Mitochondrial control of gene expression and extrinsic apoptosis

Juan Díaz-Colunga

Madrid, 2019

Mitochondrial control of gene expression and extrinsic apoptosis

Author
Juan Díaz-Colunga

Supervisors
Francisco J. Iborra and Raúl Guantes

PhD program
Condensed Matter Physics, Nanoscience and Biophysics
Universidad Autónoma de Madrid

Contents

Abstract

Even under homogeneous environmental conditions, cells with identical genotypes can display significant variations at the phenotypic level, that is, differences in size, morphology and internal state. This is a result of the genetic and signaling circuits that control cellular functions being subject to fluctuations in the levels of their components. Cell-to-cell variability is regarded as an essential agent in many key cellular activities such as development, differentiation, evolution, virus infection or cell death, and has been shown to serve a biological function in many cases. Thus, tracing back its sources is central to understand the behaviors of individual cells and ultimately act on them. In this work we use a combination of experimental, mathematical and computational tools to investigate the role of mitochondria (the main energy provider in eukaryotic cells) on the generation of gene expression noise. Gene expression is highly energy demanding and in turn determines phenotype, pointed as the cause of variable individual cellular responses in several processes, notably apoptosis. Heterogeneity in apoptotic outcomes is particularly relevant as it poses the main cause of tumor resistance to chemotherapy. Our results highlight the importance of mitochondria as a global modulator of gene expression, but also reveal its role regulating complex, non-linear processes like alternative splicing. We found that this control of gene expression is specially important in the apoptotic signaling pathway: mitochondria exhibit the power to discriminate apoptotic fates at the single cell level, making it a good candidate for a biomarker of the susceptibility of cancer cells to death-inducing treatments.

Resumen

Incluso en condiciones ambientales homogéneas, células con genotipos idénticos pueden presentar variaciones significativas a nivel fenotípico, es decir, diferencias en tamaño, morfología y estado interno. Esto es el resultado de los circutos genéticos y de señalización que controlan la función celular estando sometidos a fluctuaciones en los niveles de sus componentes. La variabilidad de célula a célula es considerada un agente esencial en multitud de actividades celulares como el desarrollo, diferenciación, evolución, infección vírica o la muerte celular. Por tanto, identificar sus fuentes es central para entender los comportamientos de células individuales y en última instancia actuar sobre ellos. En este trabajo utilizamos una combinación de herramientas experimentales, matemáticas y computacionales para investigar el papel de la mitocondria (principal proveedor de energía en células eucariotas) en la generación de ruido en expresión genética. La expresión genética es costosa energéticamente y determina el fenotipo, señalado en muchos procesos como causa de la variabilidad en las respuestas de células individuales, en particular en el de apoptosis. La heterogeneidad en los resultados apoptóticos es particularmente relevante porque supone la principal causa de resistencia de tumores a quimioterapia. Nuestros resultados resaltan la importancia de la mitocondria como modulador global de la expresión génica, pero también revelan su papel regulando procesos complejos y no lineales como el splicing alternativo. Encontramos que este control de la expresión génica es especialmente importante en la ruta de señalización apoptótica: la mitocondria muestra el potencial de discriminar destinos apoptóticos a nivel de célula única, convirtiéndola en un buen candidato a biomarcador de la susceptibilidad de células cancerígenas frente a tratamientos inductores de apoptosis.

1

Introduction

There are several aspects that motivate this work. First, the emergence of *high-throughput* technologies, along with the development of powerful statistical tools, has allowed for the systematic assessment of large sets of molecules and made quantitative information more accessible than ever for biologists. This has fueled a paradigm shift in biological research. Second, it is becoming increasingly clear that cell-to-cell differences play a critical role in many biological processes, but the origins of this differences are still not completely understood. Third, the realization that many of the factors that cause cell-to-cell variability are affected by energetic constraints has put the focus in the main source of energy for eukaryotic cells: mitochondria.

In this chapter, we will present the systems biology framework and discuss why quantitative modeling is central to it, reviewing some of the most important mathematical tools used to build models. We will also introduce cell-to-cell variability and some of its sources and implications. Finally, we will discuss how understanding and characterizing mitochondrial variability can be key to trace back the origins of cell-to-cell heterogeneity.

1.1 The Systems Biology approach

Biological networks

All biological entities have interactions with one another at many different scales, from the molecules in a cell to the species in an ecosystem. Biological systems are often represented as networks formed by nodes and links between them. The identity of these nodes and interactions is variable: biological networks exist at the genetic, metabolic, neurological or ecological levels to name a few.[1] In the context of cellular biology, networks are based on different biochemical processes serving a variety of purposes:

Metabolic networks are defined by the enzymes converting substrates into products and by the metabolites converted by such enzymes, as well as the interactions between them. Through metabolism, cells acquire the energy and materials needed for survival and reproduction.

Transcriptional networks represent the regulation of genes by transcription factors (TF). From a biochemical point of view, the structure of these networks is determined by the TF binding sites. The operation of these networks shapes the cell's gene expression landscape.

Protein-protein interaction networks represent physical interactions such as binding and complex formation, molecular modifications (mainly phosphorylation) and activation or inhibition of biochemical reactions. These networks allow cells to process information enabling, for example, rapid stress responses or intercellular communication.

Despite cellular components being highly interconnected, reductionism has dominated biological research for a long time. For many decades in the latter half of the 20th century, studies in molecular and cellular biology were devoted to the generation of information about the chemical composition and functions of individual cell components, that is, specific network nodes. The emergence and fast development of the *-omics* fields (genomics, transcriptomics, proteomics, metabolomics...) in the 21st century has greatly accelerated this process.

The rise of the *-omics* fields

The suffix *-omics* is used to designate many of the emerging fields of data-rich biology. These terms refer to a global, un-targeted assessment of large sets of molecules, rather than a study of each one of them individually. For instance, *genomics* (the first *-omics* discipline to appear) focuses on whole genomes, as opposed to *genetics* which aims to understand the role of single genes.

The increasing availability of quantitative data is due to the development of technological advances that enable fast and cheap analysis of large ensembles of molecules. These so-called *high-throughput* technologies allow for the automation and parallelization of classic cell biology methods (such as expression arrays, developed in the 90's) by incorporating techniques from chemistry, optics or image analysis among other fields. Alongside these technological advances, rigorous statistical tools and computational methods are being developed to facilitate data storage, classification, mining and analysis.

Global gene expression assays have made it possible to monitor the transcription levels of tens of thousands of genes simultaneously.[2] *Transcriptomics* is the field devoted to the examination of RNA levels genome-wide. Such examination can be qualitative (to explore which RNAs are or are not present under certain conditions) or quantitative

(how much of each transcript is expressed). RNA-seq experiments use sequencing tools to determine the amount of each type of RNA in a biological sample. They facilitate the study of, for instance, alternative splicing or changes in gene expression over time. An interesting example of the power of RNA-seq are non-coding RNAs: while typically less than 5% of a mammalian genome encodes proteins, ~80% of it can be transcribed. Modern RNA-seq studies have not just allowed for a better understanding of the complexity of the protein-coding genome, but have also revealed several important roles of non-coding RNAs.[3]

High-throughput technologies have given us the ability to produce detailed (and constantly developing) lists of biological components and their properties. But as useful as these lists can be, a more integrative approach is required to fully understand the complex behaviors that arise when those individual components work in conjunction with one another. This realization has forced a paradigm shift in biology.[4–7]

A new paradigm in biological research

Today's research is focusing not only on the components themselves but on their interactions and the states that emerge from the assembly of individual nodes into connected networks.[8–11] It is now clear that a biological function can rarely be attributed to a single molecule, instead, biological properties typically result from complex interactions between constituents like proteins, DNA or RNA. These properties are sometimes called *emergent*, since they arise from the whole despite not being associated to the individual parts.

Such a global view can give us a general impression of the performance of biological networks. Still, further progress requires that we descend in scale, since signaling pathways are typically divided into smaller, more specialized modules (often located in different regions of the cell). At the same time, the study of cells as systems provides a framework where the vast amount of data available can be integrated and displayed, and ultimately used to build computational models that bring together the many pieces of information.

1.2 Modeling in Cellular Biology

1.2.1 Modeling: why and how?

The aim of models in cellular biology

A model can be defined as an abstract representation of a set of objects or processes that explains some of their features and behaviors.[†] In a broad sense, models have a long

[†]In biology, the term *model* is also used to refer to a species that is suitable for experimentation, for example mouse or zebrafish.

history in biology. Our theories and hypotheses about biological systems are, in the end, nothing but just models. The most common form of biological model is the *word model*: many observations give rise to an abstract, wider picture that is simply stated in plain language.[12] The theory of evolution is an example of a word model.

Yet the use of *quantitative mathematical models* (which constitute an integral part of systems biology) has lied outside mainstream biological research for many years. This can be due to the mathematical approach tending to be biased toward simple, general explanations that are not necessarily suitable to describe nature, where special cases are abundant as evolution does not always select for simple solutions.[13] This makes it challenging to develop models that are faithful to biology while remaining settled on clearly interpretable mathematical principles.

Systems biology models are fundamentally based on physical laws (such as the thermodynamics of biochemical reactions) that justify their general form. In addition, experimental evidence and biochemical knowledge is needed to specifically shape a computational model. Many general biochemical *first principles* are well established. A prominent example is the so-called *central dogma* of molecular biology, that can be summarized as follows: genes code for mRNA, mRNA serves as template for proteins, and proteins perform specific tasks in the cell. However, the biochemistry of individual molecules and systems is usually unclear. Experiments can provide insight into these individual processes, while models usually aim to combine sets of experimentally tested hypotheses into a bigger coherent picture: the behavior of a complex system can rarely be understood just from knowledge of its parts.

In the context of cellular biology, the modeling of biochemical reactions involving cell components (genes, mRNAs, proteins...) can help to unveil their internal nature and dynamics. But modeling a biological system is not only useful to understand it in depth: mathematical modeling and computer simulations should reliably predict the behavior of a network, or at least of those aspects that are supposed to be covered by the model.

Cells as biochemical reactors

Generally speaking, a cell is an integrated device made of several thousands of types of interacting molecules. Out of those molecules, proteins are the nanometer-size machines that carry out specific tasks with great precision. Cells require the action of different proteins depending on the situation they encounter, for which they continuously monitor their environment and calculate the amount at which each type of protein is needed. This processing is done by signaling and transcriptional networks, the first ones eliciting rapid responses and the second ones operating more slowly, on a timescale that can be as long as the cell's generation time.[14] In both cases, networks are essentially collections of biomolecular species coupled by chemical interactions, meaning that the cell can be seen as a *biochemical reactor*.

Biochemical reactions are catalyzed by enzymes. These are specific proteins that

have a catalytic center and remain unchanged by the reaction. They often function in combination with cofactors and are typically very specific. Even complex biochemical processes are usually described as a succession of simpler, reversible binding steps and irreversible catalytic steps, each of them constituting an elementary reaction. Before the emergence of mathematical modeling in network biology, chemical reactors were being studied in different contexts. In vitro enzymatic reactions under controlled, well-mixed conditions were already of interest more than a century ago, when Michaelis and Menten developed a model for the rate of an irreversible one-substrate reaction.[15] Their work is now a fundamental part of biochemistry.

Model classification

Models are typically classified with respect to a set of criteria:

- **Qualitative** models specify the interactions among model elements, while a **quantitative** model assigns numerical values to the elements and their interactions.
- In **deterministic** models, the system evolution through all its states can be predicted from the knowledge of its current configuration. **Stochastic** models, instead, provide a probability distribution for these successive states.
- **Discrete** models take values for time, speace and state from a discrete set. In **continuous** models, values belong to a continuum.

Just like a biological object can be investigated with different experimental methods, it can be described with different types of models. At the same time, a mathematical formalism can be applied in more than one context: statistical network analysis, for example, is used to study cellular signaling pathways, transcription networks, the circuitry of nerve cells or food webs. The choice of a specific model or algorithm to describe a biological process depends on the problem and the purpose. In general, modeling should reflect the essential properties of the system studied. Different approaches provide insights into different aspects of the same problem.

Some examples of model types are boolean networks,[16] petri nets[17] or models based on ordinary differential equations (ODEs). The use of the last ones is constricted by the (usually scarce) knowledge of mechanistic details and kinetic parameters. The precision with which parameters can be identified is determined by two factors: the experimental error and the relationship between experimental observables and model parameters (*structural identifiability*).[18] However, modeling the kinetics of biochemical reactions using quantitative, continuous models expressed in terms of ODEs is widely extended because of the valuable information they provide about the network dynamics: equilibrium states and possible transitions between them, propagation of signals through time, etc.

1.2.2 Deterministic kinetic models

Networks in cellular biology are formed by many molecules that react with one another. These molecules do not usually participate in just one reaction, which can confer a high degree of complexity on these networks. To build models that represent biological processes, we first need tools to transform complex networks into mathematically tractable objects, e.g. sets of differential equations. Even though biochemical reactions are discrete by nature (a finite number of reactants produce a finite number of products), for large numbers of molecules it is justified to use averages and formulate continuous processes based on rates.

The mass action law

To simulate the elementary reactions forming a network, it is common to use models based on *mass action kinetics*. The mass action law was introduced in the 19th century by Guldberg and Waage.[19] It states that the rate of a given reaction (the speed at which reactants are converted into products) is proportional to the product of the concentrations of the reactants to the power of their molecularity, that is, the number in which they enter the reaction. Microscopically, this means the reaction rate is proportional to the probability of a collision of the reactants.

For example, for the following reaction involving three molecular species A, B and C:

$$2\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{C}$$

the reaction rate would be proportional to $[A]^2\,[B]$, where brackets indicate concentrations. The proportionality factor is called the *kinetic constant* and has units of $\text{time}^{-1} \cdot \text{concentration}^n$, where $n = 1, 0, -1, ...$ depends on the order of the reaction. If we call this constant k, the reaction rate in our example would be equal to $k\,[A]^2\,[B]$. The value of the kinetic constant is reaction-specific and, in general, it also depends on external parameters such as the temperature.

Let us now consider the corresponding reverse reaction as well, with a kinetic constant k_r:

$$2\mathrm{A} + \mathrm{B} \underset{k_r}{\overset{k}{\rightleftharpoons}} \mathrm{C}$$

Here, the rate for the reverse reaction would be $k_r\,[C]$. The overall reaction rate is given by the subtraction of the forward and backward rates, that is, $k\,[A]^2\,[B] - k_r\,[C]$. If it is positive, the reaction will move forward and vice-versa. The reaction rate quantifies the speed at which species C is produced (consumed if negative), thus we can write the following equation for the dynamics of $[C]$:

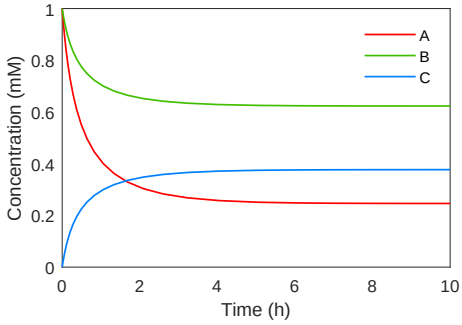


Fig. 1.1. Simple reaction dynamics modeled using mass action kinetics. Solution of equations 1.1 and 1.2, for $k = 1\text{mM}^{-1}\text{h}^{-1}$ and $k_r = 0.1\text{h}^{-1}$. Initial conditions: $[A]_0 = 1\text{mM}$, $[B]_0 = 1\text{mM}$ and $[C]_0 = 0$. Equilibrium is reached when $[A] = 0.25\text{mM}$, $[B] = 0.62\text{mM}$ and $[C] = 0.38\text{mM}$.

$$\frac{\mathrm{d}}{\mathrm{d}t}[C] = k\,[A]^2\,[B] - k_r\,[C] \tag{1.1}$$

This is also the velocity at which A and B are being *consumed* when the reaction moves forward, so the *negative* of the reaction rate will give us the dynamics of $[A]$ and $[B]$. Note that one reaction requires two molecules of A, which introduces a factor of 2 in the equation for $[A]$.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}[A] &= -2k\,[A]^2\,[B] + 2k_r\,[C]\\ \frac{\mathrm{d}}{\mathrm{d}t}[B] &= -k\,[A]^2\,[B] + k_r\,[C]\end{aligned} \tag{1.2}$$

Equations 1.1 and 1.2 form a set that fully describes the dynamics of the system. At equilibrium, concentrations stabilize and thus the expressions equal zero. This is interpreted as the rates of the forward and backward reactions becoming the same, making the system reach a steady state when

$$k\,[A]_{eq}^2\,[B]_{eq} = k_r\,[C]_{eq} \tag{1.3}$$

where the *eq* subindex indicates equilibrium. Figure 1.1 shows an example of the dynamics of this system for a specific choice of parameters and initial conditions.

Michaelis-Menten dynamics

Michaelis and Menten[15] and Briggs and Haldane[20] were the first to study enzymatic reactions, which are particularly common in molecular biology. These reactions consist of the transformation of a substrate (S) into a product (P) through the action of an enzyme (E).

$$\mathrm{S} + \mathrm{E} \underset{k_r}{\overset{k}{\rightleftharpoons}} \mathrm{S:E} \xrightarrow{k_c} \mathrm{P} + \mathrm{E}$$

The transformation has a first binding step in which the enzyme and the substrate form a complex (S:E) that can either dissociate back into the initial species or undergo a second catalytic step to produce a molecule of the product, the enzyme remaining unchanged. Using the mass action law, we can derive the following set of equations:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}[S] &= -k[S][E] + k_r[S:E] \\
\frac{\mathrm{d}}{\mathrm{d}t}[E] &= -k[S][E] + k_r[S:E] + k_c[S:E] \\
\frac{\mathrm{d}}{\mathrm{d}t}[S:E] &= k[S][E] - k_r[S:E] - k_c[S:E] \\
\frac{\mathrm{d}}{\mathrm{d}t}[P] &= k_c[S:E]
\end{aligned} \tag{1.4}$$

Even though the system in 1.4 has no analytical solution, there are some assumptions that can be made to simplify it. If the binding/unbinding steps are much faster than the catalytic one (in terms of the kinetic rates, this means $k, k_r \gg k_c$) there is a so-called *quasi-equilibrium* state between the free enzyme and the enzyme-substrate complex. In addition, $[E]_{total} \equiv [E] + [S:E]$ must remain constant (assuming no production of new molecules or degradation of existing ones), which yields the following expression for the P production rate:

$$\frac{\mathrm{d}}{\mathrm{d}t}[P] = -\frac{\mathrm{d}}{\mathrm{d}t}[S] = V_{max}\frac{[S]}{[S] + K_m} \tag{1.5}$$

where $K_m \equiv (k_r + k_c)/k$ is called the *Michaelis constant*, and $V_{max} \equiv k_c[E]_{total}$ is the maximum reaction rate, reached in the limit of $[S] \gg K_m$.

1.2.3 Probabilistic formulation of reaction kinetics

Biochemical reactions are triggered by random collisions between molecules. If these events happen a large amount of times per generation, this intrinsic randomness (*noise*) can be averaged out. Under these circumstances, deterministic kinetic models usually provide a good description of the dynamics of the system. However, many key cellular reactions occur infrequently enough for noise to become non-negligible. Additionally, fluctuations arising from these infrequent events can affect the rates of other downstream reactions, propagating through biological networks. From an experimental point of view, this idea is supported by the fact that the practical totality of studies measuring single-cell concentrations of proteins report significant variation from cell to cell: gene expression is stochastic by nature, as it depends on the random association and dissociation of transcription factors to DNA.

The master equation

In physics and other related fields, the concept of a *master equation* refers to a phenomenological set of differential equations that describe the time evolution of the probability of a system to occupy one of all its possible states.[21] Switching between states is determined by a transition rate matrix.

The state of a network formed by a subset of cellular components and reactions is determined by the abundances of said components. A molecular network consisting of L interacting species (proteins in a signaling network, metabolites in a metabolic network, etc.) can be in one of many mutually exclusive discrete states. The vector $\mathbf{n} = (n_1, n_2, \ldots, n_L)$ characterizes these states, with n_i being the number of molecules of the i-th species. Transitions between states can occur through time: as new molecules are produced, interact with each other and degrade, the values for n_i change. The probability of finding the system in a configuration $\mathbf{n}$ at time t is defined as

$$p(\mathbf{n}, t) \geq 0 \tag{1.6}$$

Additionally, the conditional probability to find the system in the state $\mathbf{n}_2$ at time t_2, granted that its previous state was $\mathbf{n}_1$ at time t_1 (with $t_2 > t_1$) is

$$p(\mathbf{n}_2, t_2 \mid \mathbf{n}_1, t_1) \geq 0 \tag{1.7}$$

Of course, one out of all the states is sure to be found at any time. Analogously, given that the system was in a configuration $\mathbf{n}_0$ at time t_0, it it is certain to have moved to another one of its possible states $\mathbf{n}$ at time $t > t_0$ (a valid particular case is $\mathbf{n} = \mathbf{n}_0$, meaning the system can remain in the same state during the interval $[t_0, t]$). Thus the normalization conditions

$$\sum_{\mathbf{n}} p(\mathbf{n}, t) = 1 \tag{1.8}$$

$$\sum_{\mathbf{n}} p(\mathbf{n}, t \mid \mathbf{n}_0, t_0) = 1 \tag{1.9}$$

must be satisfied for every t, with $\mathbf{n}$ extending to all possible configurations.

The joint probability of finding the system in the state $\mathbf{n}_1$ at t_1, followed by the state $\mathbf{n}_2$ at t_2, denoted by $p(\mathbf{n}_2, t_2; \mathbf{n}_1, t_1)$, is

$$p(\mathbf{n}_2, t_2; \mathbf{n}_1, t_1) = p(\mathbf{n}_2, t_2 \mid \mathbf{n}_1, t_1)\, p(\mathbf{n}_1, t_1) \tag{1.10}$$

The so-called *Markov assumption* postulates that the conditional probability in equation 1.7 depends exclusively on the initial state $\mathbf{n}_1$, but not on states prior to it: the system keeps no *historic memory*. The cases where the Markov assumption fails can become very complicated, but fortunately it holds most of the times at least as a good approximation. Under the Markov assumption, equation 1.10 can be generalized to

$$p\left(\mathbf{n}_2, t_2\right) = \sum_{\mathbf{n}_1} p\left(\mathbf{n}_2, t_2 \mid \mathbf{n}_1, t_1\right) p\left(\mathbf{n}_1, t_1\right) \tag{1.11}$$

for any given t_1 and t_2. Equation 1.11 shows that the probability $p\left(\mathbf{n}_1, t_1\right)$ is propagated in the time interval $[t_1, t_2]$ by the conditional probability $p\left(\mathbf{n}_2, t_2 \mid \mathbf{n}_1, t_1\right)$, named the *propagator*. Simply put, this means that knowing the conditional probability in equation 1.7 gives all the information about the whole dynamics of the system.

The master equation is a differential equation in time for the propagator. Let us consider $t_1 = t$ and $t_2 = t + \tau$, being τ an infinitesimally short time step. Equation 1.11 then becomes

$$p\left(\mathbf{n}_2, t + \tau\right) = \sum_{\mathbf{n}_1} p\left(\mathbf{n}_2, t + \tau \mid \mathbf{n}_1, t\right) p\left(\mathbf{n}_1, t\right) \tag{1.12}$$

Expanding the propagator in a Taylor series[†] with respect to the variable $t_2 = t + \tau$ at $t_2 = t$ (which is equivalent to the limit of $\tau \to 0$) yields

$$\begin{aligned} p\left(\mathbf{n}_2, t + \tau \mid \mathbf{n}_1, t\right) = p\left(\mathbf{n}_2, t \mid \mathbf{n}_1, t\right) \\ + \tau \left.\frac{\partial p\left(\mathbf{n}_2, t_2 \mid \mathbf{n}_1, t\right)}{\partial t_2}\right|_{t_2 = t} \\ + O\left(\tau^2\right) \end{aligned} \tag{1.13}$$

where $O\left(\tau^2\right)$ represents a sum of higher order terms that can be neglected for sufficiently small values of τ. As for the term $p\left(\mathbf{n}_2, t \mid \mathbf{n}_1, t\right)$ in the right side of equation 1.13, the fact that the system's states are mutually exclusive means that given the configuration $\mathbf{n}_1$ at time t, there is no chance (probability equals zero) to find a different configuration $\mathbf{n}_2$ at that same time t. Mathematically:

$$p\left(\mathbf{n}_2, t \mid \mathbf{n}_1, t\right) = \begin{cases} 1 & \text{if } \mathbf{n}_1 = \mathbf{n}_2 \\ 0 & \text{if } \mathbf{n}_1 \neq \mathbf{n}_2 \end{cases} \tag{1.14}$$

It is possible to define a *probability transition rate* from the state $\mathbf{n}_1$ to the state $\mathbf{n}_2$, $w(\mathbf{n}_2, \mathbf{n}_1)$, as

$$w(\mathbf{n}_2, \mathbf{n}_1) = \left.\frac{\partial p\left(\mathbf{n}_2, t_2 \mid \mathbf{n}_1, t\right)}{\partial t_2}\right|_{t_2 = t} \geq 0 \tag{1.15}$$

[†]The Taylor series of an infinitely differentiable function $f(x)$ with respect to the variable x at a value x_0 is

$$f(x)|_{x=x_0} = \sum_{k=0}^{\infty} \frac{1}{k!} \left(x - x_0\right)^k \left.\frac{\partial^k f(x)}{\partial x^k}\right|_{x=x_0}$$

This is a rate with dimensions of time^{-1} (transitions per unit time) and must not be confused with a probability.

If $\mathbf{n}_2 \neq \mathbf{n}_1$, equation 1.13 becomes

$$p\left(\mathbf{n}_2, t+\tau \mid \mathbf{n}_1, t\right) = \tau\, w(\mathbf{n}_2, \mathbf{n}_1) \tag{1.16}$$

while if $\mathbf{n}_2 = \mathbf{n}_1$ it reads

$$p\left(\mathbf{n}_2, t+\tau \mid \mathbf{n}_1, t\right) = 1 - \tau \sum_{\mathbf{n}_1 \neq \mathbf{n}_2} w(\mathbf{n}_2, \mathbf{n}_1) \tag{1.17}$$

where the sum covers all possible states $\mathbf{n}_1$ with the exception of $\mathbf{n}_1 = \mathbf{n}_2$. To reach expression 1.17, equation 1.9 must be differentiated with respect to t, then the rate $w\left(\mathbf{n}_2 = \mathbf{n}_1, \mathbf{n}_1\right)$ can be cleared and used in 1.13.

Let us take the limit $\tau \to 0$. Considering that

$$\frac{\mathrm{d}}{\mathrm{d}t} p\left(\mathbf{n}, t\right) = \lim_{\tau \to 0} \frac{p\left(\mathbf{n}, t+\tau\right) - p\left(\mathbf{n}, t\right)}{\tau} \tag{1.18}$$

and conveniently arranging the result of inserting expressions 1.16 and 1.17 in 1.12, the most general form of the master equation is obtained:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} p\left(\mathbf{n}, t\right) = & \sum_{\mathbf{m}} w\left(\mathbf{n}, \mathbf{m}\right) p\left(\mathbf{m}, t\right) \\ & - \sum_{\mathbf{m}} w\left(\mathbf{m}, \mathbf{n}\right) p\left(\mathbf{n}, t\right) \end{aligned} \tag{1.19}$$

Even if the mathematical formulation can be obscure at first sight, the master equation has an intuitive interpretation. The change per unit time of the probability of the state $\mathbf{n}$ is given by the sum of two terms with opposite effects (and thus opposite signs). The first term on the right side of 1.19 quantifies the probability flux from all other states $\mathbf{m}$ into the state $\mathbf{n}$. On the other hand, there is a probability flux out of the state $\mathbf{n}$ into all other states $\mathbf{m}$, given by the second term in 1.19.

The master equation is the most detailed mathematical description about the time dynamics of a system under conditions of restricted information. To obtain specific solutions, knowledge about the system is required in the form of probability transition rates (w). In the end, it is these rates that contain all the information about the process.

A simple case study

Consider a very simple biochemical system in which a molecule A is being produced at a constant rate k (concentration per unit time) and broken down in a first-order process with rate δ, for example a RNA being transcribed and degraded. This is known as a

Poisson process. Let us also assume that the reactions take place inside a cell of volume v. Schematically:

$$\phi \underset{\delta}{\overset{k}{\rightleftharpoons}} \mathrm{A}$$

The configuration of the system is fully described by the number of molecules of A, n. In this case, the vector $\mathbf{n}$ has a single component, $\mathbf{n} = (n)$, and will be treated as a scalar in the ongoing section for simplicity. The probability of finding n molecules of A at a time t is then $p(n, t)$. Let us define f_n and g_n as the probabilities per unit time of a molecule being produced or degraded, respectively, in a system with n molecules. The state n can be achieved if a molecule is produced when the system has $n-1$ molecules, or if it is degraded when it has $n + 1$. The probability fluxes can be represented as follows:

$$\boxed{p(n-1,t)} \underset{g_n}{\overset{f_{n-1}}{\rightleftharpoons}} \boxed{p(n,t)} \underset{g_{n+1}}{\overset{f_n}{\rightleftharpoons}} \boxed{p(n+1,t)}$$

Introducing these specific four fluxes in the general formulation of the master equation (1.19), it ends up reading

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} p\,(n,t) = f_{n-1}\, p(n-1,t) + g_{n+1}\, p(n+1,t) \\ - (f_n + g_n)\, p(n,t) \end{aligned} \tag{1.20}$$

It is possible to extract information from equation 1.20 without explicitly solving it. For instance, the evolution of the mean $\langle n \rangle$ of the distribution $p(n, t)$, given by

$$\langle n \rangle = \sum_n n\, p(n,t) \tag{1.21}$$

can be obtained multiplying both sides of 1.20 by n and summing over n:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \langle n \rangle = \sum_n n\, f_{n-1}\, p(n-1,t) + \sum_n n\, g_{n+1}\, p(n+1,t) \\ - \sum_n n\, (f_n + g_n)\, p(n,t) \end{aligned} \tag{1.22}$$

Because A is created at a constant rate k inside a cell of volume v, we have $f_n = k\, v$. Degradation is a first-order reaction with rate δ, so $g_n = \delta\, n$. Then equation 1.22 becomes

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \langle n \rangle = k\, v \sum_n n\, p(n-1,t) + \delta \sum_n n\,(n+1)\, p(n+1,t) \\ - k\, v \sum_n n\, p(n,t) - \delta \sum_n n^2\, p(n,t) \end{aligned} \tag{1.23}$$

Finally, making use of the fact that

$$\sum_n h(n) = \sum_n h(n \pm 1) \tag{1.24}$$

for any function $h(n)$ when the sum is carried over all n, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle n \rangle = k\, v - \delta \langle n \rangle \tag{1.25}$$

Equation 1.25 describes the dynamics of the average $\langle n \rangle$ of the probability $p(n, t)$. At equilibrium ($\mathrm{d} \langle n \rangle / \mathrm{d}t = 0$), this average equals kv/δ. The same idea can be used for all the moments of $p(n, t)$, $\left\langle n^2 \right\rangle$, $\left\langle n^3 \right\rangle$, etc. (fig. 1.2).

Dividing both sides of equation 1.25 by v, we can arrive at the following expression for the average *concentration*:

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle [A] \rangle = k - \delta \langle [A] \rangle \tag{1.26}$$

being $\langle [A] \rangle \equiv \langle n \rangle / v$. If we had described the system from a deterministic point of view using the mass action law, we would have obtained an ODE equivalent to equation 1.26. In general, the deterministic approach to a set of stochastic biochemical reactions describes the behavior of the average abundances of the molecules involved in them. However, averages alone do not always provide enough information about the system dynamics.

1.2.4 Stochastic kinetic models

The master equation is a powerful tool, but the cases where its analytical solution is tractable are very limited. In many scenarios an explicit solution for $p(\mathbf{n}, t)$ is not required and the behavior of stochastic systems can be studied through computational simulations that implement the mathematics of the master equation into algorithms.

The Gillespie algorithm

In 1977, Gillespie introduced an algorithm capable of generating exact time trajectories that describe the evolution of a system of coupled reactants.[22] In this context, *exact* means that the trajectories account for the inherent fluctuations in a way that complies with the probabilities given in the master equation.

Consider a system with L molecular species coupled by M reactions. Given the configuration $\mathbf{n} = (n_1, n_2, ..., n_L)$ at a time t, the dynamic of the system will be given by:

- The *time* needed for the next reaction to take place.
- The *type* of reaction it will be.

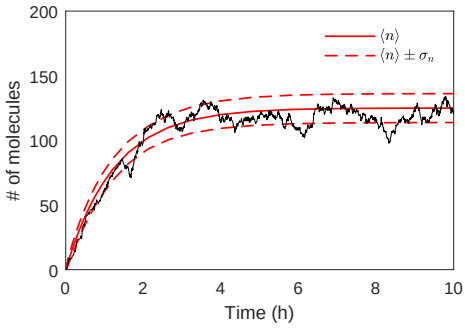


Fig. 1.2. Dynamics of a Poisson process. The master equation can be solved analytically in a system where a single molecule is being produced and degraded in a first-order reaction. Here, the solid red line represents the solution of equation 1.25 with $k = 0.1\text{molecules}/\mu\text{m}^3/\text{h}$, $\delta = 0.8\text{h}^{-1}$ and $v = 1000\mu\text{m}^3$ for the initial condition $n = 0$. Dashed lines represent the analytical solution for $\langle n\rangle + \sigma_n$ and $\langle n\rangle - \sigma_n$, where $\sigma_n^2 = \left\langle n^2\right\rangle - \langle n\rangle^2$. The black line corresponds to a trajectory simulated with the Gillespie algorithm.

The *reaction probability density* $p(\mu, \tau)$ is defined such that

$$p(\mu, \tau)\,\mathrm{d}\tau \geq 0 \tag{1.27}$$

quantifies the probability of the next reaction being of type μ ($\mu = 1, 2, ..., M$) and taking place in the interval $[t + \tau, t + \tau + \mathrm{d}\tau]$ with t being any given time. Assuming that the system is well mixed and that a reaction happens only when the reactants collide with each other in a certain manner, it is possible to use physical arguments to arrive at the following explicit form of $p(\mu, \tau)$:

$$p(\mu, \tau) = \begin{cases} a_\mu \exp(-a_0 \tau) & \text{if } \tau \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{1.28}$$

where a_μ is called the *propensity* for the μ-th reaction, and $a_0 \equiv a_1 + a_2 + ... + a_M$. Propensities are defined such that $a_\mu \mathrm{d}t$ quantifies the probability of a reaction of type μ taking place in the interval $[t, t + \mathrm{d}t]$. They depend on the configuration of the system at time t and the kinetic constants. In some cases they are scaled by the system volume plus additional factors if any reactant has a molecularity greater than 1.[22]

From the algorithm's perspective, the dynamic can be seen as a *reaction race*: given an arbitrary state of a system, each reaction has a certain probability to occur in the following time interval. It is possible to use the probability distribution in 1.28 to simulate which reaction will "win the race" (and how long it will take for it to do so), thus taking place in the system, changing its configuration and the propensities accordingly. The "race" will then be restarted with updated values for the propensities. Trajectories can be constructed in such iterative manner. An example is shown in figure 1.2.

The chemical Langevin equations

Trajectories generated with the Gillespie algorithm provide valuable information about the dynamics of stochastic biochemical systems without the need to explicitly solve the master equation. Still, a major downside is the simulation time scaling with the system complexity: the elapsed time between consecutive reactions decreases for growing a_0 (see equation 1.28), which in turn increases with the number of reactions and their propensities. This means that systems with large numbers of molecules coupled by many reactions will require a vast amount of iterations to cover a fixed time span.

