Predictability, an Orrery, and a Speciation Machine: Quest for a Standard Model of Speciation

Marius Roesti,[1,8] Hannes Roesti,[2,8] Ina Satokangas,[3] Janette Boughman,[4] Samridhi Chaturvedi,[5] Jochen B.W. Wolf,[6] and R. Brian Langerhans[7]

[1]Division of Evolutionary Ecology, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland
[2]Department of Biology, Alte Kantonsschule Aarau, 5001 Aarau, Switzerland
[3]Organismal & Evolutionary Biology Research Programme, University of Helsinki, 00014 Helsinki, Finland
[4]Department of Integrative Biology, Ecology, Evolution, and Behavior Program, Michigan State University, Lansing, Michigan 48824, USA
[5]Department of Ecology and Evolutionary Biology, Tulane University, New Orleans, Louisiana 70118, USA
[6]Division of Evolutionary Biology, Faculty of Biology, LMU Munich Biozentrum Martinsried, D-82152 Planegg-Martinsried, Germany
[7]Department of Biological Sciences, North Carolina State University, Raleigh, North Carolina 27695, USA
Correspondence: marius.roesti@unibe.ch

Accurate predictions are commonly taken as a hallmark of strong scientific understanding. Yet, we do not seem capable today of making many accurate predictions about biological speciation. Why? What limits predictability in general, what exactly is the function and value of predictions, and how might we go about predicting new species? Inspired by an orrery used to explain solar eclipses, we address these questions with a thought experiment in which we conceive an evolutionary speciation machine generating new species. This experiment highlights complexity, chance, and speciation pluralism as the three fundamental challenges for predicting speciation. It also illustrates the methodological value of predictions in testing and improving conceptual models. We then outline how we might move from the hypothetical speciation machine to a predictive standard model of speciation. Operationalizing, testing, and refining this model will require a concerted shift to large-scale, integrative, and interdisciplinary efforts across the tree of life. This endeavor, paired with technological advances, may reveal apparently stochastic processes to be deterministic, and promises to expand the breadth and depth of our understanding of speciation and more generally, of evolution.

On 4 September 2100, 12:19 p.m. local time: in Antsirabe, Madagascar, a full solar eclipse will occur. Today's astronomy can make this stunningly precise prediction, and similar predictions in the recent past held true. Arguably, speciation research is currently unable to

---

[8]These authors contributed equally to this work.

Editors: Catherine L. Peichel, Daniel I. Bolnick, Åke Brännström, Ulf Dieckmann, and Rebecca J. Safran
Additional Perspectives on Speciation available at www.cshperspectives.org

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a041456

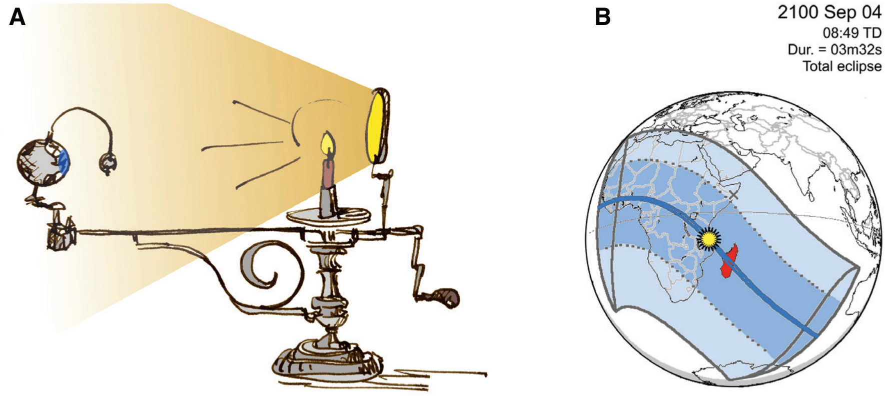


Figure 1. Predictions and models. (*A*) A planetary model of our solar system, called "orrery." In its illustrated version, the candle's light reflected by the mirror represents the Sun. The Earth and Moon follow their orbital revolution around the Sun while the Moon is moving around the Earth. The handle drives the model's mechanism, allowing us to simulate the occurrence of a solar eclipse, as indicated in blue on planet Earth. (*B*) Map for the predicted solar eclipse on 4 September 2100 with Madagascar highlighted in red. (Map edited from its original source: NASA Technical Publication; Espenak and Meeus 2006.)

make similarly accurate predictions about the origin of new species. Should we worry?

PREDICTIONS, PREDICTABILITY, AND SPECIATION

Most generally, a prediction is a statement about what is expected to happen. Its precision reflects the specificity of the predicted event (or trend), that is, the amount of detail on the "what, where, and when" of that event. An accurate prediction entails both high precision and high certainty—it renders an event highly predictable.[9] As opposed to the predictability of an event, when talking about "our predictability of something" we literally mean our ability to make accurate predictions.

Making accurate predictions is of major interest in our daily lives: we want to predict the weather, a natural disaster, a pandemic, the stock market, or the outbreak of war. In contrast to such practical interests, fundamental sciences are less interested in predictions per se, but focus more on understanding how something works. This conceptual understanding is commonly represented by a scientific model and often expressed as a mathematical formula. Yet, predictions also play an important methodological role in fundamental sciences because accurate predictions are generally accepted as a hallmark of strong scientific understanding. For example, our solar system has historically been modeled in a mechanical form called "orrery" (Fig. 1A). In its mathematical expression and fed with sufficient data, this model based on classical physics allows us to both precisely and confidently predict the solar eclipse on 4 September 2100 (Fig. 1B). In return, the model's predictive success, and hence the high predictability of solar eclipses, substantiates the classical understanding of our solar system.

Speciation research is intrinsically interested in understanding the formation of new species in the course of evolution—that is, speciation. Just as we can predict every transitory state of the ever-changing constellation of our solar system occasionally leading to a solar eclipse, the speciation process can in principle

---

[9] When we call a prediction "specific," we refer to the specificity of the event being predicted (i.e., what exactly will happen where and when). While unspecific (broad) predictions are more likely to be true than specific predictions, they are typically of lesser practical and scientific value. We note that the way we use "accuracy" in this essay—that is, reflecting not only certainty but also precision—is a common usage of this term.

also be viewed as a series of transitory states amenable to prediction. Indeed, with a full understanding of speciation, each state of this process should become accurately predictable. Our ability to predict the origin of new species should hence reflect our current understanding of this process. Despite many decades of intense research, why then does our predictability of speciation still seem so poor?

FUNDAMENTAL CHALLENGES OF PREDICTABILITY IN SPECIATION

Inspired by the orrery, let us engage in a thought experiment that will guide us throughout this essay. In this experiment, we envision an evolutionary machinery driving the tree of life. At the tips of this tree, new species occasionally form (Fig. 2A). For now, let us assume that the evolution of a new species is always the result of a general process where a common set of components in identical configuration is at work within the evolutionary machinery. These components in their standard configuration thus represent the one evolutionary process essentially generating all new species.[10]

If this hypothetical "speciation machine"[11] was rigid and simple, speciation would be deterministic, easy to understand, and hence completely and straightforwardly predictable (Fig. 2B). While this does not seem to be the case, predicting speciation neither seems completely impossible (Fig. 2H). Indeed, many biologists would argue that speciation should in principle be accurately predictable, but that making accurate predictions will always remain very difficult. Why?

