Organisms

Journal of Biological Sciences

SAPIENZA
UNIVERSITÀ DI ROMA

Vol. 7, No. 1 (2024)

ISSN: 2532-5876

Organisms

EDITORS

Editor-in-Chief

Associate Editors

Editorial Board

Managing Editor

ORGANISMS is a biannual, open access, peer-reviewed, online journal devoted to basic, translational, theoretical and clinical research. Topics of interest include but are not limited to the following areas: complexity in biology; carcinogenesis; theory of biological organization; biophysics; mathematical and computational approaches in biology and medicine; developmental processes, and degenerative diseases.

ORGANISMS' Editorial Board is involved in research programs into many aspects of normal and pathological inflammation. The Board promises to conduct and facilitate unbiased and critical reviews of manuscripts.Every proposal submitted for publication is read at least by an editor for a preliminary review. If the paper matches the editorial policies and quality level, then it is sent to reviewers for in-blind evaluation via the OJS platform.

ORGANISMS provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge. Authors of articles published in ORGANISMS retain the copyright of their works. Authors can archive pre-print, post-print, and publisher's versions.

The journal is committed to maintaining the highest level of integrity in the content published. All research must have been carried out within an appropriate ethical framework. If there is suspicion that work has not taken place within an appropriate ethical framework, Editors will follow the COPE guidelines and may reject the manuscript, and/or contact the author(s)' institution or ethics committee.

Research involving human subjects, human material, or human data, must have been performed in accordance with the Declaration of Helsinki, and must have been approved by an appropriate ethics committee. A statement detailing this, including the name of the ethics committee and the reference number where appropriate, must appear in all manuscripts reporting such research. If a study has been granted an exemption from requiring ethics approval, this should also be detailed in the manuscript (including the name of the ethics committee that granted the exemption). Further information and documentation to support this should be made available to Editors on request. Manuscripts may be rejected if the Editor considers that the research has not been carried out within an appropriate ethical framework. In rare cases, Editors may contact the ethics committee for further information.

When reporting experiments on animals, authors should indicate whether the institutional and national guide for the care and use of laboratory animals was followed. See International Committee of Medical Journal Editors ("Uniform Requirements for Manuscripts Submitted to Biomedical Journals") February 2006.

https://rosa.uniroma1.it/rosa04/organisms/index

CONTENTS

Perspective & Hypotheses

My Personal View on Science

Books

Cover image: *An alcon blue butterfly on common sainfoin*. Photo: Fran Junebug, 2020. Image from Pixabay.

Perspective & Hypotheses

Vol. 7, No. 1 (2024)
ISSN: 2532-5876

DOI: 10.13133/2532-5876/18555

Single Cell Sequencing Techniques and Biological Explanation

Caroline Angleraux[a*]

[a] *UMR 1253, iBrain, Université de Tours, Inserm, Tours, France*

***Corresponding author:** Caroline Angleraux, Email: Caroline.Angleraux@univ-paris1.fr

Abstract

The aim of this paper is to examine the type of biological explanation implied by single-cell sequencing, using established frameworks in the philosophy of biology, particularly those of new mechanical and systems biology. While investigating the extent to which new mechanistic philosophy or systems biology represent theoretical frameworks that align with single-cell sequencing, a part of -omics sequencing techniques, I claim that the objective of single-cell sequencing corresponds with the *zeitgeist* in theoretical philosophies of biology. The *zeitgeist* is a stance that advocates for a broader perspective on living organisms and that rejects reductionism. However, there remains a disparity between the scientific narrative and the practical capabilities of single-cell sequencing. Consequently, the conclusion drawn in this paper is that while single-cell sequencing aligns with the *zeitgeist* in certain theoretical philosophies of biology, it also acknowledges their theoretical limitations.

Keywords: ontology, cell type, cell classification, single-cell technology

Citation: Angleraux, C 2024, "Single cell sequencing techniques and biological explanation", *Organisms: Journal of Biological Sciences*, vol. 7, no. 1, pp. 5-12. DOI: https://doi.org/10.13133/2532-5876/18555

1. Introduction

Single-cell sequencing involves sequencing the genetic material of each cell individually. This technique offers high resolution at the genetic level and highlights the heterogeneity present within cell populations of tissue samples. Consequently, it facilitates the identification of distinct cell types and the composition of various cell populations. Additionally, by adapting to various levels of resolution, it also illuminates the dynamic nature of cellular structures. It not only underscores tissue development and microenvironments but also enables the tracking of cell lineages and specifications. As synthesized by Wang and Navin, several common applications have emerged from single cell sequencing methods over the last decade: "(1) delineating population diversity, (2) tracing cell lineages, (3) classifying cell types, and (4) genomic profiling of rare cells" (Wang and Navin 2015, p. 606). These diverse applications explain why single-cell techniques are employed across a broad spectrum of research and clinical contexts, including neurobiology, tissue mosaicism, microbiology, germline transmission, embryogenesis, organogenesis, prenatal diagnosis, immunology, and cancer research.

In the past decade, the use of single-cell techniques has experienced significant growth, emerging as both a valuable and trendy method of sequencing. Moreover, its capability to merge and integrate data from other sequencing techniques led to the recognition of single-cell multimodal -omics as the "method of the year 2019" by *Nature Methods* (2020). This acknowledgment underscores the unprecedented precision of the data obtained and the integration possibilities these techniques offer. In essence, single-cell sequencing facilitates the production and the combination of diverse datasets, which helps with the elucidation of complex patterns.

The approach of single-cell sequencing, as part of multimodal -omics, seeks to foster a more integrated understanding of biological phenomena, striving for a comprehensive perspective. By preserving and managing the complexity of collected data, single-cell multimodal -omics aims for complete data integration, thereby facilitating a holistic view of biological structures (cf. for instance Lähnemann *et al.* 2020, p. 22).

In this regard, as I will show, single-cell multimodal omics aligns closely with prevailing trends in philosophies of biology, which elaborate on the theoretical foundations of biology and which advocate for a broader understanding of living organisms. In other words, various theoretical philosophies of biology, such as systems biology or new mechanistic approaches, are deeply committed to incorporating explanations of complex biological phenomena into their frameworks. And by complex biological phenomena, they mean nonlinear effects or emergent properties. As we will develop further, their will to take into account complex biological phenomena is part of an ambition to embrace a holistic approach (Bechtel 2016). Then, as a shared inclination to consider complex biological phenomena within a holistic approach, single-cell multimodal omics mainly matches with the philosophical *zeitgeist*. As Tseng and Santra write: "Over the last two decades, there has been a significant shift towards studying biological cell function in a holistic manner, rather than adhering to a reductionist scientific paradigm, thus establishing the approach known as 'systems biology' or 'systomics'" (Tseng and Santra 2016, p. IV).

While both scientific and philosophical narratives advocate for a holistic perspective on biological phenomena, in practice, their explanations typically rely on the analysis of components and their qualities, following a bottom-up approach. While there is a tendency to integrate data from various levels and construct networks within systems to provide explanations, this integration seldom involves top-down approaches, and thus does not necessarily result in a holistic view being achieved. Upon closer examination of its epistemological framework, a mismatch becomes apparent between the scientific narrative (what it is aimed at doing – Morgan and Norton Wise 2017) and actual practices (what it actually does for now). This gap underscores the disparity between theoretical aspirations and practical implementation.

2. A Mismatch Between the Scientific Narrative and Practices

Single cell sequencing techniques are used for performing crucial tasks, mainly for identifying precise cell types or cell profiles and tracing cell lineages. The way of performing these tasks varies widely, depending on the context of experiment and on available equipment. For instance, in order to isolate a cell within a sample, methods include, among others, serial dilution, laser capture microdissection (LCM), or microfluidics (Wang and Song 2017). There are also different sequencing methods, depending on what is targeted (*e.g.* genome, epigenome or transcriptome; Wang and Song 2017, p. 3). And the way of organizing and integrating data in the libraries also depends on the kinds of biological samples that have been analyzed[a]. In essence, single cell sequencing can be applied to a wide range of objects and comprises a plurality of methods. Beyond the wide spectrum of techniques and methods it encompasses, the main steps involve isolating cells, sequencing genetic material, cataloging data, and analyzing it. Single-cell sequencing generates large and dense datasets, prompting many studies to attempt combining and integrating data obtained from genomic, epigenomic, or transcriptomic layers. As Kashima *et al.* (2020) list them, different computational methods have been developed to provide an overview of single cell data sets and to achieve multiomic analyses.

In this process, from isolating samples to combining datasets, each step is shaped by a technical context as well as driven by epistemic choices. For instance, in an experimental design aimed at identifying different cell populations within a multicellular tissue sample for cancer research, different granular level used to cluster cells ends up providing different number of cell populations within the sample. Within these cell populations, the pursuit of cell types or states (cf. Gross 2023 and Trapnell 2015) influences the selection of keywords in datasets. In another example, when the objective of an experimental design is to better understand transcription processes in bacteria, making decisions such as distinguishing between technical noise and lack of gene expression, or structuring databases to integrate data from samples of different species (cf. Zhang, Gao and Wang 2018)

[a] "Different types of measurements from multiple experiments need to be obtained and integrated. Depending on the actual research question, such experiments can be different time points, tissues, or organisms. For their integration, we need flexible but rigorous statistical and computational frameworks" (Lähnemann & *al.* 2020, p. 21).

is necessary. Their decisions are epistemic choices that shape the interpretation of results. As Leonelli argues in the context of scientific datasets: "The choice and definition of keywords used to classify and retrieve data matters enormously to their subsequent interpretation. Linking diverse datasets means making decisions about the concepts through which nature is best represented and investigated." (Leonelli 2019, p. 2). Technicians and scientists conducting experiments can justify the epistemic reasons behind their experimental designs, and data scientists can explain how and why datasets are elaborated or combined in specific ways. However, scientists who use bioinformatic data without generating them often overlook the underlying epistemic choices that shape the overall design and outputs. Published papers rarely make these assumptions explicit.

As a consequence, many single-cell sequencing studies exhibit a mismatch between the experimental approach and the overarching scientific narrative. In practice, research teams often rely on genome sequencing to infer specific cellular mechanisms or, more broadly, organic processes. However, within the narrative, these same works are presented as offering a holistic understanding of the given process. As a general but accurate example of this popular call for a holistic understanding, the article by *Nature Methods* (2020) that recognizes single-cell multimodal omics as the method of the year 2019 develops claims in its subtitle that single-cell multimodal omics measurement "offers opportunities for gaining holistic views of cells one by one". Then, while *de facto* the majority of single-cell sequencing studies are grounded in a reductionist approach with bottom-up explanations, they also advocate for a holistic view of their subject of study. In this context, "reductionism" is an epistemological approach that deduces processes, behaviors, or qualities of a system from the qualities or combinations of its components; it employs a bottom-up explanation, as higher levels of the biological system are explained by properties from the lower ones. Moreover, the use of the term "holistic" implies a comprehensive yet precise understanding of the mechanism, contextualized within a specific tissue; it often involves combining data from different levels of analysis as well (one can explicit this characterization from Polychronidou and *al.* 2023 for instance). However, it seldom involves top-down explanations, which consider the impact of structural properties or functional states of the whole system on subsystems. It also rarely addresses supra-cellular levels or combines datasets from sub- and supra-cellular analyses. (On these classical distinctions, cf. for instance Gilbert and Sarkar 2000; Mazzochi 2012 or Soto and Sonnenschein 2018).

This difference between the actual practice and the narrative is quite common in the literature. For instance, Mujal and *al.* (2022) investigated the differentiation from monocyte to macrophage in kidney cancer using mouse and human tissues. They employed single-cell RNA-sequencing analysis of tumors and discovered, among other findings, that immune cell differentiation was correlated with the amount of regulatory T cells in the mouse model. They also demonstrated that heterogeneity in macrophages cancer was correlated with regulatory T cell density. They asserted: "In this way, holistic analysis of monocyte-to-macrophage differentiation creates a framework for critically different immune states." In other words, this study highlights correlations between certain types of cells, their quantities, and certain physiological characteristics (such as density) based on RNA analysis. While the authors characterize their approach as holistic (perhaps because they identify correlations and integrate various analyses), their study, in practice, remains within a reductionist framework, drawing inferences from sequencing data.

In a similar vein, Park and *al.* (2021) try to enlighten how transcriptomic landscape of individual hepatocytes is altered in response to a high-fat diet, aiming for a "holistic characterization" of hepatocytes. While single-cell transcriptome studies have revealed that hepatocyte gene expression and function vary widely across their metabolic zonation, this paper emphasizes that the patterns of transcriptome alteration depend on the metabolic zones, with some responses being independent of the zonation profile. Thus, this study relies on a single-cell RNA-sequencing dataset and uses specific markers to define metabolic zonation profiles, employing a bottom-up explanatory approach. In this context, a holistic characterization entails deducing metabolic states from transcriptomic data, even though it struggles to account for the complex structure of liver tissue. As the authors themselves acknowledge: "it is possible that this [the given method] is an oversimplification of the complex histological architecture of the liver" (Park *et al.* 2021).

Single-cell practices include the description of molecular processes and clustering, and they predominantly employ bottom-up approaches without incorporating top-down perspectives. Then, single-cell practices do not meet their narrative, their advocacy of holism. This limitation to meet this goal is mainly understood by those who use these techniques as a technical limitation. And it is

the reason why corporations like 10x which produce single-cell tools develop new sequencing and analysis techniques in order to integrate and combine better data from different biological levels. As an example, they advertise a "Visium Spatial Gene Expression" that integrates total mRNA analysis for intact tissues sections with morphological context. The point is to better identify the connection between gene expression and morphological context, which means to better correlate the connection between different biological levels. It aims to better fulfil the holistic narrative, which entails a broader intention to combine and integrate data from various biological levels of analysis, to better justify the complex structures of biological organization.

Nevertheless, the mismatch between the narrative and actual practices persists for now. While technical development plays a part, the persistence of this mismatch may also be explained by a gap between practices and the conceptual framework in which experimental results are understood. In other words, single cell practices, *de facto*, take place in a reductionist approach but develop a "holistic" narrative that targets a comprehensive approach to living processes. Then, the mismatch between practices and narrative may also result from the underdevelopment of a conceptual framework that could read data results in a theoretical context that matches the narrative. Indeed, as it currently stands, this mismatch is epistemologically questionable because it results in a situation where, in practice, data continue to accumulate, yet the theoretical framework guiding their interpretation does not align with the intended narrative of holism.

Consequently, while experimental outputs must be analyzed within a conceptual framework to provide meaningful insights and achieve a comprehensive biological explanation – not merely an accumulation of data – this conceptual framework still appears to be under development. As Krohs and Callebaut elaborate: "The huge amounts of data produced by the genome projects were in fact collected almost free of any theoretical burden; as could have been expected, they turned out to explain next to nothing". A few pages later, they add: "'Omics', however, lack a theoretical framework that would allow to use these data sets as such (rather than just tiny bits that are extracted by advanced data-mining techniques) to build explanatory models that help understand physiological processes." (Krohs and Callebaut 2007, p. 184 and p. 208). In the case of omics studies, including single-cell analysis, the general framework for attributing meaning to data and contextualizing a biological explanation is not yet fully developed. This tension can also be perceived in Lähnemann and *al.* (2020): while this paper lists challenges that single cell data science must overcome and presents them as technical issues, it actually lists epistemological challenges to overcome (*e.g.*, how to deal with errors and missing data in the identification of variation from sequencing data, how to map single cells to a reference atlas, or how to integrate data across samples, experiments and types of measurement). Consequently, the gap between practices and the scientific narrative can be attributed to the fact that the theoretical framework that provides meaning to these practices is still in development. Moreover, as we will see in the next section, this theoretical framework itself seems to exhibit the same kind of mismatch.

3. Systems Biology as a Theoretical Approach to Single Cell Practices?

Systems biology seems to be the theoretical framework favored by users of single-cell analysis. As mentioned in the introduction, Tseng and Santra (2016) assert that systems biology represents the best theoretical approach for examining biological processes in single-cell analysis and, more broadly, omics. Similarly, Veenstra (2021) elaborates on how single-cell data are employed within a systems biology approach and how omics advances within a systems biology framework.

Systems biology seeks to explore how the functional properties of a living system, such as a cell or an entire organism, are brought about by the interactions among its components or parts (Boogerd *et al.* 2007). For instance, at a cellular level, systems biology examines the relationship between molecules and cells in two ways. Firstly, cells are viewed as organizing molecular systems to understand how functional properties arise from specific interactions between molecular processes. Secondly, cells are examined through their molecular properties to explain and predict cellular behavior. Thus, systems biology seeks to integrate bottom-up and top-down approaches in studying living systems. In a top-down approach, the focus is placed on molecular behaviors within living systems, regarded as wholes. In a bottom-up approach, emphasis is on molecular properties to understand how parts of the system interact (Boogerd *et al.* 2007). This combination of both approaches aims to apprehend biological phenomena on a broad basis, assuming that component behaviors within a living system are involved in nonlinear interactions. And yet, "in nonlinear interactions, qualitatively

new properties can arise, depending on the state the system is in, as the strength of the interactions vary with that state" (Boogerd *et al.* 2007, p. 11). Then, one may study molecular properties and behaviors in relation to the overall state of the system to gain a clearer picture of emergent and non-emergent properties of the living system.

In the narrative of systems biology, we observe that understanding living systems begins with decomposing and identifying parts of the system; the aim is to identify components and then observe how they interact. Veenstra describes the research progression in systems biology as akin to assembling a jigsaw puzzle. He outlines a three-step methodology, which involves identifying the pieces, organizing them into manageable parts, and finally assembling them to reveal the complete picture of the system operation (Veenstra 2021, p. 9). The method primarily focuses on the pieces and their assembly, often overlooking the examination of the constraints of the whole system or processes that involve entities at different levels of the system – as instantiated by Cornish-Bowden and *al.* (2004) with the example of metabolism. Consequently, this focus helps explain why, in practice, works claiming to be in systems biology often prefer bottom-up approaches. As Cornish-Bowden and *al.* assert: "Despite the current vogue for 'systems biology', this term is often little more than a euphemism for gathering ever more details on an ever larger scale, and not, as it should be, the study of biological systems as systems rather than collections of components" (Cornish-Bowden *et al.* 2004, p. 713).

Moreover, the narrative of systems biology strongly emphasizes molecular analysis. Whether employing a bottom-up or a top-down approach, the focus remains on studying molecular properties or behaviors to better define the relationship between molecular structures and functions. Molecular analysis is deeply entrenched in a well-established tradition that typically employs the bottom-up approach to study biological phenomena. This way of explaining biological phenomena based on molecular analysis often raises questions regarding whether these phenomena are epistemologically reducible to physiochemical phenomena (cf. Mossio and Umerez 2014). As a result, it explains why, in practice, works claiming to be in the realm of systems biology often favor bottom-up approaches.

Thus, systems biology does not fully achieve what its scientific narrative claims. There is a mismatch within the theoretical framework intended to support single-cell analysis, single-cell omics. In other words, just as the technique and method of single-cell analysis have not yet achieved what researchers set out to do according to their own narrative, similarly, systems biology has not yet fully addressed its objectives, despite being perceived as the privileged theoretical framework for single-cell analysis. Consequently, this is a situation in which the theoretical framework evolves alongside the techniques it supports. Additionally, due to this ongoing development simultaneously, the gap between usage and narrative may be perceived as resulting from a technical obstacle: if the gap remains unabridged, it is because omics technologies continue to evolve, particularly in elucidating the evolutions of biological processes in their spatiotemporal context. This is the argument put forward by Veenstra (2021, p. 7), who subsequently adds nuances:

> "While the progress made in omics research is exciting, a complete systems biology view that enables us to accurately predict how cells and organisms respond to either internal (e.g., gene mutations) or external (e.g., drug treatment) events is still in the distant future. Sometimes it appears this capability is beyond our reach. As we learn more about known components of the cell, new classes of biological molecules are discovered that have profound effects on how the cell functions." (Veenstra 2021, p. 9)

The promise of deeper understanding in biology appears to hinge on technological advancements, particularly in the detailed analysis of genetic material. It is therefore coherent that a bottom-up approach is still preferred, given the correlation between systems biology development and technological advancements. However, while the inability to fully implement a systems biology framework is often described resulting from technical obstacles, the underlying issue may be more theoretical in nature. As Callebaut claims, relying on Cornish-Bowden, "most papers in which the words 'systems biology' appear 'have surprisingly little to do with older notions of biological systems' such as the systems theory advocated by von Bertalanffy (1969) or the work of Robert Rosen (1934-1998)" (Callebaut 2012, p. 72). While advocating for a greater emphasis on functional aspects when theorizing biological phenomena, systems biology, in practice, still closely resembles traditional ways of doing biology. Functional and organismic perspectives, as previously emphasized by early pioneers of living systems biology, are still not as extensively incorporated as might be expected.

In summary, systems biology is presented as an operational theoretical framework for single-cell omics. Systems biology includes a variety of perspectives, yet the overarching aim is generally to

integrate both bottom-up and top-down approaches. This integration seeks to consider both organismic context and system decomposition, ultimately leading to a more comprehensive understanding of biological phenomena. However, the theoretical proposals applied in practice often fall short of the ambitious claims made by systems biology. This discrepancy is one reason why Callebaut advocates for "scientific perspectivism" which integrates different perspectives to enhance scientific practice and theoretical understanding. He also suggests that his scientific perspectivism could align with the principles of new mechanistic philosophy[b]. Building on this premise, could new mechanistic philosophy offer another theoretical framework for single-cell omics?

4. New Mechanistic Philosophy as a Theoretical Approach to Single Cell Practices?

New mechanistic philosophy originates from the classical mechanistic views of the 17th century, developed by figures such as Galileo and Descartes. This approach aimed to elucidate complex phenomena by breaking them down into interactions among their constituent parts, explainable by principles of motion. While this model is suitable for reducing various phenomena to physical laws, its application to biology raises questions about the reduction from biology to biochemical processes. To avoid such reductions while maintaining mechanistic views of biological systems, new mechanistic philosophy has emerged to provide a distinct explanatory framework inferred from the conceptual underpinnings of everyday biological practice. New mechanistic philosophy is advocated by scholars such as Craver, Darden, Bechtel, or Richardson; it encompasses a diverse range of research, and not all proponents share identical claims.

However, in general, new mechanistic philosophy regards living beings as natural systems organized into subnetworks of parts. The point is to identify parts of the system and how they are organized in order to understand how the activity of the whole system results from the activity of its organized parts. In other words, by adopting mereological perspective, a mechanistic approach considers that a living system is structured into parts and the performance of the system functions and subfunctions result from the way these parts interact. Bechtel and Richardson (1993) highlight two strategies employed in biology: decomposition, which involves physically or conceptually separating system components; and by localization, which entails precisely identifying the parts and their interactions that give rise to biological phenomena. These descriptive strategies are what new mechanistic philosophy consider to be explanations, which amount to the analysis of the constitutive parts of a system. This kind of explanation allows the prediction of future behavior, thereby facilitating anticipation of potential experimental modifications.

