 RESEARCH ARTICLE

Dissecting Stochasticity in Translation: Time in Start Codon Selection as a Case-study

Marco Casali[*], Francesca Merlin[†], Alberto Vianelli[‡]

With few exceptions, biological research dealing with stochasticity in molecular and cellular biology has focused on transcriptional processes (e.g., Elowitz et al. 2002; Lipniacki et al. 2006). Consequently, philosophical analyses have also centered on its role in transcription (e.g., Casali and Merlin 2020). Nonetheless, recent empirical evidence clearly shows that stochasticity is also pervasive in translation (e.g., Boersma et al. 2019). Focusing on time and a non-canonical case of translation (i.e., standard vs. alternative start codon selection), this paper aims to disentangle a way in which stochasticity can intervene in models of this intracellular process. Having provided a definition of what we mean by stochasticity, first, we introduce two models of time in order to show where and how stochasticity can make a difference in the timing of translation processes. Second, we show how this difference can be biologically relevant for cell physiology. We conclude that, from an explanatory point of view, it is thus worth dissecting stochasticity when studying translation, gene expression and, more generally, biological processes at the molecular level.

Keywords

chance • explanation • gene expression • stochasticity • translation • start codon selection

1. Introduction

The aim of this article is to dissect a way in which stochasticity can intervene in models of translation. Translation corresponds to the second step of gene expression, where proteins are synthesized from mRNA (messenger RNA). The parameter of time and a non-canonical case of translation (i.e., standard vs. alternative start codon selection) are our focus. We show how the timing of this specific translation process can be stochastic, and how this should be taken into account because of its biological and explanatory relevance. More generally, this paper does not

[*]Institut supérieur de philosophie, Université catholique de Louvain, Belgium, marcomarcellocasali@gmail.com
[†]IHPST, CNRS & University Paris 1 Panthéon-Sorbonne, France, francesca.merlin@cnrs.fr
[‡]Department of Human Sciences and Innovation for the Territory (DiSUIT), University of Insubria, Italy, Alberto.Vianelli@uninsubria.it

Received 01 December 2023; Revised 11 August 2024; Accepted 05 December 2024
doi:10.3998/ptpbio.5321

deal with the sources of stochasticity[1] at the molecular level. Rather, it suggests extending the study of the role of stochasticity in gene expression by looking more closely at the parameters involved in its probabilistic models. The paper is focused on time, but other parameters can be analyzed by adopting the same attitude of "dissecting stochasticity," such as the amount of mRNA available (which does not affect what happens at the level of a single mRNA), the speed of translation, the spatial distribution of molecules and intracellular space (Fusco et al. 2003; Savulescu et al. 2020), and the energy availability (e.g., ATP distribution) within a single cell (Carthew 2021). The present work is thus a preliminary study, which aims at encouraging research on the many ways in which stochasticity may be involved in translation, gene expression and, more generally, molecular biological dynamics.

1.1. The Study of Gene Expression: Deterministic vs. Stochastic Scenarios

Since the 1990s, molecular biological systems and their evolution, in particular cellular differentiation, have been described by means of probabilistic models. Stochasticity has been invoked to account for the rise of variability in the fate of cells that are isogenic and grow in a homogenous and constant environment. However, the very notion of stochasticity[2] invoked in this context appears unclear, in that it is often used to characterize a vague attitude toward the system under study. The word "stochastic" is frequently used as an adjective to describe things such as "stochastic models" (Kaern et al. 2005, 452), "stochastic fluctuations" (461) and "stochastic effects" (451). Nonetheless, exactly which elements of the system (e.g., a cell) are considered as behaving stochastically is rarely spelled out (with some exceptions, see, e.g., Li et al. 2022). Even when biologists borrow probabilistic models from stochastic chemical kinetics to quantitatively describe the evolution of biological systems at the molecular level, usually a noise term is simply added to deterministic equations[3] in order to account for stochastic fluctuations. Other times, the state of the system as a whole is probabilistically modeled.[4] However, no clue is given as to which elements of the system modeled are concerned with such stochasticity (e.g., Rao et al. 2002, 232, 234).

In this paper, we further explore this point, because we consider its clarification to be crucial for understanding, first, why a system is considered to be stochastic and modeled as such; and, second, what difference its stochastic character makes in both empirical (biological) and explanatory terms. What is meant by stochasticity in molecular biological systems that are characterized as such? More precisely, what elements, corresponding to which model parameters and/or variables, can be conceived of as stochastic? And what does that mean? Note that answers to the majority of these questions are specific to the particular system under study. However, in all cases, the last question is the most fundamental and has to be answered first in order to provide

1. The sources of stochasticity are commonly ascribed to the following physical phenomena. First, thermal agitation, i.e., molecules inside the cell continuously move around and collide with each other; second, the low concentration of molecules involved in gene expression and the non-homogenous character of the intracellular environment; third, quantum events that can affect the biochemical bonds between molecules (Kaern et al. 2005; see also Casali and Merlin 2019).

2. In this paper, we use the term "stochasticity" rather than "chance" in order to follow the literature in molecular and cellular biology, where it is more commonly used.

3. Namely, the Langevin equation – a differential deterministic equation describing the evolution of the system in continuous and quantitative terms (in terms of concentration of molecules), to which is added a random variable (a noise term).

4. By using the Master equation, which describes the evolution of a probability distribution over the possible states of a system in discrete and quantitative terms (in terms of the number of molecules), or by using the Fokker-Planck equation, which describes the evolution over time of the probability density of the possible states of the system in quantitative and continuous terms (in terms of the concentration of molecules).

an answer to the others – which is why we shall start from it. What does it mean to say that "something" can be conceived of as stochastic? And what can this "something" be?

To find an answer, we propose to look at Symmons and Raj's (2016) analysis of practices in single-cell biology. These authors provide an account of stochasticity which seems to us general enough to cope with our question, and to be uncommitted to any specific view about the (in)deterministic nature of the biological systems under study. They claim that the differentiation of two cells can be modeled by means of two different scenarios, one deterministic, the other stochastic. In the deterministic scenario, cells differ because they receive different "instructions."[5] In the stochastic scenario, the "instructions" are the same for both cells, which nonetheless come to differ. In other words, a many-to-many mapping corresponds to the deterministic scenario, while a one-to-many mapping represents the stochastic (fig. 1).

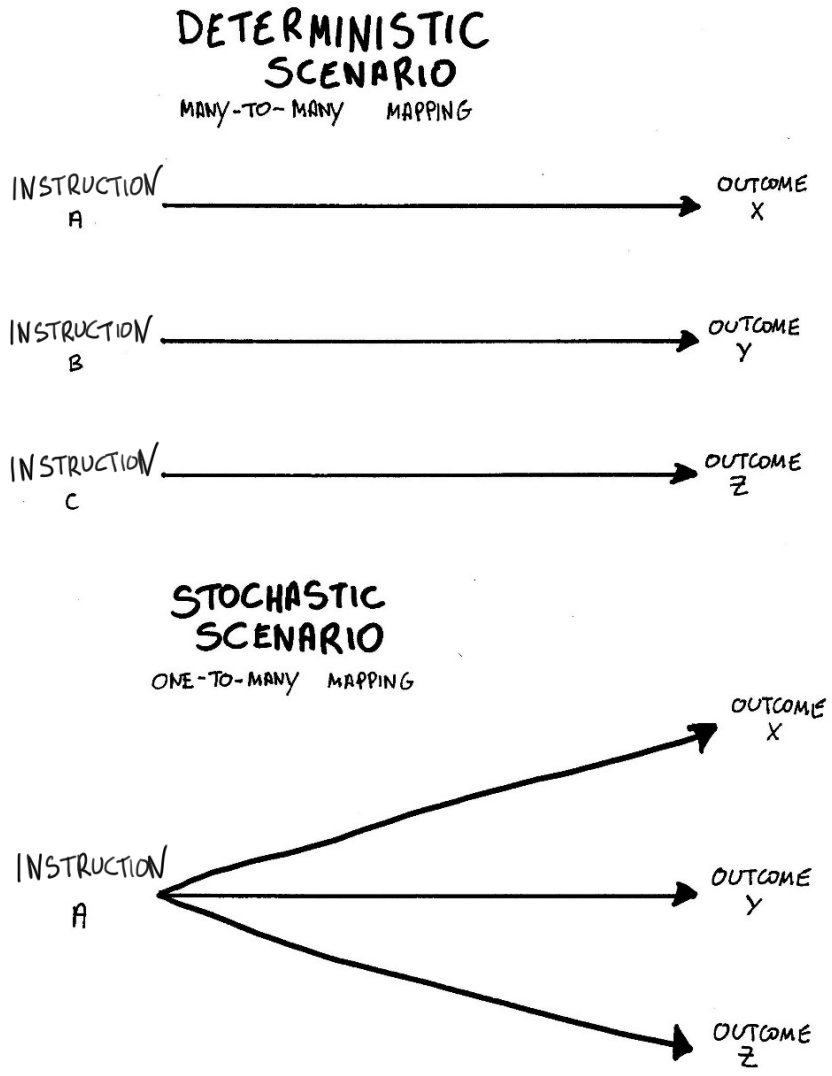

Figure 1: Top: a deterministic scenario, represented by a many-to-many mapping. Bottom: a stochastic scenario, represented by a one-to-many mapping. Figure created by Marco Casali.

5. "Instructions" is not our term but Symmons and Raj's. In using it, we do not refer to genetic information, but to the whole state of the cell at some (initial) point in time. We thus consider that information ("instructions") is diffused throughout the cell and not just contained in genes. We thank an anonymous referee for helping us to clarify this point.

Symmons and Raj thus characterize a "stochastic" situation as one in which, starting from a given situation at a given time (e.g., the same set of genes in a given [intracellular] environment just before the initiation of gene expression), different outcomes can take place at a later point in time (e.g., quantitative differences in the expression of the same gene; the expression of different genes leading to differences in cell fate). In our terms, a stochastic system is one that is modeled as follows: starting from the same subset of initial conditions, it has more than one possible future evolution. This means that, when the system actually evolves in a certain way, it could have been otherwise. This way of conceiving and representing stochasticity corresponds to the structure of Millstein's (2011) Unified Chance Concept (UCC), which she suggests as a general concept of chance in biology.[6] This is precisely how we define "stochasticity" in this paper.

Let us clarify what we mean by the expression "same sub-set of initial conditions." Again, in line with Millstein (2011), this shall not be understood, ontologically speaking, as the same complete set of conditions at a given point in time. By adopting an epistemological perspective, and thus focusing on stochasticity in models of gene expression (namely translation), we rather refer to a partial set of initial conditions. Indeed, the initial conditions that a modeler includes in her model are necessarily limited, and correspond to the subset of conditions she considers as relevant for describing and accounting for the phenomenon under study. For instance, in the case of the translation process, the subset of initial conditions in its models is defined in terms of the amount of macromolecules involved in this process, namely the amount of mRNA and ribosomes in a single cell or in a population of cells in a given *milieu*/solution (i.e., magnesium and other ions, pH, ATP), which is also assumed to be constant over a relatively short experimental time.

Note that the initial conditions in a model are always partial and depend on the modeler's choices. However, this does not imply that the stochastic character of the system under study (i.e., the fact that it is described by a stochastic scenario) is due only to epistemic reasons, and thus has no ontological value. Rather, in line with Millstein's idea (2003), which she borrows from Sober (1984) and Weber (2001), we maintain that modelers systematically choose to adopt a certain level of description and explanation to account for a phenomenon according to their epistemic interests and goals. Indeed, they more or less implicitly choose to take into account certain conditions/causes (those operating at the level chosen because they are considered to be most relevant to a certain epistemic goal) and to ignore others (those operating at other levels). Thus, our best stochastic models/scenarios tell us something real about the systems under study at the level of analysis chosen by the modelers as sufficient and relevant to their epistemic interests and goals. In short, stochasticity is a property of the models and of the target systems, too. This is why, in what follows, we do not systematically specify whether we refer to the models or the systems.