One argument for this difficulty is that speciation is highly complex. Complexity itself can have at least three different causes, which we can picture as different representations of our speciation machine. (i) "Intricacy" (Fig. 2C): the number of components—and hence their intricate interactions—can be overwhelming. Intricacy could result even if each relationship between very many components was simple. (ii) "Complicated relationships" (Fig. 2D): interactions among components may not be simple, but irregular or context-dependent. Such complicated relationships can especially reduce predictability when small variation in one component snowballs to have large and, in practice, almost unpredictable chaotic effects, or when one component qualitatively affects the relationship between other components. (iii) "Variable effect sizes" (Fig. 2E): the effects of components may vary both within and among different cases. So even with the same components in identical configuration, their relative importance within the same machine could vary strongly. Altogether, our ability to accurately predict speciation may be limited by complexity arising from numerous, complicated, or variable interactions among the components of the machine explaining all speciation.

A second, more fundamental argument for why predicting speciation is so hard, is chance. With chance, unknowable features of the universe cause true, unbreakable limits to our ability to predict speciation. While new species certainly do not only evolve through chance alone (Fig. 2H), chance could still play a critical role thereby making the speciation machine nonrigid and hence inherently stochastic. We can think of two plausible ways for how chance can arise: (i) "Contingency" refers to the interaction of two (or more) causally unrelated events. While the consequence of such an interaction is in principle foreseeable and thus predictable, its occurrence is not: it is coincidental (Fig. 2F).[12] However, what seems contingent may prove not to be as we gain more knowledge and, for example, switch the level of observation. For instance, an asteroid striking our planet seems like a contingent event to the biologist, but likely not to the modern astronomer. Contingency could thus

---

[10] Although this mechanistic metaphor bears the risk of oversimplifying or distorting biological reality, it proves useful for illustrating fundamental challenges of predictability. The power of mechanistic models lies in their mechanical representation of causal relationships.

[11] Note that we use "speciation machine" in this thought experiment as a visual heuristic to characterize speciation.

[12] We sometimes say that "Y is contingent on X." By this we mean that we cannot predict whether X will occur, but if it were to occur, we can predict it will cause Y.

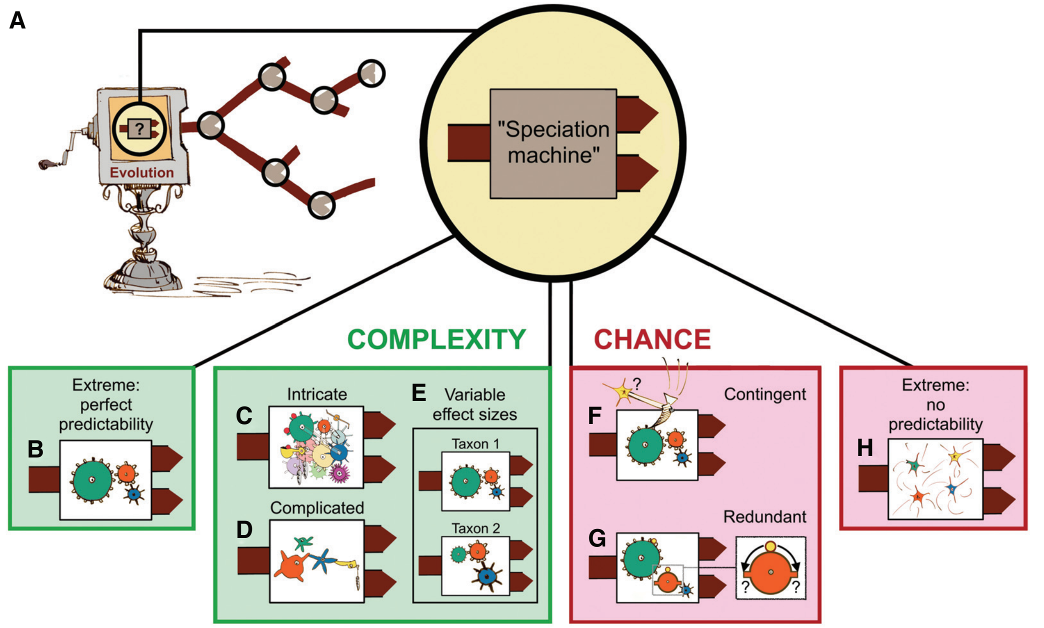


Figure 2. Thought experiment: fundamental challenges for predicting speciation. (*A*) Imagine an evolutionary machinery that produces new species at the tips of the constantly evolving tree of life. We assume that a single speciation machine drives the evolution of all new species (black circles). Speciation could be deterministic and thus in principle fully predictable (green boxes). In this case, our inability to make accurate predictions derives from complexity of the speciation machine, which should resolve as we gain greater conceptual and informational knowledge. On the other hand, speciation could involve true chance, and thus be stochastic and principally unpredictable to some extent (red boxes). In this case, even perfect knowledge could not yield perfectly accurate predictions. We posit that the speciation machine is neither (*B*) simple and rigid allowing for perfect predictability, nor (*H*) ruled completely by chance. Rather, the machine appears to be complex: different forms of complexity concern (*C*) the number of interactions (intricate), (*D*) complicated interactions (including "irregular" interactions as illustrated here), and (*E*) variation in the strength of interactions (variable effect sizes, i.e., variation in the relative importance of different factors). Apart from complexity, chance stemming from (*F*) contingency or (*G*) redundancy could also constrain us from accurately predicting speciation. Theoretically, all aspects of complexity and chance can occur simultaneously within the speciation machine.

be an artifact of ignorance.[13] (ii) "Redundancy" is another plausible way for chance to arise. Opposite to contingency, a redundant situation is in principle foreseeable and predictable, but not its outcome: it is random. Imagine a perfectly round marble on top of a wheel in our speciation machine; while we can predict this redundant situation, we are unable to foresee the marble's exact course (Fig. 2G). While chance from redundancy appears possible in theory, perfectly redundant situations may be rare or nonexistent in biological reality. Just as for contingency, ignorance may also be the source for what we deem redundant.[14]

---

[13] Ultimately, the concept of contingency requires an open system where causalities from the "outside" are possible: there is no contingency in a closed (and deterministic) system where a single initial event sparked all causality. Currently, contingent quantum effects seem to be the only known source of chance arising from contingency; however, some scientists propose some sort of "superdeterminism" in quantum mechanics and hence dismiss true contingency even here.

[14] For the sake of completeness, we should also consider "emergence." In its strong sense, we can neither predict the occurrence of an emergent phenomenon nor its consequence; indeed, up to its occurrence we may not even have considered its existence. In most cases, our ignorance may influence what we deem to be emergent. This is probably why emergent phenomena are often thought to result from high degrees of complexity (weak sense of emergence).

So far, our thought experiment has illustrated possible fundamental challenges for predicting speciation. Challenges arising from complexity are theoretically surmountable: as our understanding improves and as we gain more and better data, complexity should ultimately no longer limit our ability to predict speciation (Fig. 2C–E). While achieving this will obviously be very difficult, genuine limits may always preclude full predictability: true chance renders speciation stochastic to some degree, making predictions intrinsically uncertain (Fig. 2F,G). In reality, however, much of what we currently perceive as chance might just constitute an extreme case of complexity. As in all empirical sciences, the real limits to our predictability will only become clear as we proceed. Yet, why should we care about predictability at all?

FUNCTION AND VALUE OF PREDICTIONS

The notion that strong scientific understanding should yield highly accurate predictions requires taking a closer look at the general methodological framework in empirical sciences and the function and value of predictions therein.