While classical views about mechanism vouch for reductionist explanations of living beings, Bechtel and Richardson advocate for non-reductionist ones. They contend that, in light of the complex and non-linear effects observed in living systems, it is essential to consider how these effects emerge from the interactions among their components. Living systems are viewed as integrated systems with emergent effects[c] and a multi-level explanation is necessary to properly justify the identification of complex causal mechanisms in living organisms. Moreover, since component parts and operations can be modified by elements both within and outside the system, mechanistic explanations may also incorporate a study of the system's environment and the top-down constraints that impact the system, particularly during development. Then, in these new mechanistic approaches that attempt to depict the complexity of natural systems from a non-reductionist perspective, explanation involves describing how a biological process works and determining the causal networks that enable the process to operate (cf. Bechtel 2006, p. 34). A phenomenon is considered explained when distinguishing features are identified within specific sections of the natural system and when these features are connected through a particular causal network. As such, new mechanistic philosophy appears like a suitable theoretical framework for single-cell omics to rely on.

However, opting for new mechanistic philosophy as a theoretical framework of single cell approaches would also lead to tensions. Considering that new mechanistic philosophy is based on how biological research works, addressing issues through the mechanistic approach would equate, in a certain way, to addressing issues about how biology, as a

[b] "Scientific perspectivism inaugurates 'a methodological victory for Leibnizian organicism over a one-sided Cartesian mechanicism" (Toulmin, 1982, p. 138) – while I simultaneously believe the former can be fully cashed out in terms of the "New Mechanistic Philosophy of Science" developed by Bechtel, Darden, Glennan, and others" (Callebaut 2012, p. 75).

[c] "We suggested that such behavior could be seen as 'emergent' at least insofar as the organization of the system, rather than distinctive contributions of its constituent components, determines systemic function" (Bechtel and Richardson 1993, p. xxxv).

way of experimenting, works. In this regard, a key epistemological consideration regarding mechanisms is their relation to reductionism. To what extent does new mechanistic philosophy truly integrate bottom-up and top-down approaches? Indeed, in mechanistic explanations, the integration of the different levels of organization sometimes remains problematic, particularly as the smaller component explanatory level remains the main level of analysis. As Nicholson explains: "This heuristic fragmentation of the organism into causal mechanisms, despite being necessary for its investigation, often comes at the expense of neglecting the way in which the organism as a whole influences the behaviour of its parts" (Nicholson 2012, p. 159). In other words, while new mechanistic philosophy claims to integrate the different levels of organization for explanation, *de facto*, the focus remains primarily on the more basic parts of the natural system. Additionally, the method of decomposing the natural network depends mainly on the structural properties of the system rather than its functional properties.

Consequently, new mechanistic philosophy may also exhibit a mismatch between what they aim to achieve conceptually and their actual conceptual limitations. At least, they often present ambiguities regarding reductionism and the integration of top-down and bottom-up approaches, mirroring challenges encountered in single-cell omics. Given that new mechanistic philosophy is grounded in biological practices and highlights their theoretical foundations, it only makes sense that they encounter similar theoretical uncertainties as some of those found in systems biology.

5. Conclusion

Single-cell analysis represents an unprecedented advancement in omics research. By elucidating the heterogeneity of cell populations or lineages within a sample, they enable a unique level of inference, facilitating comprehensive studies of biological phenomena. However, upon closer examination, a mismatch emerges between the aspirations of single-cell omics—such as achieving a framework that integrates data from different levels of analysis—and their actual experimental procedures. In particular, the scientific narrative of single-cell analysis advocates for a holistic view of biological processes, emphasizing the broader intention to integrate databases across various biological levels. However, in practice, it primarily involves the description of molecular processes and clustering and predominantly relies on a bottom-up approach, neglecting to incorporate a top-down perspective. As a result, a mismatch persists between the intended narrative and the current practices in single-cell analysis.

This mismatch calls for an explanation of the theoretical background of single-cell analysis. From a biologist's perspective, this gap between practice and narrative primarily arises because of technical challenges that need to be overcome. Despite the emergence of new technical methods for processing data, such as improved integration of morphological context, dynamic views of cell specialization, and enhanced sample preservation in microbiological sequencing, the theoretical framework and its underlying assumptions remain inadequately implicit. Some authors have suggested that systems biology is perceived as a suitable theoretical framework for understanding single-cell omics. However, both systems biology new mechanistic philosophy – which has been examined as potential suitable theoretical contexts – exhibit a similar mismatch to the one observed in single-cell omics. Overall, the emphasis on holism in narratives of omics and systems biology, along with the specific attention given to nonlinear properties in systems biology and new mechanistic philosophy, primarily reflect a principled opposition to reductionism. From this initial stance against reductionism, theoretical positions are still under development.

To question the mismatch and advance the development of a suitable framework for single-cell studies, it is imperative to explicit background assumptions. From an epistemic standpoint, the goal is to delineate these assumptions to justify the level of detail to be included and to clarify its relevance to the organismic context of explanation. Explicitly stating the theoretical background and, consequently, constructing a theoretical framework that incorporates organismic and functional perspectives are crucial endeavors aimed at achieving the comprehensive biological explanations that single-cell omics seek to provide. Additionally, from a philosophical perspective, there is a need to further develop conceptual mappings of reductionism, emergentism, and organicism to continually refine the narrative aimed at comprehensively explaining biological phenomena. There are also reasons to think that this kind of gap may be a recurrent phenomenon in science. The analysis realized in this paper therefore represents a high potential of generalization that could help improve understanding of scientific dynamics beyond the boundaries of single-cell omics.

Acknowledgments

I would like to express my gratitude to François Papale for his careful proofreading and the time he spent discussing my work. I also thank the reviewers for their comments and the labex "Who am I?" for its financial support.

References

Bechtel, W 2006, *Discovering cell mechanisms: The creation of modern cell biology*, Cambridge University Press, Cambridge.

Bechtel, W, & Richardson, R 1993, *Discovering complexity: Decomposition and localization as strategies in scientific research*, Princeton University Press, Princeton.

Bechtel, W 2016, "Mechanists Must be Holists Too! Perspectives from Circadian Biology", *Journal of the History of Biology*, 49, pp. 705-731.

Boogerd, FC, *et al.* 2007, "Towards philosophical foundations of Systems Biology: introduction", in FC Boogerd, FJ Bruggeman, J-HS Hofmeyr and HV Westerhoff, (eds), *Systems Biology, Philosophical Foundations*, pp. 3-19. Elsevier, Amsterdam.

Callebaut, W 2012, "Scientific perspectivism: A philosopher of science's response to the challenge of big data biology", *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43, pp. 69-80.

Cornish-Bowden, A, *et al.* 2004, "Understanding the parts in terms of the whole", *Biology of the Cell*, 96, pp. 713-717.

Dal Molin, A, & Di Camillo, B 2019, "How to design a single-cell RNA-sequencing experiment: pitfalls, challenges and perspectives", *Briefings in Bioinformatics*, 20, pp. 1384-1394.

Gatherer, D 2010, "So what do we really mean when we say that systems biology is holistic?", *BMC Systems Biology*, 4, 22. Available from: https://bmcsystbiol.biomedcentral.com/articles/10.1186/1752-0509-4-22#citeas.

Gilbert, SF, & Sarkar, S 2000, "Embracing Complexity: Organicism for the 21st Century", *Developmental Dynamics*, 219, pp. 1-9.

Gross, F 2023, "Do Single-cell Experiments Challenge the Concept of Cell Type?", *Organisms. Journal of Biological Sciences*, 6, 2, pp. 5-9. Available from: https://rosa.uniroma1.it/rosa04/organisms/article/view/18131.

Kashima, Y, et *al.* 2020, "Single-cell sequencing techniques from individual to multiomics analyses", *Experimental & Molecular Medicine*, 52, pp. 1419-1427.

Krohs, U, & Callebaut, W 2007, "Data without models merging with models without data" in FC Boogerd, FJ Bruggeman, J-HS Hofmeyr and HV Westerhoff, (eds), *Systems Biology, Philosophical Foundations*, pp. 181-213. Elsevier, Amsterdam.

Lähnemann, D, *et al.* 2020, "Eleven grand challenges in single-cell data science", *Genome Biology*, 21, pp. 1-31.

Leonelli, S 2019, "The challenges of big data biology", *eLife*, 8, pp. 1-5. Available from: https://elifesciences.org/articles/47381.

Mazzocchi, F 2012, "Complexity and the reductionism-holism debate in systems biology", *Wiley Interdisciplinary Reviews: Systems Biology and Medecine*, 4, pp. 413-427.

Morgan, MS, & Norton Wise, M 2017, "Narrative science and narrative knowing. Introduction to special issue in narrative science", *Studies in History and Philosophy of Science*, 62, pp. 1-5.

Mossio, M 2020, *Organisation Biologique et Finalité Naturelle*, Habilitation à Diriger des Recherches (HDR), defended in Paris, unpublished.

Mossio, M, & Umerez, J 2014, "Réductionnisme, holisme et émergentisme", in T Hoquet & F Merlin, (eds), *Précis de philosophie de la biologie*, pp. 157-168, Vuibert, Paris.

Mujal, AM, *et al.* 2022, "Holistic Characterization of Tumor Monocyte-to-Macrophage Differentiation Integrates Distinct Immune Phenotypes in Kidney Cancer", *Cancer Immunology Research*, 10, 4, pp. 403-419.

Nicholson, DJ 2012, "The concept of mechanism in biology", *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43, pp. 152-163.

Papale, F, *et al.* 2023, "The evosystem: A centerpiece for evolutionary studies", *BioEssays*, 46, pp. 1-8.

Park, SR, *et al.* 2021, "Holistic characterization of single-hepatocyte transcriptome responses to high-fat diet", *Endocrinology and Metabolism*, 320, 2, pp. E244-E258.

Polychronidou, M, *et al.* 2023, "Single-cell biology: what does the future hold?", *Molecular systems biology*, 19, pp. 1-5.

Ratti, E 2015, "Big Data Biology: Between Eliminative Inferences and Explanatory Experiments", *Philosophy of Science*, 82, pp. 198-218.

Richardson, R, & Stephan, A 2007, "Mechanism and mechanical explanation in systems biology" in FC Boogerd, FJ Bruggeman, J-HS Hofmeyr and HV Westerhoff, (eds), *Systems Biology, Philosophical Foundations*, pp. 123-144, Elsevier, Amsterdam.

Soto, AM, & Sonnenschein, C 2018, "Reductionism, Organicism, and Causality in the Biomedical Sciences", *Perspectives in Biology and Medecine*, 61, pp. 489-502.

Teichmann, S 2020, "Method of the Year 2019: Single-cell multimodal omics", *Nature Methods*, 17, p. 1.

Trapnell, C 2015, "Defining cell types and states with single-cell genomics", *Genome Research*, 25, pp. 1491-1498.

Tseng, F-G, & Santra, TS 2016, *Single Cell Analysis in Biotechnology and Systems Biology*, MDPI, Basel.

Wang, J, & Song, Y 2017, "Single cell sequencing: a distinct new field", *Clinical and Translational Medicine*, 6, pp. 1-11.

Wang, Y, & Navin, NE 2015, "Advances and Applications of Single-Cell Sequencing Technologies", *Molecular Cell*, 58, pp. 598-609.

Zhang, Y, Gao, J, & Wang, J 2018, "Recent Developments in Single-Cell RNA-Seq of Microorganisms", *Biophysical Journal*, 115, pp. 173-180.

Veenstra, TD 2021, "Omics in Systems Biology: Current Progress and Future Outlook", *Proteomics*, 21, 3-4, pp. 1-11.

Perspective & Hypotheses

Vol. 7, No. 1 (2024)
ISSN: 2532-5876

DOI: 10.13133/2532-5876/18152

Polyploidization, Gene Duplication, and the Origin of Variability: Where is the Evidence?

Milton H. Gallardo[†a] *& Elkin Y. Suárez-Villota*[b*]

[a] *Instituto de Ciencias Marinas y Limnológicas, Universidad Austral de Chile, Casilla 567, Valdivia, Chile*

[b] *Instituto de Ciencias Naturales, Universidad de las Américas, Av. Jorge Alessandri 1160, 5 Piso, Campus El Boldal, Concepción, Chile*

***Corresponding author:** Elkin Y. Suárez-Villota, Email: esuarezv@gmail.com

Abstract

The emergence of molecular complexity and its impact on evolutionary organismal innovations, achieved through the duplication of ancestral states and resulting in the creation of new structures and functions, precedes the acclaimed work of Ohno *et al.* (1968) and the advent of the genome research era. The insights of Margaret O. Dayhoff (1966) let her to conclude that ferrodoxins and other proteins were derived by doubling of short peptides and that proteins' first folded domains arose by duplication, fusion, and diversification of shorter, ancestral peptides. Here, we survey some important milestones related to gene duplication and polyploidy, starting with the contribution of Dayhoff, which reveals a complex landscape where gene duplication emerges as a fundamental evolutionary force (sections 1, 2). We subsequently address polyploidy as one of the factors contributing to gene duplication, collectively driving the evolution of molecular complexity (section 3). In this context, we explore the patterns of gene/genome duplications in the expansion of Hox gene clusters, which serve as signatures of ancient polyploidization, highlighting their role in significant evolutionary transitions (section 4). We then examine the empirical findings of synthetic polyploids and the genetic variation in heat shock proteins in wheat (sections 5, 6). These investigations offer critical insights, suggesting that lineage crossings involving chromosome set duplication (allopolyploidy) are consequential phenomena in hybrid speciation, significantly contributing to the emergence of a substantial portion of macroevolutionary diversity. Allopolyploidy and ancient gene transfers among the three domains of life generate such a variation that mutation rates based on common descent lose preponderance and the notion of tree of life gets suffocated in an entangled genomic bush.

Keywords: macroevolution, genome duplication, allopolyploidy, hybridization, Hox genes

Citation: Gallardo, MH, & Suárez-Villota, EY 2024, "Polyploidization, Gene Duplication, and the Origin of Variability: Where is the Evidence?", *Organisms: Journal of Biological Sciences*, vol. 7, no. 1, pp. 13-22. DOI: https://doi.org/10.13133/2532-5876/18152

1. Margaret O. Dayhoff and the Evolution of Protein Complexity by Gene Duplication

The seminal idea that complex macromolecules and organismal novelties are derived by duplication of ancestral states predates by far the advent of genome research (Eck and Dayhoff 1966). Over 50 years ago, the pioneering work of Margaret O. Dayhoff (1925-1983) led her to conclude that functional proteins evolved through gene duplication, resulting in the the doubling and self-assembly of short peptides. Around the same time, Dayhoff *et al.* (1965) had started the Atlas of Protein Sequence and Structure, a book series, where a model of evolutionary change based on gene mutations and natural selection was

advanced (Dayhoff *et al.* 1978). Surprisingly, the idea that protein complexity is achieved by duplication of simpler units was overlooked by Ohno *et al.* (1968) in his acclaimed work on evolution from fish to mammals by gene duplication.

Dayhoff postulated that sequence homology within domains of tertiary and quaternary structures of ferrodoxins and other ancient proteins resulted from gene duplications (Eck and Dayhoff 1966; Hunt *et al.* 1974; Schwartz and Dayhoff 1978). Following duplication, the fusion of monomers leads to the observed sequence homology within protein domains, which gradually diminishes as each monomer evolves independently (Romero *et al.* 2016). Natural selection is involved in this evolutionary process, restricting sequence variation on the primary structure of proteins (Hunt and Dayhoff 1970).

Ferredoxins, studied by Dayhoff, are crucial enzymes in photosynthesis and exhibit significant internal sequence homology attributed to duplication, fusion, and peptide diversification processes. (Eck and Dayhoff 1966; Schwartz and Dayhoff 1978). Notably, clostridial-type ferredoxins provide compelling evidence for a widespread gene duplication event shared among anaerobic, heterotrophic bacteria near the root of the evolutionary tree. Comparative analyses using ferredoxins, 5S RNA, and c-type cytochromes suggest a secondary duplication event preceding the radiation of eukaryotes and involved in cellular respiration (Schwartz and Dayhoff 1978). Recent investigations into the origins of oxygenic photosynthesis confirm that the Type I photosynthetic reaction centre comprises a heterodimeric core consisting of two homologous subunits (PsaA and PsaB), arising from gene duplication (Cardona 2017). This stands in contrast to the reaction centre of anoxygenic phototrophs, which features a homodimeric core (Liebl *et al.* 1993). A compelling hypothesis regarding the evolution of a heterodimeric Type I reaction centre suggests that the gene duplication enabling the divergence of PsaA and PsaB was a response to incorporate photoprotective mechanisms against the formation of reactive oxygen species, occurring after the origin of water oxidation to oxygen (Cardona 2018). This is consistent with the hypothesis that the event originating the duplicated heterodimeric condition occurred in the early Archean, before the Great Oxidation Event, approximately two billion years ago (Cardona 2018; Oliver *et al.* 2021). In this scenario, marked by a significant increase in biological complexity, aerobic respiration preceded oxygen-releasing photosynthesis (Sousa *et al.* 2013; Cardona 2018; Soo *et al.* 2019).

Dayhoff's concepts regarding gene duplication, fusion, and diversification, along with natural selection, should serve as the foundational framework for elucidating the extensive molecular complexity found in proteins. However, the fundamental process of duplication, which underlies the emergence of complex proteins, has often been overlooked over the years and much of the research has focused on studying the adaptive scenarios of protein evolution (*e.g.* Bloom and Arnold 2009; Jayaraman *et al.* 2022). Indeed, there remains a considerable gap in our understanding of how and when short peptides and domains duplicate, originate, and combine (Buljan and Bateman 2009; Schaper and Anisimova 2015). Gene duplication is frequently regarded as a phenomenon for elucidating macromolecular complexity, functioning both as a source and a consequence of evolutionary processes. Nonetheless, comprehending the mechanisms that lead to duplications is paramount, as they represent the genuine sources (unequal recombination, replication errors, transposition, and polyploidization).

2. Protein Evolution Through Gene Duplication

Proteins are enormously diverse agents of life. Their evolution, based on sequence similarity has provided clues about gene functions. They display substantial sequence similarity and a three-dimensional structure derived by autonomous folding of units or domains (Söding and Lupas 2003). Domains are evolutionary units whose coding sequences may undergo duplication, recombination, and divergence. These processes can occur randomly and be selected not only at the domain level. Typically, small proteins contain one domain of 100 to 250 residues and large proteins contain a combination of them. Moreover, domain families contain small proteins or parts of larger ones, descended from a common ancestor (Chothia *et al.* 2003). Thus, protein domains evolved by gene duplication, forming novel and more complex proteins (Chothia *et al.* 2003; Levy *et al.* 2008). Most models of gene duplication consider genetic redundancy and predict that after doubling, the function and structure of proteins diversify (Conant and Wolfe 2008; Kuzmin *et al.* 2020; Birchler and Yang 2022). In fact, prolonged high rates of evolution would have been determined by functional properties, acquired during, or soon after a gene duplication event (Pich and Kondrashov 2014). Interestingly, while duplications contributed to the emergence of novel traits and species diversification, phylogenetic

analyses in plants have inferred 'lag-times' between duplication events and radiations. This is supported by phylogenetic asymmetries with species-rich crown groups and species-poor sister clades occurring before duplication events. Thus, the diversification of crown groups may involve not only duplication events and novel traits but also evolutionary factors such as migration events, changing environments, and differential extinction rates (Scharanz *et al.* 2012).

At the time, since homologous structures of proteins did not exist, or could not be identified, first models were constructed from scratch. This procedure, called *ab initio* modeling, was the first approach to address the riddle of protein structure (Lee *et al.* 2009). Later on, gene duplication, mutation and recombination, became more important to address the subject. In fact, the genome sequencing of the bacterium *Haemophilus influenzae* indicated that at least one-third of this protein arose after gene duplication (Brenner *et al.* 1995). This doubling process has even been reported in viruses (Shackelton and Holmes 2004; Simon-Loriere and Holmes 2013; Willemsen *et al.* 2016).

Likewise, the detection of two paralogues of the tRNA endonuclease gene of *Methanocaldococcus jannaschii* in the genome of the crenarchaeote *Sulfolobus solfataricus* led to the identification of an unrecognized oligomeric form. Both genes code for different subunits required cleaving the pre-tRNA substrate. There are three forms of tRNA endonucleases in the Archaea, namely, a homodimer and a heterotetramer (in Euryarchaea), and a third, heterotetramer endonuclease (in Crenarchaea and Nanoarchaea). It is postulated that the last one likely resulted by gene duplication (or horizontal gene transfer), and subsequent subfunctionalization (Tocchini-Valentini *et al.* 2005).

Ribosomal protein genes constitute a large class of conserved duplicated genes in mammal with a substantial number of duplicates are transcriptionally active. Selection against dominant-negative mutations would be responsible for its unexpected retention and conservation (Dharia *et al.* 2014). Ribosomal assembly proceeds by fusing two interacting subunits, to the current atomic-resolution structures of the prokaryotic 70S and the eukaryotic 80S ribosomes (Melnikov *et al.* 2012). Large and small subunits has been captured in different functional states (Yusupova and Yusupov 2017). Although inferences about ribosomal origin are speculative (Smith *et al.* 2008; Fox 2010), a model by accretion evolution has been hypothesized using 3D comparative methods. In this model, the ribosome evolved by recursively adding expansion segments, iteratively growing, subsuming, and freezing the rRNA (Petrov *et al.* 2015).

Ancestral protein reconstruction allows the characterization of ancient macromolecules by computational analyses of modern-day protein sequences. Nevertheless, the reconstruction of protein families is limited, as exemplified by the antigen receptors of jawed vertebrates, which evolved from an extinct homodimeric ancestor through gene duplication (Rouet *et al.* 2017). Similar studies supported the idea that most domain gains in animal proteins were directly mediated by gene fusion, preceded by duplication and recombination (Marsh and Teichmann 2010).

In connection with proteins assembling, Levy *et al.* (2008) demonstrated the reversal of the process through the dilution of the denaturant and/or manipulation of the ionic strength. They observed the recovery of the original homodimer in 50% of the studied complexes. They refer to this result as "a molecular analogy to Haeckel´s evolutionary paradigm of embryonic development, where an intermediate in the assembly of a complex represents a form that appeared in its own evolutionary history". It must be said that Haeckel´s biogenetic law contrasts two timescales: ontogenetics and phylogenetics. Ontogeny recapitulates in some way the phylogeny or life history of biological units. Self-assembling, in turn refers to reversible chemical accretion of symmetric multisubunit complexes occurring in the same timescale.

The presence of diverse multigene eukaryotic families underscores the significance of gene duplication followed by the diversification of functional genomes. However, the origins of these duplication events and the myriad of new functions that emerge alongside them remain less understood. There is a scarcity of cases that clearly distinguish patterns from processes, particularly when examining the evolutionary progression of protein functional diversification. Several publications continue to echo Dayhoff's framework of gene duplication, fusion, and diversification. However, the process that may be involved in the gene expansion pattern may be polyploidy, which could entail the hybridization of lineages.

3. Gene Duplication in Polyploidization and Genome Size

Polyploidization occurs when a complete set of chromosomes is added to an existing genome from the same species (autopolyploidy) or

through hybridization (allopolyploidy). These phenomena often result from errors in meiotic and mitotic segregation, leading to chromosomal endoreduplication during gamete production (Consortium 2014; Fox and Duronio 2013). Such processes induce revolutionary and evolutionary changes in the function and structure of the genome (Feldman and Levy 2012; Van de Peer *et al.* 2017). The doubling of DNA content, a consequence of polyploidization, accompanies extensive gene duplications (Van de Peer *et al.* 2017). Given that gene duplication significantly influence protein structure, as discussed in the preceding sections, hybridization processes (allopolyploidy) or autopolyploidy could be implicated in the origin of increased protein domains.