Returning to what we mean by stochasticity, we maintain that something stochastic is not simply something that, once it has occurred, "could have been otherwise." This corresponds to the way contingency is minimally and counterfactually defined in the context of the debate about the inevitability vs. the contingency of evolutionary outcomes (Beatty 2006, 2017; Desjardins 2011; Wong 2020, 2022). In fact, deterministic phenomena can be contingent too, namely in the sense of contingency as causal dependence.[7] Indeed, the outcomes of the deterministic scenario, represented in figure 1 (top), are contingent in the sense that they causally depend upon their respective initial conditions (e.g., outcome X is contingent on initial condition A; outcome Y is contingent on initial condition B, etc.). What differentiates a stochastic outcome

6. "UCC: Given a specified subset of causes, more than one future state is possible" (Millstein 2011, 428).

7. Beatty (2006), in his analysis of Gould's famous thought experiment (1989), which consists of replaying life's tape, identifies two versions of contingency – contingency as causal dependence and as unpredictability.

from a contingent deterministic one is that the former could have been otherwise starting from the same subset of initial conditions. This is clear from looking at the bottom of figure 1: from initial condition A, various different outcomes (X, Y, Z) are possible. In short, they are all contingent on the same initial condition (A); in other words, they are all causally dependent on it.

Now that we have established what it means for something to be stochastic, we shall ask what this "something" can be. Let us look again at figure 1. In the stochastic scenario, three outcomes (X, Y, Z) are possible starting from the same point (A). But to say that an outcome is stochastic just describes such a situation. If we want to account for why this is the case, we should focus our investigation at the level of the processes involved in the production of possible outcomes, and look at where and how they are stochastic. In other words, we should shift our investigation from stochastic outcomes (which we would rather call "random") to stochastic processes (possible processes starting from the same subset of initial conditions). Such processes, represented at the bottom of figure 1 by different arrows that all start from the same point, (can) lead to different possible outcomes.

This is the approach we adopt in this paper. By dissecting models of biological processes at the molecular level (gene expression, translation and, more precisely, the case of Cap-dependent initiation of translation, see sections 2.1, 2.2 and notes 18, 20) we look for what exactly can be treated as stochastic (i.e., what can be different when starting from the same subset of initial conditions, which thus leads to different possible outcomes). It is only by looking for where and how stochasticity can operate in (models of) molecular processes that the stochastic behavior of a given biological system can be accounted for, and stochasticity can thus gain explanatory power in molecular cellular biology. In other words, the kind of analysis we propose and elaborate on here allows stochasticity to be shown to make a difference, biologically speaking, in cellular systems or parts of them, and thus to be a proper and active part of the explanation of their behavior.

1.2. Zooming In: Translation and Alternative Start Codon Selection

As already mentioned, we choose to focus our analysis on the gene expression step of translation, where proteins are synthesized from RNA molecules by the ribosome; specifically, the event[8] of alternative vs. standard mRNA codon selection[9] in the initiation of protein synthesis. Why so? This question can be divided into two parts. Firstly, why did we shift the focus, with respect to the bulk of the literature, from transcription to translation; and secondly, why, among the many translation events that can be analyzed,[10] do we want to dwell on start codon selection? Let's begin by answering the first part of the question.

There are indeed many studies on stochasticity (or that deal, at least transversely, with it) in transcription (Eldar and Elowitz 2010; Eling et al. 2019; Kaern et al. 2005; Tsimring 2014). By

8. We stick to the term "event" (used in the biological literature) when we refer to the start codon selection phenomena. Nonetheless, we intend "event" to refer to a process, namely a series of activities taking place over time and producing certain results. This specification is necessary insofar as this work aims to shift the focus of the analysis of stochasticity from *results* to *processes*.

9. In this context, the term "alternative" will always refer to the activation of an "alternative" ORF, i.e., different from a "standard" one (cf. sec. 2.1).

10. There are a large number of events and processes that take place during translation, so-called "non-canonical translational events" (see, e.g., Firth and Brierley 2012; Brunet et al. 2018; Lyon et al. 2019), such as, among others, leaky scanning, non-AUG initiation, frameshifting, stop codon readthrough, and internal ribosome entry site initiation (IRES) (Blake and Wu 2019, 3; see also Sonneveld et al. 2020, 609). All these processes arise from the dynamic interaction between ribosomes and mRNA and are potentially stochastic in nature.

contrast, very few studies are conducted on stochasticity with respect to translation.[11] Why? Intuitively, since transcription and translation processes share similar physico-chemical features,[12] there is no *a priori* reason that stochasticity could not influence both. Relatedly, there is no reason stochasticity cannot produce differences in the outcomes of these two molecular processes.

In fact, this asymmetry can be explained by several factors. Let us look at two of them which, as far as we are concerned, seem to be the most important. First, until 2000, comparatively fewer techniques existed (e.g., Elowitz et al. 2002; Kaern et al. 2005) to measure stochasticity in translation compared to in transcription (recent years have in fact seen the development of important new techniques; see below). Techniques based on PCR are well suited for quantifying mRNA, but more sophisticated techniques than electrophoresis, like mass spectrometry (Brunet et al. 2018), are needed for the correct quantification of synthesized proteins (Marx et al. 2019). Sequencing and imaging of mRNA are much more effective than protein sequencing, because mRNA is present in the cell in much smaller numbers than their corresponding synthesized proteins (Milo and Phillips 2015). A typical mRNA molecule has a copy number, in unicellular organisms like yeast, in the region of 1 to 30 [13] (Zenklusen et al. 2008), whereas the number of proteins is much higher. In addition, since the number of mRNAs tends to be small, mRNA is more prone to stochastic effects, another reason that justified the focus on transcription.[14]

Second, translation is classically conceived and modeled as a deterministic or nearly deterministic process that faithfully produces polypeptide sequences that correspond to a given transcript. Putois et al. write that translation is usually seen as "a linear conversion with one predictable, unambiguous outcome" (2015, 1). Such a view makes it difficult to even raise the question of whether translation from a given transcript allows for heterogeneous outcomes in terms of the type of product produced.[15] However, these two explanations have lost their bite, as the determinist paradigm struggles to hold up even in scientific practice (Miller and Costello 2001; Strohman 2003), and recent advances in investigative tools, such as live cell single molecule fluorescence imaging of mRNA (Iwasaki and Ingolia 2016, Schmidt et al. 2020, Livingston et al. 2023) and ribosome profiling (Ingolia 2014), make it possible to study cell processes at a very fine-grained scale.[16] Building on the latest work on stochasticity in translation, this paper seeks to propose a philosophical reflection on this new theoretical and empirical exploration.[17]

11. To be fair, there are some studies that deal with the effect of transcriptional noise in translation. They analyze the relationship between the amount of transcripts and polypeptides synthesized in light of noise in transcription (e.g., Pilpel 2011). However, even these studies assume that the stochasticity is found mainly at the level of transcription, and that translation simply mirrors such stochasticity.

12. They are both intracellular processes in which macromolecular protein-based complexes of about the same size interact in their cellular environment with nucleic acid polymers (RNA and/or DNA) in an extended form.

13. Another study in multicellular animals pointed to the existence of two classes of genes: lowly expressed (a few mRNA copies, putatively non-functional but possibly having a regulatory role) and highly expressed (50–100 mRNAs) (Hebenstreit et al. 2011).

14. Thanks to an anonymous reviewer for suggesting this additional point.

15. It should be pointed out though that some works try to highlight stochasticity in translation. For example, Rao et al. go so far as to write that some evidence suggests that "most noise [i.e., stochasticity] arises during translation" (2002, 234), and Mao and Qian, though not mentioning stochasticity, write of "making sense of mRNA translational noise" (2024) which would come from noncanonical translation events.

16. We must note that the present paper focuses on studies of single cells, especially cells in culture (and thus belonging to a multicellular organism, usually human) and unicellular organisms, such as yeast. We will talk about stochasticity inside the cell, and ignore perturbations that may come from outside. This for the reason that, if the environment has any effect on stochasticity inside the cell, it might be of a general character and thus not affecting specific genes, though it might affect the probability of non-canonical translation events (Mao and Qian 2024), which do not occur on all genes.

17. Note that we are not ruling out the possibility that a more in-depth study of the various ways in which

Let us now return to our original purpose and try to answer the second question: why, in the context of translation, did we choose to focus on start codon selection? Even though stochasticity in translation has rarely been discussed from a philosophical point of view, it is currently emerging as an area of great interest to biologists.[18] In recent years, the number of studies analyzing the behavior of individual molecules over time have increased, and a great deal of work has focused on the study of RNA (Blake et al. 2024; Braselmann et al. 2020; Das et al. 2021; Hoek et al. 2019; Livingston et al. 2023; Lyon et al. 2019; Mateju et al. 2020). In this context, we noticed Boersma and colleagues' 2019 paper, in which the authors highlight stochasticity in eukaryotic translation in a way that is quite unprecedented in the literature. By developing an original method of empirical investigation, they show that stochasticity can impact the *type* of polypeptide produced – that is, the different sequences of amino acids synthesized. Let us look more closely at their work. By combining two systems of visualization for translation heterogeneity (the MoonTag and the SunTag systems), Boersma et al. (2019) trace the behavior of individual mRNA molecules (which they have engineered in such a way they possess two alternative start codons) over time, which they then analyze statistically. On this basis, they conclude that the synthesis of alternative polypeptides from the same individual mRNA molecule reflects an "inherent stochasticity in start site selection by individual ribosomes" (Boersma et al. 2019, 471). More precisely, they show that such stochastic events in translation, namely standard vs. alternative start codon selection events by ribosomes, influence the diversity (type) of polypeptide sequences translated from the same individual mRNA molecule (cf. sec. 2.2). Their result opens the door for interesting empirical and philosophical questions. Where and how does stochasticity intervene in translation? And can it be explanatory, (i.e., does it allow us to better characterize and explain molecular dynamics in translation)?

Stochasticity can characterize various parameters of a molecular process, opening up the possibility that they have different values. The fact that start codon selection events can be analyzed by focusing on individual cells (and especially on individual mRNA molecules) makes it a particularly suitable case for studying stochasticity with respect to the parameter of time. Indeed, start codon selection events naturally lead to being modeled as sequences of possible alternative events (each event leading to the production of a certain type of protein), which take place over time, on an individual single molecule of mRNA, in a single cell. This contrasts with a statistical (atemporal) analysis at the level of mRNA or cellular populations. Moreover, in Boersma and colleagues' 2019 paper, the temporal analysis of start codon selection events as a sequence has played an integral role in demonstrating its stochastic character. This is the reason why we focus our analysis on time as the stochastic element that can make a difference in the process of start codon selection.

Before deploying our analysis, we want to make one final point regarding the goal of the present work. The focus of our analysis on a single parameter (time), in order to clarify how stochasticity may have an impact on it, could possibly be seen as a reference for guiding future studies on stochasticity. What we want to emphasize with such a method is the importance of proceeding by choosing specific parameters, the analysis of which makes it possible to highlight different ways in which stochasticity can intervene in gene expression, and the different impacts

stochasticity might impact transcription processes could be of great interest too, especially in wondering, retroactively, whether stochasticity here can also play a role in the *type* of mRNA produced and not only their *amount* (i.e., number).

