Systems in Evolutionary Systems Biology

L Loewe, Laboratory of Genetics and Wisconsin Institute for Discovery, University of Wisconsin–Madison, Madison, WI, USA

Glossary

EvoSysBio A transdisciplinary framework for constructing reliable, testable, interactive overviews of nestable, dynamic, multi-dimensional fitness landscapes, which mechanistically predict: (1) changes in fitness of individual organisms when their states and environments change; (2) how populations evolve when organisms traverse fitness landscapes.

Extra-Organism Biology (EOB) Biology focusing on actions, reactions, processes, parts, and patterns in ecosystems that alter states of individual organisms as they interact with biotic or abiotic environments.

Fitness Causality Network (FCNet) A network of nodes (defined by IFTs) and links (defined by LIFTs) that describes causal influences (e.g., genotypes, environments, and initial states such as maternal methylation patterns of DNA) on consequential IFTs (e.g., survival, reproduction) in a given time period.

Fitness landscape An abstract 'landscape' defined by *causal* 'positions in a plane' and their *consequential* 'heights,' as defined by a corresponding Fitness Causality Network.

Incomplete Fitness Trait (IFT) A phenotypic trait that impacts the fitness of an organism, its offspring, or its genetic relatives and that, at least occasionally, affects rates of survival, reproduction, merging, etc., or modifies evolutionary factors like mutation rates in one or more environments.

Intra-Organism Biology (IOB) Biology focusing on actions, reactions, processes, parts, and patterns within a single, individual organism that enable it to live (grow, survive, reproduce, move, etc.) by changing its state in a given environment.

Landscape of Incomplete Fitness Traits (LIFT) An abstract representation of a fitness landscape that maps causal IFTs to consequential IFTs, thus providing a causality statement about how input governs output, comparable to probabilistic mathematical functions.

Nest Organism A single, individual organism that is the environment for one or more populations of different types of nested organisms, which may themselves be nest organisms.

Nested Organism A single, individual organism that is contained by an encapsulating nest organism, which is the nested organism's environment and may itself be nested in a larger organism of a different type.

Organism A single, individual system that consists of different parts that builds a whole, which – in biology – must be able to replicate and may be nestable.

Population-Genetics Biology (PGB) Biology focusing on heritable information (genotypes, alleles, traits, methylation patterns, etc.) that can directly or indirectly affect IFTs (survival, reproduction, etc.) when passed on by individual organisms in evolving populations.

Trans-Organism Biology (TOB) Biology focusing on integrating Extra-Organism Biology (EOB), Population-Genetics Biology (PGB), and all direct or indirect actions, processes, and patterns that otherwise fall through the disciplinary cracks of EOB or PGB (excluding the inside of organisms).

Systems approaches to biology and genome evolution are becoming increasingly important. Since the 1920s, the New Evolutionary Synthesis has been constructing an increasingly coherent view of evolution by synthesizing ideas from different parts of biology. Further progress of this relentless synthesis will increasingly depend on crossing disciplinary boundaries, complex simulations of biological systems, and reliable reproducibility. Evolutionary Systems Biology (EvoSysBio) is defined here as working towards integrated interactive overviews of dynamic multi-dimensional fitness landscapes that are testable in the real world and enable the prediction of evolution. EvoSysBio further formalizes the New Synthesis as illustrated here for cancer and antibiotics resistance evolution.

Systems Approaches to Genome Evolution

Biology has a long history with numerous independent efforts to integrate diverse aspects of biological systems in order to understand the whole. For example, R.A. Fisher (1918) resolved a big controversy about inheritance by creating an integrative systems model with a mechanism for combining the varying effects by which different genes could impact a given phenotypic trait. The rise of population genetics has since inspired mechanistic modeling of how systems with populations of organisms evolve in response to the five fundamental factors of evolution, which may vary between these biological extremes over space and time:

1. *mutation* (perfect heritability ⇆ fast change);
2. *selection* (harmful ⇆ neutral ⇆ helpful);
3. *genetic drift* (last survivor ⇆ largest finite population);
4. *recombination* (complete linkage ⇆ free segregation); and
5. *migration* (homogeneity of space ⇆ movement maximizing impact of heterogeneity).

