


Chapter 10

Paleontology: Outrunning Time

View metadata, citation and similar papers at core.ac.uk

brought to you by  CORE

John Huss

Abstract In this paper, I discuss several temporal aspects of paleontology from a philosophical perspective. I begin by presenting the general problem of “taming” deep time to make it comprehensible at a human scale, starting with the traditional geologic time scale: an event-based, relative time scale consisting of a hierarchy of chronological units. Not only does the relative timescale provide a basis for reconstructing many of the general features of the history of life, but it is also consonant with the cognitive processes humans use to think about time. Absolute dating of rocks, fossils, and evolutionary events (such as branching events on the tree of life) can be accomplished through the use of radiometric dating, chronological signals extractable from fossil growth patterns, and the “molecular clock.” Sometimes these different methods of absolute dating, which start from largely independent assumptions and evidentiary bases, converge in their temporal estimates, resulting in a consilience of inductions. At other times they fail to agree, either because fossils and molecules are giving temporal information about different aspects of nature and should not be expected to agree, or because of flawed assumptions that give rise to an inaccurate estimate. I argue that in general, despite the fact that it can be difficult to integrate disparate kinds of evidence, the principle of total evidence should be applied to the dating of evolutionary events. As a historical science, paleontology studies past events we cannot observe directly. This raises questions of epistemic access, meaning that due to the fragmentary nature of the fossil record we may find ourselves without access to the relevant traces to adjudicate between rival hypotheses about the past. The problems and prospects of epistemic access are explored through a case study of the reconstruction of the colors of dinosaurs. The paper closes with a reflection on the Darwin-Lyell metaphor of the fossil record as a highly fragmentary history book, and a call for a reconsideration of the book metaphor in favor of a systems view of the geologic and fossil records.

J. Huss (✉)

Department of Philosophy, The University of Akron, Akron 44325-1903, OH, USA
e-mail: jh76@uakron.edu

© Springer International Publishing AG 2017
C. Bouton, P. Huneman (eds.), *Time of Nature and the Nature of Time*, Boston
Studies in the Philosophy and History of Science 326,
DOI 10.1007/978-3-319-53725-2_10

211

10.1 Introduction

Within evolutionary biology, time has long been considered the province of the paleontologist. Strictly speaking, this is not true. Time itself is just as elusive for the paleontologist as it is for anybody else. Paleontologists do not traffic in time itself, but rather in events arrayed in time, in processes occurring over time, and in techniques for the measurement and representation of time. That said, paleontology does have something distinctive to contribute to our understanding of the deep past, namely its ability to draw justifiable inferences about the temporal sequence of events in the history of life. Although absolute dating based on radioactive decay of atomic nuclei and other physical methods have made it possible to get better and better estimates of the ages of rocks and fossils, paleontology throughout its history has reconstructed the history of life without explicit consideration of duration, but rather by using the relative positions and presumed temporal sequence of fossil-bearing strata, and the biological affinities among the fossil remains those strata contain. Despite refinements in absolute dating, and the introduction of molecular methods, the fossil- and event-based relative timescale has remained the backbone of the paleontological conception of time.

Time in paleontology presents a menu of philosophical issues: the phenomenology of time, the nature of relative and absolute time, and the epistemic issues arising out of attempts to reconstruct the past from its present-day traces. We will begin by contemplating the phenomenology of time in the work of the practicing paleontologist. From research in the psychology of time perception, we know that our sense of time is event-based (Resnick et al. 2012). Thus our general sense of time arises out of the first person perspective we have on our lived experience, augmented by reflection on the events of history. Yet these senses of time are inadequate to paleontology. Consider that the oldest known fossils on Earth, of anaerobic bacteria, were found in Australian rocks dated at 3.4 billion years before present (Wacey et al. 2011). The timespan over which the events studied by paleontologists has occurred is so vast as to merit its own designation, Deep Time (McPhee 1981; see also Ricqlès's Chap. 11 in this volume). Deep Time dwarfs the timescales humans are cognitively equipped to deal with. It is immeasurably long compared to the entire span of human history. So the question is whether anyone can truly grasp Deep Time. Efforts to do so nearly always make use of appropriately scaled spatial metaphors. It is a standard part of the undergraduate training of geologists to produce some spatially extended model (e.g., a ruler, a roll of toilet paper, a piano keyboard) such that the vastness of Deep Time can be compared to the mere hair's breadth (or less) of human history that blemishes the end of it. Spatializing time serves two roles. Existentially, the exercise puts us in our place: as a species, we are a relatively recent addition to Earth's biosphere. Cognitively and heuristically, space often stands in for time itself in the scientific work of the paleontologist. In diagrams, time is represented by an axis. In the field, time is represented by the succession of strata.

If you think about it for a moment, there doesn't seem to be anything special about the vastness of time. Space is vast too. Just like Deep Time, vast spatial distances used in astronomy such as the parsec (30,856,774,879,000,000 m) are also difficult to conceive of without some sort of heuristic. Strangely enough, the most intuitively graspable spatial measure used to render astronomical distances tractable—the light year—does so through an implicit time-for-space substitution. While most of us may not be able to fathom such an immense distance, we certainly have a better handle on just how long (or short) a year is. Space and time are both vast. We tame them by scaling them down to human dimensions.

Yet the spatializing of time is not simply a taming technique. The objects that paleontologists study—fossils—and the matrix in which they are embedded—rocks—are spatially extended records of processes occurring over time (see Chap. 11 in this volume). Because the geologic strata in which fossils are found are formed by a process of sedimentary deposition, younger layers sit atop older layers (Steno's Law [Miall 2004]), and as depositional environments of the past migrated due to shifts in water level and the sinking, rising, and filling of sedimentary basins, there are vertical successions of environments that were once next to one another (Walther's Law [Ebbighausen and Korn 2013]). Of course, certain of these shifts result in old deposits being eroded away too, so the geologic record is also a record of missing time. The paleontologist clambering up and down a rock outcrop to collect fossils or to note their positions is also, in a metaphorical sense, traveling through time. Indeed, the field paleontologist is always oriented to her position in time, and even in cases where rock sequences have been overturned by Earth's tectonic processes, is always keenly aware of whether she is moving backwards or forwards through geologic time. Aware of intervals of time unrepresented by rock, she also learns to "mind the gap" as they say.

Fossils too embody a spatialization of temporally extended biological processes. Consider the fossilized tusks of extinct mammals such as mastodons and gomphotheres (Fox and Fisher 2004). Like the trunks of trees, tusks contain growth rings. The chemical and isotopic signatures of these annual increments allow paleontologists to reconstruct life history, diet, and environment. Snails and corals also grow in such a way that their fossilizable hard parts embody a record of the processes that gave rise to them (Gould 1970). As we shall see, these fossils can also be used to test dates determined via radioactive decay.

Thus, the basic parameters of the human condition, of being a being whose kind recently arrived on the scene and is passing through time, combined with the nature of the objects of paleontological study, make space-for-time substitution a fundamental part of the science of paleontology. In what follows I will discuss: (1) the relative time-scale from the perspective of cognitive science; (2) the relationship between molecular and fossil-based estimates of the timing of evolutionary events; (3) the problem of epistemic access to past events and processes; and (4) Lyell's and Darwin's book metaphor for the fossil record.

10.2 Fossils, Time, and the Human Mind

A complete history of the geologic time scale is beyond my scope here, but no one would doubt that the geologic time scale has its present form for historical reasons (see Tassy's Chap. 12 in this volume). William Smith (1799) worked out the stratigraphy of England and Wales based on the lithologies of the rock layers as well as their contained fossils. Smith recognized that strata could be ordered according to the fossils they contained—the Principle of Faunal Succession (Bjornerud 2005). He also discovered that in cases where the strata were jumbled or missing, fossils were the key to correlating far-flung local sequences into one overall geologic column (Winchester 2002). Georges Cuvier and Alexandre Brongniart (1808) established the stratigraphic sequence within the Paris Basin based on similar principles (see Chap. 11 in this volume).