18. In particular, translation initiation, in which start codon selection plays a major role, is gaining center stage in studies of gene expression regulation, especially in studies of cancer (Xu and Ruggero 2020) and neurological disorders (Gleason et al. 2022). This is not surprising, given recognition of the fact that in eukaryotes (cells and the viruses infecting them), cap-dependent translation initiation (Kozak 1978; Querido et al. 2020) is the most predominant translation pathway and the main target of translational control mechanisms (Wang et al. 2020).

it can have on this process.

1.3. Plan

The paper is structured as follows. Inspired by the pioneering work of Boersma et al. (2019), in the second section we introduce our case study (start codon selection events; see also Brunet et al. 2018; Blake and Wu 2019). These events represent a case of non-canonical translation (see, e.g., Firth and Brierley 2012 and note 10 for more details) that have been shown to be stochastic in nature. We use this case study because qualitative effects are involved: stochasticity is at the origin of switches in translation initiation sites belonging to standard vs. alternative open reading frames (ORFs), which give rise to the synthesis of different polypeptides. In the third section, we use two simple models of time in order to show two ways in which stochasticity can affect this parameter of start codon selection and, more generally, in translation. In the fourth section, we show how this can make a biological difference. More precisely, we draw out the consequences of these models in terms of the temporal dynamics of protein production to show why the various ways in which stochasticity can operate in translation are possibly physiologically relevant, and thus explanatory.

2. Putting Our Molecular Case-study to Work

2.1. Open Reading Frames (ORFs)

In order to properly understand the dynamics of this non-canonical case of translation, start codon selection, we first need to define ORFs. ORFs are stretches of nucleotide sequences that are not interrupted by stop codons in a given reading frame. The length of each reading frame can be divided into triplets of nucleotides, beginning with a translation start codon (usually, but not exclusively, ATG), the codon from which a ribosome begins translation, and ending at a stop codon (the codon at which a ribosome ends translation; Sieber et al. 2018). These sequences are read by ribosomes and they usually produce a single, well-defined polypeptide, which typically undergoes one or more post-translational modifications that may alter its physicochemical and biological properties. Such ORFs are termed "standard" or "canonical," because the start/stop codons they are delimited by are deemed to be unique and, as such, are usually annotated in nucleic acid/protein databases as representatives of a certain gene or protein according to bioinformatic analysis (based on the above principles) and, especially, experimental data. This has led to a broad consensus on the existence of standard ORFs, grounded in both bioinformatic analyses and experimental evidence. However, growing evidence points to the existence of "alternative" ORFs, though the situation is more delicate than with standard ORFs as there is no common agreement as to their existence, relevance or function (Kochetov et al. 2005; Brunet et al. 2018, Xu and Zhang 2020; Wright et al. 2022). Very roughly, we can say that alternative ORFs are delimited by start/stop sites that are either not included in databases (for example, because they are considered to be due to sequence inaccuracies), or that are deemed not to be relevant. This especially happens in the case when, having their start codon (see more below) in a different frame than the standard ORF, they consequently lie out-of-frame with respect to the standard ORF, completely differing, then, in their amino-acid sequence.[19] The annotation of

19. To be clear, as we said before, the length of a frame is a multiple of three, as a triplet of nucleotides is the unit read by the ribosome. This means that if the start codon of the alternative ORF is at a distance which is not a multiple of three from the start codon of the standard frame, it will define a frame which is different from the standard one. Since each triplet corresponds to one amino acid, when the ribosome reads a series of triplets of nucleotides "out-of-frame" with respect to the series of triplets of the standard ORF, it means that the ribosome

alternative ORFs in mammals, including humans, are recent and/or not universally recognized (Brunet et al. 2018; Pavesi et al. 2018), while their existence in viruses has been demonstrated since the '70s (Barrell et al. 1976) and subsequently studied in detail (Firth 2014; Pavesi et al. 2018). Progress in bioinformatics and proteomics is rapidly advancing this field (Leblanc et al. 2024).

Biologists speak of an alternative start codon selection event when a ribosome initiates the translation process from a start codon which leads to the translation of an ORF that differs from the standard or "canonical" one. This ORF is called an "alternative" or "non-canonical" ORF. The selection of an alternative ORF's start codon initiates its translation and can be attributed to stochasticity (Alekhina and Vassilenko 2012; Brunet et al. 2018; Boersma et al. 2019). More specifically, it results from the process in which the ribosome "scans" the mRNA (Alekhina and Vassilienko 2012; Duncan and Mata 2011; Sokabe and Fraser 2019), in which alternative start codons compete stochastically with the standard start codon to initiate translation.

The majority of the scientific literature dealing with stochasticity in start codon selection analyzes its origin and consequences at the evolutionary level (i.e., focusing on ultimate causation; see Mayr 1961). In that context, stochasticity refers to the random nature of the genetic mutation(s) that lead to the formation of standard and alternative ORFs (see, e.g., Brunet et al. 2018, 613; Firth 2014). This is not our focus. Instead, we study stochasticity in start codon selection from a physiological perspective (i.e., focusing on proximate causation), by looking at what certain biologists identify as stochastic in this molecular process (e.g., Boersma et al. 2019). Stochasticity in start codon selection seems to be invoked when considering the movement of ribosomes[20] from one possible start codon (initiation site) to another. This is what Boersma and his colleagues tell us. Let's take a closer look at their work.

2.2. Our Case-study

The study by Boersma et al. (2019) seeks to investigate heterogeneity in translation, namely the possibility that from the same individual mRNA molecule, different types of polypeptides are synthesized from standard and alternative ORFs. Boersma and colleagues make this possible by designing a sequence such that the two start codons are in a different reading frame and, ending necessarily at two different stop codons, lead to the synthesis of two completely different polypeptides. In order to understand how this heterogeneity is possible, the authors use fluorescent molecules (a complex of antibody and peptide epitopes) that make polypeptides visible in real time during their synthesis. They apply this fluorescent labeling system to polypeptides synthesized by ribosomes starting from standard and alternative ORFs.

Simplifying a bit, the first fluorescent molecule they use is "SunTag," which allows for the initiation of polypeptide synthesis from the standard ORF to be made visible in real time, within a cell and at the level of an individual mRNA molecule. Scientists then add "MoonTag," which is an "orthogonal" tag capable of binding the polypeptide resulting from ribosome attachment

will translate these different triplets into different amino acids. However, as we will see later, if the difference between the two alternative start codons is a multiple of three, then the ribosome will always read the same series of triplets, in this case simply starting somewhere downstream in the mRNA. A shorter protein will be produced, having thus a different N-terminal (that is roughly, the first amino acid translated by the ribosome, the "head" of the protein; cf. sec. 2.3) which will usually give it different biological properties.

20. In cap-dependent translation initiation (see note 18) the ribosome unidirectionally scans the mRNA, stopping or not stopping (that is, just pausing in the latter case) at the start of a given ORF. In the case of IRES (internal ribosome entry site) initiation, it can also bind to the start codon of an ORF with a movement perpendicular to the mRNA (versus scanning, which occurs in a parallel direction). However, for the present argument, it suffices to refer to a generic "movement" (keeping in mind, however, that the ribosome does not go back and forth along the mRNA, it only goes in one direction; cf. fig. 2).

to the start codon of an alternative ORF. The combined use of these two systems enables, for the first time, observation of the heterogeneity of real-time translation between different ORFs in an individual mRNA molecule. Here things get interesting, because the authors observe a behavior that at a first glance seems completely disordered. It does not follow any pattern, so how can such a scenario be modeled? Boersma and colleagues suggest two ways. The first is to establish, *a priori*, the probability that protein synthesis starts from standard or alternative ORFs. Such a probability would depend on the affinity that sequences around start sites for standard and alternative ORFs have for the ribosome. In particular, RNA sequences called "Kozak"[21] can play a role in facilitating, or not facilitating, the binding of the ribosome to a certain start codon in mRNA. According to the authors, this first way of evaluating probabilities is too simplistic, because it makes mRNA sequences determine the behavior of ribosomes. Rather, they propose an approach that establishes probabilities *a posteriori*, that is, by observing the phenomenon in question. These probabilities correspond to the observed frequencies of the SunTag and MoonTag fluorescent marks during translation. The authors argue that such probabilities allow for predictions in a statistical framework, i.e., when a large number of these alternative and standard translation events are taken into account. But with respect to the single event, things get more complicated: "the selection of a translation start site by single ribosomes appeared to be *stochastic*" (Boersma et al. 2019, 471, emphasis added). That is, the individual event is not predictable and, from a given subset of initial conditions, translation can start from standard or alternative ORFs, thus leading to the synthesis of different polypeptides. On the same page, the authors stress again that the synthesis of alternative polypeptides from the same individual mRNA molecule reflects an "inherent stochasticity" in the process (Boersma et al. 2019, 471).

Boersma and colleagues' work is particularly interesting from a philosophical point of view because of the way in which stochasticity is invoked in their article. It is not used to underline a lack of understanding of the phenomenon. Rather, stochasticity is considered as a fundamental element of the best possible way to describe and account for the molecular process under investigation. We thus consider this work as particularly appropriate for addressing the philosophical questions at the core of our paper: What exactly should be conceived of as stochastic in gene expression, in particular in translation (i.e., what can have more than one possible future state from the same subset of initial conditions and thus, once it has occurred, could have been otherwise)? How is this the case, and can this be relevant from biological and explanatory points of view?

2.3. A Qualitative Model of Alternative Start Codon Selection Events

We propose figure 2 as a qualitative representation of the movement by a ribosome from a standard to an alternative ORF downstream, that is, the process that Boersma et al. describe as stochastic. This representation shows two subunits of the ribosome (fig. 2, in green) that can bind the ORF1 start codon (fig. 2, at top) or the ORF2 start codon (fig. 2, at bottom). First, the small subunit attaches to the mRNA at its leftmost end and it scans it until it encounters the ORF1 start codon, to which it can bind or not, stochastically. If it does not bind to it, it continues to scan the mRNA until it finds the ORF2 start codon, to which then it can bind. In any case, if the small subunit binds to a start codon, 1 or 2, then the large subunit can in turn bind to this complex and translation can begin. In this scenario, stochasticity does not refer to the evolutionary origin of these two ORFs (cf. Brunet et al. 2018, 613), whether standard or alternative, but to an aspect of the intracellular molecular process of initiating translation in

21. From the last name of the discoverer, the molecular biologist Marilyn Kozak who also proposed the scanning model (introduced in note 20; Kozak 1978).

which they are possible targets. Let us say that start1 is the start codon of the standard ORF1 (fig. 2, blue) and that start2 is the start codon of the alternative ORF2 (fig. 2, red), which we suppose to be in another reading frame, that is, another triplet reading. Depending on which start codon the ribosomal initiation complex (formed by the small ribosomal subunit as well as several "Initiation Factors" [eIFs] not depicted in fig. 2; Querido et al. 2020) binds to, two different proteins are produced (protein 1 or 2, not depicted in fig. 2).

The fact that the two ORFs in figure 2 (and in Boersma et al. 2019 as well, as we have seen) have different stop codons (stop1 in blue and stop2 in red) stems from the fact that in this case the two start codons lie in different reading frames. Of course, alternative start codon(s) may well be in the same reading frame as the standard one, and would thus have different starts in that frame but necessarily the same stop codon (e.g., stop1 in blue). In this case, the two alternative proteins differ only at the N-terminal side (that is to say in their "head"); this is a very common case, indeed, for alternative start codons (Benitez-Cantos et al. 2020; Xu and Zhang 2020). But even in this case, our argument would not be affected. Why? Our initial hypothesis was that stochasticity might affect the type of polypeptide synthesized, as well as the amount. Synthesized polypeptides with different starts but the same stop still count as a different "type" of polypeptide since they have different amino acid lengths and more or fewer N-terminal domains, which could correspond to a different cellular localization (in different compartments) and/or function.

