ONLINE COLLECTION

The Gene: An Appraisal

by

Keith Baverstock

Edited by

Denis Noble

Introduction

Online issue of the Journal with a Focus Article, The Gene: An Appraisal, by Keith Baverstock. The issue contains, in addition an Editorial by Denis Noble, seven commentaries by other gene experts, and a response to criticisms by Baverstock. Baverstock argues that genes should not be thought of as regulating cellular production; instead, the cellular phenotype, a gene product interactome, regulates the cell, itself, and expresses the cell's characteristics, suggesting that an appropriate metaphor is a brain. There is no one-way process from genes to phenotype as the current molecular genetic paradigm envisages. The contemporary error dates a century back to Wilhelm Johannsen's proposed 'genotype conception' which underpins population genetics and heredity today. In fact, the prior Francis Galton's statistically based ancestral law of inheritance is closer to the truth. The Editorial summarizes the commentaries, all supporting the main thrust of Baverstock's case, whereas, from the many invited to comment no response was received from senior scientists using the Genome-Wide Association (GWA) methodology. Since the articles in this issue of the Journal were published, the case has been supported by the discovery that polygenic scores based on GWA fail to predict major diseases, including cardiovascular disease and cancer. Despite spending some US$8 billion by NIMH, no gene responsible for schizophrenia has been identified either. The Editorial speaks out forcefully on the disturbing silence from those leading GWA studies, pointing to the large amount of funding consigned and the very little delivered, of clinical and public health value.

Denis Noble

EDITORIAL

Editorial for Online Collection —
The Gene: An Appraisal

by

Denis Noble

Progress in Biophysics and Molecular Biology 187 (2024) 1–4

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Editorial for online collection — The gene: An appraisal

1. Introduction

This editorial introduces an on-line Special Collection of articles centred on an appraisal of the idea of "gene" by one of the journal Board members, Keith Baverstock.

The Collection is timely, since it has come together soon after publication of an important and rigorous test of the predictive utility of polygenic scores, showing a disappointing predictive utility (Hingorani et al., 2023).

I will return to the significance of that study after introducing the Collection.

1.1. A one-sided complete silence from defenders of genome sequencing

Keith Baverstock's article, "The Gene: An Appraisal" (Baverstock, 2023a), is important since it argues that genome sequencing has generally found very low association scores for most genes in relation to the main multifactorial diseases that are resistant to a gene-centric analysis. Inevitably, that fact is also connected with the second fact, that very few strategies for curing such diseases have emerged from the results of genome sequencing. This is so despite the promise that, within a decade, such cures would automatically emerge from the human genome project (Collins, 1999). It is hard to see how anyone can fault those two conclusions. Yet, as I will now explain, the journal has received no answer whatsoever from the genomics community leaders.

An Editor-in-Chief of a journal is in a privileged position. The benefits include viewing the scientific community and its arguments through an attempt, at least, to stand back above those arguments in order to carry conviction as a relatively neutral judge. The buck stops here. In consequence an Editor is often faced with difficult decisions.

But, sometimes, remaining neutral is almost impossible. In writing this Editorial I cannot remain neutral. The reason is fundamental to any journal that prides itself on encouraging live and sometimes fierce debate. The Editor's role is to try to get opinions and arguments across the spectrum of views and interpretations.

There is no lack of such a spectrum in the case of debate about genes and their roles and effects. Opinions vary all the way from "genes for everything" to "genes for nothing" (Ball, 2014); from "genes created us body and mind" (Dawkins, 1976, p. 26) to "genes are followers of phenotype changes" (Schwander & Weimar, 2011; Noble and Phillips, 2023).

When I received Keith Baverstock's article I therefore acted as any Editor should: take a long view, solicit reactions from a wide spectrum of known opinion and expertise, then sit back and wait for the debate to happen. I therefore invited commentaries from around 15 scientists who I judged would be broadly favourable to the article, while obviously having their own criticisms from their particular standpoint. I also invited around 15 who, from their previous work, would be expected to be strongly opposed to the main thrust of the article, and some who might be in between. An overall total of 45 were invited.

Two years later, in response, the journal has received 7 articles from the first and third group of invitees, but *none whatsoever* from the second. Those invitees included leading geneticists and genomics people. Why the silence? Surely, the responsibility for the huge investments of time, money and people in genome-wide association research carries with it a responsibility for openness to criticism and questions since that funding is provided by society itself, via governments, businesses or charities. Furthermore, in the case of genetics and genomics research the stakes are very high indeed. These areas of medical research receive the lion's share of funding. Why then, over two decades since the first publications on sequencing human, and other, complete genomes, do we see so few health benefits that could begin to justify the huge investment that has been made?

Faced with a crisis of ill-health amongst the growing populations of the elderly, with multifactorial diseases notoriously unyielding to genetic interpretations, why do we continue to insist that genomics research holds the answer when the association scores with such diseases are often so abysmally low? Anyway, the association scores themselves are not a correct indication of the quantitative causal role of genes in those disease states since physiological networks are good at buffering changes at the molecular level (Noble and Hunter, 2020). As that article states, quantitative physiology is ready to come to the rescue of genomics research. Physiology measures causation, and it is often very different from association. Even a zero association score cannot prove no causal role. That is the major difficulty with Genome Wide Association scores and it has not been addressed.

I believe this one-sided silence from those responsible for managing the huge investments involved reflects badly on the scientific community. It is not in the long-term interest of science itself, for science flourishes on active debate and engagement. In the end, large scale mistakes in prioritising research will become evident. It is better that we should learn what mistakes have been made earlier rather than later. Genome-wide association research has given us masses of data but is presented, even by its own advocates, as independent of theory: "A hypothesis is a liability"! (Yanai and Lercher, 2020, 2021). On the contrary, without a theoretical guide to what to expect, we have no way to judge the significance of a piece of data. Accumulating data without interpretation is scientific 'stamp collecting' (Felin et al., 2021a, 2021b), a risky abandonment of thought in biological research.

All the commentaries on "The Gene: an Appraisal" we have received

https://doi.org/10.1016/j.pbiomolbio.2024.01.001

Available online 3 January 2024
0079-6107/© 2024 Elsevier Ltd. All rights reserved.

and published represent a strong endorsement for the journal in encouraging debate on the issue. For, while the commentaries are broadly supportive of the original article, they all make valuable points, supportive or critical, that extend what I take to be Baverstock's intentions.

As Editor in Chief for this collection of articles, I have now waited for nearly three years since the original invitations were sent. That is already too long. The journal is therefore proceeding to publish the Collection as it now is. The debate remains one-sided, but that is not the fault of this journal.

This Editorial will be more detailed than usual since the collection of articles is not being published as a separate volume but rather as an online collection. It will help readers of the collection if I summarize the main points of the commentaries. I begin by summarising the commentaries, all of which are already available online.

1.2. The genetic control paradigm

McKenna et al. (2022) write under the title "The genetic control paradigm in biology: What we say, and what we are entitled to mean" which clearly identifies the thrust of the commentary. How, in any process that consists in an interaction, in this case between phenotypes and genotypes, can we say that one is in control of the other? One way to answer that question is to note that the environment can never act *directly* on genes in an adaptive way, since the only direct effect of the environment on DNA is radiation and similar damage, causing breakage and the need for repair. All adaptive change must, surely, therefore arise via the phenotype which is in continuous interaction with the environment, including that of other organisms. Furthermore:

> "mutations can have a very large effect at the molecular level, but that effect is cancelled out or buffered by evolved homeostatic and robustness mechanisms in biochemistry, development and physiology. Cryptic genetic variation is most easily detected, documented and quantified in human diseases where genes that are characterised as risk factors for a disease by genetic epidemiologists have been well studied (Nijhout et al., 2015, 2018). Cryptic genetic variation will not be 'seen' by selection until a mutation or an environmental signal disrupts one of the stabilising mechanisms." (P. 90)

This is precisely the process that my research team found over 30 years ago in the case of the pacemaker rhythm of the heart (Noble, 2021). There is overwhelming evidence that most regulatory physiological networks are robust in this way. In the few cases where they are not, the outcome is one or other of the rare outlier genetic diseases. I cannot understand why there should still be any doubt about this. Yet, genomics research still looks for summing up all the small association scores to estimate overall genetic causation, usually called the polygenic risk score. It cannot be stated too firmly, this is nonsense. In complex interactive systems the effects are necessarily not additive.

The authors are equally firm on the use of phrases like "genetic programs":

> "authors who appeal to genetic control, programs, and blueprints seldom if ever define what exactly they mean by these terms." (P 90)

This is necessarily so, since no-one has ever identified the equivalent of IF-THEN-ELSE clauses in genome sequences. It is more than 40 years since Monod and Jacob coined the phrase "genetic program". It is high time the phrase should be relegated as highly misleading. All the important conditional processes in biological systems occur at higher levels of organisation than the genome.

The authors identify one of the key misunderstandings as the search for a master controller of what is happening:

> "Looking for a primitive causal controller in an automobile is a fool's errand. Cars are mechanical systems made-up of mutually dependent parts. Various components might be more or less important, but none are truly in control of the vehicle's overall functionality. Something similar can be said of organisms. Their genes, or more properly, their gene products, play a role in many important processes, but they are not in control of anything." (P. 91)

This is the sense in which some go so far as to say that there are no "genes *for* anything" and it is reflected in the modern view from genomics research favouring what is called the omnigenic hypothesis (Boyle et al., 2017). What is needed now is that genomics research takes seriously the need to understand the regulatory networks in organisms that enable genes to have any effects at all.

1.3. Interpretation of Johansson's "gene"

Nils Roll-Hansen (2022) writes under the title "A special role for the genotype". He declares his difference from Baverstock very early in his article.

> "Contrary to Baverstock I hold that even if the gene has become blurred the distinction between genotype and phenotype remains a foundation stone of genetics." (P. 82)

He then proceeds to a deeply scholarly analysis of Johanssen's ideas on the genotype-phenotype duality, taking issue with Baverstock on several aspects of his article, the essence of which is that Johansson's work cannot be accurately understood simply from his publication in German in 1909. He concludes his analysis of Johanssen:

> "Johanssen was by no means alone in his criticism of chromosome theory and Neo-Darwinism. He shared a holistic approach typical of German genetics and evolutionary studies in the 1920s and 1930s." (P 87)

His concluding section is the most critical of the commentaries on Baverstock, where he rejects the idea that the "zygote knows" what it will develop into independently of its genotype."

1.4. Phenotype knowledge of what?

Ken Richardson (2021) responds under the title "Genes and Knowledge", where he asks the question "knowledge of what?" and promises some "grounds for optimism".

The key to his contributions lies in the question:

> "What rules — or "rules of engagement" as KB calls them — might enable organisms to anticipate rapidly changing, constantly novel states in dynamically complex environments?" (P. 13)

A second key is

> "Robert Rosen (mentioned by KB) offered a rigorous mathematical treatise on such "anticipatory systems" in biology. Properties emerge from the deep statistical relations in networks that transcend those of independent components …..living things don't just change their "state" in response to certain conditions; they also change the "rules" by which they do so." (p. 13)

Life then is rather like a chess game in which the players change the rules. That includes the definition of how living organisms became systems:

> "When environmental change wrought on one component induced compensatory changes in another, or even changes that anticipated, nullified or amplified a future change, they became systems. System integrity over continual environmental change , at least for some period of time, is what most distinguished them from non-living molecular mixtures." (P. 14)

This is one of the most helpful definition of "systems" that I have come across. The consequence is that "living forms existed before genes … …they must have been "learning", knowledge-forming, networks

from the start ….phenotypes arrived before genes."

> "The egg includes "transcription factors, promoters , enhancers, and a rich milieu of RNAs , other proteins, fats, sugars, vitamins, metal salts, and so on. Then the sperm adds its own cargo, as well as some polarity to the ovum. In addition, epigenetic markers on offspring's genes, influencing how those gens should be used on the basis of parental experience."

I believe that statement should act as a stark warning to gene-centred theories of biology. The sheer complexity of egg and sperm need to be understood. So also do Richardson's comments on metabolism:

> "The same logic applies to the metabolism of the cell … the program of instructions comes, not from the nucleus, but rather from the metabolic structures of the host cytoplasm." (P. 14)

He concludes on a high note:
That understanding completely reverses Dawkins's prioritisation of (stable) genes over (changeable) phenotypes."

1.5. Cellular and organismal agency

Frantisek Baluska and Arthur Reber (2021) agree with Baverstock's article and suggest that

> "follow-up research needs to focus on the sensory and electrophysiology of the excitable plasma membrane which constitutes, not only a physical "smart" barrier for the cell's interior, but also allows living cells to maintain their life processes which generate and maintain ordered cellular structures. (P. 161).

Importantly, they note that

> "First cells evolved from hypothetical proto-cells. It can be speculated that these proto-cells were devoid of DNA-based digital memory and relied solely on the structural memory of their limiting membranes." (P. 161)

I agree. We can see strong evidence for that speculation in the fact that the energy factories, the mitochondria, of modern eukaryotic cells rely on their membrane potential to function. Indeed, the mitochondrial potential regulates the speed of the Krebs cycle and its ATP production (Lane, 2022, p 244, 280–284). Bacteria (from which the mitochondria evolved) die by short-circuiting their electrical potential. Life depends on Hodgkin Cycles (the interaction between membrane voltage and protein function) everywhere (Noble 2022).

They conclude:

> "Cellular membranes with associated cytoskeleton represent the primary source of the cellular agency." (P. 161)

1.6. Role of non-genetic sources of bimolecular order

Ildefonso I. De la Fuente (2021) presents a "short overview of the main non-genetic sources of bimolecular order and complexity that underline the molecular dynamics and functionality of cells." He points to several types of organisation in living organisms that are involved. These include:

Dissipative self-organisation, which generates highly ordered dynamic structures far from equilibrium, and first proposed by Ilya Prigogine in 1977. "Practically all metabolite concentrations in cells present complex oscillations and/or non-equilibrium quasi-steady states."

Molecular information processing. "An essential characteristic of the biochemistry of life is that enzymes shape modular dissipative networks, which perform fundamental relatively autonomous activities with specific and coherent catalytic patterns."

Systemic molecular turnover, which is the process by which all cellular components, including structural components are continually being renewed and controlled.

Epigenetic memory "that governs the inheritance of previously acquired new functional characteristics. This biochemical mechanism also represents a huge amount of molecular information not contained in DNA sequences.

He maintains that "enzymes not genes are the essential molecular actors of the functional architecture of life."

This commentary is rather longer than the others, running to 18 pages. It contains a valuable reference list and will form a good resource for students and researchers interested in this field.

1.7. From information to physics to biology

Giuseppe Longo (2023), at the Centre Cavailles in Paris, is a mathematician who has contributed, together with others at the Centre, in many ways to the development of the analysis of complexity in living systems. He agrees that Baverstock's article highlights many aspects of the gene-centred approach that have clearly failed to deliver what was promised, including a deeper understanding of living systems, and practical clinical applications that would cure many multi-factorial disease states. He poses the question "what was meant, and always has been, by "decoding" the genome. In general, if you have an "encoded message" …. As a sequence of signs, "decoding" means its translation into a language and context that is completely meaningful to the intelligent agent or the (biological) structure using it. Baverstock illustrates how far we are from this, that is, from associating, in general, and not in a few special cases, "DNA sequence information into the functional information that informs the phenotype."

His main criticism of the article is that "Baverstock continues to use "informational" language. Longo himself has criticised "the consequences of a terminology borrowed from other sciences." The problem this creates is that "one imports a Laplacian "structure of determination" as Turing and Schroedinger explicitly acknowledge. Longo shows that this ignores the multidimensionality of organisms. He pleads that we should avoid treating "material flows and their gradients as "information" since this by-passed "dimensions, materiality ….historicity … that is all what matters in the analysis off life."

A key section of Longo's commentary deals with the history of physics and physicalism in biology to show how we have been misled. In contradiction with some of the other commentaries, Longo maintains that "an organism is not a self-organising system. It does not emerge *spontaneously* and *necessarily* under certain boundary conditions." This arises because of the essential historicity of living organisms. "This is what we would like to add to Noble's biological relativity, the non-locality of parameters or of causal dependence: at least one of the pertinent parameters that allows/governs the new observable (the heart in embryogenesis, wings in evolution ….) depends on the entire new global structure that did not exist before."

1.8. Phenotypes and agency

Steven Rose (2023) argues that "the last half century of research has steadily chipped away at such a reductionist, unilinear trajectory, and not only because of the unexpected result of the Lenski experiment which he quotes. He is not, as it might appear from the paper, a lone heretical voice."

I think Rose is correct. There really is now a growing community of "New Trends" scientists who, to varying degrees, dissent from the Modern Synthesis and gene-centric neo-Darwinism. I believe there is real hope that the 21st century will see the rebirth of a more biological theory of biology. By that I mean treating living systems as having certain characteristics, such as purposiveness and agency as *definitive* of life, rather than as requiring reductive physical and chemical explanations. Purposive explanations are more predictive about lower-level processes than the other way round (Noble and Noble, 2022, 2023).

Rose writes "Baverstock's confinement of the term [cellular phenotype] to the cellular level is at once too broad and too restricted: the

main 'elements of biology' are expressed at several levels of complexity." Later he continues "These levels increasing complexity are not just epistemological constructs but are ontologically and irreducibly distinct, as spelled out by Joseph Needham in the 1930s."

In concluding, he writes "Phenotypes are simultaneously thing and process; the value of reductionist approaches is that they uncover thingyness; the value of process thinking is that it reinserts the 'thing' into the dynamic self-organising complexity of the living world. Baverstock's de-emphasising genes in favour of cells, I suggest, fits well within this larger theoretical framework."

1.9. Replies to the commentaries

Baverstock, 2023b careful reaction to the commentaries echoes my own interpretation of the present situation. If there existed a simple reply to the central case of the original article, can anyone doubt that at least some of those on the other side of the debate would have penned it?

Silence sometimes speaks loudest.

But this is not just an arcane academic argument between scientists. Society faces a health crisis (Yuille and Ollier, 2021), with an economic fall-out that will dwarf current preparations for health care (Scott and Gratton, 2020). If this was a war, which in a sense it is, the troops and battleships would already have been restrategised to meet the urgent task in hand, to tackle the complexity of the diseases that threaten to bring collapse to our national health systems. The nations with the largest imbalance of aged to young populations are the richest nations in the world, soon to be joined by the rapidly developing nations. Is it not the highest priority now to prepare effectively for a looming crisis? The pandemic has been bad enough. The challenge of longevity combined with intractable diseases is threatening to be even worse. The very viability of health services around the world is at stake.

1.10. Clinical Trial of the polygenic score catalog

My editorial returns to where it began: with the assessment of the performance of polygenic risk scores in screening, prediction and risk stratification, recently published by Hingorani et al. (2023). That study used the same criteria of assessment as for a Clinical Trial. It is sufficient to give the last word to their overall conclusion which showed:

> "poor performance of polygenic risk scores in population screening, individual risk prediction, and population risk stratification.The wide scope and analytical approach of our study might help to resolve the debate on the value of polygenic risk scores, and avoid unjustified expectations about their role in the prediction and prevention of disease." (P. 31)

The widely promised health benefits of genome sequencing have simply failed to materialise. We now need a careful rethink of priorities since it is clear that meeting the looming challenge of ageing populations manifesting diseases that are notoriously resistant to genetic explanations will require resources to be devoted to higher-level studies of the causes of health and disease (Yuille and Ollier, 2021). Looking at the genome level is about as useful as studying the pixels in a message, rather than the message itself. The logic of living systems is not to be found at the level of genes.

1.11. Coda

Although the journal is bringing this debate to a form of completion in publishing this on-line collection, the door remains open to anyone who wishes to respond to submit a stand-alone article. The central issues are not going to go away. The invitation to the genomics community to justify their position remains open.

References

Ball, P., 2014. There Is No "Gene for" Everything. *Prospect*. Dec 2014. https://www.prospectmagazine.co.uk/magazine/there-is-no-gene-for-everything.

Baluska, F., Reber, A.S., 2021. Cellular and organismal agency - not based on genes: a comment on Baverstock. Prog. Biophys. Mol. Biol. 167, 161–162.

Baverstock, K., 2023b. Response to the Commentaries. Progress in Biophysics and Molecular Biology please provide correct DOI.

Baverstock, K., 2023a. The Gene: An appraisal. https://doi.org/10.1016/j.pbiomolbio.2023.11.001.

Boyle, E.A., Li, Y., Pritchard, J.K., 2017. An expanded view of complex traits: from polygenic to omnigenic. Cell 169, 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038.

Collins, F.H., 1999. Shattuck lecture–medical and societal consequences of the human genome project. N. Engl. J. Med. 341, 28–37. https://doi.org/10.1056/NEJM199907013410106.

Dawkins, R., 1976. The Selfish Gene. OUP, Oxford.

De la Fuente, I.M., 2021. Correspondence: insights into the role of genes in cell functionality. Comments on "The gene: an appraisal" by K. Baverstock. Prog. Biophys. Mol. Biol. 167, 152–160.

Felin, T., Koenderink, J., Krueger, J.I., Noble, D., Ellis, G.F.R., 2021a. The data-hypothesis relationship. Genome Biol. 22, 57. https://doi.org/10.1186/s13059-021-02276-4, 2021.

Felin, T., Koenderink, J., Krueger, J.I., Noble, D., Ellis, G.F.R., 2021b. Data bias. Genome Biol. 22, 59. https://doi.org/10.1186/s13059-021-02278-2.

De la Fuente, I., 2021. Correspondence insights into the role of genes in cell functionality. Comments on "The Gene: an Appraisal" by K. Baverstock. https://doi.org/10.1016/j.pbiomolbio.2021.09.006.

Hingorani, A.D., Gratton, J., Finan, C., et al., 2023. Performance of polygenic risk scores in screening, prediction, and risk stratification: secondary analysis of data in the Polygenic Score Catalog. BMJ Medicine 2023 (2), e000554. https://doi.org/10.1136/bmjmed-2023-000554.

Lane, N., 2022. Transformer. The Deep Chemistry of Life and Death. Profile Books.

Longo, G., 2023. From information to physics to biology. Prog. Biophys. Mol. Biol. 177, 202–206.

McKenna, K.Z., Gawne, R., Nijhout, H.F., 2022. The genetic control paradigm in biology: what we say, and what we are entitled to mean. Prog. Biophys. Mol. Biol. 169–170, 89–93. https://doi.org/10.1016/j.pbiomolbio.2022.02.003.

Noble, D., 2021. The surprising heart revisited: an early history of the funny current with modern lessons. Prog. Biophys. Mol. Biol. 166, 3–11.

Noble, D., 2022. How the Hodgkin cycle became the principle of biological relativity. J. Physiol. 600, 5171–5177.

Noble, D., Hunter, P., 2020. How to Link Genomics to Physiology through Epigenomics *Epigenomics*, 4, pp. 285–287.

Noble, D., Noble, R., 2023. How purposive agency became banned from evolutionary biology. In: Corning, P., et al. (Eds.), Evolution "On Purpose". Teleonomy in Living Systems. MIT Press, Vienna series in Theoretical Biology, pp. 221–235.

Noble, D., Phillips, D., 2023. Speciation via physiological selection of environmentally-acquired traits. J. Physiol. (in press).

Noble, R., Noble, D., 2022. Physiology restores purpose to evolutionary biology. Biol. J. Linn. Soc. https://doi.org/10.1093/biolinnean/blac049.

Noble, R., Noble, D., 2023. Understanding Living Systems. Cambridge University Press, Cambridge.

Richardson, K., 2021. "Genes and knowledge: response to Baverstock, K. The gene an appraisal. Prog. Biophys. Mol. Biol. 167, 12–17.

Roll-Hansen, N., 2022. A special role for the genotype? Some comments on Keith Baverstock: "The gene: an appraisal. Prog. Biophys. Mol. Biol. 172, 82–89.

Rose, S.P.R., 2023. Commentary on Baverstock; the gene; an appraisal. Prog. Biophys. Mol. Biol. 178, 70–71. https://doi.org/10.1016/j.pbiomolbio.2023.01.004.

Schwander, T., Leimar, O., 2011. Genes as leaders and followers. Trends Ecol. Evol. 26, 143–151.

Scott, A.J., Gratton, L., 2020. The New Long Life: A Framework for Flourishing in a Changing World. Bloomsbury, London.

Yanai, I., Lercher, M., 2020. A hypothesis is a liability. Genome Biol. 21, 231, 2020.

Yanai, I., Lercher, M., 2021. The data-hypothesis conversation. Genome Biol. 22, 58. https://doi.org/10.1186/s13059-021-02277-3.

Yuille, M., Ollier, B., 2021. Saving Sick Britain: Why We Need the 'Health Society'. Manchester University Press.

Denis Noble

Department of Physiology, Anatomy & Genetics, University of Oxford, UK

E-mail address: denis.noble@dpag.ox.ac.uk.

Foreword

The idea of an *Online Collection* was proposed by Prof Denis Noble, Editor-in-Chief of Progress in Biophysics and Molecular Biology (PBMB) when it was discovered that the paper *The Gene: An Appraisal*, (*The Gene*) accepted for publication in PBMB in May 2021 had been corrupted in the publisher's proofing software and was unreadable beyond the first third. Some 45 invitations were issued to researchers cited in the paper, the work of about 15 of which I had criticised in *The Gene*. In October 2021, Elsevier agreed to an *Online Collection* comprised of the corrected version of *The Gene*, an editorial, the commentaries, and a response to criticisms. The deadline for submission of commentaries was September 2021 and by the end of the year, five had been submitted. I was, however, aware that Prof Giuseppe Longo had submitted commentary, and it was not among the five. His paper was submitted on 20 October 2021 and accepted on 16 December 2022, i.e., the paper was 'lost' in Elsevier's submission system for more than a year. Similarly, Prof Stephen Rose's paper was submitted on 7 October 2021 was also lost in the system until 6 January 2023. Thus, a project that could have been completed in a little over six months had taken 18 months due to Elsevier's incompetence. However, the problem did not end there, because at the beginning of 2023 Elsevier reneged on their promise to produce a corrected version of *The Gene*, claiming that it would be illegal to do so. Elsevier maintained this position until 9 October 2023, when I emailed Elsevier's CEO, Ms. Kumsal Bayazit. After consideration by an ethical committee, Elsevier agreed and produced a corrigendum. It then took until 20 December 2023 before a citable version of the paper was available, leaving the way clear for me to submit my paper titled *Responses to Commentaries,* which had been ready in January 2023. Incredibly, it then took until 7 February 2024 before the proofs of that paper were available to me, and until mid-March before the final version was published. The final step in the publication of the *Online Collection* was taken by Elsevier on 3 April when it posted it on the journal's website as a link to a list of papers that are mostly unavailable to the reader. Denis Noble has confirmed that this is regarded by Elsevier as the finished project.

Elsevier has demonstrated that it holds its editors, authors, and readers in contempt by maintaining that editors and authors should accept that it is legitimate for a publisher to publish in their names corrupted and unreadable texts. One wonders what Elsevier thinks the purpose of publishing is. Perhaps it is simply the money it produces. I would point out that *The Gene* produced nine papers that would not otherwise have been submitted to Elsevier and so it has gained financially from the publication of *The Gene*.

My purpose in agreeing to the issuing of invitations to submit commentaries was to open up the issues raised in *The Gene* to the widest possible discussion. Elsevier, through its incompetence, or intention to subvert (that cannot be ruled out), has made that outcome less likely and, thus, has done a disservice to science.

Three years on from the initial publication of *The Gene*, I stand by the central claim that the primary functional element in the cell is not the gene/genotype but the cellular phenotype, represented by the process of gene product interaction, in today's terminology, a gene product interactome. Where heredity is concerned, this interactome is directly inherited by offspring, in agreement with the statistical/biometric approach to heredity taken by Galton and Pearson, in the form of the Law of Ancestral Heredity, vehemently opposed by Johannsen. Thus, more than 100 years on from that Law, discussion of its relevance to biology is overdue.

Keith Baverstock

2 June 2014

FOCUS ARTICLE

The Gene: An Appraisal

by

Keith Baverstock

May 2021

Progress in Biophysics and Molecular Biology 186 (2024) e73–e88

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

The Gene: An appraisal

Keith Baverstock

Department of Environmental and Biological Sciences, University of Eastern Finland, Kuopio Campus, Kuopio, Finland

A B S T R A C T

The gene can be described as the foundational concept of modern biology. As such, it has spilled over into daily discourse, yet it is acknowledged among biologists to be ill-defined. Here, following a short history of the gene, I analyse critically its role in inheritance, evolution, development, and morphogenesis. Wilhelm Johannsen's genotype-conception, formulated in 1910, has been adopted as the foundation stone of genetics, giving the gene a higher degree of prominence than is justified by the evidence. An analysis of the results of the Long-Term Evolution Experiment (LTEE) with E. *coli* bacteria, grown over 60,000 generations, does not support spontaneous gene mutation as the source of variance for natural selection. From this it follows that the gene is not Mendel's unit of inheritance: that must be Johannsen's transmission-conception at the gamete phenotype level, a form of inheritance that Johannsen did not consider. Alternatively, I contend that biology viewed on the bases of thermodynamics, complex system dynamics, and self-organisation, provides a new framework for the foundations of biology. In this framework, the gene plays a passive role as a vital information store: it is the phenotype that plays the active role in inheritance, evolution, development, and morphogenesis.

1. Introduction

At present, much of biology is regarded as being governed, or regulated, by the genes in the genotype. From the level of the single cell, through organisms and how they develop, evolve, and function, the gene has been assigned a central role. The term is even common in discourse about aspects of human life. It is, in short, considered vital to understanding how life works. The phenotype, on the other hand, plays barely a supporting role in understanding the life *process*. I am proposing that the evidence demands the reversal of this relationship. In the early 1500s, Nicolaus Copernicus proposed reversing the positions of the Sun and the Earth, yielding the heliocentric solar system. Astronomy was simplified, and 1500 years of Ptolemaic astronomy were consigned to history. Newton's laws of motion were subsequently understood to govern the planets. I propose that the evidence dictates that the phenotype is the governor and regulator of the cell, which is the basic 'building block' of the organism. What can flow from this, I contend, is biology governed by thermodynamics and complex system dynamics and a simpler and more intuitive understanding of what life is.

My metaphor for the cellular phenotype is a brain, and for the gene, a provider of building materials, the gene products. The phenotype 'drives' and regulates the cell and the genes in the nucleus house the information for the phenotype to build and operate the cell (Nijhout 1990). Karl Popper asserts that brains and cells can acquire knowledge (Niemann 2014)[1] and I propose that the seat of that knowledge in the cell is the phenotype located in the cytoplasm.

The need for a re-thinking of biology is urgent. Huge resources are directed to the search for the genes that cause human disease. Rare inherited disease traits are often associated with a specific gene abnormality, but they affect only a few percent of the human disease burden: in this context genetics is clinically useful. Common, or so-called polygenic disease traits, potentially affecting everyone, have not yet yielded, in a clinically useable way. The reason is that genes are not responsible for common disease traits.

Explanations in science should be simple, not complicated: in his book, "*Back to Reality*" (Annila 2020), Finnish physicist Arto Annila, constantly emphasises simplicity in explaining even the most apparently intractable aspects of physics.[2] I believe the laws governing biology can be simple too, at least once some counter-intuitive aspects have been grasped.[3]

DOI of original article: https://doi.org/10.1016/j.pbiomolbio.2021.04.005.

E-mail addresses: keith.baverstock@uef.fi, keith@kbaverstock.org.

[1] See Appendix A for full text of Popper's Medawar Lecture to the Royal Society in 1986.

[2] For example, the nature of time, t: the energy, E, of a light quantum is Planck's constant, h, divided by the frequency, f, of the light. I.e., E = h/f. Therefore, h = E x t, where f = 1/t. Time is, therefore, embodied in light quanta along with energy. This is unfamiliar because it is historically not how time has been viewed: it is simple but counterintuitive. On the other hand, Newton's laws of motion, formulated in 1687, are both simple and intuitive.

[3] Unfortunately, the concept of the gene is so embedded into biological thought, and even common discourse that it now constitutes intuition. The arguments presented here, thus, appear to be counterintuitive.

https://doi.org/10.1016/j.pbiomolbio.2023.11.001

Available online 2 December 2023

0079-6107/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Abbreviations

2nd law	second law of thermodynamics
EEA	Equal environments assumption
GWA	Genome-wide association
HGP	Human Genome Project
IA model	Independent Attractor model
LTEE	Long-term evolution experiment
MS	Modern Synthesis
PGS	Polygenic score (sometimes termed polygenic risk score, or PRS)
RoE	Rules of engagement
SNP	Single nucleotide polymorphism

2. A short history of the gene

In February 1865, Gregor Mendel, Abbot of the monastery in Brno, now in the Czech Republic, introduced, in a lecture, what we now know as the gene, calling it an 'element'. He stressed the particulate nature (thingness, or 'Istikeit') of elements,[4] having noted that, in the process of inheritance they retained their unitary nature, rather than blending one with another, as Darwin had assumed.

In 1910, Danish biologist, Wilhelm Johannsen, coined the term 'gene' in a lecture, published as a paper in 1911 (Johannsen 1911).[5] He also coined the terms genotype and phenotype for what Mendel had called 'characters'. The gene quickly entered the scientific discourse of the time as the 'unit of inheritance' and it 'traded' under this guise for 50 or more years.

In 1944, Austrian physicist Erwin Schrödinger published his 1943 lecture in Dublin, "*What is Life?*", (Schrödinger 1944). His "*naïve physicist's ideas about organisms*" looked at from a quantum mechanical perspective, yielded the conclusion that the hereditary material must be a solid, he called it an 'aperiodic crystal'.

From around 1960, Petter Portin and Adam Wilkins (Portin and Wilkins 2017), report that the gene started to be viewed as a defined string of nucleobases that coded for a polypeptide: it was a material *thing*. This transformation was driven by the discovery of the structure of deoxyribonucleic acid (DNA) in 1953 by Francis Crick, James Watson, Rosalind Franklin, and Maurice Wilkins. Crick went on in 1958 to propose how proteins (more properly peptides) which yielded the phenotype, were coded in the gene's DNA sequence (Crick 1958). In 1970, Crick proposed the Central Dogma (which stipulated that information in the DNA flowed to the protein and not the reverse) and the sequence hypothesis, which stipulated that the sequence of the amino acids in a peptide determined the native and biologically active structure of the folded protein (Crick 1970). These developments in the 1950s/60s have determined how the gene has been perceived for the following 50 years: molecular genetics was born.

The prospect of sequencing the whole human genome was on the horizon by the early 1980s. Crick's assertion[6] that the 'secret of life' lay in the DNA that constituted the genes, made in the Eagle pub in Cambridge in February 1953, became increasingly convincing to biologists and the public alike. That 'secret' would be revealed in the sequence of the human genome.[7]

The Human Genome Project (HGP), aimed to sequence the 3 billion bases in the human genome, commenced on 1 October 1, 990[8] with a grant of three billion US$ from the US Congress. Initially headed by James Watson, it was brought to its conclusion in 2003 by Francis Collins, now the Director of the National Institutes of Health in Washington. In 2001, when sequencing was sufficiently advanced to announce preliminary results,[9] the human genome turned out to contain far fewer genes than the 'one gene: one polypeptide' hypothesis[10] predicted. Palaeontologist, Stephen Jay Gould, wrote in the *New York Times* under the heading "*Humbled by the Genome's Mysteries*"[11]:

> "*The general estimate [of the number of genes] for Homo sapiens … …. had stood at well over 100,000, with a more precise figure of 142,634 widely advertised and considered well within the range of reasonable expectation. Homo sapiens possesses between 30,000 and 40,000 genes, with the final tally almost sure to lie nearer the lower figure.*"

Indeed, the final figure lies between 20,000 and 25,000 protein-coding genes[12]: the HGP represented a major collision between genetics and reality.

According to Portin and Wilkins (2017), since the sequencing, several other problems have emerged with the concept of the gene: some gene sequences are not clearly delineated; the sequence of exons[13] in the gene is not necessarily reproduced at translation; a gene sequence may not be contiguous along the chromosome, and a given gene in one cell type may function differently in another. In short, the gene has proved extremely difficult to define concisely. This matters when the aim is to predict the phenotype from the genotype: which was the rationale for the HGP.[14] However, was that even a realistic aim? Take for example the

[4] Robert Olby in Mendel, Mendelism and Genetics. http://www.mendelweb.org/MWolby.html. (accessed 23.02.2021).

[5] This landmark paper was reprinted in 2014: Johannsen, W. (2014). "The genotype conception of heredity. 1911." Int J Epidemiol **43**(4): 989–1000. In this paper references are made to the original version.

[6] http://news.bbc.co.uk/2/hi/science/nature/2804545.stm (accessed 23.02.2021).

[7] Lewontin says: "… … the great panjandrum of DNA himself, James Dewy Watson, explains in an essay in the collection edited by Kevles and Hood that "he doesn't want to miss out on learning how life works" and Gilbert predicts that there will be a change in our philosophical understanding of ourselves". : Lewontin, R. C. (1992). The doctrine of DNA: Biology as Ideology. London, England, Penguin Books Ltd. p63.

[8] https://en.wikipedia.org/wiki/Human_Genome_Project (accessed 23.02.2021).

[9] The official completion date of the HGP was 14 April 2003 but a preliminary report was released in February 2001 to coincide with the birthday of Charles Darwin: Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860–921.

[10] By Archibald Garrod around 1900.

[11] https://www.nytimes.com/2001/02/19/opinion/humbled-by-the-genome-s-mysteries.html (accessed 23.02.2021).

[12] https://www.sciencedaily.com/terms/human_genome.htm (accessed 23.02.2021).

[13] The string of bases that comprise a gene is divided into exons, sections which code for gene products and introns, which are non-coding intervening sequences of bases.

[14] See: Lewontin, R. C. (1992). The doctrine of DNA: Biology as Ideology. London, England, Penguin Books Ltd. In the chapter headed "The Dream of the Human Genome" (pp 61–83) Lewontin ridicules the then much heralded idea that the sequence of the genome would tell us about the human condition and "*locate on the human chromosomes all the defective genes that plague us*" noting that some mutant genes (that for cystic fibrosis, for example) had already been located, isolated, and sequenced. A decade ago, Lewontin might have felt entirely vindicated. On 27 July 2010, Craig Venter, the entrepreneur who competed with the HGP to sequence the human genome, was interviewed by Der Spiegel under the title "We have learned nothing from the Genome". Since then, with the development of the technique of genome wide association (GWA), there has been a massive upsurge in genetic studies of common disease and behavioural traits. Despite this, Lewontin remains vindicated: as I will argue, this decade of intense research activity has not advanced our understanding of the causes of common disease and behavioural traits.

DSCAM gene found in Drosophila: it can produce 38,016[15] different peptides, (Black 2000), more than the number of genes in the human genome. According to the dogma, each peptide may fold into a different protein performing a discretely different biological function.

Despite the lack of clarity over the concept of the gene, and the unexpectedly low number of genes found by the HGP, genetic research has forged ahead in recent decades.

Traits (Mendel's characters and Johannsen's phenotypes) are classified as either monogenic (Mendelian) or polygenic. Monogenic traits, for example, the flower colour that Mendel investigated in pea plants, have been the sole basis for experimental genetics since the time of Mendel, according to American geneticist Richard Lewontin (1974). Rare inherited diseases such as Huntingdon's disease (there are thought to be ~10,000[16]), affecting less than 8% of the population, are often monogenic traits.

Rare diseases have long been diagnosed using classical genetic techniques, but success has been limited. With the benefit of knowing the human genome sequence, improvements were expected. The genomes of 85,000 UK National Health Service (NHS) patients, the majority with undiagnosed rare diseases, have been sequenced in the '100,000 Genomes Project'. Launched in 2012,[17] with sequencing completed in 2018,[18] few results have been published. The project website[19] says it has provided diagnoses in 20–25% of the cases.

Using the genome wide association (GWA) technique[20] and the human genome sequence, polygenic traits (common diseases and behavioural conditions) have allegedly been characterised by tens to hundreds of single nucleotide polymorphisms (SNPs)[21] at nearly as many loci (genes), each of very small effect. This is occurring in populations of thousands to hundreds of thousands of individuals carrying the trait. Furthermore, the total of these effects does not add up to the expected total genetic risks (or variances) of the diseases.[22] The difference is what is known as the 'missing heritability' (Manolio et al., 2009; Eichler et al., 2010; Chaufan and Joseph 2013, Blanco-Gomez et al., 2016): it is currently a major problem at the root of the genetics of common disease and behavioural traits.

GWA data *per se* are, therefore, of no clinical utility. It is, however, claimed that summing up the SNPs into a so-called polygenic score (PGS)[23] is of diagnostic value (Plomin 2018)[24]: however, this may resolve the problem of many small effects at numerous loci, but it leaves the problem that the PGS can only apply to a small fraction of the total genetic variance. The clinical utility of PGSs has yet to be proven.

Genes have been the 'material currency' of biology for 155 years. They are centrally invoked to explain inheritance, evolution, development, and morphogenesis: they have become icons of biological thought, such that it is heretical that their prominence should be challenged. Yet, they are far from well-defined, and knowing their sequences has not, so far, advanced our understanding of the most important challenge to human health, namely common disease.

3. How Mendel's elements became genes

Now I want to look in more detail at how Johannsen defined the gene. Mendel's 1865 paper lay unrecognised until 1900 when the Dutch biologist, Hugo de Vries, discovered it and re-published it. It could then be integrated with Darwin's ideas on evolution through natural selection, as laid out in "*On the Origin of Species*", which had been published in 1859 (Darwin and Kebler 1859).

The foundation stones of today's biology had been laid.

In the earliest years of the 20th Century, inheritance, or heredity, was the primary problem of the day in biology. Johannsen was opposed to the use of the above terms when applied in biology: he claimed that their everyday use, in terms of the transmission of wealth from one generation to the next, were misleading metaphors for biology.[25] The dominant theory of inheritance was Francis Galton's regression law.[26] Johannsen called it the 'transmission-conception' and regarded it as wrong: it supported Lamarckism[27] and Darwin's pangenesis concept,[28]

[15] The DSCAM gene has a total length of 61 kb (61,000 base pairs) and is divided into 24 exons. Four of those exons occur with up to 48 alternative sequences. Taking all the viable combinations of alternative splicings of the exons and alternative sequences contributing to the mRNA that can be transcribed from the gene, more than 38,000 peptides, and, therefore, proteins, can be translated.

[16] https://www.who.int/genomics/public/geneticdiseases/en/index2.html (accessed 23.02.2021).

[17] https://www.sciencemag.org/news/2012/12/uk-unveils-plan-sequence-whole-genomes-100000-patients (accessed 23.02.2021).

[18] https://www.genomicsengland.co.uk/the-uk-has-sequenced-100000-whole-genomes-in-the-nhs/(accessed 23.02.2021).

[19] https://www.genomicsengland.co.uk/about-genomics-england/the-100000-genomes-project/(accessed 23.02.2021).

[20] https://en.wikipedia.org/wiki/Genome-wide_association_study (accessed 23.02.2021).

[21] https://en.wikipedia.org/wiki/Single-nucleotide_polymorphism (accessed 23.02.2021).

[22] Typically, common diseases, at the population level, are thought to be between 10 and 70% due to genetic causes. These estimates are determined from family or twin studies. In GWA studies, typically between 5 and 15% of this risk is accounted for. The difference, or 'missing heritability', therefore, ranges up to several 10s of percentage points.

[23] https://en.wikipedia.org/wiki/Polygenic_score (accessed 23.02.2021).

[24] See Chapter 12 "The DNA fortune teller".

[25] In his 1911 paper Johannsen writes: "*The view of natural inheritance as realised by an act of transmission, viz., the transmission of the parent's (or ancestor's) personal qualities to the progeny, is the most naive and oldest conception of heredity. We find it clearly developed by Hippocrates, who suggested that the different parts of the body may produce substances which join in the sexual organs, where reproductive matter is formed.*": Johannsen, W. (1911). "The Genotype Cconception of Heredity." American Naturalist **45**: 129–159. Johannsen's main concern appears to be avoiding the inheritance of acquired characteristics. He goes on: "*The personal qualities of any individual organism do not at all cause the qualities of its offspring; but the qualities of both ancestor and descendant are in quite the same manner determined by the nature of the "sexual substances"—i.e., the gametes—from which they have developed. Personal qualities are then the reactions of the gametes joining to form a zygote; but the nature of the gametes is not determined by the personal qualities of the parents or ancestors in question. This is the modern view of heredity*." Further on he says: "*The "genotype-conception," as I have called the modern view of heredity, differs not only from the old "transmission-conception" as above mentioned, but it differs also from the related hypothetical views of Galton, Weismann and others, who with more or less effectiveness tried to expel the transmission-idea, having thus the great merit of breaking the ground for the setting in of more unprejudiced inquiries. Galton, in his admirable little paper of 1875, and Weismann, in his long series of fascinating but dialectic publications, have suggested that the elements responsible for inheritance (the elements of Galton's "stirp" or of Weismann's "Keimplasma") involve the different organs or tissue-groups of the individual developing from the zygote in question. And Weismann has furthermore built up an elaborate hypothesis of heredity, suggesting that discrete particles of the chromosomes are "bearers" of special organizing functions in the mechanism of ontogenesis, a chromatin-particle in the nucleus of a gamete being in some way the representative of an organ or a group of tissues*." Thus, Johannsen was aware of the Weismann barrier whereby the germ cells are 'insulated' from the rest of the organism and that the gametes do not carry 'personal qualities', yet he does not consider the gamete phenotypes, only their genotypes, as Mendel's 'units of inheritance'.

[26] According to Galton, an individual's traits were transmitted from their parents (50%), their grandparents (25%), their great grandparents (12.5%), and so on, with ever diminishing importance, because of the increasing number of ancestors, within whom the traits were distributed.

[27] Jean Baptiste Lamarck was a highly regarded French biologist who died in 1829. He became the professor of Zoology when the Muséum national d'Histoire naturelle opened in Paris in 1793. He advocated the idea that qualities gained during a lifetime could be passed on to future generations. This is called the inheritance of acquired characteristics.

[28] https://en.wikipedia.org/wiki/Pangenesis.

both of which implied the inheritance of acquired characteristics.

Johannsen ran an experiment with self-fertilising bean plants (a so-called 'pure line breeding' programme)[29] and recorded the dimensions of the beans produced over two growing seasons. Bean sizes were distributed according to a normal distribution, but different pure lines differed slightly in the size range of the beans they produced. Johannsen categorised these lines as 'genotypes' and the process of inheritance through the genotype he called the genotype-conception (Johannsen 1911). He found no evidence of ancestral influences in his experiments[30].

Nils Roll-Hansen (2014) says of Johannsen's presentation of his genotype-conception at the lecture in 1910, published in 1911:

> "*This lecture summed up his experimental and theoretical achievements, including a sharp analysis of the concepts of 'genotype' and 'gene'. … …. Genotype is the basic concept in Johannsen's 1910 lecture. The stability of the genotype is what makes a science of heredity possible. The concept of 'gene' is derivative. It represents an experimentally identifiable difference between genotypes*"

Thus, Johannsen's work must be credited as the basis for modern genetics and the understanding of inheritance, and the longstanding theory of evolution, the Modern Synthesis (MS),[31] since inheritance is an essential component of evolution.

The American geneticist, T. H. Morgan, writing in 1917 under the title "*The Theory of the Gene*" (Morgan 1917), defended Mendelism and confirmed the location of genes in chromosomes. Mendel's laws of inheritance, based on experiments with pea plants and Johannsen's genotype-concept, were converted into a theory using primarily the concepts of 1) two alleles (versions) per gene, each being capable of being dominant or recessive, and 2) the phenomenon of epistasis.[32] Morgan concedes:

> "*It has been said that by assuming enough genetic factors you can explain anything. This is true; and it is the greatest danger of the factorial procedure. If, for example, whenever one fails to account for a result he introduces another factor to take care of what he cannot explain he is not proving anything except that he is ingenious or only naïve*." (Morgan 1917).

Those simple concepts give considerable interpretative latitude and they have been progressively added to over the years in a manner that is perhaps not unlike epicycles in Ptolemaic astronomy. Nevertheless, genetics today is regarded as a successful and sophisticated scientific discipline. Indeed, on the 20th anniversary of the release of the draft human genome sequence in 2001, the journal *Nature* proclaimed, "*A wealth of discovery built on the Human Genome Project*" (Gates et al., 2021). The authors point out that as there is no world without the HGP it is impossible to say how much progress it represents but "*it is nonetheless clear that the HGP's catalogue* [of protein-coding genes] *catalysed the continuing genetic revolution*".

There are, however, features of genetics that should have given pause for thought.

First, likenesses between siblings, or those between parents and their offspring, which we know empirically to exist, cannot be explained intuitively in terms of the above concepts (see below).

Second, the physicist Max Delbrück defined genetics in 1935 as:

> "… …. *a far-reaching, logically closed, strict science. It is quantitative without making use of the physical measurement system*."[33] (Timoféeff-Ressovsky et al., 1935).

Delbrück acknowledges that genetics, unlike chemistry, is not based on a more fundamental physics, from where it would be possible to judge and test hypotheses. Thus, there is no more fundamental level against which to judge the genotype-conception: it is simply a theoretical model for which there is some support.

Third, in 1958 Francis Crick (1958) published his thoughts on how the information coded in the gene sequences informed the phenotype. Information coded in the base sequence needed to be transformed into information in the form of the molecular structure of a protein, the supposed biologically active molecule in the cell. He proposed the 'sequence hypothesis': which essentially posits that if the biologically inactive product of transcription and translation of a gene, the peptide, folded *itself* to a native state protein, the information in the gene would be conserved in the protein. There was no underpinning in physics for this hypothesis. However, experiments with the enzyme ribonuclease by American biologist Christian Anfinsen, (Anfinsen et al., 1961) showed that, in the test tube, denatured enzyme (that had been converted to the peptide) re-folded spontaneously to the active structure, apparently confirming Crick's hypothesis. However, conditions in the test tube[34] are very different from those in the cell (Minton 2006) and the renaturation process took far too long for it to be generally applicable in the cell. Crick reflected on the sequence hypothesis in his Nobel Lecture in 1970: this is what he wrote:

> "*Because it was abundantly clear by that time [1958] that protein had a well-defined three-dimensional structure, and that its activity depended*

[29] The 'pure line' experimental approach differed from that deployed by Mendel with pea plants. Although pea plants are self-fertile, like bean plants, Mendel did not allow self-fertilisation in his experiments. He crossed one plant with another with different characters, such as flower colour, to produce hybrids and not pure lines.

[30] "*Ancestral influence! As to heredity, it is a mystical expression for a fiction. The ancestral influences are the "ghosts" in genetics, but generally the belief in ghosts is still powerful. In pure lines no influence of the special ancestry can be traced; all series of progeny keep the genotype unchanged through long generations.*": Johannsen, W. (1911). "The Genotype Cconception of Heredity." American Naturalist **45**: 129–159.

[31] The Modern Synthesis was proposed by Julian Huxley in 1942: Huxley, J. (1942). Evolution, the modern synthesis. London, G. Allen & Unwin ltd. It is based on the earlier ideas of neo-Darwinism (see Noble, D. (2017). Dance to the tune of life: biological relativity. Cambridge; New York, Cambridge University Press pp 131–134) and it remains the backbone of evolutionary theory today. It is criticised, e.g., ibid. but it is also aggressively defended, for example, by Richard Dawkins. It has been extended: Laland, K. N., T. Uller, M. W. Feldman, K. Sterelny, G. B. Muller, A. Moczek, E. Jablonka and J. Odling-Smee (2015). "The extended evolutionary synthesis: its structure, assumptions and predictions." Proc Biol Sci **282**(1813): 20151019. More recent proposals to replace the MS and the Extended Synthesis are: Corning, P. A. (2020). "Beyond the modern synthesis: A framework for a more inclusive biological synthesis." Prog Biophys Mol Biol **153**: 5–12. and Richardson, K. (2020). "In the Light of the Environment: Evolution Through Biogrammars Not Programmers." Biological Theory **15**: 212–222. Critics of the MS are often dismissed by its advocates as Creationists. For this reason the website: https://www.thethirdwayofevolution.com/people (accessed 23.02.2021) features researchers of evolutionary biology who are neither advocates of the MS nor Creationists.

[32] Epistasis, or gene-gene interaction, is where one gene present in a genotype influences the effect of another gene in the same genotype. A gene for baldness over-rides genes for red and blonde hair, for example. This phenomenon was discovered by British biologist William Bateson and colleagues working in Cambridge, England, in 1907.

[33] "… *eine weitgebend in sich logisch geschlossene, strenge Wissenschaft. Sie ist quantitaiv, ohne vom physicalischen Maßsystem Gebrauch zu* machen.": Timoféeff-Ressovsky, N. W., K. G. Zimmer and M. Delbrück (1935). "Über die Natur der Genmutation und der Genstruktur." Nachrichten der Biologischen Gesellschaft für Wissenschaft, Göttingen **1**: 189–241.

[34] As a dilute solution in aqueous buffer, in contrast to the highly molecularly crowded environment of the cell cytoplasm: Fonin, A. V., A. L. Darling, I. M. Kuznetsova, K. K. Turoverov and V. N. Uversky (2018). "Intrinsically disordered proteins in crowded milieu: when chaos prevails within the cellular gumbo." Cell Mol Life Sci **75**(21): 3907–3929. Up to 30% of the cell cytoplasm can be occupied by peptides and proteins.

crucially on this structure, it was necessary to put the folding process to one side, and postulate that by and large the polypeptide chain folded itself up." (Crick 1970)[35]

Hardly a ringing endorsement of his own hypothesis for such a crucial feature of modern molecular genetics: peptide folding to protein is the process that transforms DNA sequence information into the functional information that informs the phenotype. It is now clear that the sequence hypothesis is invalid (Baverstock 2019a). and, therefore, the genotype-conception, cannot explain how the alleged information coded in the gene sequences informs the phenotype.

Fourth, in 1974, as already noted, Lewontin pointed to a paradox that had been inherent in experimental genetics since Mendel's time: in terms of traits, what is interesting (polygenic traits) is not measurable, and what is measurable (monogenic traits) is not interesting (Lewontin 1974). Has that situation changed? There has been a massive expansion in GWA[36] studies of polygenic traits in the past decade, but these are not leading to theories about how SNPs are related to biological effects. Indeed, the foremost advocate of the application of PGSs in the diagnosis of behavioural traits, Robert Plomin, maintains that the "predictive power of polygenic scores does not require knowing anything about the processes that lie between genes and behaviour." and notes that "success in identifying DNA differences [associated with behavioural traits] came only after the search for candidate genes selected for their possible causal connection to a trait was superseded by a hypothesis-free approach that is agnostic about the specific function of DNA variants (i.e., genome-wide association)." (Plomin and von Stumm 2022). GWA has not added anything to the theoretical basis of genetics.

Finally, hypotheses nominating variant genes (so-called candidate genes), as causes of specific common disease traits based on biological considerations, have largely failed. For example, 18 candidate genes have been hypothesised, based on their perceived biological relevance, to account for major depressive disorder. In a highly statistically significant study of a large database of patients, Border et al. (Border et al., 2019) reject all 18 genes, some of them having been prominently reported on in the past. This is not an isolated case: the failure of the candidate gene approach (based largely on classical genetics) ought to be a signal that something is very wrong: evidence has comprehensively rejected theory.

Could there be a different way to understand inheritance that fits equally well with Mendel's laws (that were, in any case, based on fudged experimental data (Elston 2018))?

Can we really say that a theory of human inheritance, based on the gene-centric genotype-conception and experimental genetics confined to monogenic traits, is scientifically secure?