All empirical sciences initially build on data gathered from nonsystematic observations (exploration). General patterns emerging from many such exploratory observations together inform a conceptual model by induction. Predictions deduced from such a model via general hypotheses are tested by targeted investigations, ideally using manipulative experiments but also through systematic observations.[15] This verifies or falsifies predictions and hence substantiates or improves the underlying model (Fig. 3A–F). This feedback loop of induction and deduction is at the heart of all empirical sciences and essentially relies on predictions as a powerful methodological tool. For instance, the targeted observation of the sky in Antsirabe on 4 September 2100 will test the prediction of a solar eclipse, thereby challenging our understanding of the solar system. Indeed, this prediction would present little scientific value if we assumed nobody could witness the eclipse. This illustrates why scientific predictability entails testability, and why high predictive success commonly functions as a hallmark of strong scientific understanding.

However, predictive success alone can be a misleading measure for the quality of our understanding. For instance, according to some historical records, Thales of Miletus successfully predicted a solar eclipse for 28 May 585 BC although he was most likely unaware of the heliocentric solar system. How to explain this paradox? (i) It is possible that Thales claimed the eclipse without any reasoning. If so, he was just lucky: his prediction was not scientific and similar future claims—perhaps more appropriately called "prophecies"—would have failed due to lack of concept. (ii) Because Thales was a passionate scientist, it seems plausible he had some sort of concept of our solar system from which he deduced his prediction. In this case, however, this concept would have been at the very least incomplete and his prediction highly uncertain.[16] Thales' high predictive success would thus still have involved luck and future predictions would likely have failed or lacked comparable accuracy. Together, these two possibilities highlight why predictability in empirical sciences involves not only testability but also repeatability: only if repeated can high predictive success of a model serve as a hallmark of strong scientific understanding.[17] (iii) It is also

---

[15] This does not mean that scientific progress only works via explicit hypothesis testing or via falsification of null hypotheses. Yet, in fundamental sciences, insights from exploration untargeted with respect to specific predictions will eventually also require an interpretation with respect to the current conceptual understanding (model).

[16] Today, we often account for uncertainty—if we are aware of it—using "probabilistic models." Indeed, such models can be used for predicting very specific events and may also have high precision in estimating the probability of an outcome. The incorporation of uncertainty in probabilistic modeling can be particularly powerful in applied situations, and more valuable to fundamental research, point to key conceptual or informational knowledge gaps. It remains however unknown, at least temporarily, to what extent the modeled uncertainty stems from ignorance and/or real chance. It is possible that probabilistic models will eventually turn into deterministic ones.

[17] Notably, even an incomplete or wrong model can sometimes produce repeated predictive success. For example, a mythological or rather crude geometrical concept explaining the movement of celestial bodies could allow for repeated and relatively accurate predictions. Upon further testing, however, the underlying model should become refined and eventu-

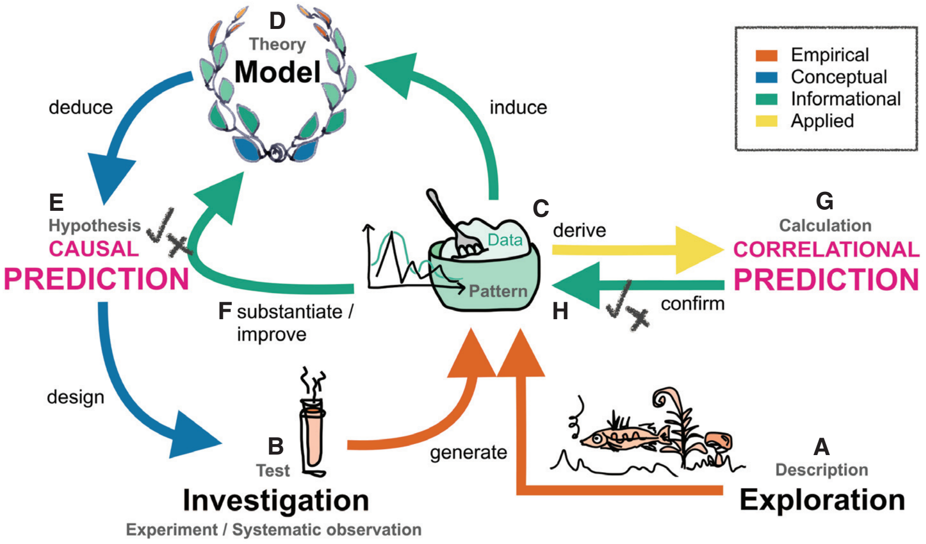


Figure 3. Role of predictions in science. (*A*) Exploration lies at the origin of all empirical sciences, creating data by description. (*B*) Data is also gained through experimental and observational investigations targeted to test general hypotheses or specific causal predictions. (*C*) Data are central in two ways. First, by inductive inference (i.e., inference from single discoveries to generally valid concepts) we detect patterns in data. Within a given theoretical framework these patterns serve us to build (*D*) a causal model. From such a model, we formulate (*E*) general hypotheses and specific causal predictions via deduction (i.e., inference from a generally valid concept to a specific case). Investigations targeted to test these predictions allow (*F*) verifying or falsifying them, thereby validating (substantiating or improving) the hypotheses and the underlying model and theory. As a whole, this constitutes the inductive–deductive feedback process central to all empirical sciences. Second, patterns detected in data allow for (*G*) correlational predictions. Correlational predictions are mostly of applied value: they can yield high predictive success even without understanding the underlying causal processes. (*H*) Testing correlational predictions through “simple” calculations may confirm a detected pattern, thereby inspiring causal reasoning and informing causal models. However, confirming predictions from correlations cannot directly validate causal models because correlations can imply causality where there is, in fact, none (i.e., spurious relationships).

possible that Thales systematically inferred his accurate prediction without any model and conceptual deduction at all. Instead, Thales might have based his prediction simply on recurrent previous observations. If so, he employed a general pattern to derive a “correlational prediction” (Fig. 3G). Correlational predictions are thus fundamentally different from “causal predictions” (Fig. 3E), which are deduced from a conceptual model.[18] Because correlational predictions allow for blind extrapolation from detected patterns, they are relatively straightforward, often function to guide practical decisions, and are hence of great applied value. Importantly, success of such predictions confirms a detected pattern (Fig. 3H), but bears the danger of suggesting an underlying—possibly even a simple and rigid—causal mechanism when there may in

---

ally be replaced with a model yielding greater accuracy across all of its predictions. Indeed, ancient precursors of the modern orrery such as the Antikythera mechanism seem to have successfully predicted solar eclipses regardless of the underlying model being geo- or heliocentric. And although powerful for predicting solar eclipses later in history, even the modern orrery based on classical (Newtonian) physics proves to be a rather crude approximation of reality when considering the “three-body problem” in classical physics or the theory of general relativity in modern physics.

[18] Like causal predictions derived from mechanical devices such as an orrery, correlational predictions also have a long history in astronomy. For instance, the Mayan Calendar was a pre-Columbian, mythological written calendar that produced relatively accurate predictions of lunar and solar eclipses by correlating the position of Venus to the Sun.

fact be none. So, while correlational predictions can inspire causal thinking, they never serve as direct tests of cause.

Although Thales' predictive success remains mysterious, his story clarifies the function and value of predictions in empirical sciences: through repeated testing, causal predictions allow validating and improving conceptual models. What does this mean for our thought experiment? The speciation machine should become represented by a generally valid model of speciation amenable to repeated empirical testing via causal predictions. Although our main interest in speciation research is to understand speciation conceptually, staying with an overly "abstract" model is hence not sufficient simply because it will not produce specific enough predictions that can be tested. How to build such a predictive "standard model of speciation"?