The chemical Langevin equations provide an approximation in which molecule numbers are represented by real-valued (rather than discrete) variables. Let $\mathbf{n}(t) = (n_1(t), n_2(t), ..., n_L(t))$ describe the state of a system formed by L reacting molecules with abundances given by $n_i(t)$ $(i = 1, 2, ..., L)$ at time t. These species are coupled by a set of M chemical reactions, with propensities $\mathbf{a} = (a_1, a_2, ..., a_M)$. Propensities depend on the state of the system at the time t, $\mathbf{n}(t)$, so in general $\mathbf{a} = \mathbf{a}(t) = \mathbf{a}(\mathbf{n}(t))$. The Langevin equation for the i-th specie would take the form:

$$\frac{\mathrm{d}}{\mathrm{d}t} n_i(t) = f_i(\mathbf{a}(\mathbf{n}(t))) + \omega_i(\mathbf{a}(\mathbf{n}(t))) \tag{1.29}$$

Here f_i represents a function analogous to the one that results from the deterministic approach, and ω_i is a noise term that depends on the propensities. Explicitly:

$$\begin{aligned} f_i(\mathbf{a}(\mathbf{n}(t))) &= \sum_{\mu} \Delta_{i\mu} a_\mu(\mathbf{n}(t)) \\ \omega_i(\mathbf{a}(\mathbf{n}(t))) &= \sum_{\mu} \Delta_{i\mu} \sqrt{a_\mu(\mathbf{n}(t))}\, \xi_\mu(t) \end{aligned} \tag{1.30}$$

where $\Delta_{i\mu}$ represents the change in the number of molecules of the i-th specie when a reaction of the μ-th type takes place, and the term $\xi_\mu(t)$ is a Gaussian white noise.

The main advantage of these equations is that they can be solved numerically with a custom $\mathrm{d}t$ that does not scale with the system complexity, as opposed to the Gillespie algorithm. It is generally accepted that the Langevin equations are valid when the numbers of molecules are high, however, a more formal condition has to do with the choice of $\mathrm{d}t$: it has to be large enough for the number of reactions in any interval $[t, t + \mathrm{d}t]$ to be much greater than one, but small enough for the propensities to remain relatively unchanged during that time, that is, $\mathbf{a}(t) \approx \mathbf{a}(t + \mathrm{d}t)$.[23] If there is no $\mathrm{d}t$ that can satisfy both conditions simultaneously, the approximation fails.

1.3 Cell-to-cell variability

Cells constantly process information regarding their enviroment (nutrient presence or absence, signals from neighbor cells, etc.) and their own internal state. Much of our

knowledge on how this processing is done is based on ensemble measurements. But ensemble behaviors are not necessarily representative of any individual, as cell-to-cell variability is always present in any population.[24]

1.3.1 Understanding cell-to-cell variability

Origins and classification

Genetic variability (often referred to as *genetic polymorphism*) is the variation in the DNA sequence across individuals of the same species. Although its determinants are not fully understood, genomics has provided a powerful tool to investigate the underlying evolutionary processes that give rise to this genetic diversity.[25] It is known that individuals of a polymorphic population can show significantly different responses to environmental changes. This has implications in, for example, human health[26] or the spread of infectious diseases.[27]

But even genetically identical (*isogenic* or *clonal*) cells in a homogeneous environment can present large differences in size, morphology, molecular components and activity. Rather than being the result of DNA sequence variations, this *non-genetic* variability is due to the circuits that regulate cellular functions being subject to stochastic fluctuations, or *noise*, in the levels of their components. At the single cell level, it is common to distinguish between *intrinsic* and *extrinsic* noise:

- Intrinsic noise is induced by the inherent stochasticity in the biochemical reactions governing signaling and gene expression networks.
- Extrinsic noise arises from fluctuations in other cellular components.

Intrinsic and extrinsic noise are are actually not independent.[28] Much of the variability in the cellular components that induce the so-called extrinsic noise is, in turn, a result of the intrinsic stochasticity of the biochemical processes in which said components are produced or degraded. At the same time, the magnitude of intrinsic noise can depend on extrinsic factors (e.g. cellular volume). Nonetheless, both contributions are typically treated separately.

An illustrative experiment was carried out by Elowitz et al.,[29] in which bacterial strains of *E. coli* were built incorporating the sequences encoding for the cyan and yellow fluorescent proteins (*cfp* and *yfp*, respectively) in the chromosome. Both reporter genes were controlled by identical promoters and their expression could be quantified through fluorescence microscopy at a single cell resolution. Cells were cultured in the presence or absence of a specific chemical (IPTG). With IPTG, the promoters behaved as constitutive ones (i.e. reporter gene expression was permanently activated), while in its absence the expression of the reporters was heavily reliant on the binding/unbinding of a transcription factor (the LacI protein), thus maximizing intrinsic noise (fig. 1.3).

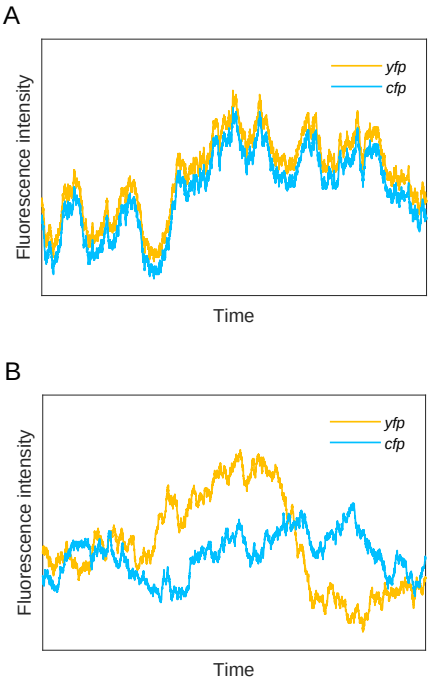


Fig. 1.3. Intrinsic versus extrinsic noise in engineered *E. coli*. In an experiment by Elowitz et al., the expression of two reporters in *E. coli* (*yfp*, yellow, and *cfp*, blue) is controlled by identical promoters. In this figure, the abundance of both *yfp* and *cfp* is shown for a single cell over time as measured with fluorescence microscopy. **A.** Under conditions of low intrinsic noise, fluctuations observed in the levels of both reporters are highly correlated, indicating that they are caused by gene-unspecific (extrinsic) factors. **B.** When the expression of each reporter depends on the infrequent binding/unbinding of a transcription factor, intrinsic noise becomes important and expression of the two genes becomes uncorrelated. Figure adapted from Elowitz et al.[29]

Similar experiments have also been carried out in eukaryotes,[30] demonstrating how low molecular copy numbers fundamentally limit gene regulation predictability.

In general, using time lapse microscopy to measure expression correlations between genes in single cells is a powerful tool to explore noise propagation through gene expression networks.[31] Simple statistical measures of dispersion, such as the coefficient of variation (CV, standard deviation relative to mean) or the Fano factor (F, variance relative to mean) are appropriate measures of variability.[32] This provides a framework in which noise at the single cell level can be quantified and modeled.

Noise in gene expression

Gene expression entails the assembly of basic molecular components (nucleotides and aminoacids) into nucleic acids and proteins. This involves the coordinated action of many different molecular steps regulated by a vast amount of enzymes and protein complexes: chromatin remodeling factors, transcription factors, polymerases, ribosomes, etc. Noise in gene expression is a result of both fluctuations in the number of these molecular constituents (metabolites and gene expression machinery) and the stochastic-

ity of the biochemical reactions in which they are involved.

Single-cell genomic and proteomic studies have shown that noise in protein abundances is dominated by the stochastic production and degradation of messenger RNAs.[33,34] Transcription depends on single molecule kinetics and often occurs in an intermittent manner, which yields proteins not trickling at uniform rates but rather being produced in *bursts*. This is usually a consequence of genes' promoters stochastically switching between long-lived active and inactive states, resulting in mRNA production bursts that are amplified at the protein level. The frequency and size of transcriptional bursts determine the activity of a gene and are affected by both gene-specific and genome-wide factors.[35] On the other hand, protein lifetimes are often longer than times elapsed between consecutive transcriptional bursts. The accumulation (buffering) of proteins tends to average out noise induced by bursting.

Significant variation in noise levels has been found across specific genes and pathways. This has been associated with their functional differences, suggesting that variability in gene expression can be either beneficial or deleterious (and thus evolutionarily selected against) depending on the context.[36,37] Stress-response genes, for instance, are particularly noisy,[33] pointing at a potential benefit of stochasticity in this case.

For the most part, noise in gene expression comes from extrinsic sources,[38] but the identity and relative contributions of said sources is still unclear. Some examples that have been characterized affect gene expression in a global manner (e.g. random partitioning of proteins at cell division[39] or cell cycle stage[40]), while others are limited to specific pathways or collections of genes (e.g. fluctuations in the abundances of upstream transcription factors[28]).

Noise in signaling networks

Signaling networks have the purpose of transmitting information about the extracellular region to downstream effectors, allowing cells to respond by adjusting their physiological state. Although biochemical noise compromises the fidelity of signal transduction, it seems to also be possible for cells to minimize fluctuations and achieve high levels of information transmission at the single-cell level.[41]

Some authors hold that the suppression of molecular fluctuations, when necessary, requires a "brute force" strategy (i.e. the dramatic increase of the number of molecules involved in a signaling pathway),[42] thus being a highly energy demanding process. Another proposed mechanism is the integration of information in time, which would compromise the rapidness of a response in favor of its reliability.[43] In general, noise suppression appears to happen when the key biological output of a process is the behavior of individual cells rather than that of a population (e.g. in chemotaxis). It has been pointed that signaling networks could exploit single-cell level noise to increase population-level information transfer.[44] This suggests a trade-off between the information transmitted in individual cells and the information effectively controlling population-level responses.

From this perspective, low levels of information transfer (induced by biochemical noise) would not be a physical limit.

Paulsson et al. even introduced the counter-intuitive concept of *stochastic focusing*.[45] This refers to the idea of biochemical networks exploiting noise (rather than suppressing it) in non-linear, multi-step processes to eventually enhance signal detection. Cai et al. also reported that the yeast calcium-response system exploits noise to increase signal transduction reliability, using the frequency of nuclear localization bursts of a transcription factor to coordinate the expression of multiple target genes.[46]

1.3.2 Implications

Identifying and characterizing the sources of cell-to-cell variability is not a merely academic question. Although noise is an impediment for the design of deterministic biological circuits, it is becoming increasingly clear that heterogeneity plays important functional roles in several key cellular processes.[24,34,47,48] In many cases, it provides critical functions that would otherwise very difficult (if not impossible) to achieve.

Both the magnitude of stochastic fluctuations and their frequency determine the extent of their consequences. Small but persistent changes in protein abundance can have significant functional implications, whereas the effect of large fluctuations may be negligible if they occur often enough.[39] In general, the time scale of intrinsic fluctuations is much shorter than that of extrinsic ones, which could explain why the latter are usually more relevant for the cell's functionality.

Noise in gene expression induces differences in gene activity that confer each cell a unique "expression fingerprint". Analogously, noise in signaling pathways can give rise to different individual responses of identical cells. This creates phenotypic heterogeneity even in isogenic populations. But to link phenotypic and molecular variability both must be simultaneously quantified and causal relationships between them need to be established.[24] This has been done in many cases, where variation at the phenotypic level has been connected to some sort of stochasticity in key proteins or genetic circuits.

Cellular decision making

Cells can acquire functionally different, heritable fates without environmental or genetic changes. This idea is often associated to the *Waddington's landscape*,[49] a depiction of cells' decision making process in which each individual is represented as a rolling ball in a fitness landscape where local minima correspond to stable states. However, this deterministic view is being challenged by an increasing body of theoretical and experimental evidence.[50] Noise seems to play a key role such that even identical cells released from the same "location" of Waddington's landscape can end up in different states due to random fluctuations.

An interesting example is found in the metabolism of *E. coli*. The lactose utilization network of this bacteria displays an "all or nothing" behavior. The well-characterized

lac promoter simultaneously controls the expression of three genes responsible for the uptake, breakdown and processing of lactose molecules. The activity of this promoter is regulated by a set of transcription factors, which results in cells being able to have the lactose utilization machinery either activated or deactivated, i.e. individual bacteria can be in one of two mutually exclusive metabolic states. Stochastic transitions between states have been observed even in cells that were originally uninduced.[51]

Cell fate choice, development and differentiation

Noise can make cells switch between alternative metastable states. State-switching systems are often based on feedback loops: stochastic fluctuations alone are typically too small to create binary switches between alternative cell fates.[52] Additional mechanisms are also needed to stabilize one specific fate choice.

It seems reasonable to assume that a precise, deterministic-like execution of the developmental program would be critical for the production of functional tissues and organs. Yet researchers have found many scenarios in which cell differentiation during development is linked to stochastic gene expression. A prominent example was reported by Serizawa et al. in a study of the mouse olfactory system development.[53] Neurons of said system have a vast amount of odorant receptor (OR) genes that are expressed in a mutually exclusive fashion, following a "one neuron-one receptor" logic. This structure is established during the mouse developmental period, and has been shown to rely on the stochastic activation of a specific OR gene (induced by intrinsic noise) followed by a negative feedback stage that shuts down all the others.

Another analogous, well characterized example of noise in gene expression affecting cell differentiation is found in hematopoietic stem cells. These express two mutually repressing genes (PU.1 and GATA1). At a certain point, that can be either induced or spontaneous (driven by intrinsic fluctuations only), one of the two takes over, repressing the other's expression and unleashing a signaling cascade that leads to the eventual differentiation of stem cells into myeloid (PU.1) or erythroid (GATA1) cells.[54]

Evolution

A very straightforward way in which noise can play a role in evolution is by expanding the phenotypic range associated with a given genotype. In this sense, increased noise can be an advantage when facing fluctuating environments. Notably, the strength of selection determines the role of fluctuations: high levels of noise seem to be beneficial when selection pressure is intense, whereas lighter pressure selects for reduced fluctuations.[55] This is consistent with the idea of increased phenotypic variation during adaptation to new, challenging conditions, followed by noise reduction when selection becomes stabilizing.

Noise can also determine which specific gene expression network topologies get favored evolutionally. Even though different architectures for a genetic circuit are often

able to perform one same task, each one can show a different distribution of stochastic fluctuations, effectively causing functional differences. This has been proposed to explain why seemingly equivalent topologies are selected for or against to implement a given cellular process.[56]

Virus infection

Virus infection is the result of the concerted action of many cellular processes, including endocytosis (cell's internalization of the viral genetic material). Significant cell-to-cell variation has been shown in these processes even in isogenic populations. This variability cannot be attributed to physical constraints on the accessibility to virus particles,[57] which points at phenotypic state as a key determinant of individual cells' responses.

Aging

The increase of noise in gene expression with age is a general observation. For instance, it has been shown that the expression of both housekeeping and strain-specific genes in murine cardiac monocytes becomes more stochastic as the organism ages.[58] Interestingly, the same effect was observed in cells isolated from young animals and treated with hydrogen peroxide, suggesting oxidative stress could be a factor. Similar observations were made in murine muscle tissue.[59]

Apoptosis

Apoptosis is a process by which cells of multicellular organisms die in response to either external (*extrinsic apoptosis*) or internal (*intrinsic apoptosis*) stimuli. Unlike necrosis (a form of cell death that results from cellular injury), apoptosis is programmed. This means that cells contain the biochemical circuits necessary to respond to apoptotic signals by unleashing a cascade of events that include cell shrinkage, chromatin condensation, nuclear fragmentation...

When a population of isogenic cells is presented with a pro-apoptotic signal, some of the individuals die while others survive in a phenomenon known as *fractional killing*. If the resistant fraction is left to grow and the reconstructed population is presented with the apoptotic signal again, fractional killing is still observed. Additionally, both the apoptotic fate (death/survival) and the time to death show significant correlation between sister cells. These findings constitute strong evidence for a non-genetic origin of fractional killing, that has instead been linked to noise-induced phenotypic heterogeneity.[60–65]

Cancer

Cancer is a disease known to involve changes in the genome. These changes provide tumor cells with functions such as apoptosis evasion, sustained angiogenesis, insensitivity to anti-growth signals, etc.[66] Even though the process is generally associated with DNA sequence alterations in specific key genes (*oncogenes*), a more general approach suggests that non-genetic variability could induce heritable state variants in cell populations that would serve as a temporary substrate for cancer progression even before mutations take place.[67]

1.4 Mitochondrial variability

The complexity of the molecular steps and machines involved in gene expression and signaling introduces sources of noise at many levels, but a common constraint to these processes is their energy dependence. Mitochondria provide most of this energy in eukaryotic cells. These cellular organelles occupy a significant fraction of the cytoplasm and can show large variation in number and functionality across cells of a clonal population, as has been observed in HeLa cells,[68,69] hematopoietic stem cells[70] and solid tumors.[71] Mitochondrial variability has been shown to induce cell-to-cell differences in energy budget[68] and pointed as an important source of extrinsic noise.[68,69,72]

1.4.1 Mitochondria and metabolism

ATP: the cellular fuel

ATP stands for *adenosine triphosphate*. It is an organic chemical found across all life forms that serves as an energy carrier and can be used to power many processes in cells.[73] Structurally, ATP consists of an adenine attached to a sugar (ribose), attached in turn to a triphosphate group.

The energy carried by an ATP molecule can be utilized in many reactions through the release of one or two of the phosphates in its triphospate group, while the adenine and sugar groups remain unchanged. The release of one or two phosphate groups converts ATP into ADP (*adenosine diphosphate*) or AMP (*adenosine monophosphate*) respectively. On the other hand, the energy stored in nutrients or obtained through photosynthesis can be used to restore the phosphate bonds, converting ADP and AMP back into ATP. A typical cell contains $\sim 10^9$ ATP molecules at any time, that get turned over (used and replaced) every 1-2 minutes.

Anaerobic glycolysis versus oxidative phosphorylation

Before nutrients can be utilized by cells, the large polymeric molecules in food must be broken down into their monomer subunits (proteins into amino acids, polysaccharides

into sugars, and fats into fatty acids and glycerol). This step (*digestion*) is carried out by specific enzymes either outside cells or in specialized organelles within them (*lysosomes*), so that only the small molecules derived from it enter the cytosol. In the cytosol, a chain of metabolic reactions known as *glycolysis* converts a molecule of glucose into two molecules of pyruvate, releasing two molecules of ATP in the process. This process is anaerobic (i.e. the reactions involved require no oxygen) and very inefficient in the sense that only ~5% of the energy potential of the glucose (38 ATP molecules) is exploited.[74] Sugars other than glucose, amino acids, fatty acids or glycerol follow similar pathways converging on the pyruvate production stage.

For most eukaryotes, glycolysis is just the beginning of the food molecule breakdown process. Pyruvate molecules are transported into mitochondria, where each one is converted into CO_2 plus an acetyl group which becomes attached to the coenzyme A (CoA) to form acetyl-CoA. In mitochondria, acetyl groups enter the *citric acid cycle* where they are oxidized to CO_2, producing large amounts of NADH, an electron carrier molecule. In the mitochondrial inner membrane (which is folded many times for increased surface), the high-energy electrons of NADH are passed along an electron-transport chain. The energy released by their transfer is used to eventually produce ATP in a series of oxygen-consuming reactions known as *oxidative phosphorilation*. Most of the cell's ATP is produced in this final stage.

Many eukaryotes, prominently normal differentiated cells in metazoan organisms, obtain the majority of their ATP through oxidative phosphorilation in mitochondria.[75] The combination of the anaerobic glycolysis and oxidative phosphorilation pathways is known as *cellular respiration*. It produces about 30 ATP molecules per glucose molecule, making it a highly efficient process.

Mitochondria and the evolution of complex life

Eukaryotic cells are generally more structured than prokaryotes, with much larger genomes and proteomes and thus a higher degree of complexity. All eukaryotes either possess mitochondria or some kind of remnant of them,[76] which makes it plausible for mitochondria and eukaryotic life to have their origin in the same event,[77] possibly a symbiotic association between an aerobic and an anaerobic bacteria that took place about 4 billion years ago. The fact that mitochondria contain their own genetic material supports this hypothesis. This mitochondrial DNA (mtDNA) encodes proteins of the respiratory electron transport chain. It consists of 37 genes in humans.

The acquisition of mitochondria, and specifically of mitochondrial genes, enabled oxidative phosphorilation raising the energy power of the cell by several orders of magnitude. This allowed cells to overcome energetic barriers and increase their genome size by a factor of roughly 200,000-fold,[78,79] with the subsequent leap in complexity.

1.4.2 The origins of mitochondrial variability

Mitochondria are variable from cell to cell in number, morphology and functionality.[68,69,80,81] Fluorescence microscopy combined with specific markers has been used to determine the abundance, size or even physical continuity of mitochondria, finding significant variation from cell to cell in isogenic populations of several cellular strains. The mitochondrial membrane potential is commonly used as a proxy for mitochondrial functionality, showing significant variability as well.

Several studies have shown that the main source for cell-to-cell heterogeneity in mitochondrial content is the asymmetric apportioning of individual mitochondria at cell division.[68,80,82] This idea is similar to the the random segregation of a given number of elements into two subsets, known as *binomial partitioning*. Such mechanism is generally accepted to explain how other cellular components (e.g. proteins or RNAs) are split between daughter cells at mitosis, and has been shown to be consistent with experimental data.[83,84]

Johonston et al. developed a model capable of recapitulating the natural variability in mitochondrial content observed in populations of clonal HeLa cells.[80] The model is formulated in terms of two coupled differential equations for the growth and segregation at mitosis of both the mitochondrial mass and the total cellular volume of a single cell. The three key variables are v, the cell's volume, n, its mitochondrial mass and f, a parameter that quantifies the efficiency of mitochondria within a cell. Specifically, f is interpreted as a proportionality factor linking mitochondrial density (n/v) and ATP concentration. The equations are:

$$\begin{aligned} \frac{\mathrm{d}v}{\mathrm{d}t} &= \alpha f n \\ \frac{\mathrm{d}n}{\mathrm{d}t} &= \beta f n \end{aligned} \tag{1.31}$$

where α and β are proportionality constants whose value was adjusted to match experimental results.

In the model, growth is deterministic (equations 1.31 are so), but randomness is introduced at cell division. When a cell reaches a threshold volume (v_{max}), mitosis takes place. Both the mitochondrial mass and the cellular volume are split between daughter cells. If we let (v, n) characterize the parent cell, and (v_1, n_1) one of the two daughters chosen arbitrarily, at cell division:

$$\begin{aligned} v_1 &= \Phi\left(\frac{v_{max}}{2}, \sigma_v\right) \\ n_1 &= \Phi\left(\frac{n}{2}, \sqrt{\frac{n}{4}}\right) \end{aligned} \tag{1.32}$$

where $\Phi(\mu, \sigma)$ represents a normal distribution with mean μ and variance σ^2. The variance of the volume distribution (σ_v^2) is chosen to match experimental data on volume

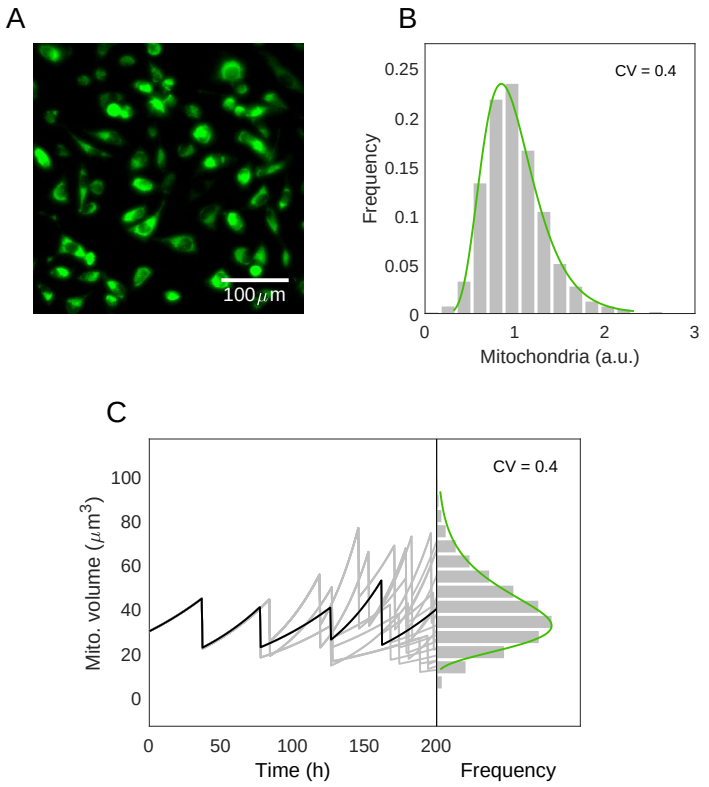


Fig. 1.4. Origins of mitochondrial variability. Mitochondrial content and functionality are variable from cell to cell. **A.** An isogenic population of HeLa cells was stained with the MitoTracker green reporter and imaged using fluorescence microscopy. MitoTracker binds to the mitochondrial membrane, thus being a good reporter of mitochondrial mass. Despite all cells being genetically identical, differences in the fluorescence intensity are appreciated. **B.** Fluorescence intensity was quantified for ~3500 cells. The distribution of mitochondrial mass (gray), in arbitrary units, was obtained from this quantification after normalizing by the population average, showing a coefficient of variation (CV) of about 0.4. The green line represents a fit to a log-normal distribution. **C.** The mathematical model developed by Johnston et al.[80] (equations 1.31 to 1.33) assumes deterministic cell growth and binomial partitioning of mitochondria between daughter cells at mitosis. We used it to reconstruct a population starting from a single cell and allowing several rounds of division until equilibrium was reached (no significant change found in the distributions of cell and mitochondrial volumes). Gray lines represent time trajectories of the cells in the simulated population. One arbitrarily chosen cell has been highlighted in black. The right panel shows the distribution of mitochondrial volumes of the population, which is well fitted by a log-normal (green line) and has a CV of 0.4, consistent with experimental findings.

partitioning. The variance of the mitochondrial distribution ($n/4$) is chosen to represent a binomial distribution in the limit of large n. Naturally, the other daughter cell would inherit a volume $v_2 = v - v_1$ and a mitochondrial mass $n_2 = n - n_1$.

Both daughter cells inherit the same mitochondrial functionality f_D, wich is resampled each time a new mitosis takes place and relates to the parent's functionality f_P according to:

$$f_D = \frac{1 + f_P}{2} + \Phi\left(0, \sigma_f\right) \tag{1.33}$$

where σ_f is chosen through optimization with constrains regarding experimental measures of cell cycle length variability in HeLa cells. The scenario of $f_D < 0$ makes no physical sense but can happen when sampling from equation 1.33. To exclude this possibility, f_D is resampled if its value falls under a small (but greater than zero) threshold.

The model is able to quantitatively reproduce the distribution of mitochondrial content found in an isogenic population of HeLa cells through fluorescence microscopy, as shown in figure 1.4. It does so by introducing noise only in the form of asymmetric (binomial) partitioning of mitochondria at cell division. Indeed, quantification of mitochondrial content after mitosis shows that the ratio of mitochondrial mass among daughter cells follows a skewed distribution, pointing towards such a mechanism of asymmetric segregation.[68,80]

Mitochondrial biogenesis is characterized by continuous cycles of fusion and fission that are faster than the cell cycle period: each mitochondrion undergoes an average of ~5 fusion/fission events per hour,[85] which likely yields a steady population of mitochondria inside the cell.[85,86] Interestingly, mitochondrial fission is enhanced during mitosis,[87] possibly facilitating passive and stochastic partitioning to daughter cells. Recent work has showed that, in yeast, mitochondria are segregated between daughter cells to achieve similar concentrations, that is, in proportion to the available cytoplasmic volume as it happens with RNAs and proteins.[88] Another study showed that mammalian epithelial stem-like cells follow asymmetric apportioning of aged mitochondria, with cells that inherit fewer old mitochondria maintaining stemness.[89]

However, other factors can make the dynamics of organelle biosynthesis vary from cell to cell, adding more potential sources for mitochondrial heterogeneity. For instance, cell cycle stage seems to act as a modulator for mitochondrial mass.[69] Some authors have described active mechanisms based on actin filament control[90] through which cells are able to undertake asymmetric segregation of mitochondria. This can be understood as a sort of "quality control".

1.4.3 Mitochondrial variability and gene expression

Energy is a common link connecting many of the factors that are variable from cell to cell and impose global constraints on gene expression.[91–93] Examples are cell's growth

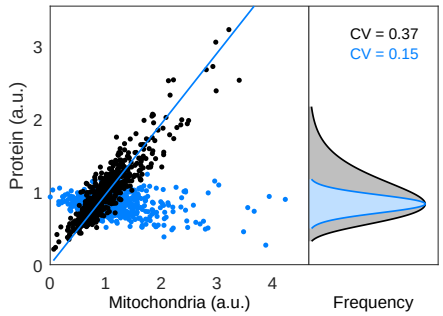


Fig. 1.5. Mitochondrial contribution to protein variability. Mitochondrial levels co-vary with the amount of total protein in HeLa cells (black dots). The contribution of mitochondria to total protein variability (gray distribution) can be quantified by de-trending the data (blue dots and distribution, see equations 1.34 to 1.36). In this case, $\mathrm{CV}_p = 0.37$ and $\mathrm{CV}_{off} = 0.15$, which yields MCV = 0.6. This is interpreted as mitochondria accounting for roughly 60% of the variability observed at the protein level.

state[94] or the availability of metabolites and enzymes.[95] The molecular machinery needed for transcription and translation (most importantly ribosomes[96]) also requires energy to be synthesized. In addition, many of the steps of the gene expression cycle depend upon the overcome of free energy barriers.[97]

All of this makes gene expression a highly energy demanding process, taking up to ~80% of the cellular ATP.[78,79] Mitochondria, being the main provider of energy in eukaryotes, plays important roles in several steps of the gene expression cycle. Other global factors that affect gene expression also undergo fluctuations that are, to some extent, modulated by mitochondrial content.

Quantifying mitochondrial and gene expression variability

To investigate the relationship between mitochondrial variability and cell-to-cell differences in gene expression, one can simultaneously measure the amount of mitochondria and proteins, RNAs or different markers of transcription and translation activities at the single cell level (e.g. using fluorescent antibodies, see figure 1.4).[69] The protein distribution for a whole population can have a large width (i.e. a high coefficient of variation), but much of this variability comes from cell-to-cell differences in mitochondrial content. The co-variation of mitochondria and protein abundances is a result of this dependence (fig. 1.5).

Given that variability in protein abundance is produced by a combination of sources, one of them being mitochondria, we have:

$$\sigma_p^2 = \sigma_m^2 + \sum_i \sigma_i^2 \tag{1.34}$$

where σ_p^2 represents the total variance at the protein level and σ_m^2 is the part of that variance that comes from mitochondrial heterogeneity. The sum over i corresponds to every other source of variability (including intrinsic noise and other extrinsic sources independent of mitochondria). The relative contribution of mitochondria to the total variability at the protein level can be quantified by representing protein abundance as a function of mitochondria abundance and measuring the correlation and the off-diagonal dispersion of the data. This is equivalent to removing the co-variation across the diagonal in figure 1.5. Such off-diagonal dispersion can be attributed to sources of noise that do not depend on mitochondrial mass, that is, it corresponds to the summation in equation 1.34. Thus, if we call

$$\sigma_{off}^2 = \sum_i \sigma_i^2 \tag{1.35}$$

we can obtain the variance associated to mitochondrial variability ($\sigma_m^2 = \sigma_p^2 + \sigma_{off}^2$) and also to all other sources (σ_{off}^2). The coefficients of variation, CV_p, CV_m and CV_{off} can be extracted by simply dividing by the respective means. It is useful to define the *Mitochondrial Contribution to Variability* (MCV) as:

$$MCV \equiv 1 - \frac{CV_{off}}{CV_p} \tag{1.36}$$

The MCV quantifies the fraction of the variability observed at the protein level that can be attributed to cell-to-cell differences in mitochondrial content. This parameter provides a simple, intuitive way of separating the contribution of mitochondria to protein variability from other sources of noise, in a way that is somehow analogous to the decomposition of total variability into its intrinsic and extrinsic components (fig. 1.3).[29] Using specific antibodies, it is possible to obtain MCVs for individual proteins. Values found span between ~25% and ~75%, changing across groups of proteins involved in different cellular processes.[69]

Chromatin organization

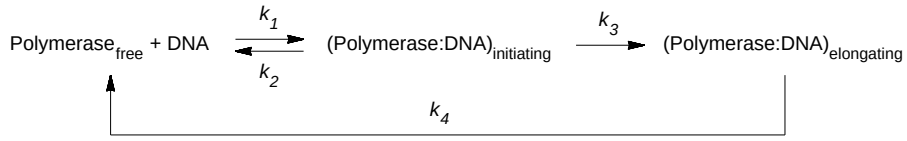
Nucleosomes are the basic repeating units of DNA packaging, consisting of a segment of DNA in sequence around eight histone protein cores. They compete for DNA binding with many transcription factors, such that these must induce nucleosome reorganization in order to interact with genomic regulatory elements. There is a variety of models for how this arrangement is done, but the general view is that it is a dynamic, ATP-dependent process.[98] According to the *assisted loading mechanism*, short-lived chromatin conformations would be induced by the action of remodeling complexes, allowing transient "windows of access" for secondary transcription factors to access their binding sites.[99] These transient windows would operate within periods spanning minutes to hours,[68,100] while many secondary DNA binding proteins have binding/unbindig times of the order

of seconds. The relatively slow time scale for nucleosome rearrangement produces long periods of gene activation and silencing, which results in transcriptional bursting[30,101] and noise in gene expression.[102] While the size of these bursts seems to be characteristic of each promoter,[103] their frequency appears to be determined by nucleosome positioning alone.[104,105]

Many molecular machines consume ATP to change nucleosome configuration. Furthermore, acetyl-CoA is an essential cofactor for histone modifications that is produced in large amounts in mitochondrial respiration. At the same time, the NAD+ coenzyme (the oxidated form of NADH, also an essential part of the cytric acid cycle) catalyzes histone deacetylation.[106] These are all mechanisms through which mitochondria could affect chromatin remodeling and thus transcriptional bursting. Indeed, mitochondrial levels have been shown to largely co-vary with chromatin modification marks related to transcriptional activation such as the histone methylation mark H3K4me3.[69]

Transcription

The transcription process consists of a first stage in which large protein complexes assemble at a gene's promoter (*initiation*), and a second one in which said machinery starts moving along the DNA while polymerizing the transcribed RNA strand (*elongation*). The cycle is completed when transcription is finished and the polymerase, together with the rest of the molecular elements, is released from the DNA.



The rate constants for each of these processes (k_1 to k_4) can be extracted from the fluorescence loss in photobleaching experiments[†] that use polymerase fluorescent reporters.[107] Many studies have made experimental observations of initiation and elongation events in different cellular strains. They found fast turnover rates of the transcription machinery, which typically assembles and dissembles at the promoter every few seconds to minutes. Multiple incomplete rounds of initiation can take place before elongation starts, each one of them keeping no memory of previous history.[107,108] Transcription initiation is thus an inefficient, stochastic process. Elongation, on the other hand, seems to be more deterministic and energy dependent, with a speed that can change as much as 3-fold from cell to cell.[107]

[†] *Photobleaching* refers to the photochemical alteration of a fluorophore molecule to permanently prevent it from fluorescing.

Measuring cellular levels of BrU[‡] and mitochondrial content simultaneously in populations of HeLa cells shows that both co-vary. Moreover, the same experiment in cells with depleted ATP demonstrates that the elongating fraction of RNA Polymerase II and the elongation rate are highly dependent on cellular energy budget, while the rates of binding/unbinding of the polymerase complex to DNA (i.e. the fraction of initiating polymerase and the initiation rate) are relatively ATP-independent.[69] The transcription machinery stochastically binds and releases promoters, but requires energy to start and sustain the elongation phase. This is consistent with a picture where cells with increased mitochondrial content have a higher number of both genes actively transcribed and RNA polymerases engaged in elongation.

Variability in mitochondrial mass would thus induce transcriptional noise. Additionally, intercellular differences in the abundance of RNA polymerases have been shown to be an extrinsic source of noise in gene expression,[109] and could potentially be linked to mitochondrial heterogeneity as well. Further work is needed to elucidate the molecular mechanisms linking mitochondria to transcription initiation/elongation.