These factors are fundamental, since they affect *all* real-world populations of organisms with a phenotype capable of reproducing heritable variations; the complexity of evolutionary

outcomes (see this Encyclopedia) is fueled by variations of these factors over space and time at nested levels of replication. Simple interactions are well understood, but complex, dynamic patterns in heterogeneous, nested, multi-locus systems pose many questions that will keep defining the cutting-edge of population genetics for a long time to come (Loewe and Hill, 2010a,b). There is no better theory of evolution than population genetics, the 'auto-mechanics of evolution' (Singh and Krimbas, 2000, p.1); yet it has much room for growth by integrating expertise from biochemical reaction networks to ecological interactions. For example, biochemistry was instrumental for generally understanding dominance in heterozygous genes (Kacser and Burns, 1981); now more detailed, dynamic models could explain patterns observed in specific genes. Population genetics can be abstract, ignoring many details (Haldane, 1964), yet its core abstractions are powerful enough for mechanistically connecting intra-organism and trans-organism biology and their environments with the phylogenetic patterns they govern together (see **Table 1**). The power of these abstractions places population genetics in a remarkably central role in biology and makes it very difficult to ignore whenever populations of replicating organisms are involved.

This article points to various systems approaches in biology and how they connect to population genetics, starting with Current Systems Biology (CSysBio) and increasingly important aspects of formal and computational modeling. Section 'Defining EvoSysBio' uses abstractions from population genetics to define EvoSysBio without loss of generality. Section 'Fitness Landscapes' discusses **L**andscapes of **I**ncomplete **F**itness **T**raits (LIFTs), which help to define EvoSysBio and represent the best albeit fractured glimpses of true fitness landscapes that will be available for a long time. The last section presents examples of five important milestones for EvoSysBio.

Current Systems Biology

Systems approaches in biology emerged from a confluence of several broad trends:

(i) Researchers have always known that cellular molecules had a complex context, but this complexity came into focus only after enough details accumulated from studying isolated molecules.
(ii) Increases in computing power inspired developers of quantitative and computational methods in biology, ushering in the era of bioinformatics and genomics.
(iii) As genome sequencing matured, some perceived its data-driven discoveries to lack inspiring transformational hypotheses; this fueled a desire for more hypothesis-driven quantitative models and opened biology up to many physicists and engineers excited about modeling.
(iv) Medical interest in understanding how varied diseases work in humans motivated the Human Genome Project ('*read* the blueprint'), but locating candidate disease genes rarely satisfied curiosity ('*understand* the blueprint'). As differences between genomes were revealed, they stimulated interest in personalizing medicine ('understand *my* blueprint' (Hood *et al.*, 2004)). A theoretical framework could guide the long journey from wishing to actually understanding biology.

These trends supported the rise of CSysBio, which can employ data from genomics and other -omics to construct a more integrated view of molecular interactions, ideally by simulating them in dynamic mechanistic models that yield predictions to be tested by perturbations in high-throughput experiments (Ideker *et al.*, 2001b). It was initially contrasted to (1) a deep-but-narrow focus on functions of isolated molecules, and (2) a broad-but-shallow, data-driven genomic search for candidate genes with functional annotations, but without corresponding mechanistic models (Ideker *et al.*, 2001a). CSysBio started in molecular biology (Westerhoff and Palsson, 2004), and since it has been expanding its scope to include cells, physiology, and more, it still lacks a generally agreed-upon definition. Some see it as a cycle that could characterize *any* system (Kitano, 2002a,b; Ideker *et al.*, 2001a):

(i) collect all details possibly relevant for a given system and research question;
(ii) define a quantitative model with all parts and interactions relevant to the question;
(iii) predict from the model properties of observable real-world perturbations;
(iv) test the model by comparing its predictions to relevant real-world data;
(v) refine the model if the distance between predictions and real-world data is too large (as usual), or expand it by including new questions.

This definition reassuringly echoes the Scientific Method; absence of system types studied, methods applied or questions asked, it is also too general for any field, which risks defining CSysBio only in 'the eye of the beholder.' Despite difficulties of definition and adoption of more thorough systems perspectives (Cornish-Bowden, 2006), CSysBio has been producing some impressive mechanistic models of systems ranging from the molecular to the physiological, including: circadian clocks (Dodd *et al.*, 2005; Huang *et al.*, 2012), metabolic fluxes in microbes (Edwards *et al.*, 2001; King *et al.*, 2015), viral replication (Endy *et al.*, 2000; Lim and Yin, 2009), full intracellular dynamics over a simple cell cycle (Karr *et al.*, 2012), human heart (Noble, 2011), rat physiology (Virtual Rat, see Section Relevant Websites), and more, occasionally even venturing beyond Intra-Organism Biology (IOB) into topics such as evolution.