Since polyploidy has occurred repeatedly throughout evolutionary history, genome size increases accordingly, although not in an ideal geometrical progression. Derived from genetic redundancy, molecular and cytological adjustments lead to gene losses and gains, altered regulatory genetic and epigenetic pathways (Markov and Kaznacheev 2016; Van de Peer *et al.* 2017). Genome evolution has followed varied evolutionary pathways as indicated by gradual and quantum shifts accounting for the differences in DNA content (Gallardo *et al.* 2003; Gregory and Hebert 1999). One of the most significant shifts in genome size is exemplified by the transition from haploidy to diploidy (and later, to polyploidy). Indeed, the staggering genome size variation across eukaryotes, mounts to over 64,000-fold whereas in land plants it ranges around 2,400-fold (Gregory 2024; Pellicer *et al.* 2018). Genome size has been widely recognized as instrumental to understand genome evolution (Levasseur and Pontarotti 2011; Wolfe 2006), molecular novelties (Conant and Wolfe 2008; Deng *et al.* 2010) and organismal complexity (Ferguson *et al.* 2014; McLysaght *et al.* 2002; Panopoulou and Poustka 2005). Thus, genome size variation is a complex topic with ongoing debates surrounding its evolutionary origins and impacts. Questions persist regarding whether genome size is a neutral trait or subject to selective pressures, and to what extent these pressures shape evolutionary outcomes. Duplications and deletions of genic regions can have immediate phenotypic effects, while changes in non-coding DNA may have longer-term consequences. Advances in sequencing technologies offer new insights into these mechanisms, but integrating them with traditional evolutionary experiments could provide a comprehensive understanding of genome size evolution and resolve existing debates (Blommaert 2020).

4. Genome Duplication and Hox Genes

Hox genes, a subfamily of homeobox-containing transcription factors, specify cell fate along the anterior-posterior axis of bilaterian animals (Mallo and Alonso 2013). Whole-genome duplication (WGD) is the most widely accepted explanation for the numerical increase in *Hox* gene clusters coincident with the origin of vertebrates and gnathostomes (Amores *et al.* 1998; Holland *et al.* 1994; Pascual-Anaya *et al.* 2013). In fact, the structure and gene content of the amphioxus genome corroborated the existence of two genome-wide duplications and subsequent reorganizations in the vertebrate lineage (Putnam *et al.* 2008). The first round of genome duplication would have predated the Cambrian explosion while the second would have occurred in the early Devonian (2R hypothesis). A fish-specific round of WGD is proposed to have occurred by the late Devonian (Meyer and Schartl 1999). Phylogenetic analyses suggest that tandem duplication of a *protoHox* gene produced a four-gene cluster, which was duplicated producing a four-gene *Hox* cluster, and a four-gene Para*Hox* cluster on a different chromosome (Brook *et al.* 1998). It is argued that these genome duplications were causally associated with quantum jumps in morphological complexity, body design, and adaptive radiations (reviewed in Taylor and Raes 2004). Apparently, vertebrates have undergone significant modifications since the last common ancestor of the chordates. Although some anterior genes are dated back to the ancient divergence between protostomes and deuterostomes, others have been lost from the vertebrate lineage more recently (Butts *et al.* 2010; Furlong and Holland 2002; Zhong and Holland 2011). Indeed, the family of posterior *Hox* genes is claimed to probably originated through independent tandem duplication events at the origin of each of the ambulacrarian, cephalochordate and vertebrate/urochordate lineages (Pascual-Anaya *et al.* 2013).

Phylogenetic analysis of ambulacrarian posterior genes (*Hox* 9 to 13), indicates a lack of correlation and multiple polytomies between this cluster and the posterior genes from cephalochordates and vertebrates (Ferrier *et al.* 2000). This lack of correlation pattern (one-to-one orthology assignments) is referred to as deuterostome posterior flexibility (Ferrier *et al.* 2000; Amemiya *et al.* 2008). Its far from understood causality is claimed to have resulted from dilution of selective constraints (Ferrier *et al.* 2000).

Thus, the genomic organization of bilaterian animals, reflected by a shared set of *Hox* genes is

rather confusing. In some distantly-related species, *Hox* genes are collinearly clustered, but not in others. This suggests that the urbilaterian ancestor had a *Hox* gene set with clustered genomic organization that was subsequently either maintained or lost (Duboule 2007). When discussing the reasons and mechanisms behind *Hox* gene clustering and collinear developmental organization, the phenomenon of allopolyploidy becomes crucial. Indeed, *Hox* gene clusters arranged on different chromosomes serve as a cytogenetic evidence supporting the underlying causal process of polyploidy.

5. Genome Duplication and Synthetic Polyploids

Wild and cultivated allopolyploids are well adapted and stable. Synthetic (man-made) allopolyploids are cytogenetically unstable at the beginning, exhibiting in some cases homeotic transformation (Murai *et al.* 2002; Murai 2013), but eventually leading to the establishment of biological novelties (Chester *et al.* 2012; Comai 2000). Chromosomal rearrangements, changes in chromatin constitution, fluctuations, and distribution in repeats of repetitive DNA accompany the newly synthesized allopolyploids (Liu *et al.* 1998a; Liu *et al.* 1998b). Retrotransposition is activated following polyploidization in several syntethic plants (Parisod *et al.* 2010). Moreover, regulatory abnormalities derive from ploidy changes and/or incompatible interactions between parental genomes (Jones and Pasakinskiene 2005). In this way, it has been suggested that intergenomic incompatibilities play the major role in the generation of a fertile organism (Comai 2000). Epigenetic expression patterns are altered as well as chromatin remodeling, affecting promoter´s response in the new cellular environment (Wendel *et al.* 2016). On the other hand, the impacts of polyploidy on the genomic processes of natural *Arabidopsis* populations are subtle yet far-reaching. These effects encompass reduced purifying selection efficiency, variations in linked selection, and extensive gene flow from diploids. Polyploidy initially conceals harmful mutations, accelerates nucleotide substitution rates, and facilitates interploidy introgression (Monnahan *et al.* 2019).

The importance of hybridizing the paternal species of naturally-occurring polyploids (2n, 4n, 6n) is that their genome dynamics and the phylogenetic pattern already evolved in nature (H_0), could be compared with its homologous synthetic allopolyploid combination, produced and maintained in laboratory conditions (H_1). All parameters of genetic or ecological interest can be accurately studied and any empirical result, comparatively validated. Thus, sequence elimination after polyploidization, genomic differentiation, and diploid-like meiotic behavior of the synthetic counterpart turn into predictive empirical questions that support conceptually transformative hypotheses. In this context, no inferences are made or needed since those predictions reflect the underlying process directly.

Micro and macroevolutionary changes in newly synthesized amphiploids of *Triticum* and *Aegilops* can become fixed in few generations and could give rise to evolutionary novelties (Liu *et al.* 1998a; Liu *et al.* 1998b; Mason and Wendel 2020). Nevertheless, the most synthetic allo- and autopolyploids are meiotically unstable, as evidenced by high frequencies of chromosome rearrangements in young allotetraploid species such as *Tragopogon miscellus* (Mason and Wendel 2020). Extensive karyotype variation has been observed in these species, including clear products of homoeologous recombination between the subgenomes (Chester *et al.* 2012). Additionally, studies comparing individuals and populations of synthetic lines with natural populations of the recently formed allotetraploids *Tragopogon mirus* and *T. miscellus* have detected extensive chromosomal polymorphisms (Lim *et al.* 2008). These included monosomic and trisomic individuals for particular chromosomes, intergenomic translocations, and variable sizes and expression patterns of individual rDNA loci. Chromosomal translocations, gene loss, and meiotic irregularities (i.e., quadrivalents) were detected in both synthetic lines and sibling plants (Lim *et al.* 2008). These patterns point to an explanatory meaning for cytogenetic variation and indicate that chromosomal adjustments, chromatin remodeling and elimination occur rapidly following polyploidization (Wendel *et al.* 2016). The lineage giving rise to the *Arabidopsis* genus has experienced three rounds of genome duplication in the last 250 Ma (De Bodt *et al.* 2005). Its synthetic allotetraploids also exhibit rapid epigenetic changes including gene silencing via heterochromatization and have preferentially retained development genes an others involved in signal transduction pathways (Bomblies and Madlung 2014; Del Pozo and Ramirez-Parra 2015; Shi *et al.* 2015). Thus, it appears that the species´ genetic redundancy is responsible for its rapid diversification (De Bodt *et al.* 2005; Couvreur *et al.* 2010; Schranz *et al.* 2012). In contrast to the 'genome shock' observed in synthetic polyploids, characterized by genome reorganization, altered expression, and transposition, recent research

has revealed that the genome of natural polyploid *Arabidopsis suecica* remains colinear with ancestral genomes. There is no dominance of a subgenome in expression, and transposon dynamics appear stable (Burns *et al.* 2021). This suggests that domesticated polyploids may not always accurately represent natural polyploidization processes.

The formation of allopolyploid wheat has been also accompanied by rapid nonrandom changes in low-copy noncoding, and coding DNA sequences (Liu *et al.* 1998a; Liu *et al.* 1998b; Levy and Feldman 2022). Indeed, newly synthesized amphiploids of different ploidy levels showed disappearance of parental hybridization fragments, and appearance of novel fragments. Pattern variations among individual plants of the same amphiploid level and between several synthetic and natural amphiploids occurred at random (Feldman and Levy 2012). Moreover, intergenomic recombination triggered DNA methylation and modified expression levels that led to meiotic diploidization, gene-dosage compensation and increasing variation among amphidiploid plants (Liu *et al.* 1998a; Liu *et al.* 1998b; Li *et al.* 2021). These evolutionary changes observed during the lifespan of allopolyploids increase intra-specific genetic diversity. Consequently, this enhancement leads to greater fitness and competitiveness (Feldman and Levy 2009). The scientific value of synthetic polyploids allows us to realize that the above-mentioned duplicated genomic patterns and adjustments are derived from interspecific hybridizations, precisely dated and available in the greenhouse. Thus, synthetic polyploids provide empirical tests of enormous predictive capabilities to address the otherwise overlooked transcendental evolutionary role of interlineage hybridization.

6. Small Heat Shock Proteins in Wheat

Bread wheat originated from hybridization involving genera *Triticum* and *Aegilops* to give rise to the allotetraploid emmer wheat (*Triticum turgidum*; AABB). The second hybridization event between emmer wheat and *Aegilops tauschii* (DD), occurred around 0.4 Ma and gave rise to allohexaploid wheat (*Triticum aestivum*; AABBDD). Thus, the present-day genome of wheat is a product of multiple, cyclic rounds of genome duplications (Marcussen *et al.* 2014).

Comparative analysis of small heat shock proteins (sHSPs) in bread wheat has pointed out massive intrachromosomal expansions and expression pattern diversity with polyploidization (Wang *et al.* 2017). The number of sHSPs in tetraploid wheat and in its diploid progenitors was similar, although gene copy number much higher and enriched in specific chromosome fragments of hexaploids. In fact, 25 to 31 sHSP genes were identified in diploid and tetraploid relatives whereas 117 were identified in the bread wheat; many more than the 56 to 70 copies of its tetraploid progenitors. Further genomic comparisons revealed remarkable sHSPs expansion in subgenomes A and B, but not in subgenome D, consistent with its stable gene content after tetraploidization (Wang *et al.* 2017). These findings underscore the significance of hexaploidization, alongside segmental and tandem duplications, in explaining the rise in sHSP numbers. This relationship between polyploidy and intrachromosomal segmental and tandem duplications, which contribute to sHSPs gene expansions, is also evident in *Arabidopsis* (Waters *et al.* 2008), rice (Sarkar *et al.* 2009), and soybean (Lopes-Caitar *et al.* 2013).

A detailed partitioning of chromosome 3B of bread wheat indicated that its 2,216 genes greatly surpass the gene number of homologues in rice and sorghum (Choulet *et al.* 2014). Additionally, 46% of these duplicated genes are tandemly repeated, while 56% are dispersed duplicates, resulting in an intriguingly even split. Additionally, more than twice as many duplicate genes are retained after intrachromosomal duplication relative to other grass species. The finding that 94% of the conserved genes in those grass relatives are also present in chromosome 3B indicated limited gene loss after polyploidization. Indeed, the reduction of the basic chromosome number from 12 to 7 in Triticeae proceeded by the telomeric insertion of one chromosome into a centromeric break of another; a process unaffecting gene content (Luo *et al.* 2009). The 23% of syntenic dispersed duplicates (those located at their ancestral locus) have originated from recent intrachromosomal and interchromosomal duplications at a much higher comparative rate. Interestingly, interchromosomal duplicates were evenly distributed along chromosome 3B whereas the increase of tandem duplications is only telomeric, suggesting the existence of two superimposed mechanisms of gene duplication (Choulet *et al.* 2014). These findings underscore the intricate interplay between polyploidization, gene duplication, and genomic evolution in shaping the genetic landscape of bread wheat.

7. Evolutionary Significance of Polyploidy

While polyploidization represents one of the most dramatic mutations known to occur, it is also a widespread and common phenomenon among eukaryotes, serving as a source for evolutionary innovation and species diversification (Otto and Whitton 2000; Otto 2007; Van de Peer *et al.* 2021). Indeed, the majority of flowering plants and vertebrates have descended from polyploid ancestors. Up to seven rounds of ancestral polyploidy have been suggested in major angiosperm phyla (Jaillon *et al.* 2007), and three have been proposed in the lineage that gave rise to chordates (Holland *et al.* 1994; Pascual-Anaya *et al.* 2013). Ancient polyploidies are widely recognized as events with important roles in the origin of evolutionary novelties in plants and animals, such as the origin of seeds and flowers (Clark and Donoghue 2017; Jiao *et al.* 2011), as well as the emergence of limbs and jaws (Holland 1998; Pascual-Anaya *et al.* 2013). Nevertheless, polyploidy is thought to have had a lesser role in animal evolution (Otto and Whitton 2000). The distinction between the numbers of validated polyploids in plants and animals is indeed substantial. While attributing solely to specific factors may oversimplify this issue, the increased developmental plasticity in plants, the absence of the Weisman barrier, and differences in meiotic processes that prevent rapid solutions to high crossover rates in animals post-WGD could all play crucial roles (Mable *et al.* 2004 and literature therein). Nevertheless, this difference gets blurred under genomic scrutiny. In fact, the conventional assertion that polyploidy is less feasible in animals has been reverted, since insects the most speciose class of invertebrates, has experienced massive polyploidization and extensive genome duplication (Li *et al.* 2018).

In this exposition, we explore genetic and genomic data concerning duplication, investigating the origin and evolutionary patterns of *Hox* genes. We also explore studies of synthetic and natural plant polyploids to glean insights into their evolutionary trajectories. Our survey suggests that polyploidy plays a significant role in generating genetic variability, driving protein evolution, and facilitating the emergence of macroevolutionary diversity. Moreover, contributes to the activation of transposons and the formation of tandem duplicates in diverse organisms. These intricate molecular processes, stemming from gene duplication and polyploidy, challenge conventional evolutionary paradigms and enrich our understanding of macroevolutionary diversity.

Acknowledgments

This perspective is in memory of Dr. Milton H. Gallardo (1947-2019). Several parts of the text remained intact according to the original notes found in Dr. Gallardo's house after his death. EY. Suárez-Villota would particularly like to acknowledge Mónica Barbet-Claus and Paola Solis for allowing him to take some of these notes.

The author also wishes to thank the Editor in Chief, Dr. Andràs Paldi, for encouraging him to complete and revise the article, as well as the anonymous reviewer for their detailed review and suggestions for improving the manuscript.

References

Amemiya, CT, Prohaska, SJ, Hill-Force, A, Cook, A, Wasserscheid, J, *et al.* 2008, "The amphioxus *Hox* cluster: characterization, comparative genomics, and evolution", *J Exp Zool B Mol Dev Evol*, 310, 5, pp. 465-77.

Amores, A, Force, A, Yan, YL, Joly, L, Amemiya, C, *et al.* 1998, "Zebrafish *Hox* clusters and vertebrate genome evolution", *Science*, 282, 5394, pp. 1711-4.

Birchler, JA, & Yang, H 2022, "The multiple fates of gene duplications: deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation", *Plant Cell*, 34, 7, pp. 2466-74.

Blommaert, J 2020, "Genome size evolution: towards new model systems for old questions", *Proc Biol Sci*, 287, 1933, p. 20201441.

Bloom, JD, & Arnold, FH 2009, "In the light of directed evolution: pathways of adaptive protein evolution", *Proc Natl Acad Sci U S A*, 16, 106, pp. 9995-10000.

Bomblies, K, & Madlung, A 2014, "Polyploidy in the *Arabidopsis* genus", *Chromosome Res*, 22, 2, pp. 117-34

Brenner, SE, Hubbard, T, Murzin, A, & Chothia, C 1995, "Gene duplications in *H. influenzae*", *Nature*, 378, 6553, p. 140.

Brooke, N, Garcia-Fernàndez, J, & Holland, P 1998, "The Para*Hox* gene cluster is an evolutionary sister of the *Hox* gene cluster", *Nature*, 392, pp. 920-22.

Buljan, M, & Bateman, A 2009, "The evolution of protein domain families", *Biochem Soc Trans*, 37, Pt 4, pp. 751-5.

Butts, T, Holland, PW, & Ferrier, DE 2010, "Ancient homeobox gene loss and the evolution of chordate brain and pharynx development: deductions from amphioxus gene expression", *Proc Biol Sci*, 277, 1699, pp. 3381-9.

Burns, R, Mandáková, T, Gunis, J, Soto-Jiménez, LM, Liu, C, Lysak, MA, Novikova, PY, & Nordborg, M 2021, "Gradual evolution of allopolyploidy in *Arabidopsis suecica*'. *Nat Ecol Evol*, 5, pp. 1367-81.

Cardona, T 2017, "Photosystem II is a chimera of reaction centers", *Journal of Molecular Evolution*, 84, 2-3, pp. 149-51.

Cardona, T 2018, "Early archean origin of heterodimeric photosystem I", *Heliyon*, 4, 3, p. e00548.

Chester, M, Gallagher, JP, Symonds, VV, Cruz da Silva, AV, Mavrodiev, EV, Leitch, AR, Soltis, PS, & Soltis, DE 2012, "Extensive chromosomal variation in a recently formed natural allopolyploid species, *Tragopogon miscellus* (Asteraceae)", *Proc Natl Acad Sci U S A*, 109, 4, pp. 1176-81.

Chothia, C, Gough, J, Vogel, C, & Teichmann, SA 2003, "Evolution of the protein repertoire", *Science*, 300, 5626, pp. 1701-3.

Choulet, F, Alberti, A, Theil, S, Glover, N, Barbe, V, *et al.* 2014, "Structural and functional partitioning of bread wheat chromosome 3B", *Science*, 345, 6194, p. 1249721.

Clark, JW, & Donoghue, PCJ 2017, "Constraining the timing of whole genome duplication in plant evolutionary history", *Proc Biol Sci*, 284, 1858, p. 20170912.

Comai, L 2000, "Genetic and epigenetic interactions in allopolyploid plants", *Plant Mol Biol*, 43, 2-3, pp. 387-99.

Conant, GC, & Wolfe, KH 2008, "Turning a hobby into a job: how duplicated genes find new functions", *Nat Rev Genet*, 9, p. 938.

Consortium, TIWGS 2014, "A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome", *Science*, 345, 6194, p. 1251788.

Couvreur, TL, Franzke, A, Al-Shehbaz, IA, Bakker, FT, Koch, MA, & Mummenhoff, K 2010, "Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae)", *Mol Biol Evol*, 27, 1, pp. 55-71.

Dharia, AP, Obla, A, Gajdosik, MD, Simon, A, & Nelson, CE 2014, "Tempo and mode of gene duplication in mammalian ribosomal protein evolution", *PLoS One*, 9, 11, p. e111721.

Dayhoff, MO, Eck, RV, Chang, MA, & Sochard, MR 1965, *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Silver Spring (Maryland).

Dayhoff, MO, Schwartz, RM, & Orcutt, BC 1978, "A model of evolutionary change in proteins", in MO Dayhoff, (ed), *Atlas of Protein Sequence and Structure*, pp. 345-52, National Biomedical Research Foundation, Silver Spring (Maryland).

De Bodt, S, Maere, S, & Van de Peer, Y 2005, "Genome duplication and the origin of angiosperms", *Trends Ecol Evol*, 20, 11, pp. 591-7.

Del Pozo, JC, & Ramirez-Parra, E 2015, "Whole genome duplications in plants: an overview from *Arabidopsis*", *J Exp Bot*, 66, 22, pp. 6991-7003.

Deng, C, Cheng, CHC, Ye, H, He, X, & Chen, L 2010, "Evolution of an antifreeze protein by neofunctionalization under escape from adaptive conflict", *Proc Natl Acad Sci U S A*, 107, 50, pp. 21593-21598.

Duboule, D 2007, "The rise and fall of *Hox* gene clusters", *Development*, 134, 14, pp. 2549-60.

Eck, RV, & Dayhoff, MO 1966, "Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences", *Science*, 152, 3720, p. 363.

Feldman, M, & Levy, AA 2009, "Genome evolution in allopolyploid wheat – A revolutionary reprogramming followed by gradual changes", *Journal of Genetics and Genomics*, 36, 9, pp. 511-18.

Feldman, M, & Levy, AA 2012, "Genome evolution due to allopolyploidization in wheat", *Genetics*, 192, 3, pp. 763-74.

Ferguson, L, Marletaz, F, Carter, JM, Taylor, WR, Gibbs, M, Breuker, CJ, & Holland, PW 2014, "Ancient expansion of the *Hox* cluster in lepidoptera generated four homeobox genes implicated in extra-embryonic tissue formation", *PLoS Genetics*, 10, 10, p. e1004698.

Ferrier, DEK, Minguillón, C, Holland, PWH, & Garcia-Fernàndez, J 2000, "The amphioxus *Hox* cluster: deuterostome posterior flexibility and *Hox14*", *Evolution & Development*, 2, 5, pp. 284-93.

Fox, DT, & Duronio, RJ 2013, "Endoreplication and polyploidy: insights into development and disease", *Development*, 140, 1, pp. 3-12.

Fox, GE 2010, "Origin and evolution of the ribosome", *Cold Spring Harb Perspect Biol*, 2, 9, p. a003483.

Furlong, RF, & Holland, PWH 2002, "Were vertebrates octoploid?", *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357, 1420, pp. 531-44.

Gallardo, MH, Bickham, JW, Kausel, G, Kohler, N, & Honeycutt, RL 2003, "Gradual and quantum genome size shifts in the hystricognath rodents", *J Evol Biol*, 16, 1, pp. 163-9.

Gregory, TR 2024, "Animal genome size database", [online]. Available from: http://www.genomesize.com [Accessed 23 February 2024].

Gregory, TR, & Hebert, PD 1999, "The modulation of DNA content: proximate causes and ultimate consequences", *Genome Res*, 9, 4, pp. 317-24.

Holland, PWH 1998, "Major transitions in animal evolution: a developmental genetic perspective", *Integrative and Comparative Biology*, 38, 6, pp. 829-42.

Holland, PWH, Garcia-Fernàndez, J, Williams, NA, & Sidow, A 1994, "Gene duplications and the origins of vertebrate development", *Development*, 1994, Supplement, p. 125.

Hunt, LT, Barker, WC, & Dayhoff, MO 1974, "Epidermal growth factor: internal duplication and probable relationship to pancreatic secretory trypsin inhibitor", *Biochem Biophys Res Commun*, 60, 3, pp. 1020-28.