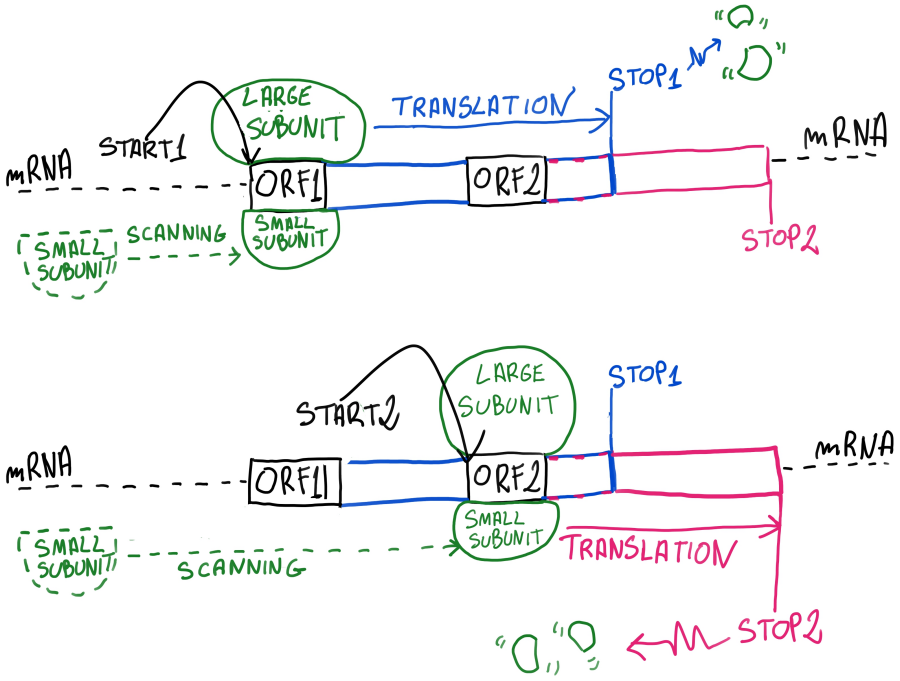

Figure 2: A qualitative model of ribosome movement between a standard ORF (ORF1) an alternative ORF (ORF2). The ribosome's "decision" to initiate translation from ORF1 or ORF2 is considered to be "stochastic" by some biologists (see Boersma et al. 2019). Figure created by Marco Casali.

One might wonder, in the context of this paper, if the final step of the translation process is affected or not by stochasticity. There is no evidence for this, though Boersma et al. rightly remind us that "ribosomes can bypass stop codons under certain conditions to generate C-terminally extended proteins" (2019, 458). That is, the bypass of a stop codon can occur, affecting the C-terminal of the protein (that is to say in their "tail"), formally (and not in terms of mechanism as far as it is known) as a symmetrical counterpart to the effect of an alternative

in-frame start codon which can generate a N-terminally shortened protein. This bypass process, technically known as stop-codon readthrough (Palma and Lejeune 2020), is considered in some cases an infrequent and usually deleterious error (Zhao et al. 2021) possibly having, however, an evolutionary role (Kosinski and Masel 2020); in others, a programmed efficient mechanism to generate variant proteins (Palma and Lejeune 2020; Zhao et al. 2021). In the future, it might then be interesting to investigate the possible stochastic nature of stop codon readthrough.

To return to the stochastic character of alternative start codon selection, selection by the ribosome among the initiation codons of the two frames (the possible initiation codons can sometimes even exceed two) happens stochastically. This means that, starting from the same intracellular context just before translation activation (i.e., the same subset of initial conditions in terms of mRNA, ribosomes and, more generally, intracellular micro-environment), ribosomes can initiate translation from ORF1 or from ORF2, thus leading to the synthesis of different genetic products/proteins (protein 1 or protein 2). This characterization merely considers a single event of start codon selection. However, since translation takes place over time, in order to more deeply understand the way in which stochasticity operates in this molecular process, we have to consider an entire (that is, during the entire lifetime of an individual mRNA molecule) sequence of events. Therefore, in the next section, we introduce the parameter of time into the picture, thus focusing on several successive events of different (standard or alternative) start codon selection. Metaphorically speaking, we do not ground our reflection on a "picture" but rather a "movie"[22] of (alternative) start codon selection events. This enables us to assess the possible biological implications of temporal sequences of stochastic events in a single cell.

3. Stochasticity in Time

The way stochasticity can play a role in start codon selection (fig. 2), and thus in the translation process, can be evaluated with regard to various parameters used to describe and model the system under consideration. Indeed, an investigation into how stochasticity affects a specific parameter (e.g., time, space, energy, etc.), can produce different analyses according to the way in which the parameter in question is represented and integrated into the model. This shows that the system can vary stochastically in this respect, i.e., relative to a given parameter and starting from the same subset of initial conditions of the system. In our case, the system can evolve in more than one possible way according to the state of the intracellular molecular *milieu*.

As stated above, this paper focuses on the parameter of time, in particular, two simple models of it: time as order and time as duration. The purpose of this section is to show how considering these two linear models of time and applying them to sequences of start codon selection (i.e., initiation) events can be relevant for appreciating some possible ways in which stochasticity can influence translation and, consequently, cell physiology, thus making it explanatorily relevant. We talk of "time as the order of events in a sequence" to refer to the order of start codon selection events over time, and of "time as the duration between events in a sequence" to refer to the time intervals between successive start codon selection events.[23] We show that these two temporal features of protein synthesis are stochastic (i.e., they can be different, starting from a given initial

22. Moore (2012) stresses that efforts by structural biochemists to develop "movie[s] of proteins synthesis" are misplaced since "all the functionally significant movements of the ribosome, both internal and external, are biased random walks, and it is most unlikely that any given ribosome will ever do exactly the same thing twice as it elongates some polypeptide" (Moore 2012, 8). Being aware of this critique, our use of the notion of a movie is not intended as a claim of any sort of regularity and/or repeatability, but rather as an attempt to take into account the temporal dimension of stochastic molecular processes.

23. These two models of time (with some modifications) are discussed in Gottlieb (1993) and Nicoglou (2017).

subset of initial conditions).

3.1. Time as the Order of Events in a Sequence

Time modeled as the order of events in a sequence refers to the way events follow one another in a linear ordering (e.g., A B C D E). Applied to our case study, this model of time describes the linear sequence of different ORF activation events (otherwise known as start codon selection events) by ribosomes, namely the temporal order in which translation is initiated at a certain ORF (e.g., ORF1 ORF2 ORF2 ORF1) and the corresponding proteins which are synthesized (e.g., P1 P2 P2 P1). The linear sequence of translations from different ORFs should not be understood in a classical sense of cause and effect, or a causal chain, such as A causing B causing C, etc. As a matter of fact, however, the processes of protein synthesis from each ORF are not causally independent, given the fact that, as we have seen, mRNA is linearly scanned by the ribosome. If the first start codon encountered is used, that is, the ribosome binds to it, this will inhibit the same ribosome from using the second, downstream start codon. Thus, the chance of ORF2 being translated is dependent on whether ORF1 was not used, making the translation of ORF1 a stochastic event, as well as that of ORF2. If we place ourselves within the temporal dimension of milliseconds (the timescale at which a translation step occurs), then the physical encumbrance of the ribosome is relevant: in fact, a second ribosome cannot start, without delay, a new translation of ORF1 or, in particular, an alternative ORF2 downstream because of the presence of the first ribosome.[24] However, if we consider a different timescale (e.g., minutes), this steric hindrance loses its relevance, since the time is long enough for all ribosomes to translate the ORFs with which they interact (for example, in a growing *E. coli* bacterium, where an individual mRNA molecule is used to make 10–100 proteins, the degradation time is typically three minutes; Milo and Phillips 2015).

This temporal dimension of translation reveals an important way in which stochasticity can intervene in this intracellular process: it can be the source of the *alternative* orders in sequences of start codon selection events. This means that, for instance, such order can randomly[25] be one of the following:

[1] ORF1 ORF2 ORF1 ORF1 ORF2 …
or
[2] ORF2 ORF2 ORF1 ORF2 ORF2 …
or
[3] ORF2 ORF1 ORF2 ORF2 ORF1 …
(and so on)

In other words, having sequence 1, 2, 3, or any other, each characterized by its specific order of events, is a matter of chance. With this simple outline of the temporal order in mind, we can thus specify an initial way in which stochasticity can have an impact on the temporal dimension of this non-canonical translation process. To be more specific, temporally speaking, starting from the same subset of initial conditions (in terms of the intracellular molecular *milieu*) at a given point in time, this process can take different possible paths. The order of start codon selection/ORF activation events over time and, even more precisely, the sequence of ORF1 and

24. Note that, on the contrary, activation of ORF2 is not expected to significantly delay activation of ORF1, if ORF2 is not too close to ORF1, due to the unidirectional movement of the ribosome during scanning.

25. In this context, by "random" we refer to processes (in this case, the various possible series of ORF translation start events) that most of the time are unpredictable or do not follow regular patterns, but that, nonetheless, can (potentially) be described by a probabilistic law or distribution (see Boersma et al. 2019) – as is expected for any stochastic process.

ORF2 activation on the same individual mRNA molecule (assuming that a series of events can be conceived of as a process), can be different. Thus, once it has occurred, an ordered sequence *could have been otherwise*. These stochastic processes (can) lead to random outcomes.[26] These outcomes are the order in which proteins are synthesized from a specific sequence of ORFs and, more broadly, their impact on the physiology of the cell (we focus specifically on the proteins produced in sec. 4.1).

3.2. Time as Duration Between Successive Events

Let us consider now how stochasticity in translation can also modify time intervals between successive start codon selection events in a sequence. Time modeled as duration between successive events considers the time that elapses between successive events. More explicitly, this second way in which time can be represented and integrated into models considers both the linear order of events and the relative relations between events. Let's focus on the following sequence:

(S) A——B——————C———D

In sequence (S), the letters A, B, C and D are not equally spaced from one other, which indicates the duration of time between events. Thus, this model describes the linear order of the events, as well as their relative temporal relation.

When applied to our case study, time as duration describes the temporal intervals between ORF activation events, leading to the translation of the same or different ORFs (e.g., ORF1—ORF2–ORF1——ORF1–ORF2). With the variable distance representing different durations between consecutive start codon selection events, each sequence represents then possible alternative sequences and time intervals for this sort of event. This model describes an important second way in which stochasticity can intervene in this intracellular process: it can be the source of the different durations of time between different start codon selection events (and thus ORF activation events). This means that, for instance, we can randomly have the following alternative scenarios:

[1] ORF1—ORF2–ORF1——ORF1–ORF2–…
or
[2] ORF1–ORF2—ORF1–ORF1——ORF2–…
or
[3] ORF1–ORF2—ORF1–ORF1———ORF2–…
(and so on)

As before, it is a matter of chance whether sequence 1 or 2 or 3, or any other, occurs, with their specific durations between ORF selection events. Again, this second linear model allows us to identify a second way in which stochasticity can have an impact on the temporal dimension of this non-canonical translation process. Specifically, this process, temporally speaking, can start from the same subset of initial conditions (in terms of the intracellular molecular *milieu*) and take different possible paths. The duration of time between different start codon selection events can differ, ORF activation events can be different and, once they have occurred, could have been otherwise. These stochastic processes may in turn give rise to random outcomes, namely different timings in the production of different proteins and, more broadly, their consequences on cell physiology.

26. To be clear, the event is the selection of the start codon, the process is the sequence of such events, and the outcome is the synthesized proteins.