TOWARD A STANDARD MODEL OF SPECIATION

The quest for a standard model of speciation is not far-fetched as similar endeavors are known from other fields of research. Apart from the search for the "theory of everything" in physics, meteorology provides a useful example as it employs a standard model of atmospheric dynamics for making specific weather forecasts. Although probabilistic in nature, continuous improvement of this model for more than a century has resulted in better short- and long-term meteorological predictions. To bring our imaginary speciation machine to life, what could a similar standard model look like in speciation research?

A standard model of speciation should capture all components of the speciation machine and their configuration. This should allow us to successfully predict not only generative mechanisms at all levels of biological organization leading to lineage branching (species birth), but also destructive mechanisms leading to lineage extinction (species death), both forward and backward in time. This model should also predict evolution within lineages relevant to speciation, such as species persistence, and inform us on net-speciation rates as well as spatial and temporal patterns of species diversity (see Box 1).

An obvious first step toward this model is the identification of the components making up the speciation machine. In a second step, these components and their configuration should become parameterized and operationalized. At best, this would eventually result in a mathematical expression of the model. Importantly, in addition to consensus on the explanatory parameters and their interactions allowing for empirical assessment, the speciation-metric(s) to be predicted requires clarification. This concerns not only the long-debated question of how to define "species," but also what it is we want to know about their origin. To date, there is no clear consensus on what the metric(s) across all of life might be, other than it should capture some aspect of lineage splitting. We may desire a simple (qualitative) metric of lineage status across all time points (e.g., presence/absence/split) or a multivariate (quantitative) speciation index capturing many aspects of lineage dynamics.[19] Regardless, we require testable predictions that are, ideally, operational across different timescales and biological systems. Ultimately, a standard model of speciation should capture and predict all aspects and metrics relevant to lineage branching, persistence, and extinction.[20]

In more practical terms, how could this endeavor be structured? It seems intuitive to first collect and systematically catalog our current knowledge on speciation. Because different research fields and disciplines (e.g., population genetics, genomics, (macro-)ecology, biogeography, paleontology) inform speciation from different perspectives or at different biological levels, this catalog would entail diverse and complementary insights. Besides identifying compo-

---

[19] Candidates include reproductive isolation (RI), the order of evolution of isolating barriers, genetic and phenotypic divergence, the probability of irreversible divergence of lineages, the timing, rate, or likelihood of speciation, and binary metrics for complete splitting (yes/no) and extinction (yes/no) of lineages. Note that RI, for instance, is not a suitable measure of speciation in asexual life forms.

[20] This model should also predict all known major patterns of RI, including the two known "rules of speciation" (Haldane's Rule and "the large X effect"; Coyne 2018), if they were involved in the speciation process (and did not accumulate after speciation), and should allow identifying the diversity and frequency of their underlying causes.

BOX 1. REMARKS ON PREDICTING EVOLUTION VERSUS SPECIATION, AND ON PREDICTION VERSUS RETRODICTION

Just as predicting evolution, predicting speciation generally evokes the idea of making proper forecasts into the (distant) future. In both endeavors, such forecasts can be made for applied reasons, such as in medicine, agriculture, or conservation (Carroll et al. 2014; Smith et al. 2014; Wortel et al. 2023). Yet, when asking scientists about the value of predictions, many—and especially the ones working in basic research—highlight predictions as a measure of confidence in our current concepts through predictive success (see Function and Value of Predictions section).

Predicting Evolution versus Speciation

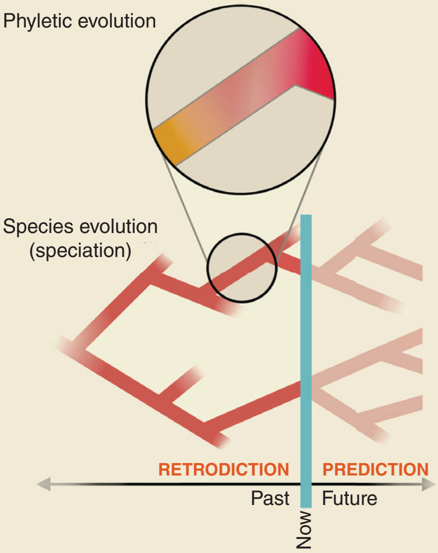
Reflections on the "predictability of evolution" have a long-standing tradition in biology. Paleontologists sparked debate on contingency versus predictability of broad-scale patterns of biodiversity ("replaying life's tape") (Gould 1990; Conway Morris 2003) that continues to this day (Losos 2018). Research today is often framed in the context of adaptation (parallelism and convergence) and centers around short- to mid-term predictions of phenotypic values or allele frequencies (Grant and Grant 2002; Shaw 2019; Nosil et al. 2020; Reiskind et al. 2021) and genetic components underlying phenotypic change (Stern and Orgogozo 2009). Does the "predictability of speciation" deserve special consideration? While the major goals for predicting evolution and speciation are the same, the two pursuits and their challenges are not identical. First, even though required for both evolution and speciation, time poses a greater practical challenge for speciation research: while evolution can be readily witnessed and tested in only a few generations, full speciation commonly takes much longer. Second, evolutionary change can occur without speciation: evolution happening within lineages—termed "phyletic evolution" (see figure in this box)—has been and continues to be ongoing throughout all of life. But speciation cannot occur without evolution, as all speciation involves an evolutionary mechanism driving lineage divergence (Langerhans and Riesch 2013; Schuler et al. 2016). Third, speciation as a special case of evolution is a relative concept inherently involving multiple lineages—the term "species" is tied to comparing at least two lineages. Thus, unlike phyletic evolution, speciation is not only relative to time but also to closely related biological entities. It therefore involves multiplication of organismal diversity, generally understood as lineage dynamics illustrated by phylogenetic trees depicting species birth (bifurcation), species persistence, species death (extinction), or even the merging of species (lineage fusion). This means that speciation always occurs with temporal (but not necessarily spatial) coexistence of the branching lineages: during initial splitting of a lineage, both lineages must persist, at least momentarily, for speciation to occur. Thus, predicting evolution includes within-lineage evolutionary change while predicting speciation explicitly centers on multilineage evolution. How much of evolution is irrelevant for speciation? Perhaps very little, as so many aspects of evolutionary change can influence speciation (Coyne and Orr 2004). A greater understanding of the predictability of speciation should help uncover this answer.

Prediction versus Retrodiction

Testing the predictability of speciation is challenged by the time required for speciation to occur for many taxa relative to human generation time (Blount et al. 2018). This can be partly overcome by studying organisms with short generation times, using in silico experiments, or by conducting long-term experiments that persist over the course of many researchers' lives. A great potential to tackle the challenge of time in speciation research stems from paleontology: scientific predictions derived from causal models can not only be tested by forecasting future events ("prediction" proper) but also by examining past events ("retrodiction") (Lloyd 2015; Roesti 2021). In fact, together with sufficient high-quality data, a generally valid standard model of speciation should accurately predict both past and future speciation (see figure in this box), reminiscent of an all-knowing creature such as Laplace's Demon (Laplace 1820). In a deterministic universe, this demon can pre- and retrodict every state, with time providing no constraint. Because of the waiting-time for future speciation events to

Continued

potentially occur, nearly all speciation research has so far focused on retrodictions. However, a challenge of retrodictions lies in our inability to properly consider chance and its importance when we only look at the past. This is because chance appears inevitable in retrospect. For example, if the Chicxulub asteroid impact 66 million years ago reflected a truly contingent event, we have the benefit of hindsight in knowing of its occurrence. If chance plays a critical role in speciation, predicting future events will prove more useful in assessing its importance. On the other hand, if true chance is rare, *retro*dictions could provide tests of model predictions that are just as useful as *pre*dictions. Because the role of chance in speciation remains unclear, and because testing future speciation over the span of millions of years is out of reach, a combination of retrodiction and prediction should provide the best approach for testing and refining a standard model of speciation.