Over time, stratigraphers worked out the succession of strata locally in various places around the world, and out of the ongoing attempt to correlate local stratigraphy globally came the stratigraphic record (see Chap. 12 in this volume). Paleontology has always played an important role in temporally correlating one local section with the next, as fossil species are identified that are reliable and geographically widespread markers of particular intervals in earth history (these are known as “index fossils”). That the geologic time scale is a product of history is also clear from the current and former names of units of rock and their corresponding units of geologic time. “Cretaceous” evokes the chalk deposits of the Paris Basin studied by Cuvier and Brongniart. The last stage in the Cretaceous—the Maastrichtian—is named for the stratotype found at Maastricht, Netherlands. The Devonian recalls Devonshire, England, where it was first defined (Sedgwick and Murchison 1839; Rudwick 1985). Until they were recently respectively renamed Paleogene and Neogene, the Tertiary and Quaternary periods were vestiges of John Phillips's (1829) periodization: Primary-Secondary-Tertiary-Quaternary. It is important to note that these historical boundaries coincide with evolutionary and stratigraphic discontinuities in the fossil and rock records, often but not always global in scale, and not always strictly contemporaneous at all locations (Newell 1959, 1966). In other words, the geologic time scale reflects not only the human history of its construction, but also the events comprising the history of the earth (e.g., flooding of continents) and of life (e.g., mass extinctions). The relative time scale is fundamentally historical in these two very different senses.

The question is why, given improvements in radiometric dating and better ability to correlate stratigraphic sequences worldwide using additional geophysical data such as the record of reversals in the earth's magnetic polarity recorded in rocks (magnetostratigraphy), geologists and paleontologists continue to cling to the geologic time scale. Why not move to a system of absolute dating of rocks, fossils, and events? In other words, instead of speaking of a Devonian trilobite, why not give the most accurate and precise estimate possible—in years—of when the particular specimen (or species) lived? Can't paleontology just “go digital”? The answer is yes, it can. But even so, there are still good reasons for retaining the

relative time scale. Despite all of the history of science frozen into the names of the subdivisions of the geologic time scale, it would be a mistake to explain the persistence of the geologic timescale as nothing more than a curiosity. Not only does its structure reflect major events in the physical and biological history of the earth, but its staying power can also be explained by cognitive science, as we shall see in the next section.

10.3 Paleontology and Geologic Time: The Relative Time Scale

Humans are cognitively ill-adapted to reasoning about interval scales of linear time, but are well adapted to thinking of time in terms of events (Resnick et al. 2012). Thus it should not be surprising that even though multiple means of absolute dating of rocks and fossils have been developed, paleontologists and geologists still use the traditional, event-based relative time scale (i.e., the geologic time scale) to situate events in time, and are constantly trying to date event boundaries more precisely. In other words, the relative time scale tends to remain more or less fixed. The dates of the defining event boundaries—evolutionary events such as mass extinctions of cosmopolitan fauna or geological events such as episodes of widespread continental flooding—are what get revised. For example, the Cretaceous-Paleogene (K-Pg) boundary is defined in the marine stratigraphic record by the extinction of certain species of foraminifera (a microscopic shelled marine organism), and debates over how to define the boundary stem from the problem that foraminifera are not found in deposits from terrestrial environments, in which case it may be argued that instantaneous traces of some alternative event—such as shifts in ocean geochemistry or magnetic shifts recorded in rocks—be used to define the boundary (see Keller 2011 for discussion). Yet what one does not see are proposals to do away with the event-based timescale. Rather, absolute dates are used to improve the temporal resolution of what is essentially an event-based timescale.

Two features of the geologic time scale are worth pointing out in this regard: it is hierarchical and it is event-based. Eons are composed of eras, which are composed of periods, which are composed of epochs, which are composed of stages, and the boundaries between these subdivisions are defined by events in earth history. Its event-based and hierarchical nature helps explain why the geologic time scale is still in use: it happens to correspond to one cognitive model—the category adjustment model—of how humans retrieve information when making temporal estimates. In the category adjustment model, a person retrieves information at the level in a hierarchy of categories that seems appropriate given the question, in light of the location of adjacent boundaries (Huttenlocher et al. 1988). They then use whatever metrical information they may have to adjust their estimate relative to a known category boundary (Resnick et al. 2012). For example, confronted with the task of estimating when in geologic time a given event occurred, a person will default to

the nearest boundary of which they are aware, which may be at a coarser scale in the hierarchy than required by the question (Resnick et al. 2012). If we restrict our expectations about “grasping the magnitude of geologic time” to being able to provide accurate estimates of the dates of events in the fossil or rock records, then clearly the hierarchical system of the geologic time scale fits well with a tried and true cognitive strategy. In working with students to give them a better sense of events in the history of life, pedagogical research has shown that students give more accurate estimates when they are taught to treat intervals of geologic time as event-bounded categories within which can be embedded finer scale event-bounded categories (Resnick et al. 2012). If the cognitive scientists are right, then the geologic time scale is too useful to be discarded anytime soon. Nonetheless, for purposes of calculating rates of evolution and extinction, and for testing models of extinction dynamics, it is crucial to be able to improve the accuracy and precision of the ages of fossils, measured in years (Erwin, 2014). In fact, efforts in this regard which involve collaboration among geochemists, geophysicists, geochronometers, paleontologists, and molecular biologists (such as the EARTHTIME project: earth-time.org) provide a fine example of scientific unification that does not involve reducing one scientific theory to another, but rather is a form of methodological integration (Grantham 2004).

10.4 Absolute Time

10.4.1 Radiometric Dating

Radiometric dating relies on the decay of radioactive isotopes whose decay constants (which can be expressed in terms of the time it takes for fifty percent of the parent isotope to decay—the isotope’s “half-life”) are known. It works by identifying systems, assumed to be chemically closed, containing radioactive isotopes and their daughter products, and uses the ratio between parent isotope and daughter isotope to compute the age of the dated object (often mineral grains from volcanic ash beds). Carbon-14, Uranium-Thorium, and Potassium-Argon are well known systems, whose mode of occurrence and half-lives suit them each to different age ranges. Dates established using radiometric methods are where the buck stops in dating events in the geologic past. But are such dates accurate? How do we know? Is it possible that physical changes in the universe cause alterations in the rate of radioactive decay, possibly by altering the underlying physical constants? Is it possible that changes in pressure and temperature, such as occur in many geological processes, alter decay constants? Other than uniformitarian assumptions about the rate of radioactive decay for various isotopes, how do we know that radioactive decay constants are, well, constant?

10.4.2 Using Fossils to Check Radiometric Dates

Ever since John Wells (1963) first proposed the idea, it has been an ongoing project to develop methods that take advantage of growth features of fossils to check the accuracy of radiometric dates. This is because many marine invertebrates incorporate annual, seasonal, (possibly) monthly, and daily growth bands into their skeletons/shells. This holds out the prospect that information about astronomical time can be extracted from fossils. Working with exquisitely preserved corals found in rocks of Devonian age, Wells (1963) proposed a way to use daily growth bands as a (cheaper) alternative to radiometric dating for estimating the absolute age of fossils and the strata in which they are found. Wells started from the assumption that tidal forces continually put the brakes on the earth's rotation around its polar axis, lengthening the day, and thus that the further into the past we look, the faster the rotation of the earth and the greater the number of days in a year. Astrogeophysicists can calculate the rate at which tidal friction has slowed down the earth's axial rotation (about 2 s per 100,000 years), which yields the number of days per year at points in the past. Wells counted the number of daily growth bands for particularly well-preserved coral specimens from the Middle Devonian of New York and Ontario. The counts ranged between 385 and 410—not very precise—but consistent with a year of 393–396 days as calculated based on astrogeophysical models of changes in the rate of the earth's rotation. This provided proof in principle that growth bands on fossils could be used to date the strata in which the fossils were found, though subsequent research has emphasized some of the practical and methodological difficulties of using fossils in this way, including the possibility that the growth bands may not be a strictly daily phenomenon (Scrutton and Hipkin 1973; Hughes 1985). Nonetheless, paleontologist Stephen Jay Gould (1970, p. 104) considered it a noteworthy reversal of the disciplinary pecking order: “Somehow, the thought that eminent physicists (Runcorn 1966a, b) are studying humble corals does wonders for our self-respect.”

Let us assume that some of the difficulties with inferring astronomical time from fossils can be (biologically and mathematically) worked out (Sisterna and Vucetich 1994). Perhaps the growth lines reflect lunar time, or some other calculable periodicity (Pompea et al. 1979). Earth-moon dynamics from the deep past can be independently estimated from sedimentary successions of thinly bedded sandstones, siltstones and mudstones known as rhythmites, which are thought to record the effects of lunar tide periodicities on sedimentation (Mazumder and Arima 2005). We then would have independent lines of evidence for the age of certain strata and their contained fossils, a consilience of inductions. As Forber and Griffith (2011) have argued—following in the footsteps of William Whewell (1840, 1858)—consilience has evidential import in sciences such as paleontology and geology. The estimation of the age of a stratum from growth bands in one of its contained fossils relies on a set of auxiliary assumptions completely independent from those of radiometric dating (paleoastronomy based on invertebrate fossils and that based on rhythmites do share at least one auxiliary assumption in common—

the constancy of Newton's universal gravitational constant G [Mazumder and Arima 2005]). In any case, to the extent that past dates can be estimated by multiple means whose auxiliary assumptions differ substantially, or are entirely independent, the resulting date garners the evidential support conferred by a consilience of inductions (Whewell 1858; Wimsatt 1981; Forber and Griffith 2011).