Current Systems Biology Meets Evolution

Most CSysBio works toward empirically supported models of processes that help genotypes and environments shape phenotypes. However, if CSysBio models include replicating entities with traits that are both heritable and mutable, then these need to be treated as organisms in their own right. Heritable variation of such traits inevitably activates the five factors with all their complicated dynamics. If evolutionary processes occur within multicellular organisms, they are easily overlooked even if they are highly relevant for medicine, such as the evolution of tumor cell populations in cancer (Stearns and Koella, 2008). Similarly, fast evolution of bacteria

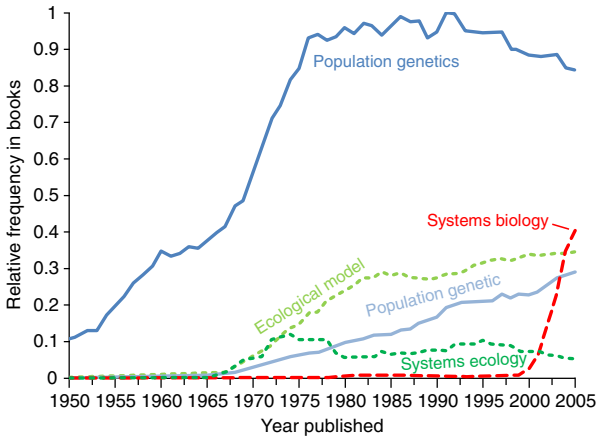

Figure 1 Rise of some terms related to systems approaches in biology as tracked by changes of frequency in books digitized by Google (terms are case insensitive; in relative units of $1 \approx 0.000\,037\%$; all big data cautions apply; Michel *et al.*, 2011). Data credit: https://books.google.com/ngrams.

(Edwards *et al.*, 2001) could be important to understand for industrial production plants.

Understanding population genetics mechanisms helps to recognize evolutionary processes and select the right types of tools and theorems for modeling them. Without some grasp of past population genetics work, researchers risk needlessly re-deriving past results instead of building on them. Trends in Figure 1 might provide rough indicators of this rich history (all caveats about 'big data' apply).

Ecological Systems Biology

CSysBio contrasts with an earlier era of 'systems theory in biology,' which inspired ecologists in the 1970s to build causal models of ecological systems (Wolkenhauer, 2001). As a discipline, ecology has learned important lessons about how to deal with extreme uncertainty and has accumulated statistical expertise in the art of developing knowledge by using empirical tests against real-world data to reduce uncertainty in models of complex systems. CSysBio can learn from Ecology's experience in quantifying uncertainty (Kirk *et al.*, 2015).

Advances in computing and the availability of 'big data' for *some specific* parts of systems provide opportunities for improving models from cells to ecosystems (e.g., high-resolution images inside cells and more (Editorial, 2015; Dietze *et al.*, 2013; Mackechnie *et al.*, 2011; Seksik and Landman, 2015; Sleeman, 2013; Xu and Rhee, 2014; Kuhlbrandt, 2014; Aronson and Rehm, 2015). However, this does not solve all modeling challenges. Integrating big data into mechanistic models can be difficult or infeasible; critical model details sometimes exist only as diffuse and diverse datasets in myriads of formats, and unknown parameters may only be amenable to 'guesstimating.' Such challenges for an integrated understanding can be substantial; appropriate advanced modeling techniques may help, but without special measures it can be surprisingly difficult to reproduce results from increasingly complex computations and statistics (James *et al.*, 2015; Stodden, 2015). These trends matter for complex models of evolution and their corresponding sub-models.

Systems Genetics

The new field of systems genetics builds on genetics, genomics, 'phenomics,' more -omics, CSysBio, and other fields to better integrate insights about how genotypes shape phenotypes in different environments (Hughes, 2010; Markowetz and Boutros, 2015) with huge medical implications (Aronson and Rehm, 2015). It combines into genome-wide association studies (GWAS) as many triplets as possible, each containing an individual genome, respective quantitative phenotypes (molecular and organismal), and a record of environmental conditions that shaped the genotype–phenotype map of these organisms. It aims to obtain integrated models that highlight which genes have probably been affected by selection on organisms with a specified phenotype in a given environment. Advanced statistical techniques used in GWAS provide empirical estimates about how much a given gene might have contributed to a particular trait of interest, thus identifying candidate genes that may cause important phenotypic functions. Systems genetics methods work because patterns of DNA sequence diversity contain information on the strength of the environment's selection of certain phenotypic traits, which affects the survival probabilities of organisms with corresponding mutations. Patterns in modern genomes echo a long history of selection. Examples include studies in the fruit fly *Drosophila melanogaster* (Anholt and Mackay, 2015; Huang *et al.*, 2014; Mackay *et al.*, 2009; Harbison *et al.*, 2009; Ayroles *et al.*, 2009).

