It's Free Insight! Arbitrage of Theoretical Frameworks and Analytical Methods from Biology to Evolutionary Computation

Matthew Andres Moreno^{1,2,3,4,*}, Emily Dolson^{5,6,7}, and Luis Zaman^{1,2,4}

¹Department of Ecology and Evolutionary Biology ³Michigan Institute for Data and AI in Society

²Center for the Study of Complex Systems

⁴University of Michigan, Ann Arbor, USA ⁵Department of Computer Science and Engineering⁶Program in Ecology, Evolution, and Behavior ⁷Michigan State University, East Lansing, USA *corresponding author: morenoma@umich.edu

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

1 Introduction

1

The metaphor of biological evolution so deeply pervades 2 evolutionary computation (EC) that difficulty can arise in 3 4 discerning the threshold where bio-inspiration ends and algorithm engineering begins. Although some work delves well outside 5 the biological metaphor (Hansen & Ostermeier, 2001; Munteanu 6 & Lazarescu, 1999; Miller & Turner, 2015), core aspects of 7 EC prevail within a bio-inspired framing - whether originally 8 conceived as such, or subsequently co-opted to an analog in nature. 9 Table 1 highlights prominent examples of this phenomenon. 10

Indeed, given the prevalence of bio-inspiration within evolution-11 12 ary computing (EC), the merit of bio-inspiration has constituted 13 a longstanding point of discussion. In addition to imparting useful 14 new ideas, some highlight value from biological metaphors simply as a source for convenient, intuitive, self-consistent vocabulary 15 (Sörensen, 2015; Banzhaf et al., 2006a). Another draw of 16 17 bio-inspiration is the charisma it lends evolutionary computation (Lehman et al., 2020). The complexity and emergent intelligence 18 19 found throughout the natural world provides a stunning display of evolution's profound creative power, and the prospect of capturing 20 any fraction of that power and creativity is fascinating. Indeed, the 21 demonstrated viability of natural precedent lends bio-inspiration 22 credence as a blueprint for ambitious research agendas in 23 24 computing (Miikkulainen & Forrest, 2021; Banzhaf et al., 2006b). 25 Some have questioned, however, whether charisma of the 26 evolutionary approach might also be serving as a preoccupying distraction (Moore, 2023; Sörensen, 2015). In particular, given 27 the broader tendency for machine learning researchers to adhere 28 29 to their preferred models (Domingos, 2012), there is risk that evolutionary approaches have become something of a pursuit for 30 31 its own sake, rather than a means to an end (Woodward et al., 32 2016; Yampolskiy, 2018).

33 Moreover, to a skeptical reader, pervasive appeals to bioinspiration can come across as window dressing on recycled ideas 34 or as an instance of the naturalistic fallacy (Wortmann & Fischer, 35 36 2020; Sörensen, 2015). Such perceptions contribute to a more 37 fundamental concern that evolutionary computation lacks a cohesive, rigorous theoretical framework with first-principles grounding 38 39 (Worzel & Riolo, 2003). (Concerns in this vein, in part, also persist from more general and longstanding contention between "scruffy" 40 and "neat" philosophies (Jones, 2008; Minsky, 1991, p. 16).) The 41 ad hoc nature of evolutionary computation has also been impli-42

cated in hindering adoption, on account of numerous configurable and tunable elements confronting new users (O'Neill et al., 2010).

One possible response to these concerns is to favor deemphasizing the evolutionary metaphor, with the goal of deepening and diversifying first principles footing of evolutionary computation (Moore, 2023). While we agree that evolutionary computation needs sounder theory, we propose the opposite --- that this challenge can best be addressed by leaning *deeper* into the evolutionary metaphor. Historically, evolutionary computation researchers have primarily drawn on bio-inspiration to improve the performance of runtime algorithms (Banzhaf et al., 2006a; Kumar & Bentley, 2003; McPhee et al., 2009). This approach has been effective, but we argue that it can be taken a step further by looking to the science arising around biological systems, rather than just the biological systems themselves, for inspiration — to gain methods and theory that characterize and explain EC algorithms, not just improve their performance. Because evolutionary computation operates on the same principles as biological evolution, substantial amounts of literature on biological theory can be brought to bear on evolution in silico (Belew et al., 1996). As we will review, substantial progress has already been made in this vein — however, much untapped value remains to be "arbitraged" from biological literature. Subsequently, we will evaluate additional yet-to-be-explored opportunities that might further springboard understanding of evolutionary algorithms off of work in evolutionary biology. Discussion, in particular, highlights connections between understanding gained from arbitrage of theory and analysis and - of ultimate importance application-oriented objectives of evolutionary computation.

At the most basic level, evolutionary biology and allied fields of population genetics, developmental biology, and ecology complement the focus of evolutionary computation in seeking to explain, rather than optimize (though not always (Cobb et al., 2013; Carroll et al., 2014)). These life sciences fields have seen radical advances in recent years due to technological leaps in data acquisition and analysis capability (Mathé et al., 2018; Deshpande et al., 2024). Also notable is increasing traction gained by the subfield of experimental evolution within evolutionary biology. Experimental evolution approaches, which observe evolution under controlled conditions, allow detailed inquiry leveraging detailed data collection, replay capabilities, and systematic experimental manipulations (Kawecki et al., 2012).

One concern in transposing knowledge from biology to

Table 1: Conceptual analogies between evolutionary computation (EC) mechanisms and biological processes.

EC Mechanism	Biological Analogy	References
Diversity maintenance	Negative frequency-dependent selection (ecology)	(Dolson & Ofria, 2018; Dolson et al., 2018)
Reciprocal selection	Co-evolution	(Lehman & Stanley, 2010; Harper, 2012; Garbus et al., 2024; Koza et al., 1991; Kala, 2012; Wang et al., 2019; Miikkulainen et al., 2024)
Inexact referencing (e.g. tags)	Active-site recognition in biomolecular interactions	(Spector et al., 2011; Moreno et al., 2023b; Lalejini & Ofria, 2018; Downing, 2015)
Search-space transforms	Genotype-phenotype maps	(Lehman et al., 2023; Moreno et al., 2018; Bentley et al., 2022; Gaier et al., 2020; Wittenberg et al., 2023) TODO

evolutionary computation is the extent to which specificity of 123 85 86 theory and methods to biological life might make them poorly 124 87 applicable to evolutionary computation. While this is a valid 125 88 concern, owing to existing needs in biology to stretch abstractions 126 89 across vast and diverse domains of biological life, we have found 127 90 there to be a good amount of work that is useful to evolutionary 128 91 computation. Indeed, computational artificial life approaches, a 129 92 close cousin of evolutionary computation, are a popular technique 130 93 for exploring the generalizability of biological concepts beyond 131 94 life-as-we-know-it (Cleland, 2013; Langton, 1989; Pennock, 2007). 132 95 Insofar as content applicable to evolutionary computation exists 133 96 within biological literature, however, applying that knowledge 134 97 requires sufficient domain knowledge (1) to identify it and (2) 135 to navigate subtleties in aligning appropriate correspondences 136 98 with EC. Given the vastness, context-dependence, and — at times 13799 polyonymy of biological theory, interdisciplinary collaborations 138 100can be highly productive (Banzhaf et al., 2020). 101 139 With investment of effort, engaging theory from evolutionary 140 102 biology can yield substantial value. Indeed, a good amount of 141 103 existing research has already taken such an approach (Table 2). 142 104 105 Our review highlights examples across three themes, 143 • genotype-phenotype maps and fitness landscapes (Section 2); 144 106 145 107 ecology assembly and coexistence theory (Section 3); and 146 108 phylogeny for prediction/analysis (Section 4). 147 We also review practical work geared at overlapping evolutionary 109 148 computation with bioinformatics infrastructure, 110 149 · sampling-based approaches, which are friendly for decentralized 111 150 infrastructure (Section 5); and 112 151 113 • interoperation with bioinformatics infrastructure (Section 6). 152 114 Additional opportunities for theory arbitrage remain entirely 153 unexplored, however — which we also outline (Section 7). 115 154 116 Despite its promise, borrowing from biology should not be 155 taken as a silver bullet, as it is subject to fundamental limitations. 117 156 Namely, substantial blind spots remain in the explanatory and 118 119 predictive power of analyses and theory in ecology and evolution 157 (Houlahan et al., 2016; Catford et al., 2022; Yates et al., 2018) 120 158 although this is best considered as the exception, rather than 121 122 the rule (Lynch, 2025). Not unlike evolutionary computation, 159

deficiency in theory has habituated perennial consternation in evolutionary biology (Welch, 2017). Among other factors, blame includes overabundance of identifiable causal factors and the existence of exceptions or complicating factors to nearly every generalization. Failures to recognize and build on existing work have also been broadly highlighted (Lynch, 2025; Beer, 2024).

These limitations notwithstanding, biological equivalencies seem likely for the preponderance of what explanatory power is possible within EC. So long as evolutionary algorithms comprise populations of discrete individuals with heritable traits, they will, in some literal sense, instantiate evolutionary processes (Pennock, 2007). Therefore, any systematic theory generalizable across evolutionary computation would likely also hold substantial explanatory power for aspects of biology and vice versa. Notable exceptions, though, arise in unique capabilities within digital evolution to entirely strip out mechanistic elements (e.g., environmental heterogeneity, indirect genotype-phenotype map, etc.) and to achieve perfect, direct observability. As such, theory established in-house for EC could plausibly lead biology in certain areas, a point we return to in our concluding remarks. Although EC will doubtlessly continue in contributing new ideas and perspectives on evolution, reciprocal exchange should be expected — as it seems unlikely for EC, as a smaller field, to profoundly outpace biologists in achieving truly general or cohesive theory.