Box Figure. Phyletic evolution versus speciation, and prediction versus retrodiction. The magnifying glass shows evolutionary change in a lineage without speciation ("phyletic evolution"). The standard model of speciation should primarily predict "species evolution" (i.e., speciation including lineage branching, persistence, and extinction)—both forward ("prediction" proper) and backward ("retrodiction") in time. Shown is a simplified version of a species tree in which, for example, lineage fusions are not depicted.

nents essential to speciation, the catalog may already contain potential "submodels" of the standard model, empirical parameters, commonly observed patterns (sometimes formulated as "rules") hinting to possible direct or indirect causalities, and possibly even mathematical formulas (Fig. 4A).

In a next step, these diverse and complementary insights should be integrated into a prototype of the standard model of speciation (Fig. 4B). Because this prototype may still be quite conceptual and thus contain relatively abstract, idealized, or coarse speciation components not directly or easily measurable, perhaps the biggest challenge will be to resolve the conceptual components through continuous refinement of the model into concrete, measurable parameters (Fig. 4C).[21] To understand how the interactions of parameters generate lineage dynamics, we would likely have to consider also other fields traditionally not directly linked to speciation, such as physics, chemistry, or meteorology. Overall, this promises to resolve (apparent) ambiguities, uncover blind spots, and clarify previously unclear causal relationships and "rules" in speciation, both conceptually and empirically. It is further possible that some parameters (or

---

[21] Indeed, researchers may have different perspectives on the status of speciation components: what some may conceptualize as a highly abstract component, others may already operationalize as a concrete empirical parameter. For example, "phenotype" or "fitness" may be perceived as very abstract and idealized concepts by some, but as concrete and real (and mathematically useful) parameters by other researchers. Diverse opinions could derive from different definitions of involved components and parameters, respectively, and constructive debate should facilitate the operationalization of the standard model of speciation.

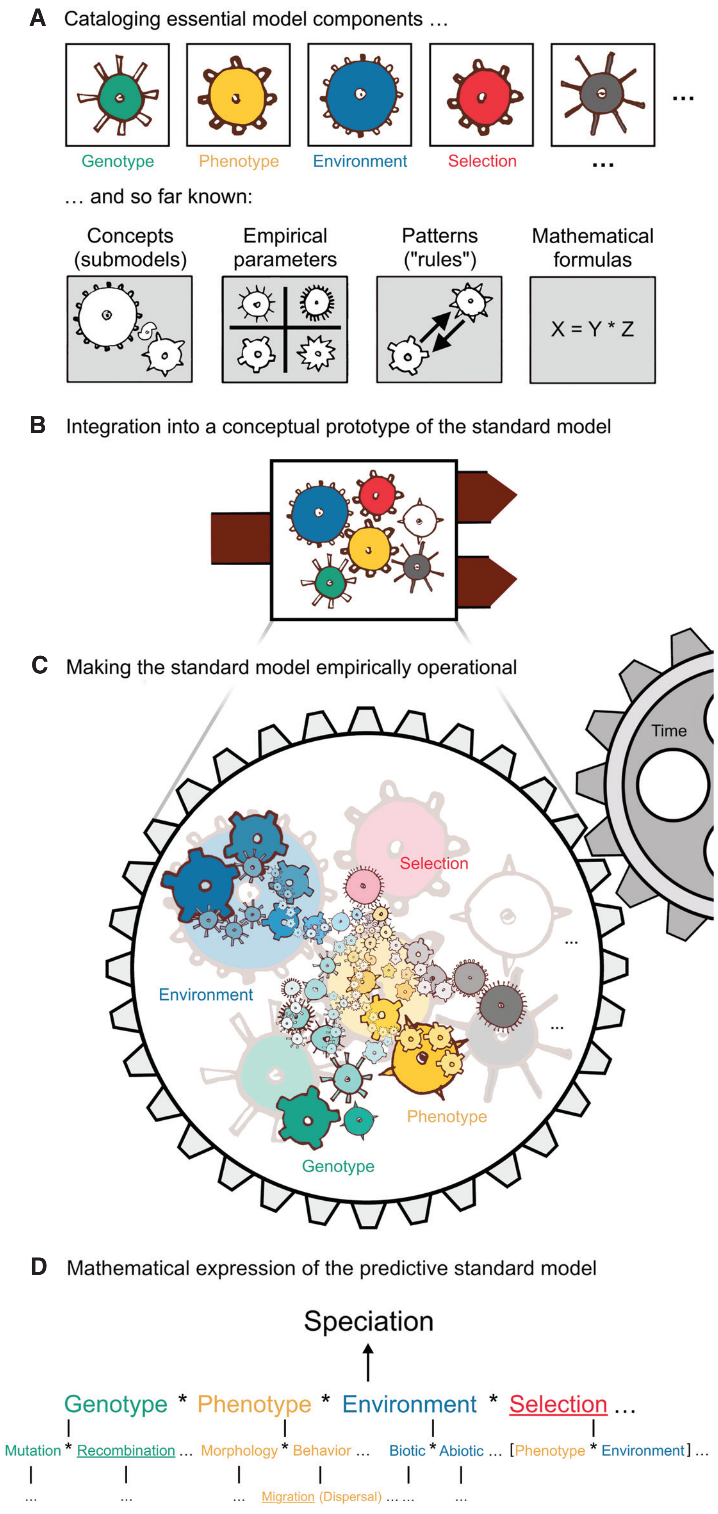


Figure 4. (*See following page for legend.*)

components) assumed essential to speciation or operationally useful will become understood to emerge from causal interactions of others. While such compound parameters (or components)—whether or not intentionally installed for practical reasons—could resolve into more fundamental ones or vanish,[22] others might need to be integrated into the standard model.

Finally, the operationalization of these parameters, ideally in the form of mathematical expressions, should generate testable predictions (Fig. 4D). The best way to mathematically express the standard model remains to be determined and could take a variety of forms. Hierarchically nested equations (inspired by the standard model in particle physics) or a 3D interaction network could prove useful to explain how parameters interact with one another, not only within but also across different levels of biological organization even in highly complex ways. Initially, the standard model may produce rather unspecific (broad) and/or imprecise predictions and concern short timescales, necessitating the incorporation of much uncertainty in the form of "probabilistic modeling" (see footnote 16). The model may also first be tested by