Systems genetics provides a different perspective on the complex gene regulatory networks that govern phenotypes, which are also studied in CSysBio; thus in principle, both could combine their strength. For example, many unknowns often exist in detailed, mechanistic models of complex gene regulatory networks that could, in principle, predict phenotypes also amenable to systems genetics. Thus, GWAS may constrain overly large parameter spaces in mechanistic models, while CSysBio simulations may add mechanistic details about genetic architecture by highlighting genes with effects too small to detect in GWAS.

Integration

Better understanding evolution weakens boundaries between diverse systems approaches in biology, as evolution is its biggest system. Molecular systems biology can only study systems shaped by evolution (excludes extinct species), while systems ecology and evolution critically depend on biochemistry; systems genetics relies on all to shape the statistical patterns of genetic diversity in populations, and initiatives in 'personalized precision medicine' aiming to predict clinical consequences of genetic variants (Aronson and Rehm, 2015) will hardly be able to do so without computational models (Iyengar *et al.*, 2015), which will both inform and be informed by a profound understanding of evolution (Stearns and Koella, 2008). Continuing progress on this integrative trajectory increasingly depends on modeling skills and leads to the mechanistic view of EvoSysBio discussed in the next section.

Systems Science: The Whole is Different from the Sum of Its Parts

The realization that a system as a whole can be very different from a heap of its parts is at least as old as Aristotle (-344 ± 22). It also has long guided Gestalt theory, arguably an early form of systems science (Wagemans *et al.*, 2012b) used to study sensory perception (Wagemans *et al.*, 2012a), which undoubtedly affects the evolution of species reliant on learned mental models for navigating their world. Investigating how elements combine into a 'Gestalt' also inspired the thought that – at least in principle – the whole world could be understood with a single mathematical equation, albeit of extravagant composite complexity (von Ehrenfels, 1890, p. 292).

Computational models in systems science essentially provide such 'world-equations,' if only for the small, closed worlds described in the composite formal structures called 'programs.' Modeling investigates the whole by studying relevant parts, their interactions, and systemic properties. Good models usually start with a curiosity, informal questions, and a hunch. They grow as researchers integrate more observations and are eventually formalized as small, computable worlds that mimic relevant interactions of parts and the emergent properties of their whole. The main results of systems science are empirically tested models that become more valuable as they survive increasingly difficult challenges to mechanistically predict non-trivial phenomena. To develop such models, it is essential to actually look at the whole repeatedly and from different perspectives, which seems to be challenging even for systems biologists (Cornish-Bowden, 2006). It helps to know what types of systems others have already seen and which models were developed as a result.

Types of models can be classified in many broad categories using various criteria, such as their approach to randomness in the system, how they are computed, which values they allow, etc.; the myriad modeling approaches needed for understanding evolution could easily fill an encyclopedia the size of this one (and indeed fill many pages in this Encyclopedia too). Examples include approaches that are

- *deterministic*: assumes no chance exists, so recomputing always produces the same result;
- *stochastic*: chance exists, so recomputing always produces variable outcomes as computers use random number generators to choose between possible outcomes;
- *probabilistic*: effects of chance are modeled as probability distributions that can form complicated networks, which can be analyzed deterministically or stochastically;
- *equilibrium*: no effective changes over time can be observed, for example, in steady-state balance;
- *nonequilibrium*: the system is dynamic and changes over time.