Are experiments with self-fertilising pea and bean plants a secure enough foundation to be able to generalise to human heredity? The genotype-conception is a theoretical model: it may be correct, but it is not grounded in science. The pieces of the genetic jigsaw puzzle can be made to fit well enough together, but can we be sure they give us the correct picture of human inheritance?

We will return to genes in inheritance after considering the role of genes in evolution.

4. The gene in evolution

Now we consider the gene's purported role in evolution in providing selectable variation as spontaneous mutations to genes. In 1930 Ronald Fisher published "*The Genetical Theory of Natural Selection*" (Fisher 1930) in which he derives the following law:

"*The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time*."

Fisher, however, does not include the mutations arising during the course of evolution as contributing to variance, stating:

"*The rate at which a mutation increases in numbers at the expense of its allelomorph will indeed depend on the selective advantage it confers, but the rate at which a species responds to selection in favour of any increase or decrease of parts depends on the total heritable variance available, and not on whether this is supplied by large or small mutations. There is no limen of appreciable selection value to be considered.*" (Fisher 1930).[37]

Given the prominence that Fisher's theory holds in the MS and the importance in that attached to *new* variation arising from mutations, it is perhaps surprising that Fisher's theory is still considered important. According to evolutionary biologist, Ernst Mayr (2001) [38], mutations are the principal source of *new* variation for natural selection to act on.

The reason for this apparent anomaly is the way Fisher frames his theory: his law only applies to an instant in time and includes other processes that advance evolution (Grafen 2003). Basener and Sandford (Basener and Sanford 2018) have set out to correct this deficit by modifying Fisher's model to include mutations that occur during the evolutionary process. Empirical evidence of how evolution occurs over a significant number of generations is needed to test their models and that is in short supply.

However, the Long-term Evolution Experiment (LTEE), which has been running since 1988 under the direction of Richard Lenski and colleagues, has racked up over 66,000 generations of a population of *E. coli* bacteria (Lenski 2017). An initial 12 cultures, drawn from a single source of starved genetically pure *E. coli*, have been grown in medium containing a limiting concentration[39] of glucose as the only readily accessible carbon source to bacteria grown under aerobic conditions. A single aliquot of each culture is inoculated into a new flask of medium every 24 hours: by that time the bacteria will have used up the glucose. Periodically, the *fitness*, in terms of the rate of growth of the bacteria compared to that of the founder bacteria, and the *body size/cell volume* of the bacteria, are assessed. Mutations are measured by gene sequencing less frequently.

This is by no means a 'natural experiment': it is adaptation to life in the inside of a laboratory flask, in a medium with a single accessible carbon source, in an incubator. However, the overwhelmingly most likely source of new variation arising in the cultures is new mutations (Lenski 2017) resulting from errors in replication. It has been proposed that epimutations[40] have a role in evolution (Jablonka 2017) but these are generally induced by stress from the environment. The environment of the bacteria in the LTEE is constant across the generations. Therefore, there should be few complications in interpreting the results in the context of expectations based on the MS. The LTEE has, however, been started without any prior hypothesis: rather a set of questions such as, 'is there a limit to extent of adaptation?' and 'how repeatable is adaptation?' (Lenski 2017).

When 10,000 generations had accrued, it was clear that all 12 samples were evolving similarly with respect to fitness but not with respect to cell volume. Both showed initially rapidly increasing trajectories that tailed off and could be fitted by hyperbolic curves (Lenski and Travisano 1994).

In 2009, one of the 12 samples which had accrued 40,000

[35] Interestingly, Crick does not mention Anfinsen's supportive experiment in this paper and neither does Anfinsen reference Crick's papers. Arguably, in his Nobel lecture, Anfinsen backs away from the idea that his work on ribonuclease is relevant to the cell: Anfinsen, C. B. (1973). "Principles that govern the folding of protein chains." Science **181**(4096): 223–230.

[36] A search on PubMed on the terms 'GWA' OR 'genome-wide association' in the 'Title/Abstract' search field returned nearly 31,000 publications since 2003.

[37] pp 15–16.

[38] p 279.

[39] The concentration of glucose is limited such that the bacteria will have consumed it within the 24 hours and returned to the starving state.

[40] For example, methylation of DNA.

generations, was analysed in detail. Fitness followed a hyperbolic growth curve up to 20,000 generations and mutations were accrued linearly (Barrick et al., 2009).[41] See Fig. 1. When extended to 50,000 generations it was clear that the fitness curve was better fitted by a power, rather than a hyperbolic, law (Wiser et al., 2013). See Fig. 2.

The primary results from the LTEE are:

1) an *identical*, initially steep, evolution of fitness in the 12 experiments with a linear evolution of mutations. See Fig. 1.
2) 12 *non-identical*, but increasing profiles for cell volume, (see legend to Fig. 1) and
3) evolution of fitness up to 50,000 generations according to a power law. See Fig. 2.

At around 30,000 generations, one of the 12 samples acquired the ability to metabolise citrate. Citrate is a component of the medium but is inaccessible under *aerobic*, but not in *anaerobic*, conditions. Since *E. coli* have a citrate transport system and so can metabolise citrate (Hall 1982, Van Hofwegen et al., 2016), the LTEE has not yielded a novel capability for the bacteria, as has been claimed (Dawkins 2009).[42]

What is to be made of this experiment? How can these results be explained at an intuitive level? Judging from the report on a single sample in 2009 (Barrick et al., 2009), the results seem to have been a conundrum for Lenski and colleagues:

> "*The simplest hypothesis that could explain the discrepancy between the nearly constant rate of genomic change and the sharply decelerating fitness trajectory posits that only a small fraction of all substitutions are beneficial, whereas most are neutral or nearly so. Accordingly, the beneficial substitutions would be concentrated in the early phase of rapid adaptation to the conditions of the experiment, but over time that initial burst would be swamped by the constant accumulation of neutral mutations by drift. However, four lines of evidence allow us to reject this explanation.*"

There was no ready explanation for the non-correspondence between the genomic evolution and the evolution of adaptation, or of the initial steep increase in fitness. What would be the probability of 12 samples (Galapagos Islands?) following identical fitness trajectories if the cause was *randomly* acquired mutations? This feature seems to rule out an effect of acquired mutations on fitness.

According to Fisher's theorem (with no mutations acquired during the evolutionary process) any increase in fitness would be due to the selection of beneficial alleles already in the founder population. This might account for the 12 experiments having the same initial slopes but why would not the same argument apply to cell volume, and is it feasible that it would apply to fitness over as many as 50,000 generations?

Richard Dawkins is one of the foremost advocates of Darwinian evolution and the MS today. He devotes about 15 pages to the LTEE in his book "*The Greatest Show on Earth: the evidence for evolution*", which was published in 2009 (Dawkins 2009). At the state of the LTEE when Dawkins was writing, 45,000 generations had accrued, and he was aware that all 12 samples were evolving in a similar way in terms of fitness.[43] In the conclusion of his discussion of the experiment, he writes this:

> "*Lenski's research shows, in microcosm and in the lab, massively speeded up so that it happened before our very eyes, many of the essential components of evolution by natural selection, random mutation followed by non-random natural selection; adaptation to the same environment by separate groups independently; the way successive mutations build on their predecessors to produce evolutionary change; the way genes rely for effects on the presence of other genes. It all happened in a tiny fraction of the time evolution normally takes.*"

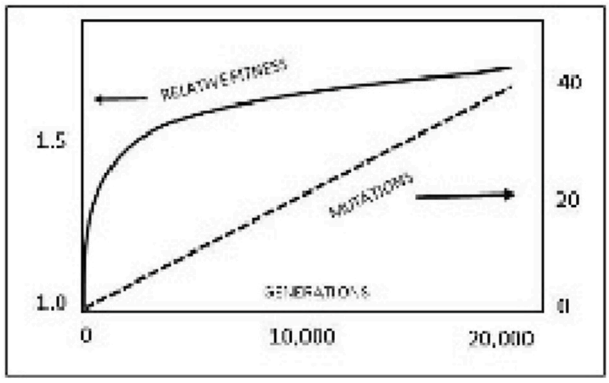

Fig. 1. To be inserted here.

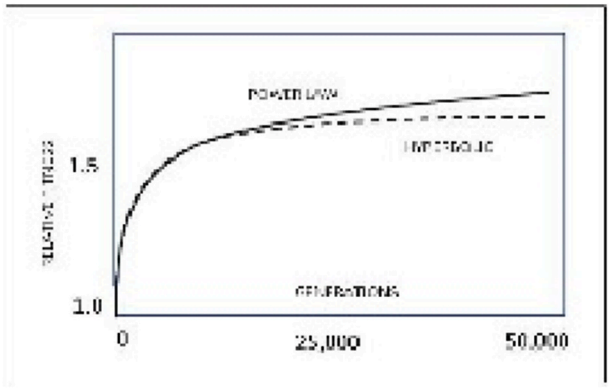

Fig. 2. To be inserted here.

Dawkins apparently has no qualms about the rapid increase in fitness in the early stages of the experiment when few mutations have arisen[44] and most of those would be deleterious, not beneficial (see below). He sees no problem with all 12 experiments being in lockstep as far as their gain in fitness is concerned. In connection with the increase in cell volume, he evidently expects that each tribe, as he calls each experiment, will take a different course.[45]

[41] after which fitness increased dramatically and the population was deemed to be hypermutable.

[42] See pp 127-130.

[43] Dawkins seems to have based his assessment of the experiment on: Lenski, R. E. and M. Travisano (1994). "Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations." Proc Natl Acad Sci U S A **91**(15): 6808–6814.

[44] At approximately 2000 generations in one experiment, relative fitness had increased to 1.4, just under half the gain at 50,000 generations but only 5 mutations had accrued: Interpolated from Fig. 2 in: Barrick, J. E., D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh, D. Schneider, R. E. Lenski and J. F. Kim (2009). "Genome evolution and adaptation in a long-term experiment with Escherichia coli." Nature **461**(7268): 1243–1247.

[45] "You can see [referring to a graph on p. 123, which shows similarly shaped profiles to that for fitness] that most of the increase in body size occurred in the first 2000 or so generations. The next interesting question is this. Given that all 12 tribes increase in body size over evolutionary time did they all increase in the same way by the same genetic route? No, they didn't. And that is the second interesting result. The graph at the top of page 123 is for one of the 12 tribes now look at the equivalent hyperbolic best fits for all 12. Look how spread out they are. They all seem to be approaching a plateau but the highest of the 12 plateaus is almost twice as high as the lowest. And the curves have different shapes: the curve that reaches the highest by generation 10,000 starts by growing more slowly than some of the others and then overtakes them before generation 7000.": Dawkins, R. (2009). The greatest show on Earth: the evidence for evolution. New York, Free Press.

Can Dawkins have it both ways: evolution of fitness is identical in all 12 experiments, while in terms of cell volume, each of the samples evolves independently of the others and both are dependent *only* on randomly acquired mutations during the evolutionary process?

Dawkins' enthusiasm for the LTEE leads him to take a stab at creationists[46] as the experiment offers no support for intelligent design. However, the LTEE does not, at face value, offer support for the MS and evolution driven by genetic variation provided by mutations.

For the naive observer, one thing stands out in this experiment: the power law governing the identical increases in fitness in all 12 experiments for over 50,000 generations (Wiser et al., 2013). Intuitively, randomly acquired mutations arising during the experiment cannot account for this. It is a very significant result, and I will return to it after seeing whether modifications of Fisher's theorem (mentioned above) can explain the LTEE results.

Lenski and his colleagues have modelled an explanation of their experiment (Good et al., 2017) and they have considerable leeway in terms of the number of variables with unknown effects that they can exploit. They can surely concoct a jigsaw puzzle where the available evidence fits relatively well. As I noted earlier, there was no prior hypothesis for the LTEE to test, only a list of questions. Constructing hypotheses after the results of the experiment are known, as Lenski and colleagues have done, is called 'harking' (Kerr 1998) and it is exactly what the Ptolemaic astronomers did for 1500 years.

4.1. Does Fisher's theory of natural selection help explain the LTEE results?

Basener and Sanford modify Fisher's law of natural selection to include the new mutations arising during the evolutionary process saying:

> "*Our goal is to correct and re-apply Fisher's Theorem, such that it is consistent with real biology*." (Basener and Sanford 2018)[47]

But first, they consider Fisher's condition of no new mutations arising. That would be the situation in the LTEE at the very earliest times, where there is a steep rise in fitness in all 12 samples. Over a relatively short period (Fisher's law is for an instantaneous relationship between variance and change in fitness) fitness increases *linearly* at a rate proportional to the variance, which remains constant. That is Fisher's law of natural selection and it is not consistent with a power law.

Next, they show that extending Fisher's model to long times, still with no mutations occurring, results in a population initially increasing linearly in fitness but levelling off at maximum fitness with zero variance, i.e., no further scope for increase in fitness. This is not consistent with a power law either.

Basener and Sanford then consider two different distributions for beneficial and deleterious mutations acquired during the evolutionary process. The first model assumes an initial distribution of beneficial and deleterious mutations in equal measure, normally distributed in terms of effect size. Under this model, fitness increases initially slowly, gradually increasing in rate over time. That is not consistent with a power law observed. As Basener points out,[48] their simulations give the average response to the distribution of mutations; individual experiments could vary due to the random mutational events that may happen at various times (during evolution). Maddamsetti et al. confirm that:

> "…. . *the replicate populations of the LTEE have largely diverged in their mutation rates and biases, even though they have adapted to identical abiotic conditions*" (Maddamsetti and Grant 2020).

Therefore, we sh ould not expect all 12 experiments to have identical fitness profiles.

This first distribution of acquired mutations, however, is not a realistic model: there is a strong consensus that deleterious mutations outnumber beneficial mutations. The second distribution Basener and Sanford modelled assumes that detrimental mutations (small in individual effectiveness) will out-number beneficial ones by a margin of a thousand to one. This is, in fact, conservative, as other estimates including by the Japanese geneticist, Kimura and by Lenski, suggest a margin of one million to one. This model shows an initial reduction in fitness, which decelerates with time, again quite contrary to observation and the observed power law relationship.

Thus, neither Fisher's longstanding and widely accepted theorem of natural selection unmodified, nor the two distributions of mutations that Basener and Sanford used to modify the theorem, fit the observations derived from the LTEE.

Now we consider the significance of the power law and the evolution in lockstep of fitness of the 12 independent experiments. Wiser et al. write:

> "*The power law describes the fitness trajectories well, but it is not explanatory*." (Wiser et al., 2013).

This is not necessarily so: natural *processes*, which inevitably involve energy dissipation, are described by power laws (Makela and Annila 2010), and evolution is one such natural process requiring nothing more than an energy input (in terms of nutrient). Evolution of fitness is an expression of the principle of least action, synonymous with the 2nd law of thermodynamics (hereinafter, the 2nd law), i.e., physics, not biology. Whatever it is in the processes in the functioning of the cell that leads to the increases in fitness is qualitatively different from that which leads to cell volume increases. Unlike fitness, being fat is not directly related to survival and is not, therefore, under the control of the principle of least action. So, the case for mutations in genes being the source of new variation to drive evolution is unsupported by the LTEE.

Thus, based on the LTEE, there is no influence on the evolution of fitness from mutations acquired during the process: Fisher was correct when he wrote of mutations: "*There is no limen of appreciable selection value to be considered*", but for the wrong reason. In terms of the MS, new variation needs to be created if evolution is to lead to the nearly infinite range of diversity we see.

However, the LTEE is not the only example of a *failure* to observe the expected response to a selective pressure in the presence of what is assumed to be genetic variation. For example, Pujol et al. note:

> " …. . *evidence for responses to selection that match predictions are often missing in quantitative genetic studies of wild populations*." (Pujol et al., 2018)

They propose that multiple biological mechanisms can unlink genetic variation from the response to selection, but equally, it may be that what they take to be genetic variation is not that, and another process (the principle of least action) is driving evolution.

If it is not the selection of genetic variance by natural selection that is driving evolution, we must accept that the gene is not Mendel's unit of inheritance: this must be the cellular phenotype.

5. History has neglected the phenotype

My argument is that the cellular phenotype, *not* the gene/genotype, plays the leading role in both inheritance and evolution. I will demonstrate (in the next section) that the cellular phenotype has the essential character of 'thingness' which is needed to be Mendel's unit of inheritance and, that the direct inheritance of the phenotype is viable. Further, I will argue that the cellular phenotype has independent agency in the process of evolution and is its biological component.

The cellular phenotype emerges from a *process* fuelled by the interaction of the products of transcription and translation of genes: it is

[46] Ibid pp 133/4.
[47] p1596.
[48] Private communication: William Basener.

represented by a quasi-stable attractor state.[49] This is the consequence of the cell being a thermodynamically open complex dissipative system. Quasi-stability means that up to a point, the attractor is robust to internal and external (environmental) stresses but beyond that point makes randomly determined transitions to variant attractor states/phenotypes. (Baverstock and Rönkkö 2008; Baverstock and Karotki 2011). In such a transition, phenotypic properties may be lost or gained. Empirical support for the instability of the phenotype (more commonly termed 'genomic instability') being initiated by low doses of ionising radiation (and other environmental stresses) is robust (Morgan 2003a, 2003b). Genomic/phenotypic instability is a biological phenomenon, unrecognised until the early 1990s (Kadhim et al., 1992), which cannot be incorporated into the Mendelian molecular genetic paradigm (Baverstock 2000; Karotki and Baverstock 2012). Furthermore, phenotypic instability *in vivo* acts over several generations as demonstrated by Huumonen et al. with *C. elegans* (Huumonen et al., 2012). A founder population of the 2nd generation offspring of irradiated worms (i.e., not directly irradiated) showed highly significantly greater diversity of gene expression (in 400 probes) after several generations of culture, compared to sham irradiated worms. Quasi-stability is a crucially important physical property of the cellular phenotype.

The cellular phenotype is governed by rules of engagement (RoE) (Baverstock and Rönkkö 2008)[50] which determine the evolution in time of the gene product composition of the interactional *process* from which the phenotype emerges. The phenotype is its own regulator: a metaphor for its role in the cell is a brain, as already noted. As pointed out by the philosopher Karl Popper in his 1986 lecture to the Royal Society in London (Niemann 2014), cells and brains can acquire knowledge, both by trial and error and from stored information. As early as the early 1900s learning behaviour was observed in single-celled organisms such as the *Stentor*. As Dennis Bray points out, that, and much other evidence, has been systematically ignored in mainstream biology (Bray 2009). Over the last two decades things are changing and, for example, the mechanisms of learning in single cells is being explored (Csermely et al., 2020).

The cellular phenotype, seen as an emergent quasi-stable state of a complex dissipative system, is, therefore, quite a different entity from that traditionally envisaged: it endows responsiveness to the environment and is a seat of 'knowledge' that gives independent agency to the cell, including in its own evolution. (We will see later that it also harbours the information that determines morphological features of multicellular organisms.)

By seeking, even in primitive ways, to improve its adaptation to its environment, an organism can modify both itself and its environment. One of the most important aspects of such modifications is improved access to nutrients. This was noted in 1835, by the British selective breeder, Edward Blyth, nearly 20 years before Darwin published, *On the Origin of Species*". Blyth wrote:

> "*[A]mong animals which procure their food by means of their agility, strength, or delicacy of sense, the one best organized must always obtain the greatest quantity; and must, therefore, become physically the strongest, and be thus enabled, by routing its opponents, to transmit its superior qualities to a greater number of offspring.*" (Blyth 1835).

Both the ability of an organism to modify its environment (e.g., find new nutrient sources, or through niche construction) and for its phenotype to be modified directly by its environment (e.g., through phenotypic instability), are necessary conditions for evolution to produce the infinite variety of species it clearly has (Waddington 2008). In evolution, everything affects everything else except in the case of organisms that live in unchanging and unchangeable environments, e.g., *E. coli* in the LTEE and, the naked mole-rat that lives only in caves.

Having independent agency means being able to choose or decide to influence the future. Humans, of course, believe that they can do that with their brains. Many will agree that other species can also do that but often a line is drawn below which this will not be the case. However, empirical evidence should lead us to expect that independent agency is an essential part of all life (Baluska and Levin 2016). For example, Darwin proposed that the tip of a plant root functioned like a brain (Baluska and Mancuso 2009). Intelligence is found in plants (Calvo et al., 2020), and many primitive organisms,[51] including single-celled microbes (Bray 2009). Brains are ubiquitous in biology.

Agency can be viewed on two levels, the macroscopic and the microscopic. The former is the actual action of the agent on its environment and the latter has to do with the processes that enable agency, within the agent. The action of choosing, being an irreversible action, entails an increase in entropy[52] to be in accordance with the 2nd law. The implications for the microstate of the organism will be addressed below.

To return to the issue of inheritance: Johannsen writes exclusively of the *genotypical* constitution of the gametes as the basis for heredity. He states:

> "*Particular resemblances between an ancestor and one or more of his descendants depend—so far as heredity is responsible—on corresponding particular identities in the genotypical constitution, and, as we have urged*

[49] Attractor states are emergent properties of complex dissipative, or non-equilibrium, systems. They are the product of self-organisation and involve dynamic steady states, where two or more counteracting processes are balanced one against the other. A simple example is the soliton: https://en.wikipedia.org/wiki/Soliton. I As a wave in water, the soliton is a dynamic steady state between the wave's tendency to dissipate, or break up into smaller waves, and to 'break' in shallow water (as seen on the shoreline). When these two processes are balanced, the soliton is formed. Such solitons were discovered by naval architect, John Scott Russell, on the Union Canal in Edinburgh in 1834. Solitons have highly counterintuitive emergent properties, for example, the velocity of the wave is proportional to its amplitude. Such non-equilibrium states are described as quasi-stable because they have what is called a boundary of attraction within which they are stable and outside of which they cease to exist. The concept of the boundary of attraction is better seen in another attractor state, the 'bicycle/rider' system. The rider keeps the bicycle in the upright position by shifting the centre of gravity of the system to the right or the left and turning the front wheel to the left or right. The system is only stable if all four 'dynamic dimensions' are within certain limits (the boundary of attraction), and freely accessible within those limits. It is true that the stability of this system is aided by the gyroscopic effect of the rotating wheels but it is possible to maintain the upright position while stationary and it is impossible to ride a bicycle with the front wheel in a tram track. The quasi-stability of the cellular phenotype is critical to the functioning of the cell.

[50] The dynamic state of the cell (the phenotype it is expressing) is governed by ongoing 'rules of engagement' (RoE) applicable to the gene products resident in the cell. The RoE form a nonholonomic record of the history of the species to which the cells belong. If, at a particular time, the RoE require a gene product that is available, the attractor state (phenotype) is stable: if it is not available a transition to a variant attractor and, therefore, phenotype, can occur – a direct transition between phenotypes (that is without modifying the genotype, as is the case in cell differentiation). This is possible because of the physically quasi-stable nature of the phenotype. The RoE constitute the *syntax* of the cell/system. See also: Baverstock, K. (2016). "Genes without prominence." Inference: International Review of Science **2(2)**.; specifically, under "Gradualism is not an option".

[51] For example, there is extensive interest in amoeba, where the ability to solve mazes has been demonstrated: Nakagaki, T., H. Yamada and A. Toth (2000). "Maze-solving by an amoeboid organism." Nature 407(6803): 470. Amoeba also exhibit memory in relation to nutrient sources: Kramar, M. and K. Alim (2021). "Encoding memory in tube diameter hierarchy of living flow network." Proc Natl Acad Sci U S A **118**(10) and a continuous spectrum of behvioural states: Fleig, P. et al. (2020) "Emergence of behaviour in a self-organised living matter network"; https://doi.org/10.1101/2020.09.06.285080.

[52] https://arxiv.org/abs/2007.05300 (accessed 20.02.2021).

here, perhaps to excess, the genotype is not a function of the personal character of any ancestor. … … The genotype-conception is thus an 'ahistoric' view of the reactions of living beings—of course only as far as true heredity is concerned." (Johannsen 1911)

But neither do the phenotypes of the gametes reflect the personal character of ancestors. So, why cannot the *phenotypical* constitution of the gametes be the basis for inheritance? I am proposing that the zygote is the product of the fusion of the two parental gamete phenotypes.

The genotypes and the phenotypes of both parents are present in the zygote, so can their roles be separated from one another? Yes, if mutations are not responsible for the variance that drives natural selection in evolution, then it cannot be the genotype/gene that is the unit of inheritance. If agency is what is moving evolution forward the phenotype must be responsible. We will come back to discuss the implications of this in the Discussion.

In the animal kingdom inheritance involves the fusion of the sperm and egg. Human genetics is the concern here so I will pursue the argument further in the context of human, rather than plant or other organism reproduction.

6. Human inheritance

A first point to make is that there are ancestral similarities in the lines of inheritance of humans: an example is the chin and nose in the Hapsburg line, which have been obvious over many generations. However, the Norwegian geneticist, Stig Omholt, wrote in 2013:

"*There is no a priori reason why an offspring, arising from the random sorting of chromosome pairs plus genetic recombination and the subsequent immense number of highly complex and nonlinear processes making the individual, should on average resemble its parents more than a randomly drawn couple from the population. We have no theory that tells us why this would not give rise to a quite unpredictable parent-offspring relationship*." (Omholt 2013)

This view sharply contradicts that advanced in the very influential and still highly regarded paper on heredity based on pure Mendelian inheritance by Fisher in 1918 (Fisher 1918). So, who is correct?

Ken Richardson in his book "*Genes, Brains and Human Potential: the science and ideology of intelligence*", (Richardson 2017) heavily criticises Fisher's approach of applying the principles of Mendelian inheritance to continuously varying traits. That inherited factors that vary continuously, like human height, were likely to involve several genes, was recognised by Fisher. His solution was to treat these additively as if they were independent of one another. Richardson points out that it is extremely unlikely that several genes acting together to produce a trait, would do so independently of one another, i.e., without gene-gene or gene-environment interactions. This view is confirmed by Rice and Borecki who write:

"*Resolving the various sources of familial resemblance entails other issues [than additive variance]. For example, there may be major gene effects that are largely or entirely nonadditive, temporal or developmental trends, and gene-gene (epistasis) and gene-environment interactions*." (Rice and Borecki 2001).

Genes acting independently underlies Fisher's approach: he notes:

"*… throughout this work it has been necessary not to introduce any avoidable complications*" (Fisher 1918).

One such was interaction between genes, i.e., epistasis, was, in Fisher's mind, an avoidable complication. Furthermore, Fisher's approach predicted very little impact of the environment on inheritance, but we know that cannot be true. For example, a study of birth cohorts from 1886 to 1994 in 143,390 twin pairs estimates heritability to be between 0.69 and 0.84 for men and between 0.54 and 0.78 for women (Jelenkovic et al., 2016). Fisher's view that Mendelian inheritance is sufficient to account for human inheritance can, thus, be rejected.

Omholt and his colleagues were able to construct a complex mathematical model introducing some new concepts, for example, monotonicity (Gjuvsland et al., 2013) to get out of this inconvenient hole but is it not simpler and more intuitive to say that inheritance is direct phenotype to phenotype inheritance, i.e., Johannsen's transmission-conception at the level of the phenotypes of the gametes?

To see how this will work we need to address first the processes of development and morphogenesis from the formation of the zygote to adulthood.

In the male and female gametes, complex changes occur in the genotypes before fertilisation of the egg by the sperm. However, sperm and egg are both functioning cells with phenotypes represented by attractor states. The sperm, although in partially suspended animation through protamine condensation of much of its chromosome content, contains gene products essential for successful fertilisation, as does, of course, the egg (Krawetz 2005). Thus, when the sperm head enters the egg, two *functional* complex systems occupy the egg cytoplasm. As attractor states they are discrete (behave like particles due to their boundaries of attraction) and do not *blend* one with the other.

As already noted, one of the simplest examples of an attractor state is a soliton or solitary wave. They are a useful metaphor for the physical aspect of the cellular phenotype *albeit* infinitely simpler. The soliton can be described thus: a non-equilibrium dissipative dynamic steady state that adopts (self-organises into) an attractor state. As waves in water, the dynamic steady state is between dispersion and breaking of the wave. These counteracting processes effectively trap the excess energy in the wave. The environment is important if the wave is not infinitely broad as is the case for a tsunami. In a canal, river, or long water tank, the energy in the wave is prevented from escaping laterally by the banks or walls of the containment. The environment thus plays a crucial role in the stability of the soliton and indeed any self-organised state, including the cellular phenotype.

Solitons are discrete states in whatever medium they occur, and they exhibit counter-intuitive emergent properties. For example, if two solitons collide, they simply pass through one another without exchanging material: they do not blend. Like solitons, attractor states (representing the phenotypes) have the properties that Mendel stipulated for the units of inheritance, i.e., particle- or thing-like and unitary.[53]

Where two dynamical systems co-exist in proximity, they tend to synchronise (Yang 1999). This behaviour was observed in the mid-1600s by the Dutch physicist, Christian Huygens when he was experimenting with pairs of pendulum clocks to measure time (fix longitude) on ships at sea in relation to that at their home ports (Oliveira and Melo 2015). A situation that is more relevant to zygote formation arises in artificial intelligence, where two (or more) artificial neural networks (attractor states) can be merged, retaining the properties of both. Another more intuitive example might be the merging of two companies manufacturing related products. The two could become one, operating from a combined manufacturing base: the full range of products from both companies could continue to be produced. Synchronisation in natural systems is a natural process because it minimises the energy of the coupled system.

Thus, the production of the zygote from two cellular attractors, without losing or diluting the individual characters of each (i.e., not blending), is certainly not ruled out as the basis of inheritance.

The human zygote is a single cell and, in the process of development, must be able to differentiate and proliferate into at least some 230 tissue or organ specific cell types in the human body: almost all[54] will have the same genome sequence. Today, it is believed that these different cell

[53] Robert Olby in Mendel, Mendelism and Genetics: http://www.mendelweb.org/MWolby.html. (accessed 24:07:2020).

[54] Some exceptions are some cells in the immune system, and red blood cells that have no nucleus, and thus no genomic DNA.

types derive their tissue identity from so called 'marking' on the chromatin,[55] primarily methylation of DNA. These marks serve to allow the expression/repression of specific genes that characterise the cells of specific organs or tissues. Either before, or immediately after fertilisation, the participating genomes are 'cleaned' of their marks,[56] and, after fertilisation reprogramming of the marking starts (Reik et al., 2001). The problem is: what is the source of the information that re-programmes the genome? That information cannot be coded in the DNA.[57] This is a major problem for the conventional theory of development.

The so-called epigenome, the specific patterns of marks on specific genes, is crucial at the zygote stage: the zygote cannot be pluripotent unless all marks that influence cell fate have been erased. As the genome sequences for several species, e.g., human, mouse, bonobo, etc. are remarkably similar, the question: "how does the zygote know 'I am a mouse', or 'I am a human', or 'I am a bonobo?" is a valid question if it is to rely on the gene sequences alone. The origin of the information that places the vital 'marks' is a major open question for conventional biology, even in respect of mitosis of somatic cells, when the marks have to be replicated on the daughter cells:

> "*Cellular specialization during development is based on the ability to establish, maintain, and execute different gene expression programs. How transcriptional programs are established during development and maintained in cycling cells is a fundamental question in biology. Chromatin organization plays a fundamental role in this process, but it remains unclear how specific chromatin states are stably inherited from a mother cell to its daughters.*" (Alabert et al., 2020)

The ignorance here is fundamental and critical to any theory of development that involves genes as the unit of inheritance.

Artificial life studies, however, point the way to an alternative way of looking at development and morphogenesis. Mauno Rönkkö, a Finnish computer scientist has developed an artificial ecosystem based on a *deterministic* particle system. The ecosystem consists of soil, grass, water pools, rain, worms, and beetles, etc. (Ronkko 2007). When the programme is run, the grass grows when it rains and releases a scent which the worms are attracted to: they eat the grass, and the beetles hunt them. *All* these 'components' of the ecosystem are composed of individual particles each with information relating to how they interact with other particles immediately adjacent to them. The remarkable thing is that, out of what is a fully deterministic system, emergent and lifelike properties are realised. Rönkkö writes:

> "*We analyzed the dynamics of six nontrivial scenarios: formation of rivers and ponds, grass growing in rain, worms finding edible grass, a beetle jumping and correcting its orientation, beetles hunting worms, and the environment affecting the global dynamics. Each of these scenarios exhibited distinct emergent dynamics, and in each scenario the dynamics showed nonmechanistic, unpredictable, and sometimes even spontaneous characteristics.*" (Ronkko 2007).

The point is that the information that enables these dynamic features to emerge is distributed across the whole ecosystem attached to the numerous individual particles. There is a very interesting parallel here with the most recent understanding of the process of cellular metabolism. According to De la Fuente et al. the numerous enzymes responsible for metabolism, self-organise into a global 'metabolic network attractor'. The authors say:

> "*The self-organization of cooperating enzymes into multienzyme complexes, seem to be central features of cellular metabolism, crucial for the functional activity, regulation and efficiency of biomolecular processes and fundamental for understanding the molecular architecture of cell life.*" (De la Fuente et al., 2013).

Enzymes/proteins are, of course, information-carrying molecules hence the parallel with Rönkkö's ecosystem is close.

The life-like character of the artificial ecosystem led us to propose that the individual cells in the body of a multicellular organism harbour the necessary information, embedded in their phenotypes, to construct, at all stages of development, the bodily structure of an organism (Baverstock and Ronkko 2014). All the spatiotemporal complexity of the process is bound up in what is essentially self-organisation based on knowledge,[58] individual cells being informed by the RoE.[59]

The flocking of birds and the shoaling of fish are two very simple metaphors for this phenomenon. Individuals (the equivalent of the cells) are obeying local rules about their position in relation to their immediate neighbours. This *alone* leads to a coherence in dynamical emergent behaviour across the flock or shoal. The information in Rönkkö's artificial ecosystem is of exactly this character. It elicits a coherence in the dynamical behaviour across the ecosystem, which gives the emergent quality of 'life-like': it is, of course, not 'life'; it just has that character, but it shares with 'life' the phenomenon of emergence despite its deterministic origin.

The conventional explanation for morphogenesis is a 'toolkit' of homologous *hox* genes which:

> " …. . *being highly conserved among phyla; … … generate the patterns in time and space which shape the embryo, and ultimately form the body plan of the organism*"[60].

It is difficult to imagine how genes, composed of DNA, a passive molecule, can generate "*patterns in space and time*".

The British scientist Alan Turing (1990) proposed a mechanism for self-organised morphogenesis, (described in detail (Schweisguth and Corson 2019)) which is essentially self-organisation based on a reaction-diffusion mechanism. Schweisguth and Corson, reviewing the evidence for self-organisation in morphogenesis, conclude that there is:

> " …. . *unambiguous evidence for self-organisation in tissue patterning*". (Schweisguth and Corson 2019).

It is, therefore, reasonable to postulate that, following the transmission of the phenotypes of the gametes of two parents to their offspring, that an offspring develops into an adult in such a way that is *not* driven by the transferred parental genes but rather, is properly seen in terms of self-organisation based on local 'knowledge' retained in

[55] Acetylation of chromatin, methylation of DNA and structural features of the histones comprising chromatin.

[56] The zygote starts life as a single, so called, pluripotent stem cell capable of differentiating into any other more specialised and functional, cell type. Therefore, all inherited marks from the parental genomes must have been removed.

[57] If the DNA carried the instructions to apply the marking, it would have to carry at least 230 variants for the ~230 cell types in the human but how could it know, in any specific instance, which one of those to express? It could not. So, another source of information, independent of the DNA, is required. This is the fallacy that underpins the current paradigm for biology, based on a genetic regulatory network (GRN): Baverstock, K. (2011). "A comparison of two cell regulatory models entailing high dimensional attractors representing phenotype." Prog Biophys Mol Biol **106**(2): 443–449.

[58] The philosopher Karl Popper draws a distinction between knowledge and information *per se*: knowledge is either derived through trial, error, and elimination, or derived *from* stored information. Information without a contextual framework within which to evaluate it is valueless: Niemann, H.-J. (2014). Karl Popper and the two new secrets of life: including Karl Popper's Medawar lecture 1986 and three related texts. Tübingen, Mohr Siebeck. In this case 'knowledge' is the more appropriate term because it derives from information held in the gene sequences that specify the gene products.

[59] Consider the development of a ball and socket joint. Throughout the development process these two separate components must match one another very closely. This requires that both components share complementary spatial and temporal knowledge from which to build the joint.

[60] https://en.wikipedia.org/wiki/Evolutionary_developmental_biology (accessed 23.02.2021).

individual cells and deriving from the RoE. (Baverstock and Ronkko 2014). The zygote 'knows' what it will develop into quite independently of its genotype. However, the cell can only know that if the genotype can provide the appropriate materials in the form of gene products (Nijhout 1990). Genes, as 'builder's merchants', must provide gene products of the highest integrity, hence the lengths the cell goes to, to preserve the integrity of its DNA.[61]

That self-organisation commences at a very early stage in human embryogenesis has been empirically established (Deglincerti et al., 2016; Shahbazi et al., 2016; Shahbazi et al., 2019). In the period of 7–14 days after fertilisation, the period following implantation, critical re-modelling of the embryo occurs. Sharbazi et al. report:

> "… *events at this stage of human development are embryo-autonomous highlighting the remarkable and unanticipated self-organising properties of human embryos*." (Shahbazi et al., 2016)

These results were obtained with an *in vitro* implantation system; there was, therefore, no input from the mother at this critical stage. This is the clearest possible evidence that development and morphogenesis are processes of self-organisation and are not dependent on genes except for the required gene products.

There are, therefore, *two* sources of information that specify the organism: 1) the RoE, which are species specific and inform all cells in the organism, and 2) the genomic DNA sequence. The former acts as the formal system syntax and the latter provides the gene products upon which the system draws to yield the phenotype. This means that the nucleus of the cell housing the genes is placed outside the system: it is treated as an organelle: a 'planet' in the cell's 'solar system' with the cytoplasm as the 'sun'.

What is described above is a system view of life based on a hierarchical self-organised structure out of which life emerges. This is very different from the prevailing view, even when that is presented as 'system biology'. Writing in 2011 in the journal *Cell*, under the title, "*The Cell in an Era of Systems Biology*", Paul Nurse[62], wrote:

> "*Our view is that scientific explanations and methodologies are essentially reductionist in nature. However, although it is difficult to imagine a scientific enquiry or explanation that is not reductionist, it is important to keep a focus on the behavior of whole systems in biology and to understand how the interactions and processes brought about by component parts acting at lower levels in a system are constrained by overall functions acting at higher levels*." (Nurse and Hayles 2011).

Reductionism might be viable for simple systems (machines) but it is totally inappropriate for complex systems.

Nobelist and physicist, the late P. W. Anderson warned in1972 that while it is possible to reduce a system to the basic laws that govern it:

> " …. *in general, the relationship between the system and its parts is intellectually a 'one-way street'. Synthesis [from underlying laws] is expected to be all but impossible; analysis, on the other hand may be not only possible but fruitful in all kinds of ways*." (Anderson 1972).

7. Discussion

In this paper, I have analysed how, under the appropriate branches of physics, cells would function at the single and multicellular organism levels. Two branches of physics are involved, namely, thermodynamics and complex dissipative system dynamics. The cell, the basic building block of organisms, is self-evidently a thermodynamically open, complex dissipative system. I have called the model of the cell based on this foundation in physics, the "Independent Attractor", or IA model, the term 'independent' representing the model's relative independence from the gene as an active functional element.

The analysis reveals several features that are at odds with the prevailing view of biology, most notably:

1. No empirical support for gene mutations providing selectable variance to drive evolution,
2. Therefore, cellular phenotypes, not genes, must be Mendel's units of inheritance,
3. Environmental stress can trigger unscheduled direct phenotype to phenotype transitions (phenotypic instability),
4. The LTEE proves that natural selection is a physical manifestation of the 2 nd law,
5. Development and morphogenesis in multicellular organisms are processes of self-organisation based on knowledge carried by the cellular phenotype.

Relying on a decision, which, among other things, might have been influenced by a specific plant breeding technique (Johannsen 1911) and two highly influential, but flawed, works by Fisher 1918, 1930 but with no compelling biological or physical insight, the genetic community has adopted Johannsen's genotype-conception and the role of the gene in that, as the functional basis for the four main elements of biology, namely, inheritance, evolution, development, and morphogenesis. The gene is, therefore, fundamental to biology as it is viewed today. There is a famous and widely accepted statement made in 1964 by the evolutionary biologist and founding influence on the MS, Theodosius Dobzhansky: "*nothing in biology makes sense except in the light of evolution*". (Dobzhansky 1964). If Lenski's LTEE is anything to go by, evolution does not make sense in terms of genes, so then, neither does biology. Evolution has undoubtedly taken place (Dawkins 2009) and mutations to the genomic DNA have undoubtedly been accrued (Lynch 2010) but the two, based on the results of the LTEE, are apparently not connected. Had Lenski adopted a more conventional approach and set up the LTEE to test a clear hypothesis,[63] Fisher's 1930 theory of natural selection would have been rejected by evidence as early as the mid-1990s.

If genetic variance does not drive evolution, then genes are not the units of inheritance. The alternative is the gamete phenotype which, being an attractor state with a boundary of attraction, has the 'particle like' character that Mendel demanded.[64] The inheritance of likenesses based on genes being the units of inheritance, as claimed in Fisher's 1918 paper, based on purely additive Mendelian inheritance (Fisher 1918), is not credible but is uncritically lauded as the basis of the modern genetic technology of GWA (Visscher and Bruce Walsh, 2019; Visscher and Goddard 2019). Additionally, Crick's crucial 'sequence hypothesis' (governing the transfer of information from the DNA sequence to the gene products which generate the phenotype) is invalid (Baverstock 2019a), even apparently in the view of Crick himself in 1970 (Crick 1970) [see above]. Omholt maintains that the genotype to phenotype map for complex traits is far from straightforward. He writes:

[61] DNA, under physiological conditions is subject to continual degradation due to hydrolysis and oxidation. Baverstock, K. (1991). "DNA instability, paternal irradiation and leukaemia in children around Sellafield." Int J Radiat Biol **60** (4): 581–595. Considerable resources of the cell are devoted to detecting and repairing this damage to maintain the integrity of the gene products that the cell requires to function correctly.

[62] Formerly President of the UK Royal Society and now Chief Executive and Director of the Francis Crick Institute, London.

[63] Such as: "fitness will evolve according to the theory of natural selection proposed by Fisher, with due modification to allow for the influence of mutations occurring during the evolutionary process".

[64] Robert Olby in Mendel, Mendelism and Genetics: http://www.mendelweb.org/MWolby.html. (accessed 24:07:2020).

> "*It should be noted that … … … …. there is no direct causal arrow from genotype to phenotype in the sense that DNA is responsible for exerting a direct effect as a sub-system on the system dynamics.*" (Omholt 2013).[65]

That absent causal connection between alleles (or mutations), and the phenotype is essential if the currently extremely active technology of GWA is to measure traits. However, GWA does measure something. Typically, as already noted, a large population, bearing a common disease or behavioural trait, will exhibit hundreds to thousands of SNPs, at almost as many loci, each contributing a very small effect to the genetic risk, when compared with a control population of similar size. Even though mostly there is no biological rationale as to why these SNP/loci would be associated with the trait, the small contributions from each SNP are added up into a PGS, from which it is claimed diagnoses of the trait can be made (Plomin 2018). Population stratification, geographical and social, could confound associations between SNPs and traits, (Sanderson et al., 2021) if those existed. As it appears that they don't, the effect of stratification is to produce false positive results, thus, falsely increasing confidence in PGSs as diagnostic tools. For example, in a study in Finland, PGSs for five common diseases and three complex traits were calculated for 2376 individuals whose parents had lived in a known specific geographical location. Within Finland, there is a well-defined genetic population structure, with an east to west divide (Kerminen et al., 2017). For all but one of the five disease traits and one of the three complex traits, the PGSs detected the geographic structure (indicating where the individual was born) and not the distribution of the trait (Kerminen et al., 2019). In most studies the background genetic structure is not as well-known as it is in Finland: it is, therefore, in most cases, not possible to discriminate between measurements of genetic background and false positives (Richardson and Jones 2019).

Since 1983 at least, when Barbara McClintock presented her Nobel Prize lecture entitled "*The Significance of Responses of the Genome to Challenge*", it has been known that environmental stress or shocks, can induce phenotypic changes in terms of karyotypic rearrangements in the genomes of maize (McClintock 1984). This same phenomenon has been established in many other experimental systems with diverse endpoints, initially in radiobiology, as radiation-induced genomic instability (Morgan 2003, Morgan 2003; Kadhim et al., 2013), or simply, genomic instability, as it can also be induced by other environmental stresses (Karotki and Baverstock 2012). Genomic, more appropriately, phenotypic, instability, a direct (without involving the genotype) transition from one cell phenotype to another, is in effect unscheduled and undirected cell differentiation. There is, therefore, no rational basis for GWA studies, or, therefore, for PGSs. (Baverstock 2019b).

Therefore, an important implication of the phenomenon of phenotypic instability is that common diseases and behavioural traits are not genetic, but rather are caused solely by environmental stress. The justification for assuming a substantial genetic component in such traits derives from studies of twins (Plomin 2018), which are, in any case, compromised by the flawed 'equal environments assumption' (Joseph, 2015). However, if genes are not the units of inheritance, then estimates of heritability based on genotypes are meaningless.

That the fitness profiles of the 12 independent experiments of the LTEE are not compatible with Fisher's law of natural selection, even when the impact of ongoing mutation, modelled by (Basener and Sanford 2018), is incorporated, and the fact that the 12 fitness profiles are identical, although each experiment is acquiring different mutations (Maddamsetti and Grant, 2020), shows that gene mutation is not the source of variance that is being selected to improve fitness. On the other hand, the fit to a power law of the evolution of fitness in 12 independent experiments over more than 50,000 generations indicates that natural selection is driven by physics, as proposed by Sharma and Annila (2007). The physics is the principle of least action (De Maupertuis, 1746), which is synonymous with the 2nd law, with nutrient the driving source of free energy. Blyth's remark in 1835, drawing attention to the importance of nutrient for the survival of organisms (Blyth 1835), taken together with Maupertuis' principle would seem to be a sufficient explanation for evolution in the LTEE and likely in evolution generally. That cell volume in the LTEE bacteria does not increase identically in all colonies should not be a surprise. It has been clearly demonstrated that individual bacteria in a genetically pure colony differ significantly in their chemotactic behaviour (Salek et al., 2019a,b). The heterogeneity in behaviour is said to derive from variation in gene expression. It can be assumed that cell volume, unlike fitness, is not governed by the principle of least action and is not, therefore, a selectable property.

Maupertuis' principle dictates that a free energy disequilibrium (in this case between the environment or ecosystem and an organism) will be levelled as efficiently as local conditions permit. In the LTEE, at the start, the 12 experimental flasks contain equal numbers of genetically identical bacteria, provide identical environments, and identical quantities of available nutrient, specifically the same limiting concentration of glucose. In each succeeding 24 hours the glucose is exhausted more rapidly, at least up to 60000 generations, but that difference diminishes every day, and unless the governing power law is truncated at some point that pattern of behaviour will continue indefinitely. The LTEE is an evolving system that has 'nowhere to go' because the environment is unchanging, and unchangeable by the bacteria. In the real world, as noted above, organisms modify their environment through, for example, niche construction, and the environment modifies phenotypes through stress and phenotypic instability. The former is the exercise of agency by organisms, which push back on the local conditions to maximise, as much as possible, the flow of free energy from the environment: this is the biological component of evolution.

The evidence of autonomous self-organisation at the earliest stages of embryogenesis (Shahbazi et al., 2016; Shahbazi et al., 2019) is a compelling indication that the whole process of development and morphogenesis is based on self-organisation. The *information* for that process derives from the RoE and is stored in the cellular phenotype but is contingent on the environment in which it takes place. The life-like behaviour of Rönkkö's fully deterministic virtual ecosystem (Ronkko 2007), when considered holistically, provides a powerful argument for the proposed role of self-organisation in development (Baverstock and Ronkko 2014).

Thus, starting with the self-organised metabolic network that extracts energy from nutrient (De la Fuente et al., 2013), through the self-organised gene-product network from which cellular phenotype emerges (Baverstock and Rönkkö 2008), to the self-organised structure of cells that is the organism (Baverstock and Ronkko 2014); a self-similar 'matryoshka' on three hierarchical levels is revealed. Being alive, including consciousness, independent agency, and what they entail, is rooted at the cellular level and is an emergent property of interacting gene products. This is the system/cell microstate that allows agency. Free energy is dissipated, and the *entropy* in the system is increased,

[65] Omholt's paper is highly innovative, building on ideas going back a few decades. He acknowledges that thermodynamically open systems can include both living and non-living systems: "*Living systems do not have exclusive ownership to phenomena like self-assembly, self-organisation, emergence, two-way causation between lower- and higher-level system dynamics features, and order creation through local reduction of entropy*." (p76). However, it is a misapprehension that the creation of order requires a local reduction in entropy: see Annila, A. and K. Baverstock (2016). "Discourse on order vs. disorder." Commun Integr Biol **9**(4): e1187348. It does not. Secondly, the genotype, the DNA, is invoked as being what makes the difference between the living and the non-living, in that it enables the living system to self-transcend its morphological constraints. However, Rönkkö demonstrated that artificial life beetles, embedded in their artificial life ecosystem, transcended their morphological constraints as deterministic objects, comprised of information bearing particles: Ronkko, M. (2007). "An artificial ecosystem: emergent dynamics and lifelike properties." Artif Life **13**(2): 159–187. There is no distinction between the animate and the inanimate. To claim such is vitalism: Annila, A. (2020). Back to Reality. New York, Privus Press. See p. 237.

appearing in the system as growth and information (Annila and Baverstock 2016). Recall that Boltzmann said that organisms sought entropy first and foremost (Boltzmann 1974).[66]

Ken Richardson points out that it is not the case that evolution is taking place in the stable environment envisaged by the MS, or according to genetic programmes but rather to inducible covariation grammars ('biogrammars') (Richardson 2020).[67] This is what Conrad Waddington foresaw in his "*Paradigm for an* [open-ended] *Evolutionary Process*" published originally in 1969 (Waddington 2008). This is also what Fisher sought, based on genetics, in 1930 (Fisher 1930) but failed to achieve (Basener and Sanford 2018). Richardson's ideas on biogrammars are in tune with Robert Rosen's theory of organisms as anticipatory systems where:

> "*an anticipatory system is a natural system that contains an internal predictive model of itself and of its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a later instant*" (Louie 2010).

As already noted, genes are like the merchants that provide the necessary materials to build a house: they are neither the architect, nor the builder but, without them, the house cannot be built. Put more formally, genes are neither the formal cause (the blueprint), nor the efficient cause (the builder) of the cell, nor of the organism: they provide the material cause, the gene products (Nijhout 1990). The *formal* cause is embedded in the RoE and the *efficient* cause in the phenotype.

The fundamental problem with the traditional 'gene-centred' view of biology that prevails today is that the information in the DNA sequence is taken to be both the formal *and* the efficient causes of the organism. In his modelling relationship, Robert Rosen emphasises the necessity of having both syntax and semantics (Rosen 1991). He asks if it is possible to define a language in terms of a formal syntax alone: it is not; a semantic component (vocabulary) is needed.[68] This is the issue that came to light with the problem of how the pluripotent zygote, stripped of its chromatin markings, could 'know' what it was from its DNA sequence alone. A second independent source of information was required and is assumed to be chromosome marking, but the origin of the information that places the marks is not known. This is the same problem as Gödel's incompleteness theorem and Turing's halting problem. The view of biology based on genetic regulatory networks alone is fundamentally flawed (Baverstock 2011).

Information is widely regarded as a key feature of biology. For example, Nurse, in a lecture[69] in Oxford in March 2020, refers to "*life as information*", linking it to "*complex systems, their control and purpose.*" However, Nurse's idea of a complex system is very different from that being discussed here. Control, in the form of homeostasis, he proposes, is imposed through feedback mechanisms as originally proposed by Jacques Monod and François Jacob in connection with gene regulation in the *lac* operon[70] (Jacob and Monod 1961). Ludwig van Bertalanffy, in his "*General System Theory*", while acknowledging a role for feedback in biology, proposes that where biological organisation and homeostasis are concerned, his principle of *equifinality* applies (Bertalanffy 1969). This principle is precisely what we have discussed here in terms of attractor states and self-organisation. Feedback is primarily a feature of machine-based, or *complicated*, systems, whereas equifinality is a feature of *complex*[71] systems, although the former may be found embedded in the latter. In the IA model homeostasis is an emergent phenomenon arising in a complex system.

Since Boltzmann formulated his molecular theory of entropy in 1877 many physicists and biologists have assumed that increasing entropy *inevitably* means increasing disorder. This, however, only applies in closed systems: organisms are thermodynamically open, and increasing entropy is fully compatible with increasing order and complexity (Annila and Baverstock 2016). There is a long history of physicists proposing complicated ruses to lower the internal entropy of an animate system to allow it to 'get around the 2nd law'.[72] Perhaps the best known is Schrödinger with his concept of negative entropy, or negentropy.[73] Many think that his book, "*What is Life*?" (Schrödinger 1944), has had a profound influence on biological thinking. He was, however, mistaken, not only in the context of the role of entropy but also in his metaphor for an organism as a 'clockwork'[74] or complicated system. As already noted, but worth repeating: given the right components, the appropriate environment, and an injection of free energy, self-organisation is a natural consequence, be it solitons, life, or consciousness (at the cellular level).

The preliminary announcement of the results of the HGP in 2001 decisively ended the 'one gene: one polypeptide' assumption. The result was unexpected by the genetics community[75]: it signalled that they had been working under a serious false assumption. In a Commentary in the journal *Nature* in February 2021 (Gates et al., 2021), the authors list the benefits they see from the HGP. These include the catalogue of protein coding genes but, they point out, not a clearer definition of what constitutes a gene. This paper is a perspective of data scientists: it does not address what the HGP has taught us about the processes underlying biology, how it works, and how we can better understand common diseases. Particularly in advancing the understanding of common disease, I would argue that the HGP has not realised what was promised in the early 1990s.

Finally, the introduction of agency as a property of organisms is a challenge to how biological phenomena can be investigated. Biology is

[66] "The general struggle for existence of animate beings is therefore not a struggle for raw materials—these, for organisms, are air, water, and soil, all abundantly available—nor for energy which exists in plenty in any body in the form of heat (albeit unfortunately not transformable), but a struggle for entropy, which becomes available through the transition of energy from the hot sun to the cold earth".

[67] Richardson draws a parallel with how human speech is understood. Several features of speech sounds, including their timing, pitch, duration, etc., covary to define a word. The trained brain is adept at decoding these features simultaneously. Similarly, the organism is receiving information about its environment on several levels and using what Richardson calls biogrammars to ensure survival, to optimise available resources and to anticipate threats.

[68] See p 44 in: Rosen, R. (1991). Life Itself: a Comprehensive Inquiry into the Nature, Origin and Fabrication of Life. New York, Columbia University Press.

[69] https://www.youtube.com/watch?app=desktop&v=92oMfkuOIlA&ucbcb=1 (accessed 10.04.2021).

[70] https://en.wikipedia.org/wiki/Lac_operon (accessed on 10.04.2021).

[71] The distinction between the terms 'complex' and 'complicated' is important. It is often said that the mark of a complex system is that its output is 'more than the sum of its parts' and this results from the more than additive interaction between the components of the complex system. On the other hand the outut of a complicated system is simpy the sum of iits compoment parts. The concept of feedback, applicable in both, is prominent in cybernetics and, in the context of biology, can give rise to homeostasis, as, for example, in regulating body temperature. Bertalanffy's principle of equifinality refers to the concept of an attractor state where, if the initial state of a system lies within the boundary of attraction, it will reach a specific final state from wherever it starts. This applies only in thermodynamically open systems and yields homeostasis as an emergent property. This is the origin of the order that Stuart Kauffman discusses in his book: Kauffman, S. A. (1993). The Origins of Order: Self Organisation and Selection in Evolution. Oxford, Oxford University Press, and is at the root of the self-organisation proposed in the IA model, through the phenomenon of emergence.