Hunt, LT, & Dayhoff, MO 1970, "The occurrence in proteins of the tripeptides Asn-X-Ser and Asn-X-Thr and of bound carbohydrate", *Biochem Biophys Res Commun*, 39, 4, pp. 757-765.

Jaillon, O, Aury, J-M, Noel, B, Policriti, A, Clepet, C, *et al.* 2007, "The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla", *Nature*, 449, p. 463.

Jayaraman, V, Toledo-Patiño, S, Noda-García, L, & Laurino, P 2022, "Mechanisms of protein evolution", *Protein Sci*, 31, 7, p. e4362.

Jiao, Y, Wickett, NJ, Ayyampalayam, S, Chanderbali, AS, Landherr, L, *et al.* 2011, "Ancestral polyploidy in seed plants and angiosperms", *Nature*, 473, 7345, pp. 97-100.

Jones, N, & Pasakinskiene, I 2005, "Genome conflict in the gramineae", *The New phytologist*, 165, 2, pp. 391-409.

Kuzmin, E, Vandersluis, B, Nguyen, AN, Wang W, Koch, EN, *et al.* 2020, "Exploring whole-genome duplicate gene retention with complex genetic interaction analysis", *Science*, 368, 6498, p. eaaz5667.

Lee, J, Wu, S, & Zhang, Y 2009, "*Ab initio* protein structure prediction", in DJ Rigden, (ed), *From Protein Structure to Function with Bioinformatics*, pp. 3-25, Springer Netherlands, Dordrecht.

Levasseur, A, & Pontarotti, P 2011, "The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics", *Biol Direct*, 6, p. 11.

Levy, AA, & Feldman, M 2022, "Evolution and origin of bread wheat", *Plant Cell*, 34, 7, pp. 2549-67.

Levy, ED, Boeri Erba, E, Robinson, CV, & Teichmann, SA 2008, "Assembly reflects evolution of protein complexes", *Nature*, 453, 7199, pp. 1262-5.

Li, Z, McKibben, MTW, Finch, GS, Blischak, PD, Sutherland, BL, & Barker, MS 2021, "Patterns and processes of diploidization in land plants", *Annu Rev Plant Biol*, 72, pp. 387-410.

Li, Z, Tiley, GP, Galuska, SR, Reardon, CR, Kidder, TI, Rundell, RJ, & Barker, MS 2018, "Multiple large-scale gene and genome duplications during the evolution of hexapods", *Proc Natl Acad Sci U S A*, 115, 18, pp. 4713-18.

Liebl, U, Mockensturm-Wilson, M, Trost, JT, Brune, DC, Blankenship, RE, & Vermaas, W 1993, "Single core polypeptide in the reaction center of the photosynthetic bacterium *Heliobacillus mobilis*: structural implications and relations to other photosystems", *Proc Natl Acad Sci U S A*, 90, 15, pp. 7124-8.

Lim, KY, Soltis, DE, Soltis, PS, Tate, J, Matyasek, R, Srubarova, H, Kovarik, A, Pires, JC, Xiong, Z, & Leitch, AR 2008, "Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (Asteraceae)", *PLoS One*, 3, 10, p. e3353.

Liu, B, Vega, JM, & Feldman, M 1998a, "Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. II. Changes in low-copy coding DNA sequences", *Genome*, 41, 4, pp. 535-42.

Liu, B, Vega, JM, Segal, G, Abbo, S, Rodova, M, & Feldman, M 1998b, "Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. I. Changes in low-copy noncoding DNA sequences", *Genome*, 41, 2, pp. 272-77.

Lopes-Caitar, VS, De Carvalho, MC, Darben, LM, *et al.* 2013, "Genome-wide analysis of the *Hsp20* gene family in soybean: comprehensive sequence, genomic organization and expression profile analysis under abiotic and biotic stresses", *Bmc Genomics*, 14, p. 577.

Luo, MC, Deal, KR, Akhunov, ED, Akhunova, AR, Anderson, OD, *et al.* 2009, "Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in Triticeae", *Proc Natl Acad Sci U S A*, 106, 37, pp. 15780-5.

Mable, BK 2004, "Why polyploidy is rarer in animals than in plants: myths and mechanisms", *Biol J Linn Soc Lond*, 82, 4, pp. 453-66.

Mallo, M, & Alonso, C.R 2013, "The regulation of *Hox* gene expression during animal development", *Development*, 140, 19, p. 3951.

Marcussen, T, Sandve, SR, Heier, L, Spannagl, M, Pfeifer, M, Jakobsen, KS, Wulff, BBH, Steuernagel, B, Mayer, KFX, & Olsen, OA 2014, "Ancient hybridizations among the ancestral genomes of bread wheat", *Science*, 345, 6194, p. 1250092.

Markov, AV, & Kaznacheev, IS 2016, "Evolutionary consequences of polyploidy in prokaryotes and the origin of mitosis and meiosis", *Biol Direct*, 11, p. 28.

Marsh, JA, & Teichmann, SA 2010, "How do proteins gain new domains?", *Genome biology*, 11, 7, pp. 126-26.

Mason, AS, & Wendel, JF 2020, "Homoeologous exchanges, segmental allopolyploidy, and polyploid genome evolution", *Front Genet*, 11, p. 1014.

Melnikov, S, Ben-Shem, A, Garreau de Loubresse, N, Jenner, L, Yusupova, G, & Yusupov, M 2012, "One core, two shells: bacterial and eukaryotic ribosomes", *Nat Struct Mol Biol*, 19, p. 560.

McLysaght, A, Hokamp, K, & Wolfe, KH 2002, "Extensive genomic duplication during early chordate evolution", *Nature Genetics*, 31, p. 200.

Meyer, A, & Schartl, M 1999, "Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions", *Curr Opin Cell Biol*, 11, 6, pp. 699-704.

Monnahan, P, Kolář, F, Baduel, P, Sailer, C, Koch, J, *et al.* 2019, "Pervasive population genomic consequences of genome duplication in *Arabidopsis arenosa*", *Nat Ecol Evol*, 3, pp. 457-68.

Murai, K 2013, "Homeotic genes and the ABCDE model for floral organ formation in wheat", *Plants (Basel)*, 2, 3, pp. 379-95.

Murai, K, Takumi, S, Koga, H, Ogihara, Y 2002, "Pistillody, homeotic transformation of stamens into pistil-like structures, caused by nuclear-cytoplasm interaction in wheat", *Plant J*, 29, 2, pp. 169-81.

Ohno, S, Wolf, U, & Atkin, NB 1968, "Evolution from fish to mammals by gene duplication", *Hereditas*, 59, 1, pp. 169-87.

Oliver, T, Sánchez-Baracaldo, P, Larkumx, AW, Rutherford, AW, & Cardona, T 2021, "Time-resolved comparative molecular evolution of oxygenic photosynthesis", *Biochim Biophys Acta Bioenerg*, 1862, 6, p. 148400.

Otto, SP 2007, "The evolutionary consequences of polyploidy", *Cell*, 131, 3, pp. 452-62.

Otto, SP, & Whitton, J 2000, "Polyploid incidence and evolution", *Annu Rev Genet*, 34, pp. 401-37.

Panopoulou, G, & Poustka, AJ 2005, "Timing and mechanism of ancient vertebrate genome duplications – the adventure of a hypothesis", *Trends Genet*, 21, 10, pp. 559-67.

Parisod, C, Alix, K, Just, J, Petit, M, Sarilar, V, Mhiri, C, Ainouche, M, Chalhoub, B, & Grandbastien, MA 2010, "Impact of transposable elements on the organization

and function of allopolyploid genomes", *New Phytol*, 186, 1, pp. 37-45.

Pascual-Anaya, J, D'Aniello, S, Kuratani, S, & Garcia-Fernandez, J 2013, "Evolution of *Hox* gene clusters in deuterostomes", *BMC Dev Biol*, 13, p. 26.

Pellicer, J, Hidalgo, O, Dodsworth, S, & Leitch, IJ 2018, "Genome size diversity and its impact on the evolution of land plants", *Genes (Basel)*, 9, 2, p. 88.

Petrov, AS, Gulen, B, Norris, AM, Kovacs, NA, Bernier, CR, *et al.* 2015, "History of the ribosome and the origin of translation", *Proc Natl Acad Sci U S A*, 112, 50, pp. 15396-401.

Pich, IRO, & Kondrashov, FA 2014, "Long-term asymmetrical acceleration of protein evolution after gene duplication", *Genome Biol Evol*, 6, 8, pp. 1949-55.

Putnam, NH, Butts, T, Ferrier, DE, Furlong, RF, Hellsten, U, *et al.* 2008, "The amphioxus genome and the evolution of the chordate karyotype", *Nature*, 453, 7198, pp. 1064-71.

Romero, ML, Rabin, A, & Tawfik, DS 2016, "Functional proteins from short peptides: Dayhoff's hypothesis turns 50", *Angewandte Chemie. International Ed. In English*, 55, 52, pp. 15966-71.

Rouet, R, Langley, DB, Schofield, P, Christie, M, Roome, B, Porebski, BT, Buckle, AM, Clifton, BE, Jackson, CJ, Stock, D, & Christ, D 2017, "Structural reconstruction of protein ancestry", *Proc Natl Acad Sci U S A*, 114, 15, p. 3897.

Sarkar, NK, Kim, YK, & Grover, A 2009, "Rice *sHsp* genes: genomic organization and expression profiling under stress and development", *Bmc Genomics*, 10, p. 393.

Schaper, E, & Anisimova, M 2015, "The evolution and function of protein tandem repeats in plants", *New Phytol*, 206, 1, pp. 397-410.

Schranz, ME, Mohammadin, S, & Edger, PP 2012, "Ancient whole genome duplications, novelty and diversification: the WGD radiation lag-time model", *Curr Opin Plant Biol*, 15, 2, pp. 147-53.

Schwartz, RM, & Dayhoff, MO 1978, "Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts", *Science*, 199, 4327, p. 395.

Shackelton, LA, & Holmes, EC 2004, "The evolution of large DNA viruses: combining genomic information of viruses and their hosts", *Trends Microbiol*, 12, 10, pp. 458-65.

Shi, X, Zhang, C, Ko, DK, & Chen, ZJ 2015, "Genome-wide dosage-dependent and -independent regulation contributes to gene expression and evolutionary novelty in plant polyploids", *Mol Biol Evol*, 32, 9, pp. 2351-2366.

Simon-Loriere, E, & Holmes, EC 2013, "Gene duplication is infrequent in the recent evolutionary history of RNA viruses", *Mol Biol Evol*, 30, 6, pp. 1263-69.

Smith, TF, Lee, JC, Gutell, RR, & Hartman, H 2008, "The origin and evolution of the ribosome", *Biol Direct*, 3, p. 16.

Soo, RM, Hemp, J, & Hugenholtz, P 2019, "Evolution of photosynthesis and aerobic respiration in the cyanobacteria", *Free Radic Biol Med*, 140, pp. 200-5.

Söding, J, & Lupas, AN 2003, "More than the sum of their parts: on the evolution of proteins from peptides", *BioEssays*, 25, 9, pp. 837-46.

Sousa, FL, Shavit-Grievink, L, Allen, JF, & Martin, WF 2013, "Chlorophyll biosynthesis gene evolution indicates photosystem gene duplication, not photosystem merger, at the origin of oxygenic photosynthesis", *Genome Biol Evol*, 5, 1, pp. 200-16.

Taylor, JS, & Raes, J 2004, "Duplication and divergence: the evolution of new genes and old ideas", *Annu Rev Genet*, 38, 615-43.

Tocchini-Valentini, GD, Fruscoloni, P, & Tocchini-Valentini, GP 2005, "Structure, function, and evolution of the tRNA endonucleases of Archaea: an example of subfunctionalization", *Proc Natl Acad Sci U S A*, 102, 25, pp. 8933-8.

Van de Peer, Y, Ashman, TL, Soltis, PS, & Soltis, DE 2021, "Polyploidy: an evolutionary and ecological force in stressful times", *Plant Cell*, 33, 1, pp. 11-26.

Van de Peer, Y, Mizrachi, E, & Marchal, K 2017, "The evolutionary significance of polyploidy", *Nat Rev Genet*, 18, 7, pp. 411-24.

Wang, X, Wang, R, Ma, C, Shi, X, Liu, Z, Wang, Z, Sun, Q, Cao, J, & Xu, S 2017, "Massive expansion and differential evolution of small heat shock proteins with wheat (*Triticum aestivum* L.) polyploidization", *Sci Rep*, 7, 1, p. 2581.

Waters, ER, Nguyen, SL, Eskandar, R, Behan, J, & Sanders-Reed, Z 2008, "The recent evolution of a pseudogene: diversity and divergence of a mitochondria-localized small heat shock protein in *Arabidopsis thaliana*", *Genome*, 51, 3, pp. 177-86.

Wendel, JF, Jackson, SA, Meyers, BC, & Wing, RA 2016, "Evolution of plant genome architecture", *Genome Biol*, 17, 37.

Willemsen, A, Zwart, MP, Higueras, P, Sardanyés, J, & Elena, SF 2016, "Predicting the stability of homologous gene duplications in a plant RNA virus", *Genome Biol Evol*, 8, 9, pp. 3065-82.

Wolfe, KH 2006, "Comparative genomics and genome evolution in yeasts", *Philos Trans R Soc Lond B Biol Sci*, 361, 1467, pp. 403-12.

Yusupova, G, & Yusupov, M 2017, "Crystal structure of eukaryotic ribosome and its complexes with inhibitors", *Philos Trans R Soc Lond B Biol Sci*, 372, 1716, p. 20160184.

Zhong, YF, & Holland, PW 2011, "The dynamics of vertebrate homeobox gene evolution: gain and loss of genes in mouse and human lineages", *BMC Evol Biol*, 11, 169.

Perspective & Hypotheses

Vol. 7, No. 1 (2024)
ISSN: 2532-5876

DOI: 10.13133/2532-5876/18136

Cancer, Cytologism and the Kinase Inhibitors Saga

Andras Paldi[a] *& Carlos Sonnenschein*[b*]

[a] *Ecole Pratique des Hautes Etudes PSL Research University Paris France*
[b] *Tufts University School of Medicine, Boston, MA USA*

***Corresponding author:** Carlos Sonnenschein, Email: carlos.sonnenschein@tufts.edu

Abstract

The role of kinases in the cell cycle of unicellular eukaryotes and cells of multicellular organisms has been the object of numerous studies involving normal and cancer cells. These studies described in detail how two daughter cells are ge-nerated from a single normal or cancer mother cell. Among the thousands of participants in the cell cycle, kinases play a crucial role in the dynamic aspects of the cell cycle thanks to the phosphorylation of substrates with which they react. Inhibitors of these kinases have figured prominently among the strategies to treat cancers. However, evidence shows that the benefits that cancer patients accrued from this therapeutic approach have been of a limited degree. In this arti-cle, we review the rationale for adopting such a strategy and the factors that contribute to its shortcomings.

Keywords: cancer, somatic mutation theory, tissue organization field theory, tyrosine kinases, kinase inhibitors, cell cycle, carcinogenesis

Citation: Paldi, A, & Sonnenschein, C 2024, "Cancer, Cytologism and the Kinase Inhibitors Saga", *Organisms: Journal of Biological Sciences*, vol. 7, no. 1, pp. 23-28. DOI: https://doi.org/10.13133/2532-5876/18136

1. Introduction

The proliferation of somatic cells in multicellular organisms is accomplished through a rather strictly regulated process called the cell cycle (Alberts *et al.* 2014; Weinberg 2014). The completion of the cell cycle in somatic cells of multicellular organisms takes a variable amount of time during which those cells traverse four stages arbitrarily called G0/G1 (G stands for gap), S (S stands for DNA synthesis), G2 and M (M stands for mitosis). Consistently, at the end of the cell cycles a "mother" cell generates two similar but not identical "daughter" cells. These stages occur regardless of whether the cells are normal or neoplastic (Nurse 1990).

The description of the myriad of molecular/biochemical interactions taking place during the cell cycle in cells from multicellular organisms has clarified to a great extend *how* cells accomplish their reproductive function. Those steps are not much different from those happening in unicellular eukaryotes, like yeast (Rew and Wilson 2000; Alberts *et al.* 2014). In fact, the characterization of those steps in yeast have enriched the detailed roles played by the cell cycle components in cells of multicellular organisms regardless, again, of whether those cells were of normal or cancer origin. Intriguingly, however, textbooks of both normal and cancer cell biology, as well as research papers in these areas have claimed for several decades that the signaling pathways happening during the cell cycle in cancer cells are qualitatively altered when compared with those in normal cells (Hunter 1998; Blume-Jensen and Hunter 2001). More specifically, under the notion that there are qualitative differences between the cell cycles of normal somatic cells and their cancerous counterparts, it has been widely reported that cyclin-dependent kinase (CDK) dysregulation, directly or indirectly, plays

an essential role in carcinogenesis (Malumbres and Barbacid 2009). In addition, this notion that cyclin-dependent kinase (CDK) dysregulation underlie other diseases has not been restricted to the carcinogenesis area alone: comparable views regarding altered intracellular signaling processes involving kinases have been extended to virtually all major diseases, such as immunological, inflammatory, degenerative, metabolic, cardiovascular, infectious diseases, epilepsis, and even mental retardation (Ferguson and Gray 2018; Maury *et al.* 2024).

All along, it has been reported that there are anywhere between 518 and over a thousand kinases encoded in the human genome that are responsible for the phosphorylation of a third of its proteome. As a result, the ubiquity of kinases makes testing the specific role of each of those mutated enzymes singly or in combination a challenging task. However, empirical evidence indicates that, either singly or in combination, mutated kinases do not deleteriously influence cell cycle steps to the extent that can be empirically verified downstream through altered cell counts when compared with non-mutated cells (Rew and Wilson 2000).

In addition, although it is seldom mentioned explicitly, the rationale behind aggressively studying the cell cycle of somatic cells and the role of kinases in it relates to the two assumptions on which the somatic mutation theory of carcinogenesis (SMT) is based, namely, 1) cancer is a cell-based disease, and 2) carcinogenesis is due to an accumulation of somatic mutations in a multitude of genes, included those involved in intracellular signaling, in an initially normal somatic cell that eventually due to the intracellular disruptions caused by those mutations will generate a neoplasm. Due to multiple incongruencies (Soto and Sonnenschein 2020; Sonnenschein and Soto 1999), the SMT has been the object of multiple course-corrections (see below). As with other intracellular molecular targets (genes, transcriptional and translational components) (De Magalhães 2022), and structural organelles (mitochondria, nucleolus, chromosomes), kinases have also been singled out as prominent targets of carcinogens and therefore, as a result of this assumed interaction, a consensus in this field adopted the notion that they are responsible for the unwieldy behavior of cancer cells. Based on this inference, it was concluded that kinase inhibitors (KIs) represented promising therapeutic agents for cancer patients (Suski *et al.* 2021). It is noteworthy, however, that first, normal cells can proliferate as fast or even faster than cancer cells; examples of rapidly proliferating normal cells are those following egg fertilization, cells in the epithelium of the intestinal tract, and hematopoietic cells. Second, mutated cell cycle kinases do not show distinctive proliferative phenotypes. And third, the original argument proposing that KIs exert their therapeutic effects by targeting mutated kinase-coding genes (BCL/etc translocation) has been subject to criticism based on the acknowledged argument that therapeutic drugs (for cancer and other diseases) have pleiotropic effects, a feature that prevents assigning accurate causation to these drugs. Separately, statistical analysis of the effects of cancer treatment in the last decades suggest that aggressive efforts in this direction have not significantly affected the overall survival of cancer patients (Unni and Arteaga 2019; Settleman *et al.* 2018; Carlisle *et al.* 2020; Tiwari *et al.* 2024).

2. KIs in Cancer Therapy

For several decades now, based on the previously mentioned idea that there are qualitative differences in the signaling pathways utilizing kinases in general and more specifically those of the cell cycle of normal and cancer cells, most cancer researchers agreed with the notion that KIs should occupy a prominent role in the strategy to effectively treat the disease. To develop such a therapeutic strategy, researchers concentrated on two main areas: a) one aimed at strictly defining the biochemical and biophysical properties of those enzymes (Hunter 1998; Blume-Jensen and Hunter 2001; Mortuza *et al.* 2018), and b) another one aimed at examining the roles of kinases in functional cellular events which affect the dynamics of the cell cycle and how to deal with alleged kinase malfunctions (Suski *et al.* 2021; Besson *et al.* 2008).

Finding small size KIs to use as therapeutic tools has been intensively pursued for a period long enough to allow for a fair evaluation of the outcome of this strategy (Prasad 2020). In fact, many compounds initially reported to be therapeutically effective were subsequently shown to lack potency, selectivity and/or be toxic (Goel *et al.* 2018; Jiang *et al.* 2023; Tiwari *et al.* 2024).

The lack of significant benefits from this therapeutic approach invites the proposal of alternative plausible explanations to either refine, if possible, the unproductive strategy, or else to abandon it altogether if proven ineffective or damaging to the patient's wellbeing. To further explore the subject, we focused our attention on i) the epistemology of carcinogenesis and ii) the rationale of designing the therapeutical approaches aimed at effectively "curing" or, at least, arresting the progress of this disease.

Table 1: Control of cell proliferation in the context of theories of carcinogenesis.

	The somatic mutation theory (SMT)	***The tissue organization field theory (TOFT)***
Implicit premise	Default state: *quiescence*	----------
Explicit premises		Default state: *proliferation*
	Neoplasms due to mutations in cell cycle and cell proliferation regulatory genes	
	Proliferation stimulated by exogenous growth factors	Proliferation controlled by exogenous and endogenous inhibitory factors
	Cell cycle is affected by oncogenes, suppressor genes, cyclins, inhibitory factors	
	Carcinogenesis occurs at the cellular level of biological organization	Carcinogenesis occurs at the tissue level of biological organization (tissue-tissue interactions)
	Control of cell proliferation and control of the cell cycle are often conflated	
	Neoplasms are monoclonal	
Corollary	Cancer is irreversible	Cancer is reversible

3. Cancer theories

Currently, there are two main theories of carcinogenesis. They are: i) the still hegemonic cell-based SMT proposed by Theodor Boveri in 1914 and ii) the Tissue Organization Field Theory (TOFT) proposed in 1999 (Sonnenschein and Soto 1999). Of note, the TOFT differs from the SMT in two fundamental criteria; first, while the SMT assumes that the default state of cells in multicellular organisms is *quiescence*, the TOFT explicitly posits instead that *proliferation* is the default state of *all* cells (Shomar *et al.* 2022). And second, while the SMT considers cancer a *cell-based* disease, the TOFT considers it as a *tissue-based* one (Table 1). Empirical evidence generated over the years when applying the strategy promoted by the SMT (i.e., cell killing, inhibition of cell proliferation) encountered multiple examples of lack of fit. When addressing these inconsistencies, researchers siding with the SMT incorporated course corrections to this theory's original version (Sonnenschein and Soto 2020). Among them, the microenvironment surrounding the original "normal" cell was added as a supplemental target that also accumulated somatic mutations or affected cancer epithelial cells through epigenetic modifications. This *ad hoc* course correction represent a "compromise" involving the original SMT plus the role played in this instance by the stroma that surrounds the primary epithelial tumor cells; this alternative was already proposed in the 1930s by J. Needham and by C. Waddington. Altogether, despite the incorporation of this and other theoretical manipulations as add-ons to the SMT, these "compromises" retain the assumption of a causal carcinogenic role for the somatic mutations accumulated in normal epithelial cells that eventually may become neoplastic (while generating mostly carcinomas). Essentially, regardless of which of these *ad hoc* modifications is adopted, the consensus that cancer is a *cell-based*, genetic, molecular disease remains unaltered. Interestingly, these conclusions are still considered meritorious by most cancer researchers even when it has been reproducibly shown that clones of normal cells present in several organs that will not generate cancers carry alleged cancer-causing "driver" mutations and that "cancer cells" that are part of a neoplasia do not carry those same "driver" mutations (Martincorena and Campbell 2015; Dou *et al.* 2018; Martincorena *et al.* 2018; Kakiuchi and Ogawa 2021). The incorporation of novel powerful and less expensive sequencing technologies resulting from the significant contributions of the Molecular Biology Revolution has contributed to unexpectedly clarify that the genome of normal cells carried comparable cancer "driver" gene mutations to those thought to be unique to cancer cells (Martincorena and Campbell 2015; Dou *et al.* 2018; Martincorena *et al.* 2018; Kakiuchi and Ogawa 2021). If anything, as pointed out above, the data now collected through deep sequence probes suggest instead that those alleged cancer-causing "driver" mutations are also present in cells which are considered normal (meaning non-cancer cells) (Naxerova 2021; Colom *et al.* 2021). This new development justifies K. Naxerova's pondering: "These new insights invite us to reconsider how we genetically define cancer. If having multiple driver mutations does not make a cancer, what does?".