Up to this point, we have proposed a conceptual reflection on the possible ways of investigating stochasticity in translation, focusing on the parameter of time. More specifically, by considering two linear models of time[27] and applying them to our case study, we showed that the order of start codon selection events over time and/or the temporal spacings between them can be stochastic.[28] This is a valuable theoretical advance, because it allows us to precisely identify something in translation that, starting from the same subset of initial conditions, can be different (and thus can make a difference, as we will see in sec. 4). Moreover, the result of our analysis up to now is a first step towards reconsidering the role that stochasticity can play at the molecular level, in particular in the process of gene expression (more about the biological relevance of stochasticity in sec. 4 below).

3.3. The Effect of Stochasticity May Be Relevant Only with Small Numbers

This section makes two clarifications related to the fact that, when applying the models of time as order and time as duration to our case study, we implicitly assume that each sequence of start codon selection events is produced from a single mRNA molecule.[29] This is in fact what it is measured by Boersma et al. (2019) in their experimental setting, by isolating the fluorescence emission of individual mRNA molecules. We have to note, however, that in real-life biology, there can be multiple copies of the same type of mRNA molecule in a cell, which can then be translated simultaneously. This makes the scenario more complex, insofar as it increases the possibility of overlaps between sequences of ORF activation events. As a consequence, each sequence of start codon selection events can result from different dynamics. Multiple translations[30] from two different ORFs (e.g., ORF1 and ORF2) in an individual mRNA molecule, or the sum of successive translations events from different ORFs in multiple mRNA molecules of the same type, can both be at the origin of a given sequence of initiation events in translation.

To shed light on these two additional levels of complexity, let us look at sequence [Actual 1] below, which represents an actualized temporal order of, and duration between, ORF activations:

[Actual 1] ORF1—ORF2–ORF1——ORF2

If we do not stick to the experimental setting of Boersma and colleagues, and have in mind real-life biology (i.e., we imagine lumping together all the mRNA present in the cell or in a particular region of it), this sequence [Actual 1] is not necessarily the result of the translation of an individual mRNA molecule by multiple ribosomes, but could also be due to the translation

27. An anonymous referee objected that the ordering model of time is irrelevant if one also considers the duration model. We understand the point but disagree for the following reasons. First, we present the two models separately in order to introduce these two aspects of time (i.e., order and duration) analytically. Second, the duration model might not explicitly mention the specific ORFs that are activated over time, but only the moment of each activation. In this case, it would not include the ordering model.

28. The two models can obviously be integrated. The various sequences may vary in both order and temporal spacing between the different ORFs.

29. Processes that involve a small number of molecules, and that are therefore subject to the small number effect, are acknowledged in the biological literature. In mouse embryos, for instance, cell differentiation is a finely tuned process, but small differences in the concentration of certain transcription factors "can tilt the balance" (Zernicka-Goetz and Huang 2010, 743). Another example comes from yeast, where it has been reported that "the abundance of proteins ranges from fewer than 50 to more than 10^6 molecules per cell. Many of these molecules, including essential proteins and most transcription factors, are present at levels that are not readily detectable by other proteomic techniques nor predictable by mRNA levels or codon bias measurements" (Ghaemmaghami et al. 2003).

30. By "multiple activations" we are referring to the possibility that the ORFs of a single RNA molecule are translated more than once. In this case, ribosomes assemble and run on ORF1 and ORF2 more than once.

of multiple neighboring RNAs by multiple ribosomes. To simplify a little, figure 3 illustrates these two possible ways in which we can arrive at the same sequence [Actual 1].

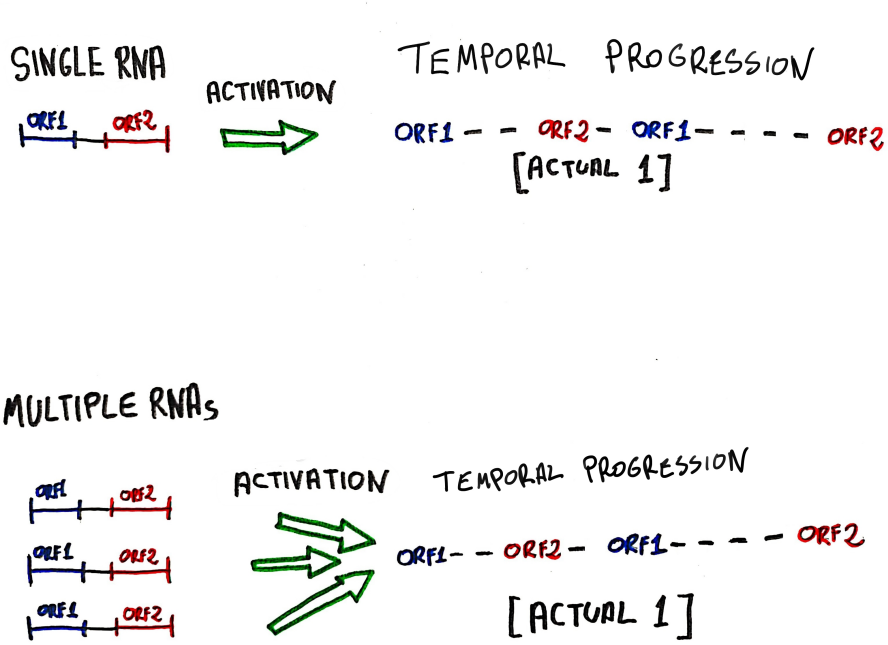

Figure 3: Depiction of two possible scenarios for the temporal order of and duration between activations of individual ORFs (in this case, [Actual 1]). At top: activation of two different ORFs (i.e., ORF1 and ORF2) from an individual mRNA molecule. At bottom: activation of two different ORFs (i.e., ORF1 and ORF2) from multiple (but identical) mRNA molecules. The green arrows represent the ORF activation process producing possible alternative sequences of ORF activation events (ribosomes and the molecular process of translation are not represented). Figure created by Marco Casali.

What matters here is to recognize that the actual sequence of different ORF activation events – which constitutes [Actual 1] and is characterized by a specific linear order and specific time intervals – is the result of successive stochastic events of start codon selection by ribosomes, which can be realized by either of the scenarios depicted in figure 3. The first (at the top) involves just one individual mRNA molecule; the second (at the bottom), more than one (three in this example, but there could be more).

Can the second scenario at the bottom of figure 3, which depicts multiple RNAs and multiple ribosomes producing a sequence of ORF activation events (e.g., [Actual 1]), be framed in terms of our two models of time? Our answer is yes, it can, at least if it fulfills certain constraints. Even if the number of individual mRNA molecules from which our [Actual 1] sequence is generated is more than one (cf. the lower part of fig. 3 in which we suppose the presence of three mRNA molecules), our modeling can be applied so long as the activation of ORFs and the appearance of proteins can be represented in a linear sequence. That is, it is necessary not to have overlap in either activation or protein appearance. If, on the other hand, the activation of ORFs and the appearance of proteins occur too often, becoming too numerous with the increasing possibility of temporal overlaps, then our model is no longer applicable and the importance of stochasticity as we have characterized it could lose its relevance. To better understand this consideration, let us now focus on a scenario in which there are many mRNA molecules (Zenklusen et al. 2008, cf. Section 1.2) in which the activations of different ORFs overlap in time, and where the linearity of the ORF activation sequences, and consequently that of the proteins produced, is lost.

In this case, the actual sequences of ORF activation events and the duration between each event in each sequence is no longer relevant for the resulting translation in the cell. In fact, the large number of ORF activations is likely to produce overlaps, rendering their timing, biologically speaking, potentially irrelevant and their representation in terms of our temporal linear models useless. Consequently, the way stochasticity can play a role in this non-canonical case of translation, namely by having an impact on the temporal sequence and duration of ORF activation events, loses relevance too. The fewer the number of proteins produced, the more relevant their representation through a linear model of ORF activation events over time, and the more important the potential effect of stochasticity (cf. sec. 3; see also Raj and van Oudenaarden 2008). This is what is called the "small number effect".[31] Thus, when a few molecules are involved in a process, their presence or absence and the timing of their actions are all of potentially great importance for cellular physiology. The relevance of our analysis thus rests on the assumption that molecular biological processes taking place inside the cell (in particular, translation initiation) can involve a small number of molecules (individual mRNA molecules in our case). This could be the case if a factor such as the spatial distribution of mRNA isolates single molecules or a small number of them, if there are also limits to the speed of diffusion in the intracellular *milieu* of the synthesized proteins (due to macromolecular crowding, for example).

4. The Impact of Stochasticity in Translation: Biologically Plausible Scenarios

Are these two ways in which stochasticity intervenes in the timing of start codon selection events biologically relevant? To show whether our analysis above is relevant for the study of cell physiology, we need to shift it from the level of ORFs (in mRNA molecules) to the level of the proteins produced and their action. The aim here is to determine whether the order of appearance of proteins, as well as the time intervals between their synthesis, can have an impact on cell physiology, and thus be explanatorily relevant. In the next section we provide an explicit argument for this change of focus.

4.1. Order of Translation Initiation vs. Order of Full Protein Completion

In the models we have so far proposed, we focused on ORF activation events, in which ribosomes bind with ORF start codons, and the synthesis of polypeptides begins. We talked about two ORFs – ORF1 and ORF2. Now, we add a further level of complexity, since the order of ORF translation initiation does not necessarily reflect the order in which proteins appear. By the "appearance of the protein," we mean the moment at which the ribosome finishes synthesizing a polypeptide, making it "ready" to perform its role.[32] Let us consider the translation initiation of two ORFs in the same RNA, which we can call "overlapping":[33] ORFx which

31. The "small number effect" originates from statistics: when fewer elements are involved, the difference between possible outcomes because of variation between those elements grows in magnitude. Hacking (2001) explains this statistical property with a simple example. Males and females are born in roughly equal numbers. If we have two hospitals, one in which many babies are born and one in which few babies are born, Hacking asks (or rather that he poses to the reader) in which hospital we will find unusual weeks in which babies, female or male, exceed 55% of births? The answer is the small one, because large deviations are more probable, and so more common, in small populations (Hacking 2001, 250).

32. We are fully aware that while still being synthesized, as a protein is exiting from the ribosome in the cytoplasmic membranes (Nyathi et al. 2013), it could already interact with other molecules (Duncan and Mata 2011). Though these interactions are very important (Schwarz and Beck 2019), we limit our analysis by referring to the interactions and relevance of polypeptides only once they are fully synthesized.

33. The term overlapping (first mentioned by Barrell et al. 1976) refers to the phenomenon whereby a single sequence of genetic material (DNA or RNA in the case of many viruses) is shared fully or in part by two ORFs,

encodes for a very long protein, and ORFy which encodes for a very short polypeptide. If the time it takes to synthesize protein X is longer than the total time it takes to activate ORFy and synthetize protein Y, then the order of ORF activation does not necessary reflect the order in which the proteins appear.[34] Let us try to make this idea clearer with a figure.

