

# What is a Lineage?

Celso Neto

**Abstract:** This paper defends lineage pluralism; the view that biological lineages are not a single, unified type of entity. I analyze aspects of evolutionary theory, phylogenetics, and developmental biology to show that these areas appeal to distinct notions of lineage. Based on this information, I formulate two arguments for lineage pluralism. These arguments undercut the main motivations for lineage monism, namely, the view that biological lineages are a single, unified type of entity. Though this view is rarely made explicit, it is often assumed in philosophy and biology. Hence, this paper sheds light on the implicit monistic assumption, and shows why lineage pluralism should be adopted instead.

Keywords: Lineages, Pluralism, Lateral Gene Transfer, Prokaryotes, Stem Cells

## 1. Introduction

Philosophers of biology rarely discuss what lineages are, but they often refer to lineages in their analyses of natural selection, species, and other topics (e.g., Hull 1980; Ereshefsky 1992; Haber 2012). Many of these analyses would benefit from a careful discussion of the defining properties of lineages. For example, philosophers assume that an evolutionary individual is an entity that can form lineages, but they do not discuss what lineages are and how they are formed (Godfrey-Smith 2011; Clarke 2016). Without discussion of lineages, it is hard to pinpoint what an evolutionary individual is, and to what extent philosophers disagree about that issue.

Discussing the properties of lineages matters because lineages are important for biology. As I will show in this paper, lineages play significant roles in evolution and development (Hull 1980; Stent 1987). They are also key to the practice of phylogenetic reconstruction (Haber 2012). If philosophers aim to characterize the entities (i.e. ontological commitments) present in biological

theories and practices, then they should investigate lineages as much as they investigate genes, species, and biological individuals.

In this paper, I investigate what lineages are. I defend lineage pluralism, which is the view that lineages are not a single, unified type of entity. There are many correct notions of lineages, and these notions correspond to distinct types of lineages. In section 2, I introduce the opposite view, namely lineage monism, as an implicit assumption held by philosophers and biologists. In the sections 3 and 4, I discuss various uses of ‘lineage’ in the contexts of evolutionary theory, phylogenetics, and developmental biology. Different uses of ‘lineage’ favor a pluralistic conception of lineages. In section 5, I show how these arguments undermine basic motivations behind lineage monism.

## **2. Lineages Monism**

Lineages are sequences of biological entities connected by ancestry-descent relationships (Hull 1980). A sequence containing myself, my father, and my grandfather is a lineage because it is a single, direct line of descent among organisms. But biologists do not only discuss lineages of organisms. They are also interested in lineages of genes, cells, and species (Stent 1987; de Queiroz 1999; Degnan and Rosenberg 2006). The common feature among these lineages is that they are continuous lines of descent.

Prominent characterizations of evolution suggest that lineages are a single type of entity (Lewontin 1970; Hull 1980). Lineages are a single type of entity because they play a single role in evolution, namely, units of evolution. Lineages are units that ‘evolve’ in the sense that they accumulate and manifest evolutionary changes over time. This idea of lineage as units of evolution is supported by how biologists and philosophers describe natural selection and other

evolutionary processes (Ereshefsky 2001). These processes require reproduction and trait transmission. Biological entities form lineages, which allow traits to pass along and selected variations to accumulate through time. Significant evolutionary changes happen at different generations of a lineage, and this suggests that lineages are the entities evolving.

Peter Godfrey-Smith's approach to natural selection suggests that lineages are a single type of entity (2009). Godfrey-Smith characterizes selection by appealing to "Darwinian individuals." These individuals are reproducing biological entities, and reproduction is a complex phenomenon involving bottleneck, integration, and germ-soma separation (2009, 91-95). Each reproductive feature varies in degree from one individual to the other. The degree to which individuals satisfy those three features determines the degree to which they form clearly individuated lineages (Booth 2014). The more clearly individuated a lineage is, the more clearly it functions as a unit of evolution. Hence, relevant differences among lineages in natural selection are differences in degree, not type. Lineages are composed of the same type of reproducing entity (Darwinian individuals) and are units of evolution to some degree.