4. Cell Proliferation and Effective KI Activity in Cancer Therapy

Toward the end of last century, aberrant tyrosine phosphorylation was being considered as an important hallmark in cancer initiation (Hunter 1998). The expected effectiveness of KIs was predicated on a series of inferences that supported the rationale that these drugs would have selectively slowed down the speed of the cell cycle of cancer cells. Those expectations have not been fulfilled as anticipated because, among other reasons, a lack of specificity. Separately, after two decades of insisting that the core of the carcinogenic event can be attributed to a dysregulation of the cell cycle of cancer cells due to "either overexpression of cyclin D1, loss of p16Ink4a, the mutation of CDK4 to an Ink4-refractory state, or the loss of Rb itself", no plausible alternatives in the form of novel KIs are being offered by academic research or BigPharma labs. Notwithstanding, all along, the notion that cell cycle signaling defects in cancer cells are central to the difference between normal and cancer cells is still being promoted (Classon and Harlow 2002; Klein *et al.* 2018; Prasad 2016; Naxerova 2021; Colom *et al.* 2021).

During a standard human lifetime, it is estimated that the average person will undergo about 10^{14} cell cycles in an uneven timescale. That is, some cells in some tissues proliferate during embryonal, fetal, and early childhood and then enter a period during which proliferation is rather minimal or totally absent (neurons, fibroblasts, etc.; meanwhile, cells in other tissues proliferate incessantly (bone marrow, intestinal system, skin). Simultaneously, epithelial cells in other organs of metazoans proliferate at different speeds while moving (streaming) and expressing their "differentiated" functions. So called "differentiated" cells are continuously subject to changes in their respective local morphogenetic units and other changes imposed on them by extemporaneous homeostatic conditions, e.g. – adult stem cell trans-differentiation. In other words, under physiological conditions, cells performing "specialized" functions (such as hepatocytes secreting albumin, small intestine epithelial cells absorbing nutrients, epithelial cells in glandular organs secreting milk, saliva, enzymes, etc.) nonetheless do continue to proliferate and move unperturbed. This uncontested feature implies that there is no obligatory linkage between the ability of cells to proliferate and move on the one hand, and the ability of those same cells to concurrently synthesize and/or secrete a variety of cell products (collagen, albumin, sex hormones, etc.), on the other (Sonnenschein and Soto 2021).

As summarized above, kinases have been claimed to be causally involved in the carcinogenic process. Based on this premise emerged the notion of small molecular KIs as a potentially powerful class of effective drugs in cancer therapy (Ferguson and Gray 2018; Zhou *et al.* 2016). How can the discrepancy between promising pre-clinical effects and the lack of equivalent results in clinical tests be explained? In addition to the lack of specificity argument allude to above, several possible explanations were proposed by the defenders of the "kinase inhibitor" therapeutic strategy. Recently, pioneers in this field conceded that dozens of published articles on a leucine zipper-containing serine/threonine kinase called MELK lacked credibility (Settleman *et al.* 2018). It was also claimed that MELK is activated during the cell cycle and is important for maintaining proper asymmetric division of stem cells (Ganguly *et al.* 2015). As a result of this inference, kinases were then considered worthy, potential therapeutic targets in human cancers. However, researchers recently claimed that the experimental criteria used to validate candidate cancer therapeutic targets was subject to serious methodological faults (Settleman *et al.* 2018). The main factors responsible for the credibility gap were considered technical, namely, the use of cancer cells in culture conditions or the use of RNA interference for target validation. Notwithstanding, it is equally plausible that additional technical factors contribute to the failure to validate the candidate kinase inhibitors. Off-target deleterious effects of TKs could be considered as valid explanations for the therapeutic failures (Gyawali *et al.* 2021).

At first glance, off-target effects and poor selectivity may appear as an issue of poor inhibitor design and/or unanticipated pleiotropy of action. For example, while the expression level of the above-mentioned MELK has been strongly correlated with the mitotic activity in human cancers and remains one of the main predictors of the patient mortality in a variety of tumors, cancer cells with a loss-of-function MELK mutation still proliferate at wild-type levels. In addition, the known targets of MELK also become phosphorylated in these cells. This and similar examples of the lack of biological specificity of alleged chemically specific enzyme inhibitors illustrate the difficulty of achieving a selective effect by targeting redundantly acting enzymes. Beyond the practical implications, functional redundancy casts doubt on the possibility that a single mutation in a single kinase gene by itself may induce cancer. The initial

enthusiasm generated by biochemists and molecular biologists for the specificity of KIs obscured the concept of redundancy and pleiotropy omnipresent in living organisms.

Inhibitors of essential kinases represent another example of the difficulty of achieving biologically selective effects using high-potency enzyme inhibitors. This difficulty is conceptual and not just technical. The development of mitotic kinase inhibitors was based in part on the idea that targeting only cycling cells will minimize the toxicity on post-mitotic cells. Unfortunately, these inhibitors lack an effective therapeutic window because of the high toxicity observed on non-neoplastic tissues with a high cellular proliferation rate exposed to KIs (Zhou *et al.* 2016). As explained above, it is a common misconception that cancer cells proliferate more rapidly than cells in normal tissues. Also, as mentioned above, several normal tissues have higher proliferation rates than most tumors. Proliferating cells in these normal tissues are also targeted and affected by KIs.

Another example illustrating the confusion between conceptual and technical difficulties is provided by the inhibitors of PI3K kinases (Vasan and Cantley 2022). PI3Ks are involved in a wide variety of pathways linked to cell growth, proliferation and differentiation through their role in the regulation of metabolic and insulin pathways. This class of kinases were and still are considered as potential candidates for therapy. Several drugs targeting the PI3K pathway have received approval. However, as in all other cases with KIs, achieving a therapeutic window that maximizes efficacy and minimizes adverse effects has proven to be a major barrier to an effective therapeutic use of Pl3K inhibitors (Prasad 2020; Tiwari *et al.* 2024).

5. Conclusions

The example of kinase inhibitors highlights how the systematic use of *ad-hoc* explanations to account for unexpected results may hide important conceptual problems that eventually canalize the research into dead-ends. The failure of kinase inhibitors as effective anti-cancer drugs challenges first, the decades-old assumption that cancer is a *cell-based* disease as suggested by the SMT. And second and of comparable, if not greater importance, the failure of those therapeutic approaches aimed at correcting those proposed, but yet-to-be rigorously documented cell cycle signaling defects, obscures productive avenues aimed at both preventing carcinogenesis and to offer effective therapeutic options based on alternative theoretical approaches. These shortcomings were already noticed over 60 years ago by David Smithers who, based on rigorous clinical data he collected, offered an organicist-based alternative to the cytologism that then began to dominate experimental and clinical cancer research (Soto and Sonnenschein 2020). Additional evidence accumulated since then -4point to the need to switch attention to theoretical and empirical alternatives that, as the TOFT proposes, are based on solid evolutionary-based premises, such as those related to the default state of cells and the merits of considering cancer as a *tissue-based* disease.

The rationale of remaining loyal to a thoroughly mistaken theory and to the diagnostic, prognostic and therapeutic mishaps that have followed as a result, does not benefit the wellbeing of cancer patients or the prestige of the scientific enterprise. Abandoning the SMT and its wrongheaded implications over cancer diagnoses and therapies is long overdue (Sonnenschein and Soto 2000).

Acknowledgments

We thank Ana M. Soto and Cheryl Schaeberle for their critical reading of the manuscript. This work was funded by the National Institute of Environmental Health Sciences United States grants ES030045. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

Alberts, B, Johnson, A, Lewis, J, Morgan, D, Raff, M, Roberts, K, & Walter, P 2014, *Molecular Biology of the Cell*, Garland Science, New York.

Besson, A, Dowdy, SF, & Roberts, JM 2008, "CDK inhibitors: cell cycle regulators and beyond", *Dev Cell*, 14, pp. 159-69.

Blume-Jensen, P, & Hunter, T 2001, "Oncogenic kinase signalling", *Nature*, 411, pp. 355-365.

Carlisle, BG, Zheng, T, & Kimmelman, J 2020, "Imatinib and the long tail of targeted drug development", *Nat Rev Clin Oncol*, 17, pp. 1-3.

Classon, M, & Harlow, E 2002, "The retinoblastoma tumour suppressor in development and cancer", *Nat Rev Cancer*, 2, pp. 910-917.

Colom, B, Herms, A, Hall, MWJ, Dentro, SC, King, C, Sood, RK, Alcolea, MP, Piedrafita, G, Fernandez-Antoran, D, Ong, SH, Fowler, JC, Mahbubani, KT, Saeb-Parsy, K, Gerstung, M, Hall, BA, & Jones, PH 2021, "Mutant clones in normal epithelium outcompete and eliminate emerging tumours", *Nature*, 598, pp. 510-514.

De Magalhães, JP 2022, "Every gene can (and possibly will) be associated with cancer", *Trends Genet*, 38, pp. 216-217.

Dou, Y, Gold, HD, Luquette, LJ, & Park, PJ 2018, "Detecting Somatic Mutations in Normal Cells", *Trends Genet*, 34, pp. 545-557.

Ferguson, FM, & Gray, NS 2018, "Kinase inhibitors: the road ahead", *Nat Rev Drug Discov*, 17, pp. 353-377.

Ganguly, R, Mohyeldin, A, Thiel, J, Kornblum, HI, Beullens, M, & Nakano, I 2015, "MELK-a conserved kinase: functions, signaling, cancer, and controversy", *Clin Transl Med*, 4, p. 11.

Goel, S, Decristo, MJ, Mcallister, SS, & Zhao, JJ 2018, "CDK4/6 Inhibition in Cancer: Beyond Cell Cycle Arrest", *Trends Cell Biol*, 28, pp. 911-925.

Gyawali, B, Rome, BN, & Kesselheim, AS 2021, "Regulatory and clinical consequences of negative confirmatory trials of accelerated approval cancer drugs: retrospective observational study", *Bmj*, 374, p. n1959.

Hunter, T 1998, "The Croonian Lecture 1997. The phosphorylation of proteins on tyrosine: its role in cell growth and disease", *Philos Trans R Soc Lond B Biol Sci*, 353, pp. 583-605.

Jiang, M, Gu, D, Liu, F, Lin, C, & Tian, L 2023, "Transforming cancer cells for long-term living with cancer: An inspiring new approach", *Chin J Cancer Res*, 35, pp. 108-125.

Kakiuchi, N, & Ogawa, S 2021, "Clonal expansion in non-cancer tissues", *Nat Rev Cancer*, 21, pp. 239-256.

Klein, ME, Kovatcheva, M, Davis, LE, Tap, WD, & Koff, A 2018, "CDK4/6 Inhibitors: The Mechanism of Action May Not Be as Simple as Once Thought", *Cancer Cell*, 34, pp. 9-20.

Malumbres, M, & Barbacid, M 2009, "Cell cycle, CDKs and cancer: a changing paradigm", *Nat Rev Cancer*, 9, pp. 153-166.

Martincorena, I, & Campbell, PJ 2015, "Somatic mutation in cancer and normal cells", *Science*, 349, pp. 1483-1489.

Martincorena, I, Fowler, JC, Wabik, A, Lawson, ARJ, Abascal, F, Hall, MWJ, Cagan, A, Murai, K, Mahbubani, K, Stratton, MR, Fitzgerald, RC, Handford, PA, Campbell, PJ, Saeb-Parsy, K, & Jones, PH 2018, "Somatic mutant clones colonize the human esophagus with age", *Science*, 362, pp. 911-917.

Maury, EA, Walsh, CA, Kahle, KT 2024, "Neurosurgery elucidates somatic mutations", *Science*, 10.1126/science.adj2244.

Mortuza, GB, Hermida, D, Pedersen, AK, Segura-Bayona, S, López-Méndez, B, Redondo, P, Rüther, P, Pozdnyakova, I, Garrote, AM, Muñoz, IG, Villamor-Payà, M, Jauset, C, Olsen, JV, Stracker, TH, & Montoya, G 2018, "Molecular basis of Tousled-Like Kinase 2 activation", *Nat Commun*, 9, p. 2535.

Naxerova, K 2021, "Mutation fingerprints encode cellular histories", *Nature*, 597, pp. 334-336.

Nurse, P 1990, "Universal control mechanism regulating onset of M-phase", *Nature*, 344, pp. 503-508.

Prasad, V 2016, "Perspective: The precision-oncology illusion", *Nature*, 537, p. S63.

Prasad, V 2020, *Malignant*, Johns Hopkins University Press, Baltimore (Maryland).

Rew, DA, & Wilson, GD 2000, "Cell production rates in human tissues and tumours and their significance. Part II: clinical data", *Eur J Surg Oncol*, 26, pp. 405-417.

Settleman, J, Sawyers, CL, & Hunter, T 2018, "Challenges in validating candidate therapeutic targets in cancer", *Elife*, 7.

Shomar, A, Barak, O, & Brenner, N 2022, "Cancer progression as a learning process", *iScience*, 25, p. 103924.

Sonnenschein, C, & Soto, AM 1999, *The Society of Cells: Cancer and Control of Cell Proliferation*, Springer Verlag, New York.

Sonnenschein, C, & Soto, AM 2000, "The somatic mutation theory of carcinogenesis: Why it should be dropped and replaced", *Molecular Carcinogenesis*, 29, pp. 1-7.

Sonnenschein, C, & Soto, AM 2020, "Over a century of cancer research: Inconvenient truths and promising leads", *PLoS Biol.*, 18, p. e3000670.

Sonnenschein, C, & Soto, AM 2021, "Control of cell proliferation: is the default state of cells quiescence or proliferation", *Organisms*, 5, pp. 33-42.

Soto, AM, & Sonnenschein, C 2020, "Revisiting D.W. Smithers's 'Cancer: An Attack on Cytologism' (1962)", *Biological Theory*, 15, pp. 180-187.

Suski, JM, Braun, M, Strmiska, V, & Sicinski, P 2021, "Targeting cell-cycle machinery in cancer", *Cancer Cell*, 39, pp. 759-778.

Tiwari, A, Kumari, B, Nandagopal, S, Mishra, A, Shukla, KK, Kumar, A, Dutt, N, Ahirwar, DK 2024, "Promises of Protein Kinase Inhibitors in Recalcitrant Small-Cell Lung Cancer: Recent Scenario and Future Possibilities", *Cancers*, 16, p. 963. https://doi.org/10.3390/cancers16050963.

Unni, N, & Arteaga, CL 2019, "Is Dual mTORC1 and mTORC2 Therapeutic Blockade Clinically Feasible in Cancer?", *JAMA Oncol*, 5, pp. 1564-1565.

Vasan, N, & Cantley, LC 2022, "At a crossroads: how to translate the roles of PI3K in oncogenic and metabolic signalling into improvements in cancer therapy", *Nat Rev Clin Oncol*, 19, pp. 471-485.

Weinberg, RA 2014, *The Biology of Cancer*, Garland Science, New York.

Zhou, Y, Shen, JK, Hornicek, FJ, Kan, Q, & Duan, Z 2016, "The emerging roles and therapeutic potential of cyclin-dependent kinase 11 (CDK11) in human cancer", *Oncotarget*, 7, pp. 40846-40859.

Perspective & Hypotheses

Vol. 7, No. 1 (2024)
ISSN: 2532-5876

DOI: 10.13133/2532-5876/18466

The Role of Biological Plasticity in Model-based Translational Research

Jessica A. Bolker[a*]

[a] *Department of Biological Sciences, University of New Hampshire, 216 Rudman Hall, 46 College Road, Durham (NH), USA, 03824*

***Corresponding author:** Jessica A. Bolker, Email: Jessica.Bolker@usnh.edu*

Abstract

Reports of low replicability and translatability of biomedical research have called the value of animal models into question. The problems are real, but abandoning animal research is not the solution. Rather, improving the translatability of model-based research requires attention to relevant differences between humans and models, and to attributes of the models themselves that are essential to both robust science and effective translation. One is biological plasticity, the responsiveness of individual organisms to complex and variable environments. Though under-represented in model systems (for both historical and practical reasons), plasticity is central to human biology. While there are good reasons to minimize environmentally-induced variation in model-based research, doing so may undermine its translatability by eliding the kinds of external influences that are critical to human development, health, and disease. Accounting for plasticity can strengthen both the replicability and the translatability of model-based studies; this paper identifies strategies for doing so at each stage of the research process.

Citation: Bolker, JA 2024, "The role of biological plasticity in model-based translational research", *Organisms: Journal of Biological Sciences*, vol. 7, no. 1, pp. 29-39. DOI: https://doi.org/10.13133/2532-5876/18466

1. Introduction

In recent years, emerging information about low replicability and translatability of biomedical research in animal models has prompted some to question their utility (Greek and Menache 2013; Pound and Ritskes-Hoitinga 2018). While the problems are real, the answer is not to give up animal studies. Rather, improving the translatability of model-based research requires paying attention to attributes of the models themselves that are essential to both robust science and effective translation (Domínguez-Oliva *et al.* 2023; Garner *et al.* 2017; Pallocca, Rovida, and Leist 2022; Robinson *et al.* 2019; Ferreira *et al.* 2020).

One such attribute is biological plasticity, the responsiveness of individual organisms to complex and variable environments. The effects of plasticity may be adaptive, negative, or neutral (for example, learning and acquired immunity are adaptive; PTSD and anaphylaxis are not). Plasticity is ubiquitous, and recognition of its importance in ecology, evolution, conservation, and medicine is now widespread (Gilbert and Epel 2015; Guidi *et al.* 2021; Levis *et al.* 2018; McCarthy and Birney 2021; Nobile, Di Sipio Morgia, and Vento 2022; Uher 2014; Sultan 2015; West-Eberhard 2003). In particular, environmental influences generate critical variations in development, health, and disease between individuals and across human populations.

In contrast, model systems used for biomedical research are constructed to minimize variation. By deliberate design and as a consequence of their history (which species are chosen, and what happens to them once they enter the research ecosystem; Bolker 1995; Krinke 2000; Logan 2002), models typically show limited phenotypic variation and

Box 1: Plasticity and validity

Study design, including model choice, dictates what form of validity can be claimed for the results. (For discussion and definitions of validity, see Garner *et al.* 2017; van der Staay, Arndt, and Nordquist 2009; Würbel 2017).

Structural validity in a study of plasticity requires that the model and target species share mechanisms for transducing environmental signals into phenotypic changes. The cues and outcomes may differ. For example, while the adrenocortical stress response is conserved across vertebrates, the identity of stressors and external manifestations of stress are shaped by each species' evolution and by individuals' prior experience.

Predictive validity, especially in a biomedical context, does not require that the cue and mechanism match precisely between model and target: a bioassay or screening study can yield useful outcomes (such as predictions about efficacy or toxicity of a drug candidate) even if we do not fully understand how it works. However, the range over which predictions are reliable is difficult to assess without some understanding of mechanism.

Internal validity – replicability and robustness of results – can theoretically be achieved by standardizing or controlling all possible variables. To account for the possibility of plasticity in the study system, it is important to record standardized (and even presumably irrelevant) factors as well as deliberately controlled or manipulated variables. Such background information can also support *post hoc* analysis if experiments stop working or cannot be reproduced in other labs.

External validity corresponds to exportability or translatability, and is often assumed for models where structural validity has been established. Structural validity alone does not guarantee successful translation. To warrant claims of external validity, model-based studies should describe support factors as well as focal mechanisms, noting that support factors for a shared mechanism may differ between species.

relatively little plasticity: they are inherently robust to environmental variation, and embedded in systems of standardized husbandry, genetics, and research practices. For some models, generations of breeding and selection in laboratory environments – the process of "laboratorization" (Robinson 1965) – may have rendered them even less flexible than their wild relatives. These attributes have many benefits: they can increase statistical power, reduce animal numbers and costs, streamline husbandry, and facilitate replication. Such practices are key to stabilizing phenotypes, especially traits that might vary in response to environmental factors.

However, deliberately removing plasticity from model systems has epistemic as well as biological implications: how we use models can weaken their external validity by eliding external influences that are critical to human biology (Voelkl *et al.* 2020b; 2020a; 2020b; Voelkl and Würbel 2016; 2021). While standardized models maintained in constant environments are excellent tools for studying molecular pathways and other internal mechanisms, they are poorly suited for questions where plasticity matters – or might matter. Researchers may fail to recognize the existence or importance of environmental influences simply because their models render such effects invisible. To counter this bias and increase the chance that results from a model system will translate to humans, it is essential to consider plasticity at each stage of the research process.

2. Planning: Is Plasticity a Question or a Challenge?

Choosing a suitable model – the right tool for the job – depends on the research goal: what the question is, and what sort of answer is desired (Bolker 2014; Clarke and Fujimura 1992). Articulating what role the model will play in addressing the question is central to identifying criteria for model choice, as well as assessing the strengths and limitations of whatever model is selected (Bolker 2009).

If the question is about plasticity, then the model needs to match the target with respect to relevant plasticity-related traits. To begin with, it is helpful if the degree of environmental responsiveness is broadly similar between species: using an inflexible model to represent a highly plastic target is not ideal.[a] (If doing so is unavoidable, the implications of this disparity need to be recognized and addressed.) Not everything needs to match: the mechanism that transduces environmental information into a shift in the phenotype or biology of the model need not be identical in model and target, unless that is what the study is about (Box 1). Conversely, if the research centers on transduction mechanisms, it does not matter if the cues or specific outcomes are different provided they operate via the same pathways: structural validity requires similarity of mechanism, not identical inputs and outputs.

[a] This is especially tricky in translational research because humans are much more plastic than most common animal models.

Table 1: What does it mean to "know your species"?