[72] Other examples: Nicolis, G. and I. Prigogine (1989). Exploring complexity: an introduction. New York, W.H. Freeman; Penrose, R. (2011). Cycles of time: an extraordinary new view of the universe. New York, Alfred A. Knopf.

[73] Which he regrets at some length in a note at the end of Chapter 6 in: Schrödinger, E. (1944). What is life? Cambridge, Cambridge University Press.

[74] Ibid. See p 89.

[75] In a sweepstake organised in 2000 and drawn in 2003, the winning prediction, the lowest prediction out of 460 bets, was 25,947. The current estimate is between 22,000 and 25,000.

governed by physics and implemented by biochemistry. In his 1986 lecture, Popper (Niemann 2014) was challenged by Max Perutz on his assertion that biochemistry was not reducible to chemistry (Rose 1988).[76] Rose was initially opposed to Popper's assertion but subsequently came to agree with it, noting that:

> "…. *while the problems of chemistry concern molecular structures in their own right, those of biochemistry concern the function of those molecules within a system and neither the system nor the function of the molecule within it can be explained merely from a study of the molecule itself.*"

What Nurse, from his reductionist standpoint, calls the "*component parts acting at lower levels of the system*", is the crowded milieu of "*intrinsically disordered proteins*" where chaos prevails (Fonin et al., 2018) but out of which, according to the framework described here, the phenotype emerges and the cell is regulated, i.e., Nurse's "*functions acting at higher levels*" emerge.[77] The gene products are not the tidily folded native protein structures that chemists and reductionists envisage, and Crick predicted in his sequence hypothesis. Popper was correct in that such a 'biochemical' system cannot be reduced to conventional chemistry. Further, the intrinsically disordered proteins, and regions of proteins, are thought to play a key role in the process of learning (Csermely et al., 2020), which is essential to knowledge acquisition and, therefore, to agency. Furthermore, the knowledge upon which agency is based is internal to the system and cannot be inferred from external observations. The framework presented here requires that a very different approach be taken to gain an understanding of biology than has heretofore been applied.

In invoking agency as the biological component of evolution, the inheritance of acquired characteristics is invoked. Blyth speaks of an organism's "*agility, strength, or delicacy of sense,"* as being characteristics important in acquiring nutrient. These are characteristics that can be acquired, and for evolution to advance, need to be inherited. Is this a fatal flaw in the IA model given the Weismann Barrier and Crick's Central Dogma? Johannsen was at pains to exclude the inheritance of acquired characteristics (see EN 25) and opted for the genotype/gene as the inherited component. In his book on biological relativism: "*Dance to the tune of life: biological relativity"* (Noble 2017), Denis Noble notes that Weismann's 'surgical' evidence for his barrier is weak.[78] Noble points out that Conrad Waddington's experiments on 'genetic assimilation' in response to environmental changes (Waddington 1942), although dismissed by Neo-Darwinists as phenotypic plasticity, is the inheritance of acquired characteristics under specific environmental conditions. Furthermore, he suggests that Crick's Central Dogma (Crick 1970) is "*better represented as an important chemical fact about coding* …" rather than a universal principle of biology. Therefore, agency in evolution, as invoked by the IA model is not precluded by evidence.

Further, the framework based on self-organisation has a plausible explanation for abiogenesis (Annila and Baverstock 2014) in terms of Alexander Oparin's theory of the origin of life (Oparin 1953), as modified by Freeman Dyson (1999). The kind of complex self-organised system proposed, with its emergent properties, is one that Murray Gell-Mann would expect to exhibit consciousness (Gell-Mann 2001) and, therefore, agency. When did consciousness arise? In the context of the two stage Oparin/Dyson model (protein only proto-life, followed by the acquisition of RNA/DNA to code for peptide sequences), it must have been present at the proto-life stage, since the second stage requires agency and, therefore, must be the product purely of protein chemistry (Baluska et al., 2016).

8. Conclusions

Crick chose the word 'dogma' to name what is, in fact, a hypothesis: the true 'dogma' in biology is the concept of the gene. The gene does have the role that, according to Roll-Hansen, Johannsen proposed in 1910, namely, to represent "*an experimentally identifiable difference between genotypes*" but it has, over more than a century, acquired a much greater prominence than its true role deserves. Arguably, defending it as that central functional feature of biology has distorted the scientific method and rejected important empirical evidence.

At the outset, I said I would follow the example of Annila in his book "*Back to Reality*" (Annila 2020) and look for simple and intuitive explanations for how the complex system that is the cell, works. I would argue that the dogma of the gene and the dominance of a complicated, or machine-oriented, rather than a complex, model for biological systems, are among the impediments to appreciating the simplicity of self-organisation. In his book "*The Nonlinear Universe: Chaos, Emergence and Life*" (Scott 2007), Alwyn Scott reminds us that the natural world is replete with nonlinear phenomena, including self-organisation. I believe the explanations I have proposed for inheritance, evolution, development, and morphogenesis should be more intuitive than molecular genetics based on Johannsen's genotype-conception, and are if the focus on the gene does not obscure the vision of the real world.

And finally, for what reason did life originate some 3.5 billion years ago? To help to equilibrate the free energy disequilibria, according to Maupertuis' principle of least action, caused on planet Earth, by its sun shining in the cold of the universe?

Acknowledgements

I dedicate this piece to the late Robert (Bob) Cundall, mentor, collaborator, and friend. I owe a debt of gratitude to the late Alwyn C Scott, who introduced Bob and me, in the 1980s, to the soliton and the importance of nonlinearity. I am grateful to two anonymous reviewers for their valuable suggestions. I am also grateful to the members of the Critical Genetics Forum for stimulating discussions on practical aspects of behavioural genetics, and to John Berriman, Hans Edelmann, and Jill Sutcliffe for valuable discussions.

[76] After the lecture, Perutz questioned Popper, who did not hear the question clearly and the Chairman repeated it for him: "'*Dr Perutz wants to know why you think biochemistry cannot be reduced to chemistry'. 'Ah, yes', Popper finally replied benignly, 'that surprised me too, but I suggest you go away and think about it for an evening, and you will see that I am right*.'" This encounter led to a long-running dispute between Popper and Perutz, Perutz remaining unconvinced at the time of Popper's death in 1994.

[77] By any standard of comparison with what can be simulated in a laboratory test tube, the cellular cytoplasm is extraordinary. As well as housing organelles, such as the mitochondria and the nucleus, up to 40% of the 'aqueous volume' is comprised of dissolved macromolecules, the gene products. Yet despite this high concentration, the cytosol is translucent with a viscosity roughly equivalent to a 10% solution of glycerine. This can be determined by centrifuging a suitable injected pellet, which moves freely through the cytosol: Hillman, H. and P. Sartory (1980). The Living Cell: a rexamination of its fine structure. Chichester, Packard Publishing Ltd. In the test tube, the folding of peptides to proteins can only be achieved in very dilute solution otherwise the peptides form aggregates. Partially denatured proteins are also stably present in the cytoplasm and may play an important role in that state, as already noted. One solution being suggested to better understand how chemistry can be so specific in the cytoplasm is liquid-liquid phase separation Li, X. H., P. L. Chavali, R. Pancsa, S. Chavali and M. M. Babu (2018). "Function and Regulation of Phase-Separated Biological Condensates." Biochemistry **57**(17): 2452–2461. Here, membrane-less condensates ordering gene products are posited to play a role in regulation of the cell. What is clear is that the chemistry taking place in the cytoplasm will not be able to be replicated in the test tube. This, however, does not really matter because that chemistry is causing an emergent property, the cellular phenotype, and, as such, its causes cannot be deduced.

[78] See pp 126–127.

References

Alabert, C., Loos, C., Voelker-Albert, M., Graziano, S., Forne, I., Reveron-Gomez, N., Schuh, L., Hasenauer, J., Marr, C., Imhof, A., Groth, A., 2020. Domain model explains Propagation dynamics and stability of histone H3K27 and H3K36 methylation landscapes. Cell Rep. 30 (4), 1223–1234 e1228.

Anderson, P.W., 1972. More is different. Science 177 (4047), 393–396.

Anfinsen, C.B., Haber, E., Sela, M., White Jr., F.H., 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Natl. Acad. Sci. U. S. A. 47, 1309–1314.

Annila, A., 2020. Back to Reality. Privus Press, New York.

Annila, A., Baverstock, K., 2014. Genes without prominence: a reappraisal of the foundations of biology. J. R. Soc. Interface 11 (94), 20131017.

Annila, A., Baverstock, K., 2016. Discourse on order vs. disorder. Commun. Integr. Biol. 9 (4), e1187348.

Baluska, F., Levin, M., 2016. On having No head: cognition throughout biological systems. Front. Psychol. 7, 902.

Baluska, F., Mancuso, S., 2009. Deep evolutionary origins of neurobiology: turning the essence of 'neural' upside-down. Commun. Integr. Biol. 2 (1), 60–65.

Baluska, F., Yokawa, K., Mancuso, S., Baverstock, K., 2016. Understanding of anesthesia - why consciousness is essential for life and not based on genes. Commun. Integr. Biol. 9 (6), e1238118.

Barrick, J.E., Yu, D.S., Yoon, S.H., Jeong, H., Oh, T.K., Schneider, D., Lenski, R.E., Kim, J. F., 2009. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature 461 (7268), 1243–1247.

Basener, W.F., Sanford, J.C., 2018. The fundamental theorem of natural selection with mutations. J. Math. Biol. 76 (7), 1589–1622.

Baverstock, K., 2000. Radiation-induced genomic instability: a paradigm-breaking phenomenon and its relevance to environmentally induced cancer. Mutat. Res. 454 (1–2), 89–109.

Baverstock, K., 2011. A comparison of two cell regulatory models entailing high dimensional attractors representing phenotype. Prog. Biophys. Mol. Biol. 106 (2), 443–449.

Baverstock, K., 2019a. Crick's sequence hypothesis: a review. Commun. Integr. Biol. 12 (1), 59–64.

Baverstock, K., 2019b. Polygenic scores: are they a public health hazard? Prog. Biophys. Mol. Biol. 149, 4–8.

Baverstock, K., Karotki, A.V., 2011. Towards a unifying theory of late stochastic effects of ionizing radiation. Mutat. Res. 718 (1–2), 1–9.

Baverstock, K., Ronkko, M., 2014. The evolutionary origin of form and function. J. Physiol. 592 (11), 2261–2265.

Baverstock, K., Rönkkö, M., 2008. Epigenetic regulation of the mammalian cell. PLoS One 3 (6), e2290.

Bertalanffy, L.v., 1969. General System Theory; Foundations, Development, applications. New York, G. Braziller.

Black, D.L., 2000. Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. Cell 103 (3), 367–370.

Blanco-Gomez, A., Castillo-Lluva, S., Del Mar Saez-Freire, M., Hontecillas-Prieto, L., Mao, J.H., Castellanos-Martin, A., Perez-Losada, J., 2016. Missing heritability of complex diseases: enlightenment by genetic variants from intermediate phenotypes. Bioessays 38 (7), 664–673.

Blyth, E., 1835. An attempt to classify the 'varieties' of animals, with observations on the marked seasonal and other changes which naturally take place in various British species, and which do not constitute varieties. The Magazine of Natural History 8, 40–53.

Boltzmann, L., 1974. Theoretical Physics and Philosophical Problems : Selected Writings. Reidel Pub. Co, Dordrecht ; Boston.

Border, R., Johnson, E.C., Evans, L.M., Smolen, A., Berley, N., Sullivan, P.F., Keller, M.C., 2019. No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. Am. J. Psychiatr. 176 (5), 376–387.

Bray, D., 2009. Wetware : a Computer in Every Living Cell. Yale University Press, New Haven ; London.

Calvo, P., Gagliano, M., Souza, G.M., Trewavas, A., 2020. Plants are intelligent, here's how. Ann. Bot. 125 (1), 11–28.

Chaufan, C., Joseph, J., 2013. The 'missing heritability' of common disorders: should health researchers care? Int. J. Health Serv. 43, 281–303.

Crick, F., 1958. On protein synthesis. Symp. Soc. Exp. Biol. 12, 138–163.

Crick, F., 1970. Central dogma of molecular biology. Nature 227 (5258), 561–563.

Csermely, P., Kunsic, N., Mendik, P., Kerestely, M., Farago, T., Veres, D.V., Tompa, P., 2020. Learning of signaling networks: molecular mechanisms. Trends Biochem. Sci. 45 (4), 284–294.

Darwin, C., Kebler, L., 1859. On the Origin of Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for life. London, J. Murray.

Dawkins, R., 2009. The Greatest Show on Earth : the Evidence for Evolution. Free Press, New York.

De la Fuente, I.M., Cortes, J.M., Pelta, D.A., Veguillas, J., 2013. Attractor metabolic networks. PLoS One 8 (3), e58284.

De Maupertuis, P.-l. M., 1746. Les Loix du mouvement et du repos deduites d'un principe metaphysique. Histoire de l'Acad Roy Sci Belles Lett, pp. 267–294.

Deglincerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., Brivanlou, A.H., 2016. Self-organization of the in vitro attached human embryo. Nature 533 (7602), 251–254.

Dobzhansky, T., 1964. Biology, molecular and organismic. Am. Zool. 4, 443–452.

Dyson, F.J., 1999. Origins of Life. Cambridge University Press, Cambridge [England] ; New York.

Eichler, E.E., Flint, J., Gibson, G., Kong, A., Leal, S.M., Moore, J.H., Nadeau, J.H., 2010. Missing heritability and strategies for finding the underlying causes of complex disease. Nat. Rev. Genet. 11 (6), 446–450.

Elston, R.C., 2018. Fisher's influence on me. Genet. Epidemiol. 42 (8), 849–853.

Fisher, R.A., 1918. The correlation between relatives on the supposition of Mendelian Inheritance. Trans. R. Soc. Edinb. 52, 399–433.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. The Clarendon press, Oxford.

Fonin, A.V., Darling, A.L., Kuznetsova, I.M., Turoverov, K.K., Uversky, V.N., 2018. Intrinsically disordered proteins in crowded milieu: when chaos prevails within the cellular gumbo. Cell. Mol. Life Sci. 75 (21), 3907–3929.

Gates, A.J., Gysi, D.M., Kellis, M., Barabasi, A.L., 2021. A wealth of discovery built on the Human Genome Project - by the numbers. Nature 590 (7845), 212–215.

Gell-Mann, M., 2001. Consciousness, reduction, and emergence. Some remarks. Ann. N. Y. Acad. Sci. 929, 41–49.

Gjuvsland, A.B., Wang, Y., Plahte, E., Omholt, S.W., 2013. Monotonicity is a key feature of genotype-phenotype maps. Front. Genet. 4, 216.

Good, B.H., McDonald, M.J., Barrick, J.E., Lenski, R.E., Desai, M.M., 2017. The dynamics of molecular evolution over 60,000 generations. Nature 551 (7678), 45–50.

Grafen, A., 2003. Fisher the evolutionary biologist. The Statistician 42 (3), 319–329.

Hall, B.G., 1982. Chromosomal mutation for citrate utilization by Escherichia coli K-12. J. Bacteriol. 151 (1), 269–273.

Huumonen, K., Immonen, H.K., Baverstock, K., Hiltunen, M., Korkalainen, M., Lahtinen, T., Parviainen, J., Viluksela, M., Wong, G., Naarala, J., Juutilainen, J., 2012. Radiation-induced genomic instability in Caenorhabditis elegans. Mutat. Res. 748 (1 - 2), 36–41.

Jablonka, E., 2017. The evolutionary implications of epigenetic inheritance. Interface Focus 7 (5), 20160135.

Jacob, F., Monod, J., 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. 3 (3), 318–356.

Jelenkovic, A., Hur, Y.M., Sund, R., Yokoyama, Y., Siribaddana, S.H., Hotopf, M., Sumathipala, A., Rijsdijk, F., Tan, Q., Zhang, D., Pang, Z., Aaltonen, S., Heikkila, K., Oncel, S.Y., Aliev, F., Rebato, E., Tarnoki, A.D., Tarnoki, D.L., Christensen, K., Skytthe, A., Kyvik, K.O., Silberg, J.L., Eaves, L.J., Maes, H.H., Cutler, T.L., Hopper, J. L., Ordonana, J.R., Sanchez-Romera, J.F., Colodro-Conde, L., Cozen, W., Hwang, A. E., Mack, T.M., Sung, J., Song, Y.M., Yang, S., Lee, K., Franz, C.E., Kremen, W.S., Lyons, M.J., Busjahn, A., Nelson, T.L., Whitfield, K.E., Kandler, C., Jang, K.L., Gatz, M., Butler, D.A., Stazi, M.A., Fagnani, C., D'Ippolito, C., Duncan, G.E., Buchwald, D., Derom, C.A., Vlietinck, R.F., Loos, R.J., Martin, N.G., Medland, S.E., Montgomery, G.W., Jeong, H.U., Swan, G.E., Krasnow, R., Magnusson, P.K., Pedersen, N.L., Dahl-Aslan, A.K., McAdams, T.A., Eley, T.C., Gregory, A.M., Tynelius, P., Baker, L.A., Tuvblad, C., Bayasgalan, G., Narandalai, D., Lichtenstein, P., Spector, T.D., Mangino, M., Lachance, G., Bartels, M., van Beijsterveldt, T.C., Willemsen, G., Burt, S.A., Klump, K.L., Harris, J.R., Brandt, I., Nilsen, T.S., Krueger, R.F., McGue, M., Pahlen, S., Corley, R.P., Hjelmborg, J.V., Goldberg, J.H., Iwatani, Y., Watanabe, M., Honda, C., Inui, F., Rasmussen, F., Huibregtse, B.M., Boomsma, D.I., Sorensen, T.I., Kaprio, J., Silventoinen, K., 2016. Genetic and Environmental Influences on Adult Human Height across Birth Cohorts from 1886 to 1994, vol. 5. Elife.

Johannsen, W., 1911. The genotype cconception of heredity. Am. Nat. 45, 129–159.

Johannsen, W., 2014. The genotype conception of heredity. 1911. Int. J. Epidemiol. 43 (4), 989–1000.

Joseph, J., 2015. The Trouble with Twin Studies : a Reassessment of Twin Research in the Social and Behavioral Sciences. Routledge, Taylor & Francis Group, New York, NY.

Kadhim, M.A., Macdonald, D.A., Goodhead, D.T., Lorimore, S.A., Marsden, S.J., Wright, E.G., 1992. Transmission of chromosomal instability after plutonium alpha-particle irradiation. Nature 355 (6362), 738–740.

Kadhim, M., Salomaa, S., Wright, E., Hildebrandt, G., Belyakov, O.V., Prise, K.M., Little, M.P., 2013. Non-targeted effects of ionising radiation–implications for low dose risk. Mutat. Res. 752 (2), 84–98.

Karotki, A.V., Baverstock, K., 2012. What mechanisms/processes underlie radiation-induced genomic instability? Cell. Mol. Life Sci. 69 (20), 3351–3360.

Kerminen, S., Havulinna, A.S., Hellenthal, G., Martin, A.R., Sarin, A.P., Perola, M., Palotie, A., Salomaa, V., Daly, M.J., Ripatti, S., Pirinen, M., 2017. Fine-scale genetic structure in Finland. G3 (Bethesda) 7 (10), 3459–3468.

Kerminen, S., Martin, A.R., Koskela, J., Ruotsalainen, S.E., Havulinna, A.S., Surakka, I., Palotie, A., Perola, M., Salomaa, V., Daly, M.J., Ripatti, S., Pirinen, M., 2019. Geographic variation and bias in the polygenic scores of complex diseases and traits in Finland. Am. J. Hum. Genet. 104 (6), 1169–1181.

Kerr, N.L., 1998. HARKing: hypothesizing after the results are known. Pers. Soc. Psychol. Rev. 2 (3), 196–217.

Krawetz, S.A., 2005. Paternal contribution: new insights and future challenges. Nat. Rev. Genet. 6 (8), 633–642.

Lenski, R.E., 2017. What is adaptation by natural selection? Perspectives of an experimental microbiologist. PLoS Genet. 13 (4), e1006668.

Lenski, R.E., Travisano, M., 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc. Natl. Acad. Sci. U. S. A. 91 (15), 6808–6814.

Lewontin, R.C., 1974. The Genetic Basis of Evolutionary Change. Columbia University Press, New York.

Louie, A.H., 2010. Robert Rosen's anticipatory systems. Foresight 12 (3), 18–29.

Lynch, M., 2010. Rate, molecular spectrum, and consequences of human mutation. Proc. Natl. Acad. Sci. U. S. A. 107 (3), 961–968.

Maddamsetti, R., Grant, N.A., 2020. Divergent evolution of mutation rates and biases in the long-term evolution experiment with Escherichia coli. Genome Biol. Evol. 12 (9), 1591–1603.
Makela, T., Annila, A., 2010. Natural patterns of energy dispersal. Phys. Life Rev. 7 (4), 477–498.
Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S., Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F., McCarroll, S.A., Visscher, P.M., 2009. Finding the missing heritability of complex diseases. Nature 461 (7265), 747–753.
Mayr, E., 2001. What Evolution Is. Basic Books, New York.
McClintock, B., 1984. The significance of responses of the genome to challenge. Science 226 (4676), 792–801.
Minton, A.P., 2006. How can biochemical reactions within cells differ from those in test tubes? J. Cell Sci. 119 (Pt 14), 2863–2869.
Morgan, T.H., 1917. The theory of the gene. Am. Nat. 51, 513–544.
Morgan, W.F., 2003a. Non-targeted and delayed effects of exposure to ionizing radiation: I. Radiation-induced genomic instability and bystander effects in vitro. Radiat. Res. 159 (5), 567–580.
Morgan, W.F., 2003b. Non-targeted and delayed effects of exposure to ionizing radiation: II. Radiation-induced genomic instability and bystander effects in vivo, clastogenic factors and transgenerational effects. Radiat. Res. 159 (5), 581–596.
Niemann, H.-J., 2014. Karl Popper and the Two New Secrets of Life : Including Karl Popper's Medawar Lecture 1986 and Three Related Texts. Tübingen, Mohr Siebeck.
Nijhout, H.F., 1990. Metaphors and the role of genes in development. Bioessays 12 (9), 441–446.
Noble, D., 2017. Dance to the Tune of Life : Biological Relativity. Cambridge University Press, Cambridge ; New York.
Nurse, P., Hayles, J., 2011. The cell in an era of systems biology. Cell 144 (6), 850–854.
Oliveira, H.M., Melo, L.V., 2015. Huygens synchronization of two clocks. Sci. Rep. 5, **11548** 3–11.
Omholt, S.W., 2013. From sequence to consequence and back. Prog. Biophys. Mol. Biol. 111 (2–3), 75–82.
Oparin, A.I., 1953. The Origin of Life. Dover Publications, New York.
Plomin, R., 2018. Blueprint : How DNA Makes Us Who We Are. The MIT Press, Cambridge, MA.
Plomin, R., von Stumm, S., 2022. Polygenic scores: prediction versus explanation. Mol. Psychiatr. 27 (1), 49–52.
Portin, P., Wilkins, A., 2017. The evolving definition of the term "gene. Genetics 205 (4), 1353–1364.
Pujol, B., Blanchet, S., Charmantier, A., Danchin, E., Facon, B., Marrot, P., Roux, F., Scotti, I., Teplitsky, C., Thomson, C.E., Winney, I., 2018. The missing response to selection in the wild. Trends Ecol. Evol. 33 (5), 337–346.
Reik, W., Dean, W., Walter, J., 2001. Epigenetic reprogramming in mammalian development. Science 293 (5532), 1089–1093.
Rice, T.K., Borecki, I.B., 2001. Familial resemblance and heritability. Adv. Genet. 42, 35–44.
Richardson, K., 2017. Genes, Brains, and Human Potential : the Science and Ideology of Intelligence. Columbia University Press, New York.
Richardson, K., Jones, M.C., 2019. Why genome-wide associations with cognitive ability measures are probably spurious. New Ideas Psychol. 55, 35–41.
Richardson, K., 2020. In the light of the environment: evolution through biogrammars not Programmers. Biol. Theor. 15, 212–222.
Roll-Hansen, N., 2014. Commentary: Wilhelm Johannsen and the problem of heredity at the turn of the 19th century. Int. J. Epidemiol. 43 (4), 1007–1013.
Ronkko, M., 2007. An artificial ecosystem: emergent dynamics and lifelike properties. Artif. Life 13 (2), 159–187.
Rose, S., 1988. Reflections on reductionism. Trends Biochem. Sci. 13 (5), 160–162.
Rosen, R., 1991. Life Itself: a Comprehensive Inquiry into the Nature, Origin and Fabrication of Life. Columbia University Press, New York.
Salek, M.M., Carrara, F., Fernandez, V., Guasto, J.S., Stocker, R., 2019a. Bacterial chemotaxis in a microfluidic T-maze reveals strong phenotypic heterogeneity in chemotactic sensitivity. Nat. Commun. 10 (1), 1877.
Salek, M.M., Carrara, F., Fernandez, V., Stocker, R., 2019b. Bacterial maze runners reveal hidden diversity in chemotactic performance. Microb. Cell 6 (8), 370–372.
Sanderson, E., Richardson, T.G., Hemani, G., Davey Smith, G., 2021. The use of negative control outcomes in Mendelian randomization to detect potential population stratification. Int. J. Epidemiol.
Schrödinger, E., 1944. What Is Life? Cambridge University Press, Cambridge.
Scott, A.C., 2007. The Nonlinear Universe: Chaos, Emergence and Life. Springer, Berlin.
Schweisguth, F., Corson, F., 2019. Self-organization in pattern formation. Dev. Cell 49 (5), 659–677.
Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.N.M., Campbell, A., Devito, L., Ilic, D., Khalaf, Y., Niakan, K.K., Fishel, S., Zernicka-Goetz, M., 2016. Self-organization of the human embryo in the absence of maternal tissues. Nat. Cell Biol. 18 (6), 700–708.
Shahbazi, M.N., Siggia, E.D., Zernicka-Goetz, M., 2019. Self-organization of stem cells into embryos: a window on early mammalian development. Science 364 (6444), 948–951.
Sharma, V., Annila, A., 2007. Natural process–natural selection. Biophys. Chem. 127 (1–2), 123–128.
Timoféeff-Ressovsky, N.W., Zimmer, K.G., Delbrück, M., 1935. Über die Natur der Genmutation und der Genstruktur, vol. 1. Nachrichten der Biologischen Gesellschaft für Wissenschaft, Göttingen, pp. 189–241.
Turing, A.M., 1990. The chemical basis of morphogenesis. 1953. Bull. Math. Biol. 52 (1–2), 153–197 discussion 119-152.
Van Hofwegen, D.J., Hovde, C.J., Minnich, S.A., 2016. Rapid evolution of citrate utilization by Escherichia coli by direct selection requires citT and dctA. J. Bacteriol. 198 (7), 1022–1034.
Visscher, P.M., Bruce Walsh, J., 2019. Commentary: Fisher 1918: the foundation of the genetics and analysis of complex traits. Int. J. Epidemiol. 48 (1), 10–12.
Visscher, P.M., Goddard, M.E., 2019. From R.A. Fisher's 1918 paper to GWAS a century later. Genetics 211 (4), 1125–1130.
Waddington, C.H., 1942. Canalisation of development and the inheritance of acquired characters. Nature 150, 563–565.
Waddington, C.H., 2008. Paradigm for an evolutionary process. Biol. Theor. 3 (3), 258–266.
Wiser, M.J., Ribeck, N., Lenski, R.E., 2013. Long-term dynamics of adaptation in asexual populations. Science 342 (6164), 1364–1367.
Yang, X., 1999. Concepts of synchronization in dynamical systems. Phys. Lett. 260 (5), 340–344.

SEVEN COMMENTATIES

Cellular and organismal agency - Not based on genes: A comment on Baverstock.
F. Baluska and A.S. Reber.

The genetic control paradigm in biology: What we say, and what we are entitled to mean.
Kenneth Z. McKenna Richard Gawne & H. Frederik Nijhout.

Genes and knowledge: Response to Baverstock, K. the gene an appraisal.
Ken Richardson.

Correspondence insights into the role of genes in cell functionality.
Comments on “The gene: An appraisal” by K. Baverstock
Ildefonso M. De la Fuente.

A special role for the genotype? Some comments on Keith Baverstock: “The gene: An appraisal”
Nils Roll Hansen.

From information to physics to biology.
Giuseppe Longo.

Commentary on Baverstock; the gene; an appraisal.
S.P.R. Rose

Progress in Biophysics and Molecular Biology 167 (2021) 161–162

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Cellular and organismal agency – Not based on genes: A comment on Baverstock

Keith Baverstock's stimulating paper criticises, besides other issues, the current gene-centric biology as missing an important aspect of biology (Baverstock 2021). We agree and suggest that follow-up research needs to focus on the sensory and electrophysiology of the excitable plasma membrane which constitutes, not only a physical "smart" barrier for the cell's interior, but also allows living cells to maintain their life processes via processes which generate and maintain ordered cellular structures. These act as analogous memory devices via their structural templates and, as argued elsewhere (Reber, 2019; Baluška and Reber 2019), support cellular sentience.

All life is based on cells (Lyons 2020). Unicellular organisms include prokaryotic archaea and bacteria as well as complex eukaryotic algae and protozoa. Ancient archaea and cyanobacteria 'invented' photosynthesis some 3,4 billion years ago (Fournier et al., 2021) and changed the Earth's climate conditions, creating the oxygen-rich atmosphere promoting the emergence and evolution of complex life forms based on genuine multicellularity (Crowe et al., 2013; Sánchez-Baracaldo and Cardona 2020). Diverse protists, embracing both protozoa and algae, have been enjoying unicellular life-style ever since the first eukaryotes evolved from the prokaryotic organisms (Richmond 1989; Sleigh 1991; Reynolds 2008; Butterfield 2015; Baluška and Lyons 2021) some 2-1,7 billion years ago (Bengtson et al., 2017; Porter 2020). Although unicellular, these organisms act as a kind of 'swimming neuron' (Brette 2021) showing signs of learning, memory, and other complex behaviours (Jennings 1906; Machemer 2001; Ginger et al., 2008; Kunita et al., 2016; Dexter et al., 2019; Trinh et al., 2019; Boussard et al., 2021; Gershman et al., 2021) – implying both cellular sentience and a primitive form of cognition (Reber 2019; Baluška and Reber 2019; Reber and Baluška 2021; Baluška et al., 2021).

Cells are the basic units of life. They generate and maintain the cellular order of their structures from energy-rich nutrients. Owing to their dynamic structure and organization, cells are highly sensitive and unstable, vulnerable to perturbations in their environment, which is in constant flux. Importantly, the cellular processes safeguarding survival of cells take place on time scales of seconds or even micro-seconds and are localized to the excitable plasma membranes which handle both energy fluxes and sensory information (Gatenby 2019; Gatenby and Frieden 2017; Frieden and Gatenby 2019, 2020). These fast and active processes, organized by the excitable membranes enclosing the cellular interior, generate cellular agency based on cell-specific sentience (for details, see Baluška and Miller, 2018; Reber 2019; Baluška and Reber 2019; Reber and Baluška 2021; Baluška et al., 2021; Miller et al., 2020a,b).

First cells evolved from hypothetical proto-cells. It can be speculated that these proto-cells were devoid of any DNA-based digital memory and relied solely on the structural memory of their limiting membranes. This analogous memory is still used by modern cells but it is complemented with the digital memory of RNA and DNA polymers. In order to survive, all organisms must take ultra-fast actions at the nanosecond time-order which precludes any direct involvements of genes based on the gene expression. To sustain their living condition, all cells must also take continuously cognitive actions, ones which minimize their disorder, free energy and surprise 'shock' events (Friston 2010; Bruineberg et al., 2018). Importantly, the long-term memory stored digitally within nucleotide sequences of DNA and RNA is not useful for the short-term cognitive decisions, the ones which constitute the cellular-organismal agency safeguarding their survival. Interestingly, this cellular-organismal agency is based on excitable plasma membranes that are targets of anaesthetics (Baluška et al. 2016, 2021) and assemble cellular consciousness and cognition (Reber 2019; Baluška and Reber 2019; Reber and Baluška 2021; Baluška et al., 2021). Cellular sentience is also closely associated with cellular circadian clocks which, in evolution, were initially assembled around the limiting and excitable plasma membranes and only later implemented genes in their control (Baluška and Reber 2021). Importantly, cellular circadian clocks are also closely integrated with molecules and processed underlying cellular sentience.

In contrast to a single living cell, a single gene is not capable of autonomous existence based on self-organization and self-replication. While some cells, such as red blood cells, can function without genes, genes without cells are not functional. Outside of living cells DNA and RNA molecules are inert and non-living macromolecular assemblies. Living cells use their genes as a digital source of biological information stored within the nucleotide sequences for their use in satisfying the cell's demands to construct specific protein-based macromolecules. But life is more than DNA-based software-like system (Baluška and Witzany 2015). In other words, genes serve cells as a kind of recipe for generation of ordered macromolecules based on specific proteins. DNA of chromosomal domains represent analogues of *books* and large sets of genomes act as kind of *libraries*, storing survival-relevant knowledge that cells accumulated during their evolution. Cellular membranes with the associated cytoskeleton represent the primary source of the cellular agency.

Declaration of competing interest

The authors declare that they have no known competing

https://doi.org/10.1016/j.pbiomolbio.2021.11.001
0079-6107/© 2021 Elsevier Ltd. All rights reserved.

financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Baluška, F., Witzany, G., 2015. Can subcellular organization be explained only by physical principles? Commun. Integr. Biol. 8, e1009796.

Baluška, F., Miller Jr., W.B., 2018. Senomic view of the cell: senome versus genome. Commun. Integr. Biol. 11, 1–9.

Baluška, F., Reber, A., 2019. Sentience and consciousness in single cells: how the first minds emerged in unicellular species. Bioessays 41, e1800229.

Baluška, F., Reber, A., 2021. CBC-clock theory of life. Integration of cellular circadian clocks and cellular sentience is essential for cognitive basis of life. Bioessays 43, e2100121.

Baluška, F., Lyons, S.L., 2021. Archaeal origins of eukaryotic cell and nucleus. Biosystems 203, 104375.

Baluška, F., Yokawa, K., Mancuso, S., Baverstock, K., 2016. Understanding of anesthesia - why consciousness is essential for life and not based on genes. Commun. Integr. Biol. 9, e1238118.

Baluška, F., Miller Jr., W.B., Reber, A.S., 2021. Biomolecular basis of cellular consciousness via subcellular nanobrains. Int. J. Mol. Sci. 22, 2545.

Baverstock, K., 2021. The gene: an appraisal. Prog. Biophys. Mol. Biol. 164, 46–62.

Bengtson, S., Sallstedt, T., Belivanova, V., Whitehouse, M., 2017. Three-dimensional preservation of cellular and subcellular structures suggests 1.6 billion-year-old crown-group red algae. PLoS Biol. 15, e2000735.

Boussard, A., Fessel, A., Oettmeier, C., Briard, L., Döbereiner, H.G., Dussutour, A., 2021. Adaptive behaviour and learning in slime moulds: the role of oscillations. Philos. Trans. R. Soc. Lond. B Biol. Sci. 376, 20190757.

Brette, R., 2021. Integrative neuroscience of Paramecium, a "swimming neuron". eNeuro 8 (2021), 0018-21.

Bruineberg, J., Rietveld, E., Parr, Th, van Maanen, L., Friston, K.J., 2018. Free-energy minimization in joint agent-environment systems: a niche construction perspective. J. Theor. Biol. 455, 161–178.

Crowe, S.A., Døssing, L.N., Beukes, N.J., Bau, M., Kruger, S.J., Frei, R., Canfield, D.E., 2013. Atmospheric oxygenation three billion years ago. Nature 501, 535–538.

Dexter, J.P., Prabakaran, S., Gunawardena, J., 2019. A complex hierarchy of avoidance behaviors in a single-cell eukaryote. Curr. Biol. 29, 4323–4329.

Fournier, G.P., Moore, K.R., Rangel, J.G., Payette, J.G., Momper, L., Bosal, T., 2021. The Archean origin of oxygenic photosynthesis and extant cyanobacterial lineages. Proc. Biol. Sci. 288, 20210675.

Frieden, B.R., Gatenby, R.A., 2019. Signal transmission through elements of the cytoskeleton form an optimized information network in eukaryotic cells. Sci. Rep. 9, 6110.

Frieden, B.R., Gatenby, R.A., 2020. Ion-based cellular signal transmission, principles of minimum information loss, and evolution by natural selection. Int. J. Mol. Sci. 21, 9.

Friston, K.J., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138.

Gatenby, R.A., 2019. The role of cell membrane information reception, processing, and communication in the structure and function of multicellular tissue. Int. J. Mol. Sci. 20, 3609.

Gatenby, R.A., Frieden, B.R., 2017. Cellular information dynamics through transmembrane flow of ions. Sci. Rep. 7, 15075.

Gershman, S.J., Balbi, P.E., Gallistel, C.R., Gunawardena, J., 2021. Reconsidering the evidence for learning in single cells. eLife 10, e61907.

Ginger, M.L., Portman, N., McKean, P.G., 2008. Swimming with protists: perception, motility and flagellum assembly. Nat. Rev. Microbiol. 6, 838–850.

Jennings, H.S., 1906. Behavior of the Lower Organisms. Columbia University Press, New York.

Kunita, I., Yamaguchi, T., Tero, A., Akiyama, M., Kuroda, S., Nakagaki, T., 2016. A ciliate memorizes the geometry of a swimming arena. J. R. Interface 13, 20160155.

Lyons, S.L., 2020. From Cells to Organisms: Re-envisioning Cell Theory. University of Toronto Press.

Machemer, H., 2001. The swimming cell and its world: structures and mechanisms of orientation in protists. Eur. J. Protistol. 37, 3–14.

Miller Jr., W.B., Torday, J.S., Baluška, F., 2020a. The N-space Episenome unifies cellular information space-time within cognition-based evolution. Prog. Biophys. Mol. Biol. 150, 112–139.

Miller Jr., W.B., Baluška, F., Torday, J.S., 2020b. Cellular senomic measurements in cognition-based evolution. Prog. Biophys. Mol. Biol. 156, 20–33.

Porter, S.M., 2020. Insights into eukaryogenesis from the fossil record. Inter. Focus 10, 20190105.

Reber, A.S., 2019. The First Minds: Caterpillars, Karyotes, and Consciousness. Oxford University Press, NY.

Reber, A.S., Baluška, F., 2021. Cognition in some surprising places. Biochem. Biophys. Res. Commun. 564, 150–157.

Reynolds, A., 2008. Amoebae as exemplary cells: the protean nature of an elementary organism. J. Hist. Biol. 41, 307–337.

Richmond, M.L., 1989. Protozoa as precursors of metazoa: German Cell Theory and its critics at the turn of the century. J. Hist. Biol. 22, 243–276.

Sánchez-Baracaldo, P., Cardona, T., 2020. On the origin of oxygenic photosynthesis and cyanobacteria. New Phytol. 225, 1440–1446.

Sleigh, M.A., 1991. Protozoa and Other Protists. Cambridge University Press.

Trinh, M.K., Wayland, M.T., Prabakaran, S., 2019. Behavioural analysis of single-cell aneural ciliate, *Stentor roeseli*, using machine learning approaches. J. R. Interface 13, 20190410.

František Baluška*
Institute of Cellular and Molecular Botany, University of Bonn, Germany

Arthur S. Reber
Department of Psychology, University of British Columbia, Vancouver, Canada
E-mail address: areber@brooklyn.cuny.edu.

* Corresponding author.
E-mail address: baluska@uni-bonn.de (F. Baluška).

24 August 2021
Available online 4 November 2021

Progress in Biophysics and Molecular Biology 169-170 (2022) 89–93

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

The genetic control paradigm in biology: What we say, and what we are entitled to mean

ABSTRACT

We comment on the article by Keith Baverstock (2021) and provide critiques of the concepts of genetic control, genetic blueprint and genetic program.

© 2022 Elsevier Ltd. All rights reserved.

> "If we say that the abnormal constituent *caused* the observed abnormality all we are entitled to mean is that it was one element in a complex which is essential to the result observed."
>
> –(Woodger, 1930)

1. Introduction

Genes play a major part in the developmental process, but just how important are they? To most biologists, this is tantamount to questioning the need for a pilot on a trans-Atlantic flight. Genetic factors are widely regarded as the causally primitive 'prime movers' of development, and this way of thinking has dominated the field for nearly a century (Keller, 2000). However, dissenting voices have been present from the outset (Russell, 1930; Waddington, 1957; Woodger, 1929), and the recent analysis by Baverstock echoes many legitimate concerns raised by successive generations of biologists. The aim of Baverstock's article is to provide an alternative understanding of development that is founded upon the idea that the genetic contributions to this process have been overstudied and overstated. Claims such as these are bound to raise hackles, but to his credit, the author pushes in his chips without hesitation. Indeed, he goes so far as to suggest that "defending [the gene] as that central functional feature of biology has distorted the scientific method and rejected important empirical evidence" (Baverstock, 2021). Some will find ideas such as these refreshing, while others will be inclined to reject them out of hand. In either case, it is worthwhile to revisit our understanding of the role genes play in development and reflect on why we hold these beliefs.

The support Baverstock offers in favor of his assertions is multifaceted, and includes biological, historical, philosophical and physical considerations of genes and their relationship to phenotypes. He makes the case that genes are necessary but *passive* players in the production of phenotypes. This is a point well taken, and we would like to single out three attributes of genes and genomes that he discusses for further elaboration. These are the claims that genes *control* the properties of living things, that genes and genomes serve as *blueprints* for cells, tissues and organisms, and that they are, or contain, *programs* for development that direct biological processes to specific ends. These three ideas permeate how we speak, how we teach Biology, and how we convey our ideas and the results of our research to the public. All three of them are wrong, as we will explain.

2. The primacy of phenotypes

Baverstock argues for the primacy of phenotypes, not genes, in evolution and development. This way of thinking resonates well with the 'phenotype first' perspective on the origin of novel characters in evolution (Gawne et al., 2018; Levis and Pfennig, 2016; Nijhout et al., 2021; Suzuki et al., 2020; West-Eberhard, 2003, 2005). In contrast to the classic Fisherian view, which suggests that new traits originate by rare mutations of large effect that are gradually refined by the accumulation of mutations of smaller effect, the phenotypes first view emphasizes the importance of phenotypic plasticity and sensitivity to environmental variables in the origin of novelties. Some phenotypic plasticity is neutral, some is deleterious, and some is advantageous (Suzuki et al., 2020). If an environment induces an advantageous phenotypic variant, and if that environment recurs, there will be selection to stabilize and refine that new phenotype in that new environment. The adaptation is genetic but does not rely on new mutations. Rather, it depends on existing cryptic genetic variation (Gibson and Dworkin, 2004; Gibson and Reed, 2008; McGuigan and Sgrò, 2009; Moczek, 2007, 2008; Nijhout et al., 2017; Paaby and Rockman, 2014; Schlichting, 2008; Schneider and Meyer, 2017; Suzuki and Nijhout, 2008).

Cryptic genetic variation arises from the many developmental and physiological mechanisms that have evolved to stabilize phenotypes against mutations that would otherwise reduce fitness (Nijhout et al., 2018, 2021; Paaby and Rockman, 2014; Suzuki et al., 2020). If phenotypes are resistant to mutations this will naturally lead to the accumulation in a population of mutations that

https://doi.org/10.1016/j.pbiomolbio.2022.02.003
0079-6107/© 2022 Elsevier Ltd. All rights reserved.

have little or no effect on the phenotypes. Some of those mutations can have a very large effect at the molecular level, but that effect is cancelled out or buffered by evolved homeostatic and robustness mechanisms in biochemistry, development and physiology. Cryptic genetic variation is most easily detected, documented and quantified in human diseases where genes that are characterized as risk factors for a disease by genetic epidemiologists have been well studied (Nijhout et al., 2015, 2018). Cryptic genetic variation will not be 'seen' by selection until a mutation or an environmental signal disrupts one of the stabilizing mechanisms (Gibson, 2009; Nijhout et al., 2015; Rutherford and Lindquist, 1998; Waddington, 1942, 1956). When this occurs, some of these cryptic mutations are no longer buffered, allowing them to become expressed as phenotypic variation on which selection can act. The crucial point here is that the mutation (or environmental stressor) that disrupts the robustness mechanism is what facilitates the evolution of the novelty. This mutation is part of a larger regulatory network that affects the development of the trait in question, but it does not necessarily lie directly within the pathways that result in its growth or patterning.

In development, phenotypes are also primary, meaning they must develop from preexisting phenotypes. Gene expression in development is controlled by transcriptional regulators that are spatially and temporally patterned in the preexisting developmental fields. It is well known that the effect of a gene product in development depends, not on the gene, but on the developmental context in which the gene is expressed. The same gene expressed in different phenotypic contexts can lead to dramatically different developmental results (Brunetti et al., 2001; Carroll et al., 1994; Wittkopp et al., 2003). This is one of the great findings in evo-devo: there is a relatively limited repertoire of transcription factors (gene products) that is used and reused in embryology, prenatal, and postnatal development to pattern different tissues, with very different tissue-specific and species-specific outcomes associated with the expression of the same genes. Once again, the implication is that genes themselves are causally inert, and only acquire functional significance when they are embedded in the appropriate phenotypic context.

3. Are genes in control?

Genes code for the sequence of nucleotides in RNA. That's it. Everything else about an organism plays out at higher levels of organization, where RNAs make essential but circumscribed contributions, mostly through the production of proteins. This basic fact has been known for a very long time and is codified as the Central Dogma of molecular biology. Yet, research published in technical journals regularly ascribes special properties to genes and genomes that greatly exceed their actual mandate of coding for proteins. Among other things, they are said regularly said to 'control' various biological parts and processes, ranging from other genes to complex morphologies. Similarly, genes and genomes are also said to contain 'programs' and 'blueprints' for cells, tissues, organs, behaviors, and even entire organisms. These claims have been commonplace for decades (Bang and Posakony, 1992; Boll and Noll, 2002; de Navas et al., 2006; Marand and Schmitz, 2022; Peter and Davidson, 2011; Pijuan-Sala et al., 2020; Srivastava and Olson, 2000), yet authors who appeal to genetic control, programs, and blueprints seldom—if ever—define what exactly they mean by these terms. It has long been recognized that these terms are actually metaphors that, perhaps, need no definition because they describe processes that need no definition because they are commonly used in day-to-day life (Gawne et al., 2018; Nijhout, 1990). The biological processes they attempt to describe are extremely complicated, but the use of metaphors allows us to talk about them efficiently, understanding, among ourselves, that they stand for processes that we actually do not know but would like to understand. Unfortunately, to the outside world these metaphors suggest a level of knowledge and understanding that really isn't there.

We use metaphors because doing biology is difficult. Identifying the causal interactions that take place within developing organisms is seldom straightforward, and as soon as we do manage to make progress, the goal posts shift. When we first begin studying an organism, simply documenting the existence of its parts is an important achievement. This temporarily satisfies our desire to understand the system, but having a catalog of 'key players' is rarely enough. The question of what these components actually do quickly becomes a haunting obsession. However, there might be tens, hundreds, or even thousands of parts in whatever system we happen to be examining. Grasping the basic functions of a gene product, cell type, or organ often requires years of effort, and it can take a lifetime or more to unravel the causal dynamics that occur within a more complex signaling cascade.

Confronted with the enormity of our task, it is natural to start looking for simplifications and time-saving work-arounds. Maybe all the system's parts are not of equal importance. In that case, some components might not require close examination. Or better yet, what if some parts of developing organisms are in 'control' of others? This would make the task of unraveling the cascade of ontogenetic causation much more manageable. Rather than trying to describe each component's functional role in the developmental process, we could focus our attention on the 'master regulators' that govern the behavior of all downstream parts and processes. If controllers of this type can be identified, knowing how they work could be all that is needed to obtain a complete understanding of character formation.

This approach is not new. Biologists have been searching for master regulators for centuries. Immaterial entities such as Aristotle's 'souls' (Solmsen, 1955) and subsequent forms of vitalism (Driesch, 1908) have given way to fully materialist paradigms where lower-level molecular processes are deemed to be causally primitive. In particular, the idea that genes control development is now widely accepted, and often simply taken for granted. This understanding of what genes do has regularly been called into question, but research practices have seldom changed as a result. One reason why criticisms of the genetic control paradigm have been ineffective is that efforts to dethrone genes from their seat of causal primacy have often been accompanied by calls to install some other purported controller in their place (Waggoner and Uller, 2015). This simply trades one king for another. An alternative way to approach the problem is to reject the notions of control and master regulation entirely. Although genes play an essential role in development, it remains unclear whether they should be regarded as *the* fundamental components that all others are dependent upon. There is no development without gene activity, but the same could be said about metabolism, respiration, and countless other biological phenomena. All processes that occur during morphogenesis are important, but it is not clear whether it makes sense to rank their significance by singling out one or more of them as the primitive, or most important, controller(s) of the process.

In his paper, Baverstock uses the metaphors of 'tools' and 'brains' to clarify the genotype-phenotype relationship. Here, we will use a simpler reductio analogy involving automobiles for this purpose. Regardless of whether organisms actually *are* or *are not* machines (Bongard and Levin, 2021; Nicholson, 2013, 2018, 2019), they are similar in that both are composed of parts whose activities jointly contribute to certain types of functional output. This implies that we could, in principle, attempt to understand how a mechanical device such as a car works by using the same methods we employ

when studying developing organisms in the laboratory. It is interesting to imagine how such an exercise might unfold. For most people, the logical first step would be to open the vehicle's hood. Once it is propped-up, the engine comes into view. Everyone knows that a car won't run without this component, and as such some will be tempted to declare that this must be the 'master controller' of the system. However, closer inspection reveals that the engine is composed of smaller parts such as pistons, rods, bearings, a crankshaft, and other elements that jointly contribute to its functionality. This means that the conglomerate we refer to singularly as the vehicle's engine is actually a complex mechanical network. Perhaps there is a specific 'node' within this system of parts that is the controller?

If we study a modern engine in detail, we see that it contains numerous electrical components, such as a knock sensor, camshaft position sensor and manifold pressure sensor. We can continue our search for a master regulator with these parts. The sensors in question play a crucial role in keeping the engine and the car itself functional – this can easily be demonstrated by unplugging them in a series of 'knock-out' style experiments. However, like other electrical devices, the sensors–or the systems that receive their output signal–only work if they have a power source. In automobiles, the battery plays this role. Here we have another absolutely crucial component of the system. If we remove one of the battery leads, the electrical sensors stop working, and more importantly, the engine itself soon stalls out. This seems to suggest that the engine is dependent upon the battery, which might cause some to conclude that the latter is the master regulator of the former, and perhaps the locus of control for the entire vehicle.

Yet, with additional scrutiny we see that the battery is charged by the alternator, which is itself powered by the engine. This adds yet another layer of complexity to our search for the vehicle's master regulator. The battery is rapidly depleted if it is not charged, causing the engine to stop. But, at the same time, the engine is dependent upon a functional electrical system. Does this mean that it is the combined interacting network of engine plus electrical system that controls the vehicle? Or, is there some other component of the system that is the chief regulator?

Questions such as these can be posed indefinitely, and there is a reason for this. Looking for a primitive causal controller in an automobile is a fool's errand. Cars are mechanical systems made-up of mutually dependent parts. Various components might be more or less important, but none are truly in control of the vehicle's overall functionality. Something similar can be said of organisms. Their genes, or more properly, their gene products, play a role in many important processes, but they are not in control of anything. Genes can only function in the appropriate cellular environments, and the cells themselves are powerless outside of the setting provided by the body. Again, this highlights the importance of *phenotypic context* in development.

Phenotypes exist at all levels of organization within an organism, and create local environments that coax highly specialized behaviors from generic morphogenetic processes. Phenomena that occur at one level can exert an influence on those that occur at other levels, and many are highly interdependent. Singling one level, or one component of one level as the master regulator of development is a conceptual error that stems from a tendency to prefer isolated causes over more involved systems-style explanations. Parsimony is a scientific virtue, and we are right to seek it. However, simple hypotheses are only desirable when they provide an accurate representation of biological reality, and the genetic control paradigm does not do so.

4. Are genomes blueprints or programs?

'Blueprint' is another metaphor that permeates the biological literature, and it goes hand-in-hand with the idea of genetic control. The use of this device is motivated by the assumption that genes specify what body parts are produced during development, and where they will be placed relative to one another. If we take this idea literally, the implication is that genes are responsible for creating body plans, understood here as recurrent morphological 'types' exhibiting a specific set of characters that are arranged in a highly stereotyped manner (Hall, 1992; Woodger, 1945).

The establishment of fixed body plans is one of the most fascinating outcomes of multicellularity. Somehow, the inheritance of genetic material leads to the inheritance of a taxon-specific body plan that can be distinguished from other morphological types through simple visual observation. For over a century, biologists have been interested in understanding what body plans exist, how they were brought about evolutionarily, and what proximate factors have contributed to their persistence over time (Hall, 1996; He and Deem, 2010; Raff, 2008). However, these issues have only recently been approached from a purely genetic perspective, and it remains unclear whether this approach should be regarded as an advancement, or an inadvertent step in the wrong direction. We believe that a strong case can be made for the latter possibility. Although it continues to be widely employed, the genomes-as-blueprints metaphor and its associated research practices unintentionally distort the facts by force-fitting them into a biologically unrealistic causal framework. More specifically, the blueprint analogy exaggerates the significance of these documents in non-developmental contexts, and then attributes similarly overstated capabilities to genes.

Returning to the automotive theme discussed above, each time we distinguish between a car and a motorcycle, we are effectively confirming the existence of two distinct body plans. The number of wheels, operator position, spatial orientation of the steering systems and numerous other mechanical parts differ dramatically between the vehicles. Crucially, however, if we wanted to better understand why the form of the two machines differ in these ways, there is little use in examining the relevant blueprints. These documents would likely help us arrive at a more complete understanding of what parts compose the systems, and could help us gain new knowledge of details like bolt lengths, torque specs, and thread pitches. Yet, it is a mistake to think they would provide us with an understanding of how the vehicles were actually constructed, or why they exhibit their representative 2- or 4-wheel form.

Blueprints leave countless questions about assembly unanswered, and the fact that this information is missing is hardly trivial. In manufacturing contexts, they provide some insight into the parts needed, and where they go, but they are unequivocally *not* a protocol for the construction of an object. As a consequence, we are unable to look at these documents after the fact and reverse engineer every step involved in the assembly process. The manufacturing of a vehicle is a highly coordinated affair where a single error can have cascading effects that render the final product non-functional. Every task needs to be completed in the correct temporal sequence, and every piece needs to be put in the right spatial location, or else production fails.

The genome-as-blueprint, genome-as-program, and genetic control metaphors suggest that development and morphogenesis are forward-looking processes, where every action at every organizational level is carried out in order to reach a predetermined, targeted, phenotypic goal. However, development is much more retrospective than forward looking. The behavior of cells at any given stage of ontogeny depends not only on the cellular environment in which they are presently immersed, but on how they themselves developed. If we insist on having a metaphor for explaining what genes and genomes do during morphogenesis,

they are best viewed as a haphazardly organized drawer cabinet of parts and tools, rather than a blueprint. Some of these parts and tools are useful and even required in specific contexts, such as the cell cycle, the deposition of bone, or the initiation of pigment synthesis in a specific part of the body, while others such as orphan nuclear receptors appear to have no function at all. Like a toolbox stocked with mechanical devices, genomes are passive objects that are best conceived as 'repositories of genes' whose developmental potentials are entirely dependent upon external actors and context.