With expanded perspective, the evolution metaphor in EC has value to offer not only in devising new approaches but also in understanding how and why they work. Looking past biology to harness the science that has arisen around it should continue to be prioritized in advancing the rigor and transparency of evolutionary computation. By highlighting notable existing steps in this direction, we hope to catalyze continued progress toward these goals. Calling attention to several yet-untapped correspondences with theory in ecology and evolutionary biology, we hope, also contributes towards this end.

2 Fitness

Landscapes and Genotype-Phenotype Maps

key citations (TODO):

	Theme	Торіс			Reference
Genotype-Phenotype Maps	Quantifying Deception: A Cas Evolution of Antimicrobial R	se Stuc Resistar	dy in the nce	(Eppstein & Ogbunugafor, 2016)	
	TBD			TBD	
	Ecology	Reachability Analysis for Lex via Community Assembly Gr	icase S raphs	Selection	(Dolson & Lalejini, 2024)
		Ecological theory provides evolutionary computation	insight	ts about	(Dolson & Ofria, 2018)
	Phylogeny Analysis Sampling and Tracking Partial Observability	What can phylogenetic metric useful diversity in evolutionar	cs tell u ry algo	us about prithms?	(Hernandez et al., 2022)
		Untangling phylogenetic dive evolutionary computation us diagnostic fitness landscapes	ersity's sing a	s role in suite of	(Shahbandegan et al., 2022)
		Interactions between learning	g and e	volution	(Ackley & Littman, 1991)
		Methods for Rich Phylogen Over Distributed Sexual Popu	netic Ir ulation	nference Is	(Moreno, 2024)
		A Guide to Tracking Phyloger and Distributed Agent-bas Models	nies in sed E ⁻	Parallel volution	(Moreno et al., 2024c)
	Interoperation	Data Standards for Artificial	Life So	oftware	(Lalejini et al., 2019)
Infrastructure	alifedata-phyloinformatics-co	onvert		(Moreno & Papa, 2024)	
Qua Ant The	antifying Deception: A Case Study in the Evolution of timicrobial Resistance (Eppstein & Ogbunugafor, 2016)			3.1 Ec 3.2 Ec	ology Assembly Theory ology Coexistence Theory
Inc	ps can steer populations to 1	local optima (Schaper & Louis,	181		4 Phylogeny Analysis
maŗ	2014) Evolution in the light of fitness landscape theory (Fragata et al., 2019)		182		
maţ 201 Evo 201	(4)(4)(4)(4)(4)(5)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)(4)<l< td=""><td>landscape theory (Fragata et al.,</td><td>182 183 184</td><td>key refereEcology Signatu</td><td>ences: /, Spatial Structure, and Selection Pressure Induce St res in Phylogenetic Structure (Moreno et al., 2024</td></l<>	landscape theory (Fragata et al.,	182 183 184	key refereEcology Signatu	ences: /, Spatial Structure, and Selection Pressure Induce St res in Phylogenetic Structure (Moreno et al., 2024
maj 201 Evo 201 Gen maj	 (4) (4) (4) (5) (5) (6) (7) (7)	landscape theory (Fragata et al., ion of the genotype-phenotype	182 183 184 185 186	key refereEcology SignatuWhat c in evolu	ences: 7, Spatial Structure, and Selection Pressure Induce St res in Phylogenetic Structure (Moreno et al., 2024 an phylogenetic metrics tell us about useful dive ationary algorithms? (Hernandez et al., 2022)
maj 201 Evo 201 Gen maj TO	 (4) (4) (4) (5) (6) (7) (7)	landscape theory (Fragata et al., tion of the genotype-phenotype	182 183 184 185 186 187 188	 key refere Ecology Signatu What c in evolu Untang computi 	ences: 7, Spatial Structure, and Selection Pressure Induce St res in Phylogenetic Structure (Moreno et al., 2024 an phylogenetic metrics tell us about useful dive itionary algorithms? (Hernandez et al., 2022) ling phylogenetic diversity's role in evolutio ation using a suite of diagnostic fitness landsc
maı 201 201 Ger maı TO	 (4) (4) (4) (5) (7) (7)	landscape theory (Fragata et al., tion of the genotype-phenotype	182 183 184 185 186 187 188 189 190	 key refere Ecology Signatu What c in evolu Untang comput (Shahba) looking 	ences: y, Spatial Structure, and Selection Pressure Induce St irres in Phylogenetic Structure (Moreno et al., 2024 an phylogenetic metrics tell us about useful dive utionary algorithms? (Hernandez et al., 2022) ling phylogenetic diversity's role in evolutio ation using a suite of diagnostic fitness landsc andegan et al., 2022) at genetic change for signatures of adaptation (altho
тај 201 Еvo 201 Ger тај ТО 1	 (4) (4) (4) (5) (9) (9) (9) (9) (9) (10) (1	landscape theory (Fragata et al., tion of the genotype-phenotype blogy	182 183 184 185 186 187 188 189 190 191	 key refere Ecology Signatu What c in evolu Untang comput (Shahba) looking these an Interview 	ences: y, Spatial Structure, and Selection Pressure Induce St tres in Phylogenetic Structure (Moreno et al., 202- an phylogenetic metrics tell us about useful dive tionary algorithms? (Hernandez et al., 2022) ling phylogenetic diversity's role in evolution ation using a suite of diagnostic fitness landsc andegan et al., 2022) at genetic change for signatures of adaptation (althere re more alife-y) vactions between learning and evolution (Ackle
maı 201 Evo 201 Ger maı TO 1 y re	 (4) (4) (4) (4) (5) (6) (7) (7)	landscape theory (Fragata et al., tion of the genotype-phenotype ology	182 183 184 185 186 187 188 189 190 191 192 193	 key refere Ecology Signatu What c in evolu Untang comput (Shahba looking these at - Inte Litt 	ences: A, Spatial Structure, and Selection Pressure Induce Structures in Phylogenetic Structure (Moreno et al., 2022) an phylogenetic metrics tell us about useful diventionary algorithms? (Hernandez et al., 2022) ling phylogenetic diversity's role in evolution ation using a suite of diagnostic fitness landsc andegan et al., 2022) at genetic change for signatures of adaptation (althore re more alife-y) practions between learning and evolution (Ackleman, 1991)
maı 201 Evc 201 Ger maı TO 1 y re Rea Ass	 (4) (4) (4) (4) (5) (6) (7) (7)	landscape theory (Fragata et al., tion of the genotype-phenotype ology case Selection via Community Lalejini, 2024)	182 183 184 185 186 187 188 189 190 191 192 193 194 195	 key refere Ecology Signatu What c in evolu Untang comput (Shahb) looking these an – Inte Litt – MC TODO 	ences: A, Spatial Structure, and Selection Pressure Induce St tres in Phylogenetic Structure (Moreno et al., 2024 an phylogenetic metrics tell us about useful dive ationary algorithms? (Hernandez et al., 2022) ling phylogenetic diversity's role in evolutio ation using a suite of diagnostic fitness landsc andegan et al., 2022) at genetic change for signatures of adaptation (althere more alife-y) eractions between learning and evolution (Ackleman, 1991) DDES (Dolson et al., 2019) lead-in goes here
maq 201 Evo 201 Ger maq TO 1 y re Rea Ass Eco con	 (4) (4) (4) (4) (5) (6) (7) (7)	landscape theory (Fragata et al., tion of the genotype-phenotype ology case Selection via Community Lalejini, 2024) insights about evolutionary 2018)	182 183 184 185 186 187 188 189 190 191 192 193 194 195 196	 key refere Ecology Signatu What c in evolu Untang comput (Shahba) looking these at Litt MC TODO 4.1 Ex 	ences: A, Spatial Structure, and Selection Pressure Induce St tres in Phylogenetic Structure (Moreno et al., 202- an phylogenetic metrics tell us about useful dive tionary algorithms? (Hernandez et al., 2022) ling phylogenetic diversity's role in evolutio ation using a suite of diagnostic fitness landsc andegan et al., 2022) at genetic change for signatures of adaptation (alther re more alife-y) practions between learning and evolution (Ackler man, 1991) DDES (Dolson et al., 2019) lead-in goes here ample Subsection
maı 201 Evcc 201 Ger maı TO 1 y re Rea Asss Ecco com	 (4) (4) (4) (4) (4) (5) (6) (7) (7)	landscape theory (Fragata et al., tion of the genotype-phenotype ology case Selection via Community Lalejini, 2024) insights about evolutionary 2018)	182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197	 key refere Ecology Signatu What c in evolu Untang comput (Shahba looking these at – Inte Litt – MC TODO 4.1 Ex 5 Sam 	ences: A, Spatial Structure, and Selection Pressure Induce St tres in Phylogenetic Structure (Moreno et al., 202- an phylogenetic metrics tell us about useful dive tionary algorithms? (Hernandez et al., 2022) ling phylogenetic diversity's role in evolutio ation using a suite of diagnostic fitness landsc andegan et al., 2022) at genetic change for signatures of adaptation (althere re more alife-y) eractions between learning and evolution (Ackles man, 1991) DDES (Dolson et al., 2019) lead-in goes here ample Subsection npling and Tracking Partial Observabi

Table 2: Representative examples of evolutionary computation work that leverages theory from biology.