Aspect	***Why it matters***
Phylogenetic and taxonomic position	Informs expectations for exportability, depending on how the model is being used Identifies ancestral vs. specialized traits Documents taxon-wide characteristics (e.g. physiology, life history strategy)
Timing of lineage divergence between model and target species	Provides context for patterns of trait similarity and divergence across clades Describes how long model and target have been evolving independently
Evolutionary history and known or inferred selection pressures in the wild	Helps identify adaptations with implications for model use and/or husbandry practices Suggests behavioral and other preferences that may reflect adaptations to evolutionary niche Guides the search for models with adaptations that make them especially useful (Krogh models)
History of laboratory strains: origins, genetics, breeding, selection in lab environments	Highlights ways laboratory animals may diverge from their recent (wild) ancestors, e.g. genetic bottlenecks and intense selection for tractability under lab conditions Identifies deliberate or incidental selection pressures in research environments that may reduce plasticity in lab strains, leading to an underestimate of its importance in ancestral or "wild type" lineages
Sensory and physiological range	Informs husbandry practices that maximize well-being and reduce stress Enables the design of experimental stimuli and assays that align with subjects' sensory capabilities

How can one choose an appropriate model for studying mechanisms of plasticity, if the research objective is to discover what the mechanisms are, or details of how they operate? One approach is to consider the evolutionary origins of each species' plasticity, and the role and context of the trait with respect to species-specific natural history (Bolker 2019; Levis *et al.* 2018). This is analogous to the strategy recommended by Blanchard and Chalfin of studying functionally and ecologically relevant behaviors in model animals, rather than relying on superficial similarities to humans (Blanchard, Summers, and Blanchard 2013; Chalfin *et al.* 2014). Importantly, similar environmental cues may have disparate impacts in different species (or lab vs. wild populations) as a consequence of their different evolutionary histories – and human evolutionary history, especially in relation to health and disease, is particularly complex (Benton *et al.* 2021; Natterson-Horowitz *et al.* 2023).

One strategy for bolstering the ability of animal models to represent human targets is to assess environmental information in a species-agnostic or at least a translatable way, recognizing that different animals have different needs and different perceptions of the world (Keijer, Li, and Speakman 2019; Makowska and Weary 2019; Yong 2022). For example, "thermoneutral range," "normal social context," "expected microbial exposure," and "low-stress environment" all have species-specific values, ranges, or definitions (Garner *et al.* 2017; Gordon 2017). Performing physiological experiments within the thermoneutral range, or providing adequate nutrition, does not require that temperature or diet be the same for every species: it means that these environmental factors need to be in a species-appropriate range for each of them. This requirement extends to social arrangements. For instance, the presence of other mice improves recovery in a murine cancer model (Hermes *et al.* 2009; Kerr *et al.* 1997), but while pet mice might conceivably enhance recovery in people with cancer, what is relevant to patients is support from other humans (Kroenke *et al.* 2006).

Determining what is species-appropriate requires understanding the species' natural history and evolution [Table 1]. Knowledge of species-specific needs is already built into many husbandry protocols (*e.g.*, provision of adequate nutrition via customized commercial feeds, and physical environments that support the expression of natural behaviors). The natural history and evolution of particular models can have paradoxical implications for how we maintain and use them: for example, the evolution of mice as small ground-dwelling scavengers able to thrive in a microbe-rich environment explains their high tolerance for bacterial toxins (Mestas and Hughes 2004; Perlman 2016; Webb *et al.* 2015). In fact, the per-kilogram dose of endotoxins sufficient to trigger an inflammatory/immune response in mice is far

higher than in humans (Mestas and Hughes 2004; Webb *et al.* 2015). However, laboratory mice raised under standard husbandry practices that strictly limit pathogen exposure have immune systems that never mature to the normal level for an adult mouse (or human) (Beura *et al.* 2016; Reese *et al.* 2016) – though development in the uterus of a wild surrogate yields lab mice with normal adult immune function (Rosshart *et al.* 2019).

If the research question is not centered on plasticity, it is essential to consider whether plasticity might impact the trait or phenomenon anyway, and how to account for that possibility in the study design. Traits or systems that directly mediate an organism's interaction with its environment (via sensory, neurobehavioral, or immune systems) are especially likely to have undergone selection for plasticity (Bolker 2019). But there is no simple, predictive rule. One can look for evidence of plasticity in related species, as well as in more distant taxa with evolutionary histories shaped by similar selection pressures: while not definitive, the occurrence of plasticity in either of those groups can provide clues about its possible role in the prospective model species. Plasticity itself can evolve, certainly at highter taxonomic levels but also potentially between wild and lab-selected lineages (Krinke 2000; Levis *et al.* 2018; West-Eberhard 2003). Here, again, it is critical to know your species [Table 1].

Besides knowing about their species, researchers need to know about the environment in which animals are housed and data collected. Laboratory conditions may generate confounding variation: statistical noise can result from acoustic noise (Lauer *et al.* 2009; Parker *et al.* 2022; Pfaff 1974), or interfere with animals' normal biology in ways that increase stress and/or energy expenditure (Garner *et al.* 2017; Gaskill and Garner 2017; Gordon 2017; Lac, Tavernier, and Moro 2023; Mo, Renoir, and Hannan 2016; Toth 2015). The traditional approach has been to try to standardize everything (Festing 2014; Festing and Altman 2002) – which certainly has advantages, but (besides tending to mask plasticity that might be present) this strategy may miss factors that are not recognized *a priori* as important: the "unknown unknowns" (Mogil 2017).

A dramatic example of an unrecognized but powerful influence was the realization that "standard" commercial rodent feeds contain high and variable levels of phytoestrogens (mainly from soy) that can confound research in areas from cancer to endocrinology (Heindel 2008; Ruhlen 2008). Paradoxically, providing soy-free diets to lab rats induces obesity, likely via perturbation of fetal metabolism; Ruhlen *et al.* suggest that this unexpected result might reflect prior adaptation of lab-bred strains to high phytoestrogen intake from commercial feeds (Ruhlen 2008). Dietary levels of phytoestrogens during pregnancy and early development can have profound impacts; however, researchers purchasing animals from commercial suppliers rarely have access to information about this key environmental variable (Heindel 2008).

Trying to eliminate variation runs the risk of missing some of the sources (such as phytoestrogens in rat chow). Another approach is to deliberately introduce variation in an explicit, systematic way, or attempt to distribute preexisting variation evenly via heterogenization (van der Staay, Arndt, and Nordquist 2010; Richter, Garner, and Würbel 2009; Richter *et al.* 2010; 2011; Würbel 2000). This strategy can potentially account for sources of variation that have yet to be recognized.

In addition to increasing overall variation, plasticity in individual research subjects can lead to results that reflect environmental variation in a systematic way, and can confound or eclipse the effects of the focal experimental variable (Mo, Renoir, and Hannan 2016; Mogil 2017; Toth 2015). Randomizing the placement of subjects or treatment groups within the environment (*e.g.*, locations of plots, tanks, traps, cages on racks) is important, but cannot eliminate biases related to experimental or observational techniques *per se*. For example, experimentally-modified and control animals might differ in their susceptibility to stress from handling or administration of placebo treatments; there can also be significant differences between individuals (Andrews and File 1993; Aydin, Frohmader, and Akil 2015; Hurst and West 2010). Even within individuals, details such as the exact location of injections can have unexpectedly significant effects (Auerbach 1978).

Determining what degree of standardization is appropriate for a given study is context-dependent and difficult (van der Staay, Arndt, and Nordquist 2010). Standardization can reduce the number of animals used and enhance statistical power and the ability to detect small effects. But if the trait being studied is, itself, plastic, over-standardization can reduce external validity (especially translation to humans) and even mask the mechanisms one hopes to understand. The goal should be to "standardize, but not too much" (Bolker 2019; Richter, Garner, and Würbel 2009; Striedter 2022).

3. Performance: Collect Environmental as well as Experimental Data

Along with results of planned experiments, it is essential to document the context in which the study is carried out: environmental factors that might turn out to be relevant, or correlate with unexpected outcomes or variation (Toth 2015). For example, details of husbandry practices or characteristics of research personnel (such as their sex; Sorge *et al.* 2014), while rarely explicitly noted in study designs, can have significant effects on lab rodents and thus on study results (Mogil 2017).

What data are worth collecting? Start by considering what environmental information is known or suspected to matter to the organisms in question. Toth and Neville *et al.* survey the importance of rodent cage environments to the reproducibility of preclinical studies (Neville *et al.* 2023; Toth 2015), and Mogil (2017) reviews external factors that affect the outcomes of pain studies in mice. Notably, both the magnitude and the direction of environmental effects can vary by genetic strain (Crabbe, Wahlsten, and Dudek 1999; Crawley *et al.* 1997; Mogil 2017).

The already long, but likely still incomplete, list of environmental factors that are known to matter suggests that there are a lot of things researchers should be tracking (and describing in published methods), from animal housing and handling to data about the physical environment (Toth 2015; Neville *et al.* 2023; Sundberg and Schofield 2018). Critically, we need to be thinking about this from the perspective of the animals, and collecting data within the species' sensory range, for instance measuring acoustic noise across frequencies audible to rodents (Lauer *et al.* 2009; Parker *et al.* 2022; Turner 2020). Even if mice in a research study are serving as surrogate models for humans, they experience their environment *as mice*: what counts as a normal or a stressful noise level, temperature range, or housing situation for them is not the same as what counts for us (Fischer, Cannon, and Nedergaard 2018; Garner *et al.* 2017; Keijer, Li, and Speakman 2019; Yong 2022; Weber *et al.* 2017)... and what seems normal to a laboratory-bred rodent may differ from its wild ancestors' natural environment, given the divergent selection pressures acting on research populations (Krinke 2000).

As a start, animal facilities should incorporate routine, automatic, continuous monitoring of physical variables such as temperature, humidity, and ambient light and noise. Inexpensive data loggers can be installed in each cage or enclosure, or at least in each room where animals are used (ideally in multiple locations). Time-stamped environmental data from husbandry facilities could be collected as part of routine management, and made available to everyone who has research animals housed there.

Time, itself, can also be an important variable. The developmental stage at which animals are subjected to stressful shipping or procedures can affect their physiological response (Beery 2018). At a smaller scale, the time of day at which data are collected can determine what the data look like: circadian clocks regulate key processes ranging from behavior to cell proliferation to drug response (Andersen 2023; Lévi *et al.* 2024; Sato and Sato 2023).

The biotic environment should be tracked and accounted for as well. Perhaps the most obvious aspect is housing. Not only social vs. individual housing, but social dynamics within shared cages, significantly affect the biology of lab rodents (George, Padilla-Coreano, and Opendak 2023; Arakawa 2018; Beery *et al.* 2020; Kerr *et al.* 1997; Manouze *et al.* 2019; Mo, Renoir, and Hannan 2016; Mogil 2017; Mumtaz *et al.* 2018).

Along with intraspecific interactions, it is critical to consider the influence of other species – particularly microbes (Honda and Littman 2016). Pathogens are routinely monitored in animal facilities, but we should also track at least some of the vast array of commensal and symbiotic species. Microbial communities play crucial roles in the development and function of macroorganisms, shaping host phenotypes at both morphological and behavioral levels, and they can be an unrecognized source of variation in rodent models (Franklin and Ericsson 2017; Gilbert and Epel 2015; Honda and Littman 2016; Kim *et al.* 2017; Shin Yim *et al.* 2017; Witjes, Boleij, and Halffman 2020). Analyzing environmental DNA collected via air filters (as well as routine samples of bedding, surface swabs, etc.) could track the presence, and potentially the abundance, of different microbes at housing or research sites (Albers *et al.* 2023; Ruppert, Kline, and Rahman 2019). The rapid expansion of research on the laboratory animals' microbiomes will shed light on a key aspect of model-based research, in addition to addressing the specific questions targeted in each study (Honda and Littman 2016).

Another aspect of the biotic environment that we may underestimate is the range and impact of odors in housing and testing facilities. Humans are not very good at odor detection, but other animal species are exquisitely sensitive to chemical signals, and rely on them to modulate their behavior and physiology (Yong 2022). Engineering approaches

to odor monitoring focus primarily on chemicals that are detectable by and/or immediately relevant to humans, but in principle the technology could be modified to monitor odors that are detectable by, and may be important to, laboratory animals (Reimringer *et al.* 2022). This would, of course, depend on deciding exactly what should be monitored – which brings us back to unknown unknowns. We could start with a "candidate odor" strategy (analogous to candidate gene approaches), for example monitoring known pheromones, stress hormones, and other molecules with demonstrated impacts on recipients' biology.

Routinely collected environmental information may retrospectively identify a factor that was not intended as a variable, but that turned out to influence results. However, it is unwise to go on a fishing expedition in search of environmental correlates for otherwise unexplained outcomes, in hopes of finding a statistically significant relationship to cite as a cause. Such correlates should be treated as only preliminary or suggestive, if they are not what the study was originally designed to evaluate. For example, if the study did not set out to measure the effects of different bedding materials, but effects seem to have occurred, a subsequent experiment can be carried out to directly assess the effects of bedding under conditions that (otherwise) match those of the original study. Any significant findings from a robust study designed to test the effect of bedding may then shed light on previous work where bedding might have been an uncontrolled but significant variable (*e.g.* Kondo, Kropik, and Wong 2022; Sláma 1966).

4. Interpretation: Accounting for Plasticity as a Possible Cause of Observed Effects

If a study was designed to examine plasticity, interpretation of the results should consider not just individual and internal mechanisms, but also environmental factors that may have contributed to the observed outcome. Beyond the variables whose effect the study is designed to test, it is essential to address other aspects of the environment that may serve as support factors that enable particular outcomes (Cartwright and Hardie 2012). Rather than considering the environment as outside the frame of a study, we need to start thinking about it as part of the frame – or even part of the picture (Bolker 2014).

Considering plasticity can be crucial even for studies that are not designed to assess it, particularly when such studies yield unexpectedly variable or contradictory results (Jaric *et al.* 2022; Voelkl and Würbel 2021). Plasticity is one possible explanation for observed variation. However, caution is required when drawing conclusions about the importance – or irrelevance – of the environment from studies that were not deliberately designed to assess plasticity. While environmentally-driven mechanisms may well help explain observed variation, lack of variation does not necessarily imply lack of plasticity, because standardizing the study environment also standardizes plastic traits. This constitutes an absence of evidence for plasticity, not evidence of its absence. If statistical analysis suggests the existence of batch effects, plasticity in response to unrecognized environmental factors or biases should be considered as a possible cause (Randall *et al.* 2019). Alternative explanations (unrelated to plasticity) could include equipment calibration, variation in reagents, or other factors independent of biology.

Failures to replicate previous work can be due to unrecognized environmental factors. Details of animal husbandry protocols, handling during experiments, and microbial exposure are often omitted from published descriptions of methods because they are assumed to be constant and/or unimportant. That assumption may need to be revisited, and both the original study and the attempted replication scrutinized for potentially significant environmental factors (Jaric *et al.* 2022; Neville *et al.* 2023). (The more thoroughly such factors were monitored and recorded along the way, the easier this will be.) Even in cases where the environment plays no causal role in producing an outcome, it may still provide support factors for conserved mechanisms, thus determining the exportability or translatability of findings to other species (Cartwright and Hardie 2012). Absence of essential support factors can lead to replication failure, even if the mechanism being studied is present.

Environmental standardization is often a deliberate strength, not a weakness, of a particular study, but this approach may limit translatability. Translating results from a tightly controlled model to a highly variable target species is an epistemic challenge; bridging the gap requires understanding the scope and nature of plasticity on both sides. Plasticity need not diminish exportability: what is critical is to identify potential sources of plasticity, and either standardize them, randomize their impacts through systematic variation, or align them appropriately between model and target (Duncan and Keller 2011; Richter 2017). The premise that plasticity is not relevant to a given study needs to be explicit and justified, not just an assumption based on the use of a carefully standardized model species in a tightly controlled environment. Thinking through

ways in which plasticity *might* matter is essential to assessing the extent to which findings from such model systems may be exportable – especially how well they will translate to notoriously flexible and un-standardized humans.

5. Summary and Conclusions

Both the history and the current practice of model-based research focus on standardization and internal processes (Bolker and Brauckmann 2015; Logan 2002). This approach has yielded deep insights into traits and mechanisms that have strong genetic bases and little external connectivity. As described by Ankeny and Leonelli, the standardization, isolation, and artificiality of model species generate a form of "placelessness" that is central to their explanatory power and broad acceptance (Ankeny and Leonelli 2020). From a biological perspective, however, "place" matters a great deal. There is thus a tension between the placelessness researchers attempt to construct (and then implicitly assume) in model-based research, and their ability to draw conclusions about species or phenomena in which the environment plays an important role. Moreover, we find ourselves with a set of dominant models that are generally poorly suited for studying plastic traits (Bolker 2017).

Why does it matter how much plasticity there is in a model species, especially if it does not appear to affect the results of a given study? Many aspects of human health – from immunology (Martin *et al.* 2021), to neuropsychiatric disorders (Uher 2014; Assary *et al.* 2018), to racial disparities in pregnancy outcomes (Leimert and Olson 2020) – depend heavily on environmental factors and gene-by-environment interactions (Benton *et al.* 2021; Duncan and Keller 2011; Guidi *et al.* 2021). Research strategies that seek to understand the underlying mechanisms while ignoring or eliding plasticity are unlikely to succeed. Environmental influences, and plastic biological responses, are central to many of the questions we want to answer in humans: What are the underlying mechanisms of immunological and neuropsychiatric disorders? What causes cancer? What factors influence the onset and progression of chronic disease? What determines whether the presence of genetic risk factors ultimately leads to disease in particular individuals (McCarthy and Birney 2021)?

The solution is not to give up on these questions or on widely-used, powerful models. Rather, recognizing the potential role of environmental factors and integrating that knowledge into the design, performance, and interpretation of experiments can give us the best of both worlds. A "yes-and" approach to biomedical research means studying humans whenever and however we can, and employing animal models in ways that are most likely to yield translatable knowledge. Accounting for plasticity can both improve the translation of model-based research to humans, and expand our understanding of the fundamental biology of all species.

Acknowledgments

The author is grateful to Rick Gawne and an anonymous reviewer for their thoughtful and constructive comments on the manuscript.

References

Albers, TM, Henderson, KS, Mulder, GB, & Shek, WR 2023, "Pathogen Prevalence Estimates and Diagnostic Methodology Trends in Laboratory Mice and Rats from 2003 to 2020", *Journal of the American Association for Laboratory Animal Science*, 62, 3, pp. 229–42. Available from: https://doi.org/10.30802/AALAS-JAALAS-22-000097.

Andersen, B, Dual, J, & Swaroop Karri, S 2023, "How and Why the Circadian Clock Regulates Proliferation of Adult Epithelial Stem Cells", *Stem Cells*, 41(4, pp. 319-327. Available from: https://academic.oup.com/stmcls/article/41/4/319/7027347.

Andrews, N, & File, SE 1993, "Handling History of Rats Modified Behavioural Effects of Drugs in the Elevated Plus-Maze Test of Anxiety", *European Journal of Pharmacology*, 235, 1, pp. 109-112.

Ankeny, R, & Leonelli, S 2020, *Model Organisms*, 1st ed., Cambridge University Press. Available from: https://doi.org/10.1017/9781108593014.

Arakawa, H 2018, "Ethological Approach to Social Isolation Effects in Behavioral Studies of Laboratory Rodents", *Behavioural Brain Research*, 341, April, pp. 98–108. Available from: https://doi.org/10.1016/j.bbr.2017.12.022.

Assary, E, Vincent, JP, Keers, R, & Pluess, M 2018, "Gene-Environment Interaction and Psychiatric Disorders: Review and Future Directions", *Seminars in Cell & Developmental Biology*, 77, May, pp. 133–43. Available from: https://doi.org/10.1016/j.semcdb.2017.10.016.

Auerbach, R, Morrissey, LW, & Sidky, YA 1978, "Regional Differences in the Incidence and Growth of Mouse Tumors following Intradermal or Subcutaneous Inoculation", *Cancer Research*, 38: 1739-1744.

Aydin, C, Frohmader, K, & Akil, H 2015, "Revealing a Latent Variable: Individual Differences in Affective Response to Repeated Injections", *Behavioral Neuroscience*, 129, 5, pp. 679–82. Available from: https://doi.org/10.1037/bne0000084.

Beery, AK 2018, "Inclusion of females does not increase variability in rodent research studies", *Current Opinion in Behavioral Sciences*, 23: 143-149.

Available from: https://linkinghub.elsevier.com/retrieve/pii/S235215461730205X.

Beery, AK, Holmes, MM, Lee, W, & Curley, JP 2020, "Stress in Groups: Lessons from Non-Traditional Rodent Species and Housing Models", *Neuroscience & Biobehavioral Reviews*, 113, June, pp. 354–72. Available from: https://doi.org/10.1016/j.neubiorev.2020.03.033.

Benton, ML, Abraham, A, LaBella, AL, Abbot, P, Rokas, A, & Capra, JA 2021, "The Influence of Evolutionary History on Human Health and Disease", *Nature Reviews Genetics*, 22, 5, pp. 269–83. Available from: https://doi.org/10.1038/s41576-020-00305-9.

Beura, LK, Hamilton, SE, Bi, K, Schenkel, JM, Odumade, OA, Casey, KA, Thompson, EA, *et al.* 2016, "Normalizing the Environment Recapitulates Adult Human Immune Traits in Laboratory Mice", *Nature*, 532, 7600, pp. 512–16. Available from: https://doi.org/10.1038/nature17655.

Blanchard, CD, Summers, CH, & Blanchard, RJ 2013, "The Role of Behavior in Translational Models for Psychopathology: Functionality and Dysfunctional Behaviors", *Neuroscience & Biobehavioral Reviews*, 37, 8, pp. 1567–77. Available from: https://doi.org/10.1016/j.neubiorev.2013.06.008.

Bolker, JA 1995, "Model Systems in Developmental Biology", *BioEssays*, 17, 5, pp. 451–55. Available from: https://doi.org/10.1002/bies.950170513.

Bolker, JA 2014, "Models in Context: Biological and Epistemological Niches", in G Barker, E Desjardins, & T Pearce, (eds), *Entangled Life: Organism and Environment in the Biological and Social Sciences*, pp. 153–66, Springer Netherlands, Dordrecht. Available from: https://doi.org/10.1007/978-94-007-7067-6_8.

Bolker, JA 2009, "Exemplary and Surrogate Models: Two Modes of Representation in Biology", *Perspectives in Biology and Medicine*, 52, 4, pp. 485–99. Available from: https://doi.org/10.1353/pbm.0.0125.

Bolker, JA 2017, "Animal Models in Translational Research: Rosetta Stone or Stumbling Block?", *BioEssays*, 39, 12, p. 1700089. Available from: https://doi.org/10.1002/bies.201700089.

Bolker, JA, & Brauckmann, S 2015, "From Experimental Zoology to Big Data: Observation and Integration in the Study of Animal Development", *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 323, 5, pp. 277–91. Available from: https://doi.org/10.1002/jez.1924.

Bolker, JA 2019, "Selection of Models: Evolution and the Choice of Species for Translational Research", *Brain, Behavior and Evolution*, 93, 2–3, pp. 82–91. Available from: https://doi.org/10.1159/000500317.

Cartwright, N, & Hardie, J 2012. *Evidence-Based Policy: A Practical Guide to Doing It Better*, Oxford University Press, Oxford.

Chalfin, L, Dayan, M, Levy, DR, Austad, SN, Miller, RA, Iraqi, FA, Dulac, C, & Kimchi, T 2014, "Mapping Ecologically Relevant Social Behaviours by Gene Knockout in Wild Mice", *Nature Communications*, 5, August, p. 4569. Available from: https://doi.org/10.1038/ncomms5569.

Clarke, AH, & Fujimura, JH 1992, *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences*, Princeton University Press, Princeton (NJ).

Crabbe, JC, Wahlsten, D, & Dudek, BC 1999, "Genetics of Mouse Behavior: Interactions with Laboratory Environment", *Science*, 284, 5420, pp. 1670–72. Available from: https://doi.org/10.1126/science.284.5420.1670.