The paradigm of Cis-regulation that underwrites the concept of gene regulatory networks (GRNs) has, in fact, definitively shown that the role genes play in development is context sensitive (Davidson, 2009; Farley et al., 2015; Jindal and Farley, 2021). In GRNs 'cis-regulatory elements' (CREs) in the vicinity of specific genes are activated by transcription factors that then stimulate gene expression (Davidson, 2001). The interactions of CREs and associated transcription factors are capable of facilitating: (1) the expression of gene products that are specific to a given time of development and place in the body, and (2) the differential expression of gene products involved in the formation or maintenance of tissue types and functions (Rickels and Shilatifard, 2018).

It is important to recognize, however, that CREs are not in control of developmental patterning, but are merely passive agents that sometimes become accessible through context dependent molecular and cellular interactions. A well-known example can be found in the *Drosophila* embryo, where segments along the anterior-posterior axis are demarcated by bands of cells secreting WNT (Baker, 1988). Between and perpendicular to WNT secreting populations of cells in the imaginal disc lie a dorsal population of cells that secretes BMP, and a ventral population that secretes EGF (Kubota et al., 2000, 2003). Each of these secreted proteins forms a concentration gradient: the WNT gradient forms along the anterior-posterior axis, while the BMP and EGF gradients form along the dorsoventral axis. The BMP gradient declines ventrally, while EGF declines dorsally. This leaves a population of cells along the midline of the tissue with low levels of BMP and EGF proteins, and high levels of WNT. This diffusion gradient system is of fundamental importance because it emerges from the interaction of several *populations of cells*. Both BMP and EGF inhibit the expression of Distal-less (Dll), a gene whose transcription is required for the induction of limb outgrowth (Kubota et al., 2000). However, WNT proteins positively regulate Dll at high concentrations (Cohen et al., 1993). Thus, limb primordial cells are established as a population of Dll expressing cells adjacent to the WNT secreting cells, and are spatially separated from both the BMP and EGF positive cells. In this system, it is the interaction among cells, rather than CREs that moves embryonic development forward.

5. A final word about control

On reading the preceding paragraph, a developmental geneticist would probably say that ultimately it is really the genes that code for WNT, BMP, EGF and Dll that are truly in control, because without them, and their association through the GRN, none of the cellular interactions would work. Fair enough. But what exactly do we mean by control? Although they are never precise about this, the conceptual definitions of control that biologists use range from 'just turning something on or off' to 'processes that guide development to a desired or predetermined end'. Regardless which of the many implied definitions we choose to employ, it is essential to recognize that control cannot be exercised by a thing that is just one of many necessary components of a system. No biological system works without water, or amino acids, but no one would claim that these control anything. In a formal sense, one can think of control as a part or process that makes the crucial difference in achieving an observed endpoint. Although the identification of such a controller might seem straightforward in principle, it is seldom simple in practice. Does the light switch control the lights, or does the person that flips the switch? Why do we privilege genes as the controllers? Maybe the thing that turns a gene on is the controller. That would be a transcriptional complex. But what turned that on? Not a gene, but another transcriptional complex or complex signaling pathway. In reality it is an infinite regression of causation that includes not only gene products, but also things like hormones, environmental factors like nutrients, and enzyme cofactors like vitamins. The causal factors we choose to prioritize in that sequence is a matter of taste, convenience or practicality, not of principle.

But finding a locus of control in development is actually more fraught than that. Consider the most common processes in developmental patterning, diffusion gradient-threshold mechanisms and reaction-diffusion mechanisms. These are dynamic spatial information systems that ensure certain things happen at the 'right' location and at the 'right' time. They are probably the best examples of self-organization in development. Diffusion gradient-threshold mechanisms are the most widespread, beginning with maternal gradients of nucleotides and proteins in oocytes. New proteins are induced at particular thresholds of the gradient that, in turn, diffuse and set up new gradients that activate subsequent diffusible proteins, and so on. Does the maternal gradient therefore control all of subsequent embryonic development? Or does control reside in the processes that put the maternal gradient there in the first place? Reaction-diffusion systems depend on positive feedback (autocatalysis), lateral inhibition, differences in the rates of synthesis and breakdown, and differences in the diffusion coefficients of the reactants. As in the case of diffusion gradient-threshold systems, the physical properties of the system: the diffusion coefficients of the reactants and the sizes and shapes of the fields in which the reactions occur are critical in determining the outcomes, as is the exact amount of time the reactions are allowed to proceed. It would clearly be a mistake to isolate one component of a reaction-diffusion system and declare it to be 'the controller', but it can be more challenging to see that the same can be said of post-embryonic events, where genetic factors have historically been afforded special causal status.

It is undeniable that genes provide indispensable building materials for living things. Likewise, there is overwhelming evidence that genetic mutations often cause observable and specific differences in phenotypes. But, following Woodger's caution in the epigraph of this paper, all we are entitled to conclude from this is that the mutated gene was one element in a complex that was essential to the result observed. Nothing more. Other defects, such as an amino acid deficiency or a vitamin deficiency can have equally profound and specific effects on phenotypes. We reiterate, once more, that genes and genomes are static entities that are acted upon by the needs of a cell, tissue or organ. The challenge for Biology is not to understand how genes and gene regulatory networks operate in isolation, but why they are activated in specific patterns, and what the downstream consequences of those activations are on the form and function of phenotypes.

Declaration of competing interest

None.

Acknowledgements

We would like to thank Denis Noble for inviting us to contribute

this commentary. An anonymous reviewer provided encouraging, and thoughtful comments on an earlier draft of the manuscript. This work was supported by grant IOS-1557341 from the National Science Foundation. Contributions by K.Z.M., R.G., and H.F.N. are equal.

References

Baker, N.E., 1988. Localization of transcripts from the wingless gene in whole *Drosophila* embryos. Development 103, 289–298.
Bang, A.G., Posakony, J.W., 1992. The *Drosophila* gene Hairless encodes a novel basic protein that controls alternative cell fates in adult sensory organ development. Genes Dev. 6, 1752–1769.
Baverstock, K., 2021. The gene: an appraisal. Prog. Biophys. Mol. Biol. 164, 46–62.
Boll, W., Noll, M., 2002. The *Drosophila* Pox neuro gene: control of male courtship behavior and fertility as revealed by a complete dissection of all enhancers. Development 129, 5667–5681.
Bongard, J., Levin, M., 2021. Living things are not (20th Century) machines: updating mechanism metaphors in light of the modern science of machine behavior. Front. Ecol. Evol. 9.
Brunetti, C.R., Selegue, J.E., Monteiro, A., French, V., Brakefield, P.M., Carroll, S.B., 2001. The generation and diversification of butterfly eyespot color patterns. Curr. Biol. 11, 1578–1585.
Carroll, S., Gates, J., Keys, D., Paddock, S., Panganiban, G., Selegue, J., Williams, J., 1994. Pattern formation and eyespot determination in butterfly wings. Science 265, 109–114.
Cohen, B., Simcox, A.A., Cohen, S.M., 1993. Allocation of the thoracic imaginal primordia in the *Drosophila* embryo. Development 117, 597–608.
Davidson, E.H., 2001. Genomic Regulatory Systems: Development and Evolution. Elsevier, San Diego.
Davidson, E.H., 2009. Network design principles from the sea urchin embryo. Curr. Opin. Genet. Dev. 19, 535–540.
de Navas, L.F., Garaulet, D.L., Sánchez-Herrero, E., 2006. The Ultrabithorax Hox gene of *Drosophila* controls haltere size by regulating the Dpp pathway. Development 133, 4495–4506.
Driesch, H., 1908. The Science and Philosophy of the Organism. A & C Black, London.
Farley, E.K., Olson, K.M., Zhang, W., Brandt, A.J., Rokhsar, D.S., Levine, M.S., 2015. Suboptimization of developmental enhancers. Science 350, 325–328.
Gawne, R., McKenna, K.Z., Nijhout, H.F., 2018. Unmodern synthesis: developmental hierarchies and the origin of phenotypes. Bioessays 40, 1600265.
Gibson, G., 2009. Decanalization and the origin of complex disease. Nat. Rev. Genet. 10, 134–140.
Gibson, G., Dworkin, I., 2004. Uncovering cryptic genetic variation. Nat. Rev. Genet. 5, 681–690.
Gibson, G., Reed, L.K., 2008. Cryptic genetic variation. Curr. Biol. 18, R989–R990.
Hall, B.K., 1992. Baupläne: fundamental body plans. In: Hall, B.K. (Ed.), Evolutionary Developmental Biology. Springer, Netherlands, pp. 66–82.
Hall, B.K., 1996. Baupläne, phylotypic stages, and constraint-Why there are so few types of animals. Evol. Biol. 29 (29), 215–261.
He, J., Deem, M.W., 2010. Hierarchical evolution of animal body plans. Dev. Biol. 337, 157–161.
Jindal, G.A., Farley, E.K., 2021. Enhancer grammar in development, evolution, and disease: dependencies and interplay. Dev. Cell 56, 575–587.
Keller, E.F., 2000. Decoding the Genetic Program. Cambridge University Press, Cambridge.
Kubota, K., Goto, S., Eto, K., Hayashi, S., 2000. EGF receptor attenuates Dpp signaling and helps to distinguish the wing and leg cell fates in *Drosophila*. Development 127, 3769–3776.
Kubota, K., Goto, S., Hayashi, S., 2003. The role of Wg signaling in the patterning of embryonic leg primordium in *Drosophila*. Dev. Biol. 257, 117–126.
Levis, N.A., Pfennig, D.W., 2016. Evaluating 'plasticity-first' evolution in nature: key criteria and empirical approaches. Trends Ecol. Evol. 31, 563–574.
Marand, A.P., Schmitz, R.J., 2022. Single-cell analysis of cis-regulatory elements. Curr. Opin. Plant Biol. 65, 102094.
McGuigan, K., Sgrò, C.M., 2009. Evolutionary consequences of cryptic genetic variation. Trends Ecol. Evol. 24, 305–311.
Moczek, A.P., 2007. Developmental capacitance, genetic accommodation, and adaptive evolution. Evol. Dev. 9, 299–305.
Moczek, A.P., 2008. On the origins of novelty in development and evolution. Bioessays 30, 432–447.
Nicholson, D.J., 2013. Organisms≠Machines, studies in history and philosophy of science Part C: studies in history and. Phil. Biol. Biomed. Sci. 44, 669–678.
Nicholson, D.J., 2018. Reconceptualizing the Organism: from Complex Machine to Flowing Stream, Everything Flows. Oxford University Press, Oxford.
Nicholson, D.J., 2019. Is the cell *really* a machine? J. Theor. Biol. 477, 108–126.
Nijhout, H.F., 1990. Problems and Paradigms: metaphors and the role of genes in development. Bioessays 12, 441–446.
Nijhout, H.F., Best, J.A., Reed, M.C., 2015. Using mathematical models to understand metabolism, genes, and disease. BMC Biol. 13, 79.
Nijhout, H.F., Best, J.A., Reed, M.C., 2018. Systems Biology of Robustness and Homeostatic Mechanisms. Wiley Interdisciplinary Reviews: Systems Biology and Medicine, e1440.
Nijhout, H.F., Kudla, A.M., Hazelwood, C.C., 2021. Genetic assimilation and accommodation: models and mechanisms. Curr. Top. Dev. Biol. 141, 337–369.
Nijhout, H.F., Sadre-Marandi, F., Best, J., Reed, M.C., 2017. Systems biology of phenotypic robustness and plasticity. Integr. Comp. Biol. 57, 171–184.
Paaby, A.B., Rockman, M.V., 2014. Cryptic genetic variation: evolution's hidden substrate. Nat. Rev. Genet. 15, 247–258.
Peter, I.S., Davidson, E.H., 2011. Evolution of gene regulatory networks controlling body plan development. Cell 144, 970–985.
Pijuan-Sala, B., Wilson, N.K., Xia, J., Hou, X., Hannah, R.L., Kinston, S., Calero-Nieto, F.J., Poirion, O., Preissl, S., Liu, F., Göttgens, B., 2020. Single-cell chromatin accessibility maps reveal regulatory programs driving early mouse organogenesis. Nat. Cell Biol. 22, 487–497.
Raff, R.A., 2008. Origins of the other metazoan body plans: the evolution of larval forms. Phil. Trans. Roy. Soc. Lond. B Biol. Sci. 363, 1473–1479.
Rickels, R., Shilatifard, A., 2018. Enhancer logic and mechanics in development and disease. Trends Cell Biol. 28, 608–630.
Russell, E.S., 1930. The Interpretation of Development and Heredity: A Study in Biological Method. Clarendon Press, Oxford.
Rutherford, S.L., Lindquist, S., 1998. Hsp90 as a capacitor for morphological evolution. Nature 396, 336.
Schlichting, C.D., 2008. Hidden reaction norms, cryptic genetic variation, and evolvability. Ann. N. Y. Acad. Sci. 1133, 187–203.
Schneider, R.F., Meyer, A., 2017. How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. Mol. Ecol. 26, 330–350.
Solmsen, F., 1955. Antecedents of Aristotle's psychology and scale of beings. Am. J. Philol. 76, 148.
Srivastava, D., Olson, E.N., 2000. A genetic blueprint for cardiac development. Nature 407, 221–226.
Suzuki, Y., McKenna, K.Z., Nijhout, H.F., 2020. Regulation of phenotypic plasticity from the perspective of evolutionary developmental biology. In: Levine, H., Jolly, M.K., Kulkarni, P., Nanjundiah, V. (Eds.), Phenotypic Switching. Academic Press, pp. 403–442.
Suzuki, Y., Nijhout, H.F., 2008. Genetic basis of adaptive evolution of a polyphenism by genetic accommodation. J. Evol. Biol. 21, 57–66.
Waddington, C.H., 1942. Canalization of development and the inheritance of acquired characters. Nature 150, 563–565.
Waddington, C.H., 1956. Genetic assimilation of the bithorax phenotype. Evolution 10, 1–13.
Waddington, C.H., 1957. The Strategy of the Genes : a Discussion of Some Aspects of Theoretical Biology. Allen & Unwin, London.
Waggoner, M.R., Uller, T., 2015. Epigenetic determinism in science and society. New Genet. Soc. 34, 177–195.
West-Eberhard, M.J., 2003. Developmental Plasticity and Evolution. Oxford University Press, New York.
West-Eberhard, M.J., 2005. Phenotypic accommodation: adaptive innovation due to developmental plasticity. J. Exp. Zool. B Mol. Dev. Evol. 304B, 610–618.
Wittkopp, P.J., Williams, B.L., Selegue, J.E., Carroll, S.B., 2003. *Drosophila* pigmentation evolution: divergent genotypes underlying convergent phenotypes. Proc. Natl. Acad. Sci. U.S.A. 100, 1808–1813.
Woodger, J.H., 1929. Biological Principles: A Critical Study. Routledge & Kegan London.
Woodger, J.H., 1930. The "Concept of Organism" and the relation between embryology and genetics. Part I. Q. Rev. Biol. 5, 1–22.
Woodger, J.H., 1945. On Biological Transformations. Clarendon Press, Oxford.

Kenneth Z. McKenna
Department of Biology, University of California, San Diego, United States

Richard Gawne
Allen Discovery Center at Tufts University, United States

H. Frederik Nijhout*
Department of Biology, Duke University, United States

* Corresponding author.
E-mail address: hfn@duke.edu (H.F. Nijhout).

14 September 2021
Available online 24 February 2022

Progress in Biophysics and Molecular Biology 167 (2021) 12–17

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Genes and knowledge: Response to Baverstock, K. the gene an appraisal. https://doi.org/10.1016/j.pbiomolbio.2021.04.005

Ken Richardson

26 Castle Terrace, Berwick upon Tweed TD15 1 NZ, UK

ARTICLE INFO

Article history:
Received 10 September 2021
Received in revised form
21 September 2021
Accepted 26 October 2021
Available online 2 November 2021

Keywords:
Knowledge
Learning
Evolution
Information
Gene

ABSTRACT

This response aims to expand on some of the issues raised by Keith Baverstock's The Gene: An Appraisal, especially on the evolution and nature of knowledge in living things. In contrast to the simple associationism envisaged in "genetic information", it emphasises the dynamic complexity and changeability of most natural environments, and, therefore, predictability based on underlying statistical structures. That seems to be the basis of the "cognitive" functions increasingly being reported about cellular, as well as more evolved, functions, and of the autonomous agency of organisms thriving creatively in complex environments.

© 2021 Elsevier Ltd. All rights reserved.

Contents

1. Introduction 12
2. Knowledge of what? 13
3. With what rules? 13
4. Origins 14
5. Development 14
6. Metabolism 14
7. For example? 15
Declaration of competing interest 16
References 16

1. Introduction

Keith Baverstock (KB) describes how much "dogma" in biology is rooted in the concept of the gene and how the need for a re-think is urgent. That view reflects many other revisionary ideas, some old some more recent. Yet, the dogma remains so embedded in the scientific literature, school and university textbooks, popular books and articles and wider culture (Noble, 2015) as to seem unassailable. At least part of the problem, is the need for a coherent and pursuasive alternative story. In this response I want to explore that question further: If it's not the genes that, both in evolution across generations, and development within them, explain the origins of form and adaptive variation, then what do we have to offer instead? KB radically asserts that we have the "phenotype" acting like a "brain" to acquire relevant knowledge (implictly by some kind of learning). So the important questions seem to be, knowledge of what? how is it derived and what form does it take - in the cell's complex functions, and in the forms and functions that evolved from it? Referring to development in particular KB says, "The ignorance here is fundamental". But I hope to show there are some grounds for optimism.

E-mail address: k.richardson@mac.com.

https://doi.org/10.1016/j.pbiomolbio.2021.10.003
0079-6107/© 2021 Elsevier Ltd. All rights reserved.

2. Knowledge of what?

"The environment", and what makes it predictable, may be an obvious answer. The MS view of the environment, based on natural selection of genes, comes with two severe constraints. The first, as explained by Darwin in chapter 4 of *The Origin*, is that the environment should be recurrent across generations, or change only very slowly (relative to generation time): that is predictability from constancy (Slobodkin and Rapoport 1974). Otherwise no consistent selection of favourable gene variants could occur. The second constraint, is that the environments doing the selecting should be quite discrete. This has meant describing environments as nominal collections of independent variables: temperature, oxygen availability, salinity, pH, presence or absence of broad categories of resources, predators, parasites, and so on (Somero et al. 2017): that is, a mosaic of distinct, recurring entities, offering durable niches to which, at the end of development, stable, well-defined phenotypes are adapted.

The only "bio-knowledge" needed under such constraints is that required to form and maintain, in development and across generations, one-to-one association with durable aspects of the environment. Hence the popular view of genes as fixed codes for phenotypes correlating with those aspects in lock and key fashion. Such is the basis of Dawkins' famous remark in *The Selfish Gene* about stable genes trumping ephemeral organisms; genes "shaped" by consequences to provide a recurrent response to recurrent conditions. That is what permitted animal behaviourist B.F. Skinner to draw an analogy between Darwinian natural selection and behavioural conditioning (Smith 2019).

It presents a shallow picture of biology and knowledge. Yet belief in such genes has formed foundations, not only of biology, but also cognitive theories of Noam Chomsky, Steven Pinker, computational cognitive scientists and sociobiology. It has dominated so-called "evolutionary psychology", which also views human cognitive and social behaviour as genetic "adaptations" to "recurrent statistical regularities" in the environment (Buss and von Hippel 2018). Consequently, the target of much research has been the discovery of the genetic codes for such adaptations: for example, by raising genetically different individuals in uniform environments, permitting their undistorted "expression" (Sultan, 2017); or more recently and more directly in genome wide association studies (GWAS).

Both ways of viewing the environment we now know to be quite wrong. While some aspects of environments do remain relatively constant across generations, others change significantly (Levins, 1968), and do so in interactive, not independent ways (see below). Both constraints entail an excessive simplification of the dynamics of the real world, as KB says, and impose serious limits on understanding. That shows in the contradictions, elisions and paradoxes so often encountered. For example, increased environmental complexity in evolution has been envisaged only in terms of increased numbers of variables, or "heterogeneity" (Godfrey-Smith 2001). As for increasing complexity in organisms, Darwin reasoned that a greater coordination of greater numbers of "parts" provides greater efficiency. However, merely gathering together parts in greater numbers does not explain where the "coordination" comes from. Attempting to explain complexity in terms of relatively simple environments is something that Darwin, himself, struggled with (Ridley 2000). Explaining complexity in development has entailed eliding the genes' restricted codes into "programs" for constructing complex forms - i.e. knowledge they cannot possibly have. So we get genomes described as "blueprints" (Plomin 2018); or even as "cookbooks" (Harden 2021).

Even those subscribing to the "extended evolutionary synthesis", "niche construction", and so on, have not escaped those constraints. As Laland et al. (2014) note, niche construction "requires an ability on the part of organisms to discriminate and actively sort between environmental resources": resources spoken of in nominal terms - generally, persistent "environmental states" such as oxygen levels, temperature, humidity, and so on. Examples of niches constructed include nests, burrows, mounds, ant hills, beaver's dams, and so on. Although fascinating, these are actually canalised behavioural developments to environments recurring across generations (and still often referred to as "instincts"). Their development demands molecular buffering (see below), but not the knowledge used by an embryonic stem cell acquiring an appropriate developmental trajectory, a protozoan learning a maze, corvids learning from environmental structure to use tools to access a novel food source, contingent association learning in plants (Reber and Baluška 2020), or the pervasive strategy "allowing organisms to adapt to their environment on time scales much faster than genetic selection" (Gershman et al., 2021).

3. With what rules?

So what rules - or "rules of engagement", as KB calls them - might enable organisms to anticipate rapidly changing, constantly-novel states in dynamically complex environments? Even logically, genetic instructions, viewed as fixed cue-response associations, would be inadequate. Is there some other possible, informative, basis? Like KB, many commentators are now proposing some sort of cognitive functions in organisms, even at cellular levels (Marshall 2021; Shapiro 2020). But what form do they take? I suggest something like the following.

Predictability in most real environments is beyond simple association because they are encountered as profusions of interacting variables, spatiotemporally nested at different statistical levels, rather than discrete recurring factors, mosaics or sequences. Furthermore, organisms do not exist in isolation or independently. They survive environments on many levels, from swirling molecular mileau, through masses of inanimate objects, to ecological communities teeming with other organisms, all interacting with each other, creating and facing constant change (Fisher and Pruitt, 2020). These conditions will have intensified in evolution as species were forced to extend ranges; and in development as conditions facing offspring departed from those experienced by parents.

Fortunately, nearly all complex environments contain underlying structural regularities - covariations dependent on others at different levels or "depths". Such patterns can be rich sources of predictability. A system that can abstract them for generating adaptive variation as required will be favoured over predetermined adaptations. As cited by KB, I call such abstracted structures "biogrammars", by analogy with speech grammars (Richardson 2020). The covariation structures underlying streams of novel speech sounds permit prediction of present and future intention or meaning (Keibel et al., 2009). By assimilating environmental structure, even within single cells, "a single network can display multiple stable dynamical solutions" (Koseska and Bastiaens, 2017).

Robert Rosen (mentioned by KB) offered a rigorous mathematical treatise on such "anticipatory systems" in biology. Properties emerge from the deep statistical relations in networks that transcend those of independent components. Basically, by using such information, living things don't just change their "state" in response to certain conditions; they also change the "rules" by which they do so - in other words they learn (Carrasco-Pujante et al., 2021). Godfrey-Smith (2001) called this "second order plasticity". The assimilation of statistical structure in experience thus forms the real, dynamic knowledge that enable organisms to survive uncertain futures. Learning, and the construction of knowledge

in that generative form, and at various levels, reflect on the MS and standard views of genes and evolution in various ways. I will suggest a few that bear on issues raised by KB.

4. Origins

Genes cannot bear the kind of knowledge that organisms in real, complex environments most need for survival. Unsurprisingly, therefore, research has made clear that the vast diversity of forms and functions in organisms across phyla are not direct reflections of diversity in gene sequences. Gerhart and Kirschner (1997) have noted this as another paradox: “where we most expect to find variation, we find conservation, a lack of change”. What seems to have been favoured is genes that can be flexibly recruited by system networks for adaptability in complex changing environments. That is what is starkly implied in KB's critique of the LTEE programme: “that gene mutation is not the source of variance that is being selected to improve fitness” (p. 13); and what makes his analogy of genes as the merchants that provide the necessary materials to build a house, but are neither the architect nor the builder, so apt.

Learning as structure-abstraction also bears on origins-of-life scenarios, emphasising that life is different from a mere mixture of the right chemicals. The essential property of life seems to be the maintenance of composition and relational integrity, in the (likely) highly changeable environments in the *Archean,* as in more recent times. Original “mixtures” could only have cohered and persisted in turbulent conditions because of covariation relations. They did so because such relations were thermodynamically efficient in free energy distribution and dissipation (see further below). When environmental change wrought on one component induced compensatory changes in another, or even changes that anticipated, nullified or amplified a future change, they became systems. System integrity over continual environmental change, at least for some period of time, is what most distinguished them from non-living molecular mixtures. Through their survivability, they would have been naturally selected, and evolved further into vast self-organising molecular networks. But they must have been “learning”, knowledge-forming, networks from the start.

Accordingly, in origins of life scenarios, the “genetics first vs. metabolism first” debate has swung decisively to the latter, indicating that living forms existed before genes. Russell et al., 2010 and others have suggested that conditions around deep-sea alkaline hydrothermal vents created many metabolism-like processes, including proton gradients across membranous linings of the vents as the energy drivers. Wimmer et al. (2021) claim to have discovered an ancient core of autotrophic metabolism encompassing 404 reactions involving H_2, CO_2, ammonia (NH_3), amino acids, nucleic acid monomers, and 19 cofactors required for synthetic pathways. From these, it is suggested, RNA and then DNA were eventually formed by steps as yet unclear. Lipid micells may have randomly encapsulated such sets of compositions in the prebiotic soup into primordial cells. Either way, phenotypes arrived before genes.

5. Development

The origins of increasingly complex forms and variation become an acute problem in the study of development. During a few weeks of embryogenesis over thirty trillion cells of 200 different types, and many more sub-types, are produced. They also move to just the right places at just the right times, in an orchestrated manner. That presents the problem of how genetically identical stem cells can “know” how to become one of so many different kinds, and where to move to. As KB says, “That information cannot be coded in the DNA. This is a major problem for the conventional theory”. As he goes on, “The ignorance here is fundamental and critical to any theory of development that involves genes as the unit of inheritance".

The veil of ignorance is perhaps being lifted a little. In development, as in evolution, differentiation is based on “outside” statistical patterns taken “inside”. Gene “expression” is turning out to be gene “recruitment”, downstream from the intelligent, knowledge-constructing, processes of the cell (and emerging organism) as a whole. Those processes are realised through the fact that gametes receive much more than genes from their parents. Many of these will be gene products, of course, but of genes that have been selected in the context of the wider composition and its functions ("facilitated evolution”, as described by Laland et al., 2014). So the egg includes transcription factors, promoters, enhancers, and a rich cellular *milieu* of RNAs, other proteins, fats, sugars, vitamins, metal salts, and so on. Then the sperm adds its own cargo, as well as some polarity to the ovum. In addition, epigenetic markers have been placed on offspring's genes, influencing how those genes should be used on the basis of parental experience. All of these ingredients are unevenly distributed. In consequence, when the egg divides, some of the daughter cells (the totipotent stem cells) will contain more of some of those biochemical constituents than others, providing relational knowledge to guide responses to other signals from outside.

Those come almost immediately, as the stem cells rain storms of signals on each other. The storms contain statistical structure, the cross-cutting “morphogen” gradients, themselves spatiotemporally patterned, and turning it all from cacophony to harmony (see contributions in Small and Briscoe 2020). The four-dimensional shape of the interactions between the gradients, their timing, duration and where they reach cell surfaces, all influence what goes on within each cell, including what genes to utilize and when. All of this, along with directional “guidance factors” and rich feed-forward/feedback cycles, constitutes what the stem cell “knows”, where to go and what to become: a self-organised programme made “on the hoof” not in the genes.

What is transmitted from parents to gametes, then, are evolved sets of components, each set capable of assimilating local knowledge (statistical structure) to create suitable form and variation. The contingencies under which each set operates - the rules of engagement they form - depend on the degree to which the environment being adapted to has been recurring across generations. We know from Waddington and others that, in some cases, including predictably required anatomical, physiological and behavioural traits, it is appropriate for development to be canalised against local perturbations. The stability of the developmental trajectory is then due to the evolved component sets able to exert compensatory interactions (Manu et al., 2009). Where conditions may be different between generations, inherited sets foster developmental plasticity. And these can be lifelong, as in many behavioural traits requiring brains and cognitive systems.

6. Metabolism

The same logic applies to the metabolism of the cell. Although allusions to “genetic programmes” are still conventional, as De la Fuente (2015) puts it, the “program of molecular instructions comes, not from the nucleus, but rather from the metabolic structure of the host cytoplasm.” Since these are induced in action, often changing the rules of engagement in the process, they have been described as “cell cognition” whose function is “learning” as mentioned above. But descriptions of the nature of those functions have not always been clear. KB says, “cells and brains can acquire knowledge, both by trial and error and from stored information” (p5), but I'm not sure that conveys the creativity of systems anticipating uncertain futures.

Csermely et al. (2020, cited by KB) suggest the molecular networks correspond to a generalized Hebbian learning process. Such learning, at its simplest, means experience-based formation of responses to co-occurring events such as to make the responses more likely in future. There have been a number of extensions of Hebbian theory to accommodate associations between numerous, often non-linear, and "complex-weighted", variables as in "deep learning" ANNs (artificial neural networks). "In ANNs, learning refers to the process of extracting structure—statistical regularities—from input data, and encoding that structure into the parameters of the network" (Zador, 2019 p. 3770). Trained ANNs can respond creatively to complex stimuli, as in recognising novel (but structurally related) inputs. De la Fuente et al. (2020) provide evidence that self-organised metabolic networks seem to be governed by attractor dynamics "similar to what happens in neural networks". It is also becoming clear that learning networks in cell metabolism involve scores of elements (nodes such as proteins) and deep interactions, with multi-level dependencies, non-linearities and feedback loops (Hasson et al., 2020).

There has been debate, even around the most successful ANNs, concerning what exactly is being "learned". However, they all seem to involve covariations conditioned by others at different levels - not only the simple pairwise associations, but the deeper statistical dependencies. The "connection weights" induced thus correspond to the structural parameters (and collectively the "grammars") captured in metabolic, and in brain, networks (Michalski et al., 2014). Those are the parameters describing the structure of experience to be assimilated: the knowledge that is far more providential than the residue of simple cue-reponse associations, or trial and error learning. It is measurable in theory as the "mutual information": the amount of information, or reduction of uncertainty, afforded by the relationships (Gabrié et al., 2018).

Why have such systems evolved at so many levels in living systems? Apart from affording predictability in uncertain futures, it is probably because maximising nested covariation - or mutual information - across vast networks is conducive to thermodynamic efficiency, a driver of self-organisation maintaining continuous coupling with the ever-changing world (De la Fuente et al., 2013; Barato et al., 2014). This is the knowledge "that gives independent agency to the cell" (KB). Or, to extend, somewhat, comments by Rovelli, cited in a footnote by KB, this is, "how an entropy gradient can give rise to the behaviour we recognise as agency … a physical mechanism that transforms low entropy into information. This may be the general mechanism at the source of the whole information on which biology builds." Elsewhere, Rovelli (2020) points out that, "the mutual-information-rich structures all around us are memory structures from past thermodynamical imbalances".

Crucially, KB cites Rosen about the necessity of both syntax and semantics in living processes. Note that the fostering of predictability in a system, as just described, does impart semantics or meaning to processes - the meaning being whatever a process anticipates/predicts for other processes. This distinguishes such information, as true knowledge, from that of the basic Shannon information theory (Kolchinsky and Wolpert, 2018). Jakulin and Bratko (2004) called it "interaction information" and generalize the idea to an arbitrary number of variables. In reviewing ways in which microbes gain a sense of future states in their changing environments, Freddolino and Tavazoie (2012) note how cellular behaviour is orchestrated in response to the "meaning" of an environmental perturbation, not only its direct and immediate consequences. I think that is consistent with Omholt's (2013, p. 75) remark about enabling "living systems to self-transcend beyond those morphogenetic limits that exist for non-living open physical systems in general". In sum, investigators are probing what seems to be rich resources of "knowledge-making", for dealing with highly changeable environments, far beyond the relatively limited information in genes.

7. For example?

So far, so abstract. But such ideas are also becoming grounded in real molecular instantiations. They start with the cell surface. Cells have often been viewed as input-output devices mechanically passing on extracellular stimuli through surface receptors to internal signaling pathways, eventually triggering responses determined by genetic programs. As Con (2013) explains, the idea that cell receptors could cooperate, exchange cross talk, and adjust responses to the covariation structure of the environment was not envisaged until relatively recently. However, to a chemotactic bacterium tracking a glucose source, the environment is not a series of independent cues, but a dynamic, spatio-temporal storm of signals. Independent "hits" of nutrient molecules on surface receptors give little precise indication of source direction. That knowledge is only constructed from the spatiotemporal structure of the molecular gradient on the curved surface. There, large numbers of non-additive interactions between receptors are observed, leading to signal amplification and integration (Galstyan et al., 2019). As Shimizu et al. (2010) put it, the "chemotaxis-signaling pathway computes time derivatives of chemo-effector concentrations." So a cell membrane has been referred to as a kind of "little brain" (see Lyon, 2015; Vallverdú et al., 2018).

Examples of contingent molecular behaviour within cells are now legion. Kar, Nelson and Parekh (2012) showed how one particular signaling molecule acts as a coincidence detector, sensitive to the timing of two other signals. The response of the detector, including gene transcription, only occurs if the stimuli producing the signals are received sufficiently close together in time. Sometimes contingency-dependence is evident in the contrasting behaviour of components according to the dynamic state of the network of which they are part. There are bi-functional enzymes capable of two opposing reactions, either activating or repressing gene transcription depending on context; and cytokines that signal a cell to either proliferate or die (Hart and Alon 2013). The final specificity of transcription factors (TFs) requires the cooperative activity of numerous promotors and enhancers, reflecting spatiotemporal informational structure rather than discrete cues. Such structures, being constantly updated, enable efficient self-organised control of metabolism in changing environments.

The complexity of such contingency is writ large in major communication networks as in G-protein coupled receptors. Kapolka et al. (2021) note that, "In humans alone, over 800 GPCRs detect stimuli such as light, hormones, and metabolites to guide cellular decision-making primarily using intracellular G-protein signaling networks. This diversity is further enriched by GPCRs that function as molecular sensors capable of discerning multiple inputs to transduce cues encoded in complex, context-dependent signals". The networks positively and negatively modulate GPCR signaling. Csermely et al. (cited by KB) provide other examples of components primed to await the next signal in a pre-activated state, and how the concerted activations of such states in signaling cascades contribute to cellular learning.

Of course this knowledge-guided adaptability in metabolic networks rests in the versatility of their components. Enzymes have main functions, but many have multi-specificities producing catalytic side-effects often creating novel biochemical pathways. They generate cross-wirings between existing network parts, allowing the network to react rapidly to perturbations in metabolite or enzyme concentrations (Notebaart et al., 2014).

Biddle et al. (2021) have closely studied the integration of

transcription factors (TFs) and their many co-regulators in gene recruitment. They show how integration depends on “higher order cooperative interactions”, facilitated through allostery of components across the whole signaling network. Jacques Monod famously described allostery as “the second secret of life” (see Lorimer et al., 2018). However, the true wealth of conformational ensembles present in most cellular processes is now being revealed. They are turning out to be a fundamental property of all protein-protein interactions involving numerous components, often at long intracellular distances. Research into their statistical properties (Ghode et al., 2020) shows not only their connection with a free energy landscape but also how we need to think of dynamic “allostery ensembles”. Biddle et al. (2021) demonstrated the emergence of such ensembles as “higher-order cooperativities” (HOCs), in which binding is collectively modulated by multiple other binding events. Ligand-binding effects could not be explained by a simple parameter such as ligand concentration but seemed dependent on integration of information from many levels. For example, The Hunchback gene, which is thought to have six binding sites for the TF Bicoid, requires HOCs up to order 5 (that is, five levels of statistical dependency) to account for observed gene expression.

It is now known that intrinsially disordered proteins (IDPs), referred to by KB, play a large part in such cooperatives, and thus of learning within them. IDPs are a large class of proteins lacking definitive structure and function except through interaction with a wide range of other macromolecular partners. Christoffer and Kihara, 2019 say, “Protein-protein interactions (PPIs) involving an IDP are major players in the PPI network, comprising an estimated 15%–45% of all interactions. These disordered PPIs are prevalent in various important cellular processes and are associated with allosteric regulation, posttranslational modification, and alternative splicing”. Csermely et al. (cited by KB) suggest that 85% of human signaling proteins contain intrinsically disordered regions. It is difficult to understand the prominence of such “disorderly” proteins except as partners in higher-order learning networks, evolved to survive in complex, changeable environments. That understanding completely reverses Dawkins' prioritisation of (stable) genes over (changeable) phenotypes.

Declaration of competing interest

No conflicting interests are declared.

References

Barato, A.C., Hartich, D., Seifert, U., 2014. Efficiency of cellular information processing. New J. Phys. 16, 103024. https://doi.org/10.1088/1367-2630/16/10/103024.

Biddle, J.W., Martinez-Corral, R., Wong, F., Gunawardena, J., 2021. Allosteric conformational ensembles have unlimited capacity for integrating information. eLife 10, e65498. https://doi.org/10.7554/eLife.65498.

Buss, D.M., von Hippel, W., 2018. Psychological barriers to evolutionary psychology: ideological bias and coalitional adaptations. Arch Sci Psychol 6, 148–158.

Carrasco-Pujante, J., Bringas, C., Malaina, I., Fedetz, M., Martínez, L., Pérez-Yarza, G., Dolores Boyano, M., Berdieva, M., Goodkov, A., López, J.I., Knafo, S., De la Fuente, I.M., 2021. Associative conditioning is a robust systemic behavior in unicellular organisms: an interspecies comparison. Front. Microbiol. 12, 707086. https://doi.org/10.3389/fmicb.2021.707086.

Christoffer, C., Kihara, D., 2019. Modeling protein-protein interactions with intrinsically disordered proteins. In: Salvi, N. (Ed.), Intrinsically Disordered Proteins. Academic Press, New York, pp. pp189–pp206.

Con, P.M., 2013. Receptor-receptor interactions. Academic Press, New York.

Csermely, P., Kunsic, N., Mendik, P., Kerestely, M., Farago, T., Veres, D.V., Tompa, P., 2020. Learning of signaling networks: molecular mechanisms. Trends Biochem. Sci. 45 (4), 284–294.

Fisher, D.N., Pruitt, J.N., 2020. Insights from the study of complex systems for the ecology and evolution of animal populations. Curr Zool 66 (1), 1–14.

Freddolino, P.L., Tavazoie, S., 2012. Beyond homeostasis: a predictive-dynamic framework for understanding cellular behavior. Annu. Rev. Cell Dev. Biol. 28, 363–384. https://doi.org/10.1146/annurev-cellbio-092910-154129.

De la Fuente, I.M., 2015. Elements of the cellular metabolic structure. Front Mol Biosci 2, 16. https://doi.org/10.3389/fmolb.2015.00016.

De la Fuente, I.M., Cortes, J.M., Pelta, D.A., Veguillas, J., 2013. Attractor metabolic networks. PLoS One 8 (3), e58284. https://doi.org/10.1371/journal.pone.0058284.

De la Fuente, I.M., Bringas, C., Malaina, I., Fedetz, M., Carrasco-Pujante, J., Morales, M., Knafo, S., Martinez, L., Pérez-Samartín, A., López, J.I., Pérez-Yarza, G., Boyanop, M.D., 2020. Evidence of conditioned behavior in amoebae. Nat. Commun. 10, 3690. https://doi.org/10.1038/s41467-019-11677-w.

Gabrié, M., Manoel, A., Luneau, C., Barbier, J., Macris, N., Krzakala, F., Zdeborová, L., 2018. Entropy and mutual information in models of deep neural networks. Adv. Neural Inf. Process. Syst. 31, 1821–1831.

Galstyan, V., Funk, L., Einav, T., Phillips, R., 2019. Combinatorial control through allostery. J. Phys. Chem. B 123, 2792–2800.

Gerhart, J., Kirschner, M., 1997. Cells, Embryos and Evolution. Blackwell Science, Oxford.

Gershman, S.J., Balbi, P.E.M., Gallistel, C.R., Gunawardena, J., 2021. Reconsidering the evidence for learning in single cells. eLife 10, e61907. https://doi.org/10.7554/eLife.61907.

Ghode, A., Gross, L.Z.F., Tee, W.V., Guarnera, E., Berezovsky, I.N., Biondi, R.M., Anand, G.S., 2020. Synergistic allostery in multiligand-protein interactions. Biophys. J. 119 (9), 1833–1848. https://doi.org/10.1016/j.bpj.2020.09.019.

Godfrey-Smith, P., 2001. Environmental complexity and the evolution of cognition. In: Sternberg, R., Kaufman, J. (Eds.), The Evolution of Intelligence. Erlbaum, New York.

Harden, K.P., 2021. The Genetic Lottery. Princeton University Press, Princeton, N.J.

Hart, Y., Alon, U., 2013. The utility of paradoxical components in biological circuits. Mol. Cell 49, 213–221.

Hasson, U., Nastase, S.A., Goldstein, A., 2020. Direct fit to nature: an evolutionary perspective on biological and artificial neural networks. Neuron 105, 416–434. https://doi.org/10.1016/j.neuron.2019.12.002.

Jakulin, A., Bratko, I., 2004. Quantifying and visualizing attribute interactions: an approach based on entropy. http://arxiv.org/abs/cs.AI/0308002.

Kapolka, N.J., Rowe, J.B., Taghon, G.J., Morgan, W.M., O'Shea, C.R., Isom, D.G., 2021. Proton-gated coincidence detection is a common feature of GPCR signaling. Proc. Natl. Acad. Sci. U.S.A. 118 (28), e2100171118. https://doi.org/10.1073/pnas.2100171118.

Kar, P., Nelson, C., Parekh, A.B., 2012. CRAC channels drive digital activation and provide analog control and synergy to Ca^{2+}-dependent gene regulation. Curr. Biol. 22, 242–247.

Kiebel, S.J., Daunizeau, J., Friston, K.J., 2009. Perception and hierarchical dynamics. Front Neurodyn 3 article 20.

Kolchinsky, A., Wolpert, D.H., 2018. Semantic information, autonomous agency and non-equilibrium statistical physics. Interface Focus 8. https://doi.org/10.1098/rsfs.2018.0041, 201800412018004.

Koseska, A., Bastiaens, P., 2017. Cell signaling as cognitive process. EMBO J. 36, 568–582.

Laland, K., Odling-Smee, J., Turner, S., 2014. The role of internal and external constructive processes in evolution. J. Physiol. 592, 2413–2422. https://doi.org/10.1113/jphysiol.2014.272070.

Levins, R., 1968. Evolution in Changing Environments. Princeton University Press, Princeton, N.J.

Lorimer, G.H., Horovitz, A., McLeish, T., 2018. Allostery and molecular machines. Phil Trans R Soc *B* 373. https://doi.org/10.1098/rstb.2017.0173, 2017017320170173.

Lyon, P., 2015. The cognitive cell: bacterial behavior reconsidered. Front. Microbiol. 6. https://doi.org/10.3389/fmicb.2015.00264. Article 264.

Manu, S. Surkova, Spirov, A.V., Gursky, V.V., Janssens, H., Kim, A.-R., Radulescu, O., Vanario-Alonso, C.E., Sharp, D.H., Samsonova, M., Reinitz, J., 2009. Canalization of gene expression and domain shifts in the Drosophila blastoderm by dynamical attractors. PLoS Comput. Biol. 5 (3), e1000303. https://doi.org/10.1371/journal.pcbi.1000303.

Marshall, P., 2021. Biology transcends the limits of computation. Prog. Biophys. Mol. Biol. 2021. https://doi.org/10.1016/j.pbiomolbio.2021.04.006.

Michalski, V., Memisevic, R., Konda, K., 2014. Modeling deep temporal dependencies with recurrent grammar cells. Adv. Neural Inf. Process. Syst. 27, 1925–1933.

Noble, D., 2015. Evolution beyond neo-Darwinism: a new conceptual framework. J. Exp. Biol. 218, 7–13. https://doi.org/10.1242/jeb.106310.

Notebaart, R.A., Szappanos, B., Kintses, B., Pál, F., Á, Györkei, Bogos, B., Lázár, V., Spohn, R., Csörgő, B., Wagner, A., Ruppin, E., Pál, C., Papp, B., 2014. Network-level architecture and the evolutionary potential of underground metabolism. Proc. Natl. Acad. Sci. U.S.A. 111 (32), 11762–11767. https://doi.org/10.1073/pnas.1406102111.

Omholt, S.W., 2013. From sequence to consequence and back. Prog. Biophys. Mol. Biol. 111 (2e3), 75e82.

Plomin, R., 2018. Blueprint: How DNA Makes Us Who We Are. MIT Press, Cambridge, Mass.

Reber, A.S., Baluška, F., 2020. Cognition in some surprising places. Biochem Biophys Res Comms 564, 150–157. https://doi.org/10.1016/j.bbrc.2020.08.115.

Richardson, K., 2020. In the light of the environment: evolution through biogrammars not programmers. Biological Theory 15, 212–222.

Ridley, M., 2000. Mendel's Demon. Oxford University Press, Oxford.

Rovelli, C., 2020. Memory and Entropy arXiv:2003.06687, **[physics.hist-ph]**.

Russell, M.J., Hall, A.J., Martin, W., 2010. Serpentinization and its contribution to the energy for the emergence of life. Geobiology 8, 355–371.

Shapiro, J.A., 2020. All living cells are cognitive. Biochem Biophys Res Comms 564,

134–149.
Shimizu, T.S., Tu, Y., Berg, H.C., 2010. A modular gradient-sensing network for chemotaxis in Escherichia coli revealed by responses to time-varying stimuli. Mol. Syst. Biol. 6, 382.
Slobodkin, L.B., Rapoport, A., 1974. An optimal strategy of evolution. Q. Rev. Biol. 49, 181–199.
Small, S., Briscoe, J. (Eds.), 2020. Gradients and Tissue Patterning. Elseveier, London.
Smith, T.L., 2019. The roles of the analogy with natural selection in B.F. Skinner's philosophy. Behav. Process. 161, 139–148. https://doi.org/10.1016/j.beproc.2018.02.007.
Somero, G.N., Lockwood, B.L., Tomanek, L., 2017. Biochemical Adaptation: Response to Environmental Challenges from Life's Origins to the Anthropocene. Oxford University Press, Oxford.
Sultan, S.E., 2017. Developmental plasticity: re-conceiving the genotype. Interface Focus 7, 20170009. https://doi.org/10.1098/rsfs.2017.0009.
Vallverdú, J., Castro, O., Mayne, R., Talanov, M., Levin, M., Baluška, F., Gunji, Y., Dussutour, A., Zenil, H., Adamatzky, A., 2018. Slime mould: the fundamental mechanisms of biological cognition. Biosystems 165, 57–70.
Wimmer, J.L.E., Vieira, A., Xavier, J.C., Kleinermanns, K., Martin, W.F., Preiner, M., 2021. The autotrophic core: an ancient network of 404 reactions converts H2, CO2 and NH3 into amino acids, bases and cofactors. Microorganisms 9, 458. https://doi.org/10.3390/microorganisms9020458.
Zador, A.M., 2019. A critique of pure learning and what artificial neural networks can learn from animal brains. Nat. Commun. 10, 3770. https://doi.org/10.1038/s41467-019-11786-6.

Progress in Biophysics and Molecular Biology 167 (2021) 152–160

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Correspondence insights into the role of genes in cell functionality. Comments on “The gene: An appraisal” by K. Baverstock

A B S T R A C T

One of the most important goals of the post-genomic era is to understand the different sources of molecular information that regulate the functional and structural architecture of cells. In this regard, Prof. K. Baverstock underscores in his recent article “The gene: An appraisal” (Baverstock, 2021) that genes are not the leading elements in cellular functionality, inheritance and evolution. As a consequence, the theory of evolution based on the Neo-Darwinian synthesis, is inadequate for today's scientific evidence. Conversely, the author contends that life processes viewed on the basis of thermodynamics, complex system dynamics and self-organization provide a new framework for the foundations of Biology. I consider it necessary to comment on some essential aspects of this relevant work, and here I present a short overview of the main non-genetic sources of biomolecular order and complexity that underline the molecular dynamics and functionality of cells. These sources generate different processes of complexity, which encompasses from the most elementary levels of molecular activity to the emergence of systemic behaviors, and the information necessary to sustain them is not contained in the genome.

© 2021 Elsevier Ltd. All rights reserved.

Keywords:
Dissipative structures
Self-organization
Information processing
Hopfield dynamics
Epigenetic memory
Cellular migration
Metabolic processes

1. Dissipative self-organization, the main source of dynamic biomolecular order in the cell

All living cells exhibit highly ordered molecular dissipative processes and use the energy of nutrients to maintain non-equilibrium states, where sophisticated functional structures, complex physiological patterns, and collective molecular self-organization can emerge. These dynamic behaviors are very diverse, and fundamentally cover biomolecular oscillations, bi-rhythmicity (two stable oscillations), multi-stability (coexistence of several non-equilibrium quasi-steady states), circadian rhythms, and spatial traveling waves (De la Fuente, 2010; De la Fuente et al., 2014a; De la Fuente et al., 2021).

From a general point of view, dissipative self-organization occurs when spontaneous highly ordered dynamic structures far from thermodynamic equilibrium emerge; these dynamic behaviors are mainly characterized by coherent spatial and/or temporal patterns. Mainly due to complex nonlinear interactions among their components and driven by energy dissipation, the self-organized structures may increase the information and the structural-functional complexity of the systems (De la Fuente, 2010; De la Fuente et al., 2014a). The theoretical basis of self-organization was formulated in 1977 by the Nobel Prize Laureate in Chemistry Ilya Prigogine in his work on dissipative structures.

Enzymes are the essential molecules for biochemical life. When a set of them (for instance those belonging to a specific metabolic pathway) operates far enough from equilibrium and dissipative self-organization emerges, all these enzymatic macromolecules perform their biochemical activity as a whole, showing catalytic coordination between them (long-range coherence) in such a way that all the substrate and product concentrations spontaneously start to oscillate over time (temporal rhythms). As a consequence, thousands and thousands of molecules and ions that shape the enzymatic subsystem (substrates, products, protons, regulatory molecules, and other ions and metabolites) exhibit massive oscillatory reorganizations in their molecular concentrations. Such dynamics are mainly characterized by collective synchronized behaviors, functional correlations between molecular components separated by macroscopic distances, coherent patterns, and highly coordinated integrative processes (De la Fuente, 2010; De la Fuente et al., 2014a; De la Fuente et al., 2021).

Practically, all metabolite concentrations in cells present complex oscillations and/or non-equilibrium quasi-steady states (metabolite concentration drifts over time in a non-constant non-oscillatory way). The quantification of certain intracellular molecules through nano-biosensors in living cells has shown sophisticated transitions between non-equilibrium quasi-steady states and oscillatory behaviors whose dynamics are never constant, being able to exhibit infinite patterns of activity (De la Fuente et al., 2021). Different experimental pieces of evidence suggest that oscillatory behaviors are much more frequent than quasi-steady states in cellular conditions (De la Fuente et al., 2021).

Molecular-enzymatic oscillations were reported in practically all cellular processes, such as cytoskeletal components, cyclins, cyclic AMP, cytokines, free fatty acids synthesis, actin polymerization, biosynthesis of phospholipids, ATP, ADP and AMP nucleotides, urea cycle, proteolysis, glycolysis, metabolism of carbohydrates, intracellular glutathione, Krebs cycle, mitochondrial metabolic

https://doi.org/10.1016/j.pbiomolbio.2021.09.006
0079-6107/© 2021 Elsevier Ltd. All rights reserved.

processes, photosynthetic reactions, CO_2, protein kinase activities, transcription factors, respiratory metabolism, intracellular calcium, peroxidase-oxidase reactions, membrane receptor activities, membrane lipid oscillation, ERK/MAPK metabolism, intracellular pH, intracellular free amino acid pools, membrane potential, metabolism of mRNA, Min-proteins, beta-oxidation of fatty acids, amino acid transports, insulin secretion, etc. (De la Fuente et al., 2021; De la Fuente, 2015)

At a systemic level, a metabolic dissipative orchestration has been also observed in cells, through which the entire metabolome and most of the transcriptome dynamically oscillate. The oscillatory periods of metabolic rhythms range from milliseconds to minutes and hours, and complex periodic oscillations, including bursting rhythms and deterministic and stochastic chaos, have often been observed (De la Fuente et al., 2021).

Another class of self-organized dissipative processes is circadian rhythms, which exhibit a dynamic oscillatory period close to 24 hours (dark-light cycle during the Earth's rotation period). These endogenous autonomous oscillators allow adapting their internal metabolism to changes in the external environment (light, temperature, food availability, etc.) during 24h day/night cycles. Circadian rhythms exist in all types of cells from prokaryotes to eukaryotes, and these dissipative patterns regulate a great variety of important physiological processes. For instance, it has been observed in some cells that 80–90% of the transcriptome show a rhythmic gene expression with cycles of 24–26h (De la Fuente et al., 2021).

Lastly, a fundamental type of dissipative structures in cells is the spatial traveling waves, which consist of three-dimensional self-organized coherent oscillations of ions and metabolite concentrations that propagate progressively across the intracellular medium, reminiscent of a wave moving across water. These dynamic wave pulses of biochemical activity, moving through subcellular domains over large intracellular distances, are very fast (for instance, 5–30 $\mu m\ sec^{-1}$ for calcium waves) and represent an essential mechanism for long-range functional interconnection among different physiological processes at a global level. In fact, spatial traveling waves have a crucial role in the coordination and synchronization among numerous physiological processes and different subcellular organelles and functional structures. Some examples of spatial biochemical oscillations have been observed in calcium ions, actin dynamics during cell locomotion, apoptotic signals, mitochondrial redox, NAD(P)H, sodium ions, phosphoprotein processes, Cdk1 implicated in the cell cycle, mitotic processes, adenosine triphosphate, NAD(P)H and protons, phosphatidylinositol (3,4,5)-trisphosphate, ROS molecules, and mitochondria activity (De la Fuente et al., 2021; De la Fuente, 2015).

Intensive studies over the last five decades have demonstrated that millions and millions of molecules spontaneously self-organize at any moment of the cellular life shaping a sophisticated orchestration of different temporal and spatial patterns (De la Fuente et al., 2021; De la Fuente, 2015). These dissipative processes are the main source of the biomolecular order of cells and constitute one of the most genuine properties of the basic unit of all known forms of life. Reactive cellular processes have little to do with the Chemistry of Equilibrium. Enzymatic-physiologic activities inside the cell are regulated by complex temporal dissipative patterns mainly coordinated and synchronized by spatial molecular waves, far from thermodynamic equilibrium.

Another relevant mechanism of molecular organization in the cell is self-assembly e.g., formation of the lipid bilayer, viral capsid, protein aggregates to hold the quaternary structure, some supramolecular polymerization, etc. Self-assembly and dissipative self-organization are the fundamental pillars of the molecular organization of all living organisms (De la Fuente et al., 2021).