10.5 Rocks and Clocks: Competition vs. Consilience

Consilience is inferentially powerful, and at times sets hearts aflutter, but independently derived quantities do not always agree (not necessarily a bad thing, so long as we can localize the error, as Wimsatt [1981] has pointed out in his discussions of failures of robustness). This has been the case in the so-called “rocks and clocks” debate over the timing of evolutionary events. Paleontology and molecular biology can both be used to estimate the timing of branching events in the history of life (see Huneman's Chap. 14 in this volume). Paleontology relies on hypothesized evolutionary relationships and dates of fossil occurrences (based ultimately on radiometric dates of the geologic record), whereas molecular biology uses molecular sequence data and estimated substitution rates (calculated using dated fossil occurrences!). The resulting estimates sometimes differ greatly. For example, one molecular clock estimate (Kumar and Hedges 1998) placed the divergence time of modern eutherian mammals 64 million years earlier than the corresponding estimate based on the fossil record (Foote, et al. 1999).

The choice of the term “molecular clock,” as proposed by Emile Zuckerkandl and Linus Pauling (1965), was meant to suggest that amino acid substitutions in lineages of molecules (whether DNA or other informational molecules, or “semantides” as they put it) accumulate at a steady but stochastic rate, ticking not like a metronome, but rather more like a Geiger counter (Wilson et al. 1987). Zuckerkandl and Pauling (1962, 1965) observed that the degree of difference between hemoglobin molecules in different species (several species of primates, horse, sheep, pig, shark, bony fish, lungfish, worm) is proportional to their evolutionary distance. On that basis they proposed the molecular evolutionary clock, partly to be provocative, but primarily because they saw the possibility of bringing together molecular biology, morphological evolution, and paleontology (Morgan 1998). As a rough calibration they placed the most recent common ancestor of human and horse somewhere in the “Cretaceous or possibly Jurassic period,” resulting in a substitution rate of one amino acid substitution per 11–18 million years (Zuckerkandl and Pauling 1962, p. 201). They knew that there would be differences in the rate of evolution for different molecules, and suspected that the rate of molecular evolution might vary, but Pauling was fond of presenting the clock concept using simple equations that assumed greater constancy than they knew to be the case empirically (Morgan 1998). In addition, they recognized that “back substitutions”—a mutation back to the original base or nucleotide—would go undetected, leading to an underestimate of the number of substitution events that

had taken place (Morgan 1998). While it has been said that Zuckerkandl and Pauling did not discover the molecular clock but rather assumed it (Easteal et al. 1995), Pauling pointed out some years later that they did check their molecularly derived age estimates against those that paleontologists had derived based on macroscopic characters (Morgan 1998). At the level of imprecision at which they were working (recall how rough was their placement of the horse-human ancestor), the overall agreement between paleontological dates and molecular dates constituted consilience. Moreover, at a time when the reconstruction of the timing of evolutionary events relied exclusively upon a fragmentary fossil record, Zuckerkandl and Pauling's research constituted proof in principle that molecules could serve a useful role in reconstructing the evolutionary past.¹

As molecular biology, phylogenetics, and paleontology have become increasingly concerned with quantitative rigor in the handling of data to place evolutionary events in absolute time, there has often been conflict between temporal estimates based on the fossil record, and those derived from molecular data. Part of the problem may be sociological, as paleontologists and molecular biologists (respectively) may be more familiar with the strengths of their own data and methods, and hence distrustful or even dismissive of those of their disciplinary counterpart, but there is also the general problem—the weight of evidence problem—of determining how to assign an appropriate weight to fundamentally different kinds of (sometimes conflicting) evidence (Douglas 2010). The problem becomes especially acute when data fail to agree. If consistency is sought, it is always possible to assign outliers a weight of zero, and to work only with those data that are consistent, but this is epistemically suspect, as consistency is no guarantee of accuracy, and throwing away some data merely because they disagree with other data smacks of *ad hoc*ery (Brochu et al. 2004). Moreover, if we accept or reject a given molecular date solely because it agrees or fails to agree with current estimates, this is tantamount to saying we gain no new information from molecular data, a dubious proposition (Bromham 2006).

There are many reasons why molecular dates and those based on the fossil record may fail to agree. One reason is that the molecular clock does not generally tick at a constant rate, or even at a stochastically constant rate (i.e., it is not analogous to radioactive decay after all [Wilson et al. 1987]). There is variation in substitution rate within and between branches, with some genetic sequences more functionally constrained than others. Another reason is that, to the extent that the molecular clock can be calibrated (i.e., that the average substitution rate can be estimated) based on dates established independently (geologically or paleontologically), there has been disagreement both about which dates to use and how to use them (Conroy

¹Perhaps surprisingly, although the molecular clock is often associated with the neutral theory of molecular evolution (Dietrich and Skipper 2007), Zuckerkandl and Pauling were operating within a selectionist framework (Morgan 1998). Essentially, they assumed that functional constraints on substitution were sufficiently steady as to allow substitution rate to serve as an approximate indicator of elapsed time (Morgan 1998). For more on the neutral model and its influence on paleontology, see Huss (2009).

and van Tuinen 2003). Due to the incompleteness of the fossil record, the evolutionary origin of a taxonomic group (clade) nearly always pre-dates its first fossil. A given group is at least as old as its oldest correctly identified and assigned fossil, which represents a hard minimum constraint on the age of the group (Benton et al. 2009). That said, establishing a reliable and accurate phylogeny from fossil evidence can be quite difficult. This is both because diagnostic traits are often absent in the earliest representatives of an evolutionary group (because they were not yet evolved), and due to the fragmentary nature of fossil remains more generally. Furthermore there is disagreement over which methods of phylogenetic reconstruction are most likely to yield the “true” evolutionary tree.

Just as students of molecular evolution have attempted either to develop more realistic models of evolution or to use methods that are as insensitive as possible to deviations from the model of evolution assumed (Huelsenbeck 1995), paleontology has increasingly emphasized mathematically modeling the incompleteness of the fossil record, such that first and last fossil appearances are interpreted appropriately in estimating the timing of evolutionary events (Benton et al. 2009). When there is a large temporal gap between a molecular estimate of the age of a group and its oldest known fossil, assumptions about preservation, diversification, and extinction rates can be used to calculate how plausible it is to have such a large temporal gap in preservation (Foote et al. 1999). Ultimately, however, there has been, until recently, too strong a partisan attempt on the part of (some) proponents of molecular clock methods to downplay the usefulness of fossil evidence, and for (some) paleontologists to point to problems with molecular clock methods and the model of evolution on which they are based (Donoghue and Benton 2007). The situation finds a parallel in debates over using stratigraphic information to constrain the topology of phylogenies based on character state data (Smith 2000; Fisher et al. 2002; Grantham 2004; see also Tassy’s Chap. 12 in this volume). In the end, paleontology should have taught us a valuable lesson by now: evidence is too hard to come by to be thrown away. If our methods are causing us to discard evidence, then new methods should be devised that adhere to the Principle (trumpeted by Bayesians but more broadly applicable) of Total Evidence (Sober 2009). Increasingly this is occurring, as new methods play to the strengths of morphological, stratigraphic, and molecular data in dating the time-tree of life. The fossil record provides hard minimum and soft maximum ages to help calibrate molecular phylogenies (Benton et al. 2009). In turn, relaxed clock methods, which do not assume constancy of the molecular clock, instead allow multiple models with freely “evolving” substitution rates to compete, with the nodes of the branching tree constrained by dates from the fossil record (Hasegawa et al. 2003). Since it is automatically the case that allowing substitution rates to vary will improve goodness of fit to the data, model selection criteria, such as the Akaike Information Criterion (Forster and Sober 1994), must be used to find those models that will achieve the best fit with the fewest independently estimated substitution rates (Hasegawa et al. 2003). Overfitting the model to the data is to be avoided (Forster and Sober 1994). In addition, assessment of molecular dates should not be viewed purely as model selection in the statistical sense. The assumptions of

models of molecular evolution also need to be empirically tested and consistent with what is known of the underlying biology, and sensitivity of date estimates to violations of those assumptions needs to be assessed (Bromham 2006).