Alternative splicing

We have discussed the ability of mitochondria to amplify or decrease gene activity through the modulation of energy-dependent rate parameters of the gene expression cycle, notably transcription. But mitochondria are not just a biological "volume knob" that tunes gene expression up or down: this modulation can be heavily non-linear. A drastic example is alternative splicing.

Most eukaryotic protein-coding genes contain amino acid coding sequences known as *exons* separated by segments called *introns*. Transcripts of such genes are called *pre-mRNAs* (precursor messenger RNAs). Newly synthesized pre-mRNAs are typically processed in the nucleus to remove introns and splice the exons together into a mature mRNA that can be translated in the cytoplasm. Alternative splicing (AS) is the process through which particular exons of a gene may be included or excluded from the final processed mRNA. Studies using high-throughput technologies show that the vast majority (95-100%) of human pre-mRNAs that contain more than one exon can be transcribed into multiple alternatively spliced mRNAs (*isoforms*).[110] This explains why eukaryotic organisms can mantain proteome sizes much larger than their corresponding genome size. AS allows for genes to encode more than one protein: alternatively spliced mRNAs of the same gene yield different protein products when translated. AS appears to be driven by fluctuations in the splicing machinery in humans.[111,112] It is a major source of proteome diversity[113] and has important consequences in processes like development[114] and disease.[115]

The number and relative abundances of mRNA isoforms is highly variable[116] as

[‡]Bromouridine (BrU) is an immediate transcription precursor. When added to a cell culture it incorporates to newly synthesized RNA, thus serving as a proxy of cells' average transcriptional activity.

shown, for instance, in immune cells by single-cell transcriptomics.[117] Transcriptomic analysis of HeLa cell subpopulations with high and low mitochondrial content also reveals dramatic alterations in isoform abundance.[69] The mediation of mitochondria in AS is likely to be due to a combination of factors. Chromatin remodeling (for which energy is critical as we discussed) influences the outcome of AS by both distinguishing exon and intron recognition and regulating the recruitment of specific splcing factors.[118] Elongation speed, higly variable from cell to cell, is also linked to mitochondrial content and impacts the secondary structure of pre-mRNA, a key determinant of AS.[119]

Cell cycle and growth

The eukaryotic cell cycle is composed of four phases, out of which the first three (G_1, S and G_2) are collectively known as the *interphase* and the fourth one (M) corresponds to cell division.

- G_1 phase: cells synthesize RNA and proteins, increasing in size in preparation for subsequent steps of the cycle.
- S phase: DNA is replicated.
- G_2 phase: a period of rapid growth and protein synthesis preceding mitosis. G_2 phase is present in most, but not all, forms of eukaryotic life.
- M phase: cell growth stops and cellular energy is devoted to the division into two daughter cells (*mitosis*).

Gene expression is coupled to cell growth and division,[40,120,121] notably transcriptional activity is increased during the S/G_2/M phase. Such coupling can make it so protein and RNA concentrations (not copy numbers) remain relatively invariant along the cell cycle.[40,122] On the other hand, asymmetric partition of molecular components at cell division (see section 1.4.2 of this text) generates a "noise floor" for extrinsic noise.[28] Cell cycle also determines the timescale at which variability becomes relevant: intrinsic noise is typically faster than cell cycle periods and can be averaged out in many scenarios, while extrinsic sources often induce fluctuations that persist for time spans comparable to one cell cycle.[39]

The size of mammalian cells is coupled to cell cycle in normal tissues, so control mechanisms are required to limit variability in cell size.[123] Mitochondria may establish such control in mouse liver cells by regulating the balance between cell size and proliferation.[124] In addition, mitochondrial levels impact the cell's biosynthetic power (through modulation of transcription and translation) which determines cell growth and division.[96] These mechanisms could explain the observed relationship between mitochondrial mass and cell cycle length.[68,125]

Both mitochondria and cell cycle shape variability at the protein level. To assess these contributions simultaneously, they must be quantified along with protein abundance in single cells and the effect of each factor on the other should be studied with independence of the third one. Such analysis shows that the contribution of cell cycle to noise in gene expression is low (around 5-17% in different cell lines) compared to that of mitochondria.[69,126]

1.4.4 Implications

The current understanding is that metabolic alterations entail genome-wide changes in transcriptional activity. There is, however, evidence that the opposite scenario is also possible: any process that involves a shift of cellular energy is likely to result in a global gene expression change. Mitochondrial variability can thus play a role in cell individuality through its impact on many of the steps of gene expression due to variable ATP availability.

The immune response is a paradigmatic example of this. Metabolic reprogramming is needed before individual immune cells can achieve the correct physiological function in reaction to antigen exposure.[127,128] Hydrogen peroxide and superoxide anion, possibly originated at activated mitochondria,[129] have been shown to be produced shortly after T cell receptor antigen cross-linking.[130] Additionally, the increase in biomass needed for cellular expansion depends highly on ATP.[131] These evidences point at metabolic reprogramming preceding shifts in the gene expression program, rather than being a consequence of it.

It is not intuitive how a general system (such as metabolism) can drive changes in the highly specific genetic programs expressed by individual cells. A possible way is through chromating remodeling: when metabolic reprogramming takes place, chromatin opens allowing for the access of transcription factors to DNA. These transcription factors are expressed heterogeneously in individual cells, so their relative abundances would determine the transcription program that gets executed. There is evidence sporting the plausibility of this mechanism: studies have shown that the use of molecules affecting the epigenetic machinery that controls chromatin dynamics can increase the efficiency of somatic cell reprogramming (which requires dramatic changes in the gene expression program) by up to 100 times.[132]

There are many examples of signaling pathways that involve mitochondria and control gene expression and cell function. One of them is glucocorticoid signaling, that involves the translocation of the glucocorticoid/receptor complex to mitochondria to modulate its function.[133] Another one is the Notch pathway that regulates cell proliferation, differentiation and death in metazoans and is specially relevant in cancer stem cells.[134] Notch proteins interact with a kinase (PINK1) to modulate mitochondrial function and possibly raise ATP levels. A third example are thyroid hormones, signaling molecules that stimulate mitochondrial respiration on a timescale that comprises a few

minutes/hours. Several days after this fast response, changes in the expression of target genes lead to sustained mitochondrial biogenesis and an induction of ATP increase.[135]

Mitochondria and pathology

Mitochondria has regulatory effects on gene expression. It is then straightforward to derive that its malfunction will result in aberrant gene expression, potentially manifesting in the form of disease. Indeed, the pathophysiology of many conditions has been associated with dysfunctional mitochondria.[136]

We have previously discussed the role of mitochondria in alternative splicing. AS is critically implicated in genetic diseases (in humans, ~50% of them arise from mutations affecting AS[137]) and heart diseases like hypertrophic cardiomyopathy or sudden death.[138,139] This is consistent with the idea of aberrant AS being related to mitochondrial dysfunction, as cardiac tissues have a high dependence on ATP. In fact, the progression of cardiopathies to heart failure (HF) is always associated with a decrease in ATP (up to 40%)[140] and mitochondrial malfunction.[141] Furthermore, therapies aimed to improve mitochondrial functionality have been shown to be effective at increasing the long-term survival rates of patients with chronic HF.[142] Other contexts where aberrant AS has been reported are neurodegenerative (Alzheimer, Parkinson...) and neurodevelopmental (autism...) diseases.[143]

In cancer, the interplay between metabolism and gene expression is critical. Nutrient uptake in animal cells requires growth factors whose concentration can be limiting in the environment. In such cases, differentiated cells adopt a metabolic strategy based on oxidative phosphorilation in mitochondria to maximize the efficiency in the production of ATP. On the other hand, under abundance of growth factors cells can switch to an anabolic strategy, increasing nutrient uptake and biomass production. In cancer, growth factor signaling is permanently activated and metabolism is reprogrammed to a glycolytic phenotype.[66,75] In many cancer types, mitochondria seem to be reprogrammed for macromolecular synthetic activity.[144] It is generally accepted that these metabolic changes are originated by aberrant gene expression.[145,146] However, it has been argued that non-genetic heterogeneity can contribute to the somatic evolution of tumors.[67] Hence, it is possible that metabolic reprogramming precedes some of the genetic mutations associated with the disease, instead of simply resulting from them. Furthermore, aberrant AS is also a hallmark of tumor cells,[147] adding another potential link between mitochondrial function and cancer.

2

Mitochondrial control of gene expression

Gene expression is heterogeneous from cell to cell. As we have seen, this has many important functional implications, so identifying its sources and quantifying their contributions is a relevant problem. In this chapter we will explore some of the mechanistic aspects through which mitochondria modulate variability in gene expression, a highly energy demanding process. We will do so with a combination of mathematical models and statistical analysis of experimental data, notably RNA sequencing (RNA-seq) data. The advantage of RNA-seq is that it provides straightforward information about changes in the activity of specific genes. In addition, we have seen that a great fraction of the noise in protein abundance is generated at the mRNA level through transcriptional bursting. This justifies the study of the transcriptome to understand the connections between mitochondrial and phenotypic variability.

2.1 Global and specific constraints on gene expression

Global scaling

In response to changes in the external conditions, cells can alter the expression of thousands of genes.[148] This response can be gene-specific in some cases, but there is evidence that, for the most part, it is the consequence of changes in cellular factors affecting several genes. For instance, genome-wide measurements in yeast and *E. coli* revealed that promoter activities across different growth conditions are shifted by a global scaling factor for 60-90% of promoters.[149]

As for eukaryotes, we have discussed that cells with more mitochondrial mass have increased biosynthetic capabilities. It is then likely that differences in mitochondrial

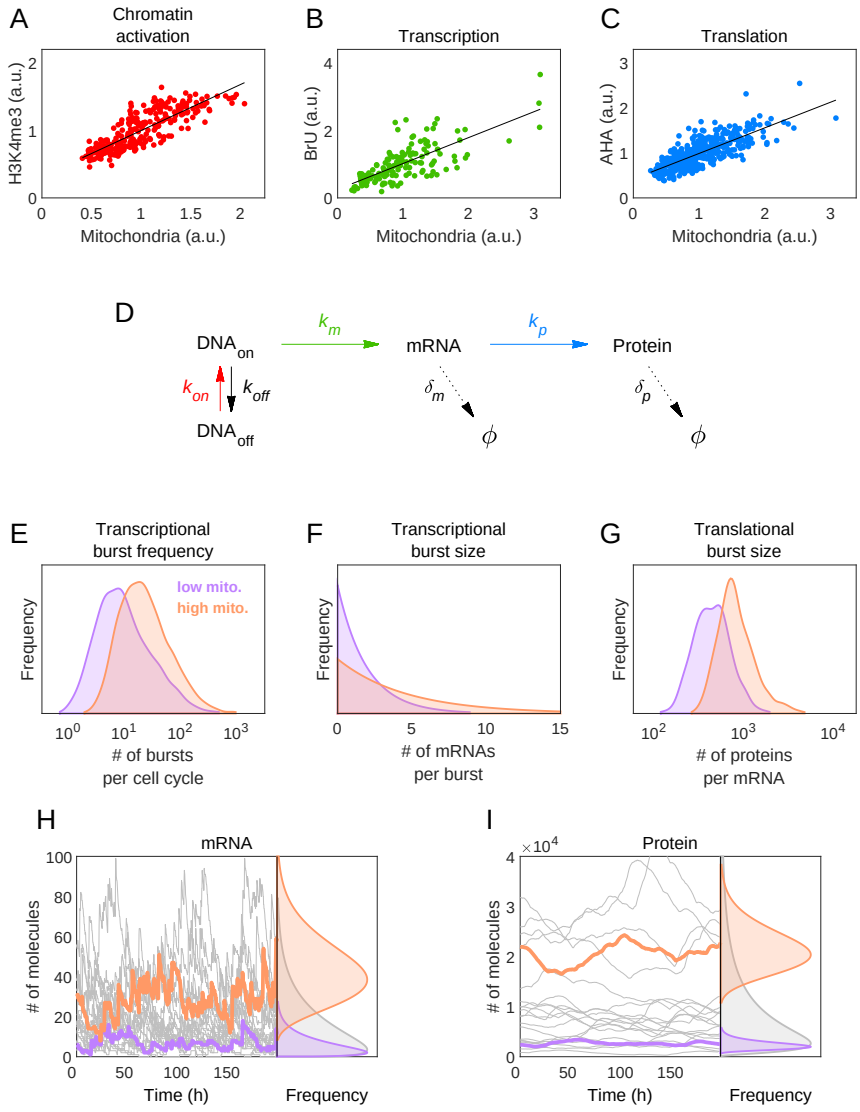

A
Chromatin activation
H3K4me3 (a.u.)
Mitochondria (a.u.)
B
Transcription
BrU (a.u.)
Mitochondria (a.u.)
C
Translation
AHA (a.u.)
Mitochondria (a.u.)
D
DNA on
DNA off
k on
k off
k m
mRNA
δ m
ϕ
k p
Protein
δ p
ϕ
E
Transcriptional burst frequency
low mito.
high mito.
Frequency
# of bursts per cell cycle
F
Transcriptional burst size
Frequency
# of mRNAs per burst
G
Translational burst size
Frequency
# of proteins per mRNA
H
mRNA
# of molecules
Time (h)
Frequency
I
Protein
# of molecules
Time (h)
Frequency

Fig. 2.1. Effect of mitochondria on gene expression variability in a central dogma model. Mitochondrial levels correlate with the amounts of: **A.** the histone methylation mark H3K4me3, a proxy for chromatin activation, **B.** nascent RNA as reported by BrU, related to transcriptional activity, and **C.** nascent protein quantified by the precursor AHA, an indicator of translational activity. Each dot represents a cell and solid lines are linear fits. Coefficients of variation: $CV_{H3K4me3} = 0.29$, $CV_{BrU} = 0.54$ and $CV_{AHA} = 0.32$. **D.** Basic *central dogma* model used for stochastic simulations. Parameter dependencies are color-coded in reference to panels A-C: promoter activation rate (k_{on}), transcription rate (k_m) and translation rate (k_p) change from cell to cell with mitochondria according to experimental measurements of variability in H3K4me3, BrU and AHA marks respectively. Parameter values: $\langle k_{on} \rangle = 0.4\text{h}^{-1}$, $k_{off} = 10\text{h}^{-1}$, $\langle k_m \rangle = 35\text{h}^{-1}$, $\langle k_p \rangle = 15\text{h}^{-1}$, $\delta_m = 0.08\text{h}^{-1}$, $\delta_p = 0.03\text{h}^{-1}$. **E.** We analyzed two subpopulations of simulated cells, with low (m within the 30% bottom values, purple) and high (m within the 30% top values, orange) mitochondrial content respectively. Fixing the values for k_m and k_p to their population averages ($\langle k_m \rangle$ and $\langle k_p \rangle$) for all cells but maintaining the co-variation of k_{on} with mitochondria, we can study the marginal effect of this dependency on gene expression: cells with higher mitochondrial content have increased transcriptional bursting frequency. **F.** Analogously, fixing k_{on} and k_p but maintaining the mitochondrial dependency of k_m reveals increased transcriptional burst sizes in the subpopulation with higher mitochondrial content. **G.** Last, fixing k_{on} and k_m and allowing cell-to-cell differences in k_p induced by mitochondria results in increased protein production in the subpopulation with higher mitochondrial levels. **H-I.** Trajectories for the mRNA and protein abundances of several cells (gray), out of which two have been highlighted (low mitochondria, purple; high mitochondria, orange). Distributions were obtained from mRNA and protein levels of the whole population (gray) or the "low"/"high" subpopulations (purple/orange). Coefficients of variation for mRNA abundance distributions: $CV_{mRNA} = 0.89$, $CV^{(low)}_{mRNA} = 0.83$, $CV^{(high)}_{mRNA} = 0.37$, and for protein abundance distributions: $CV_{Protein} = 1.09$, $CV^{(low)}_{Protein} = 0.37$, $CV^{(high)}_{Protein} = 0.21$.

content can cause changes in the expression of many genes. Indeed, it has been observed that the transcriptional activity of HeLa cells scales by a global factor across normal and ATP-depleted populations. Similar experiments also showed an analogous scaling across subpopulations of cells sorted by their mitochondrial content.[69] Interestingly, not only the average transcriptional activity but also its variability seems to be maintained across subpopulations with different mitochondrial content, and even in ATP-depleted cells. This suggests that an increased energy budget simultaneously boosts the average amount of RNA/protein and the deviations around this average at the population level.

To investigate the mechanistic aspects of this global scaling, we developed a model for the expression of a single gene based on the *central dogma* of molecular biology, which can be phrased as a series of biochemical reactions involving gene activation/inactivation, transcription initiation/elongation to synthesize mRNA strands, and translation initiation/elongation to synthesize proteins. Some of these processes are highly dependent on ATP and thus on mitochondrial content: immunofluorescence experiments where HeLa cells were tagged with MitoTracker green (for mitochondrial mass quantification) and marks for histone methylation, transcriptional activity or trans-

lational activity evidence a strong correlation of all said marks with mitochondrial levels, as shown in figure 2.1A-C.

In our model, outlined in figure 2.1D, we account for these correlations by introducing dependencies of several kinetic rates with mitochondria. Specifically, if we denote the mitochondrial content of a cell (in arbitrary units) as m, we have:

$$\begin{aligned} k_{on} &= \langle k_{on} \rangle \cdot f_1(m) \\ k_m &= \langle k_m \rangle \cdot f_2(m) \\ k_p &= \langle k_p \rangle \cdot f_3(m) \end{aligned} \tag{2.1}$$

where k_{on} is the gene activation rate, k_m the transcription rate and k_p the translation rate of that cell. Angle brackets indicate population averages of the rates, which are fixed to values within biological ranges[107, 150] (see caption of figure 2.1). The values of f_1, f_2 and f_3 are log-normally co-sampled[†] with m, with mean equal to 1 and standard deviation that complies with the variability observed experimentally for histone methylation, transcriptional activity and translational activity respectively (fig. 2.1A-C). This choice is justified because the distributions of mitochondria as well as of the marks used for the mentioned processes are well fitted by log-normal distributions (χ^2-test for goodness of fit agreed with a log-normal to a $< 5\%$ significance in all cases). Other kinetic rates (gene inactivation rate k_{off} and degradation rates, δ_m and δ_p for mRNA and protein respectively) are assumed to be independent of mitochondria.

A deterministic approach can be used to formulate the dynamics of the species' average copy numbers in terms of a set of ODEs:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\overline{DNA_{on}} &= k_{on}(m) \cdot \left(1 - \overline{DNA_{on}}\right) - k_{off} \cdot \overline{DNA_{on}} \\ \frac{\mathrm{d}}{\mathrm{d}t}\overline{mRNA} &= k_m(m) \cdot \overline{DNA_{on}} - \delta_m \cdot \overline{mRNA} \\ \frac{\mathrm{d}}{\mathrm{d}t}\overline{Protein} &= k_p(m) \cdot \overline{mRNA} - \delta_p \cdot \overline{Protein} \end{aligned} \tag{2.2}$$

where overlines indicate time averages of species' abundances within an individual cell with mitochondrial mass m.

We simulated a population of cells by first assigning each one of them a mitochondrial content sampled from a log-normal distribution of mean 1 and dispersion equal to the experimental distribution of MitoTracker green ($\mathrm{CV}_m = 0.4$, see figure 1.4). Each cell was then given individual values for k_{on}, k_m and k_p according to expression 2.1, thus accounting for the extrinsic cell-to-cell fluctuations induced by mitochondria. The dynamics of the gene expression cycle were simulated using the Gillespie algorithm to account for intrinsic noise as well. Sources of variability other than mitochondria and

[†] *Log-normal co-sampling* is the term we use to describe the sampling of a variable y that correlates with another variable x whose value is set beforehand, when both are log-normally distributed. See appendix B for more information.

intrinsic fluctuations are indirectly included in the model: rate values are sampled according to the cell-to-cell differences observed experimentally (equations 2.1 and figure 2.1A-C), where off-diagonal deviations are presumably caused by a non-characterized combination of factors.

The model reveals increased frequency and size of transcriptional bursts with mitochondrial levels through the modulation of the promoter activation and transcription rates respectively. Analogously, mitochondrial control of the translation rate leads to higher proteins produced per mRNA strand under conditions of high mitochondrial content. The combined effect of all three parameters yields increased mRNA and protein copy numbers in cells with high mitochondrial content, consistent with experimental findings of gene expression globally scaling with energy budget.[68,69]

Non-linearities and gene-specific regulation

The central dogma model we described is able to recapitulate some of the effects of mitochondria on gene expression, namely the global scaling of mRNA and protein abundances, but some features of the model are not in agreement with experimental findings:

- Experiments show a roughly linear co-variation of the per-cell protein abundance with mitochondria[68,69] (fig. 1.5), but simulations of the central dogma model yield a high degree of non-linearity (fig. 2.2C).
- The overall cell-to-cell variability in protein abundance is not quantitatively consistent with experimental data. The coefficient of variation of experimental protein distributions is much smaller (CV $\sim$ 0.4, figure 1.5) than the one obtained through simulations (CV $\sim$ 1.1, figure 2.1I).
- In the model, variability in protein levels does not scale accordingly to empirical observations: experiments simultaneously measuring mitochondrial and protein abundances in single cells show that the CV of the protein and mRNA distributions varies only slightly across subpopulations with different mitochondrial content.[68,69] But the model yields significant differences in the CV of the mRNA and protein distributions of such subpopulations.

These discrepancies are likely due to an oversimplification of the model. Gene expression is complex, and by describing transcription or translation with a single reaction rate (k_m and k_p) we have gathered together a collection of steps with many potential limiting factors. To investigate this, we developed a new model for the expression of a single gene, this time accounting for the intermediate steps of polymerase and ribosome binding and unbinding to DNA and mRNA respectively (fig. 2.2A). We assume that this binding/unbinding is a relatively energy-independent process, so we include no cell-to cell variation in the kinetic rates that characterize it (α, α_r, β and β_r). On the

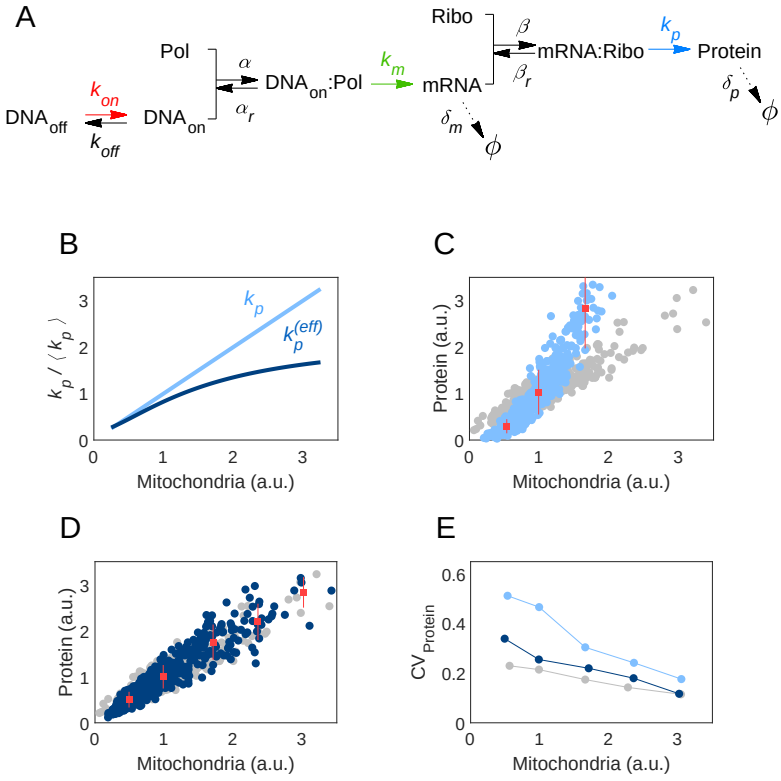


Fig. 2.2. Non-linearities and constraints on the gene expression cycle. A. Multi-step model explicitly including binding/unbinding of polyerases and ribosomes. Some parameters are taken as mitochondria-dependent (k_{on}, k_m and k_p, color coded according to figure 2.1A-C), while others are assumed to be independent of it (degradation rates and polymerase/ribosome binding/unbinding rates). Parameter values: $\langle k_{on} \rangle = 40\text{h}^{-1}$, $k_{off} = 1\text{h}^{-1}$, $\alpha = 1000\mu\text{m}^3\text{h}^{-1}\text{molecules}^{-1}$, $\alpha_r = 3500\text{h}^{-1}$, $\langle k_m \rangle = 35\text{h}^{-1}$, $\beta = 15\mu\text{m}^3\text{h}^{-1}\text{molecules}^{-1}$, $\beta_r = 0.15\text{h}^{-1}$, $\langle k_p \rangle = 15\text{h}^{-1}$, $\delta_m = 0.08\text{h}^{-1}$, $\delta_p = 0.01\text{h}^{-1}$, $Pol = 50$, $Ribo = 1000$, $v = 1000\mu\text{m}^3$ (cellular volume). **B.** Including these intermediate steps is analogous to having "effective" transcription and translation rates that vary non-linearly with mitochondrial mass. Light blue represents a linear co-variation of the translation rate k_p according to the scenario in the basic central dogma model. Dark blue represents a non-linear co-variation of the analogous effective rate in the multi-step model. **C.** The basic central dogma model yields a non-linear covariation of mitochondria and protein (light blue dots), as opposed to experimental measurements

(gray dots). Simulated cells were split into "bins" according to their mitochondrial content. The average mitochondria and protein levels of each bin are represented in red (bar indicate standard deviation within bins). **D.** On the other hand, the multi-step model (dark blue dots) reproduces the experimental trend. **E.** For both the basic (light blue) and multi-step (dark blue) models, as well as for experimental data (gray), cells were binned and the variability within each bin was quantified as the CV.

other hand, elongation rates (k_m and k_p) are co-sampled with each cell's mitochondrial content, analogously to what was done in the simplified central dogma model.

Again, we can use mass action kinetics to derive equations for the average number of molecules of the species involved. A quasi-steady state approximation for the complexes DNA_{on}:Polymerase and mRNA:Ribosome is appropriate if $\alpha, \alpha_r \gg k_m$ and $\beta, \beta_r \gg k_p$, i.e. if the binding/unbinding of polymerases and ribosomes is much faster than the elongation step, which seems to be the case in real cells.[107,108] Additionally, since we are not including explicit polymerase and ribosome production/degradation, their number must be conserved: Pol + DNA_{on}:Pol = *const*. and Ribo + mRNA:Ribo = *const*. Assuming that most polymerases and ribosomes are free in the cell (not engaged to DNA/mRNA), which is coherent with the current understanding,[151] we can finally write:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\overline{DNA_{on}} &= k_{on}(m)\cdot\left(1-\overline{DNA_{on}}\right) - k_{off}\cdot\overline{DNA_{on}} \\
\frac{\mathrm{d}}{\mathrm{d}t}\overline{mRNA} &= \alpha\,\frac{k_m(m)/\alpha_r}{1+k_m(m)/\alpha_r}\cdot Pol\cdot\overline{DNA_{on}} - \delta_m\cdot\overline{mRNA} \\
\frac{\mathrm{d}}{\mathrm{d}t}\overline{Protein} &= \beta\,\frac{k_p(m)/\beta_r}{1+k_m(m)/\beta_r}\cdot Ribo\cdot\overline{mRNA} - \delta_p\cdot\overline{Protein}
\end{aligned} \tag{2.3}$$

and defining effective transcription and translation rates as

$$\begin{aligned}
k_m^{(eff)} &\equiv \alpha\,\frac{k_m(m)/\alpha_r}{1+k_m(m)/\alpha_r}\cdot Pol \\
k_p^{(eff)} &\equiv \beta\,\frac{k_p(m)/\beta_r}{1+k_m(m)/\beta_r}\cdot Ribo
\end{aligned} \tag{2.4}$$

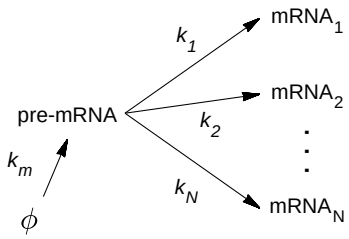
we can cast the equations in a form that is analogous to 2.2:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\overline{DNA_{on}} &= k_{on}(m)\cdot\left(1-\overline{DNA_{on}}\right) - k_{off}\cdot\overline{DNA_{on}} \\
\frac{\mathrm{d}}{\mathrm{d}t}\overline{mRNA} &= k_m^{(eff)}(m)\cdot\overline{DNA_{on}} - \delta_m\cdot\overline{mRNA} \\
\frac{\mathrm{d}}{\mathrm{d}t}\overline{Protein} &= k_p^{(eff)}(m)\cdot\overline{mRNA} - \delta_p\cdot\overline{Protein}
\end{aligned} \tag{2.5}$$

The effective transcription and translation rates of the new model, $k_m^{(eff)}$ and $k_p^{(eff)}$, are dependent on mitochondrial content but in a non-linear way, as shown in figure 2.2B. This is interpreted as the binding/unbinding steps being potentially limiting: even if a cell has enough energy to increase elongation rates, transcription and translation initiation rely on the random (not or weakly ATP-dependent) attachment of polymerases and ribosomes to DNA and mRNA. This can result in overall transcription and translation rates saturating and thus remaining relatively low even under conditions of ATP excess.

Running Gillespie simulations of this model we reconstructed an *in-silico* population of cells and quantified protein abundances and variability levels. We also sorted the cells into 5 smaller subpopulations according to their mitochondrial content, and measured the CV of the protein distributions within each subpopulation. First, we found that this new, multi-step model can recapitulate the linear co-variation of protein and mitochondria abundances (fig. 2.2D), as opposed to the basic central dogma model (fig. 2.2C). Furthermore, the variability at the protein level within the sorted subpopulations is better captured by the new model, especially at low mitochondrial content (fig. 2.2E). These results indicate that the effect of mitochondria on the gene expression cycle goes beyond a simple, linear modulation of the kinetic rates. Even though such linear modulation can explain some observations (e.g. the increase of gene products with mitochondrial mass), the complexity of gene expression and the many layers of regulation that it entails introduce a set of energetic dependencies and limiting factors. The interplay between them determines the level of expression of individual genes.

Our model suggests that the availability of polymerases and ribosomes can be a factor modulating the effect of energy budget, but many other potentially limiting elements could be taking part. For instance, the model does not account for a possible role of mitochondria on transcript or protein degradation rates, neither it includes the post-transcriptional modifications stage between mRNA production and protein synthesis. Even though the effect of alternative splicing is neglected here, it has been shown to be critical. Guantes et al. proposed a simple two-step process in which a mRNA was produced at a constant rate k_m, and then spliced into its N mature forms with rates k_1, k_2, ..., k_N. They showed that in order to explain the variability observed in transcript relative abundances across populations with different mitochondrial content, both the pre-mRNA production rate and the maturation rates had to be modulated by mitochondria.[69]



2.2 Connecting mitochondrial and transcriptional variability

Since we ultimately want to understand the connections between energy budget, gene expression and phenotypic variability, it would make sense to analyze the changes in protein abundances in cells with different mitochondrial mass. Proteins are the molecular machines that carry out specific tasks, and thus their dynamics determine each cell's phenotype. The correlation between the transcriptome and the proteome of individual cells has been shown to be relatively weak,[150] likely due to the abundance of post-transcriptional control mechanisms. However, it is generally accepted that the initiation of transcription (primary mRNA production) is the predominant form of gene expression regulation,[28–39,94,98,152] and transcriptomic analyses are still widely used to study changes in gene expression. In our case, analyzing the variability the levels of RNA expression through sequencing is justified since they represent a direct proxy of gene activity and the most relevant form of noise generation.[30–34,101,102]

2.2.1 RNA-seq and data processing

Cells stained with MitoTracker green were sorted into two subpopulations with high and low mitochondrial content respectively. The difference in mitochondrial mass across subpopulations was about 5-fold. From each population, RNA was extracted, purified and sequenced. This procedure was repeated for three cell lines: HeLa (3 biological replicates), Jurkat (3 biological replicates) and MRC-5 (2 biological replicates). More information on experimental methods such as cell culture, sorting, RNA extraction, etc. can be found in appendix A.

HeLa is the most commonly used human cell line in biological research. It is derived from cervical cancer.

Jurkat is a strain of human T lymphocytes originally obtained from the blood of a patient with T cell leukemia.

MRC-5 are fibroblasts derived from lung tissue of a human fetus.

Sequenced reads were first passed to FastQC v0.11.8[153] for quality check, finding no significant presence of adapter sequences in any sample. From sequenced reads, transcript abundance was quantified using the quasi-mapping mode of *Salmon* v0.12.0[154] with default settings, and mappings were validated using the alignment-based mode. The *Salmon* index was built using the cDNA file of the Homo sapiens genome, version GRCh38 from Ensembl.[155] Downstream analyses were performed using R v3.5.2.[156] Additional data was retrieved from the Ensembl BioMart tool with the aid of the *biomaRt*

v2.38.0 package.[157,158] Transcript counts were aggregated at the gene level using the *tximport* v1.10.1 package.[159] Differential expression analyses were done with the *DESeq2* v1.22.2 package.[160]

Expression levels were obtained in units of TPM (*transcripts per million*). TPM values in the "low" mitochondria condition were corrected by a strain-specific factor to account for the global differences in per-cell RNA abundance across subpopulations with high and low mitochondrial mass. To obtain this factor, cells were stained with MitoTracker green and the mRNA content per cell was checked by poly(T) mRNA FISH. Fluorescence images were analyzed with the ImageJ software.[161] Cells with high and low mitochondrial content were selected and the average poly(T) intensity of both subpopulations was quantified. The ratio ("low"/"high") between said intensities (equal to 0.37 in HeLa) was used to scale TPM values in the "low" condition.

Transcripts and genes expressed under a threshold were discarded. This threshold (*detection limit*, DL) was estimated as follows:[162] all features (transcripts and genes) with at least one zero and one non-zero expression value in any condition were selected. All non-zero values of this subset were listed, and the DL was taken as the median of their distribution (0.5TPM).

A note on units of expression

Understanding the units of transcript expression is key to perform consistent downstream analyses. In general, RNA-seq experiments do not provide a quantification of the RNA copy numbers per cell. The sequenced RNA is typically obtained from populations of cells (although modern single-cell sequencing techniques also exist), with a fixed library size (i.e. number of reads to be sequenced). Thus, when quantifying the expression of a given transcript from RNA-seq data, we are usually studying what *fraction* of the sequenced reads comes from each type of transcript in the sample, but obtaining absolute copy number requires additional steps. One example is the inclusion of spike-ins, transcripts of known length and well characterized quantity used to calibrate measurements in RNA-seq and other similar assays.

The most straightforward way to report the expression of a transcript is to simply give the raw number of reads that were aligned to it, that is, that came from the processing of a transcript of that type. Some common bioinformatic tools use this as input. However, there are important problems when attempting to relate these raw counts to the true level of expression of a transcript (namely the per-cell copy number):

- RNA-seq entails the breakdown of transcripts into fragments for sequencing, so longer transcripts produce increased raw fragment counts. This bias is particularly important when comparing the level of expression of two transcripts within a same sample.
- Raw fragment counts scale with library size.