Some characterizations of phylogenetic systematics also suggest that lineages are a single type of entity. The goal of this field is to represent evolutionary history; a representation that usually has the form of a phylogenetic tree, or a tree-like-structure, in which lineages branch and generate new lineages (Baum and Smith 2013). The tree-like structure is divided into 'clades,' which are groups containing a common ancestor and all its descent lineages. One assumption in phylogenetics is that evolution naturally produces clades, and thus evolutionary history is a history of clade formation (Hennig 1966, 154-155). This assumption requires not only a single notion of clade, but also a single notion of lineage according to which lineages are branches that form clades. Otherwise, if lineages are different types of entities, some of them might have

characteristics such that they cannot be part of clades. In this case, evolutionary history would not have a tree-like-structure entirely divided into clades. This would be an undesirable consequence for many phylogeneticists (Baum and Smith 2013; c.f., Doolittle and Baptiste 2007). However, this is not the only monistic assumption concerning lineages in phylogenetics.

Kevin de Queiroz also assumes that lineages are a single type of entity in phylogenetics (1999). The assumption is implicit in his defense of species monism, which is the view that there is one correct species concept in biology. De Queiroz defines species as “segments of population-level lineages,” and claims that this definition underlies all characterizations of species offered by contemporary biologists (1999, 52). All species are comparable lineage segments, and figure as branches and terminal taxa in a phylogenetic tree. De Queiroz defends this unified notion of species by assuming a unified notion of lineages. He assumes that lineages are direct lines of descent that can be segmented based on branching and terminal events. The segmentation of lineages produces comparable units (i.e. species), but this is only the case if lineages are themselves comparable and segmented in commensurable ways. In this sense, species monism depends on the idea that lineages are a single type of entity.

Notably, lineage monism is not explicitly argued for in biology or philosophy. Rather, lineage monism is either an assumption or a suggestion drawn from how Hull, Godfrey-Smith, de Queiroz, and other scholars approach natural selection, phylogenetics, species concepts, and other topics. These approaches reveal three basic motivations for lineage monism. First, lineage monism is appealing because lineages perform a single, unified role as units of evolution. This role is widely accepted in many fields of biology, including phylogenetics. Second, lineage monism is implied by the appealing idea that evolutionary history is entirely divided into clades. According to this idea, lineages should be a single type of entity that form clades. And finally,

lineage monism is appealing because the defining features of lineages are multiple-realizable. The ancestor-descent relationships that define lineages are produced in various ways and at different levels of organization. Nonetheless, this variety is irrelevant insofar it produces the same function of lineages in evolution: lineages as units of evolution. In what follows, I challenge monism by presenting two arguments in favor of lineage pluralism.

### **3. The Open Lineage Argument for Pluralism**

Prokaryotes are receiving increasing attention from biologists and philosophers (Doolittle and Zhaxinbayeva 2013). These unicellular organisms do not contain membrane-bounded nuclei and organelles. They form lineages by dividing into identical copies and become genetically diverse because of two processes: mutation and lateral gene transfer (hereafter LGT). Mutations create genetic variation in the DNA due to the insertion and deletion of nucleotides, replication errors, and various other processes. Such genetic variation is then transmitted from progenitor to offspring within a lineage. By contrast, LGT does not create genetic variation, but it promotes genetic diversity because it transmits genetic material between lineages. This transmission usually occurs by physical contact between prokaryotic organisms or using a vector (Zhaxinbayeva and Doolittle 2011).

According to Boucher and Baptiste (2009), genetic diversity in prokaryotes produces two types of prokaryotic lineages. On the one hand, some lineages accumulate more changes due to LGT than mutation over generations. The genes within these lineages do not have a common origin, and the patterns of cell divisions do not match the patterns of gene transmission. These prokaryotic lineages are called “open lineages” (2009, 532). On the other hand, some lineages accumulate more changes due to mutation than LGT over generations. Most genes within these

lineages have a common origin, such that the pattern of cell divisions matches the pattern of genetic transmission. These lineages are called “closed lineages” (2009, 532).

Boucher and Baptiste (2009) claim that the distinction between open and closed lineages explains a limitation of current phylogenetic practices in representing evolutionary history. This limitation results from phylogenetics not allowing for open lineages. Phylogenetics depends on the assumption that, for any matrix of characters, the number of homologies (i.e. characters with common origin) must exceed the number of homoplasies (i.e. characters without common origin). Open lineages can have more homoplasies than homologies. The reason is that LGT produces homoplasies, and open lineages are characterized by rampant LGT. Hence, phylogenetic methodology rules out open lineages. The consequence is that phylogenetics has an elevated risk of incorrectly understanding evolutionary relationships among prokaryotes (Boucher and Baptiste 2009, 533). Phylogenetics cannot account for the complexity of ancestry relations generated by LGT, and thus is seriously limited in its ability to represent evolutionary history.