It is worth pausing for a moment to consider the significance of the apparent failure of consilience between fossil and molecular dates. In this case the problem appears to have been a failure to recognize that fossil data and molecular data should not be expected to estimate the same quantity. While it had long been assumed that the failure of concordance of fossil-dated calibration points and branching points determined using molecular methods was undesirable, the standard for consistency in these cases is simply that the fossil date be younger than the molecular estimate, as fossils do not record branching points (Benton et al. 2009). Instead the fossil calibration point corresponds to the oldest phylogenetically secure member of the clade, dated with a precision corresponding to that of the fossil-bearing stratum in which it is found. Even so, there are conditions under which well-dated fossil remains could provide evidence against a molecular date. Suppose we have a taxon that is readily fossilized, and that the rocks of the appropriate age and environment exist to be studied, but there are no fossils found in the gap between the molecular estimate of the date of origin of the clade and the oldest known fossil from that clade. We would have good reason to suspect that the molecular estimate is overestimating the age of the clade (Foote et al. 1999; Grantham 2004).

10.6 The Importance of Absolute Dating for Paleontology

As mentioned previously, paleontology has reconstructed much of the history of life on earth using only a relative time scale. As long as the aim was to record the succession of taxa (extinctions and originations); to correlate their occurrences from place to place; to study their morphology, characteristic modes of preservation, ecological or sedimentological context, and taxonomic affinities; an absolute time scale was not needed (nor was one available). With the advent of the diverse, independent to quasi-independent methods of absolute dating mentioned earlier, and international collaboration among geochronometers, such as the EARTHTIME initiative, it is now becoming possible to begin to study events in the history of life with much greater precision and accuracy, and with this ability comes a novel capacity to ask different sorts of questions (Erwin 2014).

Paleontologist Douglas Erwin has said that to really understand the dynamics of evolution, diversification, and extinction, it is necessary to “go digital” and start using absolute dates to compute rates of evolution and extinction and to “deconvolve” specific events around mass extinction boundaries (Erwin 2014). Part of the problem comes from binning the fossil record into the units of the geologic time scale, making time resolution coarser than it needs to be. The limitation of saying a fossil is of, say, late Permian age is that all fossils from the late Permian will be binned together as if they co-existed when in fact the late

Permian spanned millions of years and any given late Permian fossil ultimately was collected from a (perhaps) more precisely datable stratigraphic bed within it. Take mass extinctions for example. Ever since Raup and Sepkoski (1984) found a 26.2 million year periodicity in the fossil record of mass extinctions, an extraterrestrial origin (such as an asteroid impact) for mass extinctions has been hypothesized. Let us set aside the possibility that the durations of stratigraphic stages themselves contain a periodic signal and thus that the periodicity of mass extinctions may be an artifact of placing last appearances in temporal bins (Stigler and Wagner 1987, 1988). Pursuit of the mass extinction periodicity hypothesis has led paleontologists to assume a uniform, recurring cause for mass extinctions, and therefore to direct their efforts toward finding certain kinds of selectivity in the traits and habitats of organisms that survived and perished at mass extinction boundaries. But increasingly it looks as if not all of the mass extinctions had extraterrestrial causes, and attention has turned to each individual extinction and its ecological dynamics (Shen et al. 2011). Here is where improved dating of extinction events becomes critically important. In order to test hypotheses that, for example, there was a disruption in the web of ecological relations, it is absolutely vital to know in some detail which taxonomic groups became extinct in what order around event boundaries (Erwin 2014). Only in this way does it become possible to differentiate between common cause explanations in which all taxa are driven directly extinct and ecological cascades in which directly caused species deletions percolate through trophic webs to cause further extinctions (Erwin 2014). Improved resolution of the timing of evolutionary events can help distinguish between rival hypotheses and improve our epistemic access to the past. The problem of epistemic access to the past is an important challenge for paleontology, one that goes beyond ascertaining the dates of evolutionary events, but to the extraction of biological features from fossil remains, a topic to which I now turn.

10.7 Epistemic Access to the Past: The Colors of the Dinosaurs

Does the study of the past, especially the deep past, present epistemic problems distinct from those of sciences that study the here and now? If so, what is the best way to characterize them? For philosophers who have turned their gaze to paleontology, a consensus has emerged that the key to understanding its epistemic challenges involves grappling with the underdetermination of theory by data (Cleland 2002; Turner 2007; Jeffares 2010; Forber 2009; Forber and Griffith 2011). Carol Cleland (2002) set the terms of the debate by invoking a pervasive time asymmetry of nature that historical sciences such as paleontology exploit in drawing causal inferences. Cleland points out that in causal sequences in general, earlier states of affairs are vastly overdetermined by later states of affairs. For example, to infer that a baseball had shattered a window would require not every

single trace of the event but perhaps a few appropriately positioned shards of glass and some fibers of horsehide (Cleland 2002). Cleland's view supports epistemic optimism about inference in historical science. With a superabundance of physical traces of effects, all that need be done is to find a disjoint combination of just the right ones—a “smoking gun”—to eliminate all but one of a number of competing causal hypotheses (Cleland 2002).

Derek Turner (2007) takes issue with Cleland's analysis of historical science. He points out that many questions raised by historical science will forever go unanswered. This is because the passing of time generally is accompanied by the degradation of evidence, and at times by the total loss of evidence (but see Sober and Steel 2014). Information-destroying processes abound (Turner 2007). Thus, even if Cleland is right about the time asymmetry of causation, for many events in the past there will remain rival hypotheses underdetermined by the available evidence. Moreover, in many cases background theories give us good reason to believe that new evidence is not forthcoming—and will not be forthcoming—to break ties between rival hypotheses (Turner 2007). Turner labels this the problem of *local underdetermination*, and it underwrites a certain epistemic pessimism about inference in historical science.

As an example of a question that will forever remain unanswered, Turner chose, “What color were the dinosaurs?” As is now widely known, paleontologists have begun to reconstruct the colors of certain dinosaurs on the basis of fossilized melanosomes, microscopic pigmentation bodies previously interpreted as traces of bacteria. In choosing his example, Turner clearly did not expect this development, and, if we are to be honest, probably most paleontologists did not expect it either. We could easily chalk this up to an ill-chosen example of an unanswerable question—if he could replay life's tape I am sure Turner would have chosen a different example—but I think we should focus our attention on what we can learn from it.

Jeffares (2010) has suggested that the mistake lies in focusing exclusively on the fact that, as time passes, evidence of the past is constantly being lost to information-destroying processes. Rather, he argues, we should think of the possibilities of reconstructing the past against the background of the totality of scientific knowledge. New theoretical, empirical, and technological developments in far-flung fields can lead us to recognize that traces exist where we may not have thought to look. In fact, it is noteworthy that in his analysis, Turner, looking to taphonomy as one of the main background theories in paleontology, characterizes it as the study of processes that distort or degrade the fossil record, whereas it could just as easily be thought of as the science of unlocking information in the fossil record (Bergmann et al. 2012). This holistic position of Jeffares gives support to epistemic optimism about our ability to know the past.

Forber (2009) has analyzed local underdetermination (*sensu* Turner 2005a) and concludes that in a large class of cases, including many in paleontology, our epistemic situation is often such that we are confronted by a set of rival hypotheses, each of which is empirically equivalent relative to a specific data set (i.e., not to all presently available data, nor to all potential data) but that we are in no position at

the present time to say what our epistemic state will be in the future. We simply do not know what evidence may or may not be forthcoming. Forber calls this type of underdetermination *contrast failure* and leaves open the possibility that the condition may be temporary.

Forber and Griffith (2011) focus on the issue of epistemic access to the past. They emphasize that one way to gain epistemic access to the past is to, in a sense, triangulate on an estimate of a particular quantity by utilizing evidentiary inferences starting from independent bodies of evidence and sharing as few auxiliary assumptions as possible. As an example, they cite various estimates of the size of the Chicxulub impact crater (caused by the asteroid thought to have been responsible for the end-Cretaceous mass extinction) that agree with each other despite utilizing distinct data sets and relying on independent auxiliary assumptions: a *consilience of inductions* (Whewell 1840, 1858).

Wimsatt (1981, 1987) has not written extensively on epistemic issues in the historical sciences, but his remarks on robustness analysis are a helpful adjunct to discussions of consilience. Forber and Griffith (2011) emphasize the epistemic support garnered by independent lines of evidence that do not overlap (or overlap as little as possible) in their auxiliary assumptions, as each line of evidence will be somewhat immune to criticisms leveled against another line. Yet Wimsatt's concept of robustness emphasizes the extent to which certain derivations will be relatively insensitive to the particular auxiliary assumptions made. Thus, multiple lines of evidence may converge because their different independent auxiliary assumptions do not bias the derivation so greatly that convergence fails to occur. In fact, if each independent line of evidence made some different, faulty assumption, and yet they somehow happened to converge on the same (accurate) estimate of some quantity, this would seem to require an enormously improbable coincidence. This lends plausibility to Wimsatt's view. Also, the existence of multiple independent lines of evidence increases the probability that at least one of the evidentiary lines does not contain a fatal error. A thorough discussion of robustness and its relationship to consilience is beyond the scope of this chapter. Wimsatt (1981) hints at it.