To enable comparisons across different transcripts in a sample and across samples with different library sizes, these effects need to be taken into account. If we denote the number of counts of the transcript i as x_i, a straightforward way to remove the effect of the library size is to use *counts per million* (CPM):

$$CPM_i = \frac{x_i}{X} \cdot 10^6 \tag{2.6}$$

being X the total number of reads. CPMs are still biased for the transcript length. In fact, the magnitude of the bias for the i-th transcript is determined by the so-called *effective length* $\tilde{l}_i$, computed as:

$$\tilde{l}_i = l_i + 1 - \mu \tag{2.7}$$

where l_i is the true length of the transcript and μ is the mean of the fragment length distribution of the library. The effective length is interpreted as the number of possible start sites at which a transcript could have generated a fragment of that length. Normalizing the raw counts by this effective length as well as for the total number of reads gives *fragments per kilobase of exon per million reads* (FPKM):

$$FPKM_i = \frac{x_i}{X \cdot \tilde{l}_i} \cdot 10^9 \tag{2.8}$$

Finally, if the normalization by the library size is done using counts scaled by their effective length (instead of raw counts), *transcripts per million* (TPM) are obtained:

$$TPM_i = \frac{x_i}{\tilde{l}_i} \cdot \left(\frac{1}{\sum_j x_j / \tilde{l}_j} \right) \cdot 10^6 = \frac{FPKM_i}{\sum_j FPKM_j} \cdot 10^6 \tag{2.9}$$

TPMs are, in principle, unbiased with respect to transcript length and library size.[163] They simply represent the fraction of transcripts of each type within a pool of RNAs (times a factor 10^6). TPMs are still not equivalent to transcript copy numbers, but independent experiments quantifying total RNA per cell can be used to make the conversion.[150,164]

A note on global scaling

When performing bulk RNA-seq, RNA is extracted from cell populations. This inevitably implies a loss of information on the states of single individuals. In our case, we know that each cell's mitochondrial content determines the abundances of RNA (as well as protein and other components), but these differences are not captured by bulk RNA-seq data. To minimize this effect, we scaled TPM values in the "low" condition by a factor extracted from independent experiments quantifying total RNA per cell. By doing so, we are neglecting the cell-to-cell differences within our subpopulations (with "high" or "low" mitochondrial content). We assume that the introduction of this global

scaling factor will yield a good estimation of the per-cell content of individual transcripts, but this is not always necessarily the case: potential sources of variability that are independent of mitochondria could induce significant cell-to-cell differences in the copy numbers of specific RNAs. For them, averaging over the whole subpopulation would provide an inaccurate description of individual cells.

In summary, even though RNA-seq is a powerful tool to investigate the effect of a variable on the gene expression landscape, we need to be aware of the bias introduced by assuming that all cells within an ensemble are identical. On the other hand, single-cell transcriptomics (which in principle bypasses this issue) has a high degree of experimental variability because of the need to sequence very small amounts of genetic material, that then has to be amplified in inherently noisy protocols.

2.2.2 Expression changes at the gene level

We performed a principal component analysis (PCA) on the expression data using samples as variables and genes as observations. We plotted the coordinates of each sample in the space of principal components, restricted to the plane formed by the first two (PC_1 and PC_2). For the cellular strains with three biological replicates (HeLa and Jurkat), the one with the highest average distance to the other two in the PC_1-PC_2 plane was discarded, and downstream analyses were carried out with two replicates per strain.

After this, a new PCA was performed with the 12 remaining samples (3 cell lines, 2 conditions —"high" and "low" mitochondria— and 2 biological replicates each, figure 2.3A). We computed the average distance across data points of different strains, finding higher values for the MRC-5 samples. This is consistent with the fact that MRC-5 cells are fibroblasts, as opposed to HeLa and Jurkat which are cancer strains. Thus, it is reasonable to expect more similarities between the last two in terms of their gene expression program. We then proceeded to identify individual genes with significant changes in their expression levels, and we checked whether such genes were conserved across strains. We selected those genes that were significantly up- or down- regulated on each strain, i.e. with an adjusted p-value < 0.05 (obtained from *DESeq2* differential expression analyses) and a fold-change (FC) either greater or smaller than the global scaling factor given by the poly(T) RNA FISH experiments, as shown in figure 2.3B for HeLa cells. For this analysis, the p-value was computed using the raw TPM data, that is, before correcting values by the factor accounting for the per-cell total RNA differences. Doing this, we identify the genes that are *significantly expressed above or below the global scaling*. Formally speaking many more genes have their expression significantly changed across conditions due to the difference in total expression induced by mitochondrial levels variation.

Figure 2.3C evidences a small degree of overlap across differentially expressed genes in distinct strains: in HeLa cells, 117 genes matched our criteria for up-/down-regulation; 171 did so in Jurkat and 87 in MRC-5. Out of them, only 4 genes were common to HeLa

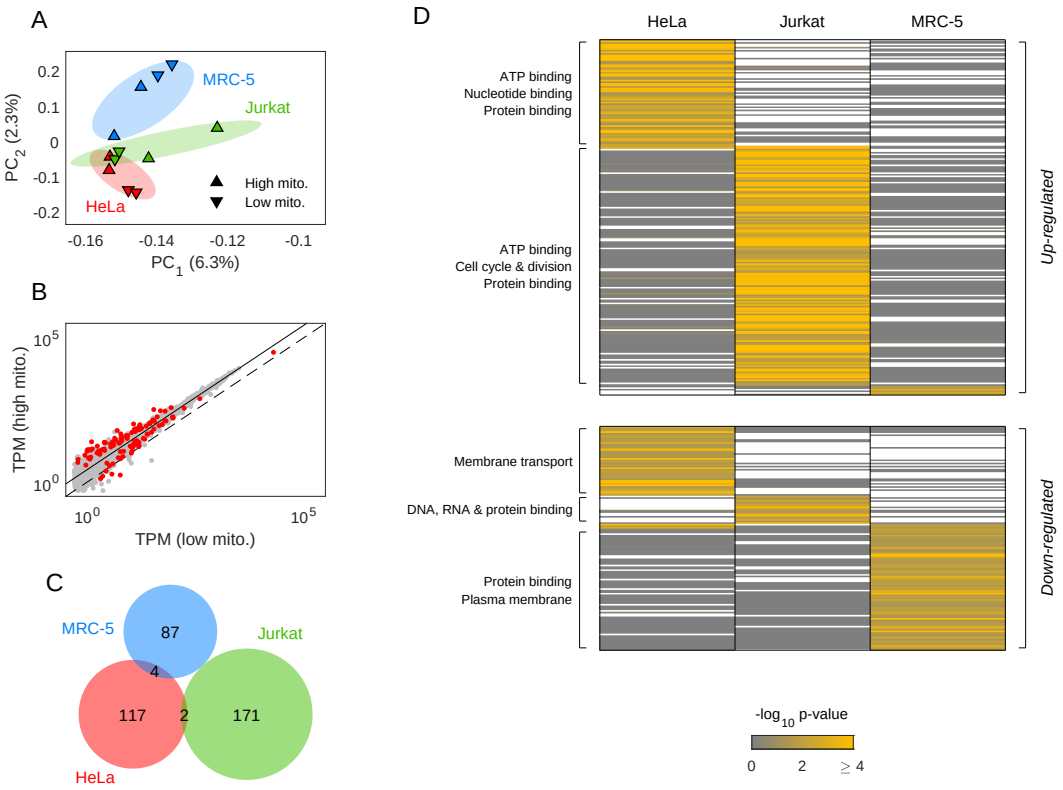


Fig. 2.3. Global effect of mitochondria on gene activity across cellular strains. A. Principal Component Analysis of the gene expression data in three cell lines: HeLa (red), Jurkat (green) and MRC-5 (blue). Two biological replicates per strain are shown. Up-pointing triangles correspond to samples with high mitochondrial content and vice-versa. Numbers in parenthesis on the axes indicate the percentage of variability explained by the first and second principal components (PC_1 and PC_2 respectively). **B.** Gene levels of expression in HeLa cells in "high" versus "low" mitochondria conditions. Values for the "low" condition have been scaled by the global factor obtained from poly(T) RNA FISH experiments, which results in the best-fit line (solid) being above the identity line (dashed). Red dots represent genes that were differentially up- or down-regulated between conditions. **C.** Overlap across the three strains in the lists of genes differentially expressed. **D.** Significance of the change in expression between the "high" and "low" conditions of a collection of genes. Each row represents a gene. Only genes that were significantly up- (upper block) or down-regulated (lower block) in at least one cellular strain are shown. Blank cells indicate that the level of expression of the gene was below detectable limits in that specific strain. Genes differentially expressed in each strain were grouped and passed to *GeneCoDis3* for annotation enrichment analysis. Groups and annotations most significantly over-represented within each are shown on the left side.

and MRC-5, 2 were common to HeLa and Jurkat and there was no overlap between the Jurkat and MRC-5 gene sets. Furthermore, many of the genes with significant changes in expression in one strain were not expressed at detectable levels in the other two, and even those that were showed strictly non-significant changes between the "high" and "low" conditions as indicated in figure 2.3D.

Even if there is little conservation in the specific up- and down-regulated gene sets across strains, there could still be pathways or functions commonly affected in all of them. To explore this, we passed the lists of genes differentially expressed in each cell line to the *GeneCoDis3* tool.[165] *GeneCoDis3* extracts annotations from a number of databases such as Gene Ontology (GO),[166] KEGG pathways[167] or PANTHER pathways[168] and computes which ones are over-represented in a list of genes. It also groups annotations based on their co-occurrence within the provided gene list, which facilitates the interpretation of the results.

Notably, even though the HeLa and Jurkat strains share only a small fraction of up-regulated genes, the "ATP binding" annotation was found to be over-represented for both of them. This term is associated to genes whose products directly interact with ATP, for example DNA helicases (responsible for gene unpacking), RNA polymerases or enzymes involved in energy utilization such as kinases, which catalyze the transfer of phosphate groups from high-energy molecules to other substrates. It specifically applies to ATPases, enzymes that mediate the decomposition of ATP into ADP for energy release. This could indicate that increased mitochondrial content does not only rise the rate of ATP production, but also boosts the expression of proteins required for its utilization.

The "protein binding" annotation is also common to the HeLa and Jurkat up-regulated gene lists. Interestingly, it is also over-represented in the set of down-regulated genes of MRC-5, which could be associated with the fact that this last one is a normal cell line while the other two are cancer strains. It is plausible that this results from mitochondrial regulation of a large subset of proteins taking part in several pathways and whose expression is aberrant in cancer cells. This is, however, a very broad statement: "protein binding" is a generic term that could be associated to many genes participating on a plethora of different processes. To know which of those processes are indeed altered in cancer, and whether abnormal mitochondrial regulation of gene expression could be attributed to such alterations, more detailed mechanistic information would be required.

Other interesting annotations that were found are "cell cycle and division" in the Jurkat up-regulated gene list (many works have indeed characterized the relationship between mitochondrial content and cell cycle[68,69,96,125]) or "membrane transport" and "plasma membrane" in the down-regulated genes of HeLa and MRC-5 respectively. These last terms could refer to the electron transport chain in the mitochondrial inner membrane that results in ATP production during oxidative phosphorilation, or more generally to the regulation of the expression of membrane transport proteins, transmembrane polypeptides that facilitate the movement of charged and polar molecules in and out of the cell's lipid bilayers (including the mitochondrial membrane but not limited to it).

Only 6 genes were found to be significantly up-regulated in MRC-5 cells, with no statistically relevant over-representation of any function. It is worth noting that one of those genes is CDIP1, a target for p53 which is a widely studied tumor suppressor. CDIP1 is involved in the cellular response response to TNF, a ligand that induces extrinsic apoptosis. The role of mitochondria on the regulation of the expression of genes involved in extrinsic apoptosis will be further studied in the last chapter of this work.

2.2.3 Alternative splicing

As we have discussed, mitochondrial levels regulate many steps of the gene expression cycle. This means that its role goes beyond that of a global modulator, instead, there are many potential effects that could modulate transcript expression individually. Our RNA-seq data allows us to quantify the expression in "high" and "low" mitochondria conditions on a transcript-by-transcript basis, so we investigated the effect of mitochondria on their relative abundances in HeLa, Jurkat and MRC-5 cells.

First, we noticed significant variation at the gene and transcript levels as evidenced in figure 2.4A. We found a fraction of genes (15%-20% depending on the strain) for which at least one transcript had a fold-change significantly different from the gene's fold-change (meaning $\left|\log_2 \mathrm{FC}_t - \log_2 \mathrm{FC}_g\right| > 1$, with FC_t and FC_g the transcript and gene's fold-changes respectively). This means that, for many genes, mitochondria does not just scale the overall expression but it also changes the relative transcript abundances: regardless of whether the total expression of a gene scales according to the expected global trend, each of its transcripts individually can behave in a unique way. For some genes, the expression pattern remains unaltered and all transcripts maintain their relative abundances in the "high" and "low" conditions (fig. 2.4B). In other cases, one or more transcripts does not obey the global scaling (fig. 2.4C).

To investigate the mechanisms of this non-linear scaling, we studied each gene's *transcript expression pattern* (TEP). We classified a transcript as *non-linearly scaling* if it satisfied (a) adjusted p-value < 0.1 and (b) fold-change such that $\left|\log_2 \mathrm{FC}_t - \log_2 \mathrm{FC}_{global}\right| > 1$ between the "high" and "low" conditions. FC_{global} refers to the global scaling and is given by the independent poly(T) RNA FISH experiment as the inverse of the factor used to correct TPM values in the "low" condition.

Transcription start site

One of the determinants of TEP variation could be transcription start site (TSS) choice:[169] we have discussed that chromatin organization is highly ATP-dependent, which can lead to mitochondrial content modulating the availability of different polymerase binding sites.

To test this hypothesis, we selected those genes with at least one linearly and one non-linearly scaling transcript. We then looped through those genes, randomly selecting two transcripts per each (one linearly and one non-linearly scaling) and comparing their

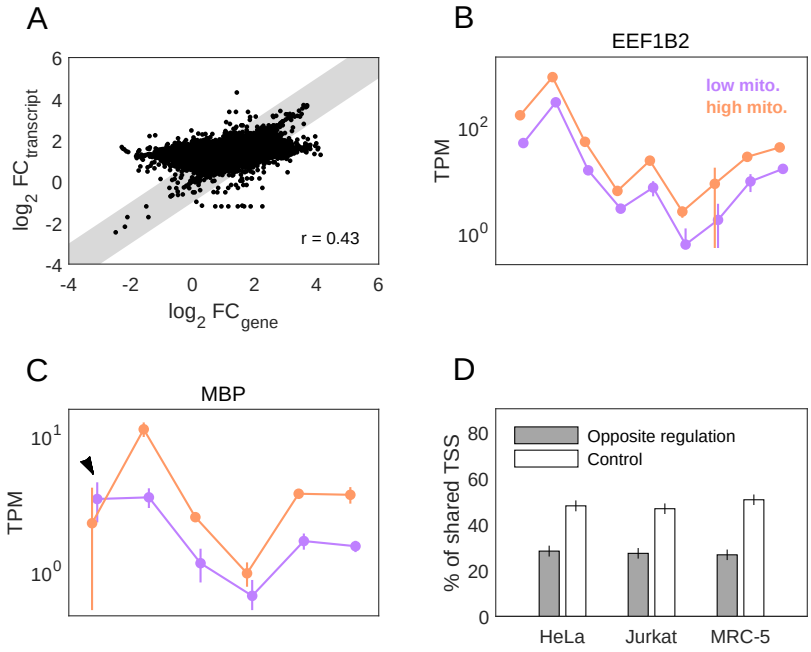


Fig. 2.4. Non-linear mitochondrial regulation of gene expression. **A.** Fold-change of the expression levels ("high"/"low") of genes (x axis) and their corresponding transcripts (y axis) in HeLa cells (Pearson's correlation coefficient $r = 0.43$). The shaded area corresponds to $|x - y| \leq 1$. Roughly 82% of the genes have all their transcripts into the shaded region, meaning there is an prominent fraction (∼18%) of genes for which at least one transcript with a FC that differs significantly from the gene's FC. **B-C.** Expression levels in the "low" (purple) and "high" (orange) mitochondria subpopulations are represented for the transcripts of two genes (EEF1B2 and MBP) in HeLa cells. Each dot is a transcript. For the first gene, expression increases with mitochondrial mass and the relative abundances of the transcripts are maintained. One of the transcripts of the second gene (indicated by the black arrow), on the other hand, shows an *inversion*: instead of following the global scaling trend, its expression is decreased in the "high" condition. Error bars are standard deviations across replicates. **D.** Fraction of shared transcription start sites across transcripts of the same gene with different regulation (gray) and for a null ensemble of transcripts chosen arbitrarily (white). Error bars represent standard deviations computed by bootstrapping.

TSS. We considered that the TSS were the same if they fell within 100bp of each other and vice-versa. As a control, we repeated this process but this time looping through all genes and selecting transcripts randomly, creating a "null" ensemble of TSS distances. We did this 20 times per strain for statistically meaningful results.

We found that for all three strains, roughly 50% of the transcripts of the same gene share their TSS, but this percentage goes down to ~30% when comparing transcripts with linear and non-linear scaling (fig. 2.4C). This shows that indeed TSS choice is, to a large extent, determining the mitochondrial control of TEPs. However, the fact that there is still a fraction of transcripts with different scaling that share the same TSS (or at least their TSS lie very close to each other) means that there must be additional layers of regulation.

Secondary structure... and transcript length?

One mechanism that has been proposed for mitochondrial control of aternative splicing is the variation of elongation speed leading to different precursor mRNA secondary structures[69,72] (fig. 2.5A-B). Secondary structure has been long regarded as a key determinant of splicing form choice.[170,171]

Addressing whether such a mechanism is plausible requires the systematic computation of pre-mRNAs secondary structures. There are many tools that aim to predict RNA folding based on nucleotide sequence.[172] In general, they work under the premise that the spatial configuration of a RNA molecule should minimize the free energy, although newer methods with different approaches and increased accuracy also exist.[173] Formally speaking, genes coding for pre-mRNAs with at least two possible secondary structures with similar free energies should be more prone to regulation through elongation speed variation. If the changes in the kinetic energy of elongating polymerases induced by mitochondrial variability are comparable to the differences in free energy between alternative pre-mRNA folding configurations, the mechanism would be consistent.

Yet addressing this question on a genome-wide scale would require an immense amount of computational resources. Additionally, the tools for secondary structure prediction generally have limited precision. A more convenient way to study this matter is by looking at transcript lengths. The number of possible secondary structures that a RNA can adopt grows exponentially with its length,[174,175] so we expect longer pre-mRNAs to have a higher chance of undergoing conformational changes induced by elongation speed variation.

We selected two subgroups of protein-coding genes with lengths within the top or bottom 10% of the complete set (fig. 2.5C). We then compared the fraction of genes within each subgroup that showed alterations in their transcript relative abundances (i.e. altered TEP) between the "high" and "low" mitochondria conditions (fig. 2.5D). We found no variation across subgroups in HeLa and Jurkat cells. This does not imply that there is no role of secondary structure in these strains but rather that, if there is, it is not

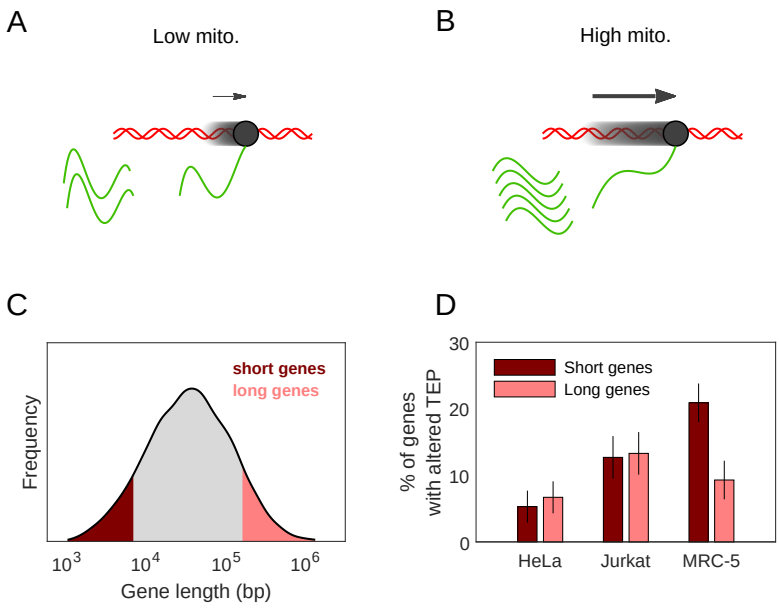


Fig. 2.5. Effect of gene length on mitochondrial modulation of expression. A-B. Mitochondrial content modulates elongation speed of polymerases (gray) through DNA strands (red), which is a determinant of mRNA (green) seconday structure. **C.** Distribution of expressed protein-coding gene lengths in HeLa. 10% shortest and longest genes are indicated in dark and light red respectively. **D.** Among the shortest (dark red) and longest (light red) genes, similar fractions have their TEPs altered by mitochondria in HeLa and Jurkat. Strikingly, in MRC-5 there is a higher fraction of genes with altered TEPs among the shortest ones, as opposed to our expectation based on the secondary structure hypothesis. Error bars were computed by bootstrapping.

prominent enough to be noticeable by just looking at precursor RNA lengths.

Surprisingly, MRC-5 cells behaved differently: non-linear mitochondrial regulation of transcript abundances was more prominent in genes with shorter lengths. It is likely that our assumption of pre-mRNA size being connected to secondary structure was too generic in this case: many other layers of energy-dependent regulation could be taking place, such as expression cost increasing with length, and neglecting them may be too rough of an approximation. In any case, the sole finding of the relationship between pre-mRNA length and mitochondrial regulation being strain-specific is remarkable and hints towards a mechanism that is less general than elongation speed variation. Further experimental work (e.g. including more cell lines or measuring the relative abundances of well-characterized, alternatively folded forms of specific pre-mRNAs) is required to characterize the effect, if any, of elongation speed on secondary structure.

2.3 Transcript-specific regulation of production and degradation rates

A possible mechanism for mitochondrial control of RNA abundance is the modulation of degradation rates. To study decay dynamics on a transcript-by-transcript basis, we performed RNA-seq on cell populations with different mitochondrial content at various times after transcription inhibition.

2.3.1 Time series RNA-seq and data processing

Cells stained with MitoTracker and sorted with respect to their mitochondrial content were treated with DRB for transcription blockage. DRB is an adenosine analogue that inhibits RNA polymerase II elongation[176] (see details in appendix A). After DRB treatment, RNA was extracted and sequenced every 2h (at 0, 2, 4 and 6h). Quantification and downstream analyses were performed according to section 2.2.1 of this chapter.

2.3.2 Quantification of transcript degradation rates

Consider a cell belonging to a subpopulation with either high or low mitochondrial content. That cell will be expressing a number of transcripts at different levels. Let $n_i(t)$ represent the number of transcripts of type i at a given time t. If transcription is blocked at $t = 0$, the abundances n_i for $t > 0$ will be determined by each transcript's degradation dynamics. Assuming all transcripts decay following an exponential law:

$$n_i(t) = n_{i,0} \exp(-\delta_i t) \tag{2.10}$$

where $n_{i,0}$ represents the abundance of the i-th transcript at $t = 0$, and δ_i its degradation rate. Note that δ_i is transcript-specific and can, in principle, vary under conditions of high and low mitochondrial content.

The total RNA in the cell, N, will be given by the sum of the abundances of all transcripts:

$$N(t) = \sum_i n_i(t) \tag{2.11}$$

The sum is carried over all transcript types expressed in the cell. Let us assume that $N(t)$ also decays exponentially when transcription is blocked:

$$N(t) = N_0 \exp(-\Omega t) \tag{2.12}$$

We have defined N_0 as the total RNA at time $t = 0$, and Ω as the *bulk degradation rate* describing the decay of the whole RNA pool. Just like the degradation rates for

individual transcripts, Ω can also be different across subpopulations with high or low mitochondrial content.

Equations 2.10 to 2.12 yield

$$\sum_i n_{i,0} \exp(-\delta_i t) \approx N_0 \exp(-\Omega t) \tag{2.13}$$

Both sides of equation 2.13 are not strictly equal, since a sum of exponentials is not necessarily an exponential itself. Yet it can serve as a good approximation under certain conditions, i.e. if the differences between the degradation rates of individual transcripts (δ_i) and the bulk degradation rate (Ω) are relatively small.

When working with RNA-seq data, we do not usually have direct information about the per-cell copy numbers of individual transcripts (namely the values of n_i). Instead, TPMs quantify the fraction of transcripts of a given type within a sample (scaled by a factor of 10^6). For the i-th transcript:

$$TPM_i(t) = f_i \frac{n_i(t)}{N(t)} \tag{2.14}$$

where f_i represents a scaling factor that is, in principle, simply equal to 10^6. However, our reasoning will hold even if f_i is different for every transcript, as long as it is time-independent. For $t = 0$, equation 2.14 becomes:

$$TPM_i(0) = f_i \frac{n_{i,0}}{N_0} \equiv TPM_{i,0} \tag{2.15}$$

Combining equations 2.10 to 2.15 we can arrive at:

$$TPM_i(t) = TPM_{i,0} \exp[-(\delta_i - \Omega) t] \tag{2.16}$$

It is useful to define a *relative degradation rate* for the i-th transcript, α_i, as

$$\alpha_i \equiv \delta_i - \Omega \tag{2.17}$$

The relative degradation rates can be positive or negative. The first case ($\alpha_i > 0$) implies that the i-th transcript degrades faster than the bulk and vice-versa. Equation 2.16 can be expressed as:

$$\log \frac{TPM_{i,0}}{TPM_i(t)} = \alpha_i t \tag{2.18}$$

Finally, calling

$$y_i(t) \equiv \log \frac{TPM_{i,0}}{TPM_i(t)} \tag{2.19}$$

turns equation 2.18 into the simple form:

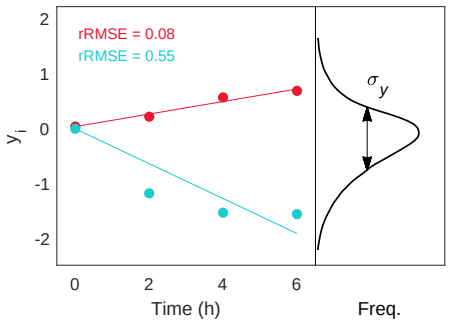


Fig. 2.6. Quantification of transcript relative degradation rates. Two transcripts have been arbitrarily chosen from our dataset out of all RNA types expressed in HeLa cells with low mitochondrial content. Dots represent the values for y_i (here, $i = 1, 2$ correspond to red and blue respectively) obtained from TPM values according to equation 2.19. Lines are fits, slopes being an estimate of each transcript's relative degradation rate (eq. 2.21). The distribution on the right panel (whose width σ_y is used to scale the RMSE of the fit, eq. 2.22) was obtained from all y_i non-zero values in the full dataset.

$$y_i(t) = \alpha_i t \tag{2.20}$$

Sequencing at different times after transcription blockage (DRB treatment), we can sample $\mathrm{TPM}_i(t)$ and thus $y_i(t)$. From the sequencing experiments we have four TPM values per transcript at times $t = 0\mathrm{h}, 2\mathrm{h}, 4\mathrm{h}, 6\mathrm{h}$, i.e. for the i-th transcript we can define a vector of observations $\mathbf{y_i} = \left(y_i^{(0\mathrm{h})}, y_i^{(2\mathrm{h})}, y_i^{(4\mathrm{h})}, y_i^{(6\mathrm{h})}\right)$. Note that in all cases $y_i^{(0\mathrm{h})} = 0$ because of the definition in equation 2.19. Fitting y_i versus t to a straight line that crosses the origin of coordinates, it is possible to get estimates for the relative degradation rates transcript-wise. The slope of this fit is given by[†]

$$\tilde{\alpha}_i = \frac{\mathbf{y_i} \cdot \mathbf{t}}{\mathbf{t} \cdot \mathbf{t}} \tag{2.21}$$

where the vector of times is defined as $\mathbf{t} = (0\mathrm{h}, 2\mathrm{h}, 4\mathrm{h}, 6\mathrm{h})$ and the dot $(\cdot)$ represents a scalar product. Figure 2.6 shows two examples of linear fits to the y_i versus t data obtained from RNA-seq of HeLa cells with high mitochondrial content.

Evaluation of the goodness of the fit

The model we have developed (equation 2.20) is based on many assumptions. Additionally, there are potential experimental, sequencing and quantification errors in our dataset.

[†]The expression for $\tilde{\alpha}_i$, the estimation of the i-th transcript's relative degradation rate, results from a least squares linear regression of y_i versus t imposing the intercept equal to zero.

We thus need a method to evaluate how consistent the data are with the model, that is, how well the fit works.

A straightforward way to do it is by calculating the *root-mean-square error* (RMSE) of the fit. The RMSE quantifies the average distance of the data points (in our case, the experimental values of y_i) to the predictions of the linear model in equation 2.20. But the RMSE can be hard to interpret when one does not have an intuitive sense of the natural scale of the dependent variable, here y_i. To bypass this issue, we will define the *relative root-mean-square error* (rRMSE) as:

$$rRMSE = \frac{RMSE}{\sigma_y} \tag{2.22}$$

Where σ_y is the standard deviation of the distribution of y_i across all transcripts, for those values that satisfy $y_i > 0$ (that is, excluding the data points at $t = 0$). The rRMSE is simply interpreted as the error of the fit relative to the dispersion of the whole data set, quantified by σ_y. The magnitude of the rRMSE will be our proxy for the goodness of the linear fit: in order to consider a fit acceptable, we will impose the condition that it has a rRMSE < 0.5 in all biological replicates. Transcripts that do not satisfy this condition will be left out from further analyses. Figure 2.6 shows an example of an accepted (red) and a rejected (blue) fit in one of the samples of HeLa cells with low mitochondrial content.

There are many factors that can make the model fail, yielding an rRMSE lower than our threshold in one or both biological replicates. Some of these factors are:

- Experimental noise or errors coming from the quantification: these errors should be less prominent for transcripts expressed at high copy numbers, since the quantification is more accurate for them.[177] However, analyzing the transcripts with expression lower than the first quartile and higher than the fourth we found no difference in the fraction of accepted fits (50%-60% depending on the strain), which in principle rules out this possibility.

- Degradation not being well described by an exponential: equation 2.18 is only valid for the i-th transcript under the assumption that it degrades in a first order process, and thus its abundance decreases according to 2.10. Although this is assumed to be the case for most transcripts, it is possible that some of the failed fits result from more complex degradation dynamics.[178]

- The experimental method used for transcription blockage being non-ideal: we have developed a mathematical framework assuming that DRB blocks transcription completely, instantaneously and homogeneously through the whole genome when it is added to the cell culture, without interfering with any other cellular process. But it is possible that this is not the case, for instance some "leaky" transcription could be still taking place (either globally of for some specific transcripts).

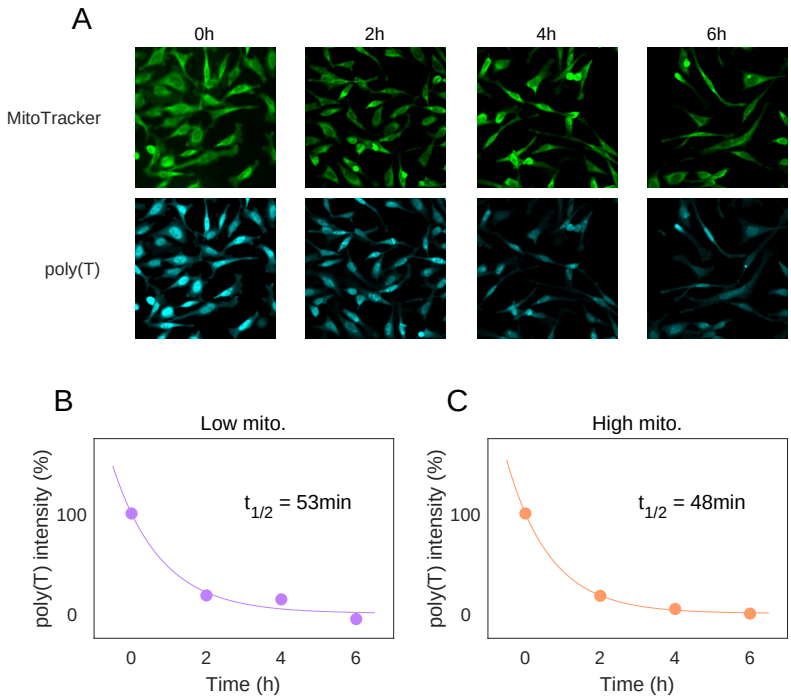


Fig. 2.7. Quantification of bulk RNA degradation rates. A poly(T) RNA FISH assay was performed on cells stained with MitoTracker green at $t = 0, 2, 4$ and 6h after DRB treatment. **A.** Microscopy images of HeLa cells through the process (green: MitoTracker and blue: poly(T) fluorescence intensities). **B-C.** After background correction, the average per-cell poly(T) intensity was fitted to an exponential curve for cells with "low" (purple) and "high" (orange) mitochondrial content. Bulk half-lives were obtained from the fits.

Yet, out model seems to work reasonably well. Roughly 50-65% (depending on the strain) of the transcripts are well fit (rRMSE < 0.5) by equation 2.20 in both replicates in at least one condition ("high" or "low").

Bulk degradation rates

The method we have described allows us to compute transcript-wise relative degradation rates. Still, to achieve true values we need the bulk degradation rate Ω (equation 2.17). This bulk rate can vary across subpopulations with different mitochondrial content (Ω_H in the "high" and Ω_L in the "low" conditions). To quantify it, we performed poly(T)

RNA FISH on cells stained with MitoTracker green and treated with DRB (see appendix A for extended experimental methods). The population was imaged every 2h. At the image processing stage, cells were classified according to their mitochondrial levels and the average per-cell RNA content was measured as the integrated poly(T) fluorescence signal. The fluorescence background was substracted and the values were scaled to the intensities at $t = 0$. Data were then fitted to an exponential curve of the type $y = \exp(-\Omega t)$. Half-lives (computed as $t_{1/2} = \log(2)/\Omega$) were found to be 48min and 53min in the "high" and "low" mitochondria conditions respectively (fig. 2.7).

Mitochondrial control of RNA decay

Once we have the bulk degradation rates (Ω) as well as the relative degradation rates (α_i) for each transcript in the "high" and "low" mitochondria conditions, we can quantify the absolute degradation rates according to equation 2.17. For the i-th transcript:

$$\begin{aligned} \delta_{i,H} &= \alpha_{i,H} + \Omega_H \\ \delta_{i,L} &= \alpha_{i,L} + \Omega_L \end{aligned} \tag{2.23}$$

where the H and L indexes indicate conditions of "high" and "low" mitochondrial content respectively.

We found a high degree of correlation between the transcript's half-lives in the "high" and "low" conditions (Pearson's correlation coefficient was between 0.7 and 0.9 depending on the strain, see figure 2.8B). Despite this correlation, we explored the possibility of individual transcripts' rates showing enough variation to explain the difference in abundances observed across subpopulations of cells with different mitochondrial content.

Similar to what we did for the TSS analysis, we filtered out all linearly scaling transcripts. We then selected one transcript per gene and computed the relative change in its degradation rate as

$$\Delta\delta_i = \frac{\delta_{i,H} - \delta_{i,L}}{\delta_{i,L}} \tag{2.24}$$

According to this definition, the relative change will be positive if the transcript degrades faster in the "high" condition and vice-versa, which is the behavior that we expect for most of them (fig. 2.8A-B). As a control, we followed the same steps but without filtering out any gene and simply selecting transcripts randomly. This was repeated 20 times for statistical significance.