The distinction between closed and open lineages also explains features of evolution. Empirical studies show how open and closed lineages have very distinct ways of evolving through time (Andam and Gogartem 2011). LGT allows genes to move in and out of open lineages through time and this movement favors highly dynamic genomes and fast evolutionary changes (Doolittle and Brunet 2016). By contrast, closed lineages constrain the ways in which genes can move, and might disfavor dynamic genomic and fast evolutionary changes. This contrast explains why some prokaryotes have a superior capacity for fast adaptation. This capacity for fast adaptation evolved in open lineages due to the high amount of LGT they are exposed to.

The distinction between open and closed lineages was originally applied to prokaryotes, but it can be very useful when expanded to include eukaryotes. Thus, the expanded distinction explains aspects of both prokaryote and eukaryote evolution. The majority – if not all – of eukaryotic lineages are closed ones because mutations are more important than LGT in these lineages. This fact explains why eukaryotic lineages usually display the evolutionary pattern of successive bifurcations, since bifurcations are produced by closed lineages. By contrast, many prokaryotic lineages generate complex evolutionary patterns because they are open lineages (Doolittle and Baptiste 2007). The difference in evolutionary patterns often observed between prokaryotes and eukaryotes exists because these organisms form open and closed lineages, respectively.

The distinction between open and closed lineages grounds my first argument for lineage pluralism. I call this argument *The Open Lineage Argument for Pluralism*. As we have seen, open and closed lineages differ in many respects. Recall that they are associated with distinct sources of genetic diversity and distinct patterns of evolution. Recognition of such distinct patterns is important to biological theories and practices. If biologists distinguish between open and closed lineages, they can explain limitations of the phylogenetic method, and the evolutionary patterns of prokaryotes and eukaryotes, among other things. The argument continues with a general claim: ontological distinctions among lineages should only be assumed insofar as they reflect important distinctions for biological theory and practice. Since the distinction between open and closed lineages is important for biology, *open and closed lineages are distinct types of lineages*. Therefore, lineage pluralism is correct. This is not the only argument one can make in favour of lineage pluralism.

#### **4. The Developmental Lineage Argument for Pluralism**

In the previous section, I discussed notions of lineage in the contexts of evolutionary biology and phylogenetics. These contexts provide a basis for how many biologists and philosophers characterize lineages. Nevertheless, lineages also appear in other fields, such as developmental biology (Stent 1987; Fagan 2013). Historical evidence shows that lineages play an important role in developmental studies since the 19<sup>th</sup> Century (Droescher 2014). These studies provide an argument in support of lineage pluralism.

Developmental biology investigates the processes and phases that organisms undergo from birth to maturity (Gilbert 2000). In this context, biologists are particularly interested in lineages at the cellular level, and they use them to analyze cellular processes and phases. This analysis relies on the practice of ‘cell lineage tracing,’ which consists of keeping track of sequences of cell divisions (Kretzschmar and Watt 2012). Biologists employ biomarkers and other techniques to trace every cell division that shares a common cellular ancestor. The basic assumption of this practice is that cell division underlies development and thus give us some information about how development occurs.

To better understand the role of cell lineage tracing in developmental biology, consider the following example. Cell lineage tracing is frequently used to describe the history of cellular changes from an undifferentiated cell (i.e. a stem cell) to a set of specialized descent cells in the mature organism. This is the case when biologists track the origin of adult epidermal cells in mammalian organisms (Jensen et al. 2010). Epidermal cells are unable to divide and are continuously lost during the adult life. For this reason, they are continuously produced by stem cells. Cell lineage tracing helps biologists understand which groups of stem cells give rise to particular epidermal cell layers. It also helps understand the properties of particular groups of

stem cells, such as their potentiality—the capacity for specialization. Therefore, cell lineage tracing is not simply a way to map cell divisions, but it is a way to study features of stem cells and how they determine development.