Crawley, JN, Collins, A, Crabbe, JC, Frankel, W, Henderson, N, Hitzemann, RJ, Maxson, SC, *et al.* 1997, "Behavioral Phenotypes of Inbred Mouse Strains: Implications and Recommendations for Molecular Studies", *Psychopharmacology*, 132, 2, pp. 107–24. Available from: https://doi.org/10.1007/s002130050327.

Domínguez-Oliva, A, Hernández-Ávalos, I, Martínez-Burnes, J, Olmos-Hernández, A, Verduzco-Mendoza, A, & Mota-Rojas, D 2023, "The Importance of Animal Models in Biomedical Research: Current Insights and Applications", *Animals*, 13, 7, p. 1223. Available from: https://doi.org/10.3390/ani13071223.

Duncan, LE, & Keller, MC 2011, "A Critical Review of the First 10 Years of Candidate Gene-by-Environment Interaction Research in Psychiatry", *American Journal of Psychiatry*, 168, 10, pp. 1041–49. Available from: https://doi.org/10.1176/appi.ajp.2011.11020191.

Ferreira, GS, Veening-Griffioen, DH, Boon, WPC, Moors, EHM, & Van Meer, PJK 2020, "Levelling the Translational Gap for Animal to Human Efficacy Data", *Animals*, 10, 7, p. 1199. Available from: https://doi.org/10.3390/ani10071199.

Festing, MFW, & Altman, DG 2002, "Guidelines for the Design and Statistical Analysis of Experiments Using Laboratory Animals", *ILAR Journal*, 43, 4, pp. 244–58. Available from: https://doi.org/10.1093/ilar.43.4.244.

Festing, MFW 2014, "Evidence Should Trump Intuition by Preferring Inbred Strains to Outbred Stocks in Preclinical Research", *ILAR Journal*, 55, 3, pp. 399–404. Available from: https://doi.org/10.1093/ilar/ilu036.

Fischer, AW, Cannon, B, & Nedergaard, J 2018, "Optimal Housing Temperatures for Mice to Mimic the Thermal Environment of Humans: An Experimental Study", *Molecular Metabolism*, 7, January, pp. 161–70. Available from: https://doi.org/10.1016/j.molmet.2017.10.009.

Franklin, CL, & Ericsson, AC 2017, "Microbiota and Reproducibility of Rodent Models", *Lab Animal*, 46, 4, pp. 114–22. Available from: https://doi.org/10.1038/laban.1222.

Garner, JP, Gaskill, BN, Weber, EM, Ahloy-Dallaire, J, & Pritchett-Corning, KR 2017, "Introducing Therioepistemology: The Study of How Knowledge Is Gained from Animal Research", *Lab Animal*, 46, 4, pp. 103–13. Available from: https://doi.org/10.1038/laban.1224.

Gaskill, BN, & Garner, JP 2017, "Stressed out: Providing Laboratory Animals with Behavioral Control to Reduce the Physiological Effects of Stress", *Lab Animal*, 46, 4, pp. 142–45. Available from: https://doi.org/10.1038/laban.1218.

George, A, Padilla-Coreano, N, & Opendak, M 2023, "For Neuroscience, Social History Matters", *Neuropsychopharmacology*, 48, 7, pp. 979–80. Available from: https://doi.org/10.1038/s41386-023-01566-8.

Gilbert, SF, & Epel, D 2015. *Ecological Developmental Biology: The Environmental Regulatiom of Development, Health, and Evolution*, Sinauer Associates Inc., Sunderland (MA).

Gordon, CJ 2017, "The Mouse Thermoregulatory System: Its Impact on Translating Biomedical Data to Humans", *Physiology and Behavior*, 179, April, pp. 55–66. Available from: https://doi.org/10.1016/j.physbeh.2017.05.026.

Greek, R, & Menache, A 2013, "Systematic Reviews of Animal Models: Methodology versus Epistemology", *Int. J. Med. Sci*, 10. Available from: https://doi.org/10.7150/ijms.5529.

Guidi, J, Lucente, M, Sonino, N, & Fava, GA 2021, "Allostatic Load and Its Impact on Health: A Systematic Review", *Psychotherapy and Psychosomatics*, 90, 1, pp. 11–27. Available from: https://doi.org/10.1159/000510696.

Heindel, JJ, & Vom Saal, FS 2008, "Meeting Report: Batch-to-Batch Variability in Estrogenic Activity in Commercial Animal Diets—Importance and Approaches for Laboratory Animal Research", *Environmental Health Perspectives*, 116, 3, pp. 389-393. Available from: https://ehp.niehs.nih.gov/doi/10.1289/ehp.10524.

Hermes, GL, Delgado, B, Tretiakova, M, Cavigelli, SA, Krausz, T, Conzen, SD, & McClintock, MK 2009, "Social Isolation Dysregulates Endocrine and Behavioral Stress While Increasing Malignant Burden of Spontaneous Mammary Tumors", *Proceedings of the National Academy of Sciences of the United States of America*, 106, 52, pp. 22393–98. Available from: https://doi.org/10.1073/pnas.0910753106.

Honda, K, & Littman, DR 2016, "The Microbiota in Adaptive Immune Homeostasis and Disease", *Nature*, 535, 7610, pp. 75-84. Available from: https://www.nature.com/articles/nature18848.

Hurst, JL, & West, RS 2010, "Taming Anxiety in Laboratory Mice", *Nature Methods*, 7, 10, pp. 825–26. Available from: https://doi.org/10.1038/nmeth.1500.

Jaric, I, Voelkl, B, Clerc, M, Schmid, MW, Novak, J, Rosso, M, Rufener, R, *et al.* 2022, "The Rearing Environment Persistently Modulates Mouse Phenotypes from the Molecular to the Behavioural Level", C Lino De Oliveira, (ed), *PLOS Biology*, 20, 10, p. e3001837. Available from: https://doi.org/10.1371/journal.pbio.3001837.

Keijer, J, Li, M, & Speakman, JR 2019, "What Is the Best Housing Temperature to Translate Mouse Experiments to Humans?", *Molecular Metabolism*, 25, July, pp. 168–76. Available from: https://doi.org/10.1016/j.molmet.2019.04.001.

Kerr, LR, Grimm, MS, Silva, WA, Andrã, W, Weinberg, J, & Emerman, JT 1997, "Effects of Social Housing Condition on the Response of the Shionogi Mouse Mammary Carcinoma (SC115) to Chemotherapy", *Cancer*, 2, 604, pp. 1124–28.

Kim, S, Kim, H, Yim, YS, Ha, S, Atarashi, K, Tan, TG, Longman, RS, Honda, K, Littman, DR, Choi, GB, & Huh, JR 2017, "Maternal Gut Bacteria Promote Neurodevelopmental Abnormalities in Mouse Offspring", *Nature*, 549(7673, pp. 528-532. Available from: https://www.nature.com/articles/nature23910.

Kondo, SY, Kropik, J, & Wong, MA 2022, "Effect of Bedding Substrates on Blood Glucose and Body Weight in Mice", *Journal of the American Association for Laboratory Animal Science*, 61, 6, pp. 611–14. Available from: https://doi.org/10.30802/AALAS-JAALAS-22-000047.

Krinke, GJ (ed) 2000, *The Laboratory Rat*, Elsevier.

Kroenke, CH, Kubzansky, LD, Schernhammer, ES, Holmes, MD, & Kawachi, I 2006, "Social Networks, Social Support, and Survival after Breast Cancer Diagnosis", *Journal of Clinical Oncology*, 24, 7, pp. 1105–11. Available from: https://doi.org/10.1200/JCO.2005.04.2846.

Lac, M, Tavernierrobinson, G, & Moro, C 2023, "Does Housing Temperature Influence Glucose Regulation and Muscle-Fat Crosstalk in Mice?", *Biochimie*, February, S0300908423000287. Available from: https://doi.org/10.1016/j.biochi.2023.01.019.

Lauer, AM, May, BJ, Hao, ZJ, & Watson, J 2009, "Analysis of Environmental Sound Levels in Modern Rodent Housing Rooms", *Lab Animal*, 38, 5, pp. 154–60. Available from: https://doi.org/10.1038/laban0509-154.

Leimert, KB, & Olson, DM 2020, "Racial Disparities in Pregnancy Outcomes: Genetics, Epigenetics, and Allostatic Load", *Current Opinion in Physiology*, 13, February, pp. 155–65. Available from: https://doi.org/10.1016/j.cophys.2019.12.003.

Lévi, FA, Okyar, A, Hadadi, E, Innominato, PF, & Ballesta, A 2024, "Circadian Regulation of Drug Responses: Toward Sex-Specific and Personalized Chronotherapy", *Annual Review of Pharmacology and Toxicology*, 64, 1, pp. 89-114. Available from: https://www.annualreviews.org/doi/10.1146/annurev-pharmtox-051920-095416.

Levis, NA, Isdaner, AJ, & Pfennig, DW 2018, "Morphological Novelty Emerges From Pre-existing Phenotypic Plasticity", *Nature Ecology & Evolution*, 2, 8, pp. 1289-1297. Available from: https://www.nature.com/articles/s41559-018-0601-8.

Logan, CA 2002, "Before There Were Standards: The Role of Test Animals in the Production of Empirical Generality in Physiology", *Journal of the History of Biology*, 35, 2, pp. 329–63. Available from: https://doi.org/10.1023/A:1016036223348.

Makowska, IJ, & Weary, DM 2019, "A Good Life for Laboratory Rodents?" *ILAR Journal*, 60, 3, pp. 373–88. Available from: https://doi.org/10.1093/ilar/ilaa001.

Manouze, H, Ghestem, A, Poillerat, V, Bennis, M, Ba-M'hamed, S, Benoliel, JJ, Becker, C, & Bernard, C 2019, "Effects of Single Cage Housing on Stress, Cognitive, and Seizure Parameters in the Rat and Mouse Pilocarpine Models of Epilepsy", *Eneuro*, 6, 4, p. ENEURO.0179-18.2019. Available from: https://doi.org/10.1523/ENEURO.0179-18.2019.

Martin, LB, Hanson, HE, Hauber, ME, & Ghalambor, CK 2021, "Genes, Environments, and Phenotypic Plasticity in Immunology", *Trends in Immunology*, 42, 3, pp. 198–208. Available from: https://doi.org/10.1016/j.it.2021.01.002.

McCarthy, M, & Birney, E 2021, "Personalized Profiles for Disease Risk Must Capture All Facets of Health",

Nature, 597, 7875, pp. 175–77. Available from: https://doi.org/10.1038/d41586-021-02401-0.

Mestas, J, & Hughes, CCW 2004, "Of Mice and Not Men: Differences between Mouse and Human Immunology", *The Journal of Immunology*, 172, 5, pp. 2731–38. Available from: https://doi.org/10.4049/jimmunol.172.5.2731.

Mo, C, Renoir, T, & Hannan, AJ 2016, "What's Wrong with My Mouse Cage? Methodological Considerations for Modeling Lifestyle Factors and Gene–Environment Interactions in Mice", *Journal of Neuroscience Methods*, 265, May, pp. 99–108. Available from: https://doi.org/10.1016/j.jneumeth.2015.08.008.

Mogil, JS 2017, "Laboratory Environmental Factors and Pain Behavior: The Relevance of Unknown Unknowns to Reproducibility and Translation", *Lab Animal*, 46, 4, pp. 136–41. Available from: https://doi.org/10.1038/laban.1223.

Mumtaz, F, Khan, MI, Zubair, M, & Dehpour, AR 2018, "Neurobiology and Consequences of Social Isolation Stress in Animal Model – A Comprehensive Review", *Biomedicine & Pharmacotherapy*, 105, September, pp. 1205–22. Available from: https://doi.org/10.1016/j.biopha.2018.05.086.

Natterson-Horowitz, B, Aktipis, A, Fox, M, Gluckman, PD, Low, FM, Mace, R, Read, A, Turner, PE, & Blumstein, DT 2023, "The Future of Evolutionary Medicine: Sparking Innovation in Biomedicine and Public Health", *Frontiers in Science*, 1, February, p. 997136. Available from: https://doi.org/10.3389/fsci.2023.997136.

Neville, V, Lind, J, Mendl, E, Cozma, NE, Paul, ES, & Mendl, M 2023, "A Mapping Review of Refinements to Laboratory Rat Housing and Husbandry", *Lab Animal*, 52, 3, pp. 63–74. Available from: https://doi.org/10.1038/s41684-023-01124-1.

Nobile, S, Di Sipio Morgia, C, & Vento, G 2022, "Perinatal Origins of Adult Disease and Opportunities for Health Promotion: A Narrative Review", *Journal of Personalized Medicine*, 12, 2, p. 157. Available from: https://doi.org/10.3390/jpm12020157.

Pallocca, G, Rovida, C, & Leist, M 2022, "On the Usefulness of Animals as a Model System (Part I): Overview of Criteria and Focus on Robustness", *ALTEX*, 39, 2, pp. 347–53. Available from: https://doi.org/10.14573/altex.2203291.

Parker, A, Hobson, L, Bains, R, Wells, S, & Bowl, M 2022, "Investigating Audible and Ultrasonic Noise in Modern Animal Facilities", *F1000Research*, 11, June, p. 651. Available from: https://doi.org/10.12688/f1000research.111170.1.

Perlman, RL 2016, "Mouse Models of Human Disease: An Evolutionary Perspective", *Evolution, Medicine, and Public Health*, April, eow014. Available from: https://doi.org/10.1093/emph/eow014.

Pfaff, J 1974, "Noise as an Environmental Problem in the Animal House", *Laboratory Animals*, 8, 3, pp. 347–54. Available from: https://doi.org/10.1258/002367774780943715.

Pound, P, & Ritskes-Hoitinga, M 2018, "Is It Possible to Overcome Issues of External Validity in Preclinical Animal Research? Why Most Animal Models Are Bound to Fail", *Journal of Translational Medicine*, 16, 1, p. 304. Available from: https://doi.org/10.1186/s12967-018-1678-1.

Randall, DW, Kieswich, J, Swann, J, McCafferty, K, Thiemermann, C, Curtis, M, Hoyles, L, & Yaqoob, MM 2019, "Batch Effect Exerts a Bigger Influence on the Rat Urinary Metabolome and Gut Microbiota than Uraemia: A Cautionary Tale", *Microbiome*, 7, 1, p. 127. Available from: https://doi.org/10.1186/s40168-019-0738-y.

Reese, TA, Bi, K, Kambal, A, Filali-Mouhim, A, Beura, LK, Bürger, MC, Pulendran, B, *et al.* 2016, "Sequential Infection with Common Pathogens Promotes Human-like Immune Gene Expression and Altered Vaccine Response", *Cell Host & Microbe*, 19, 5, pp. 713–19. Available from: https://doi.org/10.1016/j.chom.2016.04.003.

Reimringer, W, Joppich, J, Leidinger, M, Schultealbert, C, Schutze, A, & Conrad, T 2022, "A Modular Sensor System Design for Instrumental Odor Monitoring", in *2022 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN)*, 1–3, IEEE, Aveiro, Portugal. Available from: https://doi.org/10.1109/ISOEN54820.2022.9789609.

Richter, SH 2017, "Systematic Heterogenization for Better Reproducibility in Animal Experimentation", *Lab Animal*, 46, 9, pp. 343–49. Available from: https://doi.org/10.1038/laban.1330.

Richter, SH, Garner, JP, Auer, C, Kunert, J, & Würbel, H 2010, "Systematic Variation Improves Reproducibility of Animal Experiments", *Nature Methods*, 7, 3, pp. 167–68. Available from: https://doi.org/10.1038/nmeth0310-167.

Richter, SH, Garner, JP, & Würbel, H 2009, "Environmental Standardization: Cure or Cause of Poor Reproducibility in Animal Experiments?", *Nature Methods*, 6, 4, pp. 257–61. Available from: https://doi.org/10.1038/nmeth.1312.

Richter, SH, Garner, JP, Zipser, B, Lewejohann, L, Sachser, N, Touma, C, Schindler, B, *et al.* 2011, "Effect of Population Heterogenization on the Reproducibility of Mouse Behavior: A Multi-Laboratory Study", *PloS One*, 6, 1, p. e16461. Available from: https://doi.org/10.1371/journal.pone.0016461.

Robinson, R 1965, *Genetics of the Norway Rat*, Pergamon Press, New York.

Robinson, NB, Krieger, K, Khan, FM, Huffman, W, Chang, M, Naik, A, Yongle, R, *et al.* 2019, "The Current State of Animal Models in Research: A Review", *International Journal of Surgery*, 72, December, pp. 9–13. Available from: https://doi.org/10.1016/j.ijsu.2019.10.015.

Rosshart, SP, Herz, J, Vassallo, BG, Hunter, A, Wall, MK, Badger, JH, McCulloch, JA, *et al.* 2019, "Laboratory Mice Born to Wild Mice Have Natural Microbiota and Model Human Immune Responses", *Science*, 365, 6452, p. eaaw4361. Available from: https://doi.org/10.1126/science.aaw4361.

Ruhlen, RL, Howdeshell, KL, Mao, J, Taylor, JA, Bronson, FH, Newbold, RR, Welshons, WV, & Vom Saal, FS 2008, "Low Phytoestrogen Levels in Feed Increase Fetal Serum Estradiol Resulting in the 'Fetal Estrogenization

Syndrome' and Obesity in CD-1 Mice", *Environmental Health Perspectives*, 116, 3, pp. 322-328. Available from: https://ehp.niehs.nih.gov/doi/10.1289/ehp.10448.

Ruppert, KM, Kline, RJ, & Md Rahman, S 2019, "Past, Present, and Future Perspectives of Environmental DNA (eDNA) Metabarcoding: A Systematic Review in Methods, Monitoring, and Applications of Global eDNA", *Global Ecology and Conservation*, 17, January, p. e00547. Available from: https://doi.org/10.1016/j.gecco.2019.e00547.

Sato, T, & Sato, S 2024, "Circadian Regulation of Metabolism: Commitment to Health and Diseases", *Endocrinology*, 164, 7, p. bqad086. Available from: https://academic.oup.com/endo/article/doi/10.1210/endocr/bqad086/7186648.

Sláma, K, & Williams, CM 1966, "The juvenile hormone. V. The sensitivity of the bug, *Pyrrhocoris apterus*, to a hormonally active factor in American paper-pulp", *The Biological Bulletin*, 130, 2, pp. 235-246. Available from: https://www.journals.uchicago.edu/doi/10.2307/1539700.

Sorge, RE, Martin, LJ, Isbester, KA, Sotocinal, SG, Rosen, S, Tuttle, AH, Wieskopf, JS, *et al.* 2014, "Olfactory Exposure to Males, Including Men, Causes Stress and Related Analgesia in Rodents", *Nature Methods*, 11, 6, pp. 629–32. Available from: https://doi.org/10.1038/nmeth.2935.

Striedter, GF 2022, *Model Systems in Biology: History, Philosophy, and Practical Concerns*, The MIT Press, Cambridge (Massachusetts).

Sultan, SE 2015, *Organism and Environment: Ecological Development, Niche Construction, and Adaptation*, Oxford University Press. Available from: https://doi.org/10.1093/acprof:oso/9780199587070.001.0001.

Sundberg, JP, & Schofield, PN 2018, "Living inside the Box: Environmental Effects on Mouse Models of Human Disease", *Disease Models & Mechanisms*, 11, 10, p. dmm035360. Available from: https://doi.org/10.1242/dmm.035360.

Toth, LA 2015, "The Influence of the Cage Environment on Rodent Physiology and Behavior: Implications for Reproducibility of Pre-Clinical Rodent Research", *Experimental Neurology*. Available from: https://doi.org/10.1016/j.expneurol.2015.04.010.

Turner, JG 2020, "Noise and Vibration in the Vivarium: Recommendations for Developing a Measurement Plan", *Journal of the American Association for Laboratory Animal Science*, 59, 6, pp. 665–72. Available from: https://doi.org/10.30802/AALAS-JAALAS-19-000131.

Uher, R 2014, "Gene-Environment Interactions in Severe Mental Illness", *Frontiers in Psychiatry*, 5, May. Available from: https://doi.org/10.3389/fpsyt.2014.00048.

van der Staay, FJ, Arndt, SS, & Nordquist, RE 2010, "The Standardization-Generalization Dilemma: A Way Out", *Genes, Brain and Behavior*, 9, 8, pp. 849–55. Available from: https://doi.org/10.1111/j.1601-183X.2010.00628.x.

van der Staay, FJ, Arndt, SS, & Nordquist, RE 2009, "Evaluation of Animal Models of Neurobehavioral Disorders", *Behavioral and Brain Functions*, 5, 1, p. 11. Available from: https://doi.org/10.1186/1744-9081-5-11.

Voelkl, B, Altman, NS, Forsman, A, Forstmeier, W, Gurevitch, J, Jaric, I, Karp, NA, *et al.* 2020a, "Author Correction: Reproducibility of Animal Research in Light of Biological Variation", *Nature Reviews Neuroscience*, 21, 7, pp. 394–394. Available from: https://doi.org/10.1038/s41583-020-0326-y.

Voelkl, B, Altman, NS, Forsman, A, Forstmeier, W, Gurevitch, J, Jaric, I, Karp, NA, *et al.* 2020b, "Reproducibility of Animal Research in Light of Biological Variation", *Nature Reviews Neuroscience*, 21, 7, pp. 384–93. Available from: https://doi.org/10.1038/s41583-020-0313-3.

Voelkl, B, & Würbel, H 2016, "Reproducibility Crisis: Are We Ignoring Reaction Norms?", *Trends in Pharmacological Sciences*, 37, 7, pp. 509–10. Available from: https://doi.org/10.1016/j.tips.2016.05.003.

Voelkl, B, & Würbel, H 2021, "A Reaction Norm Perspective on Reproducibility", *Theory in Biosciences = Theorie in Den Biowissenschaften*, 140, 2, pp. 169–76. Available from: https://doi.org/10.1007/s12064-021-00340-y.

Webb, AE, Gerek, NZ, Morgan, CC, Walsh, TA, Loscher, CE, Edwards, SV, & O'Connell, MJ 2015, "Adaptive Evolution as a Predictor of Species-Specific Innate Immune Response", *Molecular Biology and Evolution*, 32, 7, pp. 1717–29. Available from: https://doi.org/10.1093/molbev/msv051.

Weber, EM, Ahloy Dallaire, J, Gaskill, BN, Pritchett-Corning, KR, & Garner, JP 2017, "Aggression in Group-Housed Laboratory Mice: Why Can't We Solve the Problem?", *Lab Animal*, 46, 4, pp. 157–61. Available from: https://doi.org/10.1038/laban.1219.

West-Eberhard, MJ 2003. *Developmental Plasticity and Evolution*, Oxford University Press, New York.

Witjes, VM, Boleij, A, & Halffman, W 2020, "Reducing versus Embracing Variation as Strategies for Reproducibility: The Microbiome of Laboratory Mice", *Animals*, 10, 12, p. 2415. Available from: https://doi.org/10.3390/ani10122415.

Würbel, H 2000, "Behaviour and the Standardization Fallacy", *Nature Genetics*, 26, 3, pp. 263–263. Available from: https://doi.org/10.1038/81541.

Würbel, H 2017, "More than 3Rs: The Importance of Scientific Validity for Harm-Benefit Analysis of Animal Research", *Lab Animal*, 46, 4, pp. 164–66. Available from: https://doi.org/10.1038/laban.1220.