In light of these different characterizations, let us return to the colors of the dinosaurs. How should we understand this example? As a point of clarification it should be noted that birds are part of the dinosaur clade. In common parlance, when most of us speak of dinosaurs, we are excluding birds, instead referring to non-avian dinosaurs, and these are the kinds of organisms whose reconstructed feather colors are under discussion. The primary evidentiary basis for the inferred colors of dinosaur feathers consists of the microscopic fossilized bodies that have been interpreted as melanosomes. Their size, shape, orientation, and distribution in well-preserved fossils is similar to that of feathers in living birds (Fig. 10.1), and to those of fossilized bird feathers as well.

I will focus here on the reconstruction of plumage color patterns in *Anchiornis huxleyi* (Li et al. 2010; Fig. 10.2). A specimen of this non-avian dinosaur was collected from Late Jurassic strata in the Liaoning Province of China. Its feathers

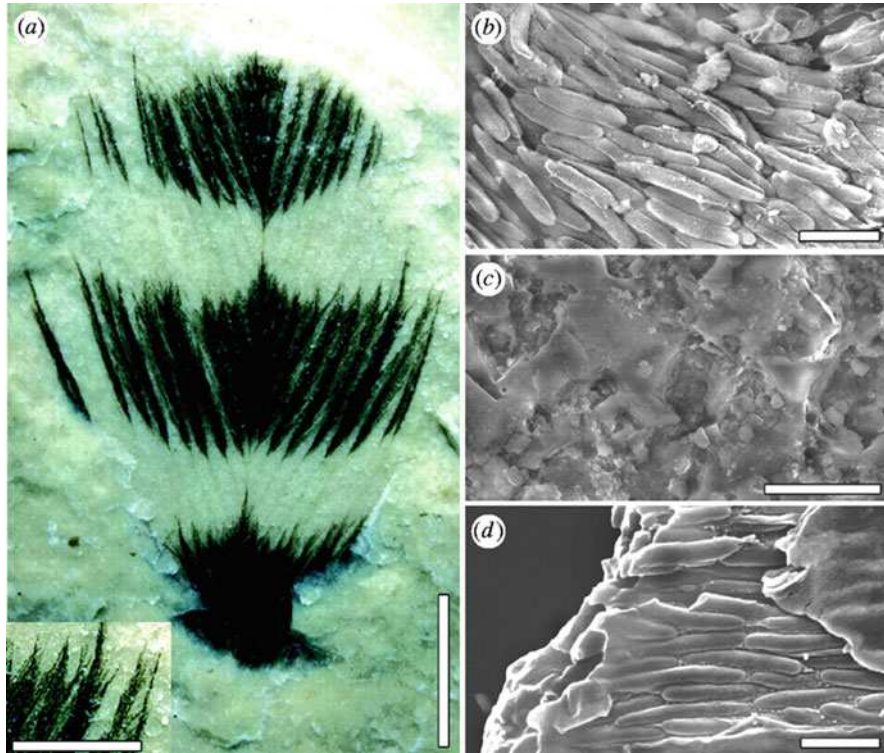


Fig. 10.1 (a) Fossil bird feather from Crato Formation, early Cretaceous, Brazil; (b) photomicrograph of dark band from fossil, showing what are interpreted as eumelanosomes; (c) photomicrograph of rock matrix; (d) photomicrograph of dark band of redwing Blackbird feather, showing eumelanosomes (Vinther et al. 2008; used with permission)

were sampled using a scalpel, mounted, and viewed using a scanning electron microscope (Li et al. 2010; Fig. 10.2).

Measurements were taken of the size, shape, and orientation of melanosomes. In order to reconstruct plumage color patterns, the researchers compiled a training set of melanosome measurements from a phylogenetically diverse set of living birds. These measurements were analyzed using discriminant analysis. Discriminant analysis works by identifying the set of measured variables that best separates samples into two or more classes, in this case, color classes: rufous (red), black and gray. The discriminant analysis was “trained” to correctly assign the feathers of extant birds to color classes with an accuracy of 88.9–91.7% (Li et al. 2010). Once the discriminant analysis had been trained on the feathers of extant birds (whose human-perceived color categories are known), the same analysis was used on the measurements of sampled feathers from *Anchiornis huxleyi* (Li et al. 2010). The discriminant analysis “predicted” (in the words of the authors) a combination of red, gray, and black feathers.

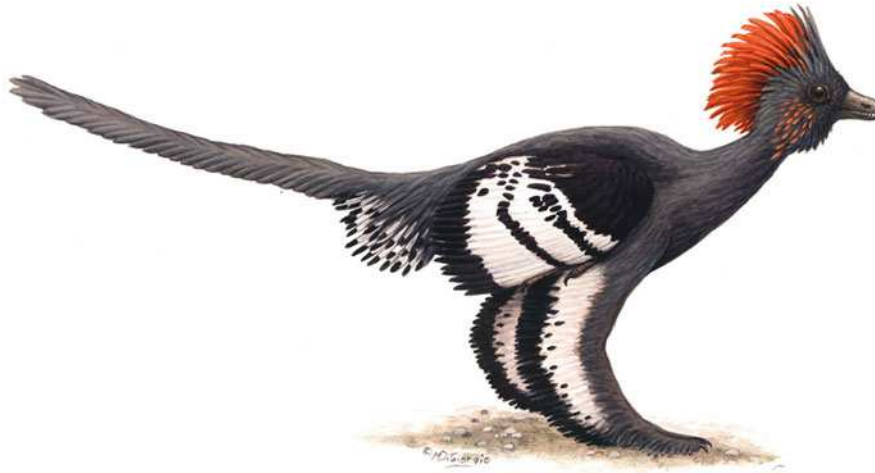


Fig. 10.2 An artist's reconstruction of *Anchiornis huxleyi* (from Li et al. 2010). Used with permission

Several sorts of auxiliary hypotheses, of varying degrees of reliability, play a role here, but immediately we can see the problem that Forber (2009) and Jeffares (2010) pointed out with local underdetermination, as articulated by Turner (2007): we were not in a position in 2007 to say what our epistemic state would be in 2010. Perhaps the most important auxiliary assumption here is that the microscopic bodies are melanosomes. Could they be fossilized bacteria instead? It has recently been argued that this possibility cannot be ruled out. In experiments, bacteria on decaying feathers align in a fashion indistinguishable from that of melanosomes (Moyer et al. 2014), but Vinther (2015, p. 6) has argued that these bacteria are far larger than any known melanosome, fossil or living. Ultimately the argument that the bodies are melanosomes rests on their distinctive morphology, their occurrence in feathers, and their similarity in size and shape to those found in feathers in extant birds. Further research on the fossilization of melanin and melanosomes will certainly shed further light on these issues (Colleary et al. 2015).

A second auxiliary assumption that stands out is the assumption that the size, shape, and distribution of melanosomes found in the fossil feathers reflects what these attributes were in life. In fact, the fossilized melanosomes are typically smaller than their corresponding voids (Vinther 2015, p. 5). Since melanosome shape (aspect ratio) is one of the most important variables in color prediction, isometric shrinkage would not affect the predicted color, but distortion of aspect ratio would (Vinther 2015, p. 5).

A third auxiliary assumption is the uniformitarian assumption that the size, shape, and distribution of melanosomes in *Anchiornis huxleyi* feathers yield the same human-perceived color as those in feathers of extant birds. It is important to differentiate three distinct bases for this assumption: analogy, homology and physiochemistry. Reconstructions based simply on simple analogy have a poor

track record (Turner 2005b). If the scientists were using modern birds simply as an *analog* for *Anchiornis huxleyi*, this would be the weakest form of support for the color inference. A somewhat better supported inference would have a phylogenetic basis (homology). But here we have a combination of support from both phylogeny and physiochemistry. *Anchiornis huxleyi* feathers are homologous with those of modern birds, and the optics, physics, and chemistry that give rise to human-perceived color in modern bird feathers is well-understood. We have very little reason to believe that the interaction of light with melanosomes of a certain shape was any different optically in the Jurassic from what it is today. That said, there are additional factors in feather coloration—structural iridescence and other pigments besides melanins—and at this stage it is unclear whether these can be recovered from fossil feathers (Vinther 2015, p. 8; but see Li et al. 2012 for a discussion of fossil evidence of iridescence). Yet, the visual effects of co-expressed pigments might well be “masked” in parts of feathers where melanin is present due to its strong light absorption (D’Alba et al. 2012).