199 • hereditary stratigraphy (Moreno et al., 2022a, 2024c)

• sampling-based knockout analyses (Moreno, 2024)

201TODO mention that best-effort approaches are of interest 258202in scaling digital evolution systems themselves (e.g., ackely, 259203darkcycle, etc.)260

204 (Moreno et al., 2022b, 2024c)

205 It has repeatedly been the case that significant steps forward in 262206 biological knowledge piggyback on innovations making new types 263 207 or quantities of data available, technological or otherwise. For 264 208 instance, evolutionary theory arose in the context of burgeoning 265 209 taxonomic collections (Winsor, 2009), core ideas in pathology 266 210 and developmental biology arose from microscopy (Turner, 1890), 267 211 high-throughput sequencing has made symbiosis a key concept 268 212 in organismal biology (e.g., gut microbiome) (Durack & Lynch, 2019), and new imaging technologies are driving new questions 269 213 about ecological interactions across continental-scale distance 270 214 (Stark et al., 2016). Although factors besides feasibility have also 271 215 catalyzed significant advances in life science (e.g., Mendel, Redi, 272 216 Semmelweis), growth in capabilities to generate and collect data 273 217 218 play a longstanding and ongoing role in enabling new biological 274 inquiry (Strasser, 2012). Contemporary biology enjoys profound, 275 219 ongoing gains in data availability (Sulston et al., 1983; Sheth et al., 276 220 221 2017; Weeks et al., 2023), nevertheless fundamental limitations 277 222 exist in data completeness, particularly with respect to historical ac- 278 223 counts of natural history (Benton et al., 2011; Delsuc et al., 2005). 279 In contrast to biology, digital evolution has, from the outset, 280 224 225 enjoyed perfect fidelity and absolute completeness in data collec- 281 tion. Indeed, such observability is a key driver of scientific interest 282 226 in using in silico models for evolution research (O'Neill, 2003). 283 227 Although some domains of genetic programming have been 284 228 highlighted for their capability to produce intuitive symbolic 285 229 expressions (Hu, 2023; Javed et al., 2022), it is also the case that 286 230 discerning the functionality of some evolved artifacts can require 287 231 extensive experiment-driven analyses. Such knockout trials - 288 232 233 in contrast to other aspects of digital evolution — are notable in 289 234 that combinatoric tractability has held back completely exhaustive 290 235 analysis for all but the smallest genomes (Nitash & Adami, 2021). 291 236 Knockout assay experiments have, therefore, typically limited 292 to single-site (Adami, 2006), all-pairs (Kumawat et al., 2023), or 293 237 iterative approaches (Langdon et al., 2014; Moreno et al., 2021). 294 238 239 Massively parallel and distributed processing power, which 295 240 has been argued crucial to future directions in digital evolution 296 (Moreno & Ofria, 2022; Taylor et al., 2016), peels away the 297 241 tractability of digital evolution's existing perfect observability 298 242 paradigm. One concern is that many-process experiments can 299 243 produce greater volumes of data than is feasible to store, much 300 244 less analyze (Klasky et al., 2021). For instance, even under serial 301 245 246 processing, maintaining full records of genetic program instruction 302 history under sexual recombination has proven to be a highly 247 technically demanding, data-intensive task (McPhee et al., 2016). 303 248 249 Parallel and distributed computing also introduces challenges 304 250 in runtime overhead from communication and synchronization 305 251 required for data collection and introduces the possibility of data 306 loss when components fail (Snir et al., 2014). Continuing with the 307 252 phylogenetic example, typical tracking approaches are sensitive 308 253 to even small amounts of data loss and, in a distributed computing 309 254 context, require runtime inter-process communication to reclaim 310 255

256 memory from extinct lineages (Moreno et al., 2024a).

257

261

Fortunately for digital evolution, research in biology, by necessity, already routinely works around issues of incomplete and imperfect data. As such, existing methods can provide a valuable foothold in scenarios where combinatorial effects or runtime multiprocessing make exhaustive direct observation impractical. This section reviews work leveraging methods borrowed from biology to negotiate data limitations on both fronts: 1) application of mark-recapture approaches from ecology to characterize fitness landscapes and 2) application of reconstruction-based approaches inspired by bioinformatics for robust, decentralized phylogeny tracing.

5.1 Mark-Recapture Estimation

Mark-recapture analysis (or capture-recapture analysis) is a widely-used and well-developed method to estimate sizes of biological populations (Amstrup et al., 2010). This method uses the proportion of individuals shared between two or more samples as a proxy to estimate the total population size that is being sampled. For large population sizes, relatively lower recapture rates are expected. It turns out, though, that idealized sampling from an urn poorly describes animal behavior. Various potential biases have been identified — ranging from the inherent disposition of certain animals to be more "trap happy" or "trap shy" to the tendency of already-captured animals to become more wary of traps — and sophisticated statistical methods have been devised to make estimation robust to them. Mark-recapture literature, therefore, provides a rich, ready-made buffet for tacking estimation problems involving repeat partial sampling.

In one application, Moreno (2024) demonstrate use of a markrecapture estimator in quantifying sites contributing to fitness that are not individually detectable due to epistatic redundancy and or small-effect contributions. Analogy to the mark-recapture scenario is established by equating sites with any potential for fitness effect — whether or not detectable through single-site knockout — to the population to be estimated. Iterative knockouts are applied to produce several "skeleton" genotypes, where no more sites can be removed without reducing fitness. Sites in each skeleton, therefore, each have a demonstrable fitness effect - but, if redundancy or small effects are at play, no skeleton contains all such sites. Each skeleton, therefore, represents a sample of sites with potential fitness effects. Crucially, though, these samples will overrepresent lower-redundancy or larger-effect sites. Application of a jackknife estimator due to Burnham & Overton (1979), however, ensures estimation accuracy remains intact. In a separate line of work, Schulte et al. (2014) have noted potential for mark-recapture methods to play a role in characterizing the extent of neutral space within multistep mutational neighborhoods of computer programs.

5.2 Reconstruction-based Phylogenetic Analysis

In addition to play-by-play accounts of extinctions, innovations, and other key events in an evolutionary run, phylogenetic analysis can provide insight into the nuts and bolts of evolutionary computation through more general characterization of the underlying mode and tempo of evolution (Moreno et al., 2023a; Hernandez et al., 2022; Shahbandegan et al., 2022; ?). Availability of an exactly accurate phylogenetic record is useful, but in most cases not strictly necessary, in accomplishing these objectives 368
(Moreno et al., 2024d). Indeed, typical biological approaches to 369
phylogenetic analysis involve inexact inference-based estimation, 370
yet such phylogenetic analysis has contributed immensely to our
understanding of biological evolution.
At the most fundamental level, modern bioinformatics accom-

plishes phylogenetic analysis by comparing traces of similarity 373 317 retained in DNA genomes under the influence of mutational 374 318 accumulation. Notably, such mutational processes occur in a com- 375 319 320 pletely decentralized manner, and reconstruction can be performed 376 321 among any number of organisms — including small subsamples of 377 322 the overall population. Disadvantageously, though, complications 378 323 arise in these analyses owing to issues of back mutation, mutational 379 saturation, selection effects, long branch attraction, and the vast 380 324 quantities of genetic sequence information required (?) In contrast 381 325 326 to biological model organisms, however, evolutionary computation 382 327 affords the capability to arbitrarily engineer genome structure - 383 328 and, therefore, affords the possibility to sidestep such challenges. 384

Hereditary stratigraphy methodology arose from such a desire 385 329 for a means to extract phylogenetic information from distributed 386 330 simulations that is efficient, robust, straightforward, and gener- 387 331 alizable across digital evolution systems. The method works by 388 332 333 bundling agent genomes with special annotations in a manner akin 389 334 to non-coding DNA (entirely neutral with respect to agent traits 390 and fitness). These annotations apply an approximate checkpoint- 391 335 336 ing mechanism to maximize reconstruction quality from a minimal 392 337 memory footprint — configurable as low as 96 bits per genome 393 (Moreno et al., 2022a). A major benefit of this approach is that it al- 394 338 339 lows the relatedness of any two organisms to be compared directly 395 without depending on global information, which opens the door 396 340 to incorporation of EC techniques that incorporate phylogenetic 397 341 information at runtime to guide evolution toward desired outcomes 398 342 (Lalejini et al., 2024a,b; Murphy & Ryan, 2008; Burke et al., 2003). 399 343 In one application, borrowing from bioinformatics has allowed 400 344 hereditary stratigraphy-enabled implementation to address chal- 401 345 lenges of scale, memory capacity, and communication bandwidth 402 346 in opening a window into digital evolution on next-generation AI 403 347 348 accelerator hardware. Moreno et al. (2024e) demonstrates tracking 404 of an island-model genetic algorithm across the 850,000 core 405 349 Cerebras Wafer-Scale Engine. Under a simple one-max equivalent 406 350 351 test regime, the strong decentralization afforded by hereditary 407 352 stratigraphy enables upwards of a quadrillion replication events to 408 353 be simulated in an hour. Moreno et al. (2024e) showed effects in 409 354 phylogenetic structure between alternate mutation operators, and 410 355 other work has demonstrated recovery of information salient to 411 356 understanding selection pressure, spatial structure, and ecological 412 dynamics (Moreno et al., 2024d). 357 413

358 For those looking to incorporate this methodology into 414 their own work, a public-facing software library ("hstrat") has 415 359 been provided to facilitate plug-and-play addition of tracking 416 360 annotations (Moreno et al., 2022b). Moreno et al. (2024c) 417 361 362 provides a step-by-step guide to configuring and using the 418 methodology. Although the core methodology ascribes an asexual 419 363 model, extensions to sexual phylogenies have been explored 420 364 (Moreno, 2024). Beyond phylogenetic tracking, underlying 421 365 algorithms developed for hereditary stratigraphy provide means 422 366 to very efficiently maintain running temporal cross-samples ("data 423 367

stream curation") (Moreno et al., 2024f), which holds potential for more general utility in reducing runtime communication and storage by support for on-demand, after-the-fact data extraction.