For HeLa and Jurkat cells, even though Welch t-tests yielded significant p-values (below 10^{-10} in both cases), the difference in the means of the $\Delta\delta$ values of the group of non-linearly scaling transcripts (fig. 2.8C, gray) and the control (fig. 2.8C, white) was below 2%. However, MRC-5 cells posed an exception again. For them, the difference in means was ~15%. Interestingly, the average $\Delta\delta$ within the non-linearly scaling transcripts group was *negative*, meaning that many transcripts degrade faster in the "low"

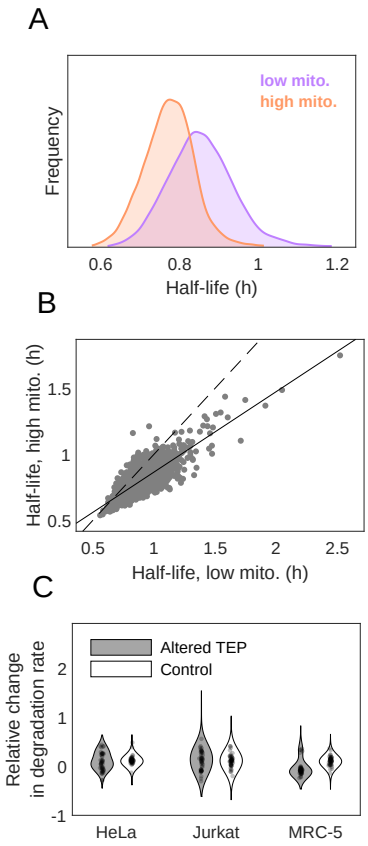


Fig. 2.8. Mitochondria-induced variation of transcript degradation rates. A. Distribution of transcript half-lives in HeLa cells with low (purple) and high (orange) mitochondrial content. The difference in averages is given by the poly(T) RNA FISH experiments after transcription inhibition (see figure 2.7). Coefficients of variation: CV_{low} = 0.11 and CV_{high} = 0.09. **B.** Correlation between transcript half-lives in HeLa cells with different mitochondrial content. Dashed line is the identity line, solid line is the best fit. **C.** We quantified the contribution of degradation rates variation to alternative splicing by selecting a subset of genes with altered TEPs (gray) and comparing the rates of the transcripts that scaled non-linearly and those of the transcripts scaling accordingly to the global factor between the "low" and "high" mitochondria subpopulations. As a control, we repeated the same procedure but selecting transcripts from all genes randomly (white). Only 50 data points per group have been plotted for clarity.

condition than they do in "high". This behavior is opposite to what we found in other strains, or in the control for MRC-5 themselves. This could indicate the existence of some mechanism acting on some specific genes and slowing down the degradation of their RNA products under conditions of increased energy budget. This modulation of the degradation rates being uneven (meaning different transcripts of the same gene scaling their degradation rates differently) can indeed give rise to alterations in the TEPs.

Although further work is required to establish causal relationships between the variation in degradation rates and transcript relative abundances, our results point towards a potential mechanism by which mitochondrial alterations of TEPs would be, to some extent, induced by an asymmetric modulation of isoform-wise degradation dynamics. The fact that this behavior doesn't seem to be present in HeLa nor Jurkat cells could mean

that the control mechanism is lost in cancer, although experiments on a wider variety of strains should be carried out before conclusions can be drawn.

2.3.3 Asymmetric scaling of transcription and degradation

Let us consider a simple Poisson process representing the production and degradation of a RNA molecule. The number of molecules at any time t is given by $n_i(t)$, and the production and degradation rates are defined as k_i and δ_i respectively. The index i covers all possible species expressed in a cell.

$$\phi \underset{\delta_i}{\overset{k_i}{\rightleftharpoons}} \text{RNA}_i$$

As we have seen, equilibrium is reached when $n_i^{eq} = k_i/\delta_i$. From our time series RNA-seq experiments we have quantified a collection of δ_i values. Transcript-wise production rates can be obtained as:

$$k_i = \delta_i \cdot n_i^{eq} \tag{2.25}$$

Formally we do not know each transcript's per cell copy number, instead, TPMs represent relative abundances. If we assume that in our data at $t = 0$ (before transcription inhibition) all species within a cell are at equilibrium, we can define

$$k_i' = \delta_i \cdot TPM_i^{eq} = \delta_i \cdot TPM_{i,0} \tag{2.26}$$

In principle k and k' values can't be compared (they do not even have the same dimensions), as they differ by the cell's total RNA content. However, we have quantified the per cell RNA ratio between the "high" and "low" mitochondria subpopulations. Having scaled TPM values by this ratio ensures that

$$\frac{k_{i,H}'}{k_{i,L}'} = \frac{k_{i,H}}{k_{i,L}} \tag{2.27}$$

where the H and L tags represent "high" and "low" mitochondria respectively.

From equations 2.25 to 2.27 we can derive:

$$\frac{k_{i,H}/k_{i,L}}{\delta_{i,H}/\delta_{i,L}} = \frac{n_{i,H}^{eq}}{n_{i,L}^{eq}} = FC_i \tag{2.28}$$

where FC_i represents the fold-change ("high"/"low") in expression for the i-th transcript. Computing production rates according to 2.26, we can represent the ratio $k_{i,H}/k_{i,L}$ as a function of the ratio $\delta_{i,H}/\delta_{i,L}$. If all transcripts scaled linearly, from equation 2.28 we would expect all data points to fall onto a straight line with a slope determined by FC_{global}, being FC_{global} the average ratio of per cell RNA content across "high" and

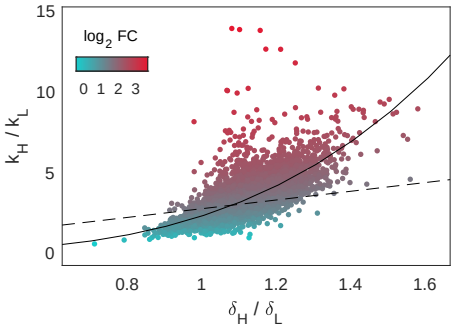


Fig. 2.9. Scaling of RNA transcription and degradation rates. Ratios ("high"/"low") of degradation and transcription rates (computed from equations 2.25 to 2.28) of individual transcripts in HeLa cells. Each dot is color coded according to the transcript's fold-change. The dashed line corresponds to $y = FC_{global} \cdot x$, and even though it represents the best possible linear fit, the data display a non-linear behavior. The solid line is the best fit of the type $y = x^n$ (fitted $n = 3$).

"low" conditions (experimentally quantified by the poly(T) RNA FISH experiments). Since we know that some transcripts scale non-linearly, we also expect a certain degree of dispersion around such line. However, we found that the data do not even display a linear trend for any of our three strains (fig. 2.9).

In principle, there is an infinite number of possibilities for k and δ that could produce a given level of expression of an RNA, both in "high" and "low" mitochondria conditions. Energetic arguments can be used to justify that, in practice, certain combinations appear selected against in nature.[179] A similar reasoning can be made to explain the different scaling of transcription and degradation rates. Under conditions of increased energy budget (i.e. "high" mitochondrial content), if the energy excess was devoted to increasing transcription and degradation rates symmetrically, the average RNA per cell would remain constant but the turnover would be faster, meaning that noise would be reduced.[28–39] But figure 2.9 shows that the scaling of the rates is in fact asymmetric, consistent with experimental evidence indicating that relative noise levels are roughly the same across cells with different mitochondrial mass, but the amount of RNA is correlated with mitochondria.[69] Combinations of parameters that imply reduced fluctuations at the cost of lower expression levels in conditions of ATP excess seem to be selected against.

Suppose that both k and δ scale with mitochondrial content (m) following a power law:

$$\begin{aligned} k &= k_0 \left(\frac{m}{m_0}\right)^{\alpha} \\ \delta &= \delta_0 \left(\frac{m}{m_0}\right)^{\beta} \end{aligned} \tag{2.29}$$

where k_0 and m_0 are the values of the rates at an arbitrary given mitochondrial content m_0, and the scaling is determined by the exponents α and β. From equations 2.29 we can derive:

$$\left(\frac{k}{k_0}\right) = \left(\frac{\delta}{\delta_0}\right)^{\alpha/\beta} \tag{2.30}$$

From the polynomial fit in figure 2.9 we infer $\alpha/\beta \approx 3$. In addition, transcription rates have been found to depend roughly linearly with mitochondrial mass[69] (fig. 2.1B), which yields:

$$\begin{aligned} k &\sim m \\ \delta &\sim m^{1/3} \end{aligned} \tag{2.31}$$

This means that RNA abundance scaling with mitochondrial mass is dominated by transcriptional rate variation. This result had already been qualitatively reported, but our findings quantify the relative contributions of both parameters.

2.4 Discussion and perspectives

Gene expression is heterogeneous from cell to cell, which represents an important source of phenotypic variation. Mitochondria is the main provider of ATP in eukaryotes, and most of this energy is invested into RNA and protein synthesis. Understanding the complexity of the gene expression cycle requires an in-depth study of the energy dependence of its many steps, and more work is needed to elucidate the molecular mechanisms connecting mitochondrial variability to processes such as chromatin remodeling or alternative splicing.

Our findings are consistent with a picture where many stages of gene expression are highly reliant on ATP availability, e.g. gene activation or transcription elongation. This leads to mitochondrial mass co-varying with the levels of RNA and protein in individual cells. Yet, gene expression entails several layers of regulation with many potential limiting factors, so this global scaling may not be representative of many individual genes. Indeed, we have found important variation across cellular strains in the specific genes subject to significant mitochondrial regulation, even though some processes seemed to be conserved. This supports the idea of mitochondria not being just a "global volume knob" amplifying the output of gene expression, but rather acting as a non-linear device yielding unique outcomes in individual genes and cell lines.

A paradigmatic example of this is alternative splicing. We have noticed an important fraction of genes for which the relative abundances of their transcripts were altered across subpopulations of cells with different mitochondrial content. This points towards a structure through which increased energy budget would favor some isoforms above

others. We have seen that transcription start site choice has an effect in this sense, possibly relating to chromatin remodeling being an ATP demanding process. Another hypothesis involves precursor mRNA secondary structure being determined by polymerase elongation speed, in turn modulated by ATP abundance. To elucidate whether this mechanism is plausible, more experimental work (possibly assessing specific mRNAs rather than the whole transcriptome) is required. Degradation rates are also affected by mitochondrial mass, both globally and for individual transcripts. We found that, in MRC-5 cells, degradation rates variation explains transcript relative abundance changes to some extent. To quantify the true relevance of this mechanism, and to check whether its prevalence in a non-cancer strain as opposed to HeLa and Jurkat is a mere artifact or is indeed related to disease, further work is necessary.

Through this chapter we have focused on a single parameter, mitochondrial mass, and its influence on gene expression through ATP budget. But structural, morphological and functional variability may also be important. In addition, mitochondrial respiration generates byproducts such as ROS and NAD+, known signaling molecules with potential roles at gene expression modulation. Understanding the contribution of mitochondria to the interplay between genotype and phenotype will require that all these factors are taken into account.

3

Mitochondrial control of extrinsic apoptosis

Mitochondrial modulation of gene expression can have important functional consequences. Several cellular processes are variable across isogenic individuals, and in many cases this variability has been attributed to differences in cells' phenotypic states. A relevant example is apoptosis. Understanding the molecular mechanisms controlling programmed cell death and dissecting the sources of variation in the apoptotic outcome has important therapeutic implications, most notably in cancer: tumor cells are able to evade apoptosis, and a common line of treatment consists in selectively triggering cell death. Through this chapter we will discuss how cell-to-cell variability in the response to death-inducing stimuli can be linked to heterogeneity in the expression of a collection of genes involved in apoptotic signaling. This expression is ultimately modulated by mitochondria, which makes it so mitochondrial content determines the apoptotic outcome in terms of cell fate and death time.

3.1 Variability in the apoptotic response

Even in genetically identical cancer cells growing in homogeneous microenvironments, *fractional killing* is observed under treatment with apoptosis-inducing chemicals. Fractional killing refers to the observation that a defined concentration of an apoptosis-inducing drug applied for a certain time span will kill a constant fraction (generally < 1) of the cells in a population regardless of the total number of cells.[180–182] This has particularly important implications in cancer,[183,184] as it poses the main cause of tumor resistance to chemotherapy.

This variability in tumor cell resistance has been traditionally associated with ge-

netic intra-tumoral heterogeneity, but it is becoming increasingly clear that non-genetic differences also play a prominent role.[67,185,186] Some context-dependent factors such as the cellular shape and the microenvironment have been pointed as key determinants of this non-genetic heterogeneity.[187–189] But minimizing this context dependence by growing isogenic cells in a homogeneous medium still shows highly variable responses to apoptosis-inducing drugs:[62] intrinsic cell-to-cell differences can elicit heterogeneous responses by themselves. Differences in the internal states of individual cells (i.e. phenotypes) also originate cell-to-cell response variation.[190] The relative contribution of external and internal sources of variability depends on the nature of each tumor and is, in general, poorly characterized.

It is then important to identify the determinants of phenotipic state heterogeneity and to study how they affect individual outcomes when identical cells are exposed to the same apoptotic stimulus. Mitochondria, controlling the cellular gene expression program in a non-linear manner, represents a good candidate for a source of phenotypic variation. In addition, these organelles are central nodes in the apoptotic signaling pathway as we will next see. Together, these observations justify the study of mitochondria as a source of variation in the cellular states and apoptotic responses.

The apoptotic signaling pathway

Anti-cancer apoptotic therapy results in the activation of two major mechanisms: the intrinsic and extrinsic pathways. The first one is triggered by signals not mediated by receptors, such as those caused by viral infection, toxins, free radicals or radiation. These stimuli induce mitochondrial outer membrane permeabilization (MOMP), considered the point of no return of the apoptotic process, and lead to the release of pro-apoptotic proteins from the mitochondrial inner space into the cytoplasm. The extrinsic route (fig. 3.1) begins with the binding of specific ligands (FAS ligand, tumor necrosis factor TNF or TNF-related apoptosis inducing ligand TRAIL) to the death receptors located on the plasma membrane. This triggers Caspase-8 activation, which in turn activates other effector caspases (responsible for chromatin condensation, DNA fragmentation and eventually cell death) but also cleaves the Bid protein inducing MOMP.

Both pathways converge at the MOMP stage, i.e. there is a crosstalk between them in which mitochondria play a central role in effector caspase activation.[191] The pro-apoptotic proteins Smac and cytochrome C are released to the cytoplasm within a few minutes after MOMP,[61,192,193] activating caspases 3 and 9. Despite the rapidness of this release, individual cells display large variability in the times elapsed between the apoptotic stimulus and MOMP, spanning a range of 4-20h depending on the stimulus type and strength.[62,194,195]

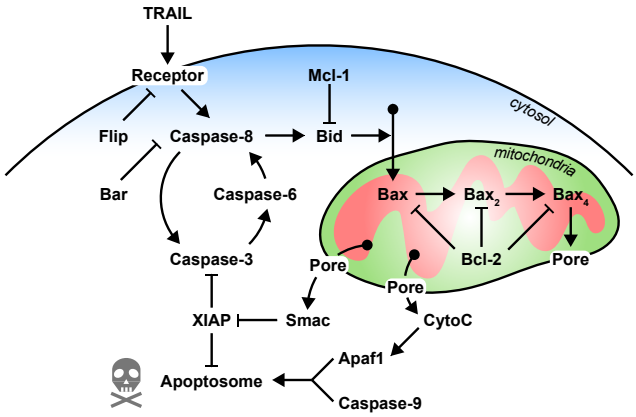


Fig. 3.1. Protein signaling network of the extrinsic apoptosis pathway. TRAIL binds to specific membrane receptors, unleashing a signaling cascade that finishes when the apoptosome complex is formed, ultimately leading to cell death. Apoptosome formation requires the release of cytochrome C (CytoC) from the mitochondria, which happens after mitochondrial outer membrane permeabilization (MOMP). Bax enters the mitochondria after the Bid protein is cleaved, where it forms tetramers that bind the mitochondrial membrane to form pores that permeabilize it. Several anti-apoptotic proteins (Flip, Bar, Mcl-1, Bcl-2, XIAP) participate in the process, slowing it down or even blocking it completely by preventing the action of their pro-apoptotic targets.

Variability in TRAIL-induced apoptosis

TRAIL is a TNF family ligand that binds specific death receptors (DR4 and DR5) on the cell surface and triggers extrinsic apoptosis. It is selective against tumor cells which makes it a promising chemotherapeutic agent, but many tumors display high rates of resistance to it severely limiting its therapeutic efficiency.[196] To identify the basis of this resistance, we used clonal populations of HeLa cells treated with variable doses of TRAIL. The fraction of dead cells after 24h of treatment was measured through visual inspection of phase contrast images (fig. 3.2A, gray bars) and by FACS using Annexin V (FITC)-PI double staining (fig. 3.2A, black dots). Both methods yield very similar response curves, with a sensitive region between ~4 and ~60ng/ml of TRAIL. Lower doses have no effect, with an outcome comparable to the control, while for larger doses the effect of TRAIL saturated at an approximately constant fraction (~35%) of dead cells, leaving a high rate of survival to treatment.

Next, we focused on the variability in death times using time-lapse microscopy of HeLa cells treated with increasing TRAIL doses. At low concentrations, a large spread with similar probabilities for long and short death times is observed, but as dose is increased average and spread of the distribution decreases as shown in figure 3.2B, even

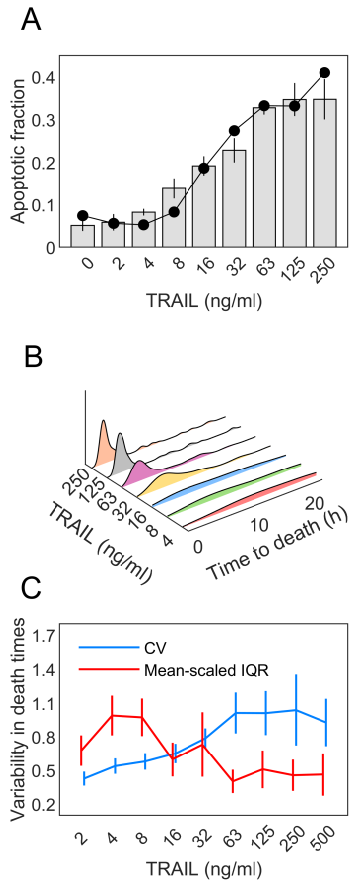


Fig. 3.2. Apoptotic variability in HeLa cells under TRAIL treatment. The response of individual HeLa cells to TRAIL is heterogeneous and dose-dependent. **A.** Apoptotic fraction of cells after 24h of treatment (0, 2, 4, 8, 16, 32, 63, 125 and 250 ng/ml). Morphological changes in apoptotic cells were identified by visual inspection of phase contrast images (grey bars). FACS using Annexin V (FITC)-PI double staining was also performed to determine apoptotic fate (black dots). Around 300 cells for each TRAIL dose were inspected. Error bars are standard deviation of three independent experiments. **B.** Distributions of death times after TRAIL treatment. Times were obtained through cell tracking in 24h time-lapse microscopy experiments. **C.** Variability in time to death at different TRAIL doses quantified as the coefficient of variation (CV, blue) and the mean-scaled inter-quartile range (IQR, red). Error bars were computed by bootstrapping.

though some cells still display long times to death. To quantify the variability in death times we used the CV (fig. 3.2C, blue) and the inter-quartile range (IQR) scaled by the mean (fig. 3.2C, red), defined as

$$\frac{t_d^{(Q3)} - t_d^{(Q1)}}{\langle t_d \rangle} \tag{3.1}$$

where t_d is the time to death, Q1 and Q3 represent the first and third quartiles respectively and angle brackets indicate a population average. Both statistics show that the variability is dose-dependent: the CV increases at large doses as a consequence of a few outliers with high apoptosis times. On the other hand, IQR (which removes the effect of outliers) is larger at low doses due to a higher degree of "flatness" in the distributions.

The observed variability has been attributed to pre-existing heterogeneity in the amounts of the proteins involved in the apoptotic pathway.[62,63] As we have discussed, variability at the protein level can be attributed to intrinsic (gene specific) or extrinsic (global) sources. Recently born sister cells (with presumably similar phenotypes conditioned by the mother's) show a high degree of correlation between their death times after being exposed to TRAIL[62,194] (fig. 3.4A), which suggests that the variability in extrinsic apoptosis must be caused, for the most part, by cellular factors that affect gene expression globally.[190]

3.2 Mitochondrial and apoptotic variability

In principle, mitochondria could play two major roles in apoptotic signaling:

- As a modulator of apoptotic gene expression: we have seen that mitochondrial content accounts for ~50% of the variability observed in cellular protein levels[69,72] (fig. 1.5), including those involved in apoptotic signaling.
- As an explicit node of the signaling network: figure 3.1 shows that the cascade of events leading to cell death involves the formation of pores in the mitochondrial membrane to allow for the release of signaling proteins from the mitochondrial matrix into the cytoplasm.

In addition, the molecular events driving apoptosis are energy-dependent.[197] Thus, the amount and/or functionality of mitochondria in individual cells could be an important factor accounting for cell-to-cell differences in death times and resistance to apoptotic signals.

3.2.1 Mitochondrial discrimination of cell fate and death time

Apoptotic fate

To study the influence of mitochondrial content on the probability of cell death, we stained HeLa cells with MitoTracker green FM (MG) and then treated them with different TRAIL doses for 24h, imaging at 15min intervals. For each cell, mitochondrial content was quantified in the initial image (at $t = 0$) and apoptotic fate was determined by manual tracking. We found clear differences in mitochondrial content across survivor and apoptotic cells, with increased mitochondria leading to higher chances of dying (fig. 3.3A). This indicates that mitochondrial mass alone can be a good marker of apoptotic cell fate.

The performance of mitochondrial content as a classifier of cell fate (death/survival) was calculated using *Receiver Operator Characteristic* (ROC) curves (fig. 3.3B). ROC curves, widely used in clinical trials, illustrate the diagnostic ability of a parameter (here

mitochondrial levels) as its discrimination threshold is changed. To create them, we set a threshold in mitochondrial content and predicted that cells below it would survive the 24h TRAIL treatment and vice-versa. We then compared the predictions with the experimental results, and counted the *true positives* (cases where cells predicted to die ended up dying) and the *false positives* (cells predicted to die surviving treatment). For each threshold we got a pair of values for the true positive rate (TPR) and the false positive rate (FPR). The ROC curve was reconstructed obtaining sets of TPR and FPR values when varying the threshold. The area under a ROC curve (AUC) summarizes the trade-off between the probability of correct and incorrect classification, and varies between 0.5 (random guessing) and 1 (perfect classifier). The AUCs we calculated indicate that mitochondrial content is a good predictor of cell fate at all TRAIL doses analyzed (fig. 3.3B).

Similar results were found for other extrinsic apoptosis inducers like TNF-α, and also for cycloheximide (CHX) and DRB, which block translation and transcription respectively causing cell damage and triggering the intrinsic apoptosis route.[198]

Death time

Next, we analyzed whether there is an influence of mitochondrial mass on death times. Such an effect was difficult to see because, on one hand, mitochondrial levels of cells committed to apoptosis are already biased towards large values (fig. 3.3A). On the other hand, death times have relatively small variability at high TRAIL doses (fig. 3.2B-C) making it challenging to identify any trend, while at low doses differences in receptor levels and activity may constitute an important source of variability in apoptosis times. In addition, low TRAIL doses result in small apoptotic fractions (fig. 3.2A), so the number of cells that would have to be analyzed to unveil any effect in death times with an acceptable degree of statistical significance is immense. To bypass these constraints, we focused on an intermediate TRAIL dose (32ng/ml) in the sensitive region of the dose-response curve (fig. 3.2A). As shown in figure 3.3C, there is a small but noticeable correlation between mitochondrial content and death times (Spearman correlation coefficient = -0.47). Apoptotic cells with mitochondrial levels below the first quartile ("low") or above the fourth ("high") exhibit significant differences in their death times (fig. 3.3D, Wilcoxon test yielded $p = 10^{-6}$). Cells with high mitochondrial content died roughly 2-5h after TRAIL addition, while cells with low mitochondrial levels displayed a wider range of times with an average of ~6h.

Ruling out other variability sources

Variability in the apoptotic outcomes of individual cells can be associated with several factors: genetic polymorphism, spatial or temporal variation in the efficiency of the apoptosis-inducing ligand, etc. Extrinsic noise sources relatively independent of mitochondria have also been pointed as regulators of the process, most notably cell cycle.[199]

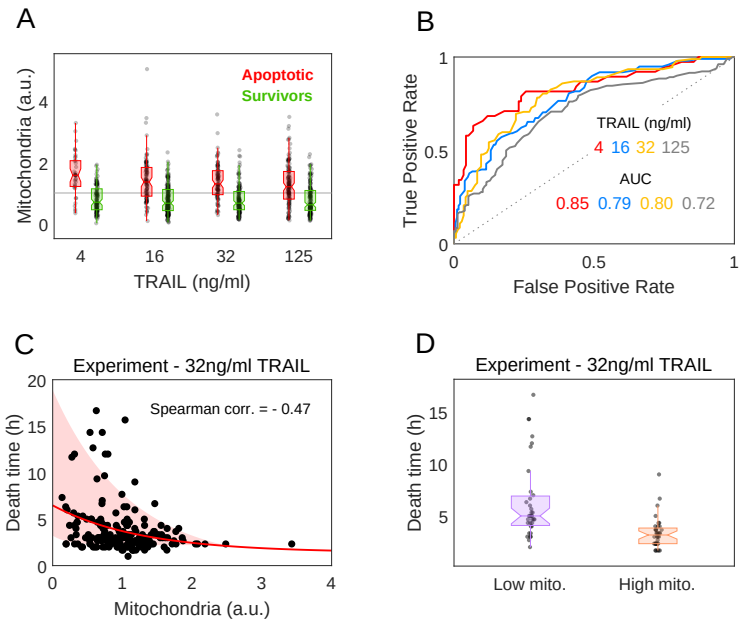


Fig. 3.3. Influence of mitochondrial content on apoptotic cell fate and death times. HeLa cells stained with MG for mitochondrial mass quantification were treated with different doses of TRAIL. After TRAIL addition, populations were imaged every 15min for 24h. For each dose, we randomly selected cells from different images, quantified their initial mitochondrial mass by integrating MG intensity at $t = 0$ and manually tracked their fate. We gathered ensembles of 200-300 cells. **A.** Mitochondrial levels of surviving (green) and dying (red) cells after 24h of treatment. Mitochondrial values are scaled by the population averages (gray line). Data are representative of six independent experiments. Boxes cover the range from the lower to the upper quartile of the data. Whiskers indicate maximum and minimum values once outliers are excluded. Horizontal lines inside the boxes represent median values, and notches indicate 95% confidence intervals for the median. **B.** Analysis of mitochondrial content as a binary classifier (death/survival) of cell fate. To quantify the performance of mitochondria as a classifier, the Receiver Operator Characteristic (ROC) curves and the area under the curves (AUC) are represented for different TRAIL doses. **C.** At 32ng/ml of TRAIL, a negative correlation between mitochondrial levels and death times in apoptotic single cells is observed. Red line is an exponential fit, for which the shaded area indicates the confidence region. **D.** Death times of HeLa cells with mitochondrial content within the first (low mito., purple) and fourth quartiles (high mito., orange).

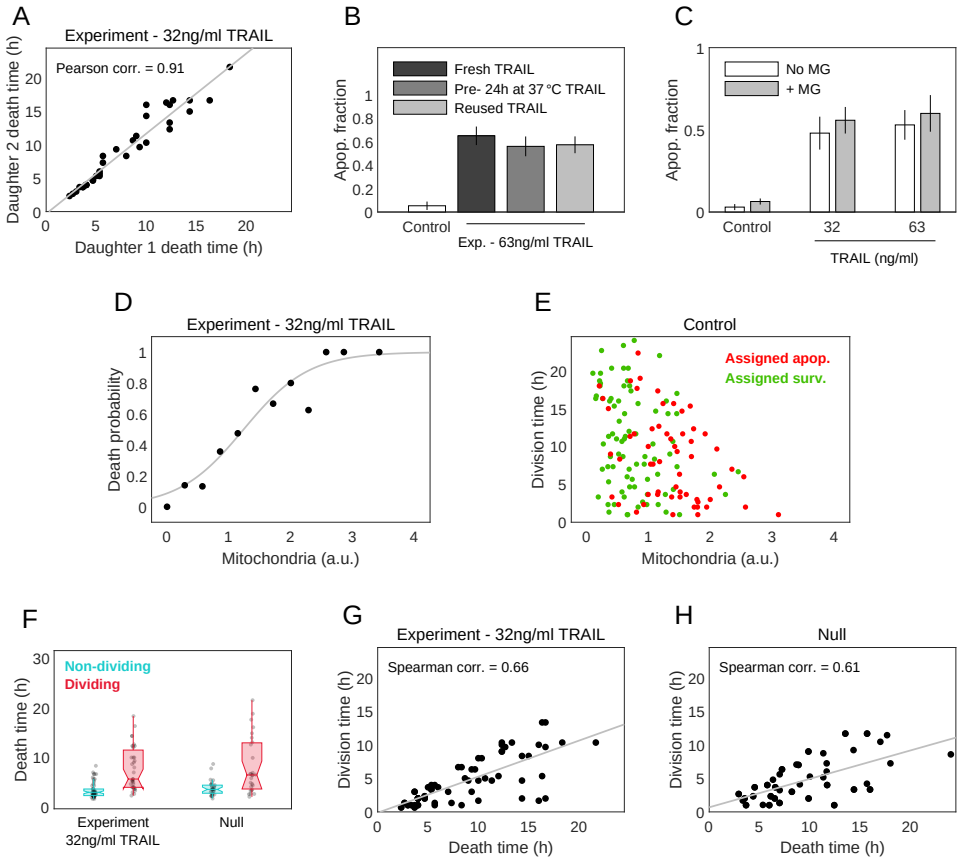


Fig. 3.4. Ruling out genetic and contextual effects on apoptotic variability. A. Correlation between apoptosis times of sister cell pairs (cells that divided after addition of 32ng/ml of TRAIL). For 16ng/ml and 125ng/ml Pearson's correlation coefficient between apoptosis times of sibling pairs were 0.84 and 0.88 respectively. **B.** Stability of TRAIL during the experimental procedure. HeLa cells stained with MG were not treated (control) or treated with fresh TRAIL at 63ng/ml, with TRAIL that had been previously incubated at 37ºC for 24h, and with the medium collected after a previous 24h experiment of TRAIL induced apoptosis. Cells were imaged for 24h every 15min and the fraction of apoptotic cells was calculated by visual inspection of phase contrast images. Error bars are standard deviations obtained by bootstrapping from different images of two biological replicates. **C**. Phototoxicity induced by MitoTracker green (MG) dye. Cells were stained with MG and then treated with 32ng/ml or 63ng/ml of TRAIL (gray), or directly treated with TRAIL without previous MG staining (white). After 24h, apoptotic fractions were determined by visual inspection of phase contrast images. **D.** Probability of dying for individual cells under TRAIL treatment. HeLa cells stained with MG were exposed to 32ng/ml of TRAIL for 24h.

Mitochondrial content was determined at $t = 0$ and cell fates were tracked manually. Then, cells were divided into bins according to their mitochondrial levels and the fraction of apoptotic cells was determined for each bin (black dots). Gray line corresponds to a sigmoidal fit. **E.** In a control experiment (with no TRAIL addition), mitochondrial mass and division times were determined for individual HeLa cells stained with MG and tracked for 24h. We then built a virtual ensemble of cells based on the null assumption of apoptotic and cell cycle programs being uncoupled. To do so, we assigned each cell in the control experiment an hypothetical apoptotic fate (green: survivors, red: apoptotic) based on their mitochondrial level and the probability given in panel D. **F.** Effect of cell cycle and division on apoptosis times. In our experiments, we noticed that cells dividing before undergoing apoptosis (red) displayed significantly longer death times than those not dividing (blue). We compared this observation with our "null" ensemble of cells (see panels D and E), finding quantitatively similar results. **G.** Correlation between death and division times for cells treated with 32ng/ml of TRAIL that divided before dying. In most sibling cells both sisters died within the observation time of 24h (70% of apoptotic dividing cells at a TRAIL dose of 16ng/ml, 75% at 32ng/ml and 86% at 125ng/ml). To correlate with division time, we took as death time the average apoptosis time of both sister cells, since they are highly correlated (panel A). **H.** Correlation between death and division times in the virtual "null" ensemble at 32ng/ml of TRAIL.

To rule out the possibility of any of these factors being behind the observed variability in our experiments, we performed a series of tests.

i. Genetic heterogeneity. Genetic heterogeneity has been traditionally attributed to variability in cancer cells response to chemotherapy, but it is now clear that non-genetic factors also play a role. Strong evidence comes from the correlation observed in the death times of recently divided sister cells[62, 195, 198] (fig. 3.4A), pointing towards a partially inheritable phenotype as the determinant of apoptotic fate even in isogenic cells. Further proof in this direction is provided by the fact that even a population reconstructed from the survivors of a first TRAIL assay will display fractional killing if presented with the same apoptotic stimulus for a second time,[62, 195] which rules out the possibility of a genetic adaptation.

ii. TRAIL degradation. To exclude the possibility that fractional killing happened due to TRAIL degradation or inactivation, the supernatants of different experiments were collected and tested for apoptotic activity. Analogously, experiments were performed with TRAIL that had been previously pre-incubated at 37°C for 24h in the absence of cells. No significant differences were found in killing efficiency in any of the cases (fig. 3.4B), meaning that cell-to-cell differences in the apoptotic response cannot be due to TRAIL being consumed or degraded through the course of the assays.

iii. Phototoxicity of the dye. For our experiments, cells were stained with Mito-Tracker green for mitochondrial content quantification prior to TRAIL treatment (see experimental methods in appendix A) and tracked through fluorescence microscopy, which in principle could induce killing by phototoxicity of the dye. Even though MG seems to slightly increase in sensitivity to TRAIL at doses of 32 and 63ng/ml (fig. 3.4C), the effect is minor and it is unlikely to be mediating the observed bias in mitochondrial content in apoptotic/survivor cells.

iv. Cell cycle. Another potential source for variability in the apoptotic outcome is cell cycle.[199] We observed that cells seemed to grow normally, many of them (both survivors and apoptotic) undergoing division after TRAIL addition. Practically all survivor cells divided within the 24h post-treatment (85-100% in the range of doses analyzed), but only a fraction of the apoptotic cells did. Furthermore, cells that divided before dying had longer death times than non-dividing apoptotic cells, and division times were correlated with death times. These observations may reflect an influence of cell division on the apoptotic outcome (e.g. delaying death times), or simply be a consequence of cells with fast commitment to death after TRAIL addition not having enough time to divide before dying. To distinguish between these two possibilities, we created a virtual ensemble of cells based on the null hypothesis of division and apoptosis being independent.

We analyzed a control experiment with no TRAIL addition, obtaining mitochondrial content and division times for all the cells in the population. Then, we manually assigned each cell a simulated apoptotic fate (death or survival, figure 3.4E) according to its mitochondrial content and based on the probability given in figure 3.4D. We gave each of the cells classified as apoptotic a hypothetical death time determined solely by its mitochondrial content (fig. 3.3C), that is, death and division times were sampled independently with mitochondrial mass as the only common nexus. Finally, we simply classify those cells with dividing times longer than the assigned death times as "non-dividing" and vice-versa.

First, we checked that the distribution of division times of the cells in the "null" ensemble classified as surviving (fig. 3.4E, green) was the same as the one observed in experiments with TRAIL addition (p-value > 0.5 in a two-sample Kolmogorov-Smirnov test). Next, comparing the death times of the dividing and non-dividing apoptotic cells of this "null" ensemble, we found the same quantitative results as in the experiment with TRAIL addition, that is, longer death times among dividing cells (fig. 3.4F) and a correlation between death and division times (fig. 3.4G-H). This indicates that apoptotic and cell cycle programs are not coupled in our system.