Forber and Griffith (2011) argued that a consilience of inductions provides epistemic support to the extent that inferences do not share auxiliaries. They point out (p. 3) that apparent disconfirmations of historical hypotheses by extant traces can often be explained away by shifting blame to suspect auxiliaries, thus the benefit of having multiple independent lines of evidence. Yet the flip side of this observation is also important. The epistemic support of a consilience of inductions is not undermined by overlap in their “secure” auxiliaries (e.g., fundamental laws of physics and chemistry). This means that special attention should be paid to identifying suspect auxiliaries and seeking alternate lines of evidence that do not share those assumptions. An example of this in the present context is the use of synchrotron X-ray fluorescence and X-ray absorption techniques, which are capable of detecting trace metals (organically chelated copper) associated with melanins (Bergmann et al. 2012) even after the melanosomes have been lost to dissolution. The problem is that organically chelated copper is not diagnostic of melanin, as other organic compounds common in nature (such as porphyrins which are common in feathers) are also capable of chelating copper (Vinther 2015).

The prospect of reconstructing the colors of dinosaurs has generated a great deal of excitement, but in our present epistemic state, some pieces of the puzzle are more secure than others. Color patterning (patterns in the presence and absence of melanosomes in particular feathers) is a fairly secure inference, the actual colors somewhat less so. Yet it is premature to say that we will never know the color of (certain) dinosaurs. One promising independent line of evidence for fossil feather coloration comes from fossil feathers encased in amber that apparently retain their original color (McKellar et al. 2011).

10.8 The Book Metaphor

Epistemic access to the past is made all the more difficult by the fact that not only do the events of interest lie at a temporal remove, but the geologic and fossil record are incomplete. Thus, it is not surprising that at least since Lyell, the dominant metaphor for the geologic record has been that of a multivolume history book—with much of the story missing. Darwin picks up on this in the *Origin*, where he writes:

For my part, following out Lyell's metaphor, I look at the natural geological record, as a history of the world imperfectly kept, and written in a changing dialect; of this history we possess the last volume alone, relating only to two or three countries. Of this volume, only here and there a short chapter has been preserved; and of each page, only here and there a few lines. Each word of the slowly-changing language, in which the history is supposed to be written, being more or less different in the interrupted succession of chapters, may represent the apparently abruptly changed forms of life, entombed in our consecutive, but widely separated, formations. On this view, the difficulties above discussed are greatly diminished, or even disappear. (Darwin 1859, pp. 310–311)

The difficulties to which Darwin is referring are rehearsed throughout Chapter Nine of the *Origin* from which this passage is drawn: the rarity of transitional forms in the fossil record (a.k.a. missing links), the (apparent) abruptness with which new species appear, and the absence of fossil-bearing strata below what was in Darwin's day called the Silurian, which included what we now call the Cambrian (since then, fossils have been found in older rocks).

The implications of the book metaphor are clear. With incompleteness at every scale—from volume to chapter to page to line to word—it is no wonder that all changes in fossil form appear to be sudden and abrupt. The long chain of intermediate living forms expected under Lyellian gradualism, and presupposed by Darwin, are words in the missing lines, or on the missing pages, or in the missing chapters or volumes, of the history of life. While the large-scale patterns of the fossil record find an explanation in Darwin's evolutionary theory, the absence of the expected fossil intermediates and the abrupt appearance of fossils at the base of the Silurian period find their explanation in Lyell's geologic theory.

The incompleteness of the geologic record and with it, the incompleteness of the fossil record, are familiar commonplaces. Sedimentary rocks come in packages, bounded by diachronous surfaces of time unrepresented by rock. The patchy nature of the fossil record led to the oft-expressed lament that evolution always seemed to be happening “off-stage,” somewhere other than where paleontologists actually find the fossils. In fact Niles Eldredge (1971) and later Stephen Jay Gould, in arriving at their theory of punctuated equilibrium, which borrowed greatly from Ernst Mayr's (1963) theory of allopatric speciation, “read” the fossil record in exactly this way (Eldredge and Gould 1972). According to punctuated equilibrium, the fossil record we see is one primarily of stasis within large populations. Speciation occurs when small subpopulations become reproductively isolated—by becoming geographically isolated according to Mayr—accumulating mutations and then eventually replacing the parent population (for more on punctuated

equilibrium, see Chap. 14 in this volume). On this view, evolution—in the form of speciation—really is happening elsewhere. As the architects of punctuated equilibrium write:

Since speciation occurs rapidly in small populations occupying small areas far from the center of ancestral abundance, we will rarely discover the actual [speciation] event in the fossil record (Eldredge and Gould 1972, p. 96)

Punctuated equilibrium was certainly received as a controversial and revolutionary theory (Sepkoski 2012, p. 137ff), perhaps largely because for a change paleontologists were seen as contributing to evolutionary theory rather than (merely) documenting evolutionary history (Maynard Smith 1984). Yet there has always been some disagreement over just how “radical” and “non-Darwinian” the theory of punctuated equilibria is: can it be easily accommodated within neo-Darwinism (Mayr 1982), or is it proposing some fundamentally different process involving “systemic mutations” and “hopeful monsters” reminiscent of Goldschmidt (1940), as Gould (1977) thought? But if we set evolutionary theory off to the side for a moment, and compare the way Darwin read the fossil record with the way Eldredge and Gould did, we see a stark contrast. Darwin looked at the fossil record and saw broad scale evidence for evolution, but explained away the absence of intermediate life forms by appealing to Lyell’s metaphor of the missing words, lines, pages, chapters, and volumes of history. One of the significant features of Eldredge and Gould’s theory of punctuated equilibria is that they interpret the face value fossil record in evolutionary terms. The absence of intermediates is not something to be explained away by the incompleteness of the fossil record, but rather what we might expect if most of the evolutionary action is taking place among small populations on the periphery of ancestral geographic ranges, and that species are in morphological stasis throughout much of their history. Small populations are statistically less likely to leave fossil remains, and gene flow within large populations keeps them in morphologic stasis. Regardless of whether one regards punctuated equilibrium as a Darwinian or non-Darwinian evolutionary process, Eldredge and Gould’s reading of the fossil record is different from what Darwin’s was.

10.9 Closing the Book Metaphor?

Taking the punctuated equilibrium of Eldredge and Gould (1972) as a point of departure, it may be time to consider the limitations of Darwin’s book metaphor (Peters 2011). David Sepkoski (2012) has rightly shown that the metaphor of reading (and re-reading) the fossil record has had heuristic value both for paleontologists and for those who wish to understand their practice. Indeed, new readings of the fossil record are constantly being proposed, and at the very least require revisions to Darwin’s book metaphor.

It is time to devote serious study to marine sedimentary rocks devoid of fossils as something other than missing pages from the book of life. Shanan Peters (2007) points out, channeling the voice of the late Jack Sepkoski, that for many fossil-free strata, rather than reflecting the loss or dissolution of once present fossils, it may be the case that in some of these units “nobody’s home” (Peters 2007, p. 165). Intensive study of such barren intervals may be helpful in understanding Earth history in terms of systematic interactions among temperature, precipitation, salinity, circulation, sedimentation, and geochemistry in marine environments (Peters 2007). In other words, it may facilitate a systems view of Earth and life history. Moreover, while there is a strong correlation between the amount of sedimentary rock available for sampling and the sampled diversity of fossil taxa over geologic time, it is not necessarily the case that measured fossil diversity is simply a result of how much rock there is to sample (Peters and Heim 2011).

An alternative hypothesis is that perturbations to the earth system are a common cause of both changes in sedimentation and extinction rate (Peters and Heim 2011). Long periods of time unrepresented by rock, such as the Great Unconformity in the Grand Canyon, where the 525 million year old Tapeats Sandstone sits atop the 1740 million year old Vishnu Schist, are not just hiatuses, but rather records of what were surely massive environmental disruptions (Peters and Gaines 2012). Peters and Gaines hypothesize that the massive influx of eroded sediments that bequeathed us the Great Unconformity altered marine geochemistry to the point where biomineralization could occur, leading to a large scale diversification in shelled fossils in the Cambrian period, a phenomenon known as the Cambrian explosion (Peters and Gaines 2012). Intriguingly, molecular dates have placed the origins of marine animals prior to this period (Erwin 2015). If these early animals lacked mineralized skeletons, then the biomineralization hypothesis of Peters and Gaines (2012) may explain why there was a time lag between the early diversification of animals and their appearance in the fossil record.