Prof. K. Baverstock is right in his article (Baverstock, 2021). All dissipative patterns, that shape the functional and molecular architecture of the cell and drive the enzymatic activity, are not stored in the genes. Genetic information and dissipative information correspond to completely different physical realities. Dissipative self-organized processes use the energy inflow to generate a negative entropy variation in the cellular open system which corresponds to an emergent positive increment in the information contained in the cell itself. Such information increases the complexity, being able to produce highly-ordered macrostructures and complex functional dynamic behaviors. Non-linear interactions and enzymatic irreversible processes may amplify fluctuations leading to a dynamic state, far from the equilibrium, in which the biochemical system increases its information and becomes spatially and temporally self-organized. Genetic information does not store any self-organized pattern. Dissipative processes cover the whole cell and, due to global impact, their quantitative repercussion and physiological importance, the information generated by self-organization processes constitute the main source of dynamic biomolecular order in all basic units of life.

2. Molecular information processing, the second fundamental source of order in the cell

An essential characteristic of the biochemistry of life is that enzymes shape modular dissipative networks, which perform fundamental relatively autonomous activities with specific and coherent catalytic patterns (De la Fuente, 2015). These networks, cornerstones of the cellular functionality, not only originate the emergence of dissipative *self-organized* patterns structured in space and time, but also are capable to produce another complex behavior by information processing such as the *self-regulatory* control of cellular activity. Studies on effective connectivity based on Transfer Entropy have made possible to quantify in bits the biomolecular information flows that emerge in the dissipative metabolic networks. These informative processes make possible that the cellular functionality behaves as a complex decentralized information processing system, similar to parallel computing, which highly increases the information and complexity of the cell system (De la Fuente et al., 2021; De la Fuente, 2015).

For instance, alternative splicing occurs during gene expression in all the kingdoms of life. This complex molecular mechanism is a highly regulated enzymatic-molecular process that allows obtaining with precision and efficiency different mRNAs and proteins from a primary transcript of mRNA or pre-mRNA. In fact, to carry out this activity, there is one of the most complex enzymatic networks in the cell, the spliceosome, which is capable to notably increase the molecular information and complexity in the cellular system. So, the Dscam1 gene in *Drosophila melanogaster* (analog of the human DSCAM gene) holds the record, up to now, for alternative splicing, being able to originate more than 38,000 different proteins from a single gene. Note that the entire *Drosophila melanogaster* genome has only 15,016 genes. It is necessary to underline also that more than 90% of all human genes undergo alternative splicing, which constitutes one of the major known sources of diversity for cellular proteome. These complex processes that increase the molecular information of cell system are not stored in any genetic sequence.

Another example of information processing is the regulation of *Escherichia coli* chemotaxis network by which bacterial cell swimmers integrate environmental information accurately to make proper decisions for their survival. So, as a result of the chemotaxis process, the signal transduction translates external molecular information into appropriate motor responses and the bacterial cells traverse gradients of chemical attractants, displaying efficient directional sensing and movement by making temporal

comparisons of ligand concentrations. In *E. coli*, membrane receptors responsible for external signal transduction assemble into large clusters of interacting proteins, shaping a complex modular network connected to flagellar motor. The system is mainly regulated by two reversible post-translational modifications: phosphorylation and carboxyl methylation. The dynamic changes in chemotaxis network functionality provoke specific molecular information processing and, as a consequence, structural and functional molecular modifications occur. They are carried out in such a way that the cell can record the recent attractant inputs using fundamentally reversible methylation processes in specific glutamic acid residues (Li and Stock, 2009). The dynamic methylation-demethylation patterns linked to *information processing* allow the detection of attractant gradients by comparing current concentrations to those encountered in the recent past. The carboxyl methylation mechanism stores information concerning environmental conditions that the bacterium has experienced, and these dynamic patterns of glutamyl modifications act as a structural dynamic memory, not stored in genes, which enables cells to respond efficiently to continual changes in attractant concentrations in the external medium (Stock et al., 2002; Stock and Zhang, 2013).

Information processing is also necessary to synthesize many complex biomolecules in the cell. So, the lipidome, the full set of lipids, may comprise over 1000 different molecular types for a single cell, and 10,000–100,000 molecular species for tissues or organisms, many of which exhibit very complex structural sequences, and all of them originating from a few hundred individual lipid classes. Moreover, some glycans also show complex molecular configurations, and their cellular repertoire is estimated to be 10–100 times larger than proteome and lipidome, depending on the species. Besides, many classes of glycoproteins and glycolipids exhibit complex structures whose sequential patterns neither are stored in the DNA sequences. In addition, most proteins can also present complex processes of post-translational modifications (PTMs) (De la Fuente, 2015) which are of utmost importance for cellular functionality. In fact, more than 200 diverse types of PTMs are currently known and the super-complex mark patterns they develop are not stored in the genes. Cells convert energy into dissipative organization and information processing, and as a result, the most complex molecules of known nature are synthesized.

Numerous examples of *information processing* can be observed in different molecular activities such as reversible phosphorylation of proteins (Thomson and Gunawardena, 2009), cell membrane (Gatenby, 2019), endoplasmic reticulum (Stroberg et al., 2019), signal transduction (Roper, 2007), transcriptional regulatory activities (Makadia et al., 2015), enzymatic systems (Katz and Privman, 2010), redox regulation (Dwivedi and Kemp, 2012), cytoskeleton (Frieden and Gatenby, 2019), transmembrane flow of ions (Gatenby and Frieden, 2017), biochemical networks (Bowsher, 2011), NF-kappaB dynamics (Tay et al., 2010), gene regulatory networks (Gabalda-Sagarra et al., 2018), intracellular signaling reactions (Purvis and Lahav, 2013; Kamimura and Kobayashi, 2012), metabolic switches (Ramakrishnan and Bhalla, 2008) and network motifs (Alon, 2007).

Information processing also highly increases the cellular complexity at a global level. *Physarum polycephalum* is a paradigmatic example of the emergence of such complex systemic behaviors. This unicellular microorganism can discover the minimum-length option between two distant points in a labyrinth (Nakagaki et al., 2000; Nakagaki, 2001). Note that finding the shortest path problem in a maze needs a rigorous mathematical solution (Miyaji and Ohnishi, 2008). *P. polycephalum* is also capable of improving the selection of the better route configuration obtained by the shortest Steiner's minimum tree connections thus developing adapted strategies to maximize its access to nutrients (Nakagaki et al., 2004a, 2004b). Even more, *P. polycephalum* achieves to solve difficult problems, for example, finding a high-quality solution to the problem of the traveling salesman, a question which is known to be NP-hard (Aono et al., 2011a, 2011b; Zhu et al., 2011). This multinucleated amoeba is capable of designing an optimal network very similar to the purpose-intended system in the Tokyo railway organization (Tero et al., 2010). During the process of adaptation to different inputs, *P. polycephalum* succeeds in memorizing changes occurring in its environment, recalling them later to adapt its behavior to the new conditions appropriately, for example, anticipating a cold-dry pattern in the environment 1 h before the change happens (Saigusa et al., 2008). It has also been observed that this protist accomplishes complex dilemmas of multi-objective foraging (Dussutour et al., 2010; Bonner, 2010; Latty and Beekman, 2011). Recently it has been shown that *P. polycephalum* is capable of developing a kind of rudimentary learning (Boisseau et al., 2016).

From the most elementary levels of enzymatic-molecular activity to the emergence of systemic behaviors, cells exhibit different orders of complexity, and the information necessary to sustain them is not contained in the genes.

3. Systemic molecular turnover, the fundamental dynamics of cell life

Unlike the reductionist cellular conception, the cell is a complex super-dynamic molecular system, extremely self-organized and self-regulated, in a permanent recycling status, which is characterized by continuous reactive dynamics of self-construction and self-destruction.

All molecular components of the cell are in this dynamic state of reactive transformations. The proteome, lipidome, glycome, metabolome, and transcriptome are synthesized and degraded continually in all cell types following sophisticated interdependent processes. Even outside the growth period, the intracellular macromolecular pool is dynamic and considerable energy is expended in the continuous processes of synthesis and degradation. For instance, only the protein turnover consumes 38–47% of the total energy produced in every cell (Lahtvee et al., 2014) (see for more details *Supplementary Material 01* in[4]).

Not only molecules but also cellular structures are subjected to cycles of construction and destruction. For instance, the endoplasmic reticulum, the largest endomembrane system, exhibits a permanent turnover; the cytoskeleton is another molecular dynamic system that undergoes continuous reorganization through the synthesis and degradation of its structural elements during the cell cycle; the mitochondria are dynamic organelles that are incessantly undergoing fusion, fission, and molecular destruction; the peroxisome turnover also takes place by autophagy-related mechanism; also, cellular membranes are highly reactive dynamic structures in nonstop recycling, so, in 30 minutes, an active cell such as a macrophage recycles, by endocytosis and exocytosis, an amount of plasma membrane that equals its complete plasma membrane (see for more details *Supplementary Material 01* in[4]).

Millions of biochemical reactive transformations happen simultaneously in every basic unit of life at any time. Nothing is inert in the cell and all the molecules and structures that make up each cell undergo complex chemical transformations. The consequences of these incessant self-organized and self-regulated dynamics of reactive molecular transformations are the adequate cellular growth, the development of all physiological processes to maintain its functional structures, the continuous adaptation to the environment, and, lastly, the mitosis. Alterations in these turnover dynamics may lead to different pathologies and if the dynamic collapses

the cell dies.

The continuous recycling and incessant chemical transformations that encompass practically all molecules and structures shape a critical and unique scenario in which cellular life is possible (De la Fuente et al., 2021). The orderly maintenance of the self-construction and self-destruction molecular dynamics is the essential element for life. This is the fundamental systemic characteristic of all basic units of life.

The millions of reactive molecular patterns of synthesis and destruction that characterized the cell functionality are not stored by genetic information; actually there is not any codified program in the genome that governs these molecular turnover dynamics. The highly complex organization of continuous reactive dynamics of self-construction and self-destruction that defies the human intellect is not stored in the genes. Any attempt to synthetically reproduce this global molecular turnover, either *in vitro* or *in silico*, has failed so far. In every cell, molecular synthesis and destruction are compensated and harmonized between them, following complex and unrepeatable reactive patterns whose laws and defining principles are still unknown.

The cell is not a mere "genoteque" (molecular box governed by genes). Each cell is a dynamic reactor in which millions of biochemical reactions tightly interrelated and integrated into sophisticated networks shape the most complex molecular system known in nature. This reactive super dynamic biochemical system is characterized by continuous molecular turnover in which self-construction and self-destruction of molecules and substructures occur following complex dissipative self-organized and self-regulated enzymatic patterns coordinated and synchronized by complex spatial traveling waves. The super complex dynamics originated by this huge molecular turnover constitute the fundamental systemic characteristic of all basic life units. Not considering the dynamics of molecular changes in the cell represents an unfortunate example of naive simplicity and reductionism still prevalent in our time.

4. Enzymes and not genes are the essential molecular actors of the functional architecture of life

The synthesis and molecular destruction that are permanently taking place in the basic units of life are carried out by enzymes. They are responsible for practically all recycling processes of the cell, and these molecular transformations are essentially chemical reactions, that is, modifications on how atoms are bonded. Enzymes are the most extraordinary macromolecular nanomachines in nature, responsible for breaking and joining covalent bonds, the fundamental chemical links in biological molecules. Through their ability to decrease the bond activation energy, enzymes permanently modify the way atoms are linked, making possible the accelerated creation and destruction of molecules (catalytic activities). Accordingly, enzymes are the main elements of this amazing super dynamic cellular biochemical reactor.

Unlike the rest of the molecular species, enzymes are the only ones that are active, capable of chemical work. They are necessary for almost all biomolecular transformations. Most enzymes are proteins, but a few RNA molecules called ribozymes also manifest catalytic activity. Enzymatic reactions are termed metabolic processes; and there are two basic types of enzymatic activities, catabolism (involved in molecule destruction) and anabolism (implicated in molecule synthesis). Together as a whole, these enzymatic activities are called cellular metabolism. In short, the cell is a complex enzyme-mediated metabolic system where reactive molecular transformations of synthesis and destruction happen unceasingly *accelerated* by catalytic activities (De la Fuente et al., 2021).

There are a lot of important biomolecules in the cell, such as ATP, nucleotides, lipids, structural proteins, etc., but enzymes constitute the unique active biomolecules capable of chemical work with the faculty to rearrange the atoms, i.e., the ability to transform some molecules into others creating or destroying bonds between atoms. These biocatalysts are responsible for the formation of all the ATP in the cell, DNA synthesis, apoptosis, mitosis, cytoskeleton organization, DNA repair, alternative splicing, chromosome regulation, gene expression, post-translational modifications, Golgi apparatus activity, etc. In short, all the main activities in the cell are mediated by enzymes (De la Fuente et al., 2021; De la Fuente, 2015). They constitute the fundamental elements of all essential physiological processes, which allow a cell to grow, multiply, and adapt to the external medium. However, although biologically essential, many issues of enzyme activities remain poorly understood which deserve further investigation.

Genes encode information about the primary structure of each enzyme, and this amino acid sequence determines its *catalytic specificity*. However, the *enzymatic functionality*, the set of different patterns of activity carried out by every enzyme, cannot be predicted from the primary structure.

In this regard, it is necessary to take into account that enzymes are not rigid molecules; they have complex dynamic conformations and are continuously undergoing a wide range of conformational fluctuations which modify their activities so that catalytic processes are not constant under biological conditions. Even in the native state, each enzyme exhibits a range of complex interconverting conformations driven by thermodynamic fluctuations (Ramanathan et al., 2014; Agarwal, 2019). Besides, the dynamic changes in the hydration shell rapidly control these conformational fluctuations of enzymes (they can be described by an energy landscape) that hinder the connection between enzyme sequence and its catalytic functionality (Fenimore et al., 2004; Dunaway-Mariano, 2008).

In addition, the activities of enzymes can be complexly regulated by specific factors. So, the enzymatic functionality primarily depends on changes in substrate concentrations, pH, temperature, ions and inhibitors or activators. Due to intricate collective interactions undergone in cellular conditions, substrate fluxes and regulatory substances shape complex dynamical topological structures in which ions and metabolite concentrations are continuously changing over time. Moreover, enzymes do not work independently one each other in the cellular molecular crowding, instead they shape different types of dissipative multienzyme associations (metabolic subsystems networks) (De la Fuente et al., 2021) and, as a consequence, the substrates and regulatory molecular flows permanently change with time exhibiting complex activity patterns with transitions between quasi-steady states and oscillatory behaviors. As said above, when a metabolic network operates sufficiently far from equilibrium, self-organized catalytic activities can emerge. Such dissipative dynamic behaviors find their roots in non-linear regulatory processes, e.g., feedback loops, cooperativity, and stoichiometric autocatalysis (Goldbeter, 2002). Therefore, enzymatic activities are essentially a function of self-organized and self-regulated dynamics that emerge in the dissipative enzymatic networks. The collective enzymatic connectivity, integrated into dissipative biochemical networks and complex dynamical topological architectures, makes highly variable the rate at which catalytic reactions proceed, being able to exhibit practically infinite patterns of enzymatic activity (De la Fuente et al., 2014a; De la Fuente, 2015). Such overflowing variety of patterns that characterize the *enzymatic functionality* are not stored in the genome.

The enzymatic activity also strongly depends on the two principles of the catalytic organization in the cell: *metabolic segregation* and *systemic metabolic integration* (De la Fuente, 2015). Both are not excluding, but complementary, and the interplay between

them represents a fundamental property for the emergence of systemic metabolism and the functional global architecture of cells. See here some examples:

All catalytic reactions that occur in living unicellular organisms are segregated functionally shaping modular enzymatic networks, with specific and coherent autonomous activities, grouped by their metabolic tasks and catalytic roles (metabolic segregation). In fact, cellular metabolism is segregated into multiple autonomous biochemical specializations originating different types of catalytic activities for instance, lysosomal metabolism, fatty acid β-oxidation, oxidative phosphorylation, signal transduction, etc. Specialized modular segregation networks have been detected in practically all known physiological activities, e.g., metabolic pathways, chaperone activities, chemotaxis, apoptosis processes, cell cycle, Golgi apparatus processes, and kinetochore organization. In particular, specific modular metabolic networks also take part in the complex *transcriptional system* such as those that regulate miRNA, transcription factors, RNA polymerase complexes, nucleosome activities and mRNA dynamics among others (De la Fuente, 2015).

While *metabolic segregation* expresses the partial interdependence of specialized metabolic networks, *metabolic integration* reflects high deviation activities from the functional autonomy of these modular complexes in such a way that the collective metabolic elements of the cell produce coherent systemic dynamics that are functionally integrated, shaping a super-complex systemic dynamic structure: the "*Cellular Metabolic Structure*" (De la Fuente, 2015). Such a *Systemic Metabolic Structure* was observed for the first time in 1999 in an exhaustive numerical analysis with several millions of different dissipative metabolic networks (De la Fuente et al., 1999). The emergent enzymatic global organization mainly consists of a small set of different dissipatively self-organized multienzyme complexes which always present active states, while the rest of dissipative multienzyme subsystems exhibit on-off active states. Under this integrative metabolic activity, each active multienzyme subsystem generates output responses with complex dynamic transitions between quasi-steady states and oscillatory patterns. The set of multienzyme complexes always within an active state are called the metabolic core (De la Fuente et al., 1999; De la Fuente et al., 2008). Later, this global enzymatic system was verified employing flux balance metabolic analysis in several prokaryotes and eukaryote cells such as *S. cerevisiae*, *H. pylori*, and *E. coli* (Almaas et al., 2004, 2005). The emergence of this systemic metabolic structure, of great importance for the cell, is not stored in genes.

Adenylate energy system constitutes another global process responsible for the functional integration of cellular metabolism. Energy is an essential element to keep the metabolic architecture of cells, and adenosine nucleotides couple all bio-energetic enzymatic processes to each other. The adenosine nucleotide levels are determined by the adenylate energy system, and such dynamic energetic functional structure is shaped by enzymatic reactions which interconvert AMP, ADP, and ATP, as well as energy-consumption processes coupled to the synthesis of ATP (De la Fuente et al., 2014b). In 1967, a primary systemic ratio between ATP, ADP, and AMP concentrations was proposed by Atkinson (Atkinson and Walton, 1967) to calculate the energetic cellular level, which was denominated the Adenylate Energy Charge (AEC). Numerous experimental cellular energy quantifications show that despite the extreme complex fluctuations in the adenine nucleotide concentrations in the cell almost all organisms appear to keep their AEC within narrow values under growth conditions, more specifically between 0.7 and 0.95. This quantitative physiological invariant represents a key property of the integrative mechanisms of energetic and functional dynamics operating at the systemic cellular level that is supported by numerous experimental researches (De la Fuente et al., 2014b). There is no genetic program for the emergence of this global energetic invariant. There are no genetic programs that govern dynamic metabolic processes at any level including systemic processes. In fact, extensive studies performed with dissipative metabolic networks (around 15,210,000 of different systemic networks), to research the mechanism behind of the emergence of the *Cellular Metabolic Structure (CMS)*, allowed to observe that this global superstructure is a property common to all metabolic systems with a *high number* of dissipative self-organized multienzyme complexes (De la Fuente et al., 2009). According to these quantitative studies, the fundamental factor that ensures the spontaneous emergence of *CMS* is *a large multiplicity of dissipative processes* inside the cell, and not the information contained in the genes (De la Fuente, 2015; De la Fuente et al., 2009).

The great redundancy of the enzymes (with an increased number of copies) that occurs inside the cell, involved in the molecular turnover, and the consequent multiplicity of dissipative self-organized multienzyme complexes that shape it, would determine the spontaneous formation of the systemic metabolic structure.

A continuous process of molecular synthesis and destruction represents thousands and thousands of dissipative metabolic reactions that happen permanently and simultaneously in the cell, ensuring the emergence, robustness and stability of the *Cellular Metabolic Structure* (De la Fuente et al., 2009).

In an endless process, thousands of dissipative anabolic processes create thousands of molecular species that feed the catabolic metabolism (molecular destruction), which supplies a permanent pool of molecular residues to start the cycles of molecular synthesis anew. As result of this process, a large number of dissipative processes are always active in the cell, which allows the spontaneous emergence and permanent maintenance of a robust *CMS* (De la Fuente et al., 2009).

Roughly 3700 million years ago, an exceptional and singular systemic metabolic structure (*CMS*) emerged from primeval matter, characterized by a high structural and functional order, improbable for the Chemistry of Equilibrium. This extraordinary molecular organization perpetuates itself by direct transmission after mitosis, and there is no scientific proof that another parallel metabolic organization has emerged "de novo".

The only possible scenario for cellular life is a dynamic system in permanent self-construction and self-destruction process, which guarantee the functionality of a high number of dissipative self-organized multienzyme complexes, permanently actives inside the cell. There is no other alternative for the life (De la Fuente et al., 2021).

The cell is essentially a super metabolic dynamic system fed back by thousands and thousands of dissipative processes involved in its own synthesis and destruction, permanently avoiding the collapse of its recycling dynamics. This sophisticated and massive dynamics of endless molecular reactive biotransformations encompasses the whole metabolic system and constitutes the most critical dynamic property, essential and definitory of cellular life.

The cell is not a molecular genoteque governed by genes. Cells are very complex super-dynamic metabolic reactors, extremely self-organized and self-regulated, characterized by continuous reactive dynamics of self-construction and self-destruction performed by enzymes. Enzymes and not genes are the essential molecular actors of the functional architecture of life.

5. Epigenetic memories

Epigenetic memory is another essential process of cellular metabolic life that governs the inheritance of previously acquired new

functional characteristics. This biochemical mechanism also represents a huge amount of molecular information not contained in DNA sequences. Note that around 75% of all CpG dinucleotides exhibit methylation marks in mammal somatic cells.

This type of non-genetic memory process can be originated mainly by chromatin-based modifications such as DNA methylation, histone post-translational modifications including acetylation, phosphorylation, methylation and ubiquitylation, incorporation of histone variants, and different chromatin modulating factors.

The role of the epigenetic memory in evolution is also fundamental, because the rates of epigenetic changes in response to experiences in previous generations are fast and easy reversed. The stable propagation of these modifications in the gene expression provides a way for variation within a species that rapidly increase during stress, providing an opportunity for adaptation to selection pressures.

Complex reprogramming of metabolic information marks in primordial germ cells according to their parental origin is a fine example of heritable biochemical patterns which are not caused by changes in the DNA sequence (Seisenberger et al., 2012; Hanna and Kelsey, 2014). The transmission of epigenetic metabolic information across generations is also observed in bacteria (Casadesús and Low, 2013; Gonzalez et al., 2014). As an interesting curiosity, a stably inherited DNA methylation pattern is reported in a variant of *Linaria vulgaris,* originally described more than 272 years ago by Linnaeus (1749); in this example, the fundamental symmetry of the flower is found to change from bilateral to radial as a result of extensive methylation of the Lcyc gene that controls the formation of the dorsal petals (Cubas et al., 1999).

In addition to epigenetic inherited marks, other memory-related processes are present in cell life. In this regard, Hopfield-like dynamics characterized by exhibiting associative memory were quantitatively verified in dissipative metabolic networks in 2013 (De la Fuente et al., 2013). Such a memory is a manifestation of emergent properties underlying the complex dynamics of the systemic metabolic networks when dissipative enzymatic self-organization and molecular information processing act together. This functional metabolic memory exhibits two main dynamic informational mechanisms (De la Fuente, 2015). The first one occurs at the level of the systemic self-organized metabolic network, in which Hopfield-like emergent dynamics have the capacity to store functional catalytic patterns that can be correctly recovered by specific input stimuli. The second mechanism occurs at the post-translational modulation level, when specific molecular information can be transferred from the functional dynamics of the metabolic networks (Hopfield-like dynamics) to the enzymatic activity involved in covalent post-translational modulation (PTMs), so that specific functional memory came from metabolic networks functionality can be reversibly embedded in multiple stable molecular marks (De la Fuente, 2015). As it is well known, these Hopfield-like dynamics in neural networks are characterized by manifesting associative memory (De la Fuente et al., 2013). Therefore, these studies were the first quantitative evidence that an individual cell can possess associative memory which would correspond to an epigenetic process (epigenetic non-inherited marks).

Recently, this thesis, cellular associative memory, has been confirmed by testing the cellular conditioning in individual amoeba cells (De la Fuente et al., 2019a). In these Pavlovian-like experiments, several hundreds of cells were capable to learn new systemic migratory behaviors and remember them over long periods relative to their cell cycle (44 minutes on average), forgetting them later. The cellular capacity of learning new adaptive systemic behaviors represents a fundamental evolutionary mechanism for cell adaptation. In a last confirmatory experiment, the cellular conditioning of more than 2000 individual cells belonging to three different species: *Amoeba proteus*, *Metamoeba leningradensis*, and *Amoeba borokensis* have been exhaustively studied (Carrasco-Pujante et al., 2021). The quantitative results allow to conclude that associative conditioning seems to be a universal characteristic of unicellular organisms (Carrasco-Pujante et al., 2021).

Regarding this type of epigenetic non-inherited memory, it should be noted that while bare DNA encodes 2 bits of information per nucleotide, the reversible chemical modifications occurring at multiple sites even in a single protein allow the encoding of a potentially large amount of information. For instance, a protein with n phosphorylated sites has an exponential number (2^n) of phospho-forms; the number of phosphorylated sites on a protein shows a significant increase from prokaryotes (with $n \leq 7$ sites) to eukaryotes, with examples having $n \geq 150$ sites, implying that a protein can encode a large amount of molecular information as a function of a varying number of covalent phospho-marks (Thomson and Gunawardena, 2009). The reversible changes by molecular marks on structural proteins through PTM (regulatory enzymatic networks) enable to store a high amount of molecular information (much superior to that of DNA), and develop complex information processing, which can originate flexibility and adaptive metabolic responses in the cell (Thomson and Gunawardena, 2009; Sims and Reinberg, 2008; Sunyer et al., 2008).

On the other hand, parallel to cell conditioning, different experiments have been carried out to analyze whether genes govern cell migration. For a wide range of cells, from prokaryotes to eukaryotes, self-locomotion is one of the most important systemic behaviors of unicellular organisms endowed with directional movements (De la Fuente and López, 2020). Free cells move in the right direction with the appropriate speeds and location, looking for adequate food, avoiding adverse conditions, and predators. In multicellular organisms, directed and precise locomotion movements are the result of a very complex systemic process, highly coordinated and carefully regulated, involved in a plethora of fundamental physiological processes, such as embryogenesis, morphogenesis, organogenesis, adult tissue remodeling, wound healing, immunological cell activities, angiogenesis, tissue repair, cell differentiation, tissue regeneration, development of the nervous system, as well as in a myriad of others important physiological activities. One of the central questions regarding the cellular directed movement is the role of the nuclear genetic information in the regulation of the locomotion system. As is well known, genes have been repeatedly considered to be the main component that governs cell migration. However, it has been recently confirmed that genetic information does not regulate this critical systemic property for cell life. So, the movement trajectories of enucleated and non-enucleated *Amoeba proteus* have been analyzed exhaustively using advanced non-linear physical-mathematical tools (mainly Statistical Mechanics) and computational methods (De la Fuente et al., 2019b). Specifically, to characterize the movements of cells and cytoplasts from a quantitative perspective, it was analyzed first the relative move-step fluctuation along their migratory trajectories by applying the root mean square fluctuation (rmsf). This approach is a classical method in Statistical Mechanics based on Gibbs and Einstein's studies that have been widely applied to quantify different time-series. Next, the Mean Square Displacement (MSD) was calculated to quantify the amount of space explored over time by the amoebas and the overall migration efficiency. This method was also proposed by Albert Einstein in his work concerning Brownian motion. Such analysis was validated by an alternative approach, the renormalization group operator (RGO) developed by the Nobel Prize Laureate in Physics Kenneth Wilson, who established the Theory of the Renormalization Group in 1971. The results of the analysis showed that both cells and cytoplasts display a very

similar kind of dynamic migration structure characterized by highly organized data sequences, non-trivial long-range positive correlations, persistent dynamics with trend-reinforcing behavior, super-diffusion, and move-step fluctuations with scale-invariant properties.

The systemic locomotion movements of cells and cytoplasts change continuously since all trajectories display random magnitudes that vary over time. But these stochastic movements shape a dynamic migration structure whose defining characteristics were preserved in all experiments. Such a dynamic migration structure characterizes the mathematical way in which the locomotion movements of enucleated and non-enucleated cells occur. Since the cytoplasts preserved the dynamic properties in their migration movements similarly to intact cells, the obtained results quantitatively confirmed that the nuclear genetic information does not significantly affect the systemic movements of amoebas in 2D environments. This conclusion, obtained from a mathematical and computational perspective, is consistent with the results reported by another group using exclusively biological techniques and different type of cells (Graham et al., 2018).

The nucleus, as a structural organelle, is important for developing appropriate mechanical responses and for regulating both contractility and mechano-sensitivity (Graham et al., 2018). More concretely, the *physical* presence, position, and material properties of the nucleus, fundamentally those related to its connections with the cytoskeleton, are essential for a broad range of cell functions. These functions mainly include intracellular nuclear movement, cell polarization, chromatin organization, cellular mechano-sensing, and mechano-transduction signaling. Eukaryotic cells require the *physical* presence of the nucleus as a necessary component of the molecular clutch involved in the regulation of their mechanical responses to the environment. The physical properties of the nucleus strongly connected with the cytoskeleton guarantee a proper cell migration when the environment displays mechanical complexities, as it happens in 3D conditions (Graham et al., 2018; Hawkins, 2018). The obtained results with enucleated and non-enucleated cells, quantitatively confirmed that the genetic information of the nucleus does not affect the control of locomotion movements. Definitively, the government of cellular migration is not directed by genes. So little genetic information cannot govern so much metabolic dynamics.

6. Conclusions

To summarize, a dissipative self-organized *Cellular Metabolic Structure* exists between the extracellular medium and the DNA (De la Fuente, 2015). This super-complex dissipative dynamic and highly refined structure behaves as a decentralized information processing system, generating sets of biochemical instructions that drive each enzymatic activity to a particular and precise dynamic of change, enabling the permanent self-regulation and adaptation to the external medium. In addition, the *CMS* permanently sends a flow of molecular signals to DNA-associated metabolism, which mainly shapes the complex transcriptional system. These molecular flows allow accurate regulation of gene expression, so that only specific polypeptides that are required for the *CMS* are synthesized (De la Fuente, 2015). The genetic expression is permanently subordinate to the metabolic needs of *CMS*, which govern the appropriate orchestration of the whole transcriptional system and, as a consequence, specific parts of the genome are activated and deactivated at any time. In accordance with the requirements of adaptive maintenance of cellular metabolism at each moment, gene expression is exactly regulated and only the genes required by *CMS* are expressed. *CMS* (which behaves as an individual metabolic entity) is able to learn and store dynamic functional patterns (Hopfield-like dynamics) and structural molecular information (in the form of post-translation marks, i.e., epigenetic non-inherited memory). Only a small part of the total molecular marks produced in the cell is inherited (epigenetic inherited memory) (De la Fuente, 2015).

Prof. K. Baverstock presents an adequate scientific approach to the role of genes in the cell (Baverstock, 2021). Life is not the result of a mere sum of genes. The cell contains much more information than its genome. And a genetic structure with so little information cannot govern a much more complex dynamic-functional system with an astounding high information value. A cell is not a molecular genoteque in evolution by random mutational changes (Noble, 2017). Cells are dissipative metabolic reactors in which millions of simultaneous self-ordered biochemical reactions are tightly interrelated and integrated into sophisticated networks that shape the most complex molecular system in nature. Their characteristics and properties are unique and breathtaking; they are capable of self-organization, self-regulation by information processing, learning, generation and storage of molecular information, as well as continuous adaptation of their functionality to the external medium. All basic units of life have the capacity to acquire new information, complexity, and new potentially adaptive physiologic characteristics, which can be transmitted for many generations. They can therefore alter their own heredity, which will be later selected by external environment pressures, but this essential evolutionary issue deserves a more detailed consideration in a broader writing and specific text.

It is truly amazing that in a reduced space of just a few microns, a super-complex dynamic system formed by millions and millions of biochemical reactions, in a permanent molecular recycling status, shapes a metabolic entity extremely self-organized and self-regulated, able to process and store large amounts of information …, and perpetuate itself in time. Over 3700 million of years, they have increased their information and complexity developing a high diversity of unusual and stunning biological adaptive forms …, up to some become self-aware. This extraordinary and singular super dynamic biochemical system represents the inexhaustible source of complexity developed by the dynamic forces underlying cellular metabolism.

Author contributions

The author confirms being the sole contributor of this work and approved it for publication.

Conflict of interest and authorship conformation form

Please check the following as appropriate:

- x All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.
- x This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.
- x The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript
- o The following authors have affiliations with organizations with direct or indirect financial interest in the subject matter discussed in the manuscript:

Declaration of competing interest

The author declares that he has no conflicts of interest and has

received no funding to support this work.

References

Agarwal, K.P., 2019. A Biophysical perspective on enzyme catalysis. Biochemistry 58, 438–449. https://doi.org/10.1021/acs.biochem.8b01004.

Almaas, E., Kovacs, B., Vicsek, T., Oltvai, Z.N., Barabási, A.L., 2004. Global organization of metabolic fluxes in the bacterium *Escherichia coli*. Nature 427, 839–843. https://doi.org/10.1038/nature02289.

Almaas, E., Oltvai, Z.N., Barabási, A.L., 2005. The activity reaction core and plasticity of metabolic networks. PLoS Comput. Biol. 1, e68. https://doi.org/10.1371/journal.pcbi.0010068.

Alon, U., 2007. Network motifs: theory and experimental approaches. Nat. Rev. Genet. 8, 450–461. https://doi.org/10.1038/nrg2102.

Aono, M., Zhu, L., Hara, M., 2011a. Amoeba-based neurocomputing for 8-city traveling salesman problem. Int. J. Unconv. Comput. 7, 463–480. https://doi.org/10.1016/j.biosystems.2013.01.008.

Aono, M., Zhu, L., Kim, S.J., Hara, M., 2011b. Performance enhancement of amoeba-based neurocomputer for 8-city traveling salesman problem. *Proc. NOLTA*. A2L-B3 104–107. https://doi.org/10.34385/proc.45.A2L-B3.

Atkinson, D.E., Walton, G.M., 1967. Adenosine triphosphate conservation in metabolic regulation. Rat liver citrate cleavage enzyme. J. Biol. Chem. 242, 3239–3241.

Baverstock, K., 2021. The gene: an appraisal. Prog. Biophys. Mol. Biol. 164, 46–62. https://doi.org/10.1016/j.pbiomolbio.2021.04.005.

Boisseau, R.P., Vogel, D., Dussutour, A., 2016. Habituation in non-neural organisms: evidence from slime moulds. Proc. R. Soc. B. 283, 20160446. https://doi.org/10.1098/rspb.2016.0446.

Bonner, J.T., 2010. Brainless behavior: a myxomycete chooses a balanced diet. Proc. Natl. Acad. Sci. U.S.A. 107, 5267–5268. https://doi.org/10.1073/pnas.1000861107.

Bowsher, C.G., 2011. Information processing by biochemical networks: a dynamic approach. J. R. Soc. Interface 8, 186–200. https://doi.org/10.1098/rsif.2010.0287.

Carrasco-Pujante, J., Bringas, C., Malaina, I., Fedetz, M., Martínez, L., Pérez-Yarza, G., Dolores Boyano, M., Berdieva, M., Goodkov, A., López, J.I., Knafo, S., De la Fuente, I.M., 2021. Associative conditioning is a robust systemic behavior in unicellular organisms: an interspecies comparison. Front. Microbiol. 12, 707086. https://doi.org/10.3389/fmicb.2021.707086.

Casadesús, J., Low, D.A., 2013. Programmed heterogeneity: epigenetic mechanisms in bacteria. J. Biol. Chem. 288, 13929–13935. https://doi.org/10.1074/jbc.R113.472274.

Cubas, P., Vincent, C., Coen, E., 1999. An epigenetic mutation responsible for natural variation in floral symmetry. Nature 401, 157–161. https://doi.org/10.1038/43657.

Dunaway-Mariano, D., 2008. Enzyme function discovery. Structure 16, 1599–1600. https://doi.org/10.1016/j.str.2008.10.001.

Dussutour, A., Latty, T., Beekman, M., Simpson, S.J., 2010. Amoeboid organism solves complex nutritional challenges. Proc. Natl. Acad. Sci. U.S.A. 107, 4607–4611. https://doi.org/10.1073/pnas.0912198107.

Dwivedi, G., Kemp, M.L., 2012. Systemic redox regulation of cellular information processing. Antioxidants Redox Signal. 16, 374. https://doi.org/10.1089/ars.2011.403.

Fenimore, P.W., Frauenfelder, H., McMahon, B.H., Young, R.D., 2004. Bulk-solvent and hydration-shell fluctuations, similar to alpha- and beta-fluctuations in glasses, control protein motions and functions. Proc. Natl. Acad. Sci. U.S.A. 101, 14408–14413. https://doi.org/10.1073/pnas.0405573101.

Frieden, B.R., Gatenby, R.A., 2019. Signal transmission through elements of the cytoskeleton form an optimized information network in eukaryotic cells. Sci. Rep. 9, 6110. https://doi.org/10.1038/s41598-019-42343-2.

De la Fuente, I.M., 2010. Quantitative analysis of cellular metabolic dissipative structures. *Int. J. Mol. Sci.* 11, 3540–3599. https://doi.org/10.3390/ijms11093540.

De la Fuente, I.M., 2015. Elements of the cellular metabolic structure. Frontiers in Molecular Biosciences 2, 16. https://doi.org/10.3389/fmolb.2015.00016.

De la Fuente, I.M., López, J.I., 2020. Cell motility and cancer. Cancers 12, 2177. https://doi.org/10.3390/cancers12082177.

De la Fuente, I.M., Benítez, N., Santamaría, A., Veguillas, J., Aguirregabiria, J.M., 1999. Persistence in metabolic nets. Bull. Math. Biol. 61, 573–595. https://doi.org/10.1006/bulm.1999.0103.

De la Fuente, I.M., Martínez, L., Pérez-Samartín, A., Ormaetxea, L., Amezaga, C., Vera-López, A., 2008. Global self-organization of the cellular metabolic structure. PLoS One 3, 1–19 e3100.

De la Fuente, I.M., Vadillo, F., Pérez-Pinilla, M., Vera-López, A., Veguillas, J., 2009. The number of catalytic elements is crucial for the emergence of metabolic cores. PLoS One 4, e7510:1–e7510:11.

De la Fuente, I.M., Cortes, Jesús, Peta, David, Juan Veguillas, 2013. Attractor metabolic networks. PLoS One 8 (3), e58284.

De la Fuente, I.M., Aon, M., Saks, V., Schlattner, U., 2014a. Metabolic dissipative structures. En: Systems Biology of Metabolic and Signaling Networks: Energy, Mass and Information Transfer. Springer Books, New York.

De la Fuente, I.M., Cortés, J.M., Valero, E., Desroches, M., Rodrigues, S., Malaina, I., et al., 2014b. On the dynamics of the adenylate energy system: homeorhesis vs homeostasis. PLoS One 9, e108676. https://doi.org/10.1371/journal.pone.0108676.

De la Fuente, I.M., Bringas, C., Malaina, I., Fedetz, M., Carrasco-Pujante, J., Morales, M., Knafo, S., Martinez, L., Pérez-Samartín, A., López, J.I., Pérez-Yarza, G., Boyano, M.D., 2019a. Evidence of conditioned behavior in amoebae. Nat. Commun. 10, 3690.

De la Fuente, I.M., Bringas, C., Malaina, I., Regner, B., Pérez-Samartin, A., Boyano, M.D., Fedetz, M., López, J.I., Pérez-Yarza, G., Cortés, J.M., Sejnovski, T., 2019b. The nucleus does not significantly affect the migratory trajectories of amoeba in two-dimensional environments. Sci. Rep. 9, 16369.

De la Fuente, I.M., Martínez, L., Carrasco-Pujante, J., Fedetz, M., López, J.I., Malaina, I., 2021. Self-organization and information processing: from basic enzymatic activities to complex adaptive cellular behavior. Front. Genet. 12, 644615. https://doi.org/10.3389/fgene.2021.644615.

Gabalda-Sagarra, M., Carey, L.B., Garcia-Ojalvo, J., 2018. Recurrence-based information processing in gene regulatory networks. Chaos 28 (10), 106313. https://doi.org/10.1063/1.5039861.

Gatenby, R.A., 2019. The role of cell membrane information reception, processing, and communication in the structure and function of multicellular tissue. Int. J. Mol. Sci. 20, 3609. https://doi.org/10.3390/ijms20153609.

Gatenby, R.A., Frieden, B.R., 2017. Cellular information dynamics through transmembrane flow of ions. Sci. Rep. 7, 15075. https://doi.org/10.1038/s41598-017-15182-2.

Goldbeter, A., 2002. Computational approaches to cellular rhythms. Nature 420, 238–245. https://doi.org/10.1038/nature01259.

Gonzalez, D., Kozdon, J.B., McAdams, H.H., Shapiro, L., Collier, J., 2014. The functions of DNA methylation by CcrM in *Caulobacter crescentus*: a global approach. Nucleic Acids Res. 42, 3720–3735. https://doi.org/10.1093/nar/gkt1352.

Graham, D.M., Andersen, T., Sharek, L., Uzer, G., Rothenberg, K., Hoffman, B.D., et al., 2018. Enucleated cells reveal differential roles of the nucleus in cell migration, polarity, and mechanotransduction. J. Cell Biol. 217, 895–914. https://doi.org/10.1083/jcb.201706097.

Hanna, C.W., Kelsey, G., 2014. The specification of imprints in mammals. Heredity 113, 176–183. https://doi.org/10.1038/hdy.2014.54.

Hawkins, R.J., 2018. Do migrating cells need a nucleus? J. Cell Biol. 217, 799–801.

Kamimura, A., Kobayashi, T.J., 2012. Information processing and integration with intracellular dynamics near critical point. Front. Physiol. 3, 203. https://doi.org/10.3389/fphys.2012.00203.

Katz, E., Privman, V., 2010. Enzyme-based logic systems for information processing. Chem. Soc. Rev. 39, 1835–1857. https://doi.org/10.1039/b806038j.

Lahtvee, P., Seiman, A., Arike, L., Adamberg, K., Vilu, R., 2014. Protein turnover forms one of the highest maintenance costs in Lactococcus lactis. Microbiology 160, 1501–1512. https://doi.org/10.1099/mic.0.078089-0.

Latty, T., Beekman, M., 2011. Speed-accuracy trade-offs during foraging decisions in the cellular slime mould *Physarum polycephalum*. Proc. Biol. Sci. 278, 539–545. https://doi.org/10.1098/rspb.2010.1624.

Li, Z., Stock, J.B., 2009. Protein carboxyl methylation and the biochemistry of memory. Biol. Chem. 390, 1087–1096. https://doi.org/10.1515/BC.2009.133.

Linnaeus, C., 1749. *De Peloria*, Diss. Ac. Amoenitates Academicae III, Uppsala.

Makadia, H.K., Schwaber, J.S., Vadigepalli, R., 2015. Intracellular information processing through encoding and decoding of dynamic signaling features. PLoS Comput. Biol. 11, e1004563. https://doi.org/10.1371/journal.pcbi.1004563.

Miyaji, T., Ohnishi, I., 2008. Physarum can solve the shortest path problem on riemannian surface mathematically rigourously. Int. J. Pure Appl. Math. 47, 353–369.

Nakagaki, T., 2001. Smart behavior of true slime mold in a labyrinth. Res. Microbiol. 152, 767–770. https://doi.org/10.1016/S0923-2508(01)01259-1.

Nakagaki, T., Yamada, H., Toth, A., 2000. Maze-solving by an amoeboid organism. Nature 407, 470. https://doi.org/10.1038/35035159.

Nakagaki, T., Kobayashi, R., Ueda, T., Nishiura, Y., 2004a. Obtaining multiple separate food sources: behavioural intelligence in the *Physarum plasmodium*. Proc. Biol. Sci. 271, 2305–2310. https://doi.org/10.1098/rspb.2004.2856.

Nakagaki, T., Yamada, H., Hara, M., 2004b. Smart network solutions in an amoeboid organism. Biophys. Chem. 107, 1–5. https://doi.org/10.1016/S0301-4622(03)00189-3.

Noble, D., 2017. Evolution viewed from physics, physiology and medicine. Interface Focus 6, 20160159. https://doi.org/10.1098/rsfs.2016.0159.

Purvis, J.E., Lahav, G., 2013. Encoding and decoding cellular information through signaling dynamics. Cell 152, 945–956. https://doi.org/10.1016/j.cell.2013.02.005.

Ramakrishnan, N., Bhalla, U.S., 2008. Memory switches in chemical reaction space. PLoS Comput. Biol. 4 (7), e1000122. https://doi.org/10.1371/journal.pcbi.1000122.

Ramanathan, A., Savol, A., Burger, V., Chennubhotla, C.S., Agarwal, P.K., 2014. Protein conformational populations and functionally relevant substates. Acc. Chem. Res. 47, 149–156. https://doi.org/10.1021/ar400084s.

Roper, S.D., 2007. Signal transduction and information processing in mammalian taste buds. Pflueg. Arch. Eur. J. Physiol. 454 (5), 759–776.

Saigusa, T., Tero, A., Nakagaki, T., Kuramoto, Y., 2008. Amoebae anticipate periodic events. Phys. Rev. Lett. 100, 018101. https://doi.org/10.1103/PhysRevLett.100.01810.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., et al., 2012. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol. Cell. 48, 849–862. https://doi.org/10.1016/j.molcel.2012.11.001.

Sims, R.J., Reinberg, D., 2008. Is there a code embedded in proteins that is based on post-translational modification? Nat. Rev. Mol. Cell Biol. 9, 815–820. https://doi.org/10.1038/nrm2502.

Stock, J.B., Zhang, S., 2013. The biochemistry of memory. Curr. Biol. 23 (17), R741–R745. https://doi.org/10.1016/j.cub.2013.08.011.
Stock, J.B., Levit, M.N., Wolanin, P.M., 2002. Information processing in bacterial chemotaxis. Sci. STKE 132, pe25. https://doi.org/10.1126/stke.2002.132.pe25.
Stroberg, W., Eilertsen, J., Schnell, S., 2019. Information processing by endoplasmic reticulum stress sensors. J. R. Soc. Interface 16, 20190288. https://doi.org/10.1098/rsif.2019.0288.
Sunyer, B., Diao, W., Lubec, G., 2008. The role of post-translational modifications for learning and memory formation. Electrophoresis 29, 2593–2602. https://doi.org/10.1002/elps.200700791.
Tay, S., Hughey, J.J., Lee, T.K., Lipniacki, T., Quake, S.R., Covert, M.W., 2010. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. Nature 466, 267–271. https://doi.org/10.1038/nature09145.
Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D.P., Fricker, M.D., et al., 2010. Rules for biologically inspired adaptive network design. Science 327, 439–442. https://doi.org/10.1126/science.1177894.
Thomson, M., Gunawardena, J., 2009. Unlimited multistability in multisite phosphorylation systems. Nature 460, 274–277. https://doi.org/10.1038/nature08102.
Zhu, L., Aono, M., Kim, S.-J., Hara, M., 2011. Problem-sizes scalability of amoeba-based neurocomputer for traveling salesman problem. Proc. NOLTA. 45, 108–111. https://doi.org/10.34385/proc.45.A2L-B4.

Ildefonso M. De la Fuente[a,b]
[a] *Department of Nutrition, CEBAS-CSIC Institute, Espinardo University Campus, Murcia, 30100, Spain*
[b] *Department of Mathematics, Faculty of Science and Technology, University of the Basque Country, UPV/EHU, Leioa-Bilbao, 48940, Spain*
E-mail addresses: ildefonso@cebas.csic.es, mtpmadei@ehu.eus.

25 August 2021
Available online 6 October 2021

Progress in Biophysics and Molecular Biology 172 (2022) 82–89

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

A special role for the genotype? Some comments on Keith Baverstock: "The gene: An appraisal"

Keywords:
Gene concept
Genotype/Phenotype distinction
Reduction of biology to physics
Wilhelm Johannsen's selection experiment

A B S T R A C T

There is at present uneasiness about the conceptual basis of genetics. The gene concept has become blurred and there are problems with the distinction between genotype and phenotype. In the present paper I go back to their role in the creation of modern genetics in the early twentieth century. The terms were introduced by the Danish botanist and geneticist Wilhelm Johannsen in his big textbook of 1909. Historical accounts usually concentrate on this book and his 1911 paper "The Genotype Conception of Heredity." His bean selection experiment of 1900–1903 is generally assumed to be the source of his genotype theory. The present paper examines the scientific context and meaning of this experiment, how it was received, and how the genotype theory became securely established by the early 1910s. I argue in conclusion that the genotype/phenotype distinction, which provides the empirical basis for Johannsen's gene, was scientifically well founded when introduced and still is. Keith Baverstock's criticism does not consider the force of the bean selection experiment at the time and as a paradigm for following investigations of heredity.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

It is fair to say that the "gene" has been the guiding idea behind a century of amazing progress in biology. From the creation of modern genetics a little more than a hundred years ago. Nevertheless, the vision of "The gene as the basis of life" (Muller, 1926) has faded during recent decades. The theory of a gene-molecule that governs the phenomena of life has met with a sea of hard questions. And the promises of benefits for human health have proved elusive. Time appears ripe for radical rethinking. It is now a widespread view that the traditional mechanistic analyses of classical and molecular genetics should be replaced by theory that emphasizes the role of whole systems, like cells, organisms, and ecosystems (see for instance, Rheinberger and Müller-Wille 2015; Noble 2013, 2021).

The terms phenotype, gene and genotype were introduced by the Danish botanist and geneticist Wilhelm Johannsen 1857–1927) in 1909. Keith Baverstock challenges the traditional understanding of these concept. He argues that genotype and gene lack sound scientific grounding and that the phenotype, not the genotype, is the seat of biological heredity. In Baverstock's view the phenotype should be seen as the "governor and regulator of the cell." He proposes a new view of heredity built on physics. In particular "thermodynamics and complex systems dynamics" can provide a "simpler and more intuitive" understanding of life. (Baverstock 2021: 46).

Contrary to Baverstock I hold that even if the gene has become blurred the distinction between genotype and phenotype remains a foundation stone of genetics. But I think his challenge is worth considering. It points to a deep ambiguity linked to the relationship of biology to physical science. Baverstock starts from a brief ("potted") history of genetics with a paper by Johannsen in the key role. However, there are significant differences between Johannsen's original concept of genotype and the view that Baverstock criticises. Johannsen's paradigmatic bean selection experiment has been widely misunderstood in the historiography of genetics. Briefly: To be properly understood and evaluated Johannsen's genotype theory must be set into the scientific problem situation of his period.

2. Preformation and epigenesis, teleology and reductionism

Biology, the science of living things, is framed by two perennial questions: What explains the generation of living organisms, as kinds and as individuals? And what explains their striking purposiveness? Or, one could say: How do we explain the origin of life, and the origin of species?

Since Antiquity preformation and epigenesis have competed in answering these questions. The similarity between parent and offspring in kind as well as in individual characters suggests preformation. Something must be transferred that determines the properties of the adult. Epigenesis on the other hand builds on observation of the development of the individual organism from an amorphous beginning to an adult form. The Aristotelian idea of formal cause gave a preformationist answer to such questions. But throughout the early modern period the use of microscopes not least greatly increased knowledge in embryology, and by the end of the eighteenth century the ideas of preformation and epigenesis had been much clarified. (Mayenschein, 2005.)

https://doi.org/10.1016/j.pbiomolbio.2022.03.005
0079-6107/© 2022 Elsevier Ltd. All rights reserved.

More than 200 years ago Immanuel Kant's concept of the living organism as a *Naturzweck* ("natural purpose") summed up the philosophical situation in a way that is still relevant to biological science. *Kritik der Urteilskraft* (The Critique of Judgment), first published in 1799, defined a Naturzweck as a thing with an internal purpose, a thing that is its own cause and effect. Like a tree it causes the reproduction of the same kind of tree, it creates itself as an individual, and the parts are mutually dependent on each other for survival. (Kant, 1926: 286.) The fundamental difference between physics and biology is expressed in the antinomy (contradiction) of teleological judgement. On the one hand the strict criteria of certain scientific knowledge, embodied in physics, demands that all material things should be explained by mechanical causes alone. On the other hand, some material products of Nature cannot be fully explained in this way, their explanation demands final causes as well. (Kant, 1926: 314.) This Kantian compromise, his accommodation of physics and biology, has been characterized as "teleomechanism" (Lenoir 1982; McKaughan 2011), or as "critical teleology" (Roll-Hansen 1979; 2011).

In Kant's view the methodological difference between physics and biology is due to limitations of the human mind. Living things are too complex for us to be understood by the true principles of mechanical physical science alone. The implication is that biology necessarily relies on the less strictly rational concepts of teleological reasoning as a framework of all its explanations. In the last instance mechanistic explanations in biology are subordinate to teleology, mechanical processes become the tools of organic purpose. (Kant, 1924; paragraphs 80–81.) The precise limits of mechanistic science was an empirical question Kant thought. But he was convinced that human reason could never "hope to understand the creation of even a grass-straw from mechanical laws alone" (Kant, 1924: 353). There could be no "Newton of the grass-straw" (p. 337). The nineteenth and twentieth centuries saw repeated discussions about such limitations, and a continuing extension of physical explanations (Roll-Hansen 1979).

Through the nineteenth century new empirical knowledge about the chemistry of living organisms, the structure and behaviour of living cells, and the processes of fertilization and reproduction provided a new basis for the material science of life. This accumulation of knowledge continued through the following century. The reach of physicochemical explanation increased, not least through the new science of genetics. By the early twenty-first century questions about preformation and epigenesis, about the origin of life and purposiveness of organisms have become much more specific and precise, though no final solution is in sight. The competition and contradiction between preformation and epigenesis has not been eliminated (Mayenschein, 2005.)

Reductionism, the idea that ultimately biological phenomena can be explained by physical science, was a major topic in the nineteenth and twentieth centuries. The general principle stimulated lively public and scientific debate because of its epistemic and religious implications. The questions were many. What physical theory is in question? The present or some future theory? Does reduction imply determinism and absence of human free will? Is reduction incompatible with religious belief? On the fundamental ontological question no agreement has been reached, unsurprisingly one might say. But in biological science a more modest version, which we can call methodological reductionism, has played an important role. A recurring question has been: Is this biological phenomenon something that we just have to take at face value, as given, or can it be analysed and explained by physical science. There have been many attempts to set limits to reduction followed by success in breaching them. This is a significant experience but hardly enough to decide the ontological question. It still appears hard to get beyond the compromise of Kant's antinomy of teleological judgement.

To better understand the interaction of reductionist and anti-reductionist ideas it is helpful to look at two salient historical cases. They show, for instance, how anti-reductionist research programmes can contribute to reductionist progress.

3. Inspiration from physics

Louis Pasteur (1822–1895) was a founder of microbiology and a main example of how science can benefit human welfare. He was the big hero of the nineteenth century enlightenment ideal of science. He also had a vision of how to bridge the gap between biology and physical science. And a major part of his research was opposed to explicitly reductionist views, of leading chemists like Justus Liebig and Paul Bert. (Dagognet 1967; Roll-Hansen 2008). Trained as a physical chemist and crystallographer, Pasteur started his illustrious career in 1844 with the discovery of enantiomers in organic chemistry. Investigating the physical chemistry of tartaric acids he found a new way that the same atomic composition could give rise to two different kinds of molecule, different by crystal structure and effect on polarized light. The crystals were mirror images and turned the polarization of light respectively to the right and the left (Geison and Secord 1988.).