Figure 4. Building the standard model of speciation. Illustrative example of how we might first build a conceptual prototype of the standard model of speciation (concept) and then move to a predictive standard model amenable to empirical testing (practice). (*A*) Cataloging existing insights on speciation from all research fields and scales of biological organization relevant to speciation. This catalog should include conceptual components essential for predicting lineage dynamics (e.g., fitness or niche; more exemplary terms are given below each box), but also already existing concepts (i.e., "submodels," e.g., for sympatric speciation), known empirical parameters (e.g., effective population size, metabolic rate, chromosome number), patterns (i.e., "rules" including correlations, see footnote 20), and mathematical models (i.e., "formulas" such as the breeder's equation). (*B*) Integrating the collected insights into a conceptual prototype of the standard model that is meant to represent the "speciation machine" in our thought experiment. (*C*) Making the standard model concrete and thus empirically operational through parameterization and operationalization. Relatively abstract components together with their configuration (shown in the background of panel *C*) should thereby resolve into interacting explanatory parameters amenable to empirical assessment. Continuous refinement and improvement of the model—perhaps structured hierarchically (illustrated by fading colors)—should result in a more and more fine-grained model producing increasingly accurate predictions. In this process, already existing and collected submodels, parameters, rules, and formulas relevant to speciation will also become integrated and refined. In this way, the standard model should describe how underlying causal parameters lead, for instance, to matrices that interact with one another to predict lineage dynamics. We note that the illustration in *C* conceptualizes "time" as an outstanding parameter; however, time could instead also be integrated on the deepest level of the model hierarchy. (*D*) The model should eventually allow for the prediction of any desirable outcome-metric tracking speciation and species extinction across the tree of life using mathematical expressions. The highly exemplary portion of the mathematical speciation model shown uses a hierarchically nested expression to highlight how the relatively conceptual components "phenotype," "genotype," "environment," and "selection" from *A* and *B*—once they are parameterized and operationalized as shown in *C*—may ultimately predict speciation. The nested structure could depict, for example, how "environment" may resolve into "biotic" and "abiotic" environment, each of which could further resolve and be parameterized as sets of biotic and abiotic factors and their specific interactions. "Selection" could perhaps be integrated and (partially) modeled as the result of the "phenotype"–"environment" interaction. This way, speciation may become predictable on different hierarchical levels because essential components "scale up" to causally drive lineage dynamics. In this process, some components or parameters initially considered essential to speciation or operationally useful may sooner or later prove to be epistemic artifacts and vanish (e.g., the white wheel in *B* and *C*). On the other hand, higher-level phenomena may emerge from interactions on lower levels (as indicated as a layer of small wheels on big wheels in *C*; see footnote 14) and thus have their own right to be considered in the standard model. The underlined terms in the depicted exemplary portion of the mathematical model in *D* capture aspects directly affecting the extent of gene exchange (e.g., effective migration rate) between populations. This involves the interaction of migration (m), recombination (r), and selection (s) (e.g., $m_e = m/(1 + s/r)$; Barton and Bengtsson 1986). Migration may be modeled in terms of an organism's behavior (which is part of its phenotype), recombination as an aspect of its genotype, and selection is influenced by interactions between an organism's phenotype and its environment. Throughout the figure, "…" denotes the existence of additional, undepicted features, and "*" stands for any possible type of mathematical operation.

predicting patterns directly measurable in nature without the need of laborious experiments, such as parallel speciation and spatial patterns of species diversity. Through integration of discoveries from different disciplines, approaches, taxa, and biological levels, the model should continuously improve and yield more and more specific and accurate predictions, also across longer timescales. Eventually, the model should bridge macro- and microevolutionary aspects of speciation and reveal how these different levels at which we study speciation relate.[23]

It is clearly beyond our capacity to present a clear and tangible route forward here; our goal is instead to highlight the epistemic value of (working toward) a standard model of speciation and stimulate discussion. While this undertaking will undoubtedly be challenging and may even feel overwhelming or unrealistic as of this writing, the last century surely taught us that the human endeavor within the scientific enterprise can accomplish remarkable feats. In fact, researchers seem to already be applying parts of a standard model of speciation by testing specific aspects of it (e.g., see Fig. 4 and footnote 20). Working toward a standard model of speciation will clearly require a highly integrative effort, possibly even warranting its own expertise and discipline. After all, we may come to conclude that the challenge of predicting speciation is primarily rooted in real-world (empirical) complexity rather than in conceptual complexity. We could also come to realize, however, that predicting speciation across the tree of life fundamentally requires more than just one model. Such insight would profoundly impact our understanding of speciation and how we research it.

ONE OR SEVERAL SPECIATION MACHINES?

We have intentionally started our thought experiment with the simple assumption of a single speciation machine underlying all speciation. Yet, what if there is more than just one speciation machine? To follow up on our thought experiment, such "speciation pluralism" would result from speciation machines built of fundamentally different sets of components and/or configurations (Fig. 5A). Opposite to a single model that captures a wide diversity of pathways to speciation in a common framework—possibly even entailing sexually and asexually reproducing taxa[24]—speciation in a pluralistic scenario would happen in fundamentally distinct manners across time or taxa. This would necessitate multiple models, and deducing predictions from a single model of speciation would be wrong. In fact, there may simply not be one explanation for *the* origin of species but multiple explanations for (many) different origins. Thus altogether, speciation pluralism, complexity, and chance present the three fundamental challenges for predicting speciation.

In the case of extreme pluralism, speciation would be fundamentally ungeneralizable.[25] While taxa obviously vary in specific properties

---

[22] For instance, genetic drift may appropriately be captured by parameters describing variance in reproductive success and population size, both of which are contained within parameters of the organismal trait space (mating system, life history traits) and their interaction with the environment. "Random" genetic drift may thus not imply true randomness in the sense of redundancy (see Fig. 2G), but rather an approximation reflecting our lack of understanding (ignorance). Similar considerations may apply, for instance, for the "randomness" of mutations or functional "redundancy" of phenotypes.

[23] Perhaps, unifying our understanding of micro- and macroevolution may be of similar importance to biology as is the unification of quantum mechanics and the theory of general relativity to modern physics.

[24] It may seem obvious that at least two models of speciation need to be conceptualized, one for sexually and one for asexually reproducing taxa. However, it is well conceivable that one and the same standard model of speciation can reflect both cases of reproduction and that in the case of asexually reproducing taxa some model parameters have simply (close to) zero effect.

[25] Here, a speciation machine and a "model" organism could only help understand speciation within a very narrow parameter space (e.g., within a particular lineage). As a consequence, this would cause "model" organisms to effectively only model themselves and not speciation in general. This illustrates why speciation pluralism does not reflect some extreme form of a previously described challenge (e.g., dramatically different effect sizes of components among taxa, or chance playing a critical role in some taxa but not in others; see Fig. 2), but would instead comprise a distinct phenomenon challenging our ability to predict (and understand) speciation.

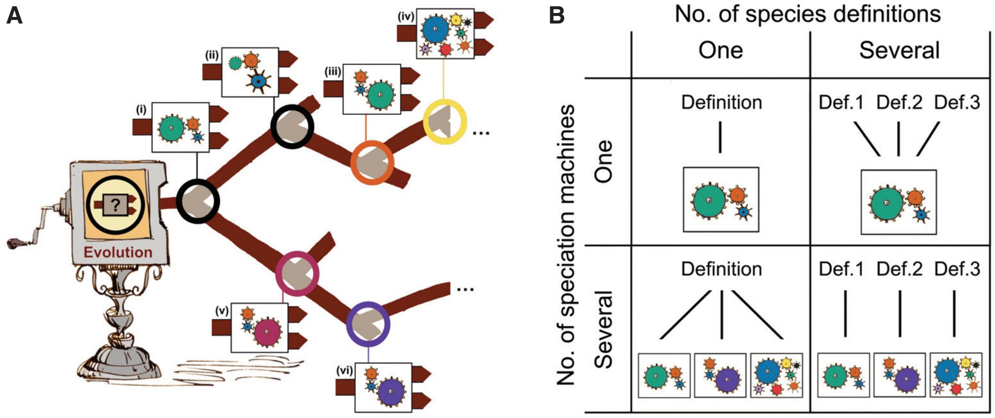


Figure 5. Speciation pluralism. (*A*) Variation in the importance of different components alone makes speciation complex, but it does not fundamentally challenge the idea of a single speciation machine underlying all speciation (compare *i* vs. *ii* [black circles]; see also Fig. 2E). However, speciation pluralism (indicated by differently colored circles) could arise for two reasons: different causalities of the same components (*i* vs. *iii*), or different components including unique causalities (e.g., *i* vs. *iv*). Theoretically, pluralism could vary in strength between machines built of quite similar to fundamentally different components and/or causalities (compare e.g., *v* vs. *vi* to *i* vs. *iv*). (*B*) In a world with only one speciation machine, one common process generates all the entities we recognize and call "species" (*upper left* panel). In this case, a single species definition would reflect biological reality. In contrast, different machines could generate what we generally recognize as "species" (*lower left* panel). In this case, species might still be united in a common definition although they would emerge in fundamentally distinct manners: one common definition of species does not imply one speciation machine. It is also possible that each of several speciation machines generates a unique type of entity, thus requiring different definitions of species (*lower right* panel). In this case, different species definitions would reflect different underlying evolutionary causes. In any case, we should not be too hasty in giving up the idea of a single speciation machine simply based on our inconsistencies and ongoing debate on how we recognize and define "species": the lack of consensus may not reflect biological reality and erroneously imply speciation pluralism where there is, in fact, none (*upper right* panel).