5.3 Inferential Observability (excerpt from EXPRESS grant)

To address these problems, we propose a paradigm shift in ABMS/PDES data collection: inferential observability. This model takes inspiration from approaches used in real-world experiments, which successfully draw scientific inferences based on a smaller and noisier sets of data than what are typically collected using ABMS/PDES. Indeed, past a certain point, precision in data from ABMS/PDES becomes of essentially negligible value, owing to arbitrary effects of stochasticity and fundamental limitations in correspondence between model and reality. Any computational resources invested in producing this excessive level of precision could be better used elsewhere. Trading a controlled amount of data precision for increased scalability and hardware accelerator compatibility would be highly worthwhile.

Historically, most research using ABMS/PDES has assumed complete observability of model state. Indeed, the ability to measure properties in silico that would be impossible to observe in vitro or in vivo is a major benefit of ABMS/PDES for scientific inquiry. However, as the scale of these models increases, the cost of data collection becomes a serious obstacle. Thus, experiments exist that are intractable in the real world because they rely on data that are physically impossible to measure, but are also intractable in digital models because storing the necessary data to answer the questions at hand is infeasible. We propose that, through careful algorithm development, we can solve this problem by recording a smaller amount of data that enables us to draw the desired scientific inferences at a fraction of the computational cost. Through this work, we will unlock the ability to perform scientific inquiry that would have been previously intractable across digital and real-world systems.

Data tracking often requires cross-referencing multiple simulation elements, which can introduce runtime communication costs under parallel and distributed computation. For example, in evolutionary models, phylogenetic relatedness (i.e., line of descent) information enables powerful analyses, but perfectly tracking this data necessitates difficult-to-scale bookkeeping to purge extinct lineages (Moreno et al., 2024b) (we solved this problem using a combination of inferential observability and space-time memory; see HStrat in Prior Research). Qualitatively similar problems occur more broadly in contexts where patterns of interaction among simulation elements must be tracked (e.g., tracking chains of pathogen transmission). Beyond slowdown from synchronization inefficiencies, a fundamental obstacle is also posed when data storage needs exceed available space.

Inferential observability aims to collect the minimal amount of data necessary to answer the scientific questions at hand. In part, we propose to achieve this goal by designing algorithms that efficiently down-sample time series data. More fundamentally, however, we suggest that modelers may be better served by exporting data only under certain circumstances and propose algorithms to support this workflow. Such an approach is particularly valuable for work with hardware accelerators such as the WSE, which have

limited input/output (I/O) bandwidth. We anticipate that users may 476 424 425 want to export data in response to certain "trigger" conditions being 477 fulfilled, through sampling processes, or — for policy-driven mod- 478 426 els — in response to real-time queries during scenario exercises. 479 427 428 Simulation is useful insofar as it is interpretable. As simulation 480 429 scales, so does the challenge of managing an exhaustive data 481 record. Recent hardware trends only exacerbate matters, growing 482 430 processing power while reducing the amount of RAM and disk 483 431 432 storage available per core — especially in accelerator-driven 484 HPC architectures (Khan et al., 2021; Gholami et al., 2024). For 485 433 434 such architectures, host-device bandwidth and latency strongly 486 435 impact performance (Kwon & Rhu, 2018). Such concerns arise 487 436 especially in work with the Cerebras WSE where only a small 488 437 fraction of peripherally-located PEs interface to the host. Our 489 438 proposed inferential observability paradigm will help mitigate 490 439 these problems. In this aim, we propose to develop, formalize, 491 440 and experimentally evaluate this approach. Additionally, we will 492 441 publish software implementations to make inferential observability 493 accessible to the ABMS/PDES community. 494 442

443 6 Interoperation 444 with Bioinformatics Infrastructure

445 key references:

- Niema's megaphylogeny software (Moshiri, 2025, 2020) and 500 taxonium (Sanderson, 2022)
- alife data standards (Lalejini et al., 2019)
- alifedata phyloinformatics convert (Moreno & Papa, 2024)
- 450 hstrat auxlib? (Moreno et al., 2022b)
- AEVOL (Daudey et al., 2024) and applications of agent-based 505
 modeling to evaluating bioinformatics? 506

453 6.1 ALIFE Data Standard

454 7 Opportunities for Future Arbitrage

- 455 future work:
- 456 trait-based phylogenetic analysis
- 457 spatial analyses (necessary due to parallel and distributed 512 computing)
 513
- detecting speciation through population genetics methods 514
 (Sukumaran et al., 2021)
- 400 (Sutumatian et al., 2021)
 461 Inferring Fitness landscapes and selection on phenotypic states 516
- 462 from single-cell genealogical data (Nozoe et al., 2017)
- 463Table 3 lists selected directions from evolutionary biology that518464could enrich EC theory.519
- 465 TODO lead-in goes here

466 7.1 Excerpt 467 from mmore500

from mmore500/hstrat-reconstruction-algo

468 In a parallel vein, the volume of data processed in bioinformatics 524 469 workflows is also increasing with continuing advances in high- 525 470 throughput sequencing technologies, enabling the construction 526 of phylogenies containing millions of taxa. As an illustrative 527 471 472 example at the cutting edge of extreme scale, Konno et al. (2022) 528 473 reports phylogeny synthesis from 235 million sequence reads 529 generated from an in silico CRISPR barcoding model - requiring 530 474 475 31 hours of compute time across 300 HPC nodes. 531

In both the context of bioinformatics and artificial life research, very large-scale phylogeny data enabled by advances in sequencing and computing technology represent a new challenge as much as an opportunity, raising the question of how best to mine this data. On a practical level, work is needed not just to push the boundaries of what can be learned from phylogenies, but also how to store, load, traverse, quantify, visualize, and manipulate very large phylogenies in an efficient manner. Indeed, projects are being developed to try to address this issue. For example, taxonium (Sanderson, 2022) is a web-based software for visualizing large phylogenies in a flexible, interactive manner, and is able to handle browsing millions of tips at a high frame rate. Other projects aim to create methods for compact, scalable phylogeny representations (Moshiri, 2025, 2020), enabling faster and more memory-efficient tree operations.

In pushing the boundaries of phylogenetic scale to billion-tip datasets, ALife research has the opportunity to contribute to an interdisciplinary ecosystem of software tools developing around working with very large-scale phylogenies. In particular, the ALife data standard, which specifies a tabular representation for phylogeny data (Lalejini et al., 2019), has strong potential to develop a backbone of a larger high-performance phylogeny processing infrastructure. Although originally envisioned as a data storage format, the tabular nature of the standard allows integrations with high-performance software tools arisen around the "Data Frame" concept, including Pandas, Polars, Dask, and data.table. These libraries provide a structured, user-friendly interface to advanced performance features such as multithreading, data streaming, query optimization, file partitioning, and column-oriented binary datafile formats (McKinney, 2010; Barrett et al., 2025; Vink et al., 2024; Rocklin, 2015). Additionally, for Python users, the columnar array format typical in data frame libraries is compatible with NumPy and Numba, readily enabling on-the-fly SIMD vectorization and Just-In-Time compilation (Harris et al., 2020; Lam et al., 2015). Indeed, this approach underlies much of the pre- and post-processing steps for end-to-end reconstruction demonstrated in this work.

495

496

497

498

499

502

503

504

507

508

509

510

511

517

520

521

522

523

8 Conclusion

The sophistication of natural organisms — still, in many cases, thoroughly unrivaled by artificial engineering — strongly evidences the creative potential of evolutionary computation, and also lends it unique charisma. The evolutionary metaphor in genetic algorithms and genetic programming has proven highly productive, both in guiding research advances and also, more simply, in capturing research attention. Despite reasonable criticisms of metaphor-driven algorithm development (Moore, 2023; Sörensen, 2015), we argue compelling potential exists to strengthen the foundations of evolutionary computation by instead leaning further into the evolution metaphor — by engaging methods and theory constructs from evolutionary biology and allied fields.

In this review, we have highlighted notable examples where reaching into the biological literature has improved visibility into application-oriented evolutionary computation. Such research, much of it very recent, reflects only first steps. Important extensions remain to be fleshed out and connected to *bona fide* real-world use cases. Further, reviewed work touches only a small fraction of promising directions for arbitrage from biology to evolutionary computation. In Section 7, we have highlighted Table 3: Selected evolutionary biology literature that may inform future theory arbitrage in evolutionary computation.