The practice of cell lineage tracing reveals a specific role for lineages in development, and this role differs from how lineages function in evolution and phylogenetics. Cell lineages track developmental phases because they track the cellular origin of particular tissues, organs, and so on. Furthermore, these lineages manifest the properties of the stem cells that initiate them (Fagan 2013). This role differs from how lineages are depicted by evolutionary biology because cell lineages are not necessarily units of evolution. In a developmental context, the role of these lineages is not to accumulate and manifest changes resulting from random mutation, selection, and other evolutionary processes. Instead, developmental biologists treat cell lineages as entities that follow certain predictable and robust sequences of changes leading to specialized adult cells.

Cell lineages in development are also very different from lineages in phylogenetics. When developmental biologists discuss cell lineages, they are interested in somatic cells (e.g., cells composing an organism's body). Phylogeneticists have little interest in these sorts of cells, as they do not play a role in the understanding of clades, species, or phylogenetic trees. Additionally, phylogeneticists and developmental biologists have different definitions of 'lineage.' Cell lineages in development are branching structures containing a common cellular ancestor and all its descents. This definition is equivalent to the notion of clades in phylogenetics, but not to the notion of lineages. Phylogeneticists define clades as groups containing an ancestor and all its descents, but they define lineages as groups containing an ancestor and some but not all of its descents (de Queiroz 1999).

This discussion of lineages in developmental biology is far from exhaustive. Nevertheless, it is sufficient to ground another argument in favor of lineage pluralism. I call this argument *The Developmental Lineages Argument for Pluralism*. We know that the role and definitions of lineages in development biology are distinct from those in evolution and phylogenetics. Developmental lineages are sequences that track development, whereas lineages in evolutionary biology and phylogenetics are both units of evolution. This difference is important for biology theory and practice. For example, because evolutionists and phylogeneticists are interested in units of evolution, they rarely focus on somatic cell lineages. Contrastively, developmental biologists focus on somatic cell lineages because development happens in these cells. The difference between developmental lineages and lineages as units of evolution also explains why developmental biologists are not interested in mutations when doing cell lineage tracing. Mutations is important when understanding how lineages evolve, but it is not important when understanding how cell lineages contribute to development. Again, ontological distinctions among lineages should only be assumed insofar as they reflect important distinctions for biological theory and practice. Since the distinction between developmental lineages and lineages as units of evolution is important, the argument infers that developmental lineages are a distinct type of lineage. Therefore, lineage pluralism is correct.

## **5. Lineage Pluralism and Lineage Monism**

My two arguments for lineages pluralism share two basic and related features. First, they are arguments for ontological, rather than epistemological, pluralism. The conclusion of these arguments is about the world, rather than about the limits of human knowledge. Second, the arguments share the premise that ontological distinctions among lineages reflect important distinctions for biological theory and practice. I have shown that biologists (i) distinguish

between open and closed lineages; and (ii) treat developmental lineages separately from lineages as units of evolution. Given the importance of these distinctions to biological theory and practice, biologists are ontologically committed to them, and thus, to lineage pluralism. Though there are three possible motivations for lineage monism, in what follows I show that such motivations fail to capture those ontological commitments.

One motivation for monism is the idea that there is a general notion of lineage (or sequences of ancestor-descent relationships), which is multiply-realized. This motivation is weak. It is not that a general definition of lineages is wrong, the problem is that the generality glosses over important distinctions found in both biological theory and practice. For example, the distinction between open and closed lineages is biologically relevant and explain facts about patterns of evolution, adaptation, and the limitations of phylogenetics. If we adopt a general notion of lineage, we run the risk of overlooking those salient differences. Likewise, we run the risk of overlooking the fact that current phylogenetic methods are not adequate for capturing most of the prokaryotic evolution (Boucher and Baptiste 2009).

Another motivation for lineage monism is the idea that evolutionary history is one of clade formation, such that clades give rise to a tree-like structure representing evolutionary history. This motivation implies that there should be one unified notion of lineages and that lineages form clades. But there are good reasons to reject this motivation for lineage monism. If one accepts the distinction between open and closed lineages, then there is a possibility that many lineages (e.g. open lineages) do not even form clades. The consequence is that an adequate representation of the evolutionary history is not exhausted by clades and lineages that form clades (Doolittle and Baptiste 2007).

Finally, monism might be motivated by the idea that lineages have certain basic roles in biological theories and practices. When using the term ‘lineage,’ evolutionary-minded biologists and philosophers might think that they all refer to the same type of entity, namely units of evolution. However, not all biology is best conceived in light of evolution. Developmental biologists treat lineages as entities manifesting non-evolutionary changes in cells, and the properties of an ancestor stem cell. Hence, they do not treat lineages as units of evolution. More generally, lineages do not play a singular role in biology. Moreover, to privilege the role of lineage as units of evolution as the defining feature of lineages is an unjustified bias (Fagan 2013).