3.2.2 Mitochondria and apoptotic gene expression

Heterogeneity in mitochondrial content accounts for roughly 50% of total protein variability[69] (fig. 1.5). We evaluated the influence of mitochondrial mass on the amounts

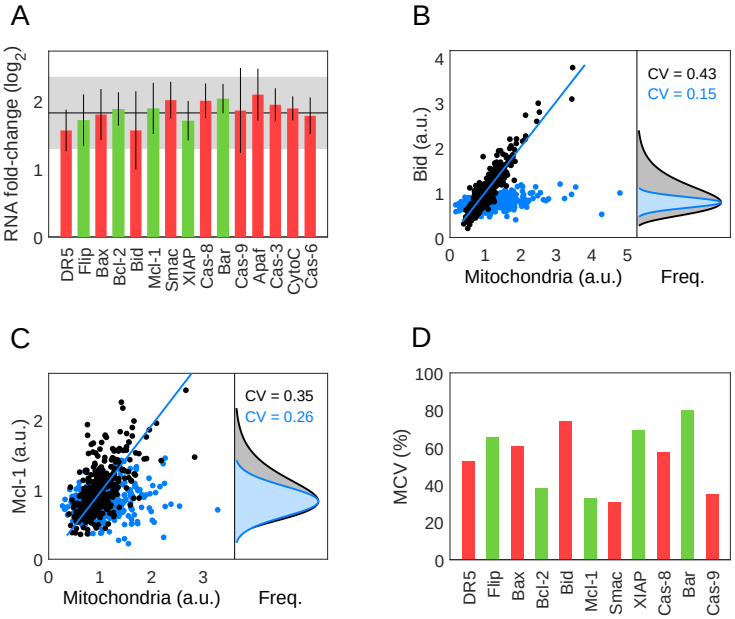


Fig. 3.5. Mitochondrial modulation of apoptotic mRNA and protein abundances. A. Logarithmic fold-change in mRNA expression of pro- (red) and anti-apoptotic (green) genes between subpopulations of cells with low and high mitochondrial content. HeLa cells labeled with MG were sorted into two subpopulations according to their mitochondrial levels. RNA was extracted and sequenced (three independent sorting experiments were performed). The solid horizontal line corresponds to the average fold-change of the whole genome (~10 000 genes). The shaded region indicates the standard deviation in the fold-change across the whole genome. Error bars are standard deviations across biological replicates. **B-C.** Scatter plots of mitochondrial mass and protein levels in single HeLa cells (black dots). Corresponding distributions of protein abundances are shown in gray. The pro-apoptotic protein Bid displays a large correlation with mitochondrial content, while the anti-apoptotic Mcl-1 has a smaller one. Co-variation with mitochondria was removed (blue dots) to estimate the protein variance not due to mitochondria (blue distributions) and calculate the mitochondrial contribution to variability (MCV, equation 1.36). **D.** Mitochondrial contribution to global variability in protein levels of several apoptotic genes. Pairs of antagonistic pro- (red) and anti-apoptotic (green) proteins are shown next to each other. Ensembles of 200-300 cells for each protein were used to estimate the MCV.

of transcripts and proteins involved in the extrinsic apoptosis pathway. To assess the impact of mitochondrial levels on transcripts, HeLa cells were sorted in two fractions with high and low levels of mitochondria, and total RNA was deep sequenced. As previously described[69,72] (see chapter 2), quantification of the per-cell average RNA content within the "high" and "low" subpopulations (see appendix A for experimental methods) revealed a global scaling, cells in the fraction with high mitochondrial levels containing around three times more RNA than cells in the low fraction. The apoptotic genes were no exception, all of them following the general trend with fold-changes in expression around the average value of the whole transcriptome (fig. 3.5A).

Transcriptome variation has a variable impact at the protein level.[150,162] We used immunolabeling to quantify the correlation between mitochondrial and protein amounts in single HeLa cells, staining untreated cells with a reporter for mitochondrial mass (MitoTracker red CMXRos)[69] and different apoptotic protein antibodies (see appendix A for extended methods). Some proteins of the route were found to be strongly correlated with mitochondria while other had weaker co-variations (fig. 3.5B-C). The mitochondrial contribution to variability (MCV) was calculated according to equation 1.36. Similarly to other protein families,[69] mitochondrial content contributed with around 50% to the total variability in the levels of apoptotic proteins (fig. 3.5D). Yet strong differences were detected in mitochondria-protein correlations between the two counterparts of of some pairs of pro- and anti-apoptotic proteins, notably the Bax/Bcl-2 and Bid/Mcl-1 pairs. These data indicate that a large fraction of the variability observed at the protein level in the apoptotic route is a consequence of cell-to-cell heterogeneity in mitochondrial content.

Mitochondria-protein correlations constrain protein-protein correlations

Given the mitochondrial mass (m) and the abundances of any two protein species P_1 and P_2 (represented by n_1 and n_2 respectively) for a cell within a population, it is possible to quantify the correlation that arises between both proteins due to each one's co-variation with mitochondria. Since mitochondrial and proteins levels are log-normally distributed, let us start by defining the transformed variables

$$\begin{aligned} x &\equiv \log m \\ y &\equiv \log n_1 \\ z &\equiv \log n_2 \end{aligned} \tag{3.2}$$

The new variables will follow normal distributions with respective means μ_x, μ_y and μ_z. We can define a vector of variables and a vector of means as

$$\begin{aligned} \mathbf{x} &\equiv (x, y, z) \\ \boldsymbol{\mu} &\equiv \left(\mu_x, \mu_y, \mu_z\right) \end{aligned} \tag{3.3}$$

and the following covariance matrix:

$$\Sigma \equiv \begin{pmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{pmatrix} \tag{3.4}$$

where σ_{ij} denotes the covariance of variables i and j (with $i = x, y, z$ and $j = x, y, z$). Since the variables are normally distributed:

$$\sigma_{ij} = \rho_{ij}\, \sigma_i\, \sigma_j \tag{3.5}$$

being ρ_{ij} the Pearson correlation coefficient between i and j. The multi-variate normal probability density function (PDF) describing the distribution of any number N of variables is given by

$$P(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}\ (\det \Sigma)^{1/2}} \exp\left[-\frac{1}{2}\,(\mathbf{x}-\boldsymbol{\mu})^T\ \Sigma^{-1}\ (\mathbf{x}-\boldsymbol{\mu})\right] \tag{3.6}$$

where the superindex T indicates a transposition, $\det \Sigma$ is the determinant of the covariance matrix and Σ^{-1} its inverse.

In the general case, $\mathbf{x}$ and $\boldsymbol{\mu}$ are N-dimensional vectors and Σ is a $N \times N$ matrix, but here we have $N = 3$. In order for 3.6 to exist, the covariance matrix must be positive definite, that is, all of its first minors must be positive. For its 2×2 minors, using 3.5 yields

$$\begin{vmatrix} \sigma_i^2 & \sigma_{ij} \\ \sigma_{ij} & \sigma_j^2 \end{vmatrix} = \sigma_i^2\, \sigma_j^2 - \sigma_{ij}^2 = \sigma_i^2\, \sigma_j^2 \left(1 - \rho_{ij}^2\right) \tag{3.7}$$

Expression 3.7 is always positive except in the trivial case of $\rho_{ij} = 1$. As for the 3×3 determinant, again using 3.5 yields

$$\begin{aligned} \det \Sigma &= \left(\sigma_x\, \sigma_y\, \sigma_z\right)^2 + 2\, \sigma_{xy}\, \sigma_{xz}\, \sigma_{yz} - \sigma_x^2\, \sigma_{yz}^2 - \sigma_y^2\, \sigma_{xz}^2 - \sigma_z^2\, \sigma_{xy}^2 \\ &= \sigma_x^2\, \sigma_y^2\, \sigma_z^2 \left[1 + 2\, \rho_{xy}\, \rho_{xz}\, \rho_{yz} - \left(\rho_{xy}^2 + \rho_{xz}^2 + \rho_{yz}^2\right)\right] \end{aligned} \tag{3.8}$$

The zeros of 3.8 are at

$$\begin{aligned} \rho_{yz}^{(0+)} &= \rho_{xy}\, \rho_{xz} + \sqrt{\left(1 + \rho_{xy}\, \rho_{xz}\right)^2 - \left(\rho_{xy} + \rho_{xz}\right)^2} \\ \rho_{yz}^{(0-)} &= \rho_{xy}\, \rho_{xz} - \sqrt{\left(1 + \rho_{xy}\, \rho_{xz}\right)^2 - \left(\rho_{xy} + \rho_{xz}\right)^2} \end{aligned} \tag{3.9}$$

It can be shown that any value for ρ_{xy} such that

$$\rho_{yz}^{(0-)} < \rho_{yz} < \rho_{yz}^{(0+)} \tag{3.10}$$

will cause $\det \Sigma$ to be positive, while all other values will make it negative. The maximum value for ρ_{yz} is reached in the middle of this interval and equals to

$$\rho_{yz}^{(1/2)} = \rho_{xy}\,\rho_{xz} \tag{3.11}$$

We have defined y and z as the log-transformed abundances of proteins P_1 and P_2 (eq. 3.2), so ρyz is the log-correlation between them, i.e. the correlation of their associated log-transformed variables. This log-correlation is constrained between to values ($\rho_{yz}^{(0-)}$ and $\rho_{yz}^{(0+)}$) that depend on the log-correlation of each respective protein with mitochondria, namely ρ_{xy} and ρ_{xz}, as expressed in equation 3.9.

Simply put, mitochondria-protein correlations restrict the available range for the corresponding protein-protein correlations. If there are no extra layers of co-regulation between the two proteins, their expected "basal" log-correlation will be given by equation 3.11. Additional co-regulation can bring this value up or down, however always within the allowed range (eq. 3.10). Table 3.1 shows some examples of log-correlation range constrains across proteins in the apoptotic signaling pathway.

Protein 1 (P_1)	Protein 2 (P_2)	ρ_{xy}	ρ_{xz}	$\rho_{yz}^{(0-)}$	$\rho_{yz}^{(0+)}$	$\rho_{yz}^{(1/2)}$
Receptor	Flip	0.79	0.87	0.39	0.99	0.69
Receptor	Caspase-8	0.79	0.86	0.37	0.99	0.58
Caspase-8	Bar	0.86	0.96	0.68	0.97	0.83
Caspase-8	Bid	0.86	0.90	0.55	0.99	0.77
Bid	Mcl-1	0.90	0.60	0.19	0.89	0.54
Bid	Bax	0.90	0.86	0.55	0.99	0.77
Bax	Bcl-2	0.86	0.51	0.00	0.88	0.44
XIAP	Smac	0.90	0.58	0.17	0.88	0.52
XIAP	Caspase-9	0.90	0.59	0.18	0.88	0.53

Table 3.1. Constraints on protein pair correlations due to global mitochondrial modulation. Legend: ρ_{xy}, ρ_{xz} log-correlations of proteins 1 and 2, respectively, with mitochondria; $\rho_{yz}^{(0-)}$, $\rho_{yz}^{(0+)}$ lower and upper bounds, respectively, for the log-correlation between proteins 1 and 2; $\rho_{yz}^{(1/2)}$ expected log-correlation between proteins 1 and 2.

3.3 Modeling apoptosis

3.3.1 The Extrinsic Apoptosis Reaction Model

The Extrinsic Apoptosis Reaction Model (EARM) is a mathematical model based on mass action representation of the well characterized extrinsic apoptotic pathway (fig. 3.1), originally built to understand the quantitative aspects of the apoptotic machinery

operation.[60] After training the EARM against experimental data, it was able to accurately reproduce the behavior of HeLa cells exposed to TRAIL. Model analysis, in agreement with experimental data, showed that the apoptotic fate as well as the elapsed time between TRAIL addition and MOMP is determined by the abundances, activities and complex interplay between the pro- and anti-apoptotic proteins participating in the signaling pathway.[60–64] Further efforts have been made to expand the EARM, for instance accounting for protein and mRNA turnover or noise induced by slow promoter dynamics.[65]

The EARM contains a total of 58 molecular species, corresponding to 18 gene products (with non-zero initial conditions) plus 40 additional species representing complexed, cleaved or differently localized forms of them. All these species react via a series of 29 biochemical reactions (listed in table 3.2) following standard Michaelis-Menten structures, with kinetic rates $k^{(\mu)}$ (forward), $k_r^{(\mu)}$ (reverse) and $k_c^{(\mu)}$ (catalytic) for the μ-th one:

$$\mathrm{A} + \mathrm{B} \underset{k_r^{(\mu)}}{\overset{k^{(\mu)}}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_c^{(\mu)}} \mathrm{D} + \mathrm{E}$$

In addition, the model includes synthesis and degradation of the molecules involved. For each cell, the synthesis rate (k_s) of the i-th molecular specie is set to

$$k_s^{(\mu)} = \delta^{(i)} n_{i,0} \tag{3.12}$$

being $n_{i,0}$ the initial number of molecules of the i-th specie and $\delta^{(i)}$ its degradation rate. The dynamics of all species can be expressed in terms of ordinary differential equations (ODEs) applying the mass action law. These equations can be numerically solved using standard computational methods.

In its original formulation, the model input is the concentration of TRAIL ligand as well as the per-cell initial levels of key proteins of the apoptotic route. Natural variability in apoptotic fate and death times arises from the random sampling of these initial protein abundances from experimental distributions.[62] Including correlations between protein pairs along the apoptotic pathway was shown to improve the predictive power of the EARM.[64] These correlations may be due to direct or indirect interactions, co-regulation by common transcription factors or other common sources of gene expression modulation such as mitochondrial content.[68,69,72]

3.3.2 Including mitochondrial regulation: the mitoEARM

We extended and modified the previous version of the EARM kinetic model[60–65] with the goal of understanding how mitochondrial variability influences apoptosis. We explicitly introduced the effect of the initial amount of mitochondrial mass into the EARM. Previous work had shown that variability in death times is a consequence of differences in the abundances of specific proteins involved in the apoptotic pathway.[62,64] On the

1. Ligand + **Receptor** $\rightleftharpoons$ Ligand : Receptor $\longrightarrow$ Ligand + Receptor*
2. Receptor* + **Flip** $\rightleftharpoons$ Receptor* : Flip
3. Receptor* + **Cas-8** $\rightleftharpoons$ Receptor* : Cas-8 $\longrightarrow$ Receptor* + Cas-8*
4. Cas-8* + **Bar** $\rightleftharpoons$ Cas-8* : Bar
5. Cas-8* + **Cas-3** $\rightleftharpoons$ Cas-8* : Cas-3 $\longrightarrow$ Cas-8* + Cas-3*
6. Cas-3* + **Cas-6** $\rightleftharpoons$ Cas-3* : Cas-6 $\longrightarrow$ Cas-3* + Cas-6*
7. Cas-6* + **Cas-8** $\rightleftharpoons$ Cas-6* : Cas-8 $\longrightarrow$ Cas-6* + Cas-8*
8. Cas-3* + **XIAP** $\rightleftharpoons$ Cas-3* : XIAP $\longrightarrow$ Cas-3*$_{Ub}$ + XIAP
9. Cas-3* + **PARP** $\rightleftharpoons$ Cas-3* : PARP $\longrightarrow$ Cas-3* + cPARP
10. Cas-8* + **Bid** $\rightleftharpoons$ Cas-8* : Bid $\longrightarrow$ Cas-8* + tBid
11. tBid + **Mcl-1** $\rightleftharpoons$ tBid : Mcl-1
12. tBid + **Bax** $\rightleftharpoons$ tBid : Bax $\longrightarrow$ tBid + Bax*
13. Bax* $\rightleftharpoons$ Bax*$_{m}$
14. Bax*$_{m}$ + **Bcl-2** $\rightleftharpoons$ Bax*$_{m}$: Bcl-2
15. Bax*$_{m}$ + Bax*$_{m}$ $\rightleftharpoons$ Bax*$_{2m}$
16. Bax*$_{2m}$ + **Bcl-2** $\rightleftharpoons$ Bax*$_{2m}$: Bcl-2
17. Bax*$_{2m}$ + Bax*$_{2m}$ $\rightleftharpoons$ Bax*$_{4m}$
18. Bax*$_{4m}$ + **Bcl-2** $\rightleftharpoons$ Bax*$_{4m}$: Bcl-2
19. Bax*$_{4m}$ + **Pore** $\rightleftharpoons$ Bax*$_{4m}$: Pore $\longrightarrow$ Pore*
20. Pore* + **CytoC$_{m}$** $\rightleftharpoons$ Pore* : CytoC$_{m}$ $\longrightarrow$ Pore* + CytoC$_{r}$
21. Pore* + **Smac$_{m}$** $\rightleftharpoons$ Pore* : Smac$_{m}$ $\longrightarrow$ Pore* + Smac$_{r}$
22. CytoC$_{r}$ $\rightleftharpoons$ CytoC
23. CytoC + **Apaf1** $\rightleftharpoons$ CytoC : Apaf1 $\longrightarrow$ CytoC + Apaf1*
24. Apaf1* + **Cas-9** $\rightleftharpoons$ Apoptosome
25. Apoptosome + **Cas-3** $\rightleftharpoons$ Apoptosome : Cas-3 $\longrightarrow$ Apoptosome + Cas-3*
26. Smac$_{r}$ $\rightleftharpoons$ Smac
27. Apoptosome + **XIAP** $\rightleftharpoons$ Apoptosome : XIAP
28. Smac + **XIAP** $\rightleftharpoons$ Smac : XIAP
29. Receptor* $\rightleftharpoons$ Ligand + **Receptor**

Table 3.2. Biochemical reactions of the EARM model. Species in bold font indicate HeLa native gene products. Asterisks represent activated species. PARP is a member of the poly-ADP-ribose polymerase protein family whose cleavage product (cPARP) can be used as a proxy for caspase activity. "Pore" is a pseudo-molecular specie representing all potential binding sites for Bax tetramers in the mitochondrial membrane (such binding produces an active "Pore*"). Subindex meanings: *Ub*, protein ubiquitinated (committed to degradation); *m*, protein localized in the mitochondrial matrix; *2* and *4*, Bax dimers and tetramers respectively; *r*, protein just released to the cytoplasm from mitochondrial matrix but not yet diffused away.

	k (h^{-1} $molec^{-1}$)	k_r (h^{-1})	k_c (h^{-1})
1.	1.30×10^{-4}	2.16×10^{-2}	2.16×10^{2}
2.	2.16×10^{-2}	2.16×10^{1}	
3.	2.16×10^{-3}	2.16×10^{1}	2.16×10^{4}
4.	2.16×10^{-2}	2.16×10^{1}	
5.	2.16×10^{-3}	2.16×10^{1}	2.16×10^{4}
6.	2.16×10^{-3}	2.16×10^{1}	2.16×10^{4}
7.	2.16×10^{-3}	2.16×10^{1}	2.16×10^{4}
8.	4.32×10^{-2}	2.16×10^{1}	2.16×10^{3}
9.	2.16×10^{-2}	2.16×10^{1}	4.32×10^{5}
10.	2.16×10^{-3}	2.16×10^{1}	2.16×10^{4}
11.	2.16×10^{-2}	2.16×10^{1}	
12.	2.16×10^{-3}	2.16×10^{1}	2.16×10^{4}
13.	2.16×10^{2}	2.16×10^{4}	
14.	3.09×10^{-1}	2.16×10^{1}	
15.	3.09×10^{-1}	2.16×10^{1}	
16.	3.09×10^{-1}	2.16×10^{1}	
17.	3.09×10^{-1}	2.16×10^{1}	
18.	3.09×10^{-1}	2.16×10^{1}	
19.	3.09×10^{-1}	2.16×10^{1}	2.16×10^{4}
20.	6.17×10^{-1}	2.16×10^{1}	2.16×10^{5}
21.	6.17×10^{-1}	2.16×10^{1}	2.16×10^{5}
22.	2.16×10^{4}	2.16×10^{2}	
23.	1.08×10^{-2}	2.16×10^{1}	2.16×10^{4}
24.	1.08×10^{-3}	2.16×10^{1}	
25.	1.08×10^{-4}	2.16×10^{1}	2.16×10^{4}
26.	2.16×10^{4}	2.16×10^{2}	
27.	4.32×10^{-2}	2.16×10^{1}	
28.	1.51×10^{-1}	2.16×10^{1}	
29.	2.16×10^{1}	0	

Table 3.3. mitoEARM parameters (I). Row numbers correspond to reactions in table 3.2. Legend: k forward reaction rate; k_r reverse reaction rate; k_c catalytic reaction rate.

		$\langle n_0 \rangle$ (molecules)	CV_p	ρ	δ (h^{-1})
1.	Ligand	$50 \times$ TRAIL [a]			0.462
2.	Receptor	500	0.36	0.79	0.139
3.	Ligand : Receptor	0			0.832
4.	Receptor*	0			0.832
5.	Flip	2 000	0.29	0.87	0.139
6.	Flip : Receptor*	0			0.832
7.	Cas-8	1 000	0.37	0.86	0.139
8.	Cas-8 : Receptor*	0			0.832
9.	Cas-8*	0			0.832
10.	Bar	1 000	0.36	0.96	0.139
11.	Cas-8* : Bar	0			0.832
12.	Cas-3	10 000	0.40 [b]	0.77 [b]	0.139
13.	Cas-8* : Cas-3	0			0.832
14.	Cas-3*	0			0.832
15.	Cas-6	10 000	0.40 [b]	0.77 [b]	0.139
16.	Cas-3* : Cas-6	0			0.832
17.	Cas-6*	0			0.832
18.	Cas-6* : Cas-8	0			0.832
19.	XIAP	100 000	0.44	0.90	0.139
20.	XIAP : Cas-3*	0			0.832
21.	PARP	100 000	0.40 [b]	0.77 [b]	0.139
22.	Cas-3* : PARP	0			0.832
23.	cPARP	0			0.832
24.	Bid	60 000	0.43	0.90	0.139
25.	Cas-8* : Bid	0			0.832
26.	tBid	0			0.832
27.	Mcl-1	20 000	0.42	0.60	0.139
28.	tBid : Mcl-1	0			0.832
29.	Bax	80 000	0.29	0.86	0.139
30.	tBid : Bax	0			0.832
31.	Bax*	0			0.832
32.	Bax^*_m	0			0.832
33.	Bcl-2	30 000	0.40	0.51	0.139
34.	Bax^*_m : Bcl-2	0			0.832
35.	Bax^*_{2m}	0			0.832

		$\langle n_0 \rangle$ (molecules)	CV_p	ρ	δ (h^{-1})
36.	Bax^*_{2m} : Bcl-2	0			0.832
37.	Bax^*_{4m}	0			0.832
38.	Bax^*_{4m} : Bcl-2	0			0.832
39.	Pore	500 000	0.40 [(c)]	1.00 [(c)]	0.139
40.	Bax^*_{4m} : Pore	0			0.832
41.	Pore*	0			2.189
42.	$CytoC_m$	500 000	0.40 [(b)]	0.77 [(b)]	0.139
43.	Pore* : $CytoC_m$	0			0.832
44.	$CytoC_r$	0			0.832
45.	$Smac_m$	100 000	0.35	0.58	0.139
46.	Pore* : $Smac_m$	0			0.832
47.	$Smac_r$	0			0.832
48.	CytoC	0			0.832
49.	Apaf	100 000	0.40 [(b)]	0.77 [(b)]	0.139
50.	Apaf : CytoC	0			0.832
51.	Apaf*	0			0.832
52.	Cas-9	100 000	0.41	0.59	0.139
53.	Apoptosome	0			0.832
54.	Apoptosome : Cas-3	0			0.832
55.	Smac	0			0.832
56.	Apoptosome : XIAP	0			0.832
57.	Smac : XIAP	0			0.832
58.	$Cas\text{-}3^*_{Ub}$	0			0

Table 3.4. mitoEARM parameters (II). Legend: $\langle n_0 \rangle$ population averages of initial protein copy numbers; CV_p coefficient of variation of the protein distributions; ρ logarithmic mitochondria-protein correlation coefficient; δ degradation rate. (a) In the original EARM, a dose of 50ng/ml of TRAIL was mimicked setting the initial Ligand copy number to 3000 molecules, i.e. 60 molecules per cell and per ng/ml of TRAIL. In the mitoEARM, we use 50 Ligand molecules per cell per each ng/ml of TRAIL. (b) The proteins for which we lack experimental data have been assumed to have $CV_p = 0.4$ and $\rho = 0.77$. These are averaged values among the proteins for which experiments are available. (c) The variable "Pore" is a pseudo-molecular specie representing all potential binding sites in the mitochondrial membrane where a pore could be formed. It relates to the mitochondrial mass in a straightforward way, which is why it has $\rho = 1$. Its coefficient of variation is given by the experimental distribution of per-cell mitochondrial content.

other hand, our experiments show that cell-to-cell variability in those proteins correlates with heterogeneity in mitochondrial levels (fig. 3.5). We thus included the effect of mitochondria-protein correlations[198] by constraining the sampling of the initial protein copy numbers used as input for the EARM. The mitoEARM does not require the explicit incorporation of protein-protein correlations, as they arise naturally from mitochondrial global modulation (see section 3.2.2).

Cell-to-cell variation in mitochondrial mass is also used to sample the variable "Pore" (see table 3.4) according to the experimental distribution of single cells mitochondrial content in HeLa[68,69,72] (fig. 1.4). In the EARM, this variable represents the number of potential binding sites for tetramers of the Bax protein in the mitochondrial membrane. When such binding happens, the variable "Pore" turns into its activated form, denoted as "Pore*" (reaction 19 in table 3.2), that represents an actual mitochondrial opening through which cytochrome C and Smac can be released into the cytosol (reactions 20 and 21 in table 3.2).

The workflow of the expanded model, which we will refer to as *mitoEARM*, can be summarized as follows (fig. 3.6A):

1. A population of cells is initialized. Each one is assigned a mitochondrial content (m) sampled from a log-normal distribution with mean and width complying with experimental data (fig. 1.4B).

2. For each cell, we log-normally co-sample (see appendix B) the abundances of all proteins involved in the apoptotic route with the assigned value of m, using individual experimental values of mitochondria-protein correlations.

3. With the sampled protein abundances as input, the mitoEARM equations are numerically solved.

4. The apoptotic fate and death times are determined from the dynamics of the species in the apoptotic signaling network (see below).

Determination of apoptotic fate. MOMP, the point of no return of the apoptotic signaling pathway, is an all-or-nothing process that takes place when a cellular threshold is overcome.[63,200] The height and rate of approach to this threshold depend on both the levels of active receptors and the cell's internal state. Recent experimental results have demonstrated that the activation speed of Caspase-8 (Cas-8) defines a threshold separating apoptotic and survivor sub-populations of HeLa cells[201] that is independent of TRAIL dose. In the mitoEARM, we fit this threshold to reproduce the probability of dying/surviving at a sensitive TRAIL dose (32ng/ml), and use the same value for all other doses tested. Cells with a maximum Cas-8 activation rate above the threshold are considered apoptotic (fig. 3.6B, red) and vice-versa (fig. 3.6B, green). Numerical simulations showed that cells assigned with larger mitochondrial levels had higher Cas-8 activity rates (fig. 3.6C).

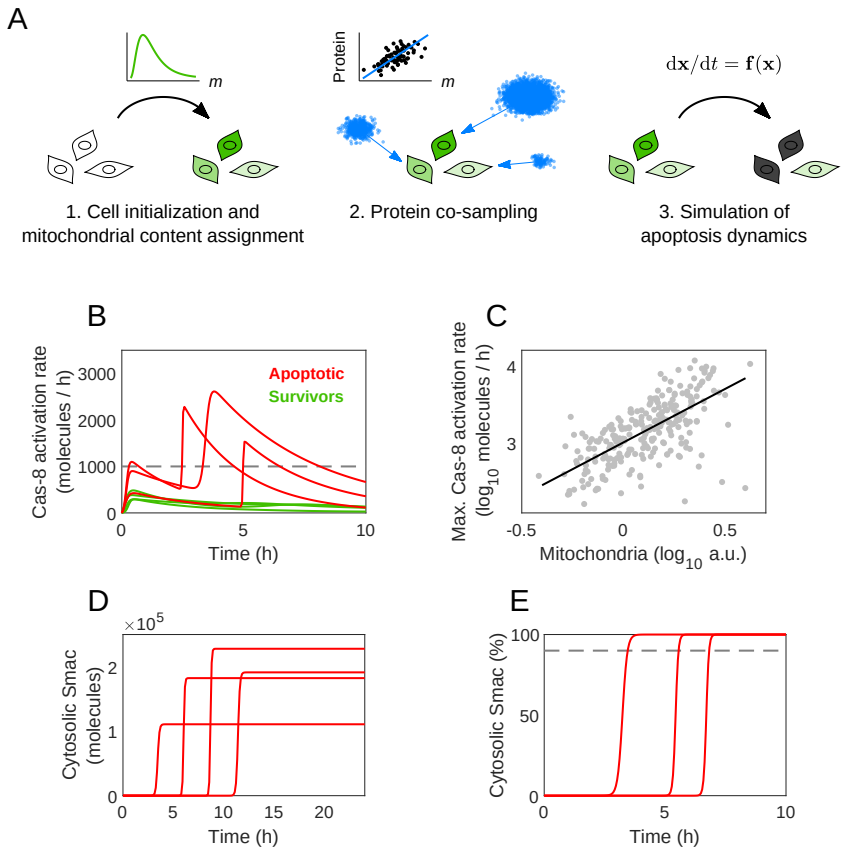


Fig. 3.6. Computational workflow and key aspects of the mitoEARM model. A. Computational model workflow. 1: Each cell is assigned a mitochondrial mass m sampled from the experimental distribution for HeLa (inset). 2: Apoptotic protein levels are co-sampled with mitochondria according to their co-variation (inset). Blue dots represent proteins. 3: The mitoEARM equations are numerically solved with the initial protein levels obtained in the previous step. Apoptotic fate and death times are determined from the dynamics of the proteins involved in the signaling pathway. Apoptotic cells are colored in gray. **B.** Caspase-8 activation rate through time for several runs of the mitoEARM (i.e. several simulated cells). The decision about the apoptotic fate (death/survival) is defined by a threshold (horizontal dashed line) in the rate of Cas-8 activation. Cells that overcome it at any time during the 24h experiment are considered apoptotic (red) and vice-versa (green). **C.** The maximum activation rate of Cas-8 depends on mitochondrial content. Each dot is a simulated cell (only cells undergoing MOMP before 24h are shown). Black line is a linear fit. **D.** The saturated levels of cytosolic Smac and its temporal dynamics are variable from cell to cell (each trajectory is a different simulated apoptotic cell). **E.** Time to death is defined in terms of the fraction of Smac released from the mitochondrial matrix to the cytosol, which reaches vales very close to 1 in all cells undergoing MOMP. Death time is quantified as the time needed for Smac to reach a cytosolic fraction of 0.9 (dashed line).

Determination of death time. The other readout of the model to be compared with experimental data is the death time of apoptotic cells. In the original EARM, time elapsed between TRAIL treatment and MOMP was taken as the time at which Smac protein reached 50% of its saturated cytosolic levels. Consistent with experimental observations, Smac release is fast in the model.[60] In the mitoEARM, we took death times as the time needed for cytosolic Smac to reach 90% of its maximum (fig. 3.6D), which agrees with the fact that most heterogeneity in death times comes from cells reaching MOMP at variable times, while the time span between MOMP and death is less variable.[60–64]

Parameter estimation and model calibration

We started the calibration of the mitoEARM with the same parameter set (population averages of protein copy numbers and kinetic rates of biochemical reactions) as the EARM version 1.3.[64] For the degradation rates, we used the values provided in Bertaux et al.[65] To reproduce the range of experimental death times, we rescaled all kinetic parameters by a common factor: since all kinetic rates appear as linear terms in the model equations,[60,64] this is equivalent to rescaling time.

To adjust the threshold in Cas-8 activation rate, we used the experimental fraction of surviving cells at 32ng/ml of TRAIL as a reference. We then adjusted the binding constant of TRAIL ligand to the death receptor DR5 (k of the first reaction in table 3.2) as well as the average Cas-8 and receptor levels. We also adjusted the number of TRAIL molecules per cell corresponding to a reference dose (table 3.4). Manual calibration of these few parameters was enough to qualitatively reproduce all our experimental observations: dose-response curve, variability in apoptosis times and mitochondrial discrimination of cell fate and death times (figures 3.3 and 3.7).

Mitochondrial and apoptotic variability in the mitoEARM

Once mitochondria-protein correlations are accounted for, the model quantitatively reproduces all of our experimental findings:

- The simulated dose-response curve follows the trend of the experimental one in the whole range of TRAIL doses , including the sensitive region between 8 and 63ng/ml of TRAIL (figs. 3.7A and 3.2A).
- The distributions of simulated death times show a large spread for low TRAIL doses, while for high ones the majority of cells die within the first ~4h after treatment (figs. 3.7B and 3.2B).
- In the mitoEARM, mitochondrial levels are able to discriminate apoptotic cell fate. Quantifying the predictive power of mitochondria as the area under the ROC curve (AUC) yields values similar to those found experimentally (figs. 3.7C and 3.3B).

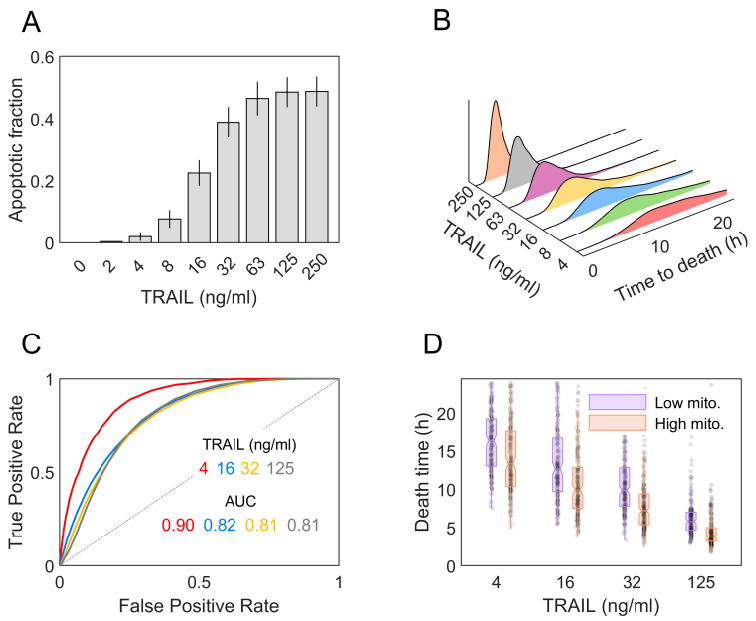


Fig. 3.7. Coupling protein and mitochondrial variability explains apoptotic outcomes. A. Fraction of simulated apoptotic cells after 24h of TRAIL treatment at the indicated doses. Error bars were computed by bootstrapping. **B.** Distributions of death times at different TRAIL doses from mitoEARM simulations. **C.** Analysis of mitochondrial content as a binary classifier (death/survival) of cell fate from model simulations. **D.** Death times of simulated apoptotic cells with mitochondrial levels within the first (high mito., orange) or the fourth quartile (low mito., purple). We simulated ensembles of 10^4 cells per dose and calculated the median, inter-quartile range and the minimum/maximum values with the whole ensembles, but only 200 data points are shown for clarity.

- In agreement with experimental data, simulated cells with low mitochondrial content display systematically longer death times than those with high mitochondrial mass (figs. 3.7D and 3.3D).

3.4 Keys to the mitochondrial regulation of apoptosis

The mitoEARM recapitulates our experimental findings regarding variability in the apoptotic response and its connection to cell-to-cell differences in mitochondrial content. It achieves this by just introducing protein and mitochondrial levels co-variations across all species involved in the apoptosis signaling pathway, indicating that the opera-

tion of said pathway is determined by the abundances of the apoptotic proteins, in turn modulated by mitochondria. To gain insight into the mechanisms through which global mitochondrial modulation induces qualitatively different apoptotic outcomes, we ran a series in-silico experiments with the mitoEARM model.

To explore the effect of mitochondria-protein correlations on death times, we simulated two extreme scenarios. First, we set all correlations in the mitoEARM to $\rho = 1$, i.e. we made it so protein abundance was deterministically regulated by mitochondrial content. Second, we set correlations to $\rho = 0$ for all proteins, making their abundances fully independent of mitochondria. Simulating apoptotic dynamics with perfect mitochondria-protein correlations shows that death times follow an inverse non-linear trend with mitochondrial mass, with longer death times corresponding to cells with the lowest mitochondrial levels (fig. 3.8A, red dots). Including mitochondria-protein correlations measured experimentally in the mitoEARM scatters death times with large deviations around the deterministic trend (fig. 3.8A, black dots), suggesting that they are very sensitive to small changes in protein variation. Therefore, any additional source of protein variability may have a noticeable impact on death times.