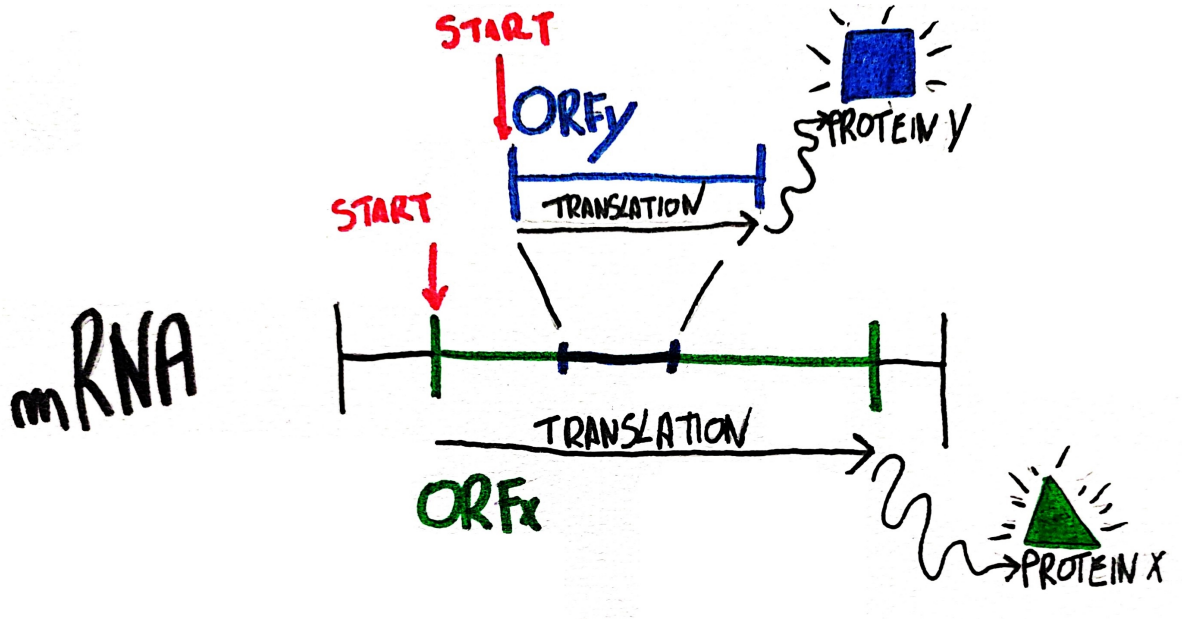

Figure 4: A representation of an mRNA molecule in which there are two overlapping ORFs, in which ORFy (in blue) completely overlaps with ORFx (in green). Although the start of ORFx is before that of ORFy (the two start arrows in red), the protein that will take less time to be synthesized is protein Y. Figure created by Marco Casali.

In figure 4, we represent an mRNA molecule in which there are two overlapping ORFs, ORFx and ORFy. In the picture, ORFx fully contains ORFy (which is "nested") and it is clear that, although the translation initiation of ORFx lies before the translation initiation of ORFy on the mRNA, protein Y, being much shorter and having its stop codon far before that of ORFy, will appear before protein X even when the latter is activated first. This example serves to make explicit that the order of appearance of proteins does not necessarily reflect the order of activation of ORF translation.

Could this possible variable symmetry between the order of translation initiation and the order of protein appearance pose a problem for the analysis of stochasticity that has, up to now, been proposed? The answer is no. On the contrary, this could be seen as a further insight into the complexity of stochasticity in translation. As discussed previously (see sec. 2.3), stochasticity can affect the timing of start and stop codon selection, for example by translating different reading frames. However, it can also act indirectly, influencing the order in which proteins appear and the intervals between their appearances.

As noted above, such timing does not necessarily reflect the order of translation initiation events nor the time intervals between them. In the case depicted in figure 4, since we know the initiation time of the synthesis of both proteins, we could reasonably assume that the order of protein appearance can be predicted from the order of ORF translation initiation. But in biological reality, these times are not always fixed and predictable. This analysis leads us to highlight another way in which stochasticity can intervene in start codon selection. It can

allowing for the synthesis of different proteins (see also Wright et al. 2022).

34. In this scenario we assume a constant rate of protein synthesis.

indirectly affect the sequence of proteins' appearance, in particular its timing (in terms of order and of duration).

4.2. Two Biologically Plausible Scenarios

Let us return to the question of the biological, and thus explanatory, relevance of our analysis. We suggest considering imagined but realistic (that is to say, plausible) biological scenarios of the temporal dynamics of protein production in order to show more precisely how stochasticity in the timing (order and duration) of translation can have a decisive impact on the dynamics of translation, and thus on its outcomes. Again, these scenarios are based on the hypothesis of small numbers of molecules involved (cf. sec. 3.3).

In the first plausible scenario, consider two different possible orders of appearance of two alternative antagonist proteins, protein 1 and protein 2, one of which inhibits (or significantly alters) the activity of the other.[35] These two proteins will thus reasonably have a different impact on cellular physiology depending on the order in which they appear (fig. 5).

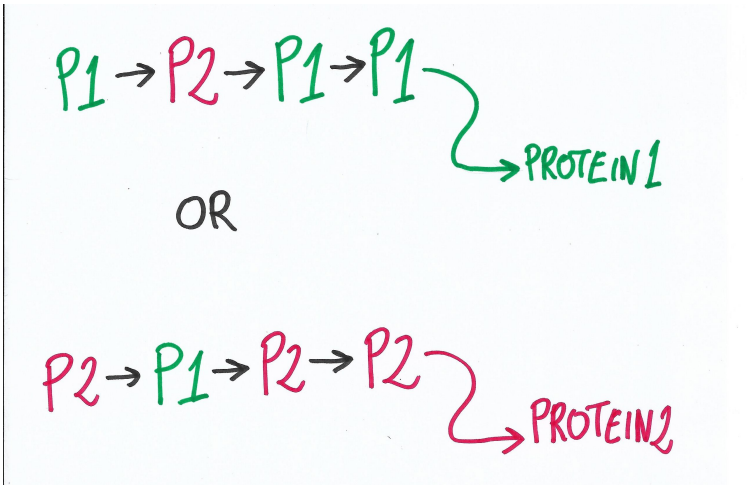

Figure 5: Representation of a possible biological scenario in which a change in the order of the appearance of antagonist proteins could affect cell physiology. "P1" represents protein 1 and "P2" represents protein 2. The protein name is written in full on the right, "protein 1" and "protein 2", indicating that after a certain sequence of protein synthesis, P1 → P2 → P1 → P1 or P2 → P1 → P2 → P2, one of the two overcomes the other, which is to say, makes a substantial difference in the cell's physiology. Figure created by Marco Casali.

In this case, the order of protein appearance could be crucial in determining cell physiology. In the upper part of figure 5, we can see that the first possible order allows mainly protein 1 to perform its role. In the lower part of figure 5, by contrast, we can see that another possible order allows mainly protein 2 to perform its role, overcoming the possible action of protein 1 (P2 is synthesized repeatedly after P1). Stochasticity in this case affects the parameter of time as order: there is more than one possible order of appearance of the two proteins, and the one that actually, and stochastically, takes place will determine which of the two proteins will prevail in performing its role for the reasons discussed above. Stochasticity can thus be biologically and

35. We can find several examples of couples of antagonist proteins derived from alternative start codon events in the same mRNA in biological literature (Freson et al. 2003; Govind et al. 2014; Mazur et al. 2008; Klemke et al. 2001). Though we have been unable to find a work explicitly linking stochasticity in translation and antagonistic proteins, we reinterpret the biological findings in terms of our novel frame of stochasticity in general and time scenarios in particular.

explanatory relevant to the extent that it can affect the organism's physiology, at least at the level of the cell.

The second plausible scenario we propose is one in which there are differences in terms of the time interval between the appearance of two antagonist proteins. In figure 6, we can see that the lapse of time between the appearance of protein 1 and the appearance of protein 2 might be crucial in determining the effects that they might have on cell physiology.[36] In the upper part of figure 6, protein 1 appears first. Since the time interval between its appearance and the appearance of protein 2 is long enough to allow protein 1 to act, protein 1 could feasibly impact cell physiology prior to the synthesis and appearance of antagonist protein 2. Seen the other way around, in the lower part of figure 6, the time interval between the appearance of the two proteins is quite short. In this scenario, we might deduce that protein 1 does not have enough time to perform its activity (especially if more than an instantaneous action is required to obtain the physiological outcome) before the synthesis and appearance of the antagonist, protein 2.

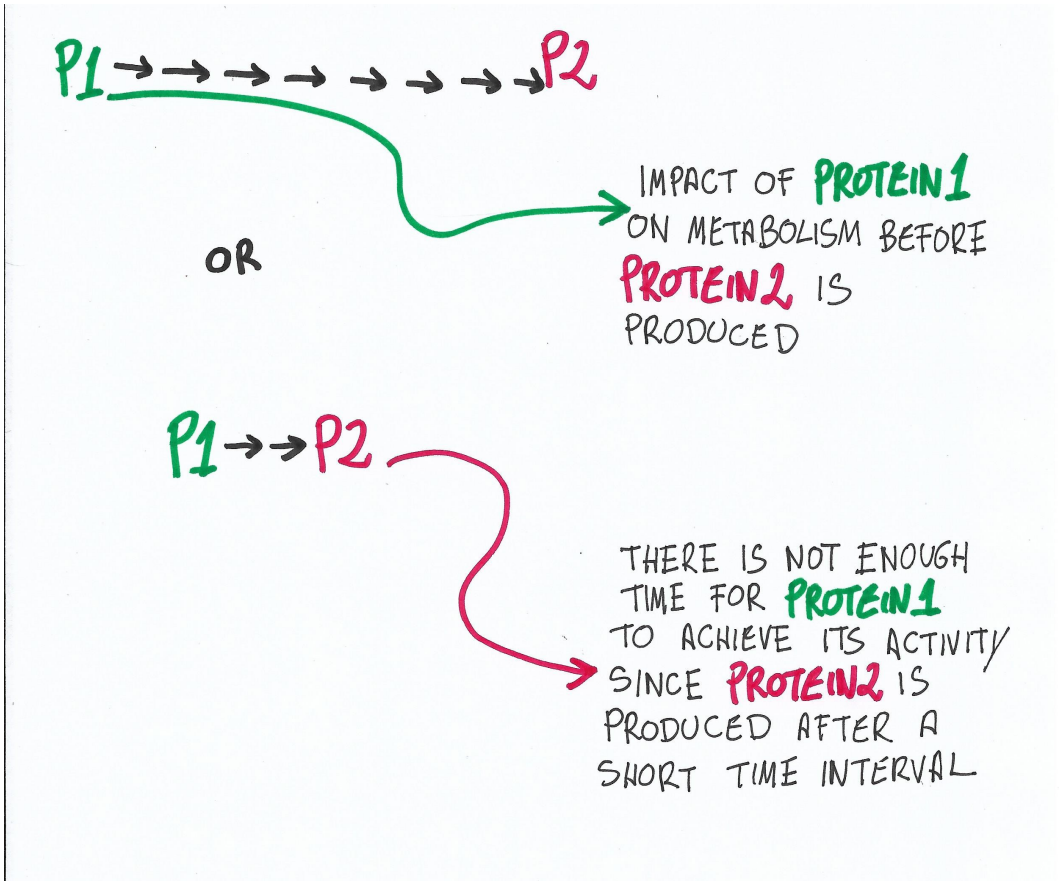

Figure 6: Representation of a plausible biological scenario in which the change in duration between the appearance of proteins could affect cell physiology. "P1" represents protein 1 and "P2" represents protein 2. Figure created by Marco Casali.

Stochasticity in this case affects the parameter of time as duration: the lapse of time between the appearance of the two proteins can vary greatly, and the one that actually, and stochastically,

36. This second scenario was suggested to us by considering the phytochrome (PhyB) system in plants, whose photo-reversibility was first skillfully and clearly shown by studying the promotion of (lettuce) seed germination in plants (Borthwick et al. 1952, 1954). Though it does not involve the synthesis of proteins, its features are an interesting example of the relevance, in time, both of the order of activation (by light, in this case) as in the first scenario, and of the time elapsed before the activation of one protein and that of its antagonist protein.

takes place will determine which of the two antagonistic proteins will act first on cellular physiology. This second plausible scenario, then, shows as well that stochasticity can have important effects on cellular physiology, and can thus be explanatorily relevant. In designing and depicting these scenarios, many factors have not been considered, and a number of approximations have been made, such as protein turnover, post-translational modification and movement of the synthesized protein. However, these approximations were necessary to provide a schematic and understandable picture. Future research into the relationship between stochasticity and these other factors may be useful to provide a more complete picture.