## **6. Conclusion**

In this paper, I provided two arguments for lineage pluralism. First, the open lineage argument for pluralism states that important explanations in biology rely on the distinction between open and closed lineages. This distinction identifies two types of lineages. Second, I appealed to uses of ‘lineage’ in developmental biology, and emphasized practices that rely on a non-evolutionary notion of lineage. Therefore, we should accept developmental lineages as a separate type of lineage from lineages as units of evolution. There is, then, no single correct answer to the question, “what is a lineage?” Rather lineage pluralism better captures the many different uses of ‘lineage’ in biology.

## **References**

- Andam, Cheryl P., and J. Peter Gogarten. 2011. "Biased gene transfer and its implications for the concept of lineage." *Biology direct* 6: 47.
- Booth, Austin. 2014. "Populations and individuals in heterokaryotic fungi: a multilevel perspective." *Philosophy of Science* 81: 612-632.

- Boucher, Yan, and Eric Bapteste. 2009. "Revisiting the concept of lineage in prokaryotes: a phylogenetic perspective." *Bioessays* 31: 526-536.
- Clarke, Ellen. 2013. "The multiple realizability of biological individuals." *The Journal of Philosophy* 110: 413-435.
- Clarke, Ellen. 2016. "Levels of selection in biofilms: multispecies biofilms are not evolutionary individuals." *Biology & Philosophy* 31: 191-212.
- De Queiroz, Kevin. 1999. "The general lineage concept of species and the defining properties of the species." In *Species: new interdisciplinary essays*. MIT Press.
- Doolittle, W. Ford, and Eric Bapteste. 2007. "Pattern pluralism and the Tree of Life hypothesis." *Proceedings of the National Academy of Sciences* 104: 2043-2049.
- Doolittle, W. Ford, and Olga Zhaxynbayeva. 2013. "What Is a Prokaryote?." In *The Prokaryotes*, pp. 21-37. Springer Berlin Heidelberg.
- Doolittle, W. Ford, and Tyler DP Brunet. 2016. "What is the tree of life?." *PLoS genetics* 12: e1005912.
- Dröscher, Ariane. 2014. "Images of cell trees, cell lines, and cell fates: the legacy of Ernst Haeckel and August Weismann in stem cell research." *History and philosophy of the life sciences* 36: 157-186.
- Ereshefsky, Marc. 1992. "Eliminative Pluralism." *Philosophy of Science* 59: 671-690.
- Ereshefsky, Marc. 2001. *The poverty of the Linnaean hierarchy: A philosophical study of biological taxonomy*. Cambridge University Press.
- Fagan, Melinda. 2013. *Philosophy of stem cell biology: Knowledge in flesh and blood*. Palgrave Macmillan.
- Gilbert, Scott F. 2000. *Developmental biology*. Sunderland, Mass: Sinauer Associates
- Godfrey-Smith, Peter. 2009. *Darwinian populations and natural selection*. Oxford University Press.
- Godfrey-Smith, Peter. 2011. "Darwinian individuals." In *From groups to individuals: evolution and emerging individuality*. Oxford University Press.
- Haber, Matthew. 2012. "Multilevel lineages and multidimensional trees: The levels of lineage and phylogeny reconstruction." *Philosophy of Science* 79: 609-623.
- Hull, David L. 1980. "Individuality and selection." *Annual review of ecology and systematics* 11: 311-332.
- Jensen, Kim B., Ryan R. Driskell, and Fiona M. Watt. 2010. "Assaying proliferation and differentiation capacity of stem cells using disaggregated adult mouse epidermis." *Nature protocols* 5: 898-911.
- Kretzschmar, Kai, and Fiona M. Watt. 2012. "Lineage tracing." *Cell* 148: 33-45.

Lewontin, Richard C. 1970. "The units of selection." *Annual review of ecology and systematics* 1: 1-18.

Stent, G. S. 1987. "Cell lineage in development." *FEBS letters* 215: 1-8.

Zhaxynbayeva, Olga, and W. Ford Doolittle. 2011. "Lateral gene transfer." *Current Biology* 21: R242-R246.