To understand the role of mitochondria-protein co-variation on cell fate, we studied whether the levels of individual proteins of the apoptotic route could discriminate fate with comparable accuracy as mitochondrial mass. We calculated the performance of each one of them as a classifier of death/survival by representing ROC curves obtained setting thresholds in specific protein abundances and quantifying the area under them (fig. 3.8B, filled bars). With the only exception of Caspase-8, whose activation rate is used as a death/survival discrimination threshold (see section 3.3.2), the performance of all other proteins in the pathway is worse than that of mitochondria, but the pro-apoptotic Bid and Bax are close to it. It is possible that the levels of these two proteins are key determinants of cell death, and that the observed good classification performance of mitochondrial mass is simply an effect of its high degree of co-variation with Bid and Bax abundances (fig. 3.5C). To investigate this possibility, we repeated the discrimination analysis but this time sampling protein levels independently of mitochondrial mass (fig. 3.8B, hollow bars). Under this circumstances, no single protein is a proper classifier by itself (AUC < 0.7 in all cases), which reinforces the view that apoptotic fate is not solely determined by a specific protein in the signaling route, but rather by the complex interplay of all of them. Mitochondrial mass is an underlying variable that globally affects all of said proteins, making it a good predictor of apoptotic fate.

Notably, the predictive power of mitochondria is lost when the correlations with the apoptotic proteins are removed (fig. 3.8B, black). Even though the mitochondrial membrane appears explicitly as a node of the apoptotic signaling pathway (MOMP is required for apoptosis completion as indicated in figure 3.1), the mitochondrial content of a cell has no effect in apoptotic fate if the co-regulation of all other proteins is not accounted for. Similarly, the absence of correlations produces a total lack of co-variation between death times and mitochondrial content (fig. 3.8A, gray dots). This is a consequence of

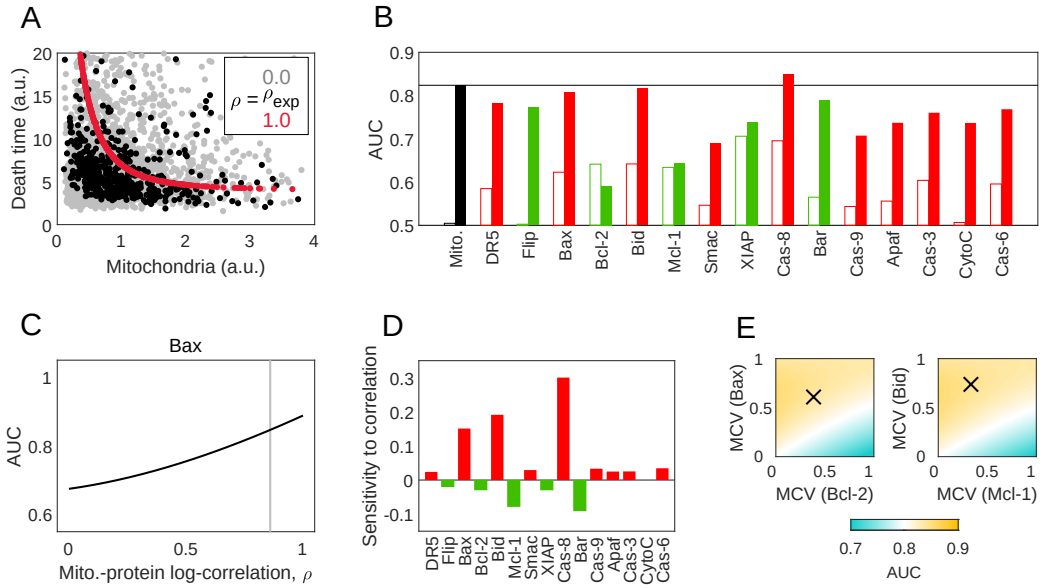


Fig. 3.8. Determinants of mitochondrial performance at apoptosis outcome prediction in the mitoEARM. We ran model simulations of populations (10^4 cells) treated with 32ng/ml of TRAIL to unveil the effect of mitochondria-protein correlations in the apoptotic outcome. **A.** Effect of global modulation of mitochondria-protein correlations on death times. Red dots: perfect correlation ($\rho = 1$) between mitochondria and protein abundances. Black dots: experimental correlation values ($\rho = \rho_{exp}$), Spearman correlation = −0.49. Gray dots: no correlation ($\rho = 0$), Spearman correlation coefficient = −0.49. **B.** Performance of the pro- (red) and anti-apoptotic (green) proteins as discriminators of apoptotic fate, quantified as the area under the ROC curves (AUC). Filled bars: including experimental mitochondria-protein correlations. Hollow bars: protein abundances sampled independently of mitochondrial levels. The horizontal line represents the discriminatory power of mitochondria as a reference. **C.** Discriminatory power of mitochondrial content (quantified as the AUC) as a function of the log-correlation (ρ) between mitochondria and the levels of Bax protein (vertical gray line indicates the experimental value for this log-correlation), with correlations of all other proteins set to their experimental values. **D.** Sensitivity of AUC to changes in individual mitochondria-protein correlations (eq. 3.13). Negative values correspond to situations where increasing correlation decreases discrimination performance. **E.** Discrimination performance (AUC) as a function of the mitochondrial contribution to protein variability (MCV, eq. 1.36) for the two pre-MOMP pairs of antagonistic pro- and anti-apoptotic proteins Bax/Bcl-2 (left panel) and Bid/Mcl-1 (right panel). Black crosses indicate experimental values of both MCVs.

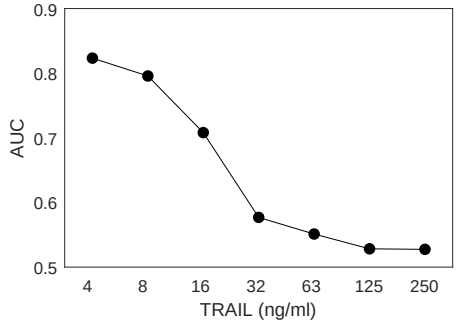


Fig. 3.9. Performance of death receptor levels as a predictor of apoptotic fate. Discriminatory power of death receptor (DR5) levels in the mitoEARM, quantified as the area under the ROC curve (AUC) for different TRAIL doses.

an excess of potential Bax binding sites, making the step of pore formation not limiting. Therefore, mitochondrial regulation of apoptotic fate seems to happen mainly through the modulation of apoptotic gene expression, with its role as a node of the network being less meaningful.

At very low TRAIL doses (4ng/ml), where death receptors are far from saturation by the ligand, the discriminatory capacity of mitochondria seems to improve (figs. 3.3B and 3.7C). Since levels of the DR5 are correlated with mitochondrial mass (fig. 3.5C), it is possible that receptor abundance plays a major role at these low doses. To test this possibility, we we repeated the discrimination analysis explained previously (fig. 3.8B) at different TRAIL doses and calculated the area under the ROC curve obtained using receptor abundance as a death/survival classifier. As shown in figure 3.9, at sensitive and saturating doses receptor levels have little discriminatory power (AUC < 0.7 for 16, 32, 63, 125 and 250ng/ml), but at low doses (4 and 8ng/ml) they are good predictors of cell fate. This indicates that, while in general the apoptotic outcome is determined by the internal state of the cell, under conditions of low ligand concentration the role of death receptors may become particularly relevant.

Finally, we investigated whether cellular fate is more sensitive to co-variation of specific proteins with mitochondrial levels. We carried out a sensitivity analysis of discrimination performance (quantified as the AUC) for each protein in the pathway, changing its correlation with mitochondria. By analogy with the local sensitivity analysis of kinetic models with respect to parameter variations,[202] we define the local sensitivity coefficient for the i-th protein as

$$s_i = \frac{\rho_i^{(exp)}}{AUC^{(exp)}} \left. \frac{\partial AUC}{\partial \rho_i} \right|_{\rho_i = \rho_i^{(exp)}} \tag{3.13}$$

where ρ_i is the logarithmic correlation of mitochondria with protein i, and the (exp) superindex indicates magnitudes at their experimentally measured values. To numerically

compute the partial derivative, we changed the correlation ρ_i (leaving all others $\rho_{j\neq i}$ at their experimental values), simulated a population of cells and calculated the resulting AUC (see fig. 3.8C for an example).

Not surprisingly, the highest sensitivity corresponds to Caspase-8, whose activity sets the threshold for cell fate discrimination in the mitoEARM (see section 3.3.2). Changes in the correlation of the pro-apoptotic proteins Bid and Bax also affects cell fate discrimination to a large extent (fig. 3.8D): tighter mitochondrial control of these proteins' expression levels improves classification performance. Interestingly, their anti-apoptotic counterparts Bcl-2 and Mcl-1 (as well as other anit-apoptotic proteins) show the opposite behavior: classification performance is decreased at higher mitochondrial-protein correlation. This indicates, on one hand, that mitochondrial control of protein abundance is especially important for Caspase-8 and the pre-MOMP pairs of pro-/anti-apoptotic proteins Bid/Mcl-1 and Bax/Bcl-2, while its influence on other nodes of the pathway may not be so relevant. On the other hand, it seems that mitochondrial control of gene expression should be more rigid for pro-apoptotic proteins than for anti-apoptotic ones (which should be freed from mitochondrial regulation) in order for mitochondrial levels to be a determinant of apoptotic fate.

An optimal cell fate discrimination by mitochondria may occur in a regime where the anti-apoptotic proteins Mcl-1 and Bcl-2 are de-correlated from mitochondrial mass, while the co-variation of mitochondria with their anti-apoptotic targets, Bid and Bax respectively, is maximal. We therefore computed the performance of mitochondria as a predictor of apoptotic fate as a function of the MCV (mitochondrial contribution to variability at the protein level, eq. 1.36). We did so by sweeping different combinations of mitochondria-protein correlations for the Bax/Bcl-2 or Bid/Mcl-1 pairs while keeping the rest of the correlations at their experimental values, and then obtaining the ROC curve and the AUC using mitochondria as a death/survival classifier. We found that optimal fate discrimination takes place at a high MCV of the pro-apoptotic protein and low MCV of its corresponding anti-apoptotic partner (fig. 3.8E). In contrast, discriminatory power substantially decreases for high MCV of the anti-apoptotic proteins. The experimental values for the MCVs are in a regime close to optimality (fig. 3.8E, black crosses).

Mitochondrial mass correlates with apoptotic protein abundances in solid tumors

With a combination of experimental results and model simulations, we have showed that mitochondrial content can predict apoptotic fate due to its influence on the expression of apoptotic proteins. We have used clonal populations of HeLa cells, which eliminates potentially important genetic heterogeneity, and homogeneous culture conditions to minimize contextual effects. In solid tumors, however, all these factors may substantially contribute to apoptotic drug resistance. Yet we tested whether the same sources of non-genetic variability as we have described in cultured HeLa cells could be found in individual cells within solid tumors.

The same immunolabeling strategy was followed to simultaneously quantify mitochondrial and protein content, their levels of cell-to-cell variability and their correlation. We stained paraffin sections from colon cancer biopsies of three individuals with antibodies against Aconitase 2 as a reporter of mitochondrial mass as well as against one of the proteins Bax, Bcl-2, Bid, Mcl-1, Smac, XIAP, Caspase-8 and Bar (fig. 3.10A). These proteins were selected because they constitute pairs of pro- and anti-apoptotic proteins that displayed high sensitivities to mitochondrial correlation in HeLa cells (fig 3.8D).

Similar to clonal HeLa cell populations, tumoral cells from colon cancer samples exhibit variability in both mitochondria and apoptotic protein levels (3.10B). We also observe a large correlation of mitochondrial mass with the abundance of specific proteins (i.e. high MCVs, figure 3.10C). Moreover, for the protein pairs Bax/Bcl-2, Mid/Mcl-1 and Smax/XIAP the pro-apoptotic protein shows a higher degree of correlation with mitochondria than its anti-apoptotic counterpart. This result suggests that the mitochondrial content may also determine variability in resistance and apoptotic fate of cells in solid tumors, and constitutes a first step towards the assessment of mitochondrial mass as a biomarker for diagnosis and prognosis in cancer.

3.5 Discussion and perspectives

Fractional killing by anti-cancer, apoptosis-inducing therapies is a major source of disease recurrence. Failure to eliminate even a small, innately drug-resistant portion of a tumor results in sustained growth and cancer relapse. This variability in the apoptotic response is, to a large extent, due to heterogeneity in the molecular signatures of cancer cells that can be caused by both genetic and non-genetic factors.[185,186,203–206] Increasing evidence points towards variability at the transcript and protein levels being strongly influenced by phenotypic state and population context.[24,34,47,48,190]

The apoptotic pathway is a complex protein network that involves non-sequential organization and the competition of molecular signals ultimately leading to a binary decision (death/survival) for each individual cell.[62,63] To address this complexity, we adopted a systems level approach combining single-cell experiments and computational modeling. Our results reveal that mitochondrial content discriminates apoptotic cell fate of single cells. To account for the role of mitochondrial content on protein variability, we modified a pre-existing model of the extrinsic apoptotic pathway.[60–64] Constraining the possible copy numbers of the apoptotic proteins by introducing their correlations with mitochondrial levels, the model reproduces all our experimental observations.

Mitochondria modulate the abundance of all proteins of the apoptotic route, but in different ways: control over the abundances of specific pro-apoptotic species that are key for MOMP triggering is tighter than that exerted over their anti-apoptotic counterparts. Our simulations indicate that the power of mitochondrial mass as a death/survival discriminator relies on these dependencies. Interestingly, these differences are also observed in cells from colon cancer tumors. In line with this finding, other works have

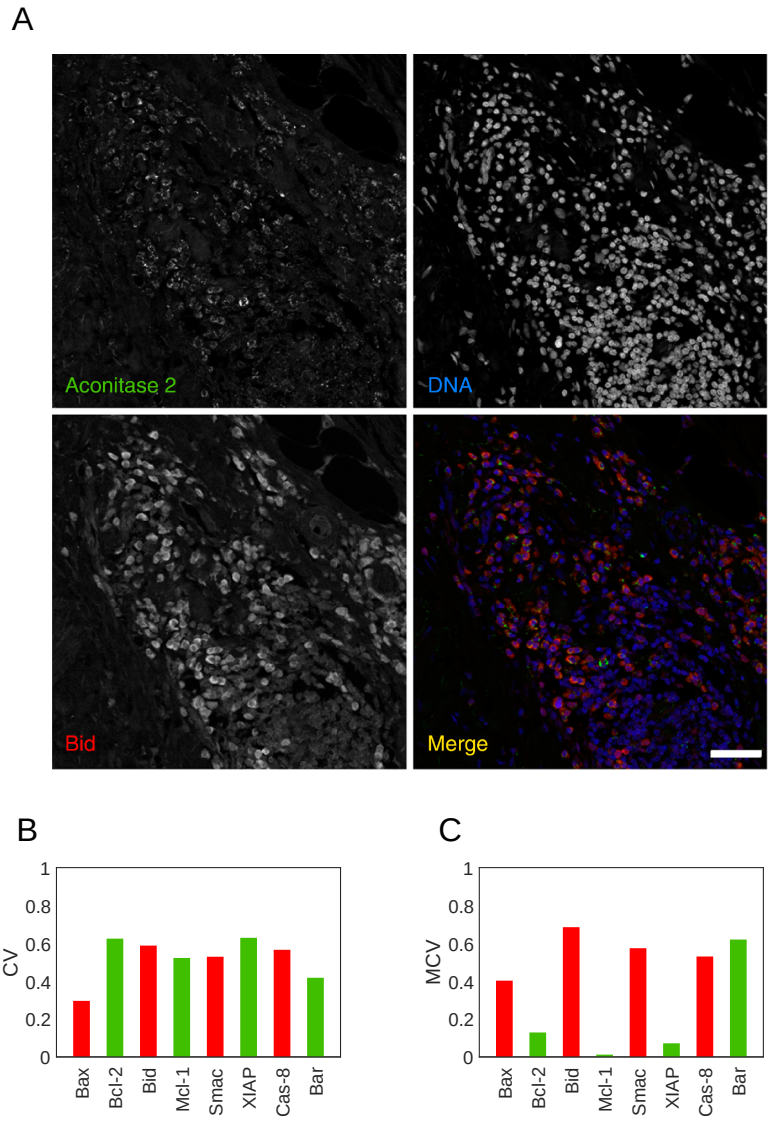


Fig. 3.10. Variability in mitochondrial and apoptotic protein levels in colon cancer samples. **A.** Colon cancer section stained with Aconitase 2, Bid and DAPI illustrating the variability in the expression of Bid and Aconitase 2. Scale bar (white): 50μm. **B.** Coefficient of variation (CV) of several proteins involved in the apoptotic pathway. **C.** Mitochondrial contribution to variability (MCV) in the levels of several apoptotic proteins. Data are representative of four independent biopsies. Statistical parameters were calculated from ensembles of 500 to 1000 cells for each protein antibody.

found that apoptotic susceptibility can be determined by the levels of pre-MOMP pro-apoptotic proteins of the BH3 family,[207] while the levels of anti-apoptotic proteins act as some kind of "buffer" to protect cells against basal levels of death-inducing signals present under normal physiological conditions.[208]

In cultured HeLa cells, mitochondria-protein correlations seem to be close to the point of optimal discrimination of apoptotic fate, suggesting that mitochondrial content being a predictor of the apoptotic outcome is not an spurious effect. There are biological reasons to single out mitochondrial mass, and possibly functionality, as a cellular determinant of programmed cell death. Apoptosis is a physiological process aimed to eliminate damaged or abnormal cells, maintaining tissue homeostasis. There is a chance that increased mitochondrial mass can induce more DNA damage, as mitochondria are the main cellular source of reactive oxygen species (ROS). Indeed, ROS levels scale with mitochondrial content in HeLa cells,[198] although experiments where cells were exposed to pro- and anti-oxidant agents prior to testing for apoptotic resistance strongly suggest that increased death rates are not due to higher ROS exposure.[198] Alternatively, mitochondria also modulate the ratios of many metabolites such as ATP/ADP, acetyl-CoA/CoA, NAD+/NADH and NADP+/NADPH, which could act as metabolic checkpoints for cell death.[209] Cells with high mitochondrial mass may be more prone to imbalances in metabolite ratios, placing them in a metabolic state that could prime them for death after a severe stress. In light of our results, this would mean that global metabolic control of programmed cell death would be achieved "indirectly", by exploiting the mitochondrial modulation of apoptotic gene expression.

There are examples in the literature connecting mitochondrial content and/or functionality to apoptotic fate and chemotherapy resistance. For instance, leukemia cells have been found to have larger mitochondrial mass, a greater mitochondrial DNA copy number and a higher rate of oxygen consumption than normal hematopoietic cells, and were selectively killed by drugs inhibiting mitochondrial protein synthesis.[210] On the other hand, down-regulating mitochondrial function by retrograde signaling promotes endothelial-mesenchymal transition (EMT),[211] which is linked to metastasis.[212] Recently, the transcriptomic analysis of 20 different types of cancers (8161 cancer and normal samples) showed that the down-regulation of mitochondrial genes was associated with the worst clinical outcome and correlated with the expression of genes promoting metastasis across many cancer types.[213]

In summary, our results suggest that mitochondrial content can be a good biomarker for the prediction of apoptotic susceptibility. Despite an overwhelming amount of studies, no optimal biomarkers have been described for some cancer types such as colorectal tumors.[214] To address this, some authors have proposed the use of the whole apoptotic profile of a tumor (rather than the expression of single markers) in order to improve the prognosis and treatment of cancer patients.[214] The reported changes in the expression of pro-apoptotic genes with mitochondrial levels in colon cancer samples raise the possibility to establish mitochondrial content as a unique biomarker, representing a readout

for the outcome of the whole apoptotic pathway. The validation of this hypothesis will require an extensive analysis of different cancer samples and their clinical response to a variety of chemotherapeutic agents.

4

Conclusions

Noise in gene expression is a key determinant of phenotypic variation and has relevant functional implications in many cellular processes. It is thus important to identify and characterize the sources of this cell-to-cell variability, which requires an in-depth understanding of the complex molecular steps of the gene expression cycle. A common constraint to many of these steps is energy dependence. In eukaryotes, most of this energy is provided by mitochondria, cellular organelles that are variable from cell to cell in number, size, morphology and functionality. Through a combination of experimental work, statistical analysis, mathematical modeling and computer simulations, we have investigated the effect of mitochondrial variability on gene expression and how it determines the heterogeneous response of individual cells to apoptosis-inducing stimuli, an essential challenge for the design of reliable chemotherapeutic strategies.

Mitochondria and gene expression

- Many stages of the gene expression cycle are energy-dependent, such as chromatin remodeling for gene activation, transcription elongation or translation elongation. These are general mechanisms that affect gene expression globally and make it so RNA and protein levels co-vary with mitochondrial mass in single cells.
- The expression of each individual gene entails many layers of regulation, including the complex interplay of specific transcription factors and molecular machines. There are many potential limiting elements (e.g. the availability of polymerases and ribosomes) that can yield a non-linear scaling of the expression of specific genes with mitochondrial content.
- Despite sharing the global trend of increased biosynthetic activity under conditions of high mitochondrial content, different cellular strains display important disparities in the modulation of individual genes.

- Mitochondria does not just work as a "global volume knob" amplifying the output of gene expression, but also as a non-linear device selectively altering the expression of individual genes. A paradigmatic example is alternative splicing: the relative abundances of the transcripts of the same gene can vary significantly across cells with different mitochondrial content.

- Mitochondrial modulation of alternative splicing is, to a large extent, driven by transcription start site choice, likely due to chromatin remodeling being an energy demanding process. Other potential mechanisms are RNA secondary structure being determined by polymerase elongation speed, or degradation rates of different splicing variants having different energy dependencies.

- mRNA production and degradation rates scale asymmetrically with mitochondrial mass. For the most part, transcript expression scaling is dominated by the first one.

Mitochondria and apoptosis

- Mitochondrial content discriminates apoptotic fate of individual cells (higher mitochondrial mass making cells more prone to death), and also has an effect on death times (cells with increased mitochondrial levels tending to die faster).

- Mitochondria modulate the abundances of the proteins participating in the apoptotic signaling pathway. This is the main determinant of its predictive power for the apoptotic outcome in single cells. The role of mitochondria as a node of the apoptotic signaling route seems to be minor in comparison.

- Optimal prediction of cell fate by mitochondrial levels happens when mitochondrial control over the abundances of pro-apoptotic proteins is tighter than control over anti-apoptotic ones, especially for the species in the pre-MOMP part of the pathway. This situation is indeed observed in cultured HeLa cells as well as in solid tumor samples.

- In colon cancer samples, pro-apoptotic proteins strongly co-vary with mitochondria, and anti-apoptotic ones display more modest correlations. These observations make mitochondrial mass a good candidate for a unique biomarker serving as a proxy for the apoptotic sensitivity of cancer cells.

Conclusiones

El ruido en expresión genética es un determinante clave de la variabilidad fenotípica y posee implicaciones funcionales relevantes en muchos procesos biológicos. Por tanto, es importante identificar y caracterizar las fuentes de esta variabilidad, lo que requiere un entendimiento profundo de las complejas etapas del ciclo de expresión génica. Una restricción común a muchas de estas etapas es su dependencia energética. En eucariotas, la mayor parte de esta energía proviene de las mitocondrias, orgánulos celulares variables de célula a célula en número, tamaño, morfología y funcionalidad. A través de una combinación de trabajo experimental, análisis estadístico, modelado matemático y simulaciones computacionales, hemos investigado el efecto de la variabilidad mitocondrial en la expresión genética y cómo determina la heterogénea respuesta de células individuales a estímulos inductores de apoptosis, un reto esencial para el diseño de terapias quimioterapéuticas fiables.

Mitocondria y expresión genética

- *Muchos pasos del ciclo de expresión genética son dependientes de energía, como la remodelación de la cromatina para la activación de genes, la elongación transcripcional o la elongación traslacional. Éstos son mecanismos generales que afectan globalmente a la expresión genética y hacen que los niveles de ARN y proteína co-varíen con la masa mitocondrial en células individuales.*

- *La expresión de genes individuales implica muchas capas de regulación, incluyendo la interacción compleja de factores de transcripción específicos y máquinas moleculares. Existen muchos elementos potencialmente limitantes (por ejemplo la disponibilidad de polimerasas y ribosomas) que pueden producir un escalado no lineal de la expresión de genes específicos con el contenido mitocondrial.*

- *Pese a compartir la tendencia global del incremento de la actividad biosintética*

bajo condiciones de alto contenido mitocondrial, diferentes líneas celulares presentan disparidades importantes en la modulación de genes individuales.

- *La mitocondria no funciona únicamente como un "controlador de volumen" que amplifica la producción de la expresión genética, sino también como un dispositivo no lineal capaz de alterar selectivamente la expresión de genes individuales. Un ejemplo paradigmático es el splicing alternativo: las abundancias relativas de los transcritos de un mismo gen pueden variar significativamente entre células con diferente contenido mitocondrial.*

- *La modulación mitocondrial del splicing alternativo está, en gran medida, dada por la elección del sitio de iniciación de la transcripción, probablemente debido a la dependencia energética del proceso de remodelación de la cromatina. Otros mecanismos potenciales son la velocidad de elongación de la polimerasa determinando la estructura secundaria del ARN, o las tasas de degradación de distintas variantes de splicing siendo dependientes de energía de forma diferente.*

- *Las tasas de producción y degradación del ARNm escalan con la masa mitocondrial de manera asimétrica. El escalado en la expresión de transcritos está mayoritariamente dominado por la primera.*

Mitocondria y apoptosis

- *El contenido mitocondrial discrimina el destino apoptótico de células individuales (a mayor masa mitocondrial, mayor probabilidad de muerte), y también afecta a los tiempos de muerte (células con más contenido mitocondrial tienden a morir más rápido).*

- *La mitocondria modula las abundancias de las proteínas que participan en la ruta de señalización de apoptosis. Este es el principal determinante de su poder predictivo de la respuesta apoptótica de células individuales. El papel de la mitocondria como nodo de la red de señalización de apoptosis parece ser menor en comparación.*

- *La predicción del destino apoptótico por parte de los niveles mitocondriales es óptima cuando el control mitocondrial sobre las anundancias de las proteínas pro-apoptóticas es más estricto que sobre las anti-apoptóticas, particularmente para aquellas especies en la parte de la ruta apoptótica anterior al MOMP. Esta situación, en efecto, se observa en células HeLa en cultivo así como en muestras sólidas de tumores.*

- *En muestras de cáncer de colon, las proteínas pro-apoptóticas co-varían fuertemente con el contenido mitocondrial, y las anti-apoptóticas presentan correlaciones más modestas. Estas observaciones hacen de la masa mitocondrial un*

buen candidato a biomarcador único como predictor de la sensibilidad a apoptosis de células cancerígenas.

Bibliography

1. Klipp E (2009). *Systems Biology: a textbook*. John Wiley & Sons Inc
2. Lockhart DJ and Winzeler EA (2000). Genomics, gene expression and DNA arrays. *Nature* **405(6788)**:827–836
3. Hasin Y, Seldin M and Lusis A (2017). Multi-omics approaches to disease. *Genome Biology* **18(1)**
4. Hartwell LH, Hopfield JJ, Leibler S and Murray AW (1999). From molecular to modular cell biology. *Nature* **402(6761supp)**:C47–C52
5. Ideker T, Galitski T and Hood L (2001). A new approach to decoding life: Systems Biology. *Annual Review of Genomics and Human Genetics* **2(1)**:343–372
6. Bray D (2003). Molecular Networks: The Top-Down View. *Science* **301(5641)**:1864–1865
7. Barabási AL and Oltvai ZN (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics* **5(2)**:101–113
8. Kitano H (2002). Systems Biology: A Brief Overview. *Science* **295(5560)**:1662–1664
9. Kitano H (2002). Computational systems biology. *Nature* **420(6912)**:206–210
10. Ge H, Walhout AJ and Vidal M (2003). Integrating omic' information: a bridge between genomics and systems biology. *Trends in Genetics* **19(10)**:551–560
11. Palsson B (2006). *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press
12. Vayttaden SJ, Ajay SM and Bhalla US (2004). A Spectrum of Models of Signaling Pathways. *ChemBioChem* **5(10)**:1365–1374
13. Alon U (2003). Biological Networks: The Tinkerer as an Engineer. *Science* **301(5641)**:1866–1867
14. Alon U (2006). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall/CRC
15. Michaelis L and Menten M (1913). Die Kinetic der Invertinwirkung. *Biochem Z* **49**:333–369
16. Wang RS, Saadatpour A and Albert R (2012). Boolean modeling in systems biology: an overview of methodology and applications. *Physical Biology* **9(5)**:055001

17. Murata T (1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE* **77(4)**:541–580

18. Chen WW, Niepel M and Sorger PK (2010). Classic and contemporary approaches to modeling biochemical reactions. *Genes & Development* **24(17)**:1861–1875

19. Guldberg C and Waage P (1879). Über die chemische Affinität. *Journal fur Praktische Chemie* **19**:69

20. Briggs GE and Haldane JB (1925). A Note on the Kinetics of Enzyme Action. *Biochem J* **19**:338–339

21. Haag G (2017). *Modelling with the Master Equation*. Springer International Publishing

22. Gillespie DT (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81(25)**:2340–2361

23. Gillespie DT (2000). The chemical Langevin equation. *The Journal of Chemical Physics* **113(1)**:297–306

24. Altschuler SJ and Wu LF (2010). Cellular Heterogeneity: Do Differences Make a Difference? *Cell* **141(4)**:559–563

25. Ellegren H and Galtier N (2016). Determinants of genetic diversity. *Nature Reviews Genetics* **17(7)**:422–433

26. Quintana-Murci L and Clark AG (2013). Population genetic tools for dissecting innate immunity in humans. *Nature Reviews Immunology* **13(4)**:280–293

27. Soares MP and Weiss G (2015). The Iron age of host-microbe interactions. *EMBO reports* **16(11)**:1482–1500

28. Volfson D, Marciniak J, Blake WJ, Ostroff N, Tsimring LS and Hasty J (2005). Origins of extrinsic variability in eukaryotic gene expression. *Nature* **439(7078)**:861–864

29. Elowitz MB, Levine AJ, Siggia ED and Swain PS (2002). Stochastic Gene Expression in a Single Cell. *Science* **297(5584)**:1183–1186

30. Raser JM (2004). Control of Stochasticity in Eukaryotic Gene Expression. *Science* **304(5678)**:1811–1814

31. Pedraza JM (2005). Noise Propagation in Gene Networks. *Science* **307(5717)**:1965–1969

32. Kærn M, Elston TC, Blake WJ and Collins JJ (2005). Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics* **6(6)**:451–464

33. Blake WJ, KÆrn M, Cantor CR and Collins JJ (2003). Noise in eukaryotic gene expression. *Nature* **422(6932)**:633–637

34. Raj A and van Oudenaarden A (2008). Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* **135(2)**:216–226

35. Sanchez A and Golding I (2013). Genetic Determinants and Cellular Constraints in Noisy Gene Expression. *Science* **342(6163)**:1188–1193

36. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y and Barkai N (2006). Noise in protein expression scales with natural protein abundance. *Nature Genetics* **38(6)**:636–643

37. Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL and Weissman JS (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441(7095)**:840–846

38. Raser JM (2005). Noise in Gene Expression: Origins, Consequences, and Control. *Science* **309(5743)**:2010–2013

39. Rosenfeld N (2005). Gene Regulation at the Single-Cell Level. *Science* **307(5717)**:1962–1965

40. Zopf CJ, Quinn K, Zeidman J and Maheshri N (2013). Cell-Cycle Dependence of Transcription Dominates Noise in Gene Expression. *PLoS Computational Biology* **9(7)**:e1003161

41. Selimkhanov J, Taylor B, Yao J, Pilko A, Albeck J, Hoffmann A, Tsimring L and Wollman R (2014). Accurate information transmission through dynamic biochemical signaling networks. *Science* **346(6215)**:1370–1373

42. Lestas I, Vinnicombe G and Paulsson J (2010). Fundamental limits on the suppression of molecular fluctuations. *Nature* **467(7312)**:174–178

43. Becker NB, Mugler A and ten Wolde PR (2015). Optimal Prediction by Cellular Signaling Networks. *Physical Review Letters* **115(25)**

44. Suderman R, Bachman JA, Smith A, Sorger PK and Deeds EJ (2017). Fundamental trade-offs between information flow in single cells and cellular populations. *Proceedings of the National Academy of Sciences* **114(22)**:5755–5760

45. Paulsson J, Berg OG and Ehrenberg M (2000). Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *Proceedings of the National Academy of Sciences* **97(13)**:7148–7153

46. Cai L, Dalal CK and Elowitz MB (2008). Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature* **455(7212)**:485–490

47. Eldar A and Elowitz MB (2010). Functional roles for noise in genetic circuits. *Nature* **467(7312)**:167–173

48. Snijder B and Pelkmans L (2011). Origins of regulated cell-to-cell variability. *Nature Reviews Molecular Cell Biology* **12(2)**:119–125

49. Waddington CH (2014). *The Strategy of the Genes*. ROUTLEDGE

50. Balázsi G, van Oudenaarden A and Collins JJ (2011). Cellular Decision Making and Biological Noise: From Microbes to Mammals. *Cell* **144(6)**:910–925

51. Ozbudak EM, Thattai M, Lim HN, Shraiman BI and van Oudenaarden A (2004). Multistability in the lactose utilization network of Escherichia coli. *Nature* **427(6976)**:737–740

52. Losick R and Desplan C (2008). Stochasticity and Cell Fate. *Science* **320(5872)**:65–68

53. Serizawa S, Miyamichi K and Sakano H (2004). One neuron–one receptor rule in the mouse olfactory system. *Trends in Genetics* **20(12)**:648–653

54. Graf T and Enver T (2009). Forcing cells to change lineages. *Nature* **462(7273)**:587–594

55. Hill WG and Zhang XS (2004). Effects on phenotypic variability of directional selection arising through genetic differences in residual variability. *Genetical Research* **83(2)**:121–132

56. Çağatay T, Turcotte M, Elowitz MB, Garcia-Ojalvo J and Süel GM (2009). Architecture-Dependent Noise Discriminates Functionally Analogous Differentiation Circuits. *Cell* **139(3)**:512–522

57. Snijder B, Sacher R, Rämö P, Damm EM, Liberali P and Pelkmans L (2009). Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature* **461(7263)**:520–523

58. Bahar R, Hartmann CH, Rodriguez KA, Denny AD, Busuttil RA, Dollé MET, Calder RB, Chisholm GB, Pollock BH et al. (2006). Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* **441(7096)**:1011–1014

59. Newlands S, Levitt LK, Robinson CS, Karpf AC, Hodgson VR, Wade RP and Hardeman EC (1998). Transcription occurs in pulses in muscle fibers. *Genes & Development* **12(17)**:2748–2758

60. Albeck JG, Burke JM, Spencer SL, Lauffenburger DA and Sorger PK (2008). Modeling a Snap-Action, Variable-Delay Switch Controlling Extrinsic Cell Death. *PLoS Biology* **6(12)**:e299

61. Albeck JG, Burke JM, Aldridge BB, Zhang M, Lauffenburger DA and Sorger PK (2008). Quantitative Analysis of Pathways Controlling Extrinsic Apoptosis in Single Cells. *Molecular Cell* **30(1)**:11–25

62. Spencer SL, Gaudet S, Albeck JG, Burke JM and Sorger PK (2009). Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459(7245)**:428–432

63. Spencer SL and Sorger PK (2011). Measuring and Modeling Apoptosis in Single Cells. *Cell* **144(6)**:926–939

64. Gaudet S, Spencer SL, Chen WW and Sorger PK (2012). Exploring the Contextual Sensitivity of Factors that Determine Cell-to-Cell Variability in Receptor-Mediated Apoptosis. *PLoS Computational Biology* **8(4)**:e1002482

65. Bertaux F, Stoma S, Drasdo D and Batt G (2014). Modeling Dynamics of Cell-to-Cell Variability in TRAIL-Induced Apoptosis Explains Fractional Killing and Predicts Reversible Resistance. *PLoS Computational Biology* **10(10)**:e1003893

66. Hanahan D and Weinberg RA (2000). The Hallmarks of Cancer. *Cell* **100(1)**:57–70

67. Brock A, Chang H and Huang S (2009). Non-genetic heterogeneity — a mutation-independent driving force for the somatic evolution of tumours. *Nature Reviews Genetics* **10(5)**:336–342