Additional incomplete and overlapping classifications of modeling approaches may serve as search terms, illustrating model diversity and jumpstarting further investigations for readers: backwards in time (e.g., coalescent models) ⇆ forwards in time (e.g., individual-based simulations); analytic ⇆ numeric ⇆ computational; binary ⇆ integer ⇆ continuous; constrained by input ⇆ output ⇆ equations; mechanistic ⇆ descriptive; absence of submodels ⇆ multilevel nesting; linear ⇆ nonlinear; time-based ⇆ event-based ⇆ static; graph-based ⇆ matrix-based ⇆ agent-based ⇆ logic-based (see Zeigler, 2012; Thiele *et al.*, 2012; E, 2011; Grimm and Railsback, 2005; Zeigler *et al.*, 2000; Law and Kelton, 2000). There are many more computer programming paradigms that also expand the list. In fact, Gödel's incompleteness theorem of mathematical formalisms suggests that infinitely many modeling and programming approaches exist, each of which is extremely powerful. Many different such approaches need to come together to enable the modeling of nested multi-scale evolution of diverse organisms from a rigorous probability theory perspective. We are very far from this goal; to make progress, we need a better grasp of the limits of various modeling approaches and how they might complement each other.

After a eureka experience of realizing how many questions could be answered by just one approach, researchers can be tempted to ignore its limits. However, even if the approach is a great hammer, not every problem is a nail. Ignoring these limits is dangerous when software assuming one model is applied to data for which the assumed model does not hold. This is particularly true for the many tools that do not automatically check for such problems (contributing much to the current reproducibility crisis). Biologists familiar with a broader range of modeling approaches are in a better position to choose a more reasonable approach for their biological question.

Model-building for any particular biological system is more efficient when researchers (1) have clear questions for their model, (2) can choose from a broad range of approaches to construct models, and (3) are well supported by corresponding tools that help to quantify uncertainty, detect modeling errors, and document results. Most approaches worth learning mirror *certain* modeling problems in biology extremely well and reasonably approximate others; however, all fail for *some* real-world systems, either due to unreasonable assumptions or computational intractability.

For any complex real-world problem, even the best models will always differ from reality and easily become misleading if interpreted from perspectives that violate their simplifying assumptions. Hence, the conclusion of the famous statistician George Box: "*All models are wrong, but some are useful*" (see Box and Draper, 2007, p. 414). Such usefulness either increases our mechanistic understanding or translates into success at prediction challenges in the real-world, where models are tested against empirical data. Estimating parameters for many biological models is a substantial challenge: the sloppy parameter sensitivities make it impossible to predict the importance of a given parameter without simulation, even in small nonlinear CSysBio models (Gutenkunst *et al.*, 2007). Difficulties compound for whole-cell models (Karr *et al.*, 2015), which suffer from the curse of dimensionality as the number of parameters increases. Aiming for reasonable reliability, complex models in CSysBio and elsewhere *must* quantify their uncertainty (Kirk *et al.*, 2015) and may want to manage their error budget (Parysow *et al.*, 2000). It is important for the reproducibility at the heart of science to avoid illusionary precision (Stodden, 2015). A combination of reasonable models and efficient tools can vastly expand a biologist's thinking capabilities and will be essential for reliable EvoSysBio analyses discussed below.

Systems Approaches Are Young

Systems modeling, big data and computational science (e.g., HPC university, see Section Relevant Websites) are young and lack the centuries of aggregated experience in weeding out unreliable results – unlike mathematics, physics, chemistry, and biology. The beginning of the previous section illustrates many problems in computational science simply by attempting to referencing Aristotle in the usual style: missing data (surname), data beyond an unreasonably narrow range (BCE), and uncertainty (year ± 22); if this were input for a scientific simulation, chances are that nobody would realize the hidden bias it introduces unless the model 'misbehaves' and produces 'unreasonable' results (as judged by competent, but not infallible experts).

The biggest challenge of the field is to improve the quality of models so they (1) describe causality chains that connect real-world input and assumptions to real-world output and conclusions in a reasonable way, (2) are complete, clear, and contradiction-free, (3) are readable and well documented for semantic reproducibility, (4) accurately quantify uncertainty of all claims for statistical reproducibility and explicitly define limits of applicability, (5) continually integrate new evidence as it becomes available, (6) run reproducible tests to reject claims that are not compatible with increasingly accurate observations, and (7) automatically trigger updates of other models that build on its output.

Reasonable models satisfy (1)–(4) for the data and evidence available at construction, but need (5)–(7) to *activate* them. State-of-the-art modeling is advanced enough to inspire such noble goals, but current model reliability is often much less clear. Much time, effort and organization will be required before such high-quality models can become standard in biology (Macklin *et al.*, 2014). Progress might be accelerated by dropping the principally impossible aims to 'validate' or 'verify' models of the natural world, which is not a closed system (Oreskes *et al.*, 1994). 'Validation' could more likely stifle further testing than encourage curious critical exploration of the 'valid.' The reliability of a model can be measured by its ability to withstand diverse and difficult tests that contrast model claims with the real world.