To summarize, the two qualitative examples depicted in figures 5 and 6 suggest that our reflection on the different ways in which stochasticity can intervene in the temporal progression of translation might be relevant to the study of cell physiology. More precisely, stochasticity can have an impact on the system under study and, as a consequence, deserves to have a role in biological explanations. These scenarios are hypothetical and do not stem from empirical evidence, though they are, in the limits of the approximations discussed above, nonetheless biologically plausible. In order to formulate them, we had to shift our attention from the order of ORF translation initiation to the order of protein appearance. Far from being a problem, this instead adds precision to our argument. The relation between the order of ORF translation initiation and the order of protein appearance and action, especially, does not always receive sufficient attention from scientists (for example, Boersma et al. 2019 do not mention it). By contrast, we showed that it can be crucial for understanding the many and different ways in which stochasticity can intervene in translation.

5. Conclusion

In this paper, we proposed a new attitude to adopt for studying stochasticity in molecular and cellular biology. We call it "dissecting stochasticity." It consists in taking a closer look at the various parameters used to describe and model a biological system of interest in order to show whether and how stochasticity can intervene in it. We focused on the parameter of time, in particular, on two ways in which it can be integrated into models of molecular processes, and in a specific non-canonical case of translation – the process of start codon selection by ribosomes. We showed in which sense and how the timing (order and duration) of this process can be stochastic, influencing its result and, as a consequence, plausibly having an impact on cell physiology. More broadly, technological advances over the past decade have enabled a deeper understanding of the temporal dimension of molecular processes within the cell (Dal Molin and Di Camillo 2018; Wang and Song 2017). This empirical awareness, however, requires parallel theoretical work. Otherwise, "we will continue with quantification *sans* justification *ad infinitum*" (Symmons and Raj 2016, 702). In fact, being increasingly able to count individual molecules and their movements in space-time does not necessarily guarantee a better understanding of cellular dynamics. This also applies to stochasticity at the molecular and cellular level, which has been mainly studied in quantitative terms (e.g., Klepàrnik and Foret 2013; Stuart and Satija 2019). Conceptual work is needed in parallel to quantitative work in order to clarify where and when stochasticity can have important effects. Our work seeks to provide a step in this direction. There are, of course, other parameters that can be analyzed by adopting this same attitude of "dissecting stochasticity," such as the amount of mRNA available (which does not affect what happens at the level of a single mRNA), the spatial distribution of molecules in intracellular space (Fusco et al. 2003; Savulescu et al. 2020), and the energy availability (e.g. ATP distribution within a single cell) within a single cell (Carthew 2021), etc. The present paper is only a preliminary study, which should encourage us to continue asking questions about the many

ways in which stochasticity may be involved in translation, gene expression and, more generally, molecular biological dynamics. This type of analysis will help to identify where and how stochasticity is biologically relevant, and thus when it can play a role in biological explanations.

Acknowledgments

We would like to thank Sanne Boersma, Roberta Millstein, Charles Pence, Katherine Waters and the anonymous reviewers for very helpful comments to earlier versions of this paper.

Literature cited

Alekhina, Oleg M., and Konstantin S. Vassilenko. 2012. "Translation Initiation in Eukaryotes: Versatility of the Scanning Model." *Biochemistry (Moscow)* 77 (13): 1465–1477.

Barrell, Bart G., Gillian M. Air, and Clyde A. Hutchison III. 1976. "Overlapping Genes in Bacteriophage $\phi\chi$174." *Nature* 264: 34–41.

Beatty, John. 2006. "Replaying Life's Tape." *Journal of Philosophy* 103 (7): 336–362

Beatty, John. 2017. "Narrative Possibility and Narrative Explanation." *Studies in History and Philosophy of Science Part A* 62: 31–41.

Benitez-Cantos, Maria S., Martina M. Yordanova, Patrick B. F. O'Connor, , Alexander V. Zhdanov, Sergey I. Kovalchuk, Dmitry B. Papkovsky, Dmitry E. Andreev, and Pavel V. Baranov. 2020. "Translation Initiation Downstream from Annotated Start Codons in Human mRNAs Coevolves with the Kozak Context." *Genome Research* 30: 974–984.

Blake, Lauren A., and Bin Wu. 2019. "One Message, Many Translations: Heterogeneity Revealed with Multicolor Imaging." *Molecular Cell* 75 (1): 3–4.

Blake, Lauren A., Ana De La Cruz, and Bin Wu. 2024. "Imaging Spatiotemporal Translation Regulation In Vivo." *Seminars in Cell and Developmental Biology* 154: 155–164.

Boersma, Sanne, Deepak Khuperkar, Bram M.P.Verhagen, Stijn Sonneveld, Jonathan B. Grimm, Luke D. Lavis, and Marvin E. Tanenbaum. 2019. "Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding." *Cell* 178 (2): 458–472.

Borthwick, Harry A., Sterling B. Hendricks, Marion W. Parker, Eben H. Toole, Vivian K. Toole. 1952. "A Reversible Photoreaction Controlling Seed Germination." *Proceedings of the National Academy of Sciences* 38: 662–666.

Borthwick, Harry A., Sterling B. Hendricks, Marion W. Parker, Eben H. Toole, Vivian, K. Toole. 1954. "Action of Light on Lettuce-Seed Germination." The Botanical Gazette 115 (3): 205–225.

Braselmann, Esther, Colin Rathbun, Erin M. Richards, and Amy E. Palmer. 2020. "Illuminating RNA Biology: Tools for Imaging RNA in Live Mammalian Cells." *Cell Chemical Biology* 27 (8): 891–903.

Brunet, Marie, Sébastien A. Levesque, Darel J. Hunting, Alan A. Cohen, and Xavier Roucou. 2018. "Recognition of the Polycistronic Nature of Human Genes Is Critical to Understanding the Genotype-Phenotype Relationship." Genome Research 28 (5): 609–624.

Carthew, Richard W. 2021. "Gene Regulation and Cellular Metabolism: An Essential Partnership." *Trends in Genetics* 37 (4): 389–400.

Casali, Marco, and Francesca Merlin. 2020. "Rethinking the Role of Chance in the Explanation of Cell Differentiation." In *Phenotypic Switching: Implications in Biology and Medicine*, edited by Herbert Levine, Mohit Jolly, Priyanka Kulkarni, and Vidyanand Nanjundiah, 23–51. Academic Press,

Elsevier.

Chia, Minghao, Cai Li, Sueli Marques, Vincente Pelechano, Nicholas M. Luscombe, and Folkert J. van Werven. 2021. "High-Resolution Analysis of Cell-State Transitions in Yeast Suggests Widespread Transcriptional Tuning by Alternative Starts." *Genome Biology* 22: 34.

Dal Molin, Alessandra, and Barbara Di Camillo. 2019. "How to Design a Single-Cell RNA-Sequencing Experiment: Pitfalls, Challenges and Perspectives." *Briefings in Bioinformatics* 20 (4): 1384–1394.

Das, Sulagna, Maria Vera, Valentina Gandin, Robert Singer, and Evelina Tutucci. 2021. "Intracellular mRNA Transport and Localized Translation." *Nature Reviews Molecular Cell Biology* 22 (7): 483–504.

Desjardins, Eric. 2011. "Reflections on Path Dependence and Irreversibility: Lessons from Evolutionary Biology." *Philosophy of Science* 78 (5): 724–738.

Duncan, Caia D. S., and Juan Mata. 2011. "Widespread Cotranslational Formation of Protein Complexes." *PLOS Genetics* 7 (12): e1002398.

Eldar, Avigdor and Michael B. Elowitz M. "Functional Roles for Noise in Genetic Circuits." *Nature* 467 (7312): 167–173.

Eling, Nils, Michael D. Morgan, and John C. Marioni. 2019. "Challenges in Measuring and Understanding Biological Noise." *Nature Reviews Genetics* 20 (9): 536–548.

Elowitz, Michael B., Arnold J. Levine, Eric D Siggia, and Peter S. Swain. 2002. "Stochastic Gene Expression in a Single Cell." *Science* 297: 1183–1186.

Firth, Andrew E. 2014. "Mapping Overlapping Functional Elements Embedded within the Protein-Coding Regions of RNA Viruses." *Nucleic Acids Research* 42 (20): 12425–12439.

Firth, Andrew E., and Ian Brierley. 2012. "Non-Canonical Translation in RNA Viruses." *Journal of General Virology* 93 (Pt 7): 1385.

Freson, Kathleen, Jaak Jaeken, Monique Van Helvoirt, Francis de Zegher, Christine Wittevrongel, Chantal Thys, Marc F. Hoylaerts, Jos Vermylen, and Chris Van Geet. 2003. "Functional Polymorphisms in the Paternally Expressed XLalphas and Its Cofactor ALEX Decrease Their Mutual Interaction and Enhance Receptor-Mediated cAMP Formation." *Human Molecular Genetics* 12 (10): 1121–1130.

Fusco, Dahlene, Nathalie Accornero, Brigitte Lavoie, Shailesh M. Shenoy, Jean-Marie Blanchard, Robert H. Singer, and Edouard Bertrand. 2003. "Single mRNA Molecules Demonstrate Probabilistic Movement in Living Mammalian Cells." *Current Biology* 13: 161–167.

Ghaemmaghami, Sina, Won-Ki Hu, Kiowa Bower, Russell W. Howson, Archana Belle, Noah Dephoure, Erin K. O'Shea, and Jonathan S. Weissman. 2003. "Global Analysis of Protein Expression in Yeast." *Nature* 425: 737–741.

Gleason, Alec C., Ghanashyam Gadge, Jin Chen, Yoshifumi Sonobe, and Raymond P. Roos. 2022. "Machine Learning Predicts Translation Initiation Sites in Neurologic Diseases with Nucleotide Repeat Expansions." *PLOS ONE* 17 (6): e0256411.

Govind, Kunduri, Arindam Bakshi, and Handanahal S. Savithri. 2014. "Interaction of Sesbania Mosaic Virus (SeMV) RNA-Dependent RNA Polymerase (RdRp) with the p10 Domain of Polyprotein 2a and Its Implications in SeMV Replication." *FEBS Open Biology* 4: 362–369.

Gould, Stephen Jay. 1989. *Wonderful Life: The Burgess Shale and the Nature of History*. New York: W. W. Norton & Company.

Gottlieb, Gilbert. 1993. "Foreword." In *Developmental Time and Timing*, edited by George Turkewitz and David Devenny, ix–x. Hillsdale: Lawrence Erlbaum Associates.

Hacking, Ian. 2001. *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Hebenstreit, Daniel, Miaoqing Fang, Muxin Gu, Varodom Charoensawan, Alexander van Oudenaarden, and Sarah A. Teichmann. 2011. "RNA Sequencing Reveals Two Major Classes of Gene Expression Levels in Metazoan Cells." *Molecular Systems Biology* 7: 497.

Hoek, Tim A., Deepak Khuperkar, Rik G.H Lindeboom, Stijn Sonneveld, Bram M.P. Verhagen, Sanne Boersma, Michiel Vermeulen, and Marvin E. Tanenbaum, 2019. "Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay." Molecular Cell 75 (2): 324–339.Ingolia, Nicholas T. 2014. "Ribosome Profiling: New Views of Translation, from Single Codons to Genome Scale." *Nature Reviews Genetics* 15 (3): 205–213.

Iwasaki, Shintaro, and Nicholas T. Ingolia. 2016. "Seeing Translation." *Science* 352: 1391–1392.

Kaern, Mads, Timothy C. Elston, William J. Blake, and James J. Collins. 2005. "Stochasticity in Gene Expression: From Theories to Phenotypes." Nature Reviews Genetics 6: 451–464.Klepárník, Karel, and František Foret. 2013. "Recent Advances in the Development of Single Cell Analysis—A Review." *Analytica Chimica Acta* 800: 12–21.