Yim, YS, Park, A, Berrios, J, Lafourcade, M, Pascual, LM, Soares, N, Kim, JY, Kim, S, Kim, H, Waisman, A, Littman, DR, Wickersham, IR, Harnett, MT, Huh, JR, & Choi, GB 2017, "Reversing Behavioral Abnormalities in Mice Exposed to Maternal Inflammation", *Nature*, 549, 7673, pp. 482-487. Available from: https://www.nature.com/articles/nature23909.

Yong, E 2022, *An Immense World: How Animal Senses Reveal the Hidden Realms around Us*, First edition, Random House, New York.

SAPIENZA UNIVERSITÀ DI ROMA

My Personal View on Science

Vol. 7, No. 1 (2024)
ISSN: 2532-5876

DOI: 10.13133/2532-5876/18557

Science Across the Abyss: Knitting Bridges with Butterflies

Caroline Bacquet Perez[a*] *& Jorge Batres*[a]

[a] *Universidad Regional Amazónica Ikiam, Tena, Ecuador*

***Corresponding author:** Caroline Bacquet Perez, Email: caroline.bacquet@ikiam.edu.ec

Abstract

Science is built on the pursuit of answers to fundamental questions and the constant expansion of our understanding of the world around us. However, this effort has not been without challenges and inequalities. This article critically examines the issue of diversity in science and the notable disparities that persist in global scientific knowledge. Throughout history, the contributions of scientists from diverse regions and cultures have been pivotal to scientific advancement. Nevertheless, significant gaps in terms of access, funding, and recognition in the global scientific community still endure. We use the concept of the "abyss," as a metaphor for the disparities in scientific practices across diverse regions of the world within the context of globalization. We seek to shed light on how the abyss influences the very essence of scientific inquiry, ranging from disparities in access to knowledge to the limitations imposed by technology and resources. This article addresses how socioeconomic, gender, and geographical disparities impact who has the opportunity to engage in and lead scientific research. The decolonization of science and the incorporation of indigenous and local perspectives in research are highlighted as crucial ways to address these disparities. Additionally, the concept of participative science is explored as an inclusive approach that allows diverse communities to take part in scientific research. Ultimately, this exploration of diversity in science and disparities in scientific knowledge seeks to inspire deeper reflection on how we can work together to ensure that science becomes a truly global and representative endeavor, enriched by a multitude of perspectives and the collaboration of people from all corners of the world.

Citation: Bacquet Perez, C 2024, "Science Across the Abyss: Knitting Bridges with Butterflies", *Organisms: Journal of Biological Sciences*, vol. 7, no. 1, pp. 41-45. DOI: https://doi.org/10.13133/2532-5876/18557

1. The Abyss

The concept of the "abyss" that underpins this article is by no means novel and has been previously used by scholars such as Boaventura de Souza Santos (2014). According to him, the Western world is organized around a divide that separates us: one side is visible, while the other dissolves into obscurity as concealed reality. This rift finds its genesis in colonial processes, which sowed the seeds of disparity between the colonizers and the colonized, resulting in contradictions concerning rights and emancipation in the realm of the former, and appropriation and violence in the sphere of the latter. This dichotomy is also mirrored in the realm of science, where the "truth" on the far side of the abyss is in perennial contention, and where scientific inquiry distances itself from others forms of knowledge prevalent on this side of the abyss. In essence, the concept of the "abyss" highlights the stark differences in circumstances and resources that make the practice of science distinct on each side of the divide, underscoring the social and geographical disparities in scientific knowledge production and access.

In this article, I will use my personal narrative to illuminate how the historical experiences of our ancestors can reverberate in our own academic and scientific journeys, giving us power to act and direction to our effort. Born in Chile in a low income

household during Pinochet's dictature, I struggled financially to complete undergraduate studies as a biochemist. Afterwards, several scholarships allowed me to earn a doctorate in France, and to conduct postdoctoral research in Hungary and England. In 2014, I came back to South America, and joined the faculty at the University Regional Amazónica Ikiam, in Ecuador. My scientific journey took me back and forth across the abyss, through different continents, countries and realities, allowing me to obtain a critical perspective of the realities and inequalities of the scientific community, how they span far and wide in space and time, and how they need to be challenged and changed.

2. Science Over the Abyss

Why do we engage in science? What are the values we recognize in it? These are fundamental questions that also deeply rooted in the social discourse about science. In many ways, science possesses an emotional quality, a subjectivity of perception that is not always recognized or acknowledged. On the other hand, scientific activity has a very real and concrete impact in society, affecting the lives of both scientists and non-scientists alike. The abyss, in this context, represents the divide in the distribution of scientific knowledge, where the interventions of science tend to serve those with resources to access it. The "why" of science often raises questions about its social equity and justice, particularly concerning accessibility to knowledge. These issues are not new, as seen in the reflections of Jean-Jacques Rousseau in 1750, where he expressed skepticism and a deep concern about its impact on human society. Rousseau believed that the advance of science and reason had led to the corruption of human nature and society. He argued that the pursuit of knowledge and technological progress often alienated people from their natural state of innocence and simplicity. Rousseau was critical of the Enlightenment's emphasis on reason and science, contending that it had contributed to inequality, social unrest, and the erosion of authentic human relationships. Instead, he advocated for a return to a more harmonious and natural way of life, where science and reason were tempered by a profound respect for the human heart and its emotions. These views are still, in different forms, being discussed today, highlighting epistemological challenges that have persisted for almost 300 years.

According to the philosopher Boaventura de Sousa Santos, the abyss represents the idea of a profound epistemological and ontological divide between different forms of knowledge. He highlights the disparity between the conventional, dominant scientific knowledge, often associated with the Global North, and the marginalized, indigenous, or local knowledge systems of the Global South. This divide, or abyss, underscores the unequal power relations and hierarchies in knowledge production. Addressing this gap is essential for creating a more inclusive and holistic approach to science, acknowledging the diversity of knowledge sources, and fostering a deeper understanding of complex global challenges like climate change, social injustice, and sustainability. Bridging this abyss requires recognizing the validity of multiple knowledge systems and promoting a more equitable distribution of epistemic authority in the scientific discourse. In this regard, I am using the butterflies in this article not only as a model organism for science, but also as emblematic entities symbolizing the ability to traverse the abyss and forge connections between disparate realities.

Irrespective of our location relative to the abyss, we grapple with comparable challenges in our scientific pursuits. The ever-increasing competition for research funding, coupled with limited resources, places immense pressure on researchers. Additionally, issues of reproducibility, inclusivity, and diversity in science persist, highlighting global systemic flaws. Balancing the demands of academia with personal life and mental health concerns adds to the difficulties scientists face. In the modern scientific landscape, researchers often find themselves in power relationships with institutions and funding bodies that may prioritize profit or specific agendas over scientific inquiry. This dynamic can lead to conflicts of interest and compromises in scientific integrity. Additionally, the hierarchical structure within research labs can create power imbalances between principal investigators and junior researchers, affecting decision-making, credit allocation, and opportunities for advancement. Hegemonic academic communities have traditionally favored uniformity and exemplary performance, leaving those who struggle to cope feeling stressed and inadequate.

Scientific research depends on funding, requiring very expensive equipment and materials that are difficult to obtain. From a geopolitical perspective, because scientific knowledge is not distributed in a socially equitable or just manner, scientific interventions in the real world will tend to serve the groups that have the resources to access this knowledge. Where are these groups? The geographical distribution of Nobel Prizes and scientific publications reveals stark inequalities in the

global scientific landscape. Historically, Nobel Prizes have disproportionately favored Western countries, with European nations and the United States claiming the majority of laureates. This inequality reflects disparities in research funding, infrastructure, and opportunities, reinforcing a perception of Western scientific dominance. Interestingly, there is also a significant gender disparity in Nobel Prize awards, with male laureates outnumbering female laureates by a substantial margin. Similarly, scientific publications exhibit geographical bias, with a significant concentration of research output in the Global North, while regions like Africa often remain underrepresented. These disparities in recognition and dissemination of scientific achievements highlight systemic challenges that hinder global collaboration and limit diverse perspectives in shaping the future of science. Addressing these inequalities is crucial for fostering a more inclusive and equitable scientific community.

Poverty, lack of education, and limited research efforts in regions rich in biodiversity are intricately interconnected challenges that perpetuate a cycle of environmental degradation and human suffering. In these biodiverse areas, many communities struggle with limited access to quality education, which hampers their ability to break free from the grip of poverty. The lack of educational opportunities often leads to unsustainable practices that exploit natural resources, further depleting the very biodiversity upon which these communities depend for their livelihoods. Simultaneously, insufficient research and scientific understanding of these ecosystems hinder the development of effective conservation strategies and sustainable resource management practices. Therefore, addressing poverty and promoting education in biodiversity-rich regions is not only a matter of human development but also a crucial step towards preserving the invaluable natural heritage that these regions hold. Effective policies and investments that target poverty alleviation, education, and scientific research can help to create a harmonious balance between human well-being and the conservation of biodiversity.

3. The Stories of Our Grandmothers

Human beings are intricate and diverse, and this as also true for scientists. Human minds are complex ecosystems of ideas, where all kinds of influences coexist. Therefore, not only rationality, but also emotions, dreams, intuition and personal narratives are integral aspects of the scientific work. As Eduardo Galeano once said, *"we are made of stories"*, our lives are narratives, woven from the threads of our experiences, cultures, and ancestors. Stories are the vessels through which we transmit knowledge, wisdom, and emotion across generations, constructing the very fabric of our identities, shaping our beliefs, values, and understanding of the world. Through storytelling we build bridges of understanding in a diverse and complex world. The stories of grandmothers hold profound importance for women scientists. These narratives, often laden with unfulfilled dreams and aspirations due to historical gender limitations, serve as a poignant reminder of the progress achieved and the journey that lies ahead. They embody resilience, determination, and a spirit of overcoming adversity, inspiring the next generation of women in science to break barriers and shatter stereotypes. Grandmothers' stories instill a sense of purpose, fostering a deep connection between past struggles and present opportunities. They empower young female scientists to not only carry forward the torch of knowledge but also to redefine the boundaries of what is achievable, ultimately contributing to a more equitable and diverse scientific landscape. I am the second woman in my whole family who could graduate from university (the first being my mother). Looking backwards, the women before me were clever, resourceful, motivated, and enjoyed learning. Their context, however, determined that once and again they were denied the opportunity to complete any studies. Growing up, these stories made me angry, but also hardened my determination to succeed in academia, not only for myself, but also as a way of achieving what my grandmothers were systematically denied. That moved me to pursue a PhD scholarship in France, and further postdoctoral stays in Hungary and Oxford. After this, I was expected to find a position in the scientific world; however, the question remained about where to establish myself and in what capacity. To me, at that moment, the academic environment seemed brutal, cold and unfulfilling. There was a huge gap between my past and what seemed to be my future as a scientist, but also, between what I considered to be meaningful and important and the alternatives laid out for me as possible jobs. More importantly, I did not see how my work as a scientist would have any impact in changing the reality of many children, and specially girls, still struggling to get an education, in the same way I did, and my mother and grandmothers did. In hindsight, I can see now that I wanted to save them, through a huge gap in space and time, and I did not know how.

4. Persisting in Science: the Power of Butterflies

From Europe, my search for a meaningful scientific project took me to the Amazon region of Ecuador, where I became involved in the establishment of Ikiam University, in 2014. Situated on the periphery of the Colonso Chalupas Biological Reserve, this university was planned to provide quality scientific education to an area previously underserved in terms of higher education. The project placed paramount importance on environmental preservation, research, and the celebration of cultural diversity. However, the university starting from scratch meant spending a long time without the basic infrastructure, equipment and materials necessary to do research. This lead some of the teachers to explore unconventional avenues within academia, and I chose to apply the art of knitting for scientific education. Weaving and textile works are practices deeply rooted in indigenous communities, with textiles mirroring their connection to the natural world. As I child, I learned crochet and knitting from my grandmother, and later in life I came back to it every time I felt sad or lost. At Ikiam, knitting emerged as a powerful tool for initiating dialogues and comprehending the interplay between indigenous communities and the local fauna. This undertaking not only empowered Kichwa women but also promoted the conservation of indigenous wildlife. Through these interactions, we succeeded in constructing a bridge between the academic sphere and the local community, surmounting the traditional barriers of titles and social hierarchies.

Once we were able to acquire some equipment and supplies, I decided to establish my research line using Heliconius butterflies. I made this choice due to the availability of genetic data, their abundance in the Amazon ecosystems, and their susceptibility to climate change. Our initial focus centered on the expression of heat shock proteins, an area of research that had remained relatively unexplored in Neotropical insects. The initial findings have unveiled potential avenues for future research, particularly concerning epigenetic responses and gene expression in response to environmental fluctuations. In reality, starting from scratch also meant I could have studied almost anything, I could have chosen any model from the great variety and diversity of species that can be found in the forest around the university. But there was another reason to chose butterflies: they can fly over the abyss. In the realm of biological symbolism, butterflies have long captivated the human imagination. Beyond their biological significance as pollinators and indicators of environmental health, butterflies have assumed a profound symbolic meaning across cultures and throughout history. Representing transformation, rebirth, and spiritual evolution, they are symbols of hope, growth, and the cyclical nature of existence. In my story, butterflies allowed me to establish important and meaningful collaborations with researchers from all over the world, which has brought forth student and academic exchanges, establishment of joint projects with international funding and organization of international scientific conferences and workshops.

Our narrative serves as a testament to the fact that, despite the obstacles and the disparities in geography and culture, science can function as a bridge connecting ostensibly distant realities. Through collaboration, mutual respect, and a willingness to embrace diverse forms of knowledge, science possesses the capacity to flourish anywhere. Science, then, emerges as a potent instrument for forging connections, spanning chasms, and fostering a world that is not only more inclusive but also more sustainable.

5. Conclusion

Our experiences on the Latin American side of the abyss have taught us that science recognizes no insurmountable boundaries. While challenges persist, ranging from resource limitations to inequities in access to knowledge, opportunities for collaboration, diversity, and inclusion are equally abundant. Science can function as a unifying force capable of transcending geographical and cultural divides. For that, our personal stories can be much more important than it has been previously thought. The world needs more science, but it needs better science, and this requires the involvement of everyone. It is not just about solving scientific questions; it is also about addressing the great challenges and dilemmas of our society. We live in a complex world, and science must address this complexity. Science needs to be open to different perspectives and needs to consider the ethical, social, and cultural aspects of its work. How can we bridge the gap between science and society? How can we make science more inclusive and participatory? How can we ensure that scientific knowledge benefits all of humanity? These are challenging questions, and there are no easy answers. But there are steps that can be taken to move in the right direction.

We need to move away from traditional, top-down, and hierarchical approaches to science and embrace more collaborative and interdisciplinary approaches that involve a diverse range of voices and perspectives in the scientific process, including those from marginalized and underrepresented communities. Bridging the gap between science and society is a complex and ongoing challenge. But it is a challenge that we must embrace if we are to address the pressing issues facing our world today.

SAPIENZA UNIVERSITÀ DI ROMA

Books

Vol. 7, No. 1 (2024)
ISSN: 2532-5876

DOI: 10.13133/2532-5876/18556

Organicist Perspectives in Newly Published Biology Textbooks

Carlos Sonnenschein[a*]

[a] *Tufts University School of Medicine, Boston, MA USA*

***Corresponding author:** Carlos Sonnenschein, Email: carlos.sonnenschein@tufts.edu

Citation: Sonnenschein, C 2024, "Organicist Perspectives in Newly Published Biology Textbooks", *Organisms: Journal of Biological Sciences*, vol. 7, no. 1, pp. 47-49. DOI: https://doi.org/10.13133/2532-5876/18556

At the beginning of the current century, in 2004 to be precise, Carl Woese, an outstanding biologist, wrote, "A society that permits biology to become an engineering discipline, that allows that science to slip into the role of changing the living world without trying to understand it, is a danger to itself. Modern society knows that it desperately needs to learn how to live in harmony with the biosphere. Today more than ever we are in need of a science of biology that helps us to do this, shows the way. An engineering biology might still show us to get there; it just doesn't know where 'there' is". A few pages later, he advised us "Let's stop looking at the organism as a molecular machine". Separately, buried in the Preface of his 1962 influential book, *The Structure of Scientific Revolutions*, Thomas Kuhn wrote "If I am right that each scientific revolution alters the historical perspective of the community that experiences it, then the change of perspective should affect the structure of postrevolutionary textbooks and research publications" (University of Chicago Press, 2nd ed, p. IX). And then he added, "One such effect… ought to be studied as a possible index to the occurrence of revolutions". Leaving aside for the moment the skeptical views of philosophers of science and of biology about paradigmatic changes and scientific revolutions, it could be acknowledged today that Thomas Kuhn was, indeed, correct in justifying substantial changes in the content of textbooks and publications that followed meaningful scientific revolutions.

The shortcomings and deficiencies attributed to the dominance of a reductionist agenda during the second half of the 20th century (the "century of the gene" as Evelyn Fox-Keller characterized it) and even the first decades of the current one on experimental and theoretical biology are increasing steadily. Despite the monumental accumulation of data generated by molecular biology technologies, the "there" Woese was referring to is continuously postponed to an indefinite future by today's scientific establishment. During the second half of the 20th century, the "next ten years" used to be their timely goal; now a days, the 'there' is being transferred to the "next generations of scientists". As a result of these confusions, a consensus is building among empirical and theoretical biologists that the onslaught of reductionism must be openly challenged by a coherent, effective strategy representing worthy alternatives offered by organicist/holistic approaches. Briefly, organicism aims at linking developmental biology and embryology with evolutionary biology from a perspective where *organisms* (both unicellular and multicellular ones) become the center of analysis and synthesis.

Organicism has a history. Since its inception, last century contributions, from among others, by J.S. Haldane, J. von Uexküll, J.H. Woodger, C. Waddington, L. von Bertalanffy, R.C. Lewontin, and S.J. Gould, organicism has been overshadowed by a greedy reductionist approach to biology that benefited from an entrepreneurial strategy that motivated many in its leadership. One of the central tools to that strategy has been the availability of textbooks for undergraduate and graduate education that successfully promoted the message that molecular biology would successfully resolve

all kinds of biological and biomedical puzzles. Textbooks like Molecular Biology of the Cell, and others like it, while educating generations of young minds in the complexity of cell structures and biochemical pathways, both implicitly and explicitly, were instrumental in promoting a reductionist gospel when interpreting biological phenomena. Sociopolitical factors nested at the academic level during second half of the last century helped as well in establishing molecular biology as the law of the land in biological research.

Admittedly, plenty of historical precedents have established that paradigmatic change is not necessarily welcomed in academic circles because of its deleterious impact on the status of scientists and institutions that prefer instead graded, incremental, easily manageable changes.

In this regard, the 2023-2024 years will probably enter the annals of the history of the biological sciences as an important turnaround milestone from the reductionist perception that has dominated academic epistemological thinking in biology in recent decades. Specifically, here are the good news. A book authored by an experienced contributor and editor of Nature Magazine, Philip Ball, entitled *How Life Works* (University of Chicago Press, 2024) and provocatively subtitled *A User's Guide to the New Biology* and another book entitled *Properties of Life* (The MIT Press, 2023) by a highly regarded philosopher of biology, Bernd Rosslenbroich, Head of the Institute of Evolutionary Biology and Morphology at Witten/Herdecke University, in Germany, are already available in libraries. In my opinion, both books qualify as worthy representatives of the organicist/holistic tradition. Philip Ball's book has already been glowingly reviewed by the admirable physiologist Denis Noble in Nature Magazine; Noble considered it as "a must-read for user's guide for biologists and non-biologists alike…". Ball's reference for a need to adopt alternative premises to evaluate the "new biology" of the 21st century is highly significant.

Rosslenbroich's book equally qualifies in my view as a "must-read opus" while being dedicated, instead, to a more scholarly qualified readership composed of graduate students, post doctorates as well as for academic biology professors. In five well-referenced Chapters, the book accurately and convincingly compels the readership to consider, even adopt, a novel approach to evaluate philosophical trends and empirical evidence about life at large. As explicitly stated in its subtitle, namely, *Toward a Theory of Organismic Biology*, Rosslenbroich concludes that the time is ripe to move Biology into an elevated stage of rigorous analysis and integration comparable to that adopted by the so-called Exact Sciences. In his own words, Rosslenbroich states that "My impression – and also my thesis – is that biology today develops, or should develop, toward such a synthesis concerning knowledge from analytical research on the one hand and an organismic understanding of life on the other" (p. 67). He further concludes that "the extensive knowledge of details in structures, functions and genetic processes provides a new opportunity to understand integrative and systemic functions. The chance for an organismic conception of life on a scientific basis has never been as good as today". This is the central message of the book. For this aim to be achieved will require the formulation and testing of theories that when proven wrong or defective be either abandoned or modified, a practice that have been ignored throughout the last half century under the epistemological and financial influence of a reductionist approach to Biology. Rosslenbroich is therefore attempting to synthesize the modern insights of Biology with an organismic conception of life.

Rosslenbroich's detailed description and scholarly treatment of a variety of biological topics qualifies as an almost-textbook for anyone interested in epistemological, historical and even sociological approaches to basic subjects in the biological sciences. Young and old biologists interested in an updated, realistic view of how organicism advantageously addresses and offer solutions to the many controversial issues enlivening basic biological phenomena will have their desire fulfilled. Rosslenbroich adopts a hierarchical interpretation/approach to levels of biological organization in an accessible language.

The theoretical and empirical contributions by Lamarck, Schwann and Schleiden, Mendel, Darwin, Virchow and their followers during the 19th century were adopted by researchers who from the very beginning of the 20th century expanded knowledge within Biology, with especial emphasis on phenomena happening at the cellular level of biological complexity. The explicit rationale for this choice of target was built around the notion that to explain biological phenomena it had to be done from the bottom-up strategy. This favored the growth of knowledge in empirical disciplines like biochemistry, genetics, and cellular structure. Despite the generous, almost extravagant, human-power and financial investments that reductionist approaches enjoyed for over half a century, their shortcomings have become obvious and, therefore, they will require a rearrangement of

funding priorities. It is time for organicism to become, again, a welcomed participant in the theoretical and empirical solutions in the biological scene.

The reductionist approach dominating theoretical and empirical biology was almost unanimous in the 20th century. Dissenting voices proposing the primacy of the organism as an alternative to the cellular, genetic, and molecular (gene-centered) explanations of phenotypes based on a one-sided reductionist perspective were offered by inspirational biologists especially in the UK and continental Europe (von Bertalanffy and Paul Weiss were the most prominent among them). Unfortunately, in some cases due to political factors (which are also a part of the constrains affecting biological development), those worthy organicist alternatives to the reductionist approach were summarily dismissed or de-emphasized. Recently, well-deserved credit has been given among others to the views of Conrad Waddington, who dealt in basic biological aspects of development in multicellular organisms in the 1940s and 50s, and to David Smithers who as early as in the 1960s presented solid arguments against what he called *cytologism* when dealing with the cancer puzzle. It is expected that professional historians of biology will soon provide an unbiased, detailed interpretation of the monumental epistemological mistakes made by leading entrepreneurial, empirical cancer researchers who influenced funders to adopt a narrow, for profit-based perspective when dealing with scientific and public health policy.

In sum, the availability of two well-conceived and written publications about the biological sciences based on the once ignored organicist perspective should allow the emergence of a new breed of scholars who would provide a more balanced view of the complexity of hierarchical levels of biological organization.

Organisms

ISSN: 2532-5876