The idea that such asymmetry somehow represented the fundamental distinguishing nature of living organisms remained with Pasteur throughout his scientific career. At first the idea guided him effectively and quite directly to discoveries in organic chemistry and fermentation, and then inspired his path-breaking experiments on the question of spontaneous generation (Roll-Hansen 2008, 2018). It remained a fundamental belief with little or no direct effect on his experimental discoveries. Stereochemical models built on the tetravalent carbon atom gave a better explanation of enantiomers than Pasteur's physical ideas. But he experimented for some time with magnets and asymmetric dynamic arrangements to control the production of enantiomers, and he continued to hold the idea that asymmetric forces on the cosmic and molecular levels somehow played an essential role in the phenomenon of life. It is still an open question why l- and not d-amino acids dominate biochemistry on Earth.

In the early twentieth century reductionism was a hot topic. Many biologists and philosophers pointed to phenomena they believed to be beyond the reach of reductionist explanation. (Roll-Hansen, 1984.) The British physiologist J.S. Haldane (1860–1936) was famous for path-breaking research on respiration. He traced the chemical uptake of oxygen by haemoglobin to the point in the capillaries where oxygen dissociates from haemoglobin. Here the complexity of the living system bars further chemical analysis, he argued. Chemical structure and process becomes inseparable from the specific biological organization: "It is life and not matter that we have before us." (Haldane, 1923: 104.) The philosopher and biologist E.S. Russell (1887–1954) and the philosopher J.H. Woodger (1894–1981) both rejected the chromosome theory of the Morgan school. This mechanistic theory was incompatible with the holistic character of the phenomena of heredity, they claimed (Russell 1930; Woodger 1929; Roll-Hansen, 1984.).

The 1920s quantum revolution stimulated anti-reductionist thinking. Could indeterminism on the atomic level somehow provide room for human free will and finally resolve the Kantian antinomy? Pascual Jordan (1902–1980), one of the pioneers of quantum mechanics, argued for a quantum biology built on his amplifier theory. The idea was that acausal behaviour of single atoms could trigger causal processes on a higher level. He used atomic radioactive disintegration as analogy. The dean of the new physics, Niels Bohr (1885–1962), also stimulated the discussion with an

intriguing lecture on, "Light and Life," about the implications of the new physics for biology. (Sloan 2011: 64–70. Beyler 2011: 113. Roll-Hansen 2011: 154–157.).

Max Delbrück (1906–1981) was a young disciple of Bohr and became drawn to questions about the physical basis of biology. Like Bohr he took a distanced view of Jordan's speculations. But the interest led him into a lifelong career in molecular genetics (Sloan 2011.). Delbrück's first major contribution was a landmark paper on the gene as a macromolecule, written together with the Russian geneticist Nikolai Timoféef- Ressovsky (1900–1981) and the German radiation physicist Karl Günther Zimmer (1911–1988) (Timoféef-Ressovsky and Delbrück, 1935.). In 1937 Delbrück moved to Caltech (California Institute of Technology) to study genetics and soon became fascinated by bacteriophage as an ideally simple experimental system to study genetics on the molecular level. (Sloane and Fogel, 2011)

Delbrück became the organizer of an international research network, the so-called phage group, that had a key role in the creation of molecular genetics. The general approach was biochemical, extending chemical explanations of biological heredity. Delbrück was inspired throughout by an idea of radical complementarity between physical science and biology. In the end biochemical analysis of heredity would reveal phenomena that demanded a new physics, somewhat like discoveries in atomic physics had led to the quantum revolution. His "waiting for the paradox" was in vain, however. When the structure of DNA was discovered in 1953 he saw some hope which soon turned to disappointment. Delbrück conceded that traditional biological and structural chemistry had succeeded in genetics. Subsequent experiments on Phycomyces were also disappointing. This fungus was unusually sensitive to light apparently able to react on single photons (Roll-Hansen 2000.). Delbrück in 1969 received the Nobel Prize in "physiology or medicine" together with the microbiologists Salvador Luria (1912–1991) and Alfred Hershey (1908–1997) "for their discoveries concerning the replication mechanism and the genetic structure of viruses." The contributions of the phage group confirmed Muller's rather mechanistic chemical vision of the gene, not a complementarity of biology and physics in the spirit of Bohr and of Delbrück.

4. Toward a modern science of heredity

By the end of the nineteenth century new knowledge in reproduction, cytology, and biological chemistry together with the break-through of Darwinian theory of evolution had prepared for a modern theory of biological heredity. Speculations abounded about the material nature of heredity, how this material was transferred from one generation to the next, and how it determined morphological properties in each individual organism. Darwinism emphasized the historical dimension and stimulated speculation on the underlying physical mechanisms. Mostly the physical basis of heredity was thought to consist in small particles - "gemmules," "pangenes," etc. – which contained primordia or determinants for specific characters or parts. An alternative to particles was hereditary factors as processual states like waves on a string.

Variation was a main topic. The heritability of variation was an important practical problem in plant and animal breeding, and it was a crucial theoretical question for the evolution of species (Roll-Hansen, 2014b). Is variation in heredity continuous or discontinuous? This became the central question for the new science of genetics in its formative period around the turn of the nineteenth century. Roughly there were two opposed views. Continuous variation was preferred by orthodox Darwinians like the British biometricians Francis Galton (1822–1911), Karl Pearson (1857–1936), and W.F.R. Weldon (1860–1906). Other students of variation like the British zoologist William Bateson (1861–1926) and the Dutch botanist Hugo de Vries (1848–1935) believed in discontinuous variation. The two views were not in principle exclusive. The question was rather how much each kind contributed.

The biometricians also developed a substantial theory called the law of ancestral heredity. Galton showed by extensive measurements on different kinds of organisms how the distribution of individual characters in a population tended to fit the formula of normal distribution. To be stable through generations this distribution demanded a continuous regression toward the mean. A certain property of children, height for instance, must on average be closer to the mean than that of their parents. By extensive measurements Galton demonstrated the presence of such regression in different species of animals and plants. And the existence of such regression became a test for the biometric view of heredity.

According to de Vries and Bateson occasional and stepwise individual variations were essential for evolution. They pointed to observed examples of discontinuous change, "mutations" as de Vries called them. And they developed experiments to support their views. The experience of practical breeders was also an important input to the discussion. It appeared that in many cases the selection for varieties with desired properties was fruitless. Instead of the general malleability promised by Darwinian evolution through natural selection the breeders saw limits set by stable biological types.

This was the context where Johannsen began his investigations of biological heredity. The breeding of cereals was his source of inspiration. He was a cultured and well-educated citizen but had no academic degree. Trained as a pharmacist he received excellent on the job training at the Carlsberg laboratory, at that time the world centre of physiology related to brewing. And rose to become a prominent and highly respected professor at the University of Copenhagen.

In the 1880s and 1890s Johannsen worked in plant breeding. He was also a prolific populariser and textbook author. His little popular book on *Arvelighed og Variabilitet* (Heredity and variability) (Johannsen, 1896) and his contributions to the authoritative textbook in general botany, *Den Almindelige Botanik* (1900–1901), show how questions of variability guided his approach to heredity. The latter book was co-authored with his teacher and mentor in botany, Eugenius Warming (1841–1924), one of the founders of ecology. Johannsen's chapters on cytology, reproduction, evolution, and heredity demonstrate his broad and up to date familiarity with contemporary scientific issues. Variability was discussed at length, starting with the "polymorphism of traditional Linnean species," followed by variation in the progeny of hybrids, including a section on "laws of Mendelian segregation." He further distinguished a strict sense of "individual of fluctuating variability" from variability due to highly different external conditions. His fifth kind was "mutations," the rare and sudden appearance of new stable forms. (Warming and Johannsen, 1900-1: 666-674). Johannsen's involvement in plant breeding made him conscious of the need to clearly distinguish different kinds of variation and the importance of precise statistical measurement. The 1896 book was focused on Galton's ideas on heredity and his development of statistical methods. Johannsen's experience in plant breeding led him toward the discontinuity view of Bateson and de Vries, noting that Galton had shared similar ideas.

The "rediscovery" of Gregor Mendel's hybridization experiments in 1900 favoured discontinuity. Mendel's experiments with peas demonstrated how qualitative characters, like colour of flowers, surface of seed, or dwarf character of plants, were inherited according to precise laws of statistical nature. Bateson immediately grasped the opportunity for experimental research on heredity, quickly developed a flourishing Mendelian program for studies in heredity, and he named the new science "genetics." Weldon and

Pearson were sceptical about the accuracy of Mendel's classification of individuals according to their type. They doubted the validity the so-called Mendel's Laws and the general implications drawn from them. These laws may fit many qualitative characters well, but it is quite unclear how they can explain the inheritance of quantitative characters like the weight of seed, and their content of starch, sugars, or proteins. Such quantitative properties were essential in agriculture and the brewing industry.

Bateson was very successful in spreading his interpretation of Mendel and building a school of experimental research on inheritance of qualitative characters, while the biometric theory fell into disrepute. Bateson was so successful that by the 1920s the established genetic theory was simply called Mendelian. Biometric contributions have largely been overlooked in history of genetics. One result is that Wilhelm Johannsen has been characterized simply as a Mendelian despite the fact that he began as a biometrician rather than a Mendelian. His bean selection experiment began in biometric spirit and it was a surprise that selection had no effect within pure lines. (Johannsen 1903). This biometric discovery supported a discontinuity view of heredity and immediately attracted much international attention. At first Johannsen's claim to have dissolved the law of ancestral heredity through internal critique met much criticism and doubt. But within a decade his experiment was generally accepted as a fundamental contribution.

5. The bean selection experiment

In 1900 Johannsen set out to test the scope of the biometric law of regression using the statistical methods of Galton and Pearson together with Vilmorin's isolation principle. He hoped in this way to decide "whether there is a real difference between mutation and fluctuating variability." That is, between the underlying stepwise change in heredity claimed by De Vries and Bateson and the continuous variation of properties as described by the normal distribution (Johannsen 1903: 8.). In the 1903 monograph he points specifically to two examples of successful pedigree breeding. The first example is the French plant breeder Louis Leveque de Vilmorin (1816–1860). He held that the best way to tell the hereditary "force" of an individual, i.e., its value as breeding material, was to isolate and analyse its offspring. Vilmorin applied this isolation principle with success to the breeding of sugar beet. Johannsen's second example is the Swedish plant breeding station in Svalöf headed by the botanist Hjalmar Nilsson (1856–1925). From the start in 1886 the Svalöf station used the ordinary mass selection method, selecting in successive generations the best individuals in a population. In the early 1890s Nilsson reformed the Svalöf breeding program by introducing pedigree method together with sophisticated botanical systematics to classify the breeding material. By the early 20th century Svalöf achieved worldwide fame as the "Mekka" of plant breeding, cereals in particular (Roll-Hansen 1978, 1986.).

The experimental material was deliberately chosen. Johannsen knew that inbreeding produced hereditary homogeneity, as Mendel's laws of segregation and distribution confirmed (Warming and Johannsen, 1900-1: 679–683). This meant that in a population with a high frequency of self-fertilization all individuals would tend to have the same set of hereditary factors. They would be homozygous, as Bateson called it. Johannsen defined "pure lines" as lineages "descending from one single self-fertilizing individual." Such lines would not be affected by variation due to hybridization and their behaviour should therefore be "the real foundation of a theory of heredity." It would be "the simplest case," he argued (Johannsen 1903: 8–9.). Johannsen aimed at a general theory covering the reproduction both of cross-fertilizing and self-fertilizing organisms, but the latter was the best place to start.

Fundamental questions of biology were at the back of Johannsen's mind. The first paragraph of his 1903 monograph begins: "In no part of biology is the unity of life more evident than in all questions concerning fertilization and heredity." The most outstanding scientists in this field had emphasize the general relevance of such knowledge. It applied equally to man and aphids, to beans and barley. This consideration ("Erwägung"), argued Johannsen, "gives me the courage - and perhaps also the justification - to give the present publication a general title" even if its objects all belong to the vegetable kingdom (Johannsen 1903:1). The Carlsberg laboratory had close links to France. Quite likely he had in mind Claude Bernard and his book on the "the phenomena common to life" (Bernard, 1878–1879).

To start the experiment Johannsen bought 8 kg of Princess Beans from the general commercial harvest of 1900. The beans had an average weight of 495 mg determined by weighing 5000 beans chosen at random. In the spring of 1901, he sowed 287 individually weighed beans. 100 were chosen with weight as as close as possible to the mean of the whole batch, 25 from the very smallest and 25 from the largest, and the rest for verious reasons. The overall harvest from the 207 plants that grew up to produce seed nicely illustrate the regression between parents and offspring in accordance with the law of ancestral heredity. Closer analysis of the first year harvest suggested limitations to this law, however. For the plants from the smallest and the largest beans the seed were weighed separately. The overall harvest from the small beans also fitted well with the theoretically calculted numbers for regression. This was, however not the case for the offsping from the largest mother beans. An irregular distribution indicated the presence of lines with different heredity. (Johannsen, 1903: 20.) His extensive breeding experiments with barley in the late 1890s using Vilmorin's pedigree method had suggested the presence of family lines with hereditary differences. (Müller-Wille, 2018) With the Princess beans Johannsen now had an excellent experimental system to pursue this question and to test the range of biometric regression, by selection within individual pure lines in the next season (Johannsen 1903: 20).

The second season of Johannsen's experiment was directly aimed at the dispute between the biometricians and de Vries. In a brief review of the first volume of de Vries' *Mutationstheorie* (1901) Johannsen had emphasized De Vries' distinction between "statistical variation" and "mutation" as a promising theoretical advance (Johannsen, 1900–1901). Weldon on the other hand published a highly critical review of the same book. He argued that de Vries' selection experiments on maize confirmed the biometric view rather than his own theory of mutations. Weldon put the burden of proof squarely on his opponents: "A clear proof that Professor Pearson's view of the facts of regression is wrong … is absolutely essential …" (Weldon, 1902: 369). Johannsen picked up the challenge (Johannsen 1903: 8–9).

In the spring of 1902 Johannsen sowed separately seed from 19 of his pure lines. In the autumn each plant was harvested separately, and all seed weighed individually. A specially constructed scales was used to efficiently weigh the total of 5494 beans. The striking result was that within each pure line he found no statistically significant difference in average weight of offspring from beans of quite different weights. This apparently complete absence of hereditary effect from selection surprised Johannsen himself.

The result was strong evidence for discontinuous change in heredity. Galton's law of regression had to be abandoned as a general law of heredity. It was valid for populations consisting of many pure lines, but it was contradicted by the behaviour of the individual pure lines. The result was, as Johannsen said in his playful manner, "at the same time a full confirmation and a complete dissolution of Galton's well-known law of regression" (Johannsen 1903: 57).

Johannsen sent his 1903 monograph to some of the leading experts. De Vries replied approvingly. Galton cautiously recognized the potential importance but found the method problematic because inbreeding was known to decrease fertility. Weldon and Pearson were not at all pleased. They rejected Johannsen's work as a case of amateurish applied statistics. Using their own statistical tools they concluded that Johannsen's results confirmed their own theory rather than his. (Weldon and Pearson 1903; Pearson 1903) To Johannsen's satisfaction another biometrician and Pearson collaborator, George Udny Yule (1871–1951), pointed out that Pearson and Weldon had misunderstood Johannsen's biological analysis. Though his statistics was not very advanced it was adequate to the occasion (Yule 1904). Pearson arrogantly brushed off Yule's correction (Pearson 1904) and never acknowledged his mistake. Many historians of genetics have taken the Pearson-Weldon criticism seriously (Provine 1971: 96–100; Sapp 2003: 144–146), but by now it is broadly accepted that it was indeed based on a misunderstanding (Bulmer 2003: 218–224; Stolzfus and Cable, 2014: 530; Shan 2021), as Yule had immediately pointed out.

Johannsen followed up with continued selection in his pure lines. The results up to 1907 fully confirmed their stability (Johannsen 1909: 145–157). He seriously believed that an exact science of heredity needed advanced statistical method, not least the biometric methods invented by Galton and further developed by Pearson. It was precisely by using such method that he was able to dissolve the law of ancestral heredity make important advance. He obviously was hurt by the dismissive arrogance of Pearson (Roll-Hansen 1989: 315–316). At the Third International Conference on Genetics in London in 1906 Johannsen started by sharply attacking Karl Pearson for his neglect of biological premises (Johannsen, 1907). This resentment continued to fire his polemics against biometric theory of heredity in the lecture to American breeders and biologists in December 1910 (Johannsen 1911). He never tired of warning against excessive formal mathematical treatment. All three editions of *Elemente* (1909, 1913, 1927) carried the same warning on page two: "We must do genetics *with* mathematics, not *as* mathematics!"[1]

To properly understand Johannsen's genotype theory and the meaning of his terms phenotype, gene, and genotype it is helpful to look more closely at the way he introduced them in the *Elemente* of 1909. The book starts with a one hundred pages course in applied statistics. He describes at length methods describing and analysing a set of data based on the normal binomial curve of variation, as introduced by the Belgian anthropologist Adolphe Quetelet (1796–1874). In chapter seven Johannsen introduced Galton's application to heredity of stature in human populations. He showed how Galton obtained his law of regression, that statistically the stature of children is intermediate between that of their parents and the mean of the whole population. In chapter eight he discussed Quetelet's statistical concept of type. Johannsen showed, using biological examples, how a population that perfectly fits the statistical concept of a unitary type, can nevertheless be composed of subpopulations each with a different type. He introduces the word phenotype for the statistically identified "appearance type." A phenotype is a real entity that can correspond to a certain biological type, but it need not do so and usually does not, he wrote (Johannsen 1909: 123).

Before we go on to testing Galton's law of regression, Johannsen continued in 1909, we must explain some basic biological premises. The difference between species, dog and cat, rose and lily, depends on corresponding differences in gametes and zygote. This "something" has been given different names. Darwin for instance called it "pangene," wrote Johannsen. I propose the convenient abbreviation "gene." When we think about a property which is conditioned by a certain gene we can for conveience talk about "the gene of the property" instead of lengthy phrases like "the gene which conditions the property." At present no conception of the "gene" is sufficiently proven. But this is not needed for progress in the study of inheritance. It is sufficient that it be securely established that such a kind of "genes" exist. We will find that for some traits the genes can easily be isolated and for others not. Johannsen eagerly accepted the results of Bateson's Mendelian research programme on hybridization at the same time as he pointed to its limits. Only this is certain: "*The individual gamete contains special, separable 'genes' for different properties*."[2] This is one of the most important results achieved through the hybridization experiments of Mendel, he claimed (Johannsen 1909:124–125.).

It is significant that the terms "phenotype", "gene," and "genotype"[3] were introduced without explicit definitions embedded in a long argument. First came "phenotype" and "gene." Only seven pages later followed "genotypic" and "genotype." (Johannsen 1909: 124-125, 130.) He made it clear from the start that the new terms belonged to a tradition of particulate theories, including Darwin's gemmules, the pangenes of de Vries, and Weismann's determinants. The convenient abbreviation "gene" could readily be combined with other words and fitted into different contexts without too much baggage of old associations. The new word simply expressed the certain fact ("sichergestellte Tatsache") that at least many properties of an organism are conditioned ("bedingt") by independent ("selbständige") factors present in the gametes. Beyond this his "gene" was not committed to any specific hypothesis. (Johannsen, 1909: 124–125). The implication is that Johannsen was solidly planted in a tradition which takes some degree of preformation to be an obvious premise.

The word "genotype" was first introduced in the adjectival form as "genotypic difference" between phenotypes (p. 126). Only in the following chapter was the substantive form "genotype" used, with the reservation that only the concept of "genotypic difference" is experimentally operational (Johannsen, 1909; p. 130). The great achievement of Johannsen's distinction was practical no less than theoretical. He invented an experimental method for clear operational distinction between genotype and phenotype.

6. Interpreting Johannsen's genotype theory

Baverstock refers to one single paper by Johannsen. It is one of very few papers that he published in English and is much used in historical analyses (Johannsen 1911). It was republished in the *International Journal of Epidemiology* in 2014, accompanied by commentaries from three historians of genetics (Sapp 2014; Falk 2014, Roll-Hansen, 2014a). Baverstock quotes Roll-Hansen's claim that the genotype rather than the gene, was the basic concept of Johannsen's theory: "The stability of genotype is what makes a science of heredity possible. The concept of gene is derivative. It represents an experimentally identifiable difference between genotypes." (Roll-Hansen, 2014a: 1007; Baverstock 2021: 49.)

The comment by Roll-Hansen described how Johannsen theory grew out of botanical systematics and experience with plant breeding inspired by Galton and de Vries. Sapp presented Johannsen as a Mendelian and his genotype theory as a corollary to Mendelism without mentioning his careful from-the-inside critique of the biometric ancestral law of heredity. Falk gave an

[1] "Wir müssen die Erblichkeitlehre zwar *mit* Mathematik, nicht *als* Mathematik treiben."

[2] Emphasis in the original.

[3] The German terms of Johannsen are: Phaenotypus, Gen, and Genotypus.

insightful presentation of the development of Johannsen's genetic thinking. The comparison of his selection experiment to Mendel's hybridization was a weak point, however. Johannsen did not, like Mendel, "inbreed his bean plants for several generations until pure lines were obtained" (Falk 2014: 1004). To start his pure lines Johannsen simply picked individual beans from the lot he bought on the market. His lines existed in nature. They were not artificially prepared for the experiment.

Johannsen's 1911 paper was originally prepared for the annual meeting of the American Naturalist Society in December 1910. He had been invited to give the key-note lecture in a symposium on "The Study of Pure Lines of Genotypes." Poor health prevented his personal presence at the symposium, but the following winter he visited the US lecturing around the country for about five months. Several participants at the symposium presented important new experimental results that supported and extended Johannsen's theory: On multiple genes in explanation of continuous variation (East, 2011a,b), on the breeding of a cross-fertilizing plant (corn) (Shull, 1911), and on genetics and breeding of an animal (chicken) (Pearl1911). There was only one strong critic present. He demanded further testing of the bean selection results in the spirit of Karl Pearson (Harris 1911).

Johannsen's tour around the US in the winter of 1911–1912 was organized by George Henry Shull (1859–1944). He was just laying the theoretical basis for hybrid corn, the first big success of modern plant breeding. Edmund Wilson (1856–1939), the world's leading cytologist, invited Johannsen to give a seminar series at Columbia University. Here was a close colleague of Thomas Hunt Morgan (1866–1945) the founder of the chromosome theory. It is not clear if Morgan himself was present at the seminars, but some members of his Drosophila group certainly were. The acclaim that met Johannsen in America was proof of the success of his theory. It was the high point in his scientific career.

It has been a widespread view that Johannsen's genotype theory and pure line experiment is contrary to Darwinism. Ernst Mayr (1904–2005), leading evolutionary biologist and influential historian of biology, has been a main source of this view. Mayr emphasized the crucial importance of the distinction between genotype and phenotype for genetics and evolutionary studies, but he found that Johannsen was confused and had not really understood his own distinction. Johannsen thought he could "demonstrate the impotence of natural selection" (Mayr 1976: 332, 338), and he was "opposed to any role of selection in evolution" (Mayr 1982: 547). Will Provine's classical account of *The Origins of Theoretical Population Genetics* is similarly misleading. He accepted the Pearson-Weldon criticism of Johannsen's bean selection experiment and found it puzzling that the genetics community so quicky accepted his genotype theory (Provine 1971: 92–108). Frederick Churchill gave a penetrating analysis of the development of Johannsen's genotypr theory from the 1903 monograph to the 1926 edition of the *Elemente* (Churchill 1974). He found that Johannsen started with a statistical concept of type which gradually developed into a fully biological concept of genotype as a property of the individual organism. The weakness of Churchill's analysis is that he overlooked Johannsen's general understanding of biology and plant breeding as it was evident for instance in his textbooks of 1896 and 1900–1901. These texts show that his idea of biological type was a property of individual organisms from the start.[4]

There is also a persistent picture of Johannsen as the pioneer who ended up rejecting progress. While Bateson accepted the chromosome theory toward the end of his life Johannsen never did (e.g., Schwartz 2008: 229–230). An examination of his writings gives a different picture.

In the first edition of *Elemente* Johannsen took the cooperation of cytology and experimental research in heredity as the obvious long-term goal. He pointed to the work of Walter Sutton and Edmund Wilson as promising examples (Johannsen 1909: 481–482.) But he warned against the morphological speculations of Weismann. In the 1911 paper Johannsen appreciated the fruitfulness of Mendelian analysis in a cytological spirit, but indicated definite limits: "the entire organization may never be 'segregated' into genes!" He even suggested that "karyokinesis, synapsis, reduction" may be the consequence rather than causes of the segregation and recombination of the "genotypic constituents" (Johannsen 1911: 153). In the second edition of *Elemente*, published a few years later, the chromosome theory and its gene concept was discussed at length. In a long footnote he applauded Morgan's use of Janssen's "Chiasmatypie" model as a working hypothesis, but once more warned against taking Weissman type cytological speculations too literally. (Johannsen, 1913: 605).

By the early 1920s Johannsen recognized the chromosome theory as the spearhead of contemporary genetic research. (Roll-Hansen, 2014b) His popular Danish introduction to genetics presented fruit fly genetics and the chromosome theory as the state of the art (Johannsen 1923b). But he was still on guard against Weissman's unhealthy influence and admitted that he himself had been "somewhat possessed with the antiquated morphological spirit" of Galton, Mendel and Weissman when he introduced the gene (Johannsen, 1923b: 136). Johannsen considered the genes that had been mapped on chromosomes as entities of superficial imoporatnce, "by far the most comprehensive and most decisive parts of the whole genotype does not seem to be able to segregate into units." He believed in "a great central 'something' as yet not divisible into separate factors." (Johannsen, 1923a: 137). Johannsen thought that genetics so far had contributed little or nothing to the theory of evolution, to answering the great question of "the origin of species." But it had at least cleared the ground of some misconceptions of Darwin and Lamarck, though the germ line theory of August Weismann was a still an influential misconception. (Johannsen 1923c: 81–83, 102–103). Johannsen was by no means alone in his criticism of chromosome theory and Neo Darwinism. He shared a holistic and systems-oriented approach typical of German genetics and evolutionary studies in the 1920s and 1930s (Harwood 1993.).

The history of genetics still struggles with the misinterpretation of Johannsen's genotype theory as Mendelian and anti-Darwinian and his concept of genotype as statistical and abstract. Jan Sapp presented Johannsen's gene concept simply as Mendelian without any discussion of the bean selection experiment (Sapp 2014). Previously he claimed that "Johannsen's pure line experiments were faulty" but gave little if any argument (Sapp 2003: 145–146). Raphael Falk's thorough book on the history of genetic thinking contains very insightful discussions of Johannsen's work. He nevertheless held that Johannsen's original analysis of the genotype "was a statistical one" (Falk 2009: 68). More recently he wrote that Johannsen similarly to Mendel "inbred his bean plants for several generations until pure lines were obtained" (Falk 2014: 1004). A recent well informed and authoritative short history of genetics still relies on Churchill's statistical interpretation and says that Johannsen saw both "phenotype and genotype as abstract entities" not to be associated with particular cellular structures (Rheinberger and Müller-Wille 2016: 41–43). And an introduction to the philosophy of genetics presents Johannsen's concepts of gene and genotype as paradigmatic examples of classical chromosomal genetics (Griffiths and Stotz 2013).

Difficulties with the particulate DNA based gene has inspired a

[4] For a detailed criticism of Mayr and Churchill see Roll-Hansen (2009).

systems theoretical approach to embryology and evolution holding that different components of the cell are more or less on par in determining heredity (e.g., Oyama et al. 2001). Lenny Moss, biochemist and historian of genetics, has compared the concept that Johannsen introduced in 1909 with the present biochemical gene concept. He finds Johannsen's critique of the chromosome theory in the 1920s is well taken (Moss 2003) and relevant to present discussions. Moss' further speculations remind of Baverstock's view. Moss finds no place for a preformed hereditary entity in the cellular biochemical machinery. The DNA gene appears as no more than a "developmental resource" indeterminate with respect to phenotype. He suggests we should "rethink the meaning of genes and phenotypes from their roots up" and aim for a theory of evolution "based upon a phenotype-centered biology." (Moss 2008: 43, 52–53.)

7. Concluding remarks on Baverstock's challenge

I judge the preceding narrative to show that even if the gen concept is problematic the distinction between phenotype and genotype is scientifically well founded. As well founded as we can demand in empirical science. As I have indicated there are nevertheless reasons to take Baverstock's theoretical challenge seriously. How adequate is our contemporary understanding of the genotype/phenotype distinction?

According to the entry on "The Genotype/Phenotype Distinction" in *Stanford Encyclopedia of Philosophy* the predominant current meaning of genotype is "some relevant part of DNA passed on to the organism by its parents," while phenotype is "(t)he physical and behavioral traits of the organism" (Taylor and Lewontin 2017).[5] It is worth noting that genotype is presented first and that phenotype is defined in a way that includes the genotype. This indicates problems in drawing a clear distinction. Is there some kind of category mistake or incommensurability between biology and physical science involved? Like Baverstock this entry refers extensively to Johannsen's 1911 paper without closer examination of his reasoning in the selection experiment or how he introduced the genotype theory in his 1909 treatise. The conceptual tension in this authoritative article indicates that the Neo-Darwinism of the so-called Modern Evolutionary Synthesis is ripe for critical revision.

Baverstock's sketch of an alternative is unconvincing, however. One problem is his ambiguous use of the word "phenotype." Baverstock refers to "cellular phenotype" without discussing the radical difference from Johannsen's phenotype, which applied to multicellular higher organisms with cellular differentiation as a main feature. His characterization of "cellular phenotype" as "a process rather than a 'thing'," and as "an emergent quasi-stable state of a complex dissipative system" is, as he says, quite different from the usual understanding of phenotypes. He needs to get beyond this rather abstract physical characterization to a concrete biological description to produce an effective biological argument. Baverstock correctly refers to Johannsen's insistence on the gametes as carriers of the genotype. But his claim that "(t)he genotypes and the phenotypes of both parents are present in the zygote" (Baverstock 2021: 53–54) indicates a conflation of the two concepts.

Baverstock appears to have left behind both Johannsen's original understanding of genotype and the modern understanding based on DNA. He turns instead to a radical form of epigenesis and overlooks the sophisticated integration of epigenetic and preformationist explanation in embryology that had been well established already a century ago. It is quite contrary to present biological thinking to hold that the development of the zygote into an adult "is *not* driven by transferred parental genes but rather, is properly seen in terms of self-organisation," and that the "zygote 'knows' what it will develop into quite independently of its genotype." (Baverstock 2021: 56.) No doubt reductionism can be a fruitful impulse to progress in biology. The history of genetics, embryology and evolution through the last couple of centuries is good evicence. But it is hard to see how such apparently inconsistent and paradoxical use of concepts from existing biological science can be of help. To paraphrase Wilhelm Johannsen: We must do genetics *with* physics, not *as* physics.

[5] The former entry (Lewontin 2011) stated a similar understanding: Phenotype is the "manifest physical properties" and the genome is the "set of physical DNA molecules."

Declaration of interest

The author declares that there is no conflict of interest.

Acknowledgment

I want to thank Staffan Müller-Wille for useful comments to this paper, and for stimulating discussions on Johannsen through many years.

References

Baverstock, Keith, 2021. The gene: an appraisal. Prog. Biophys. Mol. Biol. 164, 46–62.
Bernard, Claude, 1878. Lecons sur les phénomènes de la vie communes aux animaux et aux végétaux. Vol I Paris 1978. Vol II Paris 1979.
Beyler, Richard H., 2011. Exhuming the three-man paper: target-theoretical research in the 1930s and 1940s. In: Sloan and Fogel 2011, pp. 99–142.
Bulmer, Michael, 2003. Francis Galton. In: Pioneer of Heredity and Biometry. The Johns Hopkins University Press, Baltimore and London.
Churchill, Frederick B., 1974. Wilhelm Johannsen and the genotype concept. J. Hist. Biol. 7, 30.
Dagognet, Francois, 1967. Méthodes et doctrines dans l'oevre de Pasteur. Presses Universitaires de France, Paris.
East, Edward Murray, 2011a. A Mendelian Interpretation of Variation that is apparently continuous. Am. Nat. 44, 65–82.
East, Edward Murray, 2011b. The genotype hypothesis and hybridization. Am. Nat. 45, 160–174.
Falk, Raphael, 2009. Genetic Analysis. A History of Genetic Thinking. Cambridge University Press.
Falk, Raphael, 2014. Commentary: a century of Mendelism; on Johannsen's genotype conception. Int. J. Epidemiol. 43 (4), 1002–1007.
Geison, Gerald L., Secord, James A., 1988. Pasteur and the process of discovery: the case of Optical Isomerism." *Isis* 79: 6-36. Weimar culture and biological theory: a study of Richrd Woltereck (1877-1944). Hist. Sci. 34, 347–377.
Griffiths, Paul, Stotz, Karola, 2013. Genetics and Philosophy. An Introduction. Cambridge University press.
Haldane, J.S., 1923. Mechanism, Life and Personality. Dutton, New York.
Harris, J. Arthur, 1911. The biometric proof of the pure line theory. Am. Nat. 45, 346–363.
Harwood, Jon, 1993. Styles of Scientific Thought. The German Genetics Community 1900-1933. The University of Chicago Press, Chicago and London.
Johannsen, Wilhelm, 1896. *Om Arvelighed og Variabilitet*. København: Det Schubo-teske Forlag.
Johannsen, Wilhelm, 1900-1901. Foreløpig anmeldelse af Hugo de Vries nye Værk: 'Die Mutathionstheorie'. *Biologisk Selskabs Forhandlinger*, pp. 7–8, 1900-1901.
Johannsen, Wilhelm, 1903. Erblichkeit in Populationen und in reinen Linien. Ein Beitrag zur Beleuchtung schwebender Selektionsfragen. Verlag von Gustav Fischer, Jena.
Johannsen, Wilhelm, 1913, 1909, 1913, 1926. In: Elemente der exakten Erblichkeitslehre, Second substantially expanded edition. Verlag von Gustav Fischer, Jena. Third revised and expanded edition 1926.
Johannsen, Wilhelm, 1907. Does Hybridization Incfrease Fluctuating Variability? Report of the Third International Conference 1906 on Genetics.
Johannsen, Wilhelm, 1911. The genotype conception of heredity. Am. Nat. 45, 129–159.
Johannsen, Wilhelm, 1923a. Some remarks about units in heredity. Hereditas 4, 133–141.
Johannsen, Wilhelm, 1923b. Arvelighed I Historisk Og Eksperimentel Belysning. En Udsigt over Arvelighedsforskningens Vigtigste Resultater, fourth ed. Gyldendalske Boghandel, Copenhagen.
Johannsen, Wilhelm, 1923c. *Hundert Jahre Vererbungsforschung.» Verhandlungen der*

Gesellschaft Deutscher Naturforscher und Ärtzte. 87. Versammlung zu Leipzig 17. Vogel, Leipzig, pp. 70–104. *Bis 24 September 1922*.
Kant, Immanuel, 1924/1799. Kritik der Urteilskraft. Verlag von Felix Meiner, Hamburg.
Lenoir, Timothy, 1982. The Strategy of Life: Teleology and Mechanics in Nineteenth Century German Biology. University of Chicago Press, Chicago.
Lewontin, Richard, 2011. The genotype/phenotype distinction. Stanf. Encycl. Philos. 2011, 27.01.22.
Mayr, Ernst, 1976. Evolution and the Diversity of Life. Selected Essays. Harvard University Press, Cambridge Mass. And London.
Mayr, Ernst, 1982. The Growth of Biological Thought. Diversity, Evolution, and Inheritance. Harvard University Press, Cambridge Mass. and London.
McKaughan, Daniel J., 2011. Was Delbrück a reductionist. In: Sloan and Fogel, pp. 179–210, 2011.
Mayenschein, Jane, 2005. Epigenesis and preformationism. Stanf. Encycl. Philos. (Accessed 19 December 2021).
Moss, Lenny, 2003. What Genes Can't Do. The MIT Press, Cambridge Mass, London.
Moss, Lenny, 2008. The meanings of the gene and future of the phenotype. Genom. Soc. Pol. 4 (1), 38–57.
Muller, Herman J., 1926. The gene as the basis of life." Reprinted. In: Muller, H.J. (Ed.), Studies in Genetics. The Selected Papers of Herman Joseph Muller. Indiana University Press, Bloomington, pp. 188–204, 1962.
Müller-Wille, Staffan, 2018. Making and Unmaking Populations. Historical Studies in the Natural Sciences 48 (5), 604–615.
Noble, Denis, 2013. Physiology is rocking the foundations of evolutionary biology. Exp. Physiol. 98, 1235–1243.
Noble, Denis, 2021. The illusions of the modern Synthesis. Biosemiotics 14, 5–24.
Oyama, Susan, Griffiths, Paul E., Gray, Russell D. (Eds.), 2001. Cycles of Contingency: Developmental Systems and Evolution. The MIT Press, Cambridge MA.
Pearl, Raymond, 1911. Inheritance of fecundity in the domestic Fowl. Am. Nat. 45, 321–345.
Pearson (anonymously), Karl, 1903. Prof. Johannsen on heredity. Nature 69 (17 December), 149–150.
Pearson, Karl, 1904. Reply to Yule, anonymous, as "Reviewer". Nature 69 (7 January 1904), 224.
Provine, William B., 1971. The Origins of Theoretical Population Genetics. The University of Chicago Press, Chicago And London.
Rheinberger, Hans-Jörg, Müller-Wille, Staffan, 2015. Gene. Stanf. Encycl. Philos. (Accessed 3 January 2021).
Rheinberger, Hans-Jörg, Müller-Wille, Staffan, 2016. The Gene. From Genetics to Postgenomics. The University of Chicago Press, Chicago and London.
Roll-Hansen, Nils, 1978. The genotype theory of Wilhelm Johannsen and its relation to plant breeding and the study of evolution. Centaurus 22, 201–235.
Roll-Hansen, Nils, 1979. Reductionism in biological research: Reflections on some historical case studies in experimental biology. In: Jan, Bärmark (Ed.), *Perspectives in Metascience*. Göteborg: Kungliga Vetenskaps- Och Vitterhets-Samhället, pp. 158–172.
Roll-Hansen, Nils, 1984. E.S. Russell and J.H. Woodger: the Failure of two twentieth-century opponents of mechanistic biology. J. Hist. Biol. 17 (3), 399–428.
Roll-Hansen, Nils, 1986. Svalöf and the origins of classical genetics. In: Olsson, G. (Ed.), Svalöf 1886-1986. Research and Results in Plant Breeding. LTs förlag, Stockholm.
Roll-Hansen, Nils, 1989. The crucial experiment of Wilhelm Johannsen. Biol. Philos. 3, 303–329.
Roll-Hansen, Nils, 2000. The application of complementarity to biology: from Niels Bohr to Max Delbrück. HSPS 30 (2), 417–442.
Roll-Hansen, Nils, 2008. Pasteur, Louis. In: New Dictionary of Scientific Biography, 6. Thomson & Gale (Charles Scribner), pp. 21–30.
Roll-Hansen, Nils, 2009. Sources of Wilhelm Johannsen's genotype theory. J. Hist. Biol. 42, 457–493.
Roll-Hansen, Nils, 2011. Niels Bohr and Max Delbrück: balancing autonomy and reductionism in biology. In: Sloane and Fogel 2011, pp. 145–178.
Roll-Hansen, Nils, 2014a. Commentary: Wilhelm Johannsen and the problem of heredity at the turn of the 19th century. Int. J. Epidemiol. 43, 1007–1013.
Roll-Hansen, Nils, 2014b. Johannsen's criticism of the chromosome theory. Perspect. Biol. Med. 57 (1), 40–56.
Roll-Hansen, Nils, 2018. Revisiting the Pouchet-Pasteur controversy over spontaneous generation: understanding experimental method. Hist. Philos. Life Sci. 40, 68. Published online 01 November 2018.).
Russell, E.S., 1930. The Interpretation of Development and Heredity: A Study in Biological Method. The Clarendon Press, Oxford.
Sapp, Jan, 2003. Genesis. The Evolution of Biology. Oxford University Press, Oxford.
Sapp, Jan, 2014. Commentary: genotype does not determine phenotype. Int. J. Epidemiol. 43, 1000–1002.
Schwartz, James, 2008. In: Pursuit of the Gene. From Darwin to DNA. Harvard University Press, Cambridge, Mass, and London.
Shan, Yafeng, 2021. Beyond Mendelism and biometry. Studies in History and Philosophy of Science 89, 155–163.
Shull, George Harrison, 1911. The genotypes of maize. Am. Nat. 45, 234–252.
Sloan, Phillip R., 2011. Biophysics in berlin: the Delbrück Club. In: Sloan and Fogel 2011, pp. 61–98.
Sloane, Phillip R., Fogel, Brandon (Eds.), 2011. Creating a Physical Biology. The three-Man Paper and Early Molecular Biology. Chicago University Press, Chicago and London.
Sloane, Phillip R., Fogel, Brandon, 2011. Introduction. In: Sloan, Fogel (Eds.), Creating a Physical Biology. Chicago University Press.
Stolzfus, Arlin, Cable, Kele, 2014. Mendelian-mutationism: the Forgotten evolutionary Synthesis. J. Hist. Biol. 47, 501–506.
Taylor, Peter, Lewontin, Richard, 2017. The genotype/phenotype distinction. Stanf. Encycl. Philos. Published 6. June 2017 (Accessed 8 June 2017).
Timoféef-Ressovsky, Nikolai. Karl G. Zimmer, Delbrück, Max, 1935. Über die Natur der Genmutation und der Genstruktur. Nachrichten von der Gesellschft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse, Fachgruppe VI: Biologie 1, 189–245.
Warming, Eugenius, Johannsen, Wilhelm, 1900-1. *Den Almindelige Botanik* (General Botany). Det nordiske Forlag, Copenhagen.
Weldon, W.F.R., 1902. Professor de Vries on the Origin of Species. Biometrika 365–374.
Weldon, W.F.R., Pearson, Karl, 1903. Professor Johannsen's experiments in heredity: a review. Biometrika 2, 499–503.
Woodger, Joseph H., 1929. Biological Principles: A Critical Study. London: K. Paul, Trench, Trubner & Co. (Reprinted New York: Humanities Press, 1967 with a New Introduction by Woodger.).
Yule, George Udny, 1904. Prof. Johannsen on heredity. Nature 69 (7 January 1904), 223–224.

Nils Roll-Hansen
Department of Philosophy, Classics, History of Art and Ideas, University of Oslo, Oslo, Norway
E-mail address: nils.roll-hansen@ifikk.uio.no.

17 September 2021
Available online 01 April 2022

Progress in Biophysics and Molecular Biology 177 (2023) 202–206

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

From information to physics to biology

Giuseppe Longo

Centre Cavaillès, CNRS et Ens, Paris, France

ARTICLE INFO

Keywords:
Information
Physics
Biology
Central dogma

ABSTRACT

Commentary to "The gene: An appraisal" by Keith Baverstock.
PBMB, Volume 164, September 2021, Pages 46–62.
Note: this short and informal commentary constructively criticizes the very interesting approach in the paper by a brief survey of the work that a few of us develop since several years. I will first recall the very pertinent critique of the Modern Synthesis and the genocentric approach presented in the paper, then suggest a methodological (and theoretical) critique of the approach by K. Baverstock and hint to alternatives paths that are compatible, but "extend" the physics for biology presented by the author. The purposes and the space allowed force a limited number of references and technical details. These may be found in the references contained in the few papers quoted below that are not the most nor the only representative contributions to the that work, but are inserted as a source of references or as synthetic presentations of our views.

1. DNA, the organism and evolution in an "informational" context

Baverstock's article conducts first a careful examination of the conceptual and theoretical errors that have accompanied the various projects aimed at the knowledge of DNA, human in particular. It recalls the persistence and for a long time, until year 2000, of the myths of the bijective determination by the genomic coding of polypeptides (primary sequences that, if sufficiently long and folded in three dimensions, form the active proteins in cells). So, the top managers of the Human Genome Project, until the eve of the announcement of the decoding of the human genome, estimated at 80,000 or more the human "genes", promising at the same time miracles of knowledge and therapeutics, once the chemical structure of DNA was going to be known (Collins 1999), see also (Liang et al., 2000) for further estimates at the eve of "decoding". The hype on the potentialities of this new, and in fact very relevant knowledge, did not disappear once it was shown that human "genes" are in fact 25,000, indeed less …. What do they "determine" then, what sequences they "program" and, from there, what phenotypes? With ideological and non-scientific arrogance, too many continued to say that humanity had at last "decoded" the "book of life written by God" and that we could thus understand all or almost all of the biological dynamics, and … Cure cancer and definitively, within 15 years (!), wipe out all monogenetic diseases … Diagnose, immediately, and cure, very soon, (almost) all human diseases (see (Longo 2018) for references). Unfortunately, God seems to have a strong propensity to write books with overlapping words (or genes), to dynamically "transpose" fragments of words from one place to another under the influence of changing contexts, as observed by B. McClintock, among many other peculiarities that make their reading very difficult. And even more so, their "editing".

Baverstock then rigorously explains why the analyses of even the later projects dedicated to genome wide associations (GWA) have had little clinical utility, particularly in view of the abundant pleiotropy and polymorphism that characterize gene-protein matches (i.e., they are of the "many-to-many" type, with peaks of 38,016 different peptides for the drosophila DSCAM gene and, conversely, the possibility that the same primary structure is produced by different DNA segments). So, the author very clearly explains the abuse that led to identify the "Mendelian genes" with a segment of DNA and how this has distorted research, clinical in particular, since very rarely diseases can be associated with molecular dynamics entirely intelligible in terms of gene expression, even as large networks of genome-wide associations (as claimed by the GWA project).

One would have also to wonder what was meant, and always has been, by "decoding" the genome. In general, if you have an "encoded message", to use a fashionable molecular terminology, as a sequence of signs, "decoding"" means its translation into a language and context that is *completely meaningful* to the intelligence agent or the (biological) structure using it. Now, the "meaning", in a cell, in an organism, of a segment of DNA should at least be the *function* of its chemical structure in said context – that is what the DNA *does*. Baverstock illustrates how

E-mail address: giuseppe.longo@ens.fr.

https://doi.org/10.1016/j.pbiomolbio.2022.12.003
Received 10 October 2021; Received in revised form 15 December 2022; Accepted 16 December 2022
Available online 23 December 2022
0079-6107/© 2022 Elsevier Ltd. All rights reserved.

far we are from this, that is, from associating, in general and not in a few special cases, "DNA sequence information into the functional information that informs the phenotype."

Unfortunately, but perhaps just to conform to the dominant fashion, Baverstock continues to use "informational" language. I have criticized in several articles, including, most recently, (Longo 2019), the consequences of a terminology borrowed from other sciences. With an implicit reference to the two main theories of elaboration (Turing) and transmission (Shannon) of digital data, one imports a Laplacian "structure of determination", as Turing and Schroedinger explicitly acknowledge. That is, one thinks, first and foremost, that determination implies predictability, since this is proper to both theories, and, again, Turing and Schrödinger say this with great lucidity – Shannon assumes it, since one of the purposes of his theory is exactly to reduce noise, that is to allow to predict the result of a well determined transmission (see (Longo 2019) for references). In short, the first, Turing, stresses the Laplacian nature of his 1936 Logic Machine to process information, at the origin of computer science, in 1952; Schrödinger observes, in 1944 "What is life?", that " … In calling the structure of the chromosome fibers a code-script we mean that the all penetrating mind, once conceived by Laplace, … Could tell from their structure whether the egg would develop, under suitable conditions, into a black cock or into a speckled hen … " (p. 7). Not only that, but the information theoretic language suggests the idea that dimensionality and materiality (hardware) do not matter. In fact, both digital theories allow to encode any finite dimension in one, a fundamental property of processing and transmission of discrete data types (digital data) that allows to process and transmit as sequences of digits even three-dimensional pixelated images, in perfect fidelity: in the mathematical discrete, there are no dimensional differences, everything is completely encoded in a linear sequence, an essential property of transmission and of computers' elaboration of data.

In addition, the independence of software from hardware makes computer science possible, as a general theory of processing sequences of signs, independently of the hardware. The same can be said of the transmission of information that can be done with smoke signals, frequency modulations, drums, sequences of bits …. The living is instead of a "radical materiality", that is, it is made only of these molecules, those that make up DNA, RNA and proteins, these membranes, and not others, in short, of this only "flesh" that makes us alive. At most, the non-specificity of macromolecular interactions makes it possible to replace a molecule (a hormone, for example) with another similar one, which may interact, with more or less probabilities, with the same cellular receptor (an endocrine disruptor, typically) – yet the non-specificity of macromolecular interactions should not be confused with the independence of information from the material support, it simply stresses the prevailing stochasticity of these interactions, that must be given in probabilities whose value depend on the context.

This idealistic obliteration of the dimensionality and materiality of the living, at all levels of organization, is one of the most serious distortions of knowledge that we owe to dominant molecular biology, along with the myth of the "exact stereospecificity" of molecular interactions, which has expelled randomness from molecular analyses, an ideological absurdity among the most serious, see (Paldi 2020) for a recent introduction. Again, informational language has contributed to this expulsion, since digital information is not processed nor transmitted at random: at most, probabilities, as a frequentist analysis, can help give syntactic relevance to sequences of signs, in Shannon's approach.

Let us therefore avoid to treat material flows and their gradients as "information", while stressing the contribution that this approach has given to the most deleterious of modern mechanicisms, that of the cell as a "Cartesian mechanism", today a digital computer, a "boolean algebra", of which the Central Dogma is at the heart, although accompanied by the feedbacks described by Crick and Monod. The latter, if expressed in a digital language as they are, are forms of *recursion*, therefore perfectly described by Turing's calculus or, even better, by Church's lambda-calculus, another wonderful theory of computability with its usual structure of Laplacian determination (Longo 2018, 2019), without bothering the differential theory of control in the continuum by Wiener. The computational language is perfectly closed on discrete data types and it is fantastic for its purposes, which need to by-pass dimensions, materiality … historicity (see below), that is all what matters in the analysis life.

In short, Baverstock falls into the linguistic trap that has done so much to structure the approach from which he wants to move away. Fortunately, this does not prevent him from telling very well a famous experience that provides a robust counter-example the Modern Synthesis in Evolutionary Theories (MS) – this "synthesis" has become the core informational approach to Evolution since the '60s (Gouyon et al., 2002). It is in fact very interesting the analysis conducted in the paper concerning the Long-term Evolution Experiment (LTEE), a laboratory experience that followed 66,000 generations of a population of *E. coli* bacteria. The choice of E. Coli is part of the experimental practices, as this bacterium is "the standard model" of many laboratories. The standardization of the model is a good practice, but has also other motivations. Very few unicellular organisms survive and reproduce in solitude, as *E. coli* can do: the vast majority of them in fact needs a diverse ecosystem, populated by thousands or millions or more of different (micro-)organisms. These organisms horizontally exchange fragments of DNA, RNA, proteome … Increasing at the same time both the production of diversity, through hybridization and contagions, and the forms of channeling and mutual control of evolutionary dynamics. The excellent arguments by Baverstock that use the LTEE to criticize the Modern Synthesis are a fortiori valid, if followed one by one, in reference to natural contexts, far from the artificial, though very interesting, evolution in the LTEE flasks: on the one hand the production of diversity and novelty is much greater, on the other hand the natural ecosystems constrains Evolution by excluding the incompatible in an even more effective way. Both phenomena, in nature, induce the change that matters most in an evolutionary dynamics, as "history": the change of the "space of the possible", or of the "phase space", we will return to this. In fact, this change, unlikely, but not impossible, in a monospecies' ecosystem, is the focus of a much more radical criticism, which we will mention, of the MS and its "mathematical" heart, the simplistic Fisher's law - sometimes even called "theorem", while it is only an abstract principle, expressed in mathematical writing, and that the LTEE is sufficient to demonstrate of no empirical foundations.

2. The history of physics and physicalism in biology

I will address a methodological critique, as an a priori of any discussion concerning an analysis dealing with the *historicity* of Evolution and the *specificity* of the pertinent objects, the organisms. A critique by "facts" may only follow a perspective from which one may read the facts. Indeed, the history of physics, I will hint, proposed a major diversity of perspertives into phenomena of the inert – it is a fortiori so when adressing issues pertaining the living state of matter.

The main richness of physics in the last four centuries, which has contributed to its hegemony among the sciences, is mainly due to its many "theoretical inventions". Faced with a new phenomenon or simply a change in scale or a new look at a common phenomenon, physics has been able to invent radically original theories. It was enough to look differently at a falling stone, a body in friction on an inclined plane, and a profoundly original theory of gravitation and movement (inertial) was born, which had though nothing to do with the movements of the planets, according to Galileo. Newton will *unify* the two very different scales, falling apples and moving planets, introducing a further theoretical revolution. Faced with new thermal machines, physicists then invented a new theory with different observables, such as entropy. Again, the *unity* with particles movements will be found in Boltzmann's revolution, thanks to a third theory, based on new principles and concepts, the ergodic principle, the thermodynamic integral; both hypotheses and notions are infinite constructions, that is new asymptotic

principles that allow to "understand together" particles in motion and macroscopic thermodynamic quantities like temperature. So Maxwell and Einstein will unify different fields (optics, electricity and magnetism, the first; gravitation and inertia, the second), further revolutions. Few measurements in microphysics, and there are those who dare to propose principles and theories incompatible with classical physics, based on the indetermination and non commutativity of measurement, to invent new spaces for quantum dynamics, the Hilbert spaces. The unity, in this great theoretical diversity, even a physical and mathematical incompatibility, is given by common principles of "conservation" (of energy, of momentum …), or rather by very general principles of "symmetry", as H. Weyl will explain (1929).

Many physicists, instead, turning to the living, locus of a rather original phenomenality, an entanglement of almost all physical scales, think that there is nothing else to say, that physical theories are enough to derive its fundamental properties. Suffice it instead to think that in a cell … There is a lot of water and that the hydrodynamics of incompressible fluids is not describable in terms of either thermodynamics or quantum physics, that is, in terms of known theories of "particles". Partial results give bridges between the equations of Boltzmann and Schrödinger and those of Navier-Stokes (see Chibbaro et al. book, reviewed in Longo (2016)), but we are far from a unification, which, as always, would presuppose the invention of a new theory. Yet, Bavestock tells us that the properties of the living are *derivable* from a thermodynamic principle.

The starting point is certainly valid: an organism is at *least* a thermodynamically open system, to which certainly apply principles of thermodynamics, Prigogine style - systems far from equilibrium (Nicolis and Prigogine, 1977). However, even if an animal falls with the acceleration proposed by Galileo, very little is *derived* of its biological properties from Galileo's theory and just some more from Prigogine's thermodynamics (although both are very important sciences and do affect organisms). Beginning with the properties of water in a cell, observe that these properties need, in addition to hydrodynamics, also Quantum Electrodynamics, see (del Giudice 2007), all theories intractable from thermodynamics and long from being unified. Invoking "emergence" does not help much: the properties of incompressible fluids do not emerge from those of quanta, nor from statistical physics, just as Newton's laws or the relativistic field do not emerge from the properties of falling stones nor from the quantum field, respectively. To understand them, a new theory had or must be invented, which correlates these theories, or, better, unifies them - the reduction to an underlying level from which the "laws" of the one above would emerge does not exist in physics and the best approximation of such a practice, statistical physics, requires, as we said, asymptotic principles, that is, a new theoretical invention. In other words, when we pass from the theories of particles, be it statistical physics or quanta, to the hydrodynamics of incompressible fluids, we change theory by a conceptual transition to different pertinent observables that requires a new theoretical frame. The unity between phenomenal levels or scales, the only way we know to move among them, as between apples and planets, is to be invented - and much is being done. The notion of "emergence" masks an impossible reduction, and hides the need for a new unified frame, making the different phenomena and their physical transition intelligible.