In summary, barren zones in the fossil record may not be missing pages from the book of life, as Lyell and Darwin envisaged, but may actually be a record of the absence of life from those environments. Likewise, hiatuses in sedimentation, and intervals of time unrepresented by rock, while they may be described as missing pages from the book of life, are also records of perturbations of the earth system that may have played a major causal role in the diversification and extinction of the biota. As Peters (2011, p. 14) put it:

Geologists would do well to move beyond Darwin’s view of the incompleteness of the rock record and remember the words of the French composer Claude Debussy: “music is the space between the notes.”

10.10 Conclusion

Paleontologists, as students of the history of life, grapple both with the immensity of time and with the fact that their objects of study lived in the past, often the quite remote past. One factor in the study of the past is the construction of a framework and methods for dating past events as accurately and precisely as possible (Erwin 2014). Fossils have played, and continue to play, a central role in this, not only by serving as the basis for the relative timescale and allowing global stratigraphic correlation, but also by virtue of being, in a sense, “time recorders,” literally embodying periodic signals (Wells 1963). Studying the history of life on earth requires drawing reliable inferences about past organisms, events, and processes from present day traces, including molecules in living organisms (Easteal et al. 1995). Progress in understanding the past is limited by the available traces, but the history of science has shown that at any given time, it simply cannot be known what traces will prove to be evidentially relevant (Jeffares 2010). Inferences about the past must be indemnified against certain sorts of risks. Auxiliary assumptions involved in drawing inferences are particularly suspect, meaning that multiple inferential chains that do not share suspect auxiliary assumptions improve the reliability of inferences (Forber and Griffith 2011). So do inferential chains whose auxiliary assumptions are relatively secure. While it is true that in a sense, paleontology is outrunning time, trying to develop new and better ways to access evidence about the deep past despite the fact that traces of the past are lost to information-destroying processes (Turner 2007), it should also be borne in mind that those information-destroying processes are part of the dynamic, living earth system, and feed back into history of life itself (Peters and Gaines 2012). The earth may be like a book—an autobiography (Bjernerud 2005)—but if it is, then the missing words, pages, and chapters are a part of the story itself.

Acknowledgments This paper began as a talk given at the University of Bordeaux, and benefited from feedback received from my fellow symposiasts. My thanks go out to Christophe Bouton, Sharon Cebula, Patrick Forber, Todd Grantham, Philippe Huneman, Shanan Peters, David Sepkoski, Matt Shawkey and Joanna Trzeciak for their comments on the manuscript, and to John Alroy and Douglas Erwin for discussion. I also thank Carl Mendelson at Beloit College for introducing me to the idea of Deep Time.

References

- Benton, M.J., P.C.J. Donoghue, and R.J. Asher. 2009. Calibrating and constraining molecular clocks. In *The Timetree of Life*, ed. S.B. Hedges and S. Kumar, 35–86. Oxford: Oxford University Press.
- Bergmann, U., P.L. Manning, and R.A. Wogelius. 2012. Chemical mapping of paleontological and archeological artifacts with synchrotron X-Rays. *Annual Review of Analytical Chemistry* 5: 361–389.

- Björnerud, M. 2005. *Reading the rocks: The autobiography of the earth*. Cambridge, MA: Westview Press.
- Brochu, C.A., C.D. Sumrall, and J.M. Theodor. 2004. When clocks (And Communities) collide: Estimating divergence time from molecules and the fossil record. *Journal of Paleontology* 78: 1–6.
- Bromham, L. (2006). Molecular dates for the cambrian explosion: Is the light at the end of the tunnel an oncoming train? *Palaeontologia Electronica* 9(1.2E). http://palaeo-electronica.org/paleo/toc9_1.htm.
- Cleland, C.E. 2002. Methodological and epistemic differences between historical science and experimental science. *Philosophy of Science* 69: 447–451.
- Colleary, C., et al. 2015. Chemical, experimental, and morphological evidence for diagenetically altered melanin in exceptionally preserved fossils. *Proceedings of the National Academy of Sciences of the United States of America* 112 (41): 12592–12597.
- Conroy, C.J., and M.L. van Tuinen. 2003. Extracting time from phylogenies: Positive interplay between fossil and genetic data. *Journal of Mammalogy* 84: 444–455.
- Cuvier, G., and A. Brongniart. 1808. Essai sur la géographie minéralogique des environs de Paris. *Annales du Musée Histoire Naturelle de Paris* 11: 293–326.
- D’Alba, L., L. Kieffer, and M.D. Shawkey. 2012. Relative contributions of pigments and biophotonic nanostructures to natural color production: A case study in budgerigar (*Melopsittacus undulatus*) feathers. *The Journal of Experimental Biology* 215 (8): 1272–1277.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection*. London: John Murray.
- Dietrich, M.R., and R.A. Skipper Jr. 2007. Manipulating underdetermination in scientific controversy: The case of the molecular clock. *Perspectives on Science* 15: 295–326.
- Donoghue, P.C. and Benton, M.J., 2007. Rocks and clocks: Calibrating the Tree of Life using fossils and molecules. *Trends in Ecology & Evolution* 22(8): 424–431.
- Douglas, H. 2010. Engagement for progress: Applied philosophy of science in context. *Synthese* 177: 317–335.
- Easteal, S., C. Collet, and D. Betty. 1995. *The Mammalian Molecular Clock*. Austin: R. G. Lands.
- Ebbighausen, R., and D. Korn. 2013. Paleontology as a circumstantial evidence lawsuit. *Historical Biology* 25: 283–295.
- Eldredge, N. 1971. The allopatric model and phylogeny in Paleozoic invertebrates. *Evolution* 25: 156–167.
- Eldredge, N., and S.J. Gould. 1972. Punctuated equilibria: An alternative to phyletic gradualism. In *Models in paleobiology*, ed. T.J.M. Schopf, 82–115. San Francisco: Freeman, Cooper.
- Erwin, D.H. 2014. Temporal acuity and the rate and dynamics of mass extinctions. *Proceedings of the National Academy of Sciences* 111 (9): 3203–3204.
- . 2015. Was the Ediacaran–Cambrian radiation a unique evolutionary event? *Paleobiology* 41: 1–15.
- Fisher, D.C., M. Foote, D.L. Fox, and L.R. Leighton. 2002. Stratigraphy in phylogeny reconstruction: Comment on smith. *Journal of Paleontology* 76: 585–586.
- Foote, M., J.P. Hunter, C.M. Janis, and J.J. Sepkoski Jr. 1999. Evolutionary and preservational constraints on origins of biologic groups: Divergence times of eutherian mammals. *Science* 283: 1310–1314.
- Forber, P. 2009. Spandrels and a pervasive problem of evidence. *Biology and Philosophy* 24: 247–266.
- Forber, P., and E. Griffith. 2011. Historical reconstruction: Gaining epistemic access to the deep past. *Philosophy and Theory in Biology* 3: 1–19.
- Forster, M.R., and E. Sober. 1994. How to tell when simpler, more unified, or less Ad Hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science* 45: 1–35.
- Fox, D.L., and D.C. Fisher. 2004. Dietary reconstruction of Miocene Gomphotherium (Mammalia, Proboscidea) from the Great Plains region, USA, based on the carbon isotope composition of tusk and molar enamel. *Palaeogeography, Palaeoclimatology, Palaeoecology* 206: 311–335.
- Goldschmidt, R. 1940. *The Material Basis of Evolution*. New Haven: Yale University Press.