Theme	Торіс	Reference
Genotype-Phenotype Maps	Evolution in the light of fitness landscape theory	(Fragata et al., 2019)
	The arrival of the frequent: how bias in genotype-phenotype maps can steer populations to local optima	(Schaper & Louis, 2014)
Fcology	assembly theory	TBD
Leology	coexistence theory	TBD
Phylogeny Analysis	TBD	TBD
Sampling-based	TBD	TBD
and Partial Observability	TBD	TBD
Interoperation	Tools for exploring massive phylogenies	(Sanderson, 2022; Moshiri, 2025, 2020)
Infrastructure	TBD	TBD

563

532 several possibilities.

Productive exchange between biology and evolutionary com- 564 533 putation is a two-way street. Already, digital organisms find use 565 534 535 cases in experiments conducted in conjunction with in vivo inquiry 566 (Sanjuán et al., 2007; Wilke et al., 2001; Hindré et al., 2012). How- 567 536 ever, we believe that application-oriented genetic programming 568 537 538 and genetic algorithms have special value to offer in the realm of 569 539 genotype-phenotype maps and evolvability. Owing to its artificial 570 540 nature, genotype-phenotype maps - much less, maps with strong 571 evolvability - do not come baked into evolutionary computation 572 541 542 a priori (Kirschner & Gerhart, 1998). As such, EC research has 543 invested significant effort in teasing apart how the properties of a 573 genotype-phenotype map influence outcomes from adaptive evolu-544 574 tion (Banzhaf, 1994; Hu & Banzhaf, 2010; Whigham et al., 2017). 545 A particularly promising line of EC work has sprung up in inves-546 tigating how to harness unsupervised learning (e.g., autoencoders, 547 576 548 LLMs) to generate evolvable genotype-phenotype maps (Lehman 577 549 et al., 2023; Moreno et al., 2018; Bentley et al., 2022; Gaier et al., 550 2020; Wittenberg et al., 2023). Corresponding work in biology, 578 however, is only fledgling. Indeed, such work considering evolv-551 579 ability as an unsupervised learning process has notably already 552 580 553 been driven forward through collaboration with evolutionary com-581 554 putation practitioners (Kouvaris et al., 2017; Szilágyi et al., 2020). 555 The longstanding bidirectional exchange between evolutionary 582 556 computation and evolutionary biology is truly remarkable, particu- 583 557 larly in contrast with other areas of AI/ML where such efforts have 584 been sparser and more recent (Marblestone et al., 2016; Richards 558 et al., 2019). We look forward to seeing this exchange deepen and, 585 559 in particular, fulfill concrete objectives in better explaining evo- 586 560 lutionary computation, diagnosing failure cases, and prescribing 587 561 appropriate methods for challenging domain problems. 588 562

Acknowledgement

This research was supported by Michigan State University through the computational resources provided by the Institute for Cyber-Enabled Research and is based upon work supported by the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Futures program. This research was supported in part by funding from the NSF (DEB 1813069). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

- Ackley, D. & Littman, M. (1991). Interactions between learning and evolution. *Artificial life II*, 10, 487–509.
- Adami, C. (2006). Digital genetics: unravelling the genetic basis of evolution. *Nature Reviews Genetics*, 7(2), 109–118.
- Altenberg, L. (1995). Genome growth and the evolution of the genotype-phenotype map, 205–259. Springer Berlin Heidelberg. https://doi.org/10.1007/ 3-540-59046-3_11
- Amstrup, S. C., McDonald, T. L., & Manly, B. F. (2010). *Handbook of capture-recapture analysis*. Princeton University Press.
- Banzhaf, W. (1994). Genotype-phenotype-mapping and neutral variation — A case study in Genetic Programming, 322–332. Springer Berlin Heidelberg. https: //doi.org/10.1007/3-540-58484-6_276

- Banzhaf, W., Beslon, G., Christensen, S., Foster, J. A., Képès, 638 589
- 590 F., Lefort, V., Miller, J. F., Radman, M., & Ramsden, J. J. 639
- (2006a). From artificial evolution to computational evolution: 640 591

592 a research agenda. Nature Reviews Genetics, 7(9), 729-735. 641

- Banzhaf, W., Beslon, G., Christensen, S., Foster, J. A., Képès, 642 593
- 594 F., Lefort, V., Miller, J. F., Radman, M., & Ramsden, 643
- 595 J. J. (2006b). From artificial evolution to computational 644
- evolution: a research agenda. Nature Reviews Genetics, 7(9), 596 645
- 729-735. https://doi.org/10.1038/nrg1921 597 646
- Banzhaf, W., Cheng, B. H. C., Deb, K., Holekamp, K. E., 647 598 Lenski, R. E., Ofria, C., Pennock, R. T., Punch, W. F., 648 599
- & Whittaker, D. J., editors (2020). Evolution in Action: 649 600
- Past, Present and Future: A Festschrift in Honor of Erik 601 650
- D. Goodman. Springer International Publishing. https: 651 602
- //doi.org/10.1007/978-3-030-39831-6 603 652
- Barrett, T., Dowle, M., Srinivasan, A., Gorecki, J., Chirico, 653 604 M., Hocking, T., Schwendinger, B., & Krylov, I. 654 605 Extension of 'data.frame'. 655 (2025). data.table: 606 https://r-datatable.com. R package ver-607 1.17.99, https://Rdatatable.gitlab.io/data.table, 656 608 sion https://github.com/Rdatatable/data.table 657 609
- 658 Beer, R. D. (2024). (a)life as it could be. Artificial Life, 30(4), 659 610 611 539-545. https://doi.org/10.1162/artl a 660 612 00455 661
- 662 Belew, R. K., Mitchell, M., & Ackley, D. H. (1996). Computation 613

and the natural sciences. Adaptive Individuals in Evolving 663 614 Populations: Models and Algorithms, 26, 431–440. 615 664

- 665 Bentley, P. J., Lim, S. L., Gaier, A., & Tran, L. (2022). Evolving 666 616 Through the Looking Glass: Learning Improved Search 617 Spaces with Variational Autoencoders, 371–384. Springer 667 618
- 619 International Publishing. https://doi.org/10.668 669 1007/978-3-031-14714-2 26 620
- Benton, M. J., Dunhill, A. M., Lloyd, G. T., & Marx, F. G. 670 621 671 (2011). Assessing the quality of the fossil record: insights 622 from vertebrates. Geological Society, London, Special 672 623 673 624 Publications, 358(1), 63-94. 674
- Burke, E., Gustafson, S., Kendall, G., & Krasnogor, N. 625 675 (2003). Is increased diversity in genetic programming 626 676 beneficial? an analysis of lineage selection. The 627 677 2003 Congress on Evolutionary Computation, 2003. 628 678 629 CEC '03., volume 2, 1398-1405 Vol.2. https: //doi.org/10.1109/cec.2003.1299834 630 679
- 680 Burnham, K. P. & Overton, W. S. (1979). Robust estimation 631 of population size when capture probabilities vary among 681 632 633 animals. Ecology, 60(5), 927-936. 682
- 683 Carroll, S. P., Jørgensen, P. S., Kinnison, M. T., Bergstrom, C. T., 634
- Denison, R. F., Gluckman, P., Smith, T. B., Strauss, S. Y., 684 635
- & Tabashnik, B. E. (2014). Applying evolutionary biology 685 636
- to address global challenges. science, 346(6207), 1245993. 686 637

- Catford, J. A., Wilson, J. R., Pyšek, P., Hulme, P. E., & Duncan, R. P. (2022). Addressing context dependence in ecology. Trends in Ecology & Evolution, 37(2), 158–170. https: //doi.org/10.1016/j.tree.2021.09.007
- Cleland, C. E. (2013). Is a general theory of life possible? seeking the nature of life in the context of a single example. Biological Theory, 7, 368–379.
- Cobb, R. E., Chao, R., & Zhao, H. (2013). Directed evolution: past, present, and future. AIChE Journal, 59(5), 1432-1440.
- Daudey, H., Parsons, D. P., Tannier, E., Daubin, V., Boussau, B., Liard, V., Gallé, R., Rouzaud-Cornabas, J., & Beslon, G. (2024). Aevol_4b: Bridging the gap between artificial life and bioinformatics. The 2024 Conference on Artificial Life, Alife 2024. https://doi.org/10.1162/isal a 00716
- Delsuc, F., Brinkmann, H., & Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. Nature Reviews Genetics, 6(5), 361-375.
- Deshpande, D., Chhugani, K., Ramesh, T., Pellegrini, M., Shifman, S., Abedalthagafi, M. S., Alqahtani, S., Ye, J., Liu, X. S., Leek, J. T., Brazma, A., Ophoff, R. A., Rao, G., Butte, A. J., Moore, J. H., Katritch, V., & Mangul, S. (2024). The evolution of computational research in a data-centric world. Cell, 187(17), 4449-4457. https: //doi.org/10.1016/j.cell.2024.07.045
- Dolson, E., Banzhaf, W., & Ofria, C. (2018). Applying Ecological Principles to Genetic Programming, 73-88. Springer International Publishing. https: //doi.org/10.1007/978-3-319-90512-9 5
- Dolson, E. & Lalejini, A. (2024). Reachability analysis for lexicase selection via community assembly graphs. Genetic Programming Theory and Practice XX, 283-301. Springer.
- Dolson, E. & Ofria, C. (2018). Ecological theory provides insights about evolutionary computation. Proceedings of the Genetic and Evolutionary Computation Conference Companion, Gecco '18, 105-106. https: //doi.org/10.1145/3205651.3205780
- Dolson, E. L., Vostinar, A. E., Wiser, M. J., & Ofria, C. (2019). The modes toolbox: Measurements of open-ended dynamics in evolving systems. Artificial Life, 25(1), 50-73. https://doi.org/10.1162/artl a 00280
- Domingos, P. (2012). A few useful things to know about machine learning. Communications of the ACM, 55(10), 78-87.
- Downing, K. (2015). Intelligence Emerging: Adaptivity and Search in Evolving Neural Systems. EBL-Schweitzer. MIT Press.
- Durack, J. & Lynch, S. V. (2019). The gut microbiome: Relationships with disease and opportunities for therapy. Journal of experimental medicine, 216(1), 20-40.