Physics, or better, according to Baverstock, only one of its theories in the very rich theoretical scattering of often incompatible theories, should instead allow to derive the properties of the living on the basis of a single, in itself very interesting, principle of optimality, the principle of "least action".

3. Changing phase spaces, or the historicity of the living

Before further criticizing the thermodynamic perspective in biology, I would first like to recognize in its promoters important allies in the battle against the vision of the living as a "Cartesian mechanism", driven by the genetic program, a molecular machine for processing and transmitting information, fashionable distortion of biological knowledge. The understanding of the importance, in biology, of the thermodynamics of systems far from equilibrium is an important first step, an essential component of the intelligibility of life. But we must go further.

First, an organism is not a self-organizing system. It does not emerge *spontaneously* and *necessarily* under certain boundary conditions, like a hurricane or a flame. It is not a self-organizing system of flows, but it *uses* flows of energy and matter, starting from the most ancient cell capable of metabolism and reproduction. To say that this spontaneous emergence must have happened, once (!), at the origin of life, is like moving the problem of unification of statistical physics and hydrodynamics at the moment of the Big Bang or 7 s later and call it "emergence": it does not help much and it even prevents to give an autonomous theory that unifies or at least relates the two theories, as physics has been able to do in the many examples described above, a possible and preliminary way also to understand what happened in the 7 s after the Big Bang. So it is a good practice, closer to the method of physics, to give oneself first an adequate theory of the pertinent phenomenal level, the organism, as some attempt in many writings, see for example (Soto et al., 2016), and then try to unify, or better, to extend the pertinent physical theories, if possible, by this new theory (Longo 2020). Darwin followed the first step of this path, by a theoretical invention, and proposed a robust theory of Evolution, while avoiding to discuss the origin of life – which may require an interface with physical theories. He founded it on a "non-conservation" principle (heredity is "descent with modification"), while Hamilton was founding physics on conservation principles and deriving from them Maupertuis (optimality) principle for classical dynamics.

In a Darwinian context, time has a novel and crucial role, both in phylogenesis and ontogenesis. Every organism has a history: instead, hurricanes and flames, on Earth, are of the same type, they are thermodynamically and mathematically identical, since four billion years, they have no history and can be described in the same "phase space". The living has changed quite a bit in this time – changing or new pertinent observables and parameters must be part of the analysis. And this is a crucial point: thermodynamics offers us a theory of irreversible time, which is necessary to understand the living, well beyond the time of classical and relativistic theories (Longo 2021), but is insufficient to grasp the additional *dimension* of historical time. This further form of time, historicity, is given by the change in phase space, a change not addressed by existing physical-mathematical theories, we will return to this, and by the role of rare events in determining phylogenetic trajectories, a role unknown to physical theories, even those dealing with "large fluctuations" (Longo 2018a). I am aware that this language is foreign both to the biologists of the "genetic program" and to the best physicalist reductionism such as Baverstock's: we are a small community trying to advance in a theory *proper to* the organism and for us nothing is understandable in biology without an analysis of its historical, that is phylogenetic and ontogenetic timing. Main stream, genocentric, biology goes elsewhere, like the fantastic main stream astronomy, trigonometry and geometry of epicycles in the Arabic world in year 1000 (Longo and Mossio, 2020; Longo 2021a).

Second, let us return to the issue of "phase space". Kant beautifully framed the Newtonian practice in mathematical-physics by proposing an epistemic pre-existence of space and time to knowledge construction, as the locus of all possible trajectories, all already in potentia in space-time, or, also, he will see space and time as a condition of possibility of theorizing in physics – that is for writing equations and describing actual trajectories. The nineteenth century will extend this pre-condition of any theoretical description in physics to the "a priori" of the "space of phases", that is the choice of observables and parameters relevant to the theory intended – including space or time. A very productive theoretical boldness will lead physics to invent new spaces of phases for new theories, that is new observables, such as entropy, new pertinent parameters or to make variable the number of parameters (statistical physics), to propose spaces of infinite dimensions (Hilbert spaces in quantum

mechanics), but each theory will have, and in principle, a pre-given space of phases. The symmetries proper to the description of these spaces make them always mathematically describable in the finite and a priori, even Hilbert spaces.

This is no longer valid in a historical science, such as biology. In biology, the historicity of the processes is due precisely to the dynamics of the space of the possible, a phase space, which changes along with the co-construction of phenotypes and organisms, their niches and ecosystems: for example, the observable "the placenta" was not already there, albeit in potentia, before a retroviral infection les than 100 mln years ago (Lavialle et al., 2013). All forms of ex-aptation à la Gould propose new observables, if we reason at the relevant scale and biological observables, that of phenotypes and organisms (Gould 2002). Their reduction to the level of colliding molecules stops, we said, at least in front of the theoretical irreducibility of incompressible fluids to classical or quantum molecular dynamics. But there is much more. Biological relativity, (Noble et al., 2019), does not allow to privilege one causal level over others: the heart and the vascular system, in their example, are formed at a critical transition in which the interactions between different scales and levels of organization allow the formation of the organ and its function, at once. Indeed, the formation of an organ in embryogenesis, but even more so the establishment of an evolutionarily novel phenotype, feathers and, then, wings for flight in dinosaurs, requires an extended critical transition (Longo and Montévil, 2014), i.e., a mathematically dense cascade of changes in symmetries, which depends on the entire ecosystem context. For feathers and functional wings, it ranges from muscular and lung structure, to … Air density: the phenomenon depends on all these constraints of which the last one is a typical stable boundary condition, of physical type, the others are co-constituted at and by the evolutionary event and depend on the very formation of the possibility to fly, which, for vertebrates did not exist before – it is a new observable, depending on new observables and parameters, all of them. The new possible actually depends on the global interaction of all these constraints, some of which are due to the constitutive dynamics itself; mathematically, in my view, it depends on a *non-local variable*, that is on a parameter depending on the global network of interactions making the novelty possible, but not necessary.

This is what we would like to add to Noble's biological relativity, the non-locality of parameters or of causal dependence: at least one of the pertinent parameters that allows/govern the new observable (the heart in embriogenesis, wings in Evolution …) depends on the entire new global structure that did not exist before. As for Evolution, this is part of its historicity, which does not allow to give, a priori, a phase space. (By the way, it is possible that Cosmology, at least when dealing with the formation of novelty at the early stages after the Big Bang, such as the fundamental constants of physics, is facing a similar challenge, historicity, a major problem for physicists.)

To give a principle of optimum, instead, such as Maupertius, we know from Hamilton that it is necessary to give, a priori, a space with a partial order, in which we can speak of "extremal" values, minima or maxima, be they local or global. This is the pre-given phase space that the historicity of the living forbids. In our perspective, the analysis of the dynamics of this changing space is part of the theory of Evolution and, therefore, of the organism (similarly, no way to derive the values of the fundamental constants of physics from some "optimality principle", when these constants are not yet fixed – infinitely many values are possible, thus the "many Universes" hypothesis in Cosmology). Mathematically, some of us are trying to treat this aspect of historical dynamics in life sciences as "heterogenesis" (Sarti et al., 2019), well beyond the (very interesting) physical morphogenesis to which far from equilibrium thermodynamics belongs.

Let's put it differently, but still summarizing apodictically a lot of work: history, evolutionary history in particular, does not follow optimal trajectories, therefore specific, but possible or generic trajectories, thus not necessary ones, in spaces co-constituted with and by the trajectories themselves. This process produces specific objects, that is, historical, individuated organisms in the case of the living. The physical trajectories, instead, are specific, geodesics in the appropriate space of phases, a space for each relevant theory, of course, but always given a priori. Physical objects are then generic, definable by a finite and a-temporal list of properties. The electron is a solution of the Dirac equation, just as a weight is perfectly described by some Galilean properties, without history. In contrast, a mouse is only definable in phylogenetic terms, telling its evolutionary story (Montévil 2020). As an example of an application of a method from physics, but not a theory, I dare to recall here the work in (Bailly and Longo, 2009), where we proposed to fix a counting of some measurable features of phenotypes independently of their actual realization (number of foldings, fractal dimensions, e.g. in lungs, vascual systems …, tissue differentiations …). The aim was to model Gould's analysis of "increasing complexity" in Evolution as a-symmetric random diffusion of "complexity" over bio-mass (Gould 1996), a new notion that we called anti-entropy. The "method" is borrowed from physics, but anti-entropy is a new observable, proper to life, to be added to entropy/negentropy, which are well defined in thermodynamics; its analysis does not assume extremal principles (the existence of maxima or minima), it thus departs from this omnipresent tool in mathematical physics. The production of anti-entropy, typically, requires the production of entropy (it does not oppose to it, like negentropy) and it is related to historical changes in the phase space (organisms' phenotypes and their pertinent parameters), a non-sense in existing theories in physics.

Let's follow the method, if inspiring, not the already given theories of physics: let's propose a theory adequate to the phenomenality of living, then we may try to correlate it, possibly as a non-conservative extension to the relevant physical theories (Longo 2020) - not easy because these theories are not unified. Non-conservative means that the extension of physics by biology can also allow us to explain, for example, the physical-chemical properties of the huge networks of molecular interactions that take place only in a living cell. These networks, from the physical point of view, have almost zero probability to exist and are found only inside cells. In short, there are macromolecular, thus physical-chemical activities, made possible only by biological contexts. That is, there is no plausible physical explanation of the origin and maintenance of huge networks of chemical interactions without considering the cellular structure in which they take place. These networks contribute to continually regenerate the cell, its membranes and its other organs, which, in turn, make the networks possible, *enable* them. This is a form of "constrained process" that is eminently biological and quite different from the ones given within boundary conditions treated by physics, since the constraints are produced by the dynamics itself, including the interaction networks they enable. In my opinion, the best treatment to date of this "closure of constraints" may be found in (Montévil and Mossio, 2015). In that theoretical framework, the DNA itself is analyzed as an internal constraint of the cell, permanently reconstructed by the molecular activities that it contributes to structure and canalize, in the cell. Obviously, a change in this constraint, the DNA, will affect cellular and organismal dynamics, but does not "drive" them. No "book of life" but an amazingly important constraint, sitting within each cell, a core component of the process of protein formation, continually modified and repaired by the macromolecular dynamics within the cell – like vital tissues and organs, at their scale, are repaired in an organism. For example, double-strand breaks can be repaired by non-homologous or microhomological-mediated end joining, which may also introduce changes in the DNA. By this, the molecular networks in the cell, by their dimensions and materiality, their variations, affect the DNA itself, while pressures and torsions on the chromatine modify the opening sites and, thus, in addition to (de-)methylation, change what matters in these processes, the *function* of the DNA.

This role of the context in the functioning of the DNA, as a constraint, is outside the scope of the Central Dogma (CD) as well of its negation: it is not a matter of "information" going back from the proteins to the DNA, but of protein networks, of three-dimensional and material, physical-

chemical and biological contexts that modify the DNA or what DNA *does* in the cell. In this, I disagree with both Baverstock and Noble on their critique of the validity of the CD. The CD, in the language of information, *implicitly* says or implies that the ontogenetic information and its hereditary transmission is *completely* contained in the DNA: at most the RNA may modify this information content by a retroaction on the DNA, not the proteins nor the epigenetic context, of course. I insist: the "completeness" assumption is crucial, while hidden in the language of information, as the transmission (heredity) and the elaboration (development) *of information* define biological dynamics. Thus, the CD, as stated, in that language, is either insignificant, since the language of information (Turing, Shannon etc) is insignificant in these processes, as I claim, or it si true, since, clearly, there is no information sent back from proteins inside DNA. As René Thom writes Thom, 1991, the limit of truth is not falsehood, but insignificance – possibly by the construction of a language-frame that forces to miss what is significant. Thus, the language of information and programming, as a frame for the CD, has proposed a perspective that made insignificant for biology the many *material* processes that are essential components of how the DNA *works*, often referred to as "epigenetic", for example pressures and torsions on the chromatine, RNA – proteome's interactions and activities etc. Conversely, the CD itself is insignificant w.r. To that meaningful, material frame (including all epigenetic phenomena). By this, the CD has had a major impact on research: in a community using the language of information, it let too many disregard the physico-chemical (and biological) role of the proteome, the cell, the organism and the environment in determining biological dynamics, with serious ramifications for applied fields of biotechnology and genetic engineering in agro-food systems (GMOs are the direct children of the CD) and medicine, e.g. cancer research (Longo 2018).

Minimal references

(Longo's papers and the references therein can be found in https://www.di.ens.fr/users/longo/download.html)

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

These views have been constructed thanks to the collaborations with F. Bailly, P.A. Miquel, M. Montévil, M. Mossio, C. Sonnenschein, A. Soto, A. Vianelli.

References

Bailly, F., Longo, G., 2009. Biological organization and anti-entropy. J. Biol. Syst. 17 (1), 63–96.
Collins, F.S., 1999. Medical and societal consequences of the human genome project. N. Engl. J. Med. 341, 28–37.
Del Giudice, E., 2007. Old and new views on the structure of matter and the special case of living matter. J. Phys. Conf. 67, 012006.n.
Gould, S.J., 1996. Full House. Three Rivers Press, New York.
Gould, S.-J., 2002. The Structure of Evolutionary Theory. Harvard U. Press.
Gouyon, P.H., Henry, J.P., Arnoud, J., 2002. Gene Avatars, the Neo-Darwinian Theory of Evolution. Kluwer.
Liang, F., Holt, I., Pertea, G., et al., 2000. Index analysis of the human genome estimates approximately 120,000 genes. Nat. Genet. 25, 239–240. https://doi.org/10.1038/76126.
Lavialle, C., Cornelis, G., Dupressoir, A., Esnault, C., Heidmann, O., Vernochet, C., Heidmann, T., 2013. Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation. Phil Trans R Soc B 368, 20120507. https://doi.org/10.1098/rstb.2012.0507.
Longo, G., 2016. A review-essay on reductionism: some reasons for reading *"Reductionism, Emergence and Levels of Reality. The Importance of Being Borderline"*, a book by S. Chibbaro, L. Rondoni, A. Vulpiani. *In* Urbanomic,. https://www.urbanomic.com/document/on-the-borderline/London. May 8.
Longo, G., 2018. Information and causality: mathematical reflections on cancer biology. Organisms. Journal of Biological Sciences 2 (1).
Longo, G., 2018a. How future depends on past histories and rare events in systems of life, "how future depends on past histories and rare events in systems of life". Found. Sci. 23 (3), 443–474, 2018.
Longo, G., 2019. Information at the threshold of interpretation, science as human construction of sense. In: Bertolaso, M., Sterpetti, F. (Eds.), Will Science Remain Human? Springer, New York, pp. 67–100.
Longo, G., 2020. Naturalizing Physics. or, Embedding Physics in the Historicity and Materiality of the Living. *Deleuziana*, N. 11, Special Issue on "Differential Heterogenesis: Deleuze, Mathematics and the Creation of Forms".
Longo, G., 2021. Confusing biological rhythms and physical clocks. Today's ecological relevance of Bergson-Einstein debate on time. In: Campo, A., Gozzano, S. (Eds.), Einstein vs Bergson. An Enduring Quarrel of Time. De Gruyter.
Longo, G., 2021a. Programming evolution: a crack in science. *A Review of the book by Nobel winner, jennifer A. Doudna, and samuel H. Sternberg "A Crack in creation: gene Editing and the unthinkable Power to control evolution"* 2017, in. Organisms. J. Bio Sci. 5 (1).
Longo, G., Montévil, M., 2014. Perspectives on Organisms: Biological Time, Symmetries and Singularities. Springer, Dordrecht.
Longo, G., Mossio, M., 2020. Geocentrism vs genocentrism: theories without metaphors, metaphors without theories. Interdiscipl. Sci. Rev. 45, 380–405.
Montévil, M., 2020. Historicity at the heart of biology. Theor. Biosci. 141, 165–173. Springer Verlag.
Montévil, M., Mossio, M., 2015. Closure of constraints in biological organisation. J. Theor. Biol. 372, 179–191.
Nicolis, G., Prigogine, I., 1977. Self-organization in Non-equilibrium Systems. Wiley, New York.
Noble, R., Tasaki, K., Noble, P., Noble, D., 2019. Biological Relativity requires circular causality but not symmetry of causation: so, where, what and when are the boundaries? Front. Physiol. 10, 827.
Paldi, A., 2020. Stochastic or deterministic? That is the question. Organisms. J. Bio Sci. 4 (1).
Sarti, A., Citti, G., Piotrowski, D., 2019. Differential heterogenesis and the emergence of semiotic function. Semiotica 2019 (230).
Soto, A., Longo, G., Noble, D. (Eds.), 2016. *From the Century of the Genome to the Century of the Organism: New Theoretical Approaches*, a Special Issue of *Progress In Biophysics And Mol. Biology*, vol. 122. Elsevier, 1.
Thom, R., 1991. Predire N'est Pas Expliquer. Eshel, Paris.

Progress in Biophysics and Molecular Biology 178 (2023) 70–71

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Commentary on Baverstock; the gene; an appraisal

A B S T R A C T

I argue that Baverstock's demotion of the gene in favour of the cellular phenotype is still too limited, and that phenotypes must be considered at multiple irreducible levels. I emphasise process thinking and the significance of agency in living systems.

Baverstock identifies the 'four main elements of biology: inheritance, evolution, development, and morphogenesis; ' argues that 'the genetic community' has mistakenly adopted the gene as the functional basis for these processes, and proposes instead that it is the cellular phenotype that should be viewed as the fundamental unit of life (the brain, as he puts it - an unhelpful metaphor). It is the cellular phenotype, he claims, which possesses agency and hence is the driver of evolution, and it is on these two terms, phenotype and agency, that I focus.

Baverstock's quick trawl through one and a half centuries of genetic research plots the transformation of Mendel's hidden determinants into genes, initially as map references on chromosomes and, by the 1960s, strands of DNA whose nucleotide sequences constitute the information store on which life depends -Johannsen's genotype. In the genocentric view these 'master molecules' - the ideological resonances of the term are clear - control all four of Baverstock's 'main elements'. Morphogenesis and development are directed by sequential expression of DNA sequences; inheritance is reduced to its faithful copying into a new cell; and evolution is the process by which random mutations in DNA result in variations in the cellular phenotype and hence provide the differences in fitness on which natural selection can operate. The molecule itself is ascribed agency, memorably captured by Dawkins in his assertion that the organism is the 'selfish' gene's way of propagating itself. Such genocentrism (ultra-Darwinism) sees the cell/organism as passive, without independent agency, merely ground out from between the lower and upper millstones of gene and environment.

While Baverstock regards this as still the 'genetic community's' consensus view, the last half century of research has steadily chipped away at such a reductionist, unilinear trajectory, and not only because of the unexpected result of the Lenski experiment which he quotes. He is not, as it might appear from the paper, a lone heretical voice. To summarise:

- a single DNA sequence (AKA gene) may be used to code for many different polypeptide chains; (20,000 genes in the human genome, 100,000+ proteins), as revealed by the HGP.
- A phenotypic 'character' depends on the expression of multiple genes, but classical heritability estimates have failed to resolve the paradox of 'lost heritability' (GWAS)
- A mutation in a gene does not necessarily result in a change in the phenotype (e.g. Lenski)
- heritable phenotypic changes can occur without changes in DNA sequences (e.g. Jablonka and Lamb, 2014)

Hence specific DNA sequences are only loosely coupled to what Baverstock calls the cellular phenotype; they are necessary but not sufficient to define it.

So far, so good, and I believe that these conclusions would be uncontroversial amongst most geneticists and evolutionary biologists today (though as a neuroscientist myself I speak with some diffidence). So what of Baverstock's alternative, the cellular phenotype as a 'brain' containing the 'knowledge' and possessing the agency necessary for life? The problem, it seems to me, lies in the vague and slippery term phenotype, originally as he says, a Mendelian 'character' and later more precisely a gene product. Today the term embraces multiple levels, from a polypeptide chain to a multienzyme complex, an organelle or a cell, an organism or an ecosystem - Dawkins's 'extended phenotype'.

Baverstock's confinement of the term to the cellular level is at once too broad and too restricted; the 'main elements of biology' are expressed at several levels of complexity. Viruses – simultaneously acellular genotype and phenotype - have no cellular phenotype but can mutate and evolve under natural selection. In actively penetrating and transforming host cells into factories enabling the virus to replicate they could be said to be demonstrating agency – it has certainly felt like that during the current Covid pandemic. Unicellular organisms mutate and evolve and show agency in selecting favourable environments (e.g. food-rich) and avoiding harmful ones (e.g. too acidic or alkaline) and in doing so transform those environments (e.g. by secreting waste products). Development and morphogenesis pertain to multicellular organisms whose agency in choosing and acting upon their environment (including interacting with other life forms) is apparent. The textbook example is the dam constructed by beavers (niche construction) which limits and alters the direction of river flow, providing a fitter environment for many other species to flourish while making it less favourable for others. Thus entire ecosystems change and evolve. This is the extended evolutionary synthesis (Laland, 2017). In seeking to reduce all these levels to that of the cellular phenotype, Baverstock performs his own reductionist operation.

These levels of increasing complexity are not just epistemological constructs but are ontologically and irreducibly distinct, as spelled out by Joseph Needham in the 1930s (Needham, 1943). Lower level

https://doi.org/10.1016/j.pbiomolbio.2023.01.004
Received 7 October 2021; Received in revised form 6 January 2023; Accepted 9 January 2023
Available online 16 January 2023
0079-6107/© 2023 Elsevier Ltd. All rights reserved.

processes enable but do not determine higher level ones, whilst higher level systems constrain the freedom of lower level ones (Levins and Lewontin, 1987; Noble, 2016, Rose, 2005). The evolution of multicellularity meant that individual unicellular organisms gave up some of their freedoms, as in the case of mitochondria and chloroplasts, believed to be the descendants of once free-floating aerobic prokaryotes, engulphed by eukaryotes to become endosymbionts (Margulis). In a multicellular organism, the agency of any single cellular phenotype is constrained by its embeddedness. At the same time multicellularity demands the emergence of specialised cells, thus driving the evolution of novel cellular phenotypes.

At each of these levels of organisation of matter, there is a contrast between a phenotype as a 'thing,' as Baverstock citing Mendel emphasises, and, as he hints but does not expand on, as a 'process.' I believe this contrast is fundamental to how we are to understand living systems. Life is essentially dynamic; at every moment living systems at all levels are both *being* one thing and *becoming* another. Virtually every molecule in every cell, whether uni- or multi-cellular, is continually being broken down and resynthesised (protein half lives in the human vary hugely, ranging from minutes to days; only a very few, collagen being an example, can outlive the body in which they are located). The seeming unity of a living organism is a process unity of form which persists even as its components are replaced. Which is why the term homeostasis is so misleading. It should really be homeodynamics, for stasis means death. And over the life cycle the set points that homeo refers to change, sometimes subtly, sometimes -as in the drama of birth itself-sharply and speedily. Phenotypes are simultaneously thing and process; the value of reductionist approaches is that they uncover into the dynamic self-thingyness; the value of process thinking is that it reinserts the 'thing' into the dynamic self- organising complexity of the living world. Baverstock's de-emphasising genes in favour of cells, I suggest, fits well within this larger theoretical framework.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Jablonka, E., Lamb, M.J., 2014. In: Evolution in Four Dimensions. MIT Press.
Laland, K.N.L., 2017. Darwin's Unfinished Symphony. Princeton University Press.
Levins, R., Lewontin, R.C., 1987. The Dialectical Biologist. Harvard University Press.
Needham, J., 1943. Time, the Refreshing River. Allen and Unwin.
Noble, D., 2016. Dance to the Tune of Life: Biological Relativity. Cambridge University Press.
Rose, S, 2005. Lifelines: Life beyond the Gene. Vintage.

S.P.R. Rose[1]
School of Life, Health and Chemical Science, The Open University, Milton Keynes, MK7 6AA, UK
E-mail address: s.p.r.rose@open.ac.uk.

[1] the author declares no competing or financial interests.

RESPONSES TO COMMENTRIES

Responses to Commentaries on
The Gene: An Appraisal

by

Keith Baverstock

Progress in Biophysics and Molecular Biology 188 (2024) 31–42

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Responses to commentaries on "The gene: An appraisal"

Keith Baverstock

Department of Environmental and Biological Sciences, University of Eastern Finland, Kuopio Campus, Kuopio, Finland

ARTICLE INFO

Search terms:
2nd law of thermodynamics
Natural system
Genotype conception
Entropy

ABSTRACT

The central conclusions of "The Gene: An Appraisal" are that genetic variance does not underpin biological evolution, and, therefore, that genes are not Mendel's units of inheritance. In this response, I will address the criticisms I have received via commentaries on that paper by defending the following statements:

1. Epistasis does not explain the power-law fitness profile of the Long-Term Evolution Experiment (LTEE). The data from the evolution of *natural systems* displays the power-law form ubiquitously. Epistasis plays no role in evolution.
2. The common characteristics of living things (natural systems) are described by the principle of least action in de Maupertuis's original form, which is synonymous with the 2nd law of thermodynamics and Newton's 2nd law of motion in its complete form, i.e., $F = dp/dt$. Organisms strive to achieve free energy balance with their environments.
3. Based on an appraisal of the scientific environment between 1880 and 1911, I conclude that Johannsen's genotype conception was perhaps, the only option available to him.
4. The power-law fitness profile of the LTEE falsifies Fisher's Genetical Theory of Natural Selection, Johannsen's genotype conception, and the idea that 'Darwinian evolution' is an exception to the generic thermodynamic process of evolution in natural systems.
5. The use of the technique of genome-wide association to identify the causes and the likelihoods of inherited common diseases and behavioural traits is a 'wild goose chase' because genes are not the units of inheritance.

1. Introduction

> 'It is able to overthrow the order of things and reconceive the world time and time again'. On the nature of scientific thinking: Carlo Rovelli in *Anaximander and the Birth of Science*.

The Gene: An Appraisal, originally (Baverstock, 2021) subsequently reissued as a corrigendum (Baverstock, 2024), has two primary aims. The *first* is to put into realistic context what a gene is and what it does and does not do, and the *second* is to abstract a *framework* derived from the biological evidence accumulated over more than a hundred years of research, upon which a new theory of biology can be built. On the publication of (Baverstock, 2021), the editor of *Progress in Biophysics and Molecular Biology*, Denis Noble, asked me if I would let him issue invitations to researchers I had cited to comment on the paper; I agreed. I understand that 45 invitations were issued,[1] but only seven of the invitees accepted. My reason for agreeing to Denis's suggestion was to attract challenges to both the criticisms of the way the gene is regarded in genetics today and the proposals I and my collaborators have made for a framework of a 'new biology'. Critical commentaries were offered by Nils Roll-Hansen, a historian of genetics at the University of Oslo; Giuseppe Longo, a physicist and mathematician at CNRS in Paris and a member of a panel of researchers aiming to better define current biology in terms of physics; and Steven Rose, a neuroscientist, and author who was a researcher at the UK Medical Research Council and is now Emeritus Professor at the Open University in the UK. I will address these criticisms below; however, I stress that none of the invitees have challenged the central assertions of (Baverstock, 2024), namely that genes are not the units of inheritance, and that genetic variance does not underlie biological evolution.

I received four mainly supportive commentaries, for which I am most grateful:

Abbreviations: 2nd law, Second law of thermodynamics; GWA[S], Genome-wide association [studies]; IA model, Independent Attractor model.; LTEE, Long-term evolution experiment.; PGS, Polygenic score (sometimes termed polygenic risk score, or PRS).; RoE, Rules of engagement.

E-mail addresses: keith.baverstock@uef.fi, keith@kbaverstock.org.

[1] Invitees were issued with a corrected copy of the paper.

https://doi.org/10.1016/j.pbiomolbio.2024.02.002
Received 23 December 2023; Accepted 2 February 2024
Available online 14 February 2024
0079-6107/© 2024 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1) (Baluska and Reber, 2021): František Baluška's courageous, pioneering work on plant intelligence and consciousness inspired my thinking about the role of organismal agency. Therefore, I thank him for his contribution. I question one point though: I am sceptical of the view that DNA-free proto-cells have a memory based on the structure of their membranes. I think it is more likely that proteins can perform 'cognitive type functions' in the cellular cytoplasm, (completely independently of the DNA), by virtue of their intrinsically disordered state and their consequent folding and unfolding (Fonin et al., 2018), which converts free energy to information (Annila and Baverstock, 2014). This is, perhaps, a somewhat similar situation to where the phosphorylation of three proteins isolated from cyanobacteria are capable of reconstituting circadian rhythm *in vitro* (Nakajima et al., 2005), but without the involvement of a membrane. In our joint paper (Baluska et al., 2016) we say, 'Lipid bilayers, cellular membranes, and critical proteins emerge as the most probable primary targets of anesthetics'. On reflection, I think the target most probably is the proteins, due to their continuous folding and unfolding activity (Fonin et al., 2018).
2) (McKenna et al., 2022): H. F. Nihjout's early work on the role of DNA in development was a major influence in my questioning of the role of DNA in biology. In their commentary, McKenna and colleagues dispel any notion that DNA is exerting control over the phenotype, exposing, for example, Robert Plomin's nonsense notion of DNA as a 'fortune teller' of life's outcomes (Plomin, 2018).[2]
3) (Richardson, 2021): I view Ken Richardson's development of the concept of 'biogrammars' as closely parallel to my ideas for moving forward with a 'new biology'. His views on anticipatory systems are a vital aspect of the 'brain metaphor' for the cell phenotype that I did not touch on (Baverstock, 2024).
4) (De la Fuente, 2021): I. M. De la Fuente's first publication on the role of attractor states in metabolic systems (De La Fuente et al., 2008) was published only months after Mauno Rönkkö and I published our formalisation of an attractor state to represent the cellular phenotype (Baverstock and Ronkko, 2008). Their publication gave me confidence that I was moving in the right direction. In this paper, De la Fuente et al. go much further into the dissipative and self-organising nature of life. Surprisingly, this view is not more universally held, for example, by Longo and Roll-Hansen, who have expressed scepticism about the role of self-organisation in biology.

Besides these 'official' commentaries, I received a final rejection notice from Max Reuter, editor-in-chief of the Journal of Evolutionary Biology (JEB), in response to my submission in 2022. The rejection, supported by an 'informed and independent' member of the JEB Editorial Board, details a reason, which I will contest, why my analysis of the LTEE in Baverstock (2024) is flawed. Since my interpretation of the LTEE is central to my arguments, I will address this criticism first. However, at this juncture, I find it pertinent to mention that my interpretation has not attracted any comments from LTEE researchers, either in the form of a commentary or privately.

2. Response to the editor-in-chief of JEB, Max Reuter

I have been aware of the LTEE for some considerable time. I can tell from the notes I made at a 2010 symposium in Helsinki that I was aware of the work of Barrick et al. (Barrick et al., 2009), which reported on one of the 12 *independent* experiments that comprise the LTEE. I had written, 'The experimenters [on the LTEE project] are now faced with the uncomfortable fact that this unique experiment does not support the dogma that genetic change underlies evolutionary adaptation.' I am confident that at least one of the LTEE researchers, and likely a few other evolutionary biologists, were invited to comment on (Baverstock, 2024); the lack of any defence of mainstream Neo-Darwinian evolutionary biology is surprising, given that I am proposing that several decades of well-established dogma should be overturned, including, of course, the Modern Synthesis.[3] However, a commentary I submitted in 2022 in response to a paper in the JEB was rejected based on the following assessment made by the reviewer: 'I [also[4]] disagree that the Baverstock (2021) paper [(Baverstock, 2024)] presents information that casts doubt on the fact that mutations provide the substrate for evolution and adaptation. The results of Long-term Evolution Experiment (LTEE) from the Lenski lab do not require explanations that are outside the realm of known genetic and evolutionary processes. In fact, the Wiser et al. (2013) paper uses known genetic processes to explain the results of the LTEE trials.' These are the words of the 'Deciding Editor', appointed as a qualified and independent reviewer for my commentary by the Editor-in-Chief of JEB, Dr Max Reuter.

The above opinion is presumably that of an expert evolutionary biologist. The 'known genetic processes' referred to are 'clonal interference' and 'diminishing returns epistasis' as cited in Wiser et al. (2013). Clonal interference occurs when two or more *beneficial* mutations compete with one another for nutrients and thus delay the fixation of a mutation in the clone. Epistasis refers to 'interactions between genes/mutations', which serve to mask or enhance the effects of other genes/mutations, including features of the 'genetic background'. The term was coined in 1909 by William Bateson and has found considerable application in explaining results that do not adhere to Mendel's laws of inheritance. Typically, epistasis is invoked in the context of statistical/mathematical modelling rather than empirical evidence. This is the case for Wiser et al. (2013), who say they have generated theoretical 'mean-fitness trajectories that agree well with the experimental data'. However, constructing a *model* that yields similar results to an *experiment* is not the same thing as describing a *process* that has yielded those results. Almost 100 years after the term 'epistasis' was proposed, Jason Moore wrote in *Human Hereditary* in 2003, 'What is a proper method for detecting epistasis? The answer to this question is currently unknown ….' (Moore, 2003). I have not found any evidence that since this question was raised by Moore it has been answered. In other words, epistasis is a statistical/mathematical concept, not a known and understood genetic *process*: it is something that evolutionary biologists and geneticists believe might be happening without knowing how. Epistasis is therefore instrumentalism, not science: it has served as a convenient justification for narratives based on Johannsen's genotype conception and Mendelism.

Given the structure of chromatin, it is impossible to see how genes (segments of DNA wrapped in chromatin), or their mutations, can *physically* interact. The term 'gene interaction' must, therefore, be shorthand for 'gene product interaction', i.e., protein interaction. That is the process that in any case is deemed to yield the cellular phenotype. In the IA model it does so under the regulation of the 'rules of engagement' (Baverstock and Ronkko, 2008).

Let us look more closely at what clonal interference and epistasis might mean in the context of the LTEE, starting with epistasis. The term used by Wiser et al., is 'global diminishing returns epistasis'. The power law profile observed for the LTEE is one of diminishing returns in the context of both fitness (Wiser et al., 2013) and increasing cell size (Lenski and Travisano, 1994). However, the LTEE comprises 12 parallel and independent experiments, each acquiring different mutations

[2] For a current review of Plomin's book *Blueprint: How DNA makes us who we are*, see Joseph, J. (2022).

[3] More recently, Arto Annila and I published a concise summary of the argument developed in Baverstock (2024) that thermodynamics, rather than genetic variance, drives evolution on the Qeios website (https://www.qeios.com/read/NLISV5) on 7 September 2023. At the time of writing, 18 reviews have been submitted, none from leading mainstream geneticists and evolutionists.

[4] Refers to the Editor-in-Chief, Dr Max Reuter.

(Maddamsetti and Grant, 2020), and in all of them, epistasis (if it is having an effect) produces an *identical* effect on fitness (growth rate compared to that of the founding bacteria) but *not* on cell size, the profiles of which are *different* for each of the 12 experiments (Lenski and Travisano, 1994).[5] If mutations are responsible for both, why the difference? The literature on the effects of introducing different numbers of mutations into the genomes of bacteria is contradictory: in some cases, fitness decreases, in others it increases.

For clonal interference to influence fitness similarly in 12 independent experiments, beneficial mutations would have to be very frequent. In 1998, Gerrish and Lenski examined the fates of competing beneficial mutations in asexual organisms (Gerrish and Lenski, 1998). They concluded, based on evidence from the experimentally determined spontaneous mutation rate of *E. coli,* that one in about a million mutations would be beneficial. Thus, the chances of two beneficial mutations occurring in the same clone close in time are small, and that such an occurrence would occur in 12 independent experiments to produce the same effect on fitness is vanishingly small. Thus, the identical power law fitness profiles of the 12 experiments cannot be unequivocally explained by invoking clonal interference and epistasis. On the other hand, the increase in cell size can be easily explained. Individual, genetically pure *E. coli* bacteria can show different responses to chemotaxis (Frankel et al., 2014; Salek et al., 2019), possibly because of specific gene duplications (Bratlie et al., 2010). According to this explanation, cell size cannot decrease in relation to the founder bacteria. Having a distribution of cell sizes in a population may increase the resilience of the population but *does not alter* how rapidly the bacteria can dissipate the free energy in the nutrient. Instead, the thermodynamic principle of consuming free energy in the least time determines how well *E. coli* adapt to their environment. Thus, according to Annila, as a natural system, bacteria continually strive to improve their characteristics to attain balance with their surroundings (Sharma and Annila, 2007; Makela and Annila, 2010). The scale-free power law observed in the LTEE results from that thermodynamic principle. Therefore, the power law characterises natural processes, evolution being one of them. Annila notes, 'Evolution does not make a distinction between the living and the lifeless, the microscopic and the cosmic, or the simple and the complex, but all courses of events follow natural law instead of being the result of a random walk' (Annila, 2020). In terms of thermodynamics, microbes evolve in the *least time*, or by the most *efficient* locally available route. Because the environments in each experiment, including the nutrients, are identical, this is the same for all 12 of the independent LTEE experiments, meaning that they respond identically. Under this framework, fitness *must* increase compared to the founder population. This contrasts with the notion that evolution is driven by genetic variance. If that were the case, a) it would surely lead to fitness profiles varying from one independent experiment to another; and b) since mutations are believed to be more likely to be detrimemtal than beneficial to fitness by a million to one according to Gerrish and Lenski (1998), a decline in fitness should be more likely than an increase. Using a modified form of Fisher's law (to allow for evolution over a finite period instead of at an instant), Basener and Sanford (2018) showed that: a) reductions in fitness are a real possibility; and b) even under the most optimistic assumptions about the proportion of beneficial versus detrimental mutations, the fitness profiles of the LTEE cannot be reproduced. (See (Baverstock, 2024) for the full argument.)

Later results from the LTEE (Maddamsetti, 2021) explain the trend of increasing fitness. As *individual* gene products accumulate mutations, the resilience of the so-called *collective* protein interactome towards environmental perturbation is maintained. This result is in line with the results of Zitnik et al. (2019), who examined the resilience to environmental stresses of the protein interactomes of 1840 species, finding that the interactomes become more resilient with evolution as a result of changes in the network topology of the interactome (Zitnik et al., 2019).

To summarise the thermodynamic argument in the context of the LTEE results, as evolution proceeds, the dissipation of free energy (nutrient) leads to an increase in entropy and a reduction in internal free energy, which yields an increasingly more probable state, i.e., the process of natural selection (Sharma and Annila, 2007; Makela and Annila, 2010). This has everything to do with physics (Maupertuis' principle of least action and the 2nd Law) and nothing to do with genetic variation.

Alternatively, to attribute the LTEE results to 'genetic variation' requires a) accepting Johannsen's genotype conception as correct (see section 3 on the lack of secure underpinning empirical evidence) and b) that the experimental evidence from the LTEE cannot be taken at its face value but needs to be corrected using poorly understood 'genetic tools' such as epistasis. Further, (Kocher and Dill, 2023) point out that Darwinian evolution is 'antithermodynamic' because it is a process driving away from equilibrium and, therefore, would require a force (for which there is no empirical evidence).

Science dictates that the former explanation be accepted, and the latter abandoned.

It is, therefore, clear that Fisher's law, based as it is on genetic variation, is false, and that natural selection is an entirely different process than the one currently envisaged by mainstream evolutionary biology, the Modern Synthesis: it is not driven by genetic variation. Further, Darwin's 'struggle for existence' is, as was pointed out by Edward Blyth (1835) 20 years before Darwin published *The Origin,* in fact, the struggle for nutrients, i.e., for free energy. As increasing free energy implies increasing entropy, we can understand why it is that Boltzmann asserted that organisms seek entropy (Boltzmann, 1974).[6] Furthermore, in the case of the LTEE, and I suspect much more widely, epistasis—a 'known genetic and evolutionary process', in the words of Reuter's expert and independent reviewer—is nothing more than a 'fudge factor' to make experimental results agree with the theory. Naturally, if experimental data are interpreted on the *false* premise that the units of inheritance are genes, the data will *not* comply with the theory. For the last century, genetics, and evolutionary theories, have been based on a false premise, and that is why experiment and theory are so often in conflict and the invocation of epistasis (and likely several other so-called 'genetic processes', such as over- and underdominance) is needed to bring experiment in line with theory. That Lewontin's paradox (Lewontin, 1974), which concerned the historical *lack* of genetic experiments performed with difficult-to-measure, but interesting traits, was so prescient, yet ignored, even by Lewontin himself, is a mark of a lack of collective critical thinking on the part of geneticists over decades. As Paneth and Vermund have pointed out, there have been no measurable benefits to public health as a result of half a century of molecular genetic research, but at least 17 Nobel Prizes have been awarded for 'discoveries' in genetics (Paneth and Vermund, 2018). We now know why.

3. Response to Nils Roll-Hansen

While stating his concluding remarks in his commentary (Roll--Hansen, 2022), which authoritatively and thought-provokingly details how Johannsen came to propose the genotype conception, Nils Roll-Hansen says, 'I judge the preceding narrative to show that even if the gene concept is problematic the distinction between phenotype and genotype is scientifically well-founded' (Roll-Hansen, 2022). I agree, and Wilhelm Johannsen surely deserves the credit for this. However, what is open to question is whether the pure line bean experiment truly underpinned the genotype concept, especially now we know that it is wrong. It has so far proved to be a very expensive mistake in that the landmark molecular genetic project, the $3 billion Human Genome

[5] See Fig. 3 Illustrating the 12 individual profiles for cell size, and compare with Fig. 4, illustrating the single profile of the 12 experiments for fitness.

[6] See footnote 66 in Baverstock (2024) for fuller quotation.

Project, can now be seen as unnecessary, at least concerning the urgency with which it was undertaken.

The bean experiment did falsify Galton's law of regression (Galton, 1885), which one can argue was responsible for the eugenics movement started by Galton in 1883 and which grew to pose as a 'science' with an endowed Chair at University College, London, held by Karl Pearson. Moreover, the bean experiment unequivocally demonstrated that in a pure line, nothing additional can be gained by selection, disqualifying the then-dominant idea of continuous variation playing a role in evolution. Johannsen favoured de Vries' mutation theory (Vries, 1901), i.e., entailing non-continuous variation. Mendel's experimental work of the mid-1800s became the basis for Fisher's *Genetical Theory of Natural Selection* (1930) and Huxley's *Evolution: The Modern Synthesis* (1942), both under the appellation of Mendelism. Both are currently active in the mainstream of genetics and evolutionary theory, but both are fatally challenged.

As far as the genotype conception is concerned, I have, as a practicing physical scientist (rather than a historian); and based on my reading of other historians' accounts of Johannsen's experiment, a somewhat more pragmatic view than Roll-Hansen. To summarise, I don't think Johannsen could have opted for the other available choice, the transmission/phenotype conception, given the intellectual environment so clearly described by Roll-Hansen and his fellow historians, even though the biometric approach, which was pioneered by Galton in his 'Law of Ancestral Heredity' (Galton, 1876) and reinforced by Pearson (Pearson et al., 1903), cannot be described as anything other than convincing. There are two main reasons for this: 1) Galton's Stirp Theory and Weisman's Germ Plasm Theory, both invoking what Johannsen would describe as 'genotypical' as opposed to phenotypical, theories, were circulating widely at the time and were respected by Johannsen; and 2) it would not necessarily have been clear, in the case of the bean experiment, what the phenotype to be inherited in the phenotype conception would be: a bean or a plant. Further, at the time, the term 'phenotype' applied to the whole organism, and its use in the context of a single cell, a gamete, was probably unknown.[7] Consequently, Johannsen opted for the only option for him. What is inexcusable is that the genetics community, in the 100 or so years since then, has not questioned Johannsen's decision.[8]

It is not as if there have not been opportunities to reassess the wisdom of the genotype conception. To mention just two: 1) As already mentioned, in 1974, the late Richard Lewontin, a highly respected geneticist, drew attention to his famous paradox about *experimental* genetics being focused on measurable but uninteresting traits, rather than on difficult-to-measure but interesting traits (Lewontin, 1974). Surely, this must have indicated that something was seriously amiss in genetics. 2) In 2001, it became clear to the world that the Human Genome Project had found that the number of genes in the human genome was many fewer than previously thought. This led Stephen J Gould to write an opinion piece for the New York Times entitled *Humbled by the Genome's Mysteries*.[9] Surely this was a major collision between reality and belief.

The pointless field of genome-wide association studies (GWAS) still thrives today,[10] despite the perceived failure of the earliest large study in 2008, the Wellcome Trust Case Control Consortium (WTCCC) study, which commenced in 2005. 'While early large scale GWAS were successful in identifying large numbers of genes associated with CCDs [common complex disorders] …. it was widely agreed in the period immediately following that these findings in themselves are insufficient to ground new medical interventions' (Heeney, 2021). The reason for the failure was that the extent of the genetic component, expected to be 50 or more percent for many common so-called polygenic traits, was mostly in single figures, leading to the still unsolved problem that became known as the missing heritability problem (Manolio et al., 2009). The idea that a product of GWAS, the polygenic risk score (PGS/PRS), could be taken seriously in a clinical context was discussed in a BBC radio documentary, hosted by Adam Rutherford, a geneticist, and broadcast in December 2022.[11]

Furthermore, Jan Sapp says that the period 1900–1910 was characterised by a battle for scientific authority between the Mendelians and the biometricians; among the latter, Galton and Pearson (Sapp, 1987). Galton had formulated the Law of Ancestral Heredity (Galton, 1897) based on his theory of stirp in the germplasm and a statistical (biometrical) analysis of adult human characteristics, such as height and eye colour, over a few generations. Although Johannsen was, according to Roll-Hansen, initially supportive of the biometric approach, he was converted to Mendelism shortly after Hugo de Vries republished Mendel's work on pea plants in 1900. A parallel controversy at the time was the nature of variation: was it continuous, as the biometric school maintained, or discontinuous, as de Vries maintained based on his Theory of Mutation published in 1901 (Vries, 1901)? Thus, Johannsen's pure line experiment with beans was set against a background of fierce controversy, and on the publication of his results in a 100-page or so monograph in 1903, his first critics were British biometricians, Karl Pearson and Walter Weldon (Anonymous, 1903; Weldon and Pearson, 1903).

The criticisms were answered (though not to the satisfaction of Pearson) at the time by George Yule (1904), a mathematician, and it seemed to have settled the argument as far as the wider scientific community was concerned, but apparently, the criticisms were never addressed by Johannsen, even in his 1911 publication. Roll-Hansen identifies the Pearson criticism as the stimulus for Johannsen introducing the concepts of 'genotype' and 'phenotype' in 1909, in his textbook *Elemente der exakten Ereblichkeitslehre* (The Elements of an Exact Theory of Heredity). He published these concepts in English in a paper written in 1911 (Johannsen, 1911).

However, Johannsen's lack of response to Weldon and Pearson's criticism (Anonymous 1903) is an issue, especially since according to Provine (2001),[12] Johannsen conducted further experiments that he claimed confirmed his 1903 results, so he could have easily addressed the criticisms. Yule claimed that Pearson and Weldon had misunderstood Johannsen's claim: 'I find it difficult to understand Prof. Johannsen's book in the sense in which the reviewers have, apparently, read it. In both notices, it is stated that, if the author's views were correct the correlation between mother and daughter plants should be perfect. As I take it, however, Prof. Johannsen's view does not imply, and is not consistent with, such a hypothesis. … This misunderstanding, in my

[7] Roll-Hansen seems to be confused with the application of the term 'phenotype' to a single cell. There is an asymmetry in the definitions referred to by Roll-Hansen in the *Stanford Encyclopaedia of Philosophy*. The genotype is a cellular feature, whereas the phenotype is mostly used as an organismal feature. The latter could be viewed as a linear sum of the constituent cellular phenotypes comprising the organism. We now know that the cellular phenotype results from a continuous process of gene product interaction in what is called the protein interactome, located in the cell cytoplasm. So, it is correct to call it a 'process' and a 'thing' but we understand very little about how phenotypic characteristics can emerge from a protein interactome.

[8] I am not aware of any serious debate about the issue since Pearson questioned it in the early 1900s and maintained his critiques.

[9] https://www.nytimes.com/2001/02/19/opinion/humbled-by-the-genome-s-mysteries.html.

[10] Searching PubMed on the terms 'genome wide association', genomewide association', 'genome-wide association', and 'GWA' in the Title and Abstract field produced 44,585 publications between 2001 and 25 February 2024, with 14,404 since 2021.

[11] https://www.bbc.co.uk/sounds/play/m001gj50 Listen at 20 min 45 s in to the programme where US geneticist Kathryn Paige talks about the use of genetic information in social policy making and the role of polygenic scores. [checked 05.04.2023].

[12] See pp. 97–98.

view, is fundamental' (Yule, 1904). In response, Roll-Hansen writes, 'To attribute such a blunder to Pearson and Weldon may appear audacious on my part. They must indeed have been very superficial or single-minded in their reading of Johannsen. It is, however, the only acceptable interpretation of their text that I have been able to find.' I cannot agree, because as in their response to Yule (Anonymous, 1903), Weldon and Pearson reiterate their criticisms, and, as the matter was still unresolved in 1906 when Johannsen addressed a meeting in London. 'Still smarting from the criticisms of his pure line researches by Pearson and Weldon, Johannsen entertained a receptive audience with repeated blasts at the biometricians' (Provine, 2001).[13] I suspect that the criticism was valid, and Johannsen was aware of that, so failed to address the criticism.

According to Roll-Hansen, Johannsen's visit to the USA in 1911 was key to the acceptance of the genotype conception: 'The acclaim that met Johannsen in America was proof of the success of his theory. It was the high point in his scientific career.' Provine writes, 'Johannsen's book of 1903 has been hailed as a very important step in the history of genetics. All geneticists know that his ideas concerning heredity in pure lines were basically correct, but it is not generally known, as Pearson and Weldon pointed out, that Johannsen's data were an imperfect support for the conclusions he drew from them. The genetics literature from 1903 onwards contains rare citations to the criticisms of Pearson, Weldon, and Yule; but it contains hundreds of citations of Johannsen's 1903 data as if they proved the pure line theory' (Provine, 2001).[14]

I, for one, give the final word to Sapp. Following this somewhat disingenuous (in the light of the weakness of Johannsen's scientific position and his failure to respond to Pearson's criticism) quotation from Johannsen's 1911 paper: 'The genotype-conception is thus an "ahistoric" view of the reactions of living beings—of course only as far as true heredity is concerned. This view is analogous to the chemical view, as already pointed out; chemical compounds have no compromising ante-act, H_2O is always H_2O, and reacts always in the same manner, whatsoever may be the "history" of its formation or the earlier states of its elements. I suggest that it is useful to emphasize this "radical" ahistoric genotype-conception of heredity in its strict antagonism to the transmission-or phenotype-view.' Sapp then states: 'The fundamental basis of heredity and variation would now be hidden deep within the gametes of the organism' (Sapp, 1987). It means that the topic of heredity is out of reach of biometricians, naturalists, and breeders. This, perhaps, helps to explain why geneticists regard themselves as the 'high priests' of biology when they can be more accurately described as 'true believers' in an ideology (Lewontin, 1992) and master fudgers in making empirical data fit the ideology.

In response to my assertion in (Baverstock, 2024) that development is not driven by transferred parental genes, but rather can be properly seen in terms of self-organisation, and that the 'zygote knows' what it will develop into quite independently of its genotype, Roll-Hansen says this 'is quite contrary to present biological thinking'. Absolutely! I am proposing a quite different and, I think, unique view on development (Baverstock and Ronkko, 2014), which, in the early embryonic stages, is now being seen as based on self-organisation (Shahbazi et al., 2016; Shahbazi et al., 2019). Furthermore, the fate of a zygote cannot be determined by its genotype alone. The gene sequences of a mouse and a human are almost identical, and yet we have never seen a mouse give birth to a human or vice versa. To express a phenotype, two sources of information are required (Baverstock, 2011): in the IA model, the 'code for the gene products' and the 'rules of engagement'. The latter determines the fate of a zygote.

4. Response to Giuseppe Longo

I thank Giuseppe Longo for his deep and thoughtful constructive criticism (Longo, 2023). He makes numerous points from the standpoint of a project of a group of his colleagues, writing, 'We are a small community trying to advance in a theory proper to the organism ….' This is an aim my colleagues and I share, and it is interesting to note that starting at roughly the same time (2000), with one common initial influence—Stuart Kauffman's work on the origins of order (Kauffman, 1993)—we have landed in very different places in respect of the way we see physics as it applies to biology. Longo worked with Kauffman on aspects of the physics of what Kauffman termed the 'unprestatable' or 'unpredictable', nature of the evolution of life, contrasting that with the laws governing motion, which have characterised physics from the time of Galileo and Newton (mid-1600s). In an article in the Huffington Post,[15] Kauffman used the events that followed Turing's work, which ultimately led to the personal computer and the creation of the World Wide Web, and Facebook: an evolution over less than 100 years that could not have been predicted or prestated by Turing. Most importantly, in the view of Longo, Montevil and Kauffman, biological evolution is *lawless* (Longo et al., 2012a).

Their view contrasts with the one my colleague and collaborator, Arto Annila, and I infer from data. Namely, it is generally recognised that the same patterns in data are seen throughout nature. For example, power laws are ubiquitous. They approximate sigmoid curves that accumulate from skewed distributions. However, our point is not to argue for a particular mathematical form, say, power-law, lognormal or logistic function, but to point out that the data emerges with patterns because of systems following a common physical principle without demarcation between biological and non-biological. We argue, in many publications, that this principle is the principle of increasing entropy, equivalent to the least-time consumption of free energy, derived from statistical physics (Annila, 2020). I will return to the equation of evolution, but at this juncture, please note that the general principle is in agreement with the power-law outcome for the evolution of fitness that the unique evolutionary experiment, the LTEE, has yielded over the past 30 years.

Another important difference lies in the reasons we each undertook our projects. For me, the Independent Attractor (IA) model, the product of our group's research, commenced with trying to explain the phenomenon of genomic instability reported by former colleagues at the UK Medical Research Council's (MRC) Radiobiology Unit in 1992 (Kadhim et al., 1992). Initially, I thought this result was an artefact but subsequent experiments (Lorimore et al., 1998), which revealed the bystander effect, convinced me otherwise, and in 2000, I wrote a paper exploring how biological systems can be influenced by radiation through routes other than mutation (as, clearly, genomic instability was not caused by mutation) (Baverstock, 2000). In previous work (Baverstock and Cundall, 1988) on energy transfer in DNA, Bob Cundall and I invoked solitons as carriers of energy along the DNA molecule. A soliton is an example of a dynamic steady state and in discussions with Alwyn Scott, whose landmark book *The Non-linear Universe: Chaos, Emergence, Life*, which was in preparation at the time (Scott, 2007), solitons seemed to be a promising route to understanding energy transfer in DNA. Some years later, I started to explore the idea that the cellular phenotype might be represented by an attractor state (multiple dynamic steady states) comprised of gene products (mainly proteins).[16] A radically different model for the cell was needed because the phenomenon of genomic instability could not back then and, presently, still cannot be accommodated in the prevailing molecular genetic paradigm. Thus, we started, so to speak, with a clean slate as far as accounting, in terms of physics and chemistry, for the phenomenon of living matter.