(e.g., possess fundamentally different traits), and while different factors (e.g., different isolating barriers) can play roles of varying importance among different cases of speciation, our present insights do not support an extreme view of pluralism. Still, speciation pluralism needs serious consideration: how would the existence of several speciation machines impact how we view and study speciation? A systematic extension of the standard model of speciation would be required, including information on which specific machine is at work in which particular context. Indeed, the speciation machine itself could evolve, resulting in different (but more or less related) machines within certain parts of the tree of life (Fig. 5A). A plurality of speciation machines would not necessarily reduce the predictability of speciation if we could accurately determine the applicable machine in each case. It would, however, have deep implications for our general understanding of speciation.

To know whether and to what extent speciation pluralism challenges our quest for one standard model of speciation, we should not be tricked by the various ways we define "species" (Fig. 5B). While inconsistencies and controversies in this regard may suggest the existence of different speciation machines, the various species definitions could also lack real biological justification. Ideally, we would all operate with one and the same definition of what we recognize and call "species." Yet, this seems hardly feasible because even if we had complete consensus on the conceptual notion of species, diagnosing them across the entire tree of life and various timescales may require different definitions and methodologies. In fact, diversity in definitions and methodologies, as long as they are consistent and explicit, could facilitate our ability to identify the operationally best species definition across different approaches, build a refined model, uncover the existence/absence of pluralism, and ultimately gain an all-inclusive view on speciation.

PROMISING AVENUES TOWARD A PREDICTIVE SCIENCE OF SPECIATION

With the aforementioned challenges in mind, some avenues appear particularly promising for working toward a standard model of speciation. To construct a candidate model, a more concerted and integrated effort is required: research needs to systematically diversify with respect to the studied organisms and to unify its insights to better gauge the predictability of speciation and the possible sources of its constraints, including complexity, chance, and speciation pluralism (Fig. 6).

If thereby a single candidate model proved particularly promising, a joint conceptual, technological, and financial effort similar to the CERN enterprise in physics may prove particularly informative. Such a grand set of experiments could powerfully test and fine-tune the candidate model by gaining deep insights in what we conceptualize as essential but find difficult to discern in nature. Conducted under highly controlled settings, it could uncover blind spots in the model and assess the roles of complexity and chance. An obvious limitation of such an effort is that we could only study organisms with relatively short generation times and over relatively short timescales. Focusing on one (or few) organisms could then not only miss or underestimate speciation pluralism, but also lead to a biased or limited view of central components of the speciation machine if they vary in effect size across taxa (see Fig. 2E).

To address these limitations of CERN-like experiments, a joint effort of a large consortium of researchers testing the potential standard model in diverse taxa in field, laboratory, and computer-simulated contexts across disparate timescales should be maintained. Taxon choice would best be haphazard and representative of the tree of life. In a world governed by a single speciation machine, one model should generate similarly accurate predictions across all of life (see Fig. 6). Particularly strong tests of a model's general validity would come from predicting extreme cases of speciation, such as in taxa showing unusually high or low speciation rates or other extreme biological properties. While we must always keep in mind that by focusing too much on the uncommon we may miss the common, the study of such extreme cases seems particularly useful in resolving the question of speciation pluralism because testing a wide parameter space can quickly validate or refute model configurations. If we found strong indication for speciation pluralism, systematically studying diverse and haphazardly chosen organisms should also allow for an unbiased understanding of the frequency of different speciation machines and perhaps how they themselves evolve (see Fig. 5A).

Some may argue that faced with the many daunting challenges of long evolutionary timescales (see Box 1), irreproducible natural history, our inability to manipulate time, and the difficulty of perfectly replicating tests of model predictions —in addition to complexity, chance, and pluralism—working toward a truly predictive science of speciation is hopeless. We disagree. While directly observable patterns in nature (e.g., cases of convergent or parallel speciation) may be consistent with causal predictions and thus the underlying model, experiments paired with recent technological advances provide particularly promising and strong tests forward. Notable here are experimental (in vivo) speciation, in silico speciation, engineered speciation, the analysis of ancient DNA, as well as the various applications of artificial intelligence.[26] In any case, as long as all life in-

---

[26] Some remarks to this list: first, by "engineered speciation" we mean cases where humans create species through direct (genetic) manipulation of real organisms. This provides opportunity to directly test causal predictions. Second, although "ancient DNA" does not allow for manipulative experiments, it can serve as a time capsule and thus powerfully address questions about predictability via retrodiction (see Box 1). Finally, applications of "artificial intelligence," such as machine learning, are not only powerful tools for applied correlational and probabilistic predictions, but they also hold great potential to inform and inspire conceptual understanding. The danger of artificial intelligence-guided inferences is, however, to mistake correlational with causal predictions (see Function and Value of Predictions section).

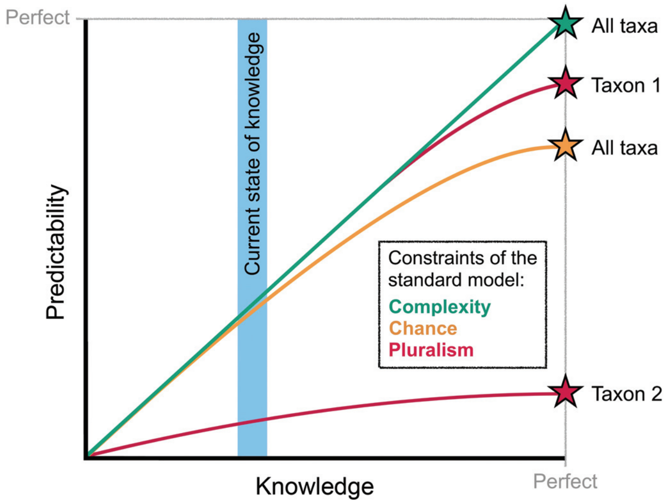


Figure 6. Expected relationships between knowledge and predictability of speciation assuming different constraints. In a deterministic world with one standard model of speciation, predictability can constantly increase (green line) and ultimately reach perfection (green star) as we gain more knowledge. In this case, limitations derive only from complexity, reflecting a lack of conceptual and informational knowledge. If chance plays an important role in speciation (orange line), predictability with a standard model would never reach perfection even with perfect knowledge (orange star). Another challenge to our predictability of speciation may be speciation pluralism, where the existence of multiple speciation machines requires multiple models (see Fig. 5A). If speciation was pluralistic, the relationship between knowledge and predictability would vary across taxa when using one common model: predictability could become (close to) perfect for one taxon (Taxon 1, *upper* red line and star) but remain poor for another (Taxon 2, *lower* red line and star). Arguably, with our current knowledge (blue bar) we are in an undecided situation how to explain our struggles in predicting speciation: the reason could be some combination of lack of knowledge (due to complexity), true chance, and speciation pluralism. (This figure is an extension of Figure 2 in Roesti 2021.)

volves a common set of causally connected features that similarly influence species birth, death, and persistence, the quest for a standard model of speciation seems viable and worthy even if it continues (far) beyond our generation of researchers. With serious crosstalk and conceptual unification within and across disciplines, concerted and integrative efforts involving large experiments and approaches from different angles toward common goals, combined with the gathering, analyzing, and synthesizing of big data and patterns using new technological approaches, we could be on the verge of entering a new era of truly predictive speciation research.