- Eppstein, M. J. & Ogbunugafor, C. B. (2016). Quantifying decep-739 687
- tion: A case study in the evolution of antimicrobial resistance. 740 688
- Proceedings of the Genetic and Evolutionary Computa-741 689 tion Conference 2016, Gecco '16, 101-108. https: 742 690
- //doi.org/10.1145/2908812.2908821 691
- Fragata, I., Blanckaert, A., Dias Louro, M. A., Liber-744 692 les, D. A., & Bank, C. (2019). Evolution in 745 693 the light of fitness landscape theory. Trends in 694 https:746 695 *Ecology & Evolution*, 34(1), 69–82.
- 747 //doi.org/10.1016/j.tree.2018.10.009 696
- 748 Gaier, A., Asteroth, A., & Mouret, J.-B. (2020). Dis- 749 697 covering representations for black-box optimization. 698 Proceedings of the 2020 Genetic and Evolutionary 750 699 700 Computation Conference, Gecco '20, 103–111. https: 751 //doi.org/10.1145/3377930.3390221 752 701
- Garbus, J., Willkens, T., Lalejini, A., & Pollack, J. (2024). 753 702 703 Accelerating co-evolutionary learning through phylogeny- 754 704 informed interaction estimation. Proceedings of the 755 705 Genetic and Evolutionary Computation Conference Companion, GECCO '24 Companion, 427-430. https: 756 706 757
- //doi.org/10.1145/3638530.3654250 707
- 708 Gholami, A., Yao, Z., Kim, S., Hooper, C., Ma-759 honey, M. W., & Keutzer, K. (2024). Ai and 709 https:760 710 memory wall. IEEE Micro, 1-5.761 711 //doi.org/10.1109/mm.2024.3373763
- 712 Hansen, N. & Ostermeier, A. (2001). Completely de- 762 713 randomized self-adaptation in evolution strategies. 763 Evolutionary Computation, 9(2), 159–195. https:764 714 //doi.org/10.1162/106365601750190398 765 715
- Spatial co-evolution: quicker, fit-767 Harper, R. (2012). 716 ter and less bloated. Proceedings of the 14th 717 718 annual conference on Genetic and evolution- 768 ary computation, Gecco '12, 759-766. 719 https: 769 //doi.org/10.1145/2330163.2330269 720
- 770 Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, 771 721 722 R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, 772 J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, 773 723 724 S., van Kerkwijk, M. H., Brett, M., Haldane, A., del 774
- Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, 725
- P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., 775 726
- Gohlke, C., & Oliphant, T. E. (2020). Array programming 776 727
- with numpy. Nature, 585(7825), 357-362. https: 777 728 778
- //doi.org/10.1038/s41586-020-2649-2 729
- 779 Hernandez, J. G., Lalejini, A., & Dolson, E. (2022). What 780 730 can phylogenetic metrics tell us about useful diversity in 731 evolutionary algorithms? Genetic programming theory and 781 732 782 practice XVIII, 63-82. Springer. 733
- 783 Hindré, T., Knibbe, C., Beslon, G., & Schneider, D. 784 734 New insights into bacterial adaptation 785 735 (2012).through in vivo and in silico experimental evolu-736 tion. Nature Reviews Microbiology, 10(5), 352-365.786 737
- https://doi.org/10.1038/nrmicro2750 738 787

- Houlahan, J. E., McKinney, S. T., Anderson, T. M., & McGill, B. J. (2016). The priority of prediction in ecological understanding. Oikos, 126(1), 1–7. https://doi.org/10.1111/oik.03726
- Hu, T. (2023). Genetic programming for interpretable and explainable machine learning. Genetic Programming Theory and Practice XIX, 81-90. Springer.

743

758

766

- Hu, T. & Banzhaf, W. (2010). Evolvability and speed of evolutionary algorithms in light of recent developments in biology. Journal of Artificial Evolution and Applications, 2010, 1–28. https://doi.org/10.1155/2010/568375
- Javed, N., Gobet, F., & Lane, P. (2022). Simplification of genetic programs: a literature survey. Data Mining and Knowledge Discovery, 36(4), 1279-1300.
- Jones, M. T. (2008). Artificial Intelligence: A Systems Approach: A Systems Approach. Jones & Bartlett Learning.
- Kala, R. (2012). Multi-robot path planning using coevolutionary genetic programming. Expert Systems with Applications, 39(3), 3817-3831. https: //doi.org/10.1016/j.eswa.2011.09.090
- Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., & Whitlock, M. C. (2012). Experimental evolution. Trends in ecology & evolution, 27(10), 547-560.
- Khan, A., Sim, H., Vazhkudai, S. S., Butt, A. R., & Kim, Y. (2021). An analysis of system balance and architectural trends based on top500 supercomputers. The International Conference on High Performance Computing in Asia-Pacific Region, HPC Asia 2021. https: //doi.org/10.1145/3432261.3432263
- Kirschner, M. & Gerhart, J. (1998). Evolvability. Proceedings of the National Academy of Sciences, 95(15), 8420-8427.
- Klasky, S., Thayer, J., & Najm, H. (2021). Data reduction for science: Brochure from the advanced scientific computing research workshop. Technical report, Office of Scientific and Technical Information (OSTI). https://doi.org/10.2172/1770192
- Konno, N., Kijima, Y., Watano, K., Ishiguro, S., Ono, K., Tanaka, M., Mori, H., Masuyama, N., Pratt, D., Ideker, T., Iwasaki, W., & Yachie, N. (2022). Deep distributed computing to reconstruct extremely large lineage trees. Nature Biotechnology, 40(4), 566-575. https: //doi.org/10.1038/s41587-021-01111-2
- Kouvaris, K., Clune, J., Kounios, L., Brede, M., & Watson, R. A. (2017). How evolution learns to generalise: Using the principles of learning theory to understand the evolution of developmental organisation. PLoS computational biology, 13(4), e1005358.
- Koza, J. R. et al. (1991). Genetic evolution and co-evolution of computer programs. Artificial life II, 10, 603-629.

- Kumar, S. & Bentley, P. J. (2003). Biologically inspired 837
 evolutionary development. *International Conference on* 838 *Evolvable Systems*, 57–68.
- ⁸⁴⁰ 791 Kumawat, B., Lalejini, A., Acosta, M., & Zaman, L. (2023). Fluc-
- tuating environments promote evolvability by shaping adap- 841
 tive variation accessible to populations. *bioRxiv*, 2023–01. 842
 843
- 794 Kwon, Y. & Rhu, M. (2018). Beyond the memory wall: 844
- 795 A case for memory-centric hpc system for deep learn- 845
- 796 ing. 2018 51st Annual IEEE/ACM International 797 Symposium on Microarchitecture (MICRO) https://www.sec.
- 797 Symposium on Microarchitecture (MICRO). https: ⁸⁴⁰
- 798 //doi.org/10.1109/micro.2018.00021 847 848
- Lalejini, A., Dolson, E., Bohm, C., Ferguson, A. J., Parsons, ⁸⁴⁹
 D. P., Rainford, P. F., Richmond, P., & Ofria, C. (2019).
 Data standards for artificial life software. *The 2019* 851 *Conference on Artificial Life*, Alife 2019, 507–514.
 852
- 803 https://doi.org/10.1162/isal_a_00213
- Lalejini, A., Moreno, M. A., Hernandez, J. G., & Dolson, 854
 E. (2024a). Phylogeny-informed fitness estimation 855
 for test-based parent selection. *Genetic Programming* 856 *Theory and Practice XX*, 241–261. Springer Interna-857
 tional Publishing. https://doi.org/10.1007/
- 809 978-981-99-8413-8 13
- Lalejini, A. & Ofria, C. (2018). Evolving event-driven programs 860
 with signalgp. *Proceedings of the Genetic and Evolutionary* 861 *Computation Conference*, Gecco '18, 1135–1142. https:
- 813 //doi.org/10.1145/3205455.3205523
- 863 814 Lalejini, A., Sanson, M., Garbus, J., Moreno, M. A., & Dolson, 864
- 815 E. (2024b). Runtime phylogenetic analysis enables 865
- 816 extreme subsampling for test-based problems. https: (/doi org/10_48550/orViv_2402_01610_866
- 817 //doi.org/10.48550/arXiv.2402.01610
- Lam, S. K., Pitrou, A., & Seibert, S. (2015). Numba: 868
 a llvm-based python jit compiler. Proceedings 869
 of the Second Workshop on the LLVM Compiler Infrastructure in HPC, Sc15, 1–6. https://doi.org/10.1145/2833157.2833162
- 822 77 401.019710.114072000107.2000102 872 823 Langdon W.B. Modat M. Petke, I. & Harman M. (2014) 873
- 823 Langdon, W. B., Modat, M., Petke, J., & Harman, M. (2014).
- Improving 3d medical image registration cuda software with 874
 genetic programming. *Proceedings of the 2014 Annual Con*-875
- *ference on Genetic and Evolutionary Computation*, 951–958. 876
- 877 Leaster C (1080) Actificial Life Values L Decending Of As
- Langton, C. (1989). Artificial Life, Volume I: Proceedings Of An
 Interdisciplinary Workshop On Synthesis And Simulation Of
- *Living Systems*. Advanced book program: Addison-Wesley. 879
 Avalon Publishing. 880
- Lehman, J., Clune, J., Misevic, D., Adami, C., Altenberg, L., 882
 Beaulieu, J., Bentley, P. J., Bernard, S., Beslon, G., Bryson, 883
- B33 D. M., et al. (2020). The surprising creativity of digital
- 834 evolution: A collection of anecdotes from the evolutionary 884
- 835computation and artificial life research communities.836Artificial life, 26(2), 274–306.886