68. das Neves RP, Jones NS, Andreu L, Gupta R, Enver T and Iborra FJ (2010). Connecting Variability in Global Transcription Rate to Mitochondrial Variability. *PLoS Biology* **8(12)**:e1000560

69. Guantes R, Rastrojo A, Neves R, Lima A, Aguado B and Iborra FJ (2015). Global variability in gene expression and alternative splicing is modulated by mitochondrial content. *Genome Research* **25(5)**:633–644

70. Romero-Moya D, Bueno C, Montes R, Navarro-Montero O, Iborra FJ, Lopez LC, Martin M and Menendez P (2013). Cord blood-derived CD34+ hematopoietic cells with low mitochondrial mass are enriched in hematopoietic repopulating stem cell function. *Haematologica* **98(7)**:1022–1029

71. Sotgia F, Whitaker-Menezes D, Martinez-Outschoorn UE, Flomenberg N, Birbe R, Witkiewicz AK, Howell A, Philp NJ, Pestell RG et al. (2012). Mitochondrial metabolism in cancer metastasis. *Cell Cycle* **11(7)**:1445–1454

72. Guantes R, Díaz-Colunga J and Iborra FJ (2015). Mitochondria and the non-genetic origins of cell-to-cell variability: More is different. *BioEssays* **38(1)**:64–76

73. Knowles JR (1980). Enzyme-Catalyzed Phosphoryl Transfer Reactions. *Annual Review of Biochemistry* **49(1)**:877–919

74. Alberts B (2014). *Molecular Biology of the Cell*. Garland Science

75. Heiden MGV, Cantley LC and Thompson CB (2009). Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **324(5930)**:1029–1033

76. van der Giezen M (2009). Hydrogenosomes and Mitosomes: Conservation and Evolution of Functions. *Journal of Eukaryotic Microbiology* **56(3)**:221–231

77. Martin W and Müller M (1998). The hydrogen hypothesis for the first eukaryote. *Nature* **392(6671)**:37–41

78. Wagner A (2005). Energy Constraints on the Evolution of Gene Expression. *Molecular Biology and Evolution* **22(6)**:1365–1374

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

80. Johnston IG, Gaal B, das Neves RP, Enver T, Iborra FJ and Jones NS (2012). Mitochondrial Variability as a Source of Extrinsic Cellular Noise. *PLoS Computational Biology* **8(3)**:e1002416

81. Collins TJ (2002). Mitochondria are morphologically and functionally heterogeneous within cells. *The EMBO Journal* **21(7)**:1616–1627

82. Johnston IG, Burgstaller JP, Havlicek V, Kolbe T, Rülicke T, Brem G, Poulton J and Jones NS (2015). Stochastic modelling, Bayesian inference, and new in vivo measurements elucidate the debated mtDNA bottleneck mechanism. *eLife* **4**:e07464

83. Huh D and Paulsson J (2010). Non-genetic heterogeneity from stochastic partitioning at cell division. *Nature Genetics* **43(2)**:95–100

84. Huh D and Paulsson J (2011). Random partitioning of molecules at cell division. *Proceedings of the National Academy of Sciences* **108(36)**:15004–15009

85. Twig G, Hyde B and Shirihai OS (2008). Mitochondrial fusion, fission and autophagy as a quality control axis: The bioenergetic view. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1777(9)**:1092–1097

86. Twig G and Shirihai OS (2011). The Interplay Between Mitochondrial Dynamics and Mitophagy. *Antioxidants & Redox Signaling* **14(10)**:1939–1951

87. Mishra P and Chan DC (2014). Mitochondrial dynamics and inheritance during cell division, development and disease. *Nature Reviews Molecular Cell Biology* **15(10)**:634–646

88. Jajoo R, Jung Y, Huh D, Viana MP, Rafelski SM, Springer M and Paulsson J (2016). Accurate concentration control of mitochondria and nucleoids. *Science* **351(6269)**:169–172

89. Katajisto P, Dohla J, Chaffer CL, Pentinmikko N, Marjanovic N, Iqbal S, Zoncu R, Chen W, Weinberg RA et al. (2015). Asymmetric apportioning of aged mitochondria between daughter cells is required for stemness. *Science* **348(6232)**:340–343

90. Higuchi R, Vevea JD, Swayne TC, Chojnowski R, Hill V, Boldogh IR and Pon LA (2013). Actin Dynamics Affect Mitochondrial Quality Control and Aging in Budding Yeast. *Current Biology* **23(23)**:2417–2422

91. Molenaar D, van Berlo R, de Ridder D and Teusink B (2009). Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular Systems Biology* **5**

92. Hui S, Silverman JM, Chen SS, Erickson DW, Basan M, Wang J, Hwa T and Williamson JR (2015). Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology* **11(2)**:e784–e784

93. Weiße AY, Oyarzún DA, Danos V and Swain PS (2015). Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences* **112(9)**:E1038–E1047

94. Gerosa L, Kochanowski K, Heinemann M and Sauer U (2014). Dissecting specific and global transcriptional regulation of bacterial gene expression. *Molecular Systems Biology* **9(1)**:658–658

95. Elf J (2003). Selective Charging of tRNA Isoacceptors Explains Patterns of Codon Usage. *Science* **300(5626)**:1718–1722

96. Scott M, Klumpp S, Mateescu EM and Hwa T (2014). Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Molecular Systems Biology* **10(8)**:747–747

97. Coulon A, Chow CC, Singer RH and Larson DR (2013). Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nature Reviews Genetics* **14(8)**:572–584

98. Voss TC and Hager GL (2013). Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nature Reviews Genetics* **15(2)**:69–81

99. Voss TC, Schiltz RL, Sung MH, Yen PM, Stamatoyannopoulos JA, Biddie SC, Johnson TA, Miranda TB, John S et al. (2011). Dynamic Exchange at Regulatory Elements during Chromatin Remodeling Underlies Assisted Loading Mechanism. *Cell* **146(4)**:544–554

100. van Royen ME, Zotter A, Ibrahim SM, Geverts B and Houtsmuller AB (2010). Nuclear proteins: finding and binding target sites in chromatin. *Chromosome Research* **19(1)**:83–98

101. Brown CR, Mao C, Falkovskaia E, Jurica MS and Boeger H (2013). Linking Stochastic Fluctuations in Chromatin Structure and Gene Expression. *PLoS Biology* **11(8)**:e1001621

102. Small EC, Xi L, Wang JP, Widom J and Licht JD (2014). Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. *Proceedings of the National Academy of Sciences* **111(24)**:E2462–E2471

103. Hornung G, Bar-Ziv R, Rosin D, Tokuriki N, Tawfik DS, Oren M and Barkai N (2012). Noise-mean relationship in mutated promoters. *Genome Research* **22(12)**:2409–2417

104. Dadiani M, van Dijk D, Segal B, Field Y, Ben-Artzi G, Raveh-Sadka T, Levo M, Kaplow I, Weinberger A et al. (2013). Two DNA-encoded strategies for increasing expression with opposing effects on promoter dynamics and transcriptional noise. *Genome Research* **23(6)**:966–976

105. Dey SS, Foley JE, Limsirichai P, Schaffer DV and Arkin AP (2015). Orthogonal control of expression mean and variance by epigenetic features at different genomic loci. *Molecular Systems Biology* **11(5)**:806–806

106. Wallace DC and Fan W (2010). Energetics, epigenetics, mitochondrial genetics. *Mitochondrion* **10(1)**:12–31

107. Darzacq X, Shav-Tal Y, de Turris V, Brody Y, Shenoy SM, Phair RD and Singer RH (2007). In vivo dynamics of RNA polymerase II transcription. *Nature Structural & Molecular Biology* **14(9)**:796–806

108. Larson DR, Zenklusen D, Wu B, Chao JA and Singer RH (2011). Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene. *Science* **332(6028)**:475–478

109. Yang S, Kim S, Lim YR, Kim C, An HJ, Kim JH, Sung J and Lee NK (2014). Contribution of RNA polymerase concentration variation to protein expression noise. *Nature Communications* **5(1)**

110. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP and Burge CB (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* **456(7221)**:470–476

111. Melamud E and Moult J (2009). Stochastic noise in splicing machinery. *Nucleic Acids Research* **37(14)**:4873–4886

112. Waks Z, Klein AM and Silver PA (2014). Cell-to-cell variability of alternative RNA splicing. *Molecular Systems Biology* **7(1)**:506–506

113. Nilsen TW and Graveley BR (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463(7280)**:457–463

114. Kalsotra A and Cooper TA (2011). Functional consequences of developmentally regulated alternative splicing. *Nature Reviews Genetics* **12(10)**:715–729

115. Cooper TA, Wan L and Dreyfuss G (2009). RNA and Disease. *Cell* **136(4)**:777–793

116. Pickrell JK, Pai AA, Gilad Y and Pritchard JK (2010). Noisy Splicing Drives mRNA Isoform Diversity in Human Cells. *PLoS Genetics* **6(12)**:e1001236

117. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, Schwartz S, Yosef N, Malboeuf C et al. (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* **498(7453)**:236–240

118. Braunschweig U, Gueroussov S, Plocik AM, Graveley BR and Blencowe BJ (2013). Dynamic Integration of Splicing within Gene Regulatory Pathways. *Cell* **152(6)**:1252–1269

119. Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E and Muñoz MJ (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nature Reviews Molecular Cell Biology* **14(3)**:153–165

120. Cookson NA, Cookson SW, Tsimring LS and Hasty J (2009). Cell cycle-dependent variations in protein concentration. *Nucleic Acids Research* **38(8)**:2676–2681

121. Zhurinsky J, Leonhard K, Watt S, Marguerat S, Bähler J and Nurse P (2010). A Coordinated Global Control over Cellular Transcription. *Current Biology* **20(22)**:2010–2015

122. Kempe H, Schwabe A, Crémazy F, Verschure PJ and Bruggeman FJ (2015). The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise. *Molecular Biology of the Cell* **26(4)**:797–804

123. Ginzberg MB, Kafri R and Kirschner M (2015). On being the right (cell) size. *Science* **348(6236)**:1245075–1245075

124. Miettinen TP, Pessa HK, Caldez MJ, Fuhrer T, Diril MK, Sauer U, Kaldis P and Björklund M (2014). Identification of Transcriptional and Metabolic Programs Related to Mammalian Cell Size. *Current Biology* **24(6)**:598–608

125. Arciuch VGA, Elguero ME, Poderoso JJ and Carreras MC (2012). Mitochondrial Regulation of Cell Cycle and Proliferation. *Antioxidants & Redox Signaling* **16(10)**:1150–1180

126. McDavid A, Dennis L, Danaher P, Finak G, Krouse M, Wang A, Webster P, Beechem J and Gottardo R (2014). Modeling Bi-modality Improves Characterization of Cell Cycle on Gene Expression in Single Cells. *PLoS Computational Biology* **10(7)**:e1003696

127. Wang R and Green DR (2012). Metabolic checkpoints in activated T cells. *Nature Immunology* **13(10)**:907–915

128. Pearce EL and Pearce EJ (2013). Metabolic Pathways in Immune Cell Activation and Quiescence. *Immunity* **38(4)**:633–643

129. Brand MD, Affourtit C, Esteves TC, Green K, Lambert AJ, Miwa S, Pakay JL and Parker N (2004). Mitochondrial superoxide: production, biological effects, and activation of uncoupling proteins. *Free Radical Biology and Medicine* **37(6)**:755–767

130. Devadas S, Zaritskaya L, Rhee SG, Oberley L and Williams MS (2002). Discrete Generation of Superoxide and Hydrogen Peroxide by T Cell Receptor Stimulation. *The Journal of Experimental Medicine* **195(1)**:59–70

131. Maciolek JA, Pasternak JA and Wilson HL (2014). Metabolism of activated T lymphocytes. *Current Opinion in Immunology* **27**:60–74

132. Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE and Melton DA (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature Biotechnology* **26(7)**:795–797

133. Du J, Wang Y, Hunter R, Wei Y, Blumenthal R, Falke C, Khairova R, Zhou R, Yuan P et al. (2009). Dynamic regulation of mitochondrial function by glucocorticoids. *Proceedings of the National Academy of Sciences* **106(9)**:3543–3548

134. Lee KS, Wu Z, Song Y, Mitra SS, Feroze AH, Cheshier SH and Lu B (2013). Roles of PINK1, mTORC2, and mitochondria in preserving brain tumor-forming stem cells in a non-canonical Notch signaling pathway. *Genes & Development* **27(24)**:2642–2647

135. Cioffi F, Senese R, Lanni A and Goglia F (2013). Thyroid hormones and mitochondria: With a brief look at derivatives and analogues. *Molecular and Cellular Endocrinology* **379(1-2)**:51–61

136. Wallace DC (2013). A mitochondrial bioenergetic etiology of disease. *Journal of Clinical Investigation* **123(4)**:1405–1412

137. Matlin AJ, Clark F and Smith CWJ (2005). Understanding alternative splicing: towards a cellular code. *Nature Reviews Molecular Cell Biology* **6(5)**:386–398

138. Kong SW, Hu YW, Ho JW, Ikeda S, Polster S, John R, Hall JL, Bisping E, Pieske B et al. (2010). Heart Failure–Associated Changes in RNA Splicing of Sarcomere Genes. *Circulation: Cardiovascular Genetics* **3(2)**:138–146

139. Ballinger SW (2005). Mitochondrial dysfunction in cardiovascular disease. *Free Radical Biology and Medicine* **38(10)**:1278–1295

140. Fillmore N and Lopaschuk GD (2013). Targeting mitochondrial oxidative metabolism as an approach to treat heart failure. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1833(4)**:857–865

141. Bayeva M, Gheorghiade M and Ardehali H (2013). Mitochondria as a Therapeutic Target in Heart Failure. *Journal of the American College of Cardiology* **61(6)**:599–610

142. Mortensen SA, Rosenfeldt F, Kumar A, Dolliner P, Filipiak KJ, Pella D, Alehagen U, Steurer G and Littarru GP (2014). The Effect of Coenzyme Q 10 on Morbidity and Mortality in Chronic Heart Failure. *JACC: Heart Failure* **2(6)**:641–649

143. Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, Mill J, Cantor RM, Blencowe BJ et al. (2011). Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474(7351)**:380–384

144. Ward PS and Thompson CB (2012). Metabolic Reprogramming: A Cancer Hallmark Even Warburg Did Not Anticipate. *Cancer Cell* **21(3)**:297–308

145. Hanahan D and Weinberg RA (2011). Hallmarks of Cancer: The Next Generation. *Cell* **144(5)**:646–674

146. Martín-Martín N, Carracedo A and Torrano V (2018). Metabolism and Transcription in Cancer: Merging Two Classic Tales. *Frontiers in Cell and Developmental Biology* **5**

147. Ladomery M (2013). Aberrant Alternative Splicing Is Another Hallmark of Cancer. *International Journal of Cell Biology* **2013**:1–6

148. Stern S, Dror T, Stolovicki E, Brenner N and Braun E (2007). Genome-wide transcriptional plasticity underlies cellular adaptation to novel challenge. *Molecular Systems Biology* **3**

149. Keren L, Zackay O, Lotan-Pompan M, Barenholz U, Dekel E, Sasson V, Aidelberg G, Bren A, Zeevi D et al. (2014). Promoters maintain their relative activity levels under different growth conditions. *Molecular Systems Biology* **9(1)**:701–701

150. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W and Selbach M (2011). Global quantification of mammalian gene expression control. *Nature* **473(7347)**:337–342

151. Jackson DA, Iborra FJ, Manders EM and Cook PR (1998). Numbers and Organization of RNA Polymerases, Nascent Transcripts, and Transcription Units in HeLa Nuclei. *Molecular Biology of the Cell* **9(6)**:1523–1536

152. McAdams HH and Arkin A (1997). Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences* **94(3)**:814–819

153. Andrews S (2010). *FastQC: a quality control tool for high throughput sequence data.* http://www.bioinformatics.babraham.ac.uk/projects/fastqc

154. Patro R, Duggal G, Love MI, Irizarry RA and Kingsford C (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14(4)**:417–419

155. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A et al. (2017). Ensembl 2018. *Nucleic Acids Research* **46(D1)**:D754–D761

156. R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, http://www.R-project.org/

157. Durinck S, Moreau Y, Kasprzyk A, Davis S, Moor BD, Brazma A and Huber W (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21(16)**:3439–3440

158. Durinck S, Spellman PT, Birney E and Huber W (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols* **4(8)**:1184–1191

159. Soneson C, Love MI and Robinson MD (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**:1521

160. Love MI, Huber W and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15(12)**

161. Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET and Eliceiri KW (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* **18(1)**

162. Jovanovic M, Rooney MS, Mertins P, Przybylski D, Chevrier N, Satija R, Rodriguez EH, Fields AP, Schwartz S et al. (2015). Dynamic profiling of the protein life cycle in response to pathogens. *Science* **347(6226)**:1259038–1259038

163. Wagner GP, Kin K and Lynch VJ (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences* **131(4)**:281–285

164. Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5(7)**:621–628

165. Tabas-Madrid D, Nogales-Cadenas R and Pascual-Montano A (2012). GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Research* **40(W1)**:W478–W483

166. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS et al. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics* **25(1)**:25–29

167. Kanehisa M, Sato Y, Furumichi M, Morishima K and Tanabe M (2018). New approach for understanding genome variations in KEGG. *Nucleic Acids Research* **47(D1)**:D590–D595

168. Thomas PD (2003). PANTHER: A Library of Protein Families and Subfamilies Indexed by Function. *Genome Research* **13(9)**:2129–2141

169. Reyes A and Huber W (2017). Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. *Nucleic Acids Research* **46(2)**:582–592

170. Solnick D (1985). Alternative splicing caused by RNA secondary structure. *Cell* **43(3)**:667–676

171. Shepard PJ and Hertel KJ (2008). Conserved RNA secondary structures promote alternative splicing. *RNA* **14(8)**:1463–1469

172. Fallmann J, Will S, Engelhardt J, Grüning B, Backofen R and Stadler PF (2017). Recent advances in RNA folding. *Journal of Biotechnology* **261**:97–104

173. Tan Z, Fu Y, Sharma G and Mathews DH (2017). TurboFold II: RNA structural alignment and secondary structure prediction informed by multiple homologs. *Nucleic Acids Research* **45(20)**:11570–11581

174. Lorenz R, Wolfinger MT, Tanzer A and Hofacker IL (2016). Predicting RNA secondary structures from sequence and probing data. *Methods* **103**:86–98

175. Waterman M and Smith T (1978). RNA secondary structure: a complete mathematical analysis. *Mathematical Biosciences* **42(3-4)**:257–266

176. Yankulov K, Yamashita K, Roy R, Egly JM and Bentley DL (1995). The Transcriptional Elongation Inhibitor 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole Inhibits Transcription Factor IIH-associated Protein Kinase. *Journal of Biological Chemistry* **270(41)**:23922–23925

177. Arora S, Pattwell SS, Holland EC and Bolouri H (2018). Uncertainty in RNA-seq gene expression data. *bioRxiv*

178. Deneke C, Lipowsky R and Valleriani A (2013). Complex Degradation Processes Lead to Non-Exponential Decay Patterns and Age-Dependent Decay Rates of Messenger RNA. *PLoS ONE* **8(2)**:e55442

179. Hausser J, Mayo A, Keren L and Alon U (2019). Central dogma rates and the trade-off between precision and economy in gene expression. *Nature Communications* **10(1)**

180. Longley D and Johnston P (2005). Molecular mechanisms of drug resistance. *The Journal of Pathology* **205(2)**:275–292

181. Flusberg DA, Roux J, Spencer SL and Sorger PK (2013). Cells surviving fractional killing by TRAIL exhibit transient but sustainable resistance and inflammatory phenotypes. *Molecular Biology of the Cell* **24(14)**:2186–2200

182. Ballweg R, Paek AL and Zhang T (2017). A dynamical framework for complex fractional killing. *Scientific Reports* **7(1)**

183. Holohan C, Schaeybroeck SV, Longley DB and Johnston PG (2013). Cancer drug resistance: an evolving paradigm. *Nature Reviews Cancer* **13(10)**:714–726

184. Settleman J (2016). Bet on drug resistance. *Nature* **529(7586)**:289–290

185. Alizadeh AA, Aranda V, Bardelli A, Blanpain C, Bock C, Borowski C, Caldas C, Califano A, Doherty M et al. (2015). Toward understanding and exploiting tumor heterogeneity. *Nature Medicine* **21(8)**:846–853

186. Marusyk A, Almendro V and Polyak K (2012). Intra-tumour heterogeneity: a looking glass for cancer? *Nature Reviews Cancer* **12(5)**:323–334

187. Sero JE, Sailem HZ, Ardy RC, Almuttaqi H, Zhang T and Bakal C (2015). Cell shape and the microenvironment regulate nuclear translocation of NF- B in breast epithelial and tumor cells. *Molecular Systems Biology* **11(3)**:790–790

188. Gligorijevic B, Bergman A and Condeelis J (2014). Multiparametric Classification Links Tumor Microenvironments with Tumor Cell Phenotype. *PLoS Biology* **12(11)**:e1001995

189. Mumenthaler SM, Foo J, Choi NC, Heise N, Leder K, Agus DB, Pao W, Michor F and Mallick P (2015). The Impact of Microenvironmental Heterogeneity on the Evolution of Drug Resistance in Cancer Cells. *Cancer Informatics* **14s4**:CIN.S19338

190. Battich N, Stoeger T and Pelkmans L (2015). Control of Transcript Variability in Single Mammalian Cells. *Cell* **163(7)**:1596–1610

191. Tait SWG and Green DR (2010). Mitochondria and cell death: outer membrane permeabilization and beyond. *Nature Reviews Molecular Cell Biology* **11(9)**:621–632

192. Rehm M, Düßmann H, Jänicke RU, Tavaré JM, Kögel D and Prehn JHM (2002). Single-cell Fluorescence Resonance Energy Transfer Analysis Demonstrates That Caspase Activation during Apoptosis Is a Rapid Process. *Journal of Biological Chemistry* **277(27)**:24506–24514

193. Goldstein JC, Waterhouse NJ, Juin P, Evan GI and Green DR (2000). The coordinate release of cytochrome c during apoptosis is rapid, complete and kinetically invariant. *Nature Cell Biology* **2(3)**:156–162

194. Bhola PD and Simon SM (2009). Determinism and divergence of apoptosis susceptibility in mammalian cells. *Journal of Cell Science* **122(23)**:4296–4302

195. Rehm M, Huber HJ, Hellwig CT, Anguissola S, Dussmann H and Prehn JHM (2009). Dynamics of outer mitochondrial membrane permeabilization during apoptosis. *Cell Death & Differentiation* **16(4)**:613–623

196. Dimberg LY, Anderson CK, Camidge R, Behbakht K, Thorburn A and Ford HL (2012). On the TRAIL to successful cancer therapy? Predicting and counteracting resistance against TRAIL-based therapeutics. *Oncogene* **32(11)**:1341–1350

197. Elmore S (2007). Apoptosis: A Review of Programmed Cell Death. *Toxicologic Pathology* **35(4)**:495–516

198. Díaz-Colunga J, Márquez-Jurado S, das Neves RP, Martinez-Lorente A, Almazán F, Guantes R and Iborra FJ (2018). Mitochondrial levels determine variability in cell death by modulating apoptotic gene expression. *Nature Communications* **9(1)**

199. Vermeulen K, Berneman ZN and Bockstaele DRV (2003). Cell cycle and apoptosis. *Cell Proliferation* **36(3)**:165–175

200. GOLDSTEIN JC, KLUCK RM and GREEN DR (2006). A Single Cell Analysis of Apoptosis: Ordering the Apoptotic Phenotype. *Annals of the New York Academy of Sciences* **926(1)**:132–141

201. Roux J, Hafner M, Bandara S, Sims JJ, Hudson H, Chai D and Sorger PK (2015). Fractional killing arises from cell-to-cell variability in overcoming a caspase activity threshold. *Molecular Systems Biology* **11(5)**:803–803

202. Zi Z (2011). Sensitivity analysis approaches applied to systems biology models. *IET Systems Biology* **5(6)**:336–346

203. Almendro V, Cheng YK, Randles A, Itzkovitz S, Marusyk A, Ametller E, Gonzalez-Farre X, Muñoz M, Russnes HG et al. (2014). Inference of Tumor Evolution during Chemotherapy by Computational Modeling and In Situ Analysis of Genetic and Phenotypic Cellular Diversity. *Cell Reports* **6(3)**:514–527

204. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, Ritchey JK, Young MA, Lamprecht T et al. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481(7382)**:506–510

205. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D et al. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* **472(7341)**:90–94

206. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344(6190)**:1396–1401

207. Chonghaile TN, Sarosiek KA, Vo TT, Ryan JA, Tammareddi A, Moore VDG, Deng J, Anderson KC, Richardson P et al. (2011). Pretreatment Mitochondrial Priming Correlates with Clinical Response to Cytotoxic Chemotherapy. *Science* **334(6059)**:1129–1133

208. Sarosiek KA, Chonghaile TN and Letai A (2013). Mitochondria: gatekeepers of response to chemotherapy. *Trends in Cell Biology* **23(12)**:612–619

209. Green DR, Galluzzi L and Kroemer G (2014). Metabolic control of cell death. *Science* **345(6203)**:1250256–1250256

210. Škrtić M, Sriskanthadevan S, Jhas B, Gebbia M, Wang X, Wang Z, Hurren R, Jitkova Y, Gronda M et al. (2011). Inhibition of Mitochondrial Translation as a Therapeutic Strategy for Human Acute Myeloid Leukemia. *Cancer Cell* **20(5)**:674–688

211. Guha M, Srinivasan S, Ruthel G, Kashina AK, Carstens RP, Mendoza A, Khanna C, Winkle TV and Avadhani NG (2013). Mitochondrial retrograde signaling induces epithelial–mesenchymal transition and generates breast cancer stem cells. *Oncogene* **33(45)**:5238–5250

212. Yang J and Weinberg RA (2008). Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis. *Developmental Cell* **14(6)**:818–829

213. Gaude E and Frezza C (2016). Tissue-specific and convergent metabolic transformation of cancer correlates with metastatic potential and patient survival. *Nature Communications* **7(1)**

214. Zeestraten EC, Benard A, Reimers MS, Schouten PC, Liefers GJ, van de Velde CJ and Kuppen PJ (2013). The Prognostic Value of the Apoptosis Pathway in Colorectal Cancer: A Review of the Literature on Biomarkers Identified by Immunohistochemistry. *Biomarkers in Cancer* **5**:BIC.S11475

215. Brown JM, Leach J, Reittie JE, Atzberger A, Lee-Prudhoe J, Wood WG, Higgs DR, Iborra FJ and Buckle VJ (2006). Coregulated human globin genes are frequently in spatial proximity when active. *The Journal of Cell Biology* **172(2)**:177–187

216. Brown JM, Green J, das Neves RP, Wallace HA, Smith AJ, Hughes J, Gray N, Taylor S, Wood WG et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *The Journal of Cell Biology* **182(6)**:1083–1097

217. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9(7)**:676–682

List of Figures

List of Tables

Appendices

Appendix A

Experimental methods

Cell culture and mitochondrial mass quantification

HeLa (ATCC CCL-2) cells were grown in DMEM (Gibco)–GlutaMAX-I supplemented with 10% fetal bovine serum (FBS, Hyclone) and penicillin–streptomycin (Sigma) in a 37°C humidified incubator with ~5% CO_2. Mitochondrial mass for in vivo experiments was measured as the integrated signal of MitoTracker green FM (MG, Molecular Probes) incorporated by individual cells. In fixed cells, mitochondrial mass was measured using MitoTracker red CMXRos (CMXRos, Molecular Probes).

Transcription and translation activities

Transcriptional activity was monitored in CMXRos stained cells by BrU incorporation after 30min as described in das Neves et al., 2010.[68] Controls were performed by incubation for 1h with 100μM DRB or for 1h with 1μg/mL actinomycin D prior to BrU incubation, abolishing BrU incorporation completely.

Translational activity in CMXRos-stained cells was monitored after 30min of incorporation of the methionine analogous L-homopropargylglycine (AHA) and the Click-iT HPG Alexa Fluor 488 Protein Synthesis Assay Kit, following manufacturer guidelines. Controls were performed by incubation for 30min with 1mM cycloheximide, which abolished AHA incorporation into nascent proteins (data not shown).

Cell sorting

HeLa cells were stained with MitoTracker green for 40min in DMEM. After staining, cells were washed twice with PBS, trypsinized and resuspended in PBS with 5mM EDTA. Then, cells were sorted on a fluorescence-activated cell sorter MoFlo XDP (Beckman Coulter) into two populations of 10^6 cells each, with high and low mitochon-

drial content respectively. The difference in mitochondrial mass across subpopulations was around 5-fold.

RNA extraction and sequencing

Total RNA was extracted using RNeasy Mini Kit (QIAGEN) according to the manufacturer guidelines. The quality of the extracted was measured by RNA Integrity Number (RIN) value from Bioanalyzer, being higher than 8 for all samples. 3μg of purified RNA were sequenced at the SNP&SEQ facility (Science for Life laboratory, Uppsala sequencing node). Total RNA was depleted from rRNA prior to library construction. One lane per sample was used in a 60bp paired-end run on an Illumina HiSeq 2500 sequencer. For each sample, over 50 million paired-end reads were sequenced.

RNA scaling factor determination

RNA FISH was performed as described in Brown et al., 2006 and 2008.[215,216] Per-cell RNA content was quantified as the integrated signal of poly(T) intensity. After quantification, cells were classified according to their mitochondrial content (as reported by MitoTracker green) and the ratio between RNA levels in the subpopulations with "high" and "low" mitochondria was obtained.

TRAIL apoptosis assays

HeLa cells were seeded in 24-well plates (Nunc) and incubated with increasing doses from 2 to 250 ng/ml of human recombinant TRAIL (Milipore) for 24h. After the treatment, both the dead-suspended cells and the live-adherent cells were collected. Then, the cells were washed twice with PBS and stained with Annexin V-FITC/PI (Propidium Iodide). Apoptotic analysis was performed using a FACSCalibur flow cytometer.

Live cell microscopy

HeLa cells were seeded in 24-well plates (Falcon) 1 day before the experiments. Prior to addition of apoptotic inducers, the cells were stained for 40min with MG and washed twice with DMEM. 15-30 min prior to the start of the movie, cells were added to the culture medium: TRAIL (at the indicated dilution), or 63ng/ml of DRB, or 2.5μg/ml of CHX, or a combination of 2.5μg/ml of CHX plus TNF at 20ng/ml. Treated cells were imaged at 15min intervals for 24h in a 37°C humidified chamber in ~5% CO_2. The cells were imaged at 20x magnification (0.4 NA HCX PL FL) on a Leica DMi6000b microscope (Leica MicroSystem) equipped with a Hamamatsu Orca-R2 digital CCD Camera, and the images were acquired using the LAS AF 2.7 software (Leica MicroSystem). Time to death was monitored by morphological changes associated with apoptosis. The images were analyzed using Fiji 2.0.0-rc-43 software.[217] Mitochondrial levels were

quantified from the first fluorescence image, and cell fates at the end of the experiment were determined by morphological changes associated with apoptosis.

Immunostaining using wide confocal cytometry

HeLa cells growing on coverslips were fixed and proteins indirectly immunolabeled using the corresponding primary antibodies. Secondary antibodies were Alexa Fluor 488, 546 or 647 donkey anti-mouse, goat or rabbit IgG (H+L) (Invitrogen). The coverslips or slides were mounted in Vectashield (Vector Laboratories). The images of the labeled cells were collected in a Leica TCS Sp5 multispectral confocal system (Leica MicroSystem), with a 20x 0.7 HCX PL APO CS, with the pinhole completely opened in order to collect the maximum amount of light emitted by the specimen. Hundreds of cells in different fields of the slide were collected. These images were exported to and analysed with MetaMorph 7.8.0.0 software (Molecular Devices).

For colon cancer immuno-histochemistry, tumour biopsies were formalin fixed and paraffin embedded. Tissue sections (5µm) were treated with EnVision FLEX Target retrieval solution low pH (DAKO) (95°C, 2min) in order to unmask the antigens. The immunolabeling was performed in the same way as that for the cultured cells.

The protein antibodies were used at dilution 1:1000 and were purchased from Abcam: Flip (ab167409), XIAP (ab137392), Aconitase 2 (ab110321 and ab99467), Bar (ab106547) and DR5 (ab8416); Santa Cruz Biotechnology: Bak (sc832) and Bax (sc493); Cell Signaling Technology: Smac/Diablo (15,108); Cusabio Biotech: Bcl-10 (CSB-PA002608ESR2HU); and from Sigma Prestige Antibodies: Casp8 (HPA005688), Casp9 (HPA001473), Bcl-2 (B3170) and Mcl-1 (HPA008455).

Human colorectal samples

Human colorectal tumor biopsies from de-identified patients were obtained with signed patient-informed consent and approval from the Human Ethics Review Committee of the Torrevieja and Vinalopó Hospitals.

Appendix B

Log-normal co-sampling

Consider two variables a and b that are log-normally distributed and correlated. If we want to sample one of them once the other is known, we need to account for this correlation. This can be done as follows: let us begin by defining a set of transformed variables as

$$\begin{aligned} x &\equiv \log a \\ y &\equiv \log b \end{aligned} \tag{B.1}$$

Since a and b are log-normally distributed, both x and y follow a normal distribution. Given the means of x and y (μ_x and μ_y), their standard deviations (σ_x and σ_y) and their Pearson's correlation coefficient (ρ), the probability of finding a pair of values (x, y) is given by the bivariate normal distribution $P(x, y)$:

$$P(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[\frac{-z}{2(1-\rho)}\right] \tag{B.2}$$

with

$$z \equiv \left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) \tag{B.3}$$

The expression for $P(x, y)$ can be expressed as

$$P(x,y) = \underbrace{\frac{1}{\sqrt{2\pi}\sigma_x} \exp\left[\frac{-(x-\mu_x)^2}{2\sigma_x^2}\right]}_{P(x)} \times \underbrace{\frac{1}{\sqrt{2\pi}\sigma_{yi}} \exp\left[\frac{-(y-\mu_{yi})^2}{2\sigma_{yi}^2}\right]}_{P(y \mid x)} \tag{B.4}$$

where we have defined

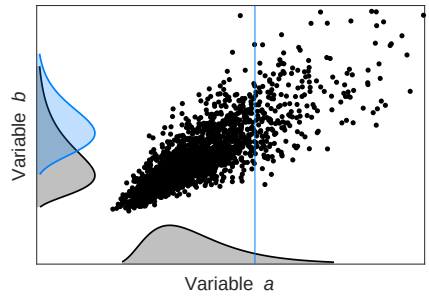


Fig. B.1. Log-normal co-sampling. Both variables a and b are log-normally distributed (gray distributions) and correlated ($\rho = 0.8$ in this example). Fixing the value of a (blue line) constrains the probability density function (blue distribution) for the subsequent sampling of b.

$$\begin{aligned} \mu_{yi} &\equiv \mu_y + \rho \frac{\sigma_y}{\sigma_x} (x - \mu_x) \\ \sigma_{yi} &\equiv \sigma_y \sqrt{1 - \rho^2} \end{aligned} \tag{B.5}$$

Expression B.4 is the product of two probabilities. The first one is identified as the the probability of obtaining a value x when sampling from a univariate normal distribution $P(x)$. The second one represents the conditional probability of obtaining a value y once that x has been fixed ($P(y|x)$). $P(y|x)$ is just a normal distribution with new mean μ_{yi} and standard deviation σ_{yi} that depend on the value of x set and the correlation between x and y.

Log-normally co-sampling a and b is just interpreted as first sampling a value for x from $P(x)$, then doing the same for y from $P(y|x)$ and finally reverting the logarithmic transformation defined in equation B.1. The mean and standard deviation of $P(y|x)$ are constrained by the initially sampled value of x, (with the constraint being tighter the higher the correlation ρ, figure B.1).