Longo's group accepted the *status quo* of biology and looked to use physics to explain biological phenomena, which they saw as needing the

[13] See p. 97.
[14] See p. 97.
[15] https://www.huffpost.com/entry/co-creating-our-world_b_2398515.
[16] There is a very fundamental dynamic steady state in biology, namely the balance between spontaneous DNA damage and cellular DNA repair processes.

application of a wide range of theories from the physics of the inanimate world: living systems 'an entanglement of almost all physical scales' (Longo, 2023), as well as a high level of complexity. For example, the number of conceivable biological macromolecules active in the simplest of cells is so great that most have never been synthesised. As Kauffman calculates (Kauffman, 2000),[17] for a 200-residue polypeptide consisting of 20 natural amino acids, there is the potential for 10^{260} unique molecules. This huge number is more than the estimated number of particles in the known universe (Kauffman, 2000). The problem only gets worse when these molecules network into a single cell, let alone a tissue or whole multicellular organism (Noble, 2017). Moreover, this so-called combinatorial problem is only a small part of that complexity. Indeed, Noble's book (just referenced) *Dance to the Tune of Life* is a readable and comprehensive account of the dimensions of that complexity.

Both Longo and I must 'make sense' of that complexity, and we have recognised three ways to achieve this.

The first option, put forward by Noble (2017), is holistic. He chooses to frame the complexity in terms of qualitative *networks*. The networks are more than the sum of their parts. A collection of networks can be 'modularised' into component networks of lesser complexity, considering the strengths of interactions between the networks.

The second option, adopted by Longo and his colleagues, is to choose a 'space' into which the cell/organism can be conceptualised in terms of observable features. The collective properties of gas molecules, e.g., temperature and pressure, can be conceptualised in a *phase space* where each molecule, at any given time, has a spatial position (coordinates x, y, and z) and three vectors for its motion, i.e., a 'phase space' of six dimensions in which every molecule is uniquely defined at an instant of time and their trajectories can be observed/computed over time. Thus, growth and evolution can be conceptualised in some kind of *phase space*.

The third option, chosen by me, initially in collaboration with a former MRC colleague, Mike Thorne, is to employ a *state space*. In contrast with phases, i.e., configurations with the same energy, the states differ in energy. Specifically, the state space of a cell is inhabited by cellular phenotypes derived from interactions between gene products, which comprise the dimensions of the state space (Baverstock and Ronkko, 2008), typically a few thousand.

Considering Longo's arguments, I acknowledge that the phase space trajectories can be mathematicised and computed. However, the conservation of symmetry means the phase space does not evolve in *energy*, whereas evolving systems consume free energy, in the case of organisms, from nutrients. Conversely, I recognise that a state space only provides a snapshot of the system in time, compared, if you like, with a video provided by a phase space approach. A state space may, therefore, appear less useful in the context of evolution and development. However, the state of a system, not in balance with its surroundings, contains free energy. Free energy is the force that drives the system from its present state to another in the quest for attaining balance with its surroundings. Consequently, it is possible to deduce from gradients in energy where the system is heading.

Conversely, Longo writes of our efforts, 'Physics, or better, according to Baverstock, only one of its theories in the very rich theoretical scattering of often incompatible theories, should instead allow to derive the properties of the living on the basis of a single, in itself very interesting, principle of optimality, the principle of "least action"' (Longo, 2023). Well, yes but Longo reasons that life is a special case, only existing in specific situations where suitable nutrients can be obtained from the environment. Obviously—it speaks for itself. However, any system, animate or inanimate, evolves to attain thermodynamic balance with its surroundings. This drive for balance manifests itself in such a way that an organism actualises its potentiality as it comes to terms with its environment. The neurologist and psychiatrist, Kurt Goldstein, stresses the importance of this holistic relationship between environment and organism, if the organism is to achieve its *full* potential, in his 1934 masterpiece *Der Aufbau des Organismus* (Goldstein, 1934). Therefore, the key to making sense of development and evolution is not in the genes but in the phenotypes interacting with their environments. As noted in Section 2, this is what was observed in the LTEE, where resilience is maintained while mutational damage affects individual proteins (Zitnik et al., 2019; Maddamsetti and Grant, 2020). I have proposed that evolution should be seen as a two-part process (Baverstock, 2022): firstly, a process based on the above physics, internal to the cells of the organism; and secondly, a process in which the agency of organisms acting in their environments can lead to macro-evolutionary changes. On this basis, we would argue that the thermodynamic principle of least action alone provides a physics base for understanding the life process at its most fundamental level.

It is common to think, as does Longo, 'It is a good practice, closer to the method of physics, to give oneself first an adequate theory of the pertinent phenomenal level, the organism, …. and then try to unify … ' (Longo, 2023), i.e., the unification of existing theories of physics is the way to progress in applying physics concepts to biology. However, why unification? After all, social media is replete with well-known physicists from elite universities and research institutes disagreeing about the value of current theories. To take one example, German physicist, Sabine Hossenfelder, speaking on the Institute of Arts and Ideas channel[18] describes the present state of physics as 'stagnation' – will unifying these theories improve the situation? This is not a new problem. In 2006, Lee Smolin reflected on the then state of physics (Smolin, 2006), noting that few major advances had happened since 1981, compared to the previous 200 years of 'explosive growth' in the subject. Little has changed, except, perhaps, for the detection of the long-proposed Higgs boson, since Smolin wrote his book.

Therefore, I am motivated to argue for the IA model because it complies with thermodynamics. Thermodynamics is a universal and fundamental theory when derived from microscopic entities using statistical physics, as Annila's publications substantiate (Annila, 2020). In its original Maupertuis form, instead of Lagrange's constant energy form, the principle of least action is of profound importance in understanding evolving systems (Annila and Baverstock, 2014).

Indeed, Kocher and Dill describe Darwinian evolution, which I assume Longo accepts at its face value, as 'antithermodynamic' i.e., driving away from equilibrium (Kocher and Dill, 2023), in contrast to what we are proposing and which applies across the natural world, the seeking of free energy balance, in this case between organism and environment.

In contrast to Longo's stance, which I detect as an underlying theme of his commentary, there is nothing, at a fundamental physics level, that distinguishes the inert from the living. Any plot of data, once the labels are removed from the axes, is indistinguishable as to whether it is from an inert or a living evolving system (Annila, 2020). In other words, it would be a revival of vitalism to claim that life possesses something special. On the contrary, it is logical to think that theories of physics have focused on stationary systems, whereas living systems are obviously evolving and changing. Indeed, the mathematical biologist, Robert Rosen, questions whether, as far as physics is concerned, life might be the general and the inert, the special case, and consequently, studying life without the encumbrance of the plethora of theories Longo refers to might lead us to physics that is missing from our understanding of the inert (Rosen, 1991).

Longo acknowledges that our starting point 'is certainly valid: an organism is at least a thermodynamically open system, to which certainly apply principles of thermodynamics', but mistakes the thermodynamics of open systems for the 'Prigogine style – systems far from equilibrium (Nicolis, 1989; Longo, 2023). We explicitly reject the concept of neg-entropy and anti-entropy which play a central role in Longo's and his colleagues'

[17] See p. 144.

[18] https://www.youtube.com/watch?v=8aUk6oi_AmM.

model (Bailly and Longo, 2009; Longo and Montevil, 2012). However, it is not that we are arguing against Longo specifically. The issue is that physicists take Boltzmann's entropy as generally valid despite it being derived from the steady-state condition (Annila and Baverstock, 2016). At the steady state, a system will lose its phase coherence and sink into disorder when in contact with disorderly surroundings. However, the disorder is not synonymous with increasing entropy (Annila and Baverstock, 2016) which is equivalent to decreasing free energy, e.g., when nutrients are consumed by organisms. In other words, Boltzmann's calculations addressed phase space configurations of thermodynamically closed systems, whereas we describe evolution in state space where a system evolves from one state to another, differing in energy. This consumption of free energy manifests as increasing complexity and is order driven when the system evolves to gain balance with an environment rich in resources. It is a natural process (Sharma and Annila, 2007). Natural processes include molecular events, such as the repair of damaged DNA, and organismal series of events, such as embryogenesis.

On a related issue, Longo says 'Unfortunately, but perhaps just to conform to the dominant fashion, Baverstock continues to use "informational" language' (Longo, 2023). I am doubtful that informational language can be avoided when it comes to biology. At this juncture, I would like to draw the reader's attention toward a recent paper by Keith Farnsworth titled *How an information perspective helps overcome the challenge of biology to physics* (Farnsworth, 2022). Farnsworth sees information as the formal cause of organisms that exist as 'dynamic patterns of matter and energy in space and time', such patterns continually being reinforced through closed-to-the-efficient-cause feedback loops. Similarly, in the IA model, we regard the formal cause as the 'rules of engagement' (i.e., information) (Baverstock, 2024). As in Longo's vision, history is a vital element of the IA model: the rules of engagement and, of course, the DNA sequence, go back to the origin of the species. Also, as we have invoked, energy and information are exchangeable, and indeed, there is an on-chip Maxwell's Demon that acts as an information-powered refrigerator, developed at Aalto University in Finland (see Davies, 2020 for further information). In this context, the presence of intrinsically disordered proteins and their continual state of folding and unfolding in the cell cytoplasm (Fonin et al., 2018), which seems to be a component of the learning process in non-neural cells (Csermely et al., 2020), is a free energy-to-information conversion process.

Longo raises the issue of the state of water in the cell, as do I in (Baverstock, 2024).[19] It is indeed a curious state that supports a high concentration of organic solute yet does not have a high viscosity. Longo suggests that quantum electrodynamics (QED) would be required to understand this phenomenon. Without question, QED provides the anomalous magnetic moment of the electron with high precision. However, as Annila points out, QED does not explain what gives rise to the moment (Lehmonen and Annila, 2022). In other words, models of modern physics are instrumentalism. They give the right numbers but do not specify what the model parameters mean in the real world.

In 2018, the Royal Society Interface published a headline review entitled *The Future of Quantum Biology* (Marais et al., 2018). They cite examples of biological phenomena that cannot be explained in terms of classical physics. However, such a claim is subject to defining classical physics. For example, Newton's second law is often said to be $\mathbf{F} = m\mathbf{a}$ (Euler's form) instead of the original form, i.e., $\mathbf{F} = d\mathbf{p}/dt = m\mathbf{a} + \mathbf{v}dm/dt$, where the change in mass, dm/dt, corresponds to the loss of energy in a chemical reaction. Leaving out dissipation certainly narrows the application of the second law to stationary systems. Likewise, the Schrödinger equation describes the phase evolution of the wave function, and not the system's evolution from one state to another.

Moreover, it is worth recalling that the key concept of quantum mechanics, the wave function, is not observable but collapses. Therefore, notions such as quantum entanglement are not falsifiable. From this perspective, we are not refuting modern physics but recognising that it is instrumentalism rather than rationalism or empiricism. John von Neumann, a pioneer in information technology, foresaw this current trend more than half a century ago: 'The sciences do not try to explain, they hardly even try to interpret, they mainly make models' (Annila, 2020).

Finally, Longo writes: 'I disagree with both Baverstock and Noble on their critique of the validity of the CD [Central Dogma].' However, as has been noted in (Baverstock, 2024), the inheritance of acquired characteristics will have to play a role in some form of organismal agency to be part of the process of evolution. We already know about another route of inheritance, namely prions (Halfmann and Lindquist, 2010). These proteins, in addition to being associated with disease, can have more than one stable folded form, which can be 'copied' by other proteins; potentially, not only can they change the phenotype of a cell without cell division, but they can also inherit those folded structures. Also, in a variety of amyloid diseases, kernels of the aberrant form can self-propagate through seeding—protein-to-protein—by themselves.[20] They are certainly violating the central dogma.

I hope I have answered most of the constructive criticisms raised by Giuseppe Longo, and I thank him for his interest in writing his commentary.

5. Response to Steven Rose

Steven Rose concludes his thoughtful commentary (Rose, 2023) with the following, more-than-welcome sentence: 'Baverstock's de-emphasising genes in favour of cells, I suggest, fits well within this larger theoretical framework', which he expertly elaborates in the body of his commentary. There are, however, some issues he raises that I would like to address.

Rose calls the term 'phenotype', as I deploy it, 'vague and slippery'. I disagree. I have deliberately defined what I am discussing in (Baverstock, 2024) as the '*cellular* phenotype' to avoid confusion with the multicellular phenotype, and the levels of complexity outlined in Rose's paragraph's beginning 'Baverstock's confinement of the term [phenotype] to the cellular level is at once too broad and too restricted.' I am, of course, fully aware of the full extent of the term's deployment in biology, and beyond to the 'extended phenotype'. I am not 'seeking to reduce all the levels [of complexity Rose outlines in that paragraph] to that of the cellular phenotype.' What I am saying is that my approach to understanding the complexity of biology *begins* by understanding the basic unit of organisms, the cell. I am down in theoretical biologist Robert Rosen's 'building basement',[21] abstracting what I find useful there to formulate a framework upon which I can build an alternative building because I believe, and I show in (Baverstock, 2024), that the present edifice is built on unsound foundations. I suggest that the cell is the lowest organismal level at which the term phenotype is applicable. The virus, as Keith Farnsworth shows, is not, in Rosen's terms, an (M,R)-system and is, therefore, not alive (Farnsworth, 2021). Hence, the complexity Rose describes would be relevant at a later stage of my quest, although Mauno Rönkkö and I have addressed development in the context of the IA model (Baverstock and Ronkko, 2014). However, it is an interesting question as to how much of that complexity arises as the result of false premises, genes being assumed to be the unit of inheritance, for example. But for sure, there is no avoiding the fact that cells and organisms are, in the words of Farnsworth, 'dynamic patterns of matter and energy in space and time' (Farnsworth, 2022), and that the cellular phenotype is a process, as well as a thing, and that the whole organismal system at all levels must be viewed in the context of the physics of complex dissipative systems and not Newtonian physics. The attractor state that my

[19] See footnote 77.

[20] https://www.ucl.ac.uk/prion/news/2023/jan/unique-how-probing-atomic-scale-diversity-among-prions-electrons.

[21] In Chapter 3 of his book *Life Itself,* Rosen speaks of scientific achievements being like a tall building. He warns that visits to the basement, however unwanted, are necessary to ensure that the foundations are secure (Rosen, 1991).

colleagues and I have proposed to represent the cellular phenotype (Baverstock and Ronkko, 2008) is a multidimensional, dynamic, quasi steady state, therefore unequivocally a process, and an entity much closer to nature than the alternative feedback-based system upon which cybernetics is based (See Bertalanffy, 1969). It is that dynamic steady state that can be referred to as homeostasis: it has nothing to do with death and everything to do with dynamics. There is another connotation of the term 'stasis' that is valid, i.e., the objective of the organism to reach a state of balance energetically with its environment, which is implicit in the principle of least action. The latter is not about death either.

Rose deems the metaphor of a brain for the cellular phenotype as 'unhelpful'. I disagree. The cellular phenotype I propose produces both Mendel's characters and regulates that process, leading to circular causality, very different from the gene-centric view. While it is true that, 'in a multicellular organism, the agency of any single cellular phenotype is constrained by its embeddedness,' no character can appear in a multicellular organism's phenotype, unless the appropriate cellular phenotype is present. Neither do I think it is the case that 'multicellularity demands the emergence of specialised cells'; any demand for specialised cells derives from the environment to which the organism is adapting.

6. Discussion

6.1. A note about the importance of face-value evidence, epistasis, and polygenic risk scores

The purpose of an experiment is to falsify doctrine, i.e., laws, theories, and hypotheses. However, when such falsification is achieved, and it is challenged or negated by invoking an ad hoc correction factor to 'rescue' the established doctrine, that is an abuse of the 'scientific process'. Instead, the proper response for those who do not wish to accept the *face value* result, for example, as in the case of the reinterpretation of the LTEE in (Baverstock, 2024), where the Editor-in-Chief of the JEB invoked epistasis to refute the falsification of Fisher's theory of natural selection, is to propose an experimentally testable hypothesis that would justify invoking epistasis. In the meantime, the falsification should be accepted at its face value.

I have addressed the issue of epistasis in the context of the interpretation of the LTEE in Section 2 above. As noted, the literature on epistasis over the past 100 years is extensive, muddled, and contradictory presumably as a result of the term being applied in several different contexts (Phillips, 2008). Fisher's term, *epistacy,*[22] coined in 1918 for any deviation from pure additivity (Fisher, 1918), has dropped out of use, and *epistasis* is now applied to any so-called 'gene interaction'. That term is potentially misleading because it cannot be the genes themselves that interact, but rather their products, foremost among them being proteins. However, protein–protein interaction is a ubiquitous feature of cell biology (Phizicky and Fields, 1995). Indeed, Rönkkö and I have equated the cellular phenotype with the attractor state of the gene product (protein) interactome (Baverstock and Ronkko, 2008). If this is correct, it would be impossible to separate epistatic interaction from the more general feature of the cell's functioning, out of which its cellular phenotype emerges (see Section 2).

Taken at its face value, the evidence (Baverstock, 2024) from the LTEE (the power law fitness trajectory exhibited identically by the 12 LTEE E. *Coli.* bacterial cultures) supports the position that evolution is not, as proposed by Fisher (1930) and is almost universally assumed to be, driven by genetic variance but rather is a manifestation of a generic, in natural systems, thermodynamic process of equilibration of free energy between the organism and its environment. This constitutes the processes of adaptation, as defined by Popper (Niemann, 2014), and natural selection (Sharma and Annila, 2007; Makela and Annila, 2010). Mainstream evolutionists argue that the power law trajectory is the result of epistasis (the interaction between existing and newly acquired mutations) to produce 'declining adaptability' (Johnson et al., 2023) or 'diminishing returns (in fitness)' (Wiser et al., 2013) but they only produce statistical arguments to the effect that the power law trajectory can be modelled to be consistent with the LTEE experimental results (Wiser et al., 2013). Such modelling could, with suitable assumptions, model almost any trajectory and it is not acceptable as a refutation of the face-value interpretation of the LTEE.

That face-value interpretation hinges on the rejection of the universally accepted genotype conception, based on the empirical evidence from Johannsen's pure line breeding experiments performed between 1900 and 1903 and reported in a lecture in 1910 (Johannsen, 1911). I have argued (Baverstock, 2024 and Section 3 herein) that Johannsen's adoption of the genotype conception over the alternative transmission/phenotype conception is unconvincing. This conclusion rests on the face value interpretation of the LTEE, i.e., that it does not support genes being Mendel's units of inheritance. I now add the following evidence, to further strengthen the case for rejecting the genotype conception.

1. Darwinian evolution, based on the genotype conception, is antithermodynamic and requires a so far elusive driving force if it is not to violate the 2nd Law (Kocher and Dill, 2023), whereas evolution viewed under the assumption of the phenotype conception is in accordance with the 2nd Law (Baverstock, 2024; Annila and Baverstock, 2014).
2. Bacterial genomes reduced *in vitro* through 'streamlining' or as a result of parasitism, mutated, i.e., acquired genetic variance, faster than the larger bacteria from which they were reduced but evolved in fitness at a similar rate (Moger-Reischer et al., 2023) thus proving that the rate of evolution is independent of genetic variance.
3. Johannsen's decisive rejection of the transmission/phenotype conception in heredity (Johannsen, 1911), i.e., the biometric approaches of Galton and Pearson,[23] and the influence of ancestors,[24] betrays an anti-scientific 'bias' in favour of the genotype conception.
4. Johannsen's arguments in favour of the genotype conception (Johannsen, 1911) are not definitive and the same pure-line experiments can also be interpreted as supporting the phenotype conception.[25]

Furthermore, the biometric approach taken by Galton in the late 1800s (Galton, 1876) and followed-up by Pearson in the early 1900s (Pearson et al., 1903), based on human biometric data and animal traits, such as coat colour in dogs and horses, provides reasonably accurate

[22] Fisher refers to 'epistacy' as a 'statistical term'.

[23] Johannsen says: 'The famous Galtonian law of regression and its corollaries elaborated by Pearson pretended [my emphasis] to have established the laws of "ancestral influences" in mathematical terms'.

[24] Johannsen writes: 'Ancestral influence! As to heredity, it is a mystical expression for a fiction [my emphasis]. The ancestral influences are the "ghosts" in genetics, but generally the belief in ghosts is still powerful. In pure lines no influence of the special ancestry can be traced; all series of progeny keep the genotype unchanged through long generations.' Clear evidence of ancestral influence exists in the portraits of the Habsburg line over generations.

[25] The beans that Johannsen sowed for the second growing season (1902), which were the progeny of 19 pure lines, i.e., 19 specific 'types', were, in effect, given two phenotypes, namely, a) their phenotype as that progeny (the average weight of progeny beans from the specific line or 'type'), and b) their individual weight on the distribution of bean weights yielded by that line or 'type'. In the experiment beans of a specific 'type' but varying in phenotype 'b', all yield beans with phenotype 'a'. Based on this result Johannsen claims that it is the 'type' (which he then deems to be the genotype) that determines the progeny phenotype 'a'. This is the basis for Johannsen's claim that the genotype conception applies to heredity. However, it is also the case that the parent beans with phenotype 'a', i.e., the progeny of a pure line, yield progeny with phenotype 'a', thus supporting the phenotype conception. I maintain that phenotype 'a' is the true phenotype of the progeny of a pure line and phenotype 'b' is not a true phenotype but a somatic, and thus non-heritable, variation and, therefore, that Johannsen's experiment supports the phenotype conception.

predictions of offspring phenotypes/traits provided the data from ancestors beyond the immediate parents is included.

Thus, the evidence in favour of replacing the genotype conception with the phenotype conception where evolution and inheritance are concerned is compelling.

Given this conclusion, epistasis plays no role and is no more than a confection 'cooked up' to explain experimental results on the basis of a false premise: it has no part to play, let alone an essential part, in the process of evolution (Phillips, 2008; Wiser et al., 2013; Johnson et al., 2023).

From my experience over more than 20 years, changing one's perspective from the 'gene-centric' to what might be called the 'phenocentric' view of biology is no small task and it is easy to slip back into gene-centric thinking. For some, I think it is an impossible transition,[26] but to continue, for example, the search for genes that will 'tell us who we are', as Robert Plomin claims as his aim in his book *Blueprint* (Plomin, 2018; Plomin and von Stumm, 2022) is not only pointless, it is dangerous because Plomin is putting his polygenic risk scores (PGSs) forward as a basis for education policy. The 'blurb' for an earlier book, *G is for Genes* (Asbury and Plomin, 2014), claims that the book 'shows how a dialogue between geneticists and educationalists can have beneficial results for the education of all children–and can also benefit schools, teachers, and society at large.' That audience is a community that is likely to take advice on trust rather than being able to recognise the fallacy they are being sold. I have already condemned PGSs from the 'genetic' perspective (Baverstock, 2019). Calle Burt has written a devastating critique entitled *Challenging the Utility of Polygenic Scores for Social Science: Environmental Confounding, Downward Causation, and Unknown Biology* (Burt, 2022), and Jay Joseph has reviewed Plomin's *Blueprint*, *A Blueprint for Genetic Determinism* (Joseph, 2022), demonstrating its lack of scientific credibility.

As I noted in Baverstock (2024), in a study in Finland, PGSs for five common diseases and three complex traits were calculated for 2376 individuals whose parents had lived in a known specific geographical location. Within Finland, there is a well-defined genetic population structure, with an east-to-west divide (Kerminen et al., 2017). For all but one of the five disease traits and one of the three complex traits, the PGSs detected the geographic structure (indicating where the individual was born) and not the distribution of the trait (Kerminen et al., 2019). Few PGSs have been subjected to such stringent testing, so this study is a 'landmark' signalling the failure of the PGS (and GWA) concept.

Most recently (Hingorani et al., 2023) have examined the performance in population screening, individual risk assessment, and population risk stratification of 926 PGSs for 310 disease traits in the *Polygenic Score Risk Catalog*. The authors describe the performance in all three contexts as 'poor' but in the specific examples that they address, e.g., breast cancer, and coronary artery disease and stroke, PGSs perform no better than existing risk indicators such as age and blood pressure. That PGSs show any correlations with traits at all is perhaps surprising since I am arguing that PGSs are based purely on noise, i.e., that SNPs don't register *any* signal in GWA studies. However, as Ken Richardson points out, albeit, in the context of educational attainment (Richardson, 2017), there are underlying 'structures' in populations, due to migration, for example, which are clustered. These clusters may be typified by higher-than-average specific disease prevalences, thus potentially biasing PGSs in which they are included and leading to weak associations between a disease trait and the genetic signal from a cluster. This is the situation in the results of Kerminen et al. (2019), where the clustering is a geographical feature of Finland.

As already noted, genome-wide association-based research has been flourishing since 2005 (see FN 10) as has research on PRS/PGSs, again, particularly in the past two years. The funding of such research needs to be stringently assessed and the UK BioBank,[27] a collaborator in many GWA studies, needs to consider its obligation to the volunteers whose data it holds, and ensure that the research in which it participates has value for humankind.

6.2. Mathematisation in science

The positive value of mathematisation is illustrated in this paper (Abstract and Section 4) with the quotation of Newton's second law of motion in its full differential form. How else would we know that a rock dropped from the top of the Leaning Tower in Pisa had lost mass, albeit a minuscule amount, when it landed at the bottom of the tower? On the negative side is the construction of mathematical and statistical models, such as the model to invoke epistasis in the LTEE by Wiser et al. (2013) that duped JEB editor, Reuter (See Section 2). Here, mathematisation is a strategy being used to prop up a false interpretation of data.

Those are the extremes of the deployment of mathematisation. In between lie the highly mathematicised models, referred to above (in section 4), which populate a lot of physics and, to a lesser extent, biology, which would disappear if an alternative underlying premise were adopted; for example, that it is not the gene that is the unit of inheritance but rather it is the phenotypes of the parental gametes. There is a common misunderstanding that if some notion or idea can be expressed in mathematics it must be true. Mathematics can say something about the internal consistency of an idea, but it says nothing about how the idea relates to the real world. Quantum mechanics, which Longo claims to have a role in biology, is a case in point: as Annila shows, abandoning the concept of the void being a vacuum in cosmology, and replacing it with paired photons in line with Bose-Einstein statistics, can account for many of the features of modern physics that we find difficult to relate to reality, including, for example, dark matter (Annila, 2020; Annila and Wikström, 2022).

6.3. How should evolutionary theory evolve now?

The LTEE has falsified Fisher's genetical theory of natural selection, the foundation for mainstream evolutionary theory as it is understood in terms of Huxley's *Evolution the Modern Synthesis* (Huxley, 1942) and The Extended Evolutionary Synthesis.[28] In his book *Darwin's Legacy: What Evolution Means Today* (Dupré, 2003), John Dupré points out that by moving from the mere fact of *evolution* to recognising *evolution by natural selection,* we have entered the realm of 'a richly articulated causal theory.' He continues; 'Moreover, natural selection remains by far the most powerful—according to many the only—theory that provides an explanation for adaption of organisms to their environments' (emphasis added). Further, Dobzhansky famously wrote: 'Nothing in Biology makes sense except in the Light of Evolution' (1973), by which he meant Neo-Darwinian evolution as he was a founder member of the Modern Synthesis. What could be the basis for such confidence in Neo-Darwinian evolution as a theory? In 1969, Conrad Waddington published an essay entitled: *Paradigm for an Evolutionary Process* (reprinted as Waddington, 2008) in which he proposed that there needed to be both direct actions of the environment on the organism's phenotype, and the reverse, for the extraordinary diversity of life to have evolved. Seen in the context introduced above, of adaptation being when organism and environment are in free energy balance, both Waddington's requirement, and Edward Blyth's stipulation, made in 1835, that '[A]mong animals which procure their food by means of their agility, strength, or delicacy of sense, the one best organized must always obtain

[26] In a devastating criticism of the standard model of cosmology https://www.youtube.com/watch?v=XmzulJsGtZ4 the Swedish philosopher and cosmologist, Børn Ekeberg, points out, when asked why a so obviously inadequate 'cosmological framework' was still the basis for modern cosmology, that 'taking a "framework" away from a scientist is the same as making them blind'. Genetics has provided such a 'framework' for biology for more than 100 years and it will be difficult to replace.

[27] https://www.ukbiobank.ac.uk/.

[28] https://en.wikipedia.org/wiki/Extended_evolutionary_synthesis.

the greatest quantity; and must, therefore, become physically the strongest, and be thus enabled, by routing its opponents, to transmit its superior qualities to a greater number of offspring' (Blyth, 1835), make perfect sense. In contrast, if evolution were driven by genetic variation and Fisher's natural selection, with strict application of Crick's Central Dogma (Crick, 1970), any advance in the diversity of life beyond bacteria would be most unlikely, as noted above (Sections 2 and 6.1). To stress the point, throughout the mainstream development of this pivotal biological process, crucial ideas have been ignored, even while the problems of interpreting the lack of a relationship between mutational damage and adaptation, demonstrated by the LTEE, i.e., direct empirical evidence, were defeating the evolutionary biology research community, including the LTEE's investigators (Barrick et al., 2009).

A theory of evolution needs to explain how, for example, the different body forms have arisen from their progenitors. The LTEE has demonstrated that genetic variation cannot explain this, but neither can gene products, i.e., proteins, in terms of their molecular structure, solve the problem. For a start, many proteins in the cell cytoplasm are in a state of flux in terms of folding and unfolding (Fonin et al., 2018). Emil Fischer's 'lock and key' concept has long been rejected as a mechanism for enzyme action (Hammes, 2008), yet the evolution of protein structure is still considered a potential explanation for transitions to new species. Recent work on cephalopods suggests that the more highly developed nervous system of the octopus, compared to its progenitor, the squid, is partly due to genome reorganisation (Albertin et al., 2022). This is in line with much earlier experiments (Kashiwagi et al., 2006) in which bacteria with an incorporated synthetic bistable switch, which would enable them to use one of two alternative nutrients, A or B, were able to adapt rapidly (within hours) to B when A was replaced by B and vice versa. This phenomenon was attributed to the cell's ability to adopt appropriate 'adaptive attractors', i.e., rearranged genomes.

In the IA model, it is virtual *rules* that determine the relationship between a cell and its neighbours, and, therefore, determine *morphology*; rules that derive from the inherited RoE (Baverstock and Ronkko, 2014). The RoE is independent of the DNA, which is a separate thread of information coding for the gene products and extending back through evolutionary history. It is interactions among the gene products from which the cellular phenotype emerges, regulated by the RoE (Baverstock and Ronkko, 2008). For evolutionary changes in organismal phenotypes to occur would require changes in the RoE but not necessarily in the DNA (genes). In what way such changes could occur is an open question, but I suspect the answer involves violating Crick's Central Dogma.

In the IA model, Kauffman's concept of a 'fitness landscape' (Kauffman, 1993) is replaced by a free energy phenotypic 'state space' in which transitions are precipitated by environmental stresses, and 'unprestatable' trajectories follow energy gradients dictated by environmental factors. In that sense, evolution is lawless, as Kauffman and Longo maintained (see Section 4), but it is lawful in the context of thermodynamics, in that increasing entropy (through ingestion of nutrients) and concomitant decreases in free energy, i.e., Maupertuis's principle of least action, work towards maintaining the balance between organism and environment. That, I have proposed, is the first part of a two-part evolutionary process (Baverstock, 2022).

The second part entails organismal agency. Notwithstanding Frantisek Balušca's preference for consciousness emerging from the cell membrane (Baluska and Reber, 2021), as noted above (Section 1), I would opt for proteins being the source of consciousness and, therefore, the source of the ability of primitive organisms to learn and memorise i.e., have agency as discussed in Section 5 of (Baverstock, 2024). I share Balušca's view that prebiotic entities had a degree of consciousness, but it was more likely to be rooted in the 'content' of the protocell, rather than the 'container'. Secondly, as pointed out by Rosen (1991), proteins are not normal molecules: they come closer than any other biomolecule, including DNA,[29] to having lifelike properties, in that these can depend on environmental factors. In addition, mis-folded proteins, commonly associated with neurological diseases, self-replicate through seeding (Soto and Pritzkow, 2018). The above properties are one of the reasons why Karl Popper's statement that biochemistry cannot be reduced to chemistry holds good (Rose, 1988).

I propose that, given natural selection in biology as we have redefined it as a thermodynamic feature, over evolutionary time scales, organismal agency, natural genetic engineering (Shapiro, 2011), and 'genomic rearrangement'[30], could provide the scope for the evolution of the complex life forms, both extinct and extant that have evolved on this planet. The 'deposed gene' can be seen *not* as the efficient cause but *rather* as the material cause of organisms, providing the 'bricks', in the form of proteins, out of which organisms are made. The human gut microbiome contains more than nine million unique gene sequences (Yang et al., 2009), and this must be a small fraction of the total number of gene sequences comprising prokaryotes, so there are, therefore, plenty of 'bricks' to be adopted and 'naturally genetically engineered' through horizontal transfer into extant organisms to provide modifications. So, in what is proposed, Darwin's descent with modification still applies but the origin of the modification is not genetic but a combination of natural processes, none of which are new, but which have been ignored by evolutionists. In this way, we can see how life, initially simple, can progressively increase in complexity.[31]

7. Afterword

The instigator of the LTEE, Richard Lenski, did not set up the experiment to test a hypothesis but rather '[I] undertook the LTEE to ask some basic questions about the process of adaptation' (Lenski, 2017, 2023), however, given the stringent conditions under which the experiment has been conducted, the twelve independent replicate cultures, and the time it would take to repeat it in the guise of such a test, it is legitimate to regard it as a) falsifying Fisher's genetical theory of natural selection (Fisher, 1930) and the almost universally accepted dogma that evolution is driven by genetic variation, i.e., is a Mendelian process based on the Neo Darwinian concept of natural selection, the Modern Synthesis, and, b) showing that genes are not Mendel's units of inheritance. These ideas have pervaded biology (and penetrated other domains of science, including psychology and even quantum mechanics[32]) for decades: they are dogma in the sense of the word as it applies to religious faith and its critics are contemptuously dismissed as 'creationists'[33] by its adherents. Yet, given the invitation to defend their dogma by writing a commentary on (Baverstock, 2024), the apostles of the Modern and Extended Syntheses were missing. When challenged privately about their positions (in circumstances where they have made public statements about their ideas) evolutionists and geneticists have simply failed to reply. As the Editorial to this collection points out, in many cases, they are pleased enough to accept public money for their research, from the MRC, Wellcome Foundation, and the ESF, for

[29] It is sometimes claimed that DNA is self-replicating, but that is not true; to replicate, DNA requires proteins. In Section 4 it is noted that under some circumstances proteins can replicate through seeding.

[30] Variant, or alternative attractors based on the same genome.

[31] Interestingly, a group of US authors have recently proposed a new law of increasing functional information: Wong et al. (2023) to account for the evolution of complex systems by looking for equivalencies among evolving systems, including life.

[32] See: https://en.wikipedia.org/wiki/Quantum_Darwinism

[33] It is because of this that the website www.thethirdwayofevolution.com was set up to illustrate that there were viable alternatives to both the Modern Synthesis and creationism.

example. Some of the most excoriating criticism of Darwinism has come from the intelligent design (ID) advocate David Berlinski (see his essays *The Deniable Darwin*,[34] and *Has Darwin met his match* (Berlinski, 2002). Berlinski abhors (rightly) the narrative character of the evolutionary biological literature and has a special place in his criticism for Richard Dawkins, the foremost evangelist for Neo-Darwinism. Dawkins's role uncritically in propagating the dogma is especially egregious as his (well-written) books are widely sold in many languages and are likely to form the introduction to biology of uncountable numbers of children for decades to come. His lack of critical thinking,[35] given the extent to which, over decades, he has immersed himself in the dogma and used it to promulgate his contempt for 'religious belief', has been exceptionally damaging to science, especially as he held the Simonyi Professorship for the Public Understanding of Science at Oxford University for several years. That chair was endowed especially for Dawkins: 'to communicate science to the public without, in doing so, losing those elements of scholarship which constitute the essence of true understanding'.[36]

I would rather end on a positive note, i.e., about the work my colleagues and I are doing toward developing a new framework for biology. Whereas, what we reject, the traditional gene-centric view of biology which tells us little beyond its immediate application, we replace with thermodynamics, Maupertuis's principle of least action, and the importance of increasing entropy: this approach will enable us to address questions such as 'what the world is, how we know about it, what is the meaning of life, and how we should live' (Annila, 2023) as well as how life works.

As Carlo Rovelli writes in: *Anaximander and the Birth of Science* (Rovelli and Rosenberg, 2023), we need to urgently '*reconceive worlds*', specifically of biology and physics.

CRediT authorship contribution statement

Keith Baverstock: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I am grateful to Arto Annila, John Berriman, Steve Fromage, Oliver James, Mike Jones, and Ken Richardson for valuable comments during the preparation of this paper. I also wish to record my gratitude to colleagues, whether they agreed with me or not, over the years as the ideas presented here developed, including, John Ashby, David Berlinski, the late Bob Cundall, Jukka Juutilainen, Andre Karotki, the late Krishnaswami Sankaranarayanan (Sankar), Mike Thorne, Dillwyn Williams, members of the Department of Environmental and Biological Sciences at UEF in Kuopio, members of the ARCH project, and members of the Critical Genetics Forum.

References

Albertin, C.B., Medina-Ruiz, S., Mitros, T., Schmidbaur, H., Sanchez, G., Wang, Z.Y., Grimwood, J., Rosenthal, J.J.C., Ragsdale, C.W., Simakov, O., Rokhsar, D.S., 2022. Genome and transcriptome mechanisms driving cephalopod evolution. Nat. Commun. 13 (1), 2427.
Annila, A., 2020. Back to Reality. Privus Press, New York.
Annila, A., 2023. Philosophy of thermodynamics. Philos. Trans. A Math. Phys. Eng. Sci. 381 (2252), 20220281.
Annila, A., Baverstock, K., 2014. Genes without prominence: a reappraisal of the foundations of biology. J. R. Soc. Interface 11 (94), 20131017.
Annila, A., Baverstock, K., 2016. Discourse on order vs. disorder. Commun. Integr. Biol. 9 (4), e1187348.
Annila, A., Wikström, M., 2022. Dark matter and dark energy denote the gravitation of the expanding universe. Friont Phys 10 (14 October 2022).
Anonymous, 1903. Prof Jahannsen on heredity. Nature 69 (17 December), 149–150.
Asbury, K., Plomin, R., 2014. G Is for Genes: the Impact of Genetics on Education and Achievement. Wiley-Blackwell, Chichester, West Sussex.
Bailly, F., Longo, G., 2009. Biological organization and anti-entropy. J. Biol. Syst. 17, 63–96.
Baluska, F., Reber, A.S., 2021. Cellular and organismal agency - not based on genes: a comment on Baverstock. Prog. Biophys. Mol. Biol. 167, 161–162.
Baluska, F., Yokawa, K., Mancuso, S., Baverstock, K., 2016. Understanding of anesthesia - why consciousness is essential for life and not based on genes. Commun. Integr. Biol. 9 (6), e1238118.
Barrick, J.E., Yu, D.S., Yoon, S.H., Jeong, H., Oh, T.K., Schneider, D., Lenski, R.E., Kim, J. F., 2009. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature 461 (7268), 1243–1247.
Basener, W.F., Sanford, J.C., 2018. The fundamental theorem of natural selection with mutations. J. Math. Biol. 76 (7), 1589–1622.
Baverstock, K., 2000. Radiation-induced genomic instability: a paradigm-breaking phenomenon and its relevance to environmentally induced cancer. Mutat. Res. 454 (1–2), 89–109.
Baverstock, K., 2011. A comparison of two cell regulatory models entailing high dimensional attractors representing phenotype. Prog. Biophys. Mol. Biol. 106 (2), 443–449.
Baverstock, K., 2019. Polygenic scores: are they a public health hazard? Prog. Biophys. Mol. Biol. 149, 4–8.
Baverstock, K., 2021. The gene: an appraisal. Prog. Biophys. Mol. Biol. 164, 46–62.
Baverstock, K., 2022. Evolution in two parts: as seen in a new framework for biology. Theor Biol Forum 115 (1–2), 71–83.
Baverstock, K., 2024. The Gene: an appraisal. Prog. Biophys. Mol. Biol. 186, e73–e88.
Baverstock, K., Cundall, R.B., 1988. Solitons and energy transfer in DNA. Nature 332 (6162), 312–313.
Baverstock, K., Ronkko, M., 2008. Epigenetic regulation of the mammalian cell. PLoS One 3 (6), e2290.
Baverstock, K., Ronkko, M., 2014. The evolutionary origin of form and function. J. Physiol. 592 (11), 2261–2265.
Berlinski, D., 2002. Has Darwin met his match. Commentary 114 (5), 31–41.
Bertalanffy, L.v., 1969. General System Theory; Foundations, Development, Applications. G. Braziller, New York.
Blyth, E., 1835. An attempt to classify the 'varieties' of animals, with observations on the marked seasonal and other changes which naturally take place in various British species, and which do not constitute varieties. The Magazine of Natural History 8, 40–53.
Boltzmann, L., 1974. Theoretical Physics and Philosophical Problems : Selected Writings. Reidel Pub. Co, Dordrecht ; Boston.
Bratlie, M.S., Johansen, J., Sherman, B.T., Huang da, W., Lempicki, R.A., Drablos, F., 2010. Gene duplications in prokaryotes can be associated with environmental adaptation. BMC Genom. 11, 588.
Burt, C.H., 2022. Challenging the Utility of Polygenic Scores for Social Science: Environmental Confounding, Downward Causation, and Unknown Biology. Behav Brain Sci, pp. 1–36.
Crick, F., 1970. Central dogma of molecular biology. Nature 227 (5258), 561–563.
Csermely, P., Kunsic, N., Mendik, P., Kerestely, M., Farago, T., Veres, D.V., Tompa, P., 2020. Learning of signaling networks: molecular mechanisms. Trends Biochem. Sci. 45 (4), 284–294.
Davies, P., 2020. Does new physics lurk inside living matter? Phys. Today August, 35–40.
De la Fuente, I.M., 2021. Correspondence insights into the role of genes in cell functionality. Comments on "The gene: an appraisal" by K. Baverstock. Prog. Biophys. Mol. Biol. 167, 152–160.
De La Fuente, I.M., Martinez, L., Perez-Samartin, A.L., Ormaetxea, L., Amezaga, C., Vera-Lopez, A., 2008. Global self-organization of the cellular metabolic structure. PLoS One 3 (8), e3100.
Dobzhansky, T., 1973. Nothing in biology makes sense except in the light of evolution. Am. Biol. Teach. 35 (3), 125–129.
Dupré, J., 2003. Darwin's Legacy : what Evolution Means Today. Oxford University Press, Oxford ; New York.
Farnsworth, K.D., 2021. An organisational systems-biology view of viruses explains why they are not alive. Biosystems 200, 104324.
Farnsworth, K.D., 2022. How an information perspective helps overcome the challenge of biology to physics. Biosystems 217, 104683.
Fisher, R.A., 1918. The correlation between relatives on the supposition of Mendelian Inheritance. Trans. R. Soc. Edinb. 52, 399–433.
Fisher, R.A., 1930. The Genetical Theory of Natural Selection. The Clarendon press, Oxford.
Fonin, A.V., Darling, A.L., Kuznetsova, I.M., Turoverov, K.K., Uversky, V.N., 2018. Intrinsically disordered proteins in crowded milieu: when chaos prevails within the cellular gumbo. Cell. Mol. Life Sci. 75 (21), 3907–3929.
Frankel, N.W., Pontius, W., Dufour, Y.S., Long, J., Hernandez-Nunez, L., Emonet, T., 2014. Adaptability of non-genetic diversity in bacterial chemotaxis. Elife 3.
Galton, F., 1876. A theory of heredity. J. Anthropol. Inst. G. B. Ireland 5, 329–348.

[34] Available from the Discovery Institute: https://www.discovery.org/a/130/.
[35] See my comments on his analysis of the LTEE in Baverstock (2024). (p e79) which adequately testify to Dawkins' lack of critical thinking.
[36] See: https://en.wikipedia.org/wiki/Simonyi_Professor_for_the_Public_Understanding_of_Science.

Galton, F., 1885. Regression towards mediocrity in hereditary stature. J Anthropol Inst 15, 246–253.
Galton, F., 1897. The average contribution of each several ancestor to the total heritage of the offspring. PrProc. R. Soc. Lond. 61, 401–413.
Gerrish, P.J., Lenski, R.E., 1998. The fate of competing beneficial mutations in an asexual population. Genetica 102–103 (1–6), 127–144.
Goldstein, K., 1934. Der aufbau des organismus; einführung in die biologie unter besonderer berücksichtigung der erfahrungen am kranken menschen. Haag, M. Mijhof.
Halfmann, R., Lindquist, S., 2010. Epigenetics in the extreme: prions and the inheritance of environmentally acquired traits. Science 330 (6004), 629–632.
Hammes, G.G., 2008. How do enzymes really work? J. Biol. Chem. 283 (33), 22337–22346.
Heeney, C., 2021. Problems and promises: how to tell the story of a genome wide association study? Stud. Hist. Philos. Sci. 89, 1–10.
Hingorani, A.D., Gratton, J., Finan, C., Schmidt, A.F., Patel, R., Sofat, R., Kuan, V., Langenberg, C., Hemingway, H., Morris, J.K., Wald, N.J., 2023. Performance of polygenic risk scores in screening, prediction, and risk stratification: secondary analysis of data in the Polygenic Score Catalog. BMJ Med 2 (1), e000554.
Huxley, J., 1942. Evolution, the Modern Synthesis. G. Allen & Unwin ltd, London.
Johannsen, W., 1911. The genotype conception of heredity. Am. Nat. 45, 129–159.
Johnson, M.S., Reddy, G., Desai, M.M., 2023. Epistasis and evolution: recent advances and an outlook for prediction. BMC Biol. 21 (1), 120.
Joseph, J., 2022. A Blueprint for genetic determinism. Am. J. Psychol. 135 (4), 442–454.
Kadhim, M.A., Macdonald, D.A., Goodhead, D.T., Lorimore, S.A., Marsden, S.J., Wright, E.G., 1992. Transmission of chromosomal instability after plutonium alpha-particle irradiation. Nature 355 (6362), 738–740.
Kashiwagi, A., Urabe, I., Kaneko, K., Yomo, T., 2006. Adaptive response of a gene network to environmental changes by fitness-induced attractor selection. PLoS One 1, e49.
Kauffman, S.A., 1993. The Origins of Order: Self Organisation and Selection in Evolution. Oxford University Press, Oxford.
Kauffman, S.A., 2000. Investigations. Oxford University Press, Oxford ; New York.
Kerminen, S., Havulinna, A.S., Hellenthal, G., Martin, A.R., Sarin, A.P., Perola, M., Palotie, A., Salomaa, V., Daly, M.J., Ripatti, S., Pirinen, M., 2017. Fine-scale genetic structure in Finland. G3 (Bethesda) 7 (10), 3459–3468.
Kerminen, S., Martin, A.R., Koskela, J., Ruotsalainen, S.E., Havulinna, A.S., Surakka, I., Palotie, A., Perola, M., Salomaa, V., Daly, M.J., Ripatti, S., Pirinen, M., 2019. Geographic variation and bias in the polygenic scores of complex diseases and traits in Finland. Am. J. Hum. Genet. 104 (6), 1169–1181.
Kocher, C.D., Dill, K.A., 2023. Darwinian evolution as a dynamical principle. Proc. Natl. Acad. Sci. U.S.A. 120 (11), e2218390120.
Lehmonen, L., Annila, A., 2022. Baryon breakdown in black hole. Front. Physiol. 10, 954439.
Lenski, R.E., 2023. Revisiting the design of the long-term evolution experiment with Escherichia coli. J. Mol. Evol. 91 (3), 241–253.
Lenski, R.E., 2017. What is adaptation by natural selection? Perspectives of an experimental microbiologist. PLoS Genet. 13 (4), e1006668.
Lenski, R.E., Travisano, M., 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc. Natl. Acad. Sci. U.S.A. 91 (15), 6808–6814.
Lewontin, R.C., 1974. The Genetic Basis of Evolutionary Change. Columbia University Press, New York.
Lewontin, R.C., 1992. The Doctrine of DNA: Biology as Ideology. Penguin Books Ltd, London, England.
Longo, G., 2023. From information to physics to biology. Prog. Biophys. Mol. Biol. 177, 202–206.
Longo, G., Montevil, M., 2012. The inert vs. the living state of matter: extended criticality, time geometry, anti-entropy - an overview. Front. Physiol. 3, 39.
Longo, G., Montevil, M., Kauffman, S.A., 2012. No entailing laws, but enablement in the evolution of the biosphere. In: Genetic and Evolutionary Computation Conference. GECCO 2012, Philadelphia, PA, USA.
Lorimore, S.A., Kadhim, M.A., Pocock, D.A., Papworth, D., Stevens, D.L., Goodhead, D.T., Wright, E.G., 1998. Chromosomal instability in the descendants of unirradiated surviving cells after alpha-particle irradiation. Proc. Natl. Acad. Sci. U.S.A. 95 (10), 5730–5733.
Maddamsetti, R., 2021. Selection maintains protein interactome resilience in the long-term evolution experiment with Escherichia coli. Genome Biol Evol 13 (6).
Maddamsetti, R., Grant, N.A., 2020. Divergent evolution of mutation rates and biases in the long-term evolution experiment with Escherichia coli. Genome Biol Evol 12 (9), 1591–1603.
Makela, T., Annila, A., 2010. Natural patterns of energy dispersal. Phys. Life Rev. 7 (4), 477–498.
Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S., Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F., McCarroll, S.A., Visscher, P.M., 2009. Finding the missing heritability of complex diseases. Nature 461 (7265), 747–753.
Marais, A., Adams, B., Ringsmuth, A.K., Ferretti, M., Gruber, J.M., Hendrikx, R., Schuld, M., Smith, S.L., Sinayskiy, I., Kruger, T.P.J., Petruccione, F., van Grondelle, R., 2018. The future of quantum biology. J. R. Soc. Interface 15 (148).
McKenna, K.Z., Gawne, R., Nijhout, H.F., 2022. The genetic control paradigm in biology: what we say, and what we are entitled to mean. Prog. Biophys. Mol. Biol. 169–170, 89–93.
Moger-Reischer, R.Z., Glass, J.I., Wise, K.S., Sun, L., Bittencourt, D.M.C., Lehmkuhl, B.K., Schoolmaster Jr., D.R., Lynch, M., Lennon, J.T., 2023. Evolution of a minimal cell. Nature 620 (7972), 122–127.
Moore, J.H., 2003. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum. Hered. 56 (1–3), 73–82.
Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T., Kondo, T., 2005. Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science 308 (5720), 414–415.
Nicolis, G., Prigogine, I., 1989. Exploring complexity : an introduction. In: W.H. Freeman, New York.
Niemann, H.-J., 2014. Karl Popper and the Two New Secrets of Life: Including Karl Popper's Medawar Lecture 1986 and Three Related Texts. Tubingen, Mohr Siebeck.
Noble, D., 2017. Dance to the Tune of Life : Biological Relativity. Cambridge University Press, Cambridge ; New York.
Paneth, N., Vermund, S.H., 2018. Human molecular genetics has not yet contributed to measurable public health advances. Perspect. Biol. Med. 61 (4), 537–549.
Pearson, K., Yule, G.U., Blanchard, N., Lee, A., 1903. The law of ancestral heredity. Buiometrika 2 (2), 211–236.
Phillips, P., 2008. Epistasis - the essential role of gene interactions in the structure and evolution of genetic systems. Nat. Rev. Genet. 9, 855–867.
Phizicky, E.M., Fields, S., 1995. Protein-protein interactions: methods for detection and analysis. Microbiol. Rev. 59 (1), 94–123.
Plomin, R., 2018. Blueprint : How DNA Makes Us Who We Are. The MIT Press, Cambridge, MA.
Plomin, R., von Stumm, S., 2022. Polygenic scores: prediction versus explanation. Mol. Psychiatr. 27 (1), 49–52.
Provine, W.B., 2001. The Origins of Theoretical Population Genetics. University of Chicago Press, Chicago.
Richardson, K., 2017. GWAS and cognitive abilities: why correlations are inevitable and meaningless: GWA studies to identify genetic factors for educational achievements largely ignore underlying social structures and dynamics. EMBO Rep. 18 (8), 1279–1283.
Richardson, K., 2021. "Genes and knowledge: response to Baverstock, K. the gene an appraisal. Prog. Biophys. Mol. Biol. 167, 12–17. https://doi.org/10.1016/j.pbiomolbio.2021.04.005.
Roll-Hansen, N., 2022. A special role for the genotype? Some comments on Keith Baverstock: "The gene: an appraisal. Prog. Biophys. Mol. Biol. 172, 82–89.
Rose, S., 1988. Reflections on reductionism. Trends Biochem. Sci. 13 (5), 160–162.
Rose, S.P.R., 2023. Commentary on Baverstock; the gene; an appraisal. Prog. Biophys. Mol. Biol. 178, 70–71.
Rosen, R., 1991. Life Itself: a Comprehensive Inquiry into the Nature, Origin and Fabrication of Life. Columbia University Press, New York.
Rovelli, C., Rosenberg, M.L., 2023. Anaximander and the Birth of Science. Riverhead Books, New York.
Salek, M.M., Carrara, F., Fernandez, V., Guasto, J.S., Stocker, R., 2019. Bacterial chemotaxis in a microfluidic T-maze reveals strong phenotypic heterogeneity in chemotactic sensitivity. Nat. Commun. 10 (1), 1877.
Sapp, J., 1987. Beyond the Gene : Cytoplasmic Inheritance and the Struggle for Authority in Genetics. Oxford University Press, New York.
Scott, A.C., 2007. The Nonlinear Universe: Chaos, Emergence and Life. Springer, Berlin.
Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.N.M., Campbell, A., Devito, L., Ilic, D., Khalaf, Y., Niakan, K.K., Fishel, S., Zernicka-Goetz, M., 2016. Self-organization of the human embryo in the absence of maternal tissues. Nat. Cell Biol. 18 (6), 700–708.
Shahbazi, M.N., Siggia, E.D., Zernicka-Goetz, M., 2019. Self-organization of stem cells into embryos: a window on early mammalian development. Science 364 (6444), 948–951.
Shapiro, J.A., 2011. Evolution : a View from the 21st Century. FT Press Science, Upper Saddle River, N.J.
Sharma, V., Annila, A., 2007. Natural process–natural selection. Biophys. Chem. 127 (1–2), 123–128.
Smolin, L., 2006. The Trouble with Physics: the Rise of String Theory, the Fall of a Science, and what Comes Next. Houghton Mifflin, Boston.
Soto, C., Pritzkow, S., 2018. Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. Nat. Neurosci. 21 (10), 1332–1340.
Vries, H.d., 1901. Die Mutationstheorie, vol. I. Leipszig, Von Weit.
Waddington, C.H., 2008. Paradigm for an evolutionary process. Biological Theory 3 (3), 258–266.
Weldon, W.F.R., Pearson, K., 1903. Inheritance in phaseolus vulgaris. Biometrika 2, 499–503.
Wiser, M.J., Ribeck, N., Lenski, R.E., 2013. Long-term dynamics of adaptation in asexual populations. Science 342 (6164), 1364–1367.
Wong, M.L., Cleland, C.E., Arend Jr., D., Bartlett, S., Cleaves 2nd, H.J., Demarest, H., Prabhu, A., Lunine, J.I., Hazen, R.M., 2023. On the roles of function and selection in evolving systems. Proc. Natl. Acad. Sci. U.S.A. 120 (43), e2310223120.
Yang, X., Xie, L., Li, Y., Wei, C., 2009. More than 9,000,000 unique genes in human gut bacterial community: estimating gene numbers inside a human body. PLoS One 4 (6), e6074.
Yule, G.U., 1904. Prof. Johannsen on heredity. Nature 69 (January 2004), 223–224.
Zitnik, M., Sosic, R., Feldman, M.W., Leskovec, J., 2019. Evolution of resilience in protein interactomes across the tree of life. Proc. Natl. Acad. Sci. U.S.A. 116 (10), 4426–4433.