- Lehman, J., Gordon, J., Jain, S., Ndousse, K., Yeh, C., & Stanley, K. O. (2023). Evolution Through Large Models, 331–366. Springer Nature Singapore. https://doi.org/10. 1007/978-981-99-3814-8_11
- Lehman, J. & Stanley, K. O. (2010). Efficiently evolving programs through the search for novelty. *Proceedings* of the 12th Annual Conference on Genetic and Evolutionary Computation, Gecco '10, 837–844. https: //doi.org/10.1145/1830483.1830638
- Lynch, M. (2025). Complexity myths and the misappropriation of evolutionary theory. *Proceedings of the National Academy of Sciences*, 122(23). https: //doi.org/10.1073/pnas.2425772122
- Marblestone, A. H., Wayne, G., & Kording, K. P. (2016). Toward an integration of deep learning and neuroscience. *Frontiers in computational neuroscience*, 10, 94.
- Mathé, E., Hays, J., Stover, D., & Chen, J. (2018). The omics revolution continues: The maturation of highthroughput biological data sources. *Yearbook of Medical Informatics*, 27(01), 211–222. https: //doi.org/10.1055/s-0038-1667085
- McKinney, W. (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*, SciPy, 56–61. https: //doi.org/10.25080/majora-92bf1922-00a

858 859

862

867

- McPhee, N. F., Crane, E., Lahr, S. E., & Poli, R. (2009). Developmental plasticity in linear genetic programming. *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*, 1019–1026.
- McPhee, N. F., Donatucci, D., & Helmuth, T. (2016). Using graph databases to explore the dynamics of genetic programming runs. *Genetic programming theory and practice XIII*, 185–201.
- Miikkulainen, R. & Forrest, S. (2021). A biological perspective on evolutionary computation. *Nature Machine Intelligence*, 3(1), 9–15. https: //doi.org/10.1038/s42256-020-00278-8
- Miikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., Raju, B., Shahrzad, H., Navruzyan, A., Duffy, N., & Hodjat, B. (2024). *Evolving deep neural networks*, 269–287. Elsevier. https://doi.org/10.1016/ b978-0-323-96104-2.00002-6
- Miller, J. & Turner, A. (2015). Cartesian genetic programming. Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation, Gecco '15, 179–198. https: //doi.org/10.1145/2739482.2756571
- Minsky, M. L. (1991). Logical versus analogical or symbolic versus connectionist or neat versus scruffy. *AI magazine*, 12(2), 34–34.

- 887 Moore, J. H. (2023). Is the evolution metaphor still necessary 935 888 or even useful for genetic programming? *Genetic* 936
- 889 *Programming and Evolvable Machines*, 24(2), 21.
- 938 890 Moreno, Dolson, & Rodriguez-Papa (2023a). Toward phyloge- 939
- 891 netic inference of evolutionary dynamics at scale. *Alife 2023*,
- 892 volume Proceedings of the 2023 Artificial Life Conference 940
- 893 of Proceedings of the 2023 Artificial Life Conference, 79. 941
- 894 https://doi.org/10.1162/isal_a_00694 942 943
- 895 Moreno, M. A. (2024). Methods for rich phylogenetic
- 896 inference over distributed sexual populations. Ge-944
- 897 netic Programming Theory and Practice XX, 125–945
- 898 141. Springer International Publishing. https:946
- 899 //doi.org/10.1007/978-981-99-8413-8_7 947
- Moreno, M. A., Banzhaf, W., & Ofria, C. (2018). Learning 948
 an evolvable genotype-phenotype mapping. *Pro-* 949 *ceedings of the Genetic and Evolutionary Compu-* 950 *tation Conference*, Gecco '18, 983–990. https:
 //doi.org/10.1145/3205455.3205597
- // 001.019/10.1145/5205455.5205557
- Moreno, M. A., Dolson, E., & Ofria, C. (2022a). Hereditary stratigraphy: Genome annotations to enable
 phylogenetic inference over distributed populations. *The 2022 Conference on Artificial Life*, 418–428.
 https://doi.org/10.1162/isal_a_00550
 957
- 910 Moreno, M. A., Dolson, E., & Ofria, C. (2022b).
 958
 911 hstrat: a python package for phylogenetic infer-959
 912 ence on distributed digital evolution populations.
 960
 913 Journal of Open Source Software, 7(80), 4866.
 961
- 914 https://doi.org/10.21105/joss.04866 962
- Moreno, M. A., Lalejini, A., & Ofria, C. (2023b). Matchmaker, 963
 matchmaker, make me a match: geometric, variational, and 964
 evolutionary implications of criteria for tag affinity. *Genetic* 965 *Programming and Evolvable Machines*, 24(1), 4. https: 966
- 919 //doi.org/10.1007/s10710-023-09448-0
- 920 Moreno, M. A. & Ofria, C. (2022). Best-effort communication 968 921 improves performance and scales robustly on conven-969
- 922
 tional hardware.
 arXiv, (arXiv:2211.10897).
 https:970

 923
 //doi.org/10.48550/arxiv.2211.10897
- 924
 Moreno, M. A. & Papa, S. R. (2024). mmore500/alifedata-972

 925
 phyloinformatics-convert. https://doi.org/10.973

 926
 5281/zenodo.10701178
- 927Moreno, M. A., Papa, S. R., & Dolson, E. (2024a). Algorithms975928for efficient, compact online data stream curation. arXiv.929https://doi.org/10.48550/arxiv.2403.977
- **930** 00266 **978**
- 979 931 Moreno, M. A., Papa, S. R., & Dolson, E. (2024b). 980
- Analysis of phylogeny tracking algorithms for se-981
 rial and multiprocess applications. *arXiv*. https:
- 934 //doi.org/10.48550/arxiv.2403.00246

Moreno, M. A., Papa, S. R., & Ofria, C. (2021). Case study of novelty, complexity, and adaptation in a multicellular system. OEE4: The Fourth Workshop on Open-Ended Evolution. https://doi.org/10.48550/arXiv. 2405.07241. Accessed: 2023-12-04

937

952

- Moreno, M. A., Ranjan, A., Dolson, E., & Zaman, L. (2024c). A guide to tracking phylogenies in parallel and distributed agent-based evolution models. https: //doi.org/10.48550/arXiv.2405.10183
- Moreno, M. A., Rodriguez-Papa, S., & Dolson, E. (2024d). Ecology, spatial structure, and selection pressure induce strong signatures in phylogenetic structure. *arXiv preprint arXiv:2405.07245*.
- Moreno, M. A., Yang, C., Dolson, E., & Zaman, L. (2024e). Trackable agent-based evolution models at wafer scale. *arXiv*, 87–98. https: //doi.org/10.48550/arxiv.2404.10861
- Moreno, M. A., Zaman, L., & Dolson, E. (2024f). Structured downsampling for fast, memoryefficient curation of online data streams. https: //doi.org/10.48550/arXiv.2409.06199
- Moshiri, N. (2020). Treeswift: A massively scalable python tree package. *SoftwareX*, 11, 100436. https://doi.org/10.1016/j.softx.2020.100436
- Moshiri, N. (2025). Compacttree: a lightweight header-only c++ library and python wrapper for ultra-large phylogenetics. *Gigabyte*, 2025. https://doi.org/10.46471/gigabyte.152
- Munteanu, C. & Lazarescu, V. (1999). Improving Mutation Capabilities in a Real-Coded Genetic Algorithm, 138–149. Springer Berlin Heidelberg. https://doi.org/10.1007/10704703_11
- Murphy, G. & Ryan, C. (2008). A simple powerful constraint for genetic programming. *Genetic Programming*, 146–157. https://doi.org/10.1007/978–3–540–78671–9_13
- Nitash, C. & Adami, C. (2021). Information-theoretic characterization of the complete genotype-phenotype map of a complex pre-biotic world: Comment on "from genotypes to organisms: State-of-the-art and perspectives of a cornerstone in evolutionary dynamics" by susanna manrubia et al. *Physics of Life Reviews*, 38, 111–114.
- Nozoe, T., Kussell, E., & Wakamoto, Y. (2017). Inferring fitness landscapes and selection on phenotypic states from single-cell genealogical data. *PLOS Genetics*, 13(3), e1006653. https: //doi.org/10.1371/journal.pgen.1006653
- 982 O'Neill, B. (2003). Digital evolution. *PLoS Biology*, 1(1), e18.