Questions are keys, models are locks. The key to assessing the reliability of a model is in the questions that motivated building it and that defined its purpose. This insight is both simple and profound. It is easy to see that the street map of a city (a type of model) does not help determine geological strata (wrong question for this model). It usually takes much more experience to identify similar model-data mismatches in 'big data' collections, which can be analyzed using unreasonable models to generate irreproducible results (Stodden, 2015). Hence remembering the question and relevant transformations is essential for avoiding results like '42,' also known as the answer to everything (Adams, 1979).

Questions that drive models are also essential for determining the level of details to include, which can be increased (*fine-graining*) or decreased (*coarse-graining*) in almost all cases. These decisions determine required parameters to be measured, indirectly estimated, collected from the literature, or imported from appropriate databases (such as BrendaEnzymes for kinetic parameters or BioNumbers for cell biology, see Section Relevant Websites). Much of the value of the abstractions developed by the New Evolutionary Synthesis since the 1920s is in their ability to guide coarse-graining and fine-graining across vast parts of biology. The formal EvoSysBio framework discussed below further refines this process.

Optimistically, and with some organization and standardization, the transition toward reliable computational modeling in biology will require far less time than the centuries it took to move from alchemy to chemistry. EvoSysBio could help catalyze these efforts, and the definitions of IOB, EOB, PGB, and TOB given in **Table 1** could be a step on the way.

Defining EvoSysBio

EvoSysBio, in its fullest form requires a union of insights from three very broad areas of research: evolutionary research (studies replicators, sometimes beyond biology), systems science (studies dynamic things consisting of parts), and biology (studies all aspects of carbon based life). Without such a union of insights and matching computational tools, progress in EvoSysBio remains a distant dream or a daunting yet diffuse challenge. This integrative reading of EvoSysBio reflects its trans-disciplinary mission to build bridges between many diverse disciplines, making it both extraordinarily broad and deep.

However, English syntax interprets combined terms as intersections, so that 'EvoSysBio' is less general than 'SysBio,' which already specializes 'Bio,' just as in 'House Keys.' This reading might see EvoSysBio as a highly specialized add-on with narrow applicability, a view hardly appropriate on principal grounds: virtually every 'thing' studied in biology can also be studied with the tools of systems science and has also been evolving or playing a role in evolution. Thus, EvoSysBio represents a panoramic perspective that matters for quantifying change over time in systems with self-replication.

This astonishing breadth of EvoSysBio implies that its results can depend on preceding breakthroughs in far-flung fields of evolution, systems biology, computing, and other disciplines, often developed by researchers without interests in EvoSysBio, evolution, or related topics. Once fully grown, EvoSysBio will be more keystone than add-on, as it brings together various tall pillars of evidence quantified with enough precision to meet in one place, essentially stabilizing the enormously complex structure of biology as a discipline.

Thus, EvoSysBio is much more than the occasional application of evolutionary insights in CSysBio or the occasional inclusion of metabolic network analyses in evolutionary genetics (see, e.g., Klipp *et al.*, 2009; Voit, 2013; Caetano-Anolles, 2010; Walhout *et al.*, 2013). Different attempts have been made to define EvoSysBio (Loewe, 2009, 2012; O'Malley, 2012; O'Malley and Soyer, 2012; Soyer and O'Malley, 2013; O'Malley *et al.*, 2015). To help clarify its contributions, the definition of EvoSysBio below integrates previous definitions and provides formal abstractions that can facilitate quantitative connections to the diverse, descriptive, experimental, and theoretical contributions necessary for building the rational models essential to EvoSysBio. Since EvoSysBio's core goal is to explain how evolution works mechanistically, it inherits the astonishing integrative abilities of evolutionary theory.