Klemke, Maerin, Ralph H. Kehlenbach, and Wieland B. Huttner. 2001. "Two Overlapping Reading Frames in a Single Exon Encode Interacting Proteins – A Novel Way of Gene Usage." *EMBO Journal* 20 (14): 3849–3860.

de Klerk, Eleonora, and Peter A. C. 't Hoen. 2015. "Alternative mRNA Transcription, Processing and Translation: Insights from RNA Sequencing." *Trends in Genetics* 31 (5): 128–139.

Kochetov, Alex V., Akinori Sarai, Igor B. Rogozin, Vladimir K. Shumny, and Nikolay A. Kolchanov. 2005. "The Role of Alternative Translation Start Sites in the Generation of Human Protein Diversity." *Molecular Genetics and Genomics* 273: 491–496.

Kosinski, Luke J., and Joanna Masel. 2020. "Readthrough Errors Purge Deleterious Cryptic Sequences, Facilitating the Birth of Coding Sequences." *Molecular Biology and Evolution* 37 (6): 1761–1774.

Kozak, Marilyn. 1978. "How Do Eucaryotic Ribosomes Select Initiation Regions in Messenger RNA?" *Cell* 15: 1109–1123.

Leblanc, Sébastien, Feriel Yala, Nicols Provencher, Jean-François Lucier, Maxim Levesque, Xavier Lapointe, Jean-François Jacques, Isabelle Fournier, Michel Salzet, Ouangraoua Aïda, Michelle S. Scott, François-Michel Boisvert, Marie A Brunet, and Xavier Roucou. 2024. "OpenProt 2.0 Builds a Path to the Functional Characterization of Alternative Proteins." *Nucleic Acids Research* 52: D522–D528.

Li, Ke, jinhui Kong, Shuo Zhang, Tong Zhao, and Wenfeng Qian. 2022. "Distance-Dependence Inhibition of Translation Initiation by Downstream Out-of-Frame AUGs Is Consistent with a Brownian Ratchet Process of Ribosome Scanning." *Genome Biology* 23: 254.

Lipniacki, Tomasz, Pawel Paszek, Anna Marciniak-Czochra, Allan R. Brasier, and Marek Kimmel. 2006. "Transcriptional Stochasticity in Gene Expression." *Journal of Theoretical Biology* 238 (2): 348–367.

Livingston, Nathan M., Jiwoong Kwon, Oliver Valera, James A Saba, Niladri K. Sinha, Pranav Reddy, Blake Nelson, Clara Wolfe, Taekjip Ha, Rachel Green, Jian Liu, and Bin Wu. 2023. "Bursting Translation on Single mRNAs in Live Cells." *Molecular Cell* 83: 2276–2289.

Lyon, Kenneth, Luis U. Aguilera, Tatsuya Morisaki, Brian Munsky, and Timothy J. Stasevich. 2019. "Live-Cell Single RNA Imaging Reveals Bursts of Translational Frameshifting." *Molecular Cell* 75 (1): 172–183.

Mao, Yuanhui, and Shu-Bing Qian. 2024. "Making Sense of mRNA Translational 'Noise'." *Seminars in Cell and Developmental Biology* 154: 114–122.

Marx, Vivien. 2019. "A Dream of Single-Cell Proteomics." *Nature Methods* 16: 809–812.

Mateju, Daniel, Bastian Eichenberger, Franka Voigt, Jan Eglinger, Gregory Roth, and Jeffrey A. Chao. 2020. "Single-Molecule Imaging Reveals Translation of mRNAs Localized to Stress Granules." *Cell* 183 (7): 1801–1812.

Mazur, Igor, Darisuren Anhlan, David Mitzner, Ludmilla Wixler, Ulrich Schubert, and Stephan Ludwig. 2008. "The Proapoptotic Influenza A Virus Protein PB1-F2 Regulates Viral Polymerase Activity by Interaction with the PB1 Protein." *Cellular Microbiology* 10 (5): 1140–1152.

Mayr, Ernst. 1961. "Cause and Effect in Biology: Kinds of Causes, Predictability, and Teleology Are Viewed by a Practicing Biologist." *Science* 134 (3489): 1501–1506.

Miller, Eleanor M., and Carrie Y. Costello. 2001. "The Limits of Biological Determinism." *American Sociological Review* 66 (4): 592–598.

Millstein, Roberta L. 2003. "Interpretations of Probability in Evolutionary Theory." *Philosophy of Science* 70 (5): 1317–1328.

Millstein, Roberta L. 2011. "Chances and Causes in Evolutionary Biology: How Many Chances Become One Chance." In *Causality in the Sciences*, edited by Phyllis McKay Illari, Federica Russo, and Jon Williamson, 425. Oxford: Oxford University Press.

Milo, Ron, and Rob Phillips. 2015. *Cell Biology by the Numbers*. New York: Garland Science.

Moore, Peter. 2012. "How Should We Think about the Ribosome?" *Annual Review of Biophysics* 41: 1–19.

Nicoglou, Antonine. 2017. "The Timing of Development." In *Time of Nature and the Nature of Time*, 359–390. Cham: Springer.

Nyathi, Yvonne, Barrie M. Wilkinson, and Martin R. Pool. 2013. "Co-Translational Targeting and Translocation of Proteins to the Endoplasmic Reticulum." *Biochimica et Biophysica Acta* 1833: 2392–2402.

Palma, Martine, and Fabrice Lejeune. 2020. "Deciphering the Molecular Mechanism of Stop Codon Readthrough." *Biological Reviews of the Cambridge Philosophical Society* 96: 310–329.

Pavesi, Angelo, Alberto Vianelli, Nicola Chirico, Yiming Bao, Olga Blinkova, Robet Belshaw, Andrew Firth, David Karlin. 2018. "Overlapping Genes and the Proteins They Encode Differ Significantly in Their Sequence Composition from Non-Overlapping Genes." *PLOS ONE* 13 (10): e0202513

Pilpel, Yitzhak. 2011. "Noise in Biological Systems: Pros, Cons, and Mechanisms of Control." In *Yeast Systems Biology*, edited by Jesus Castrillo and Stephen Oliver, Methods in Molecular Biology 759: 407–425. Totowa, NJ: Humana Press.

Putois, Olivier, François Villa, and Jonathan B. Weitzman. 2015. "Lost in Post-Translation." *EMBO Reports* 16: 671.

Querido, Jailson B., Masaaki Sokabe, Sebastian Kraatz, Yuliya Gordiyenko, Mark Skehel J., Christopher S. Fraser, and V. Ramakrishnan. 2020. "Structure of a Human 48S Translational Initiation Complex." *Science* 369: 1220–1227.

Raj, Arjun, and Alexander van Oudenaarden. 2008. "Nature, Nurture or Chance: Stochastic Gene Expression and Its Consequences." *Cell* 135: 216–226.

Rao, Christopher, Denise M. Wolf, and Adam P. Arkin. 2002. "Control, Exploitation and Tolerance of Intracellular Noise." *Nature* 420 (6912): 231–237.

Savulescu, Anca F., Caron Jacobs, Yutaka Negighi, Laurianne Davignon, and Musa M. Mhlanga. 2020. "Pinpointing Cell Identity in Time and Space." *Frontiers in Molecular Biosciences* 7: 209.

Schmidt, Andreas, Guoming Gao, Saffron R. Little, Ameya P. Jalihal, and Nils G. Walter. 2020. "Following the Messenger: Recent Innovations in Live Cell Single Molecule Fluorescence Imaging." *Wiley Interdisciplinary Reviews RNA* 11 (4): e1587.

Schwarz, Andre, and Martin Beck. 2019. "The Benefits of Cotranslational Assembly: A Structural Perspective." *Trends in Cell Biology* 29 (10): 791–803.

Sieber, Patricia, Matthias Platzer, and Stefan Schuster. 2018. "The Definition of Open Reading Frame Revisited." *Trends in Genetics* 34 (3): 167–170.

Sober, Elliott. 1984. *The Nature of Selection*. Cambridge, MA: MIT Press.

Sokabe, Masaaki, and Chris S. Fraser. 2019. "Toward a Kinetic Understanding of Eukaryotic Translation." *Cold Spring Harbor Perspectives in Biology* 11: a032706.

Sonneveld, Stijn, Bram M.Verhagen, and Marvin E. Tanenbaum. 2020. "Heterogeneity in mRNA Translation." *Trends in Cell Biology* 30 (8): 606–618.

Symmons, Orsolya, and Arjun Raj. 2016. "What's Luck Got to Do with It: Single Cells, Multiple Fates, and Biological Nondeterminism." *Molecular Cell* 62 (5): 788–802.

Strohman, Richard C. 2003. "Genetic Determinism as a Failing Paradigm in Biology and Medicine: Implications for Health and Wellness." *Journal of Social Work Education* 39 (2): 169–191.

Stuart, Tim, and Rahul Satija. 2019. "Integrative Single-Cell Analysis." *Nature Reviews Genetics* 20 (5): 257–272.

Taiz, Lincoln, Eduardo Zeiger, Ian M. Møller, and Angus Murphy. 2014. *Plant Physiology and Development*. 6th ed. Sunderland, MA: Sinauer Associates.

Tsimring, Lev S. 2014. "Noise in Biology." *Reports on Progress in Physics* 77 (2): 026601.

Wang, Jian, and Yuanlin Song. 2017. "Single Cell Sequencing: A Distinct New Field." *Clinical and Translational Medicine* 6 (10): 1–11.

Wang, Hongyun, Lexi Sun, Anthony Gaba, and Xiaohui Qu. 2020. "An In Vitro Single-Molecule Assay for Eukaryotic Cap-Dependent Translation Initiation Kinetics." *Nucleic Acids Research* 48 (1): e6.

Weber, Marcel. 2001. "Determinism, Realism, and Probability in Evolutionary Theory." *Philosophy of Science* 68 (S3): S213–S224.

Wong, Tien Y. 2020. "Sources of Evolutionary Contingency: Chance Variation and Genetic Drift." *Biology & Philosophy* 35: 36.

Wong, Tien Y. 2022. "The Neo-Gouldian Argument for Evolutionary Contingency: Mass Extinctions." The British Journal for the Philosophy of Science 73 (4): 1093–1124.

Wright, Bradley W., Zixin Yi, Jonathan S. Weissman, and Jin Chen. 2022. "The Dark Proteome: Translation from Noncanonical Open Reading Frames." *Trends in Cell Biology* 32 (3): 243–258.

Xu, Chuan, and Jianzhi Zhang. 2020. "Mammalian Alternative Translation Initiation Is Mostly Nonadaptive." *Molecular Biology and Evolution* 37 (7): 2015–2028.

Xu, Yichen, and Ruggero, Davide. 2020. "The Role of Translation Control in Tumorigenesis and Its Therapeutic Implications." *Annual Review of Cancer Biology* 4: 437–457.

Zenklusen, Daniel, Daniel R. Larson, and Robert H Singer. 2008. "Single-RNA Counting Reveals Alternative Modes of Gene Expression in Yeast." *Nature Structural and Molecular Biology* 15 (12): 1263–1271.

Zernicka-Goetz, Magdalena, and Sui Huang. 2010. "Stochasticity versus Determinism in Development: A False Dichotomy?" *Nature Reviews Genetics* 11 (11): 743–744.

Zhao, Yunpo, Bo G. Lindberg, Shiva E. Seyedoleslami, Xiongzhuo Tang, Stefano Piazza, and Ylva Engström. 2021. "Stop Codon Readthrough Alters the Activity of a POU/Oct Transcription Factor during Drosophila Development." *BMC Biology* 19: 185.

ISSN 2475-3025