaction induces random transcription profiles that are subsequently modified in an environment dependent manner by intra- and intercellular signaling. A related and medically important example of such a coordinated heterogeneity is the epithelial-to-mesenchymal transition (EMT) and its counterpart MET as a mechanism for metastasis [20]. The underlying mechanism is that a cancerous epithelial cell can transdifferentiate into a mobile mesenchymal stem cell obeying multipotent properties which subsequently leaves the tumor, travels with the blood stream to distinct tissue sites, is then able to transform back into an epithelial cell due to its adult stem cell character and finally initiate a secondary tumor.

Model cell lines like clonal HMLER cells where e.g. mesenchymal cell identity can be determined by flux cytometry using cell-surface markers such as CD44 [21] allow to study the transition experimentally. The HMLER population exhibits a heterogeneous steady-state distribution as shown in Figure 2B where a unimodal phenotype distribution (blue) generated by cell sorting is relaxing towards a stable bimodal distribution (gray) that represents a mixture of epithelial (low CD44) and mesenchymal (high CD44) cells.

Mathematical EMT/MET-models can be devised by systems of differential equations. In Ref. [22], the authors showed how the regulatory network governing EM plasticity contains a ternary switch which introduces a hybrid E/M phenotype. It is characterized by migratory as well as adhesive properties – a state of collective migration. In a next step, the authors incorporated transcription factors GRHL2 and OVOL into the model and observed that they can stabilize the hybrid phenotype [23]. This finding potentially allows predicting