Table 1 Relationships among different focuses in biology

Focus of study	*Relationship to other focuses*
Intra-Organism Biology (IOB) *intra*, Latin for 'inside.' Relevant disciplines: *Biochemistry, molecular biology, gene regulation network biology, cell biology, cancer biology, developmental biology, physiology, neurobiology, and more.*	Focuses on actions, reactions, processes, parts, and patterns within a single, individual organism that enable it to live (grow, survive, reproduce, move, etc.) by changing its state in a given environment. IOB models simplify TOB models by assuming that genotypes and environmental interactions only matter where explicitly specified for a given individual organism and the nested organisms it may contain (e.g., gut bacteria, parasites, cancer cells). IOB details depend on environments; if the latter are known, then IOB allows in principle the prediction of all relevant fitness traits by constructing a complete Fitness Causality Network (see **Table 2**).
Extra-Organism Biology (EOB) *extra*, Latin for 'outside.' Relevant disciplines: *Ecology, biogeochemistry, climate science, population ecology, community ecology, ecosystem ecology, behavioral ecology, cognitive ecology, social ecology, and more.*	Focuses on actions, reactions, processes, parts, and patterns in ecosystems that alter states of individual organisms by interacting with biotic or abiotic environments. EOB models simplify PGB models by explicitly listing all relevant genotypic differences and assuming that no others matter. EOB models implicitly simplify IOB models by providing corresponding rates of interaction between individual organisms of the same or different types, even if these interactions are ultimately governed by biochemistry. Nested organisms have nested EOB ecosystems.
Population-Genetics Biology (PGB) *genetikos*, Classic Greek for 'generative' *γένεσις*, Classic Greek for 'origin.' Relevant disciplines: *Population genetics of single and multiple loci, quantitative genetics, breeding, coalescent theory, population genomics, population epigenetics, and more.*	Focuses on heritable information (genotypes, alleles, traits, methylation patterns, etc.) that can directly or indirectly affect fitness traits like survival and reproduction when passed on by individual organisms in evolving populations. PGB models simplify EOB models by using simple demographics (e.g., constant, exponential growth) to explicitly or implicitly summarize complex ecological mechanisms. PGB models also simplify IOB models to approximate genotype–phenotype-fitness maps for a given environment (e.g., multiplicative fitness models). An extreme focus on counting alleles in a population can turn PGB into efficient but abstract 'Bean-Bag Genetics' (Haldane, 1964). Nesting works as in EOB.
Trans-Organism Biology (TOB) *trans*, Latin for 'beyond' Relevant disciplines: *Evolutionary ecology, conservation biology, coevolution, sociobiology, metagenomics, phylogenetics, and more.*	Focuses on integrating EOB, PGB, and all direct or indirect actions, processes, and patterns that otherwise fall through the disciplinary cracks of EOB or PGB (excluding the inside of organisms). TOB models simplify IOB models. Examples range from the inclusive fitness of relatives (Gardner and West, 2014) to environmental DNA abundances in metagenomics (Wooley *et al.*, 2010) and phylogenetics at any nesting level (Felsenstein, 2004) or genotype by environment interactions (Pavlicev and Wagner, 2015). Nesting works as in EOB.
Evolutionary Systems Biology (EvoSysBio) Integrating 'keystone,' not 'add-on' Relevant disciplines: *All of above and more, including data science, computational modeling, network biology, information management, integrative biology or just 'biology'*	As a field, EvoSysBio has so far mostly used various informal definitions, which are not broader than the more formal definition given in the main text. *Formal EvoSysBio*: focuses on developing more reliable real-world 'flight-simulators for fitness landscapes' by integrating IOB and TOB using the five fundamental factors of evolution. Formal EvoSysBio drops the problematic aim to *directly* quantify the complex multi-dimensional dynamics of real-world fitness landscapes. Instead, it progresses by constructing **F**itness **C**ausality **Net**works (FCNets), which quantify and connect different types of **L**andscapes of **I**ncomplete **F**itness **T**raits (LIFTs; see **Tables 2** and **3**); formal EvoSysBio can then simulate evolving populations governed by dynamic, mechanistic IFT predictions from these FCNets. *Informal EvoSysBio:* builds bridges between IOB ('Current Systems Biology') and TOB ('Evolutionary Biology'). Even if using different terminology and not motivated by fitness landscapes, informal EvoSysBio results often contribute toward a better understanding of details that might directly or indirectly help to better quantify LIFTs.

Notes: The different interests of biologists often distinguish their focuses when studying organisms (see Table 2 for definitions). To create models relevant for their focus, researchers often simplify the models created by peers with other interests. EvoSysBio aims to integrate Intra-Organism and Trans-Organism Biology by using the five fundamental factors of evolution as shared abstractions to facilitate the construction of fitness landscapes. Other biological disciplines provide integration from different perspectives, for example, those of focal organisms (e.g., botany, virology, zoology), habitats (e.g., limnology), problems (e.g., cancer biology), organizational levels (e.g., cell biology), disciplines (e.g., biophysics), and more; but it is difficult to find biological theories more general than those evolutionary theories that enable, in principle, the construction of appropriately quantified fitness landscapes, thereby requiring the integration of vast amounts of biological results (Figure 2).