- Gould, S.J. 1970. Evolutionary paleontology and the science of form. *Earth-Science Reviews* 6: 77–119.
- . 1977. The return of hopeful monsters. *Natural History* 86: 22–30.
- Grantham, T. 2004. Conceptualizing the (Dis)unity of science. *Philosophy of Science* 71: 133–155.
- Hasegawa, M., J.L. Thorne, and H. Kishino. 2003. Time scale of eutherian evolution estimated without assuming a constant rate of molecular evolution. *Genes & Genetic Systems* 78: 267–283.
- Huelsenbeck, J.P. 1995. Performance of phylogenetic methods in simulation. *Systematic Biology* 44: 17–48.
- Hughes, W.W. 1985. Planetary rotation and invertebrate skeletal patterns: Prospects for extant taxa. *Geophysical Surveys* 7: 169–183.
- Huss, J.E. 2009. The shape of evolution: The MBL model and clade shape. In *The paleobiological revolution: Essays on the growth of modern paleontology*, ed. D. Sepkoski and M. Ruse, 326–345. Chicago: University of Chicago Press.
- Huttenlocher, J., L. Hedges, and V. Prohaska. 1988. Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review* 95: 471–484.
- Jeffares, B. 2010. Guessing the future of the past. *Biology and Philosophy* 25: 125–142.
- Keller, G. (2011). Defining the Cretaceous-Tertiary boundary: A practical guide and return to first principles. In *The end-Cretaceous mass extinction and the Chicxulub impact in Texas. SEPM Special Publication 100*, eds. Gerta Keller and Thierry Adatte, 23–42.
- Kumar, S., and S.B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature* 392: 917–920.
- Li, Q., K.-Q. Gao, J. Vinther, M.D. Shawkey, J.A. Clarke, L. D’Alba, Q. Meng, D.E.G. Briggs, and R.O. Prum. 2010. Plumage color patterns of an extinct dinosaur. *Science* 327: 1369–1372.
- Li, Q., K.Q. Gao, Q. Meng, J.A. Clarke, M.D. Shawkey, L. D’Alba, R. Pei, M. Ellison, M.A. Norell, and J. Vinther. 2012. Reconstruction of Microraptor and the evolution of iridescent plumage. *Science* 335: 1215–1219.
- Maynard Smith, J. 1984. Paleontology at the high table. *Nature* 309: 401–402.
- Mayr, E. 1963. *Animal species and evolution*. Cambridge, MA: The Belknap Press at Harvard University Press.
- . 1982. Speciation and macroevolution. *Evolution* 36: 1119–1132.
- Mazumder, R., and M. Arima. 2005. Tidal rhythmites and their implications. *Earth-Science Reviews* 69: 79–95.
- McKellar, R.C., B.D.E. Chatterton, A.P. Wolfe, and P.J. Currie. 2011. A diverse assemblage of late Cretaceous dinosaur and bird feathers from Canadian amber. *Science* 333: 1619–1622.
- McPhee, J. 1981. *Basin and Range*. New York: Farrar, Straus, Giroux.
- Miall, A.D. 2004. Empiricism and model building in stratigraphy: The historical roots of present-day practices. *Stratigraphy* 1: 3–25.
- Morgan, G.J. 1998. Emile Zuckerkandl, Linus Pauling, and the Molecular Evolutionary Clock, 1959–1965. *Journal of the History of Biology* 31: 155–178.
- Moyer, A. E., Zheng, W., Johnson, E. A., Lamanna, M. C., Li D., Lacovara, K. J. & Schweitzer, M. H. (2014). Melanosomes or microbes: Testing an alternative hypothesis for the origin of microbodies in fossil feathers. *Science Reports* 4: Article number 4233.
- Newell, N.D. 1959. The nature of the fossil record. In *Commemoration of the centennial of the publication of “The Origin of Species” by Charles Darwin. Proceedings of the American Philosophical Society* 103 (2): 264–285.
- . 1966. Problems of geochronology. *Proceedings of the Academy of Natural Sciences of Philadelphia* 118: 63–89.
- Peters, S.E. 2007. The problem with the Paleozoic. *Paleobiology* 33: 165–181.
- Peters, S. E. 2011. A new view of the sedimentary rock record: Drivers of Earth system evolution are encoded by gaps and rocks. *The Outcrop*, Department of Geoscience, University of Wisconsin-Madison, 12–14.

- Peters, S.E., and R.R. Gaines. 2012. Formation of the ‘Great Unconformity’ as a trigger for the Cambrian explosion. *Nature* 484: 363–366.
- Peters, S.E., and N.A. Heim. 2011. Macrostratigraphy and macroevolution in marine environments: Testing the common-cause hypothesis. In *Comparing the geological and fossil records: Implications for biodiversity studies*, ed. A.J. McGowan and A.B. Smith, 95–104. London: Geological Society, Special Publication 358.
- Phillips, J. 1829. *Illustrations of the Geology of Yorkshire*. York: T. Wilson.
- Pompea, S.M., P.G.K. Kahn, and R.B. Culver. 1979. Paleoastronomy and nautiloid growth: A perspective. *Vistas in Astronomy* 23: 185–205.
- Raup, D.M., and J.J. Sepkoski Jr. 1984. Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Sciences of the United States of America* 81: 801–805.
- Resnick, I., Shipley, T. F., Newcombe, N., Massey, C., Wills, T. (2012). Examining the representation and understanding of large magnitudes using the hierarchical alignment model of analogical reasoning. *CSS 2012 Conference Proceedings*, 917–922.
- Rudwick, M.J.S. 1985. *The Great Devonian Controversy*. Chicago: The University of Chicago Press.
- Runcorn, S.K. 1966a. Corals as paleontological clocks. *Scientific American* 215 (4): 26–33.
- . 1966b. Middle Devonian day and month. *Science* 154: 292.
- Scrutton, C.T., and R.G. Hipkin. 1973. Long-term changes in the rotation rate of the earth. *Earth-Science Reviews* 9: 259–274.
- Sedgwick, A., and R.I. Murchison. 1839. Classification of the older stratified rocks of Devonshire and Cornwall. *Philosophical Magazine and Journal of Science Series 3*. 14: 241–260.
- Sepkoski, D. 2012. *Rereading the fossil record: The growth of paleobiology as an evolutionary discipline*. Chicago: University of Chicago Press.
- Shen, S.-Z., et al. 2011. Calibrating the end-permian mass extinction. *Science* 334: 1367–1372.
- Sisterna, P.D., and H. Vucetich. 1994. Cosmology, oscillating physics, and oscillating biology. *Physical Review Letters* 72: 454–457.
- Smith, W. 1799. *Tabular view of the order of strata in the vicinity of Bath with their respective organic remains*.
- Smith, A.B. 2000. Stratigraphy in phylogeny reconstruction. *Journal of Paleontology* 74: 763–766.
- Sober, E. 2009. Absence of evidence and evidence of absence: Evidential transitivity in connection with fossils, fishing, fine-tuning, and firing squads. *Philosophical Studies* 143: 63–90.
- Sober, E., and M. Steel. 2014. Time and knowability in evolutionary processes. *Philosophy of Science* 81: 537–557.
- Stigler, S.M., and M.J. Wagner. 1987. A substantial bias in nonparametric tests for periodicity in geophysical data. *Science* 238: 940–945.
- . 1988. Testing for periodicity of extinction. *Science* 241: 96–99.
- Turner, D. 2005a. Local underdetermination in historical science. *Philosophy of Science* 72: 209–230.
- . 2005b. Misleading observable analogues in paleontology. *Studies in History and Philosophy of Science* 36: 175–183.
- . 2007. *Making prehistory: Historical science and the scientific realism debate*. Cambridge: Cambridge University Press.
- Vinther, J. 2015. A guide to the field of palaeo colour. *BioEssays* 37(6): 643–656.
- Vinther, J., D.E.G. Briggs, R.O. Prum, and V. Saranathan. 2008. The colour of fossil feathers. *Biology Letters* 4 (5): 522–525.
- Wacey, D., M.R. Kilburn, M. Saunders, J. Cliff, and M.D. Brasier. 2011. Microfossils of sulphur-metabolizing cells in 3.4-billion-year-old rocks of Western Australia. *Nature Geoscience* 4: 698–702.
- Wells, J.W. 1963. Coral growth and geochronometry. *Nature* 197: 948–950.
- Whewell, W. 1840. *Aphorisms Concerning Ideas, Science, and the Language of Science*. London: Harrison and Company.

- . 1858. *Novum Organon Renovatum*. London: John W. Parker.
- Wilson, A.C., H. Ochman, and E.M. Prager. 1987. Molecular time scale for evolution. *Trends in Genetics* 3: 241–247.
- Wimsatt, W.C. 1981. Robustness, reliability, and overdetermination. In *Scientific inquiry in the social sciences* (a festschrift for Donald T. Campbell), ed. M. Brewer & B. Collins, 123–162. San Francisco: Jossey-Bass.
- . 1987. False models as a means to truer theories. In *Neutral models in biology*, ed. M. Nitecki and A. Hoffmann, 23–55. Oxford: Oxford University Press.
- Winchester, S. 2002. *The Map that Changed the World: William Smith and the Birth of Modern Geology*. New York: Harper Collins.
- Zuckerkandl, E., and L. Pauling. 1962. Molecular disease, evolution and genetic heterogeneity. In *Horizons in biochemistry: Albert Szent-Györgyi dedicatory volume*, ed. M. Kasha and B. Pullman, 189–225. New York: Academic.
- . 1965. Evolutionary divergence and convergence in proteins. In *Evolving genes and proteins*, ed. V. Bryson and H. Vogel, 97–166. New York: Academic.