- O'Neill, M., Vanneschi, L., Gustafson, S., & Banzhaf, W. (2010).1031 983 984 Open issues in genetic programming. Genetic Programming 1032 985 and Evolvable Machines, 11, 339–363. 1033 Pennock, R. T. (2007). Models, simulations, instantiations, and evi-1034 986 dence: the case of digital evolution. Journal of Experimental 1035 987 988 & Theoretical Artificial Intelligence, 19(1), 29–42. https://org 989 //doi.org/10.1080/09528130601116113 1037 1038 Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, 1039 990 991 R., Christensen, A., Clopath, C., Costa, R. P., de Berker, A., 992 Ganguli, S., et al. (2019). A deep learning framework for 1040 neuroscience. Nature neuroscience, 22(11), 1761–1770. 1041 993 Rocklin, M. (2015). Dask: Parallel computation with blocked 994 algorithms and task scheduling. Proceedings of the 14th¹⁰⁴³ 995 Python in Science Conference, SciPy, 126–132. https: 996 //doi.org/10.25080/majora-7b98e3ed-013 997 Taxonium, a web-based tool¹⁰⁴⁶ Sanderson, T. (2022). 998 for exploring large phylogenetic trees. 999 eLife, 11. 047 https://doi.org/10.7554/elife.82392 1000 1048 Sanjuán, R., Cuevas, J. M., Furió, V., Holmes, E. C., &1049 1001 Moya, A. (2007). Selection for robustness in mutage-1050 1002 nized ma viruses. PLoS Genetics, 3(6), e93. https: 1003 //doi.org/10.1371/journal.pgen.0030093¹⁰⁵¹ 1004 1052 Schaper, S. & Louis, A. A. (2014). The arrival of the frequent:1053 1005 How bias in genotype-phenotype maps can steer populations1054 1006 1007 to local optima. PLoS ONE, 9(2), e86635. https: //doi.org/10.1371/journal.pone.00866351055 1008 1056 1009 Schulte, E., Fry, Z. P., Fast, E., Weimer, W., & Forrest, S. (2014).1057 Software mutational robustness. Genetic Programming and 1058 1010 Evolvable Machines, 15, 281–312. 1011 1059 1060 Shahbandegan, S., Hernandez, J. G., Lalejini, A., & Dolson, E₁₀₆₁ 1012 (2022). Untangling phylogenetic diversity's role in evolution-1062 1013 ary computation using a suite of diagnostic fitness landscapes, 1063 1014 1015 Proceedings of the Genetic and Evolutionary Computation Conference Companion, Gecco '22, 2322-2325. https: 1064 1016 //doi.org/10.1145/3520304.3534028 1017 1065 Sheth, R. U., Yim, S. S., Wu, F. L., & Wang, H. H. (2017)₁₀₆₆ 1018 Multiplex recording of cellular events over time on crispr $_{1067}^{1067}$ 1019 1020 biological tape. Science, 358(6369), 1457–1461. 1068 Snir, M., Wisniewski, R. W., Abraham, J. A., Adve, S. V., Bagchi, 1021 1070 S., Balaji, P., Belak, J., Bose, P., Cappello, F., Carlson, B., 1022 et al. (2014). Addressing failures in exascale computing. 1023 *The International Journal of High Performance Computing Applications* 28(2) 129–173 1024 Applications, 28(2), 129–173. 1025 1074 Sörensen, K. (2015). Metaheuristics-the metaphor exposed. 1026 International Transactions in Operational Research, 22(1),1075 1027 1076 1028 3-18.1077
- Spector, L., Martin, B., Harrington, K., & Helmuth, T. (2011).1078
 Tag-based modules in genetic programming. *Proceedings*1079

of the 13th annual conference on Genetic and evolutionary computation, Gecco '11, 1419–1426. https: //doi.org/10.1145/2001576.2001767

- Stark, S. C., Breshears, D. D., Garcia, E. S., Law, D. J., Minor, D. M., Saleska, S. R., Swann, A. L., Villegas, J. C., Aragão, L. E., Bella, E. M., et al. (2016). Toward accounting for ecoclimate teleconnections: intra-and inter-continental consequences of altered energy balance after vegetation change. *Landscape Ecology*, 31, 181–194.
- Strasser, B. J. (2012). Data-driven sciences: From wonder cabinets to electronic databases. *Studies in History and Philosophy* of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43(1), 85–87.
- Sukumaran, J., Holder, M. T., & Knowles, L. L. (2021). Incorporating the speciation process into species delimitation. *PLoS Computational Biology*, 17(5), e1008924.
- Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic cell lineage of the nematode caenorhabditis elegans. *Developmental biology*, 100(1), 64–119.
- Szilágyi, A., Szabó, P., Santos, M., & Szathmáry, E. (2020). Phenotypes to remember: Evolutionary developmental memory capacity and robustness. *PLoS computational biology*, 16(11), e1008425.
- Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., McMullin, B., Packard, N., Rasmussen, S., Virgo, N., Agmon, E., Clark, E., McGregor, S., Ofria, C., Ropella, G., Spector, L., Stanley, K. O., Stanton, A., Timperley, C., Vostinar, A., & Wiser, M. (2016). Open-ended evolution: Perspectives from the oee workshop in york. *Artificial Life*, 22(3), 408–423. https://doi.org/10.1162/artl_a_00210
- Turner, W. (1890). The cell theory, past and present. *Journal* of anatomy and physiology, 24(Pt 2), 253.
- Vink, R., de Gooijer, S., Beedie, A., Gorelli, M. E., nameexhaustion, Peters, O., Guo, W., Burghoorn, G., van Zundert, J., Hulselmans, G., Grinstead, C., Marshall, chielP, Turner-Trauring, I., Mitchell, L., Santamaria, M., Heres, D., Magarick, J., Genockey, K., ibENPC, deanm0000, Harbeck, H., Wilksch, M., eitsupi, Koutsouris, I., Leitao, J., van Gelderen, M., Brannigan, L., & Barbagiannis, P. (2024). *pola-rs/polars: Python polars 1.16.0.* https: //doi.org/10.5281/zenodo.14244124
- Wang, R., Lehman, J., Clune, J., & Stanley, K. O. (2019). Poet: open-ended coevolution of environments and their optimized solutions. *Proceedings of the Genetic and Evolutionary Computation Conference*, Gecco '19, 142–151. https: //doi.org/10.1145/3321707.3321799

- 1080 Weeks, B. C., Zhou, Z., O'Brien, B. K., Darling, R., Dean, M., 1130
- 1081 Dias, T., Hassena, G., Zhang, M., & Fouhey, D. F. (2023).1131
- 1082
 A deep neural network for high-throughput measurement
- 1083 of functional traits on museum skeletal specimens. *Methods*
- *in Ecology and Evolution*, 14(2), 347–359.
- Welch, J. J. (2017). What's wrong with evolutionary biology? *Biology & philosophy*, 32(2), 263–279.
- 1087Whigham, P. A., Dick, G., & Maclaurin, J. (2017). On1088the mapping of genotype to phenotype in evolu-1089tionary algorithms. Genetic Programming and1090Evolvable Machines, 18(3), 353–361. https:1091//doi.org/10.1007/s10710-017-9288-x
- Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., & Adami, C.
 (2001). Evolution of digital organisms at high mutation rates
 leads to survival of the flattest. *Nature*, 412(6844), 331–333.
 https://doi.org/10.1038/35085569
- Winsor, M. P. (2009). Taxonomy was the foundation of darwin's evolution. *Taxon*, 58(1), 43–49.
- Wittenberg, D., Rothlauf, F., & Gagné, C. (2023). Denoising
 autoencoder genetic programming: strategies to control exploration and exploitation in search. *Genetic Programming and Evolvable Machines*, 24(2), 17.
- Woodward, J. R., Johnson, C. G., & Brownlee, A. E. (2016). Gp
 vs gi: if you can't beat them, join them. *Proceedings of the*2016 on Genetic and Evolutionary Computation Conference
 Companion, 1155–1156.
- Wortmann, T. & Fischer, T. (2020). Does architectural design optimization require multiple objectives? a critical analysis. 25th *International Conference on Computer-Aided Architectural Design Research in Asia, CAADRIA 2020*, 365–374.
- Worzel, B. & Riolo, R. (2003). Genetic programming: Theory and practice. *Genetic Programming Theory and Practice*, (Chapter 1, 1–10). Kluwer.
- Yampolskiy, R. V. (2018). Why we do not evolve software? analysis of evolutionary algorithms. *Evolutionary Bioinformatics*,
 14, 1176934318815906.
- Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, 1116 C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, 1117 S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, 1118 C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, 1119 1120 E. J., Halpin, P. N., Heikkinen, R. K., Heinänen, S., Jones, A. R., Krishnakumar, P. K., Lauria, V., Lozano-Montes, H., 1121 Mannocci, L., Mellin, C., Mesgaran, M. B., Moreno-Amat, 1122 E., Mormede, S., Novaczek, E., Oppel, S., Ortuño Crespo, 1123 1124 G., Peterson, A. T., Rapacciuolo, G., Roberts, J. J., 1125 Ross, R. E., Scales, K. L., Schoeman, D., Snelgrove, P., Sundblad, G., Thuiller, W., Torres, L. G., Verbruggen, 1126 H., Wang, L., Wenger, S., Whittingham, M. J., Zharikov, 1127 Y., Zurell, D., & Sequeira, A. M. (2018). Outstanding 1128 challenges in the transferability of ecological models. 1129

Trends in Ecology & Evolution, 33(10), 790–802. https: //doi.org/10.1016/j.tree.2018.08.001

11329Supplemental Material

1133 (Moreno et al., 2024e)