CONCLUDING REMARKS

Let us return to the predicted solar eclipse for 4 September 2100 in Antsirabe and ask: was it appropriate to draw parallels between speciation research and astronomy? Most speciation researchers would probably object, and we believe for good reasons.

Biology and physics are not the same. In astronomy, a simplified model of our solar system seems to suffice to conceptualize the phenomenon of a solar eclipse and generate accurate predictions despite ignoring real details. This works because the parameters essential to the model seem relatively stable over extremely long spatial and temporal scales. In comparison, biology can be highly sensitive to even tiny changes for which there are a multitude of possible causes. When compared to physical sciences, this makes biology perhaps more similar to meteorology than astronomy: it seems we need to know nearly "everything" for accurate predictions. This perhaps is why biology, just as meteorology, may rely on big data, complex models, and at least temporarily on the methodological incorporation of much uncertainty.

If we wanted to stick to drawing parallels between astronomy and speciation research, we should perhaps view our solar system as only one of a vast array of solar systems within a vast number of galaxies, while our eclipses of interest additionally range across an extremely small time span of the universe. So maybe a better analogy for predicting solar eclipses on Earth is to compare it to predicting speciation in a single taxon over very short timescales under highly standardized conditions with strong or even artificially imposed selection—something not out of reach for evolutionary biology today. On the flip side, perhaps predicting speciation across life on Earth is closer to predicting the formation of stars and planets in the entire universe—something quite challenging even to modern astronomy.

Finally, opposite to how we perceive speciation, the solar system's relatively cyclical and stable nature makes it appear highly deterministic and thus predictable. Yet, this perception might be misleading as speciation could in reality be as (non)deterministic as our solar system. For both cases, this boils down to the same fundamental question: are our limits in understanding and predicting literally in our predict*in*ability or in true *un*predictability? The quest for a standard model of speciation may not only transform our understanding of the origin of species, but also of evolution in general and perhaps even the world as a whole.

AUTHOR CONTRIBUTIONS

Overall conceptualization: M.R. and H.R. (lead), B.L. and J.W. (major), I.S. (minor). Original draft: M.R. and H.R. (lead), B.L. and J.W. (minor). Review and editing: M.R., H.R., and B.L. (lead), J.W. (major), I.S., J.B., and S.C. (minor). Conceptualization of figures: M.R. and H.R. (lead), B.L., J.W., and I.S. (minor). Visualization of figures: H.R. (lead), M.R. (major), I.S. (minor). Overall supervision: M.R.

ACKNOWLEDGMENTS

We thank Daniel Ortiz-Barrientos, Diana Rennison, Michi Tobler, and members of the Peichel Laboratory in Bern and the Wolf Laboratory in Munich for helpful comments and insightful discussions. Gary Lackman filled our lack of knowledge on meteorological forecasting with valuable insights. Daniel Bolnick, Andrew Hendry, and one anonymous reviewer provided thoughtful comments that further improved this essay. Finally, we feel obliged to everyone who has, to us knowingly or unknowingly, motivated and influenced us in writing this essay. We tried to approach our writing for this Special Volume on Speciation as free-minded as possible. The cited references were kept to a minimum and are meant to highlight some of the many writings on the topics discussed. We really are dwarfs standing on the shoulders of giants to whom we feel very much obliged.

REFERENCES

Barton N, Bengtsson BO. 1986. The barrier to genetic exchange between hybridizing populations. *Heredity (Edinb)* **57:** 357–376. doi:10.1038/hdy.1986.135

Blount ZD, Lenski RE, Losos JB. 2018. Contingency and determinism in evolution: replaying life's tape. *Science* **362:** 6415. doi:10.1126/science.aam5979

Carroll SP, Jørgensen PS, Kinnison MT, Bergstrom CT, Denison RF, Gluckman P, Smith TB, Strauss SY, Tabashnik BE. 2014. Applying evolutionary biology to address global challenges. *Science* **346:** 1245993. doi:10.1126/science.1245993

Conway Morris S. 2003. *Life's solution: inevitable humans in a lonely universe*. Cambridge University Press, Cambridge.

Coyne JA. 2018. "Two rules of speciation" revisited. *Mol Ecol* **27:** 3749–3752. doi:10.1111/mec.14790

Coyne JA, Orr HA. 2004. *Speciation*. Sinauer Associates, Sunderland, UK.

Espenak F, Meeus J. 2006. Five millennium canon of solar eclipses: −1999 to +3000. NASA Tech Publ TP-2006-214141. NASA, Greenbelt, MD.

Gould SJ. 1990. *Wonderful life: the Burgess Shale and the nature of history*. Norton, New York.

Grant PR, Grant BR. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296:** 707–711. doi:10.1126/science.1070315

Langerhans RB, Riesch R. 2013. Speciation by selection: a framework for understanding ecology's role in speciation. *Curr Zool* **59:** 31–52. doi:10.1093/czoolo/59.1.31

Laplace P. 1820. *Essai philosophique sur les probabilités* (Paris) [*A philosophical essay on probabilities*]. (transl. by FW Truscott, 1951) (ed. Truscott FW, Emory FL). Dover, New York.

Lloyd EA. 2015. Model robustness as a confirmatory virtue: The case of climate science. *Stud Hist Philos Sci A* **49:** 58–68. doi:10.1016/j.shpsa.2014.12.002

Losos JB. 2018. *Improbable destinies: fate, chance, and the future of evolution*. Riverhead, London.

Nosil P, Flaxman SM, Feder JL, Gompert Z. 2020. Increasing our ability to predict contemporary evolution. *Nat Commun* **11:** 5592. doi:10.1038/s41467-020-19437-x

Reiskind MOB, Moody ML, Bolnick D, Hanifin CT, Farrior CE. 2021. Nothing in evolution makes sense except in the light of biology. *Bioscience* **71:** 370–382. doi:10.1093/biosci/biaa170

Roesti M. 2021. Evolution: Predictability and the promise of ancient DNA. *Curr Biol* **31:** R446–R448. doi:10.1016/j.cub.2021.03.063

Schuler H, Hood GR, Egan SP, Feder JL. 2016. Modes and mechanisms of speciation. *Rev Cell Biol Molec Med* **2:** 60–93. doi:10.1002/3527600906.mcb.201600015

Shaw RG. 2019. From the past to the future: Considering the value and limits of evolutionary prediction. *Am Nat* **193:** 1–10. doi:10.1086/700565

Smith TB, Kinnison MT, Strauss SY, Fuller TL, Carroll SP. 2014. Prescriptive evolution to conserve and manage biodiversity. *Annu Rev Ecol Evol Syst* **45:** 1–22. doi:10.1146/annurev-ecolsys-120213-091747

Stern DL, Orgogozo V. 2009. Is genetic evolution predictable? *Science* **323:** 746–751. doi:10.1126/science.1158997

Wortel MT, Agashe D, Bailey SF, Bank C, Bisschop K, Blankers T, Cairns J, Colizzi ES, Cusseddu D, Desai MM, et al. 2023. Towards evolutionary predictions: Current promises and challenges. *Evol Appl* **16:** 3–21. doi:10.1111/eva.13513