This integrative power can best be understood by reviewing how the definition of EvoSysBio developed in the past decade.

Partial Approaches

CSysBio has produced many analyses of molecular interaction networks; accordingly, 'evolutionary systems biology' has been used from the very beginning to denote the comparison of such molecular networks among different biological species (Stearns *et al.*, 2003; Medina, 2005). Such use might appear to contrast with more mechanistic frameworks (e.g., Loewe, 2009), which build on simulations that directly or indirectly include network interactions along with many other details affecting a system's dynamics. Network analyses often need only small subsets of the data required by corresponding mechanistic simulations. Therefore, collecting more data on the same system can, in principle, provide a common basis for using both approaches to ask the shared question of how life evolves. It seems unnecessarily confusing to fragment EvoSysBio into many different sub-definitions; one focusing on *network analysis*, another for *mechanistic simulations*, and many more for *different methodological approaches* (e.g., flux-balance-analysis (Ibarra *et al.*, 2002; Papp *et al.*, 2011), metabolic control theory (Kacser and Burns, 1981; Keightley, 1996), systems genetics (Markowetz and Boutros, 2015), etc.), others for *different biological levels of organization* (e.g., molecular functions (Dean and Thornton, 2007), cells (Lynch *et al.*, 2014), epigenetics (Hallgrimsson and Hall, 2011), development (Carroll, 2008), ecology (Pelletier *et al.*, 2009), etc.), and others for *different biological questions* (e.g., energetics (Watt, 1985), modularity (Wagner *et al.*, 2007), robustness (Payne and Wagner, 2014), game strategies (Pacheco *et al.*, 2014; Bohl *et al.*, 2014; Hummert *et al.*, 2014), etc.), or even focusing on *different perspectives provided by the five fundamental factors*, mutation (e.g., Loewe and Hill, 2010a,b), selection (e.g., Okasha, 2006), genetic drift (e.g., Lynch, 2007), recombination (e.g., Charlesworth *et al.*, 2009), and space (e.g., Westervelt and Cohen, 2012). All these contribute different important aspects to EvoSysBio, but where would the splintering stop?

Darwin and subsequent evolutionary geneticists have convincingly demonstrated the abstract elegance and unity of evolution based on the five fundamental factors. They demonstrated, in principle and with many examples, how these factors combined into a mechanism powerful enough to create the bewildering diversity of species and biological phenomena just by variations in the patterns of strengths of its fundamental factors. Ideally, the conceptual beauty of evolutionary genetics at the core of EvoSysBio will enable different important lines of EvoSysBio inquiry to sharpen evolutionary theory by extending a single consistent set of abstractions, which will make it easier for future researchers to build on its foundation. By comparison, defining different vital aspects of EvoSysBio in conflicting but partially overlapping terms seems less desirable if no clear conceptual integration can be provided.

Thus, the question of splintering EvoSysBio is related to the question of whether it is essential to extend evolutionary theory in any unusual way, for example, to integrate TOB interactions between genotypes and environments (Laland *et al.*, 2014). There is no doubt that evolutionary theory needs to be extended in many ways or else research would stop, but it seems unnecessary to extend it in any *unusual* way beyond its continued extension 'through relentless synthesis' (Wray *et al.*, 2014). EvoSysBio can add to this relentless New Synthesis in many ways, ideally by quantifying aspects that are necessary for mechanistically predicting evolution, one of the highest goals in evolutionary biology since Fisher, Wright, and Haldane laid the groundwork for the New Evolutionary Synthesis. The wish to avoid the splintering of disciplines has also inspired the wish to reintegrate all biological subdisciplines into a 'New Biology' (National Research Council USA, 2009).

Pragmatic Definitions

EvoSysBio has been defined as an effort to build bridges between CSysBio and evolutionary biology, integrating theoretical tools, experimental methods, and datasets from multiple disciplines into an evolutionary framework (see **Figure 2**; Soyer and O'Malley, 2013; Loewe, 2009). Comparable in breadth to CSysBio, this broad definition easily encompasses the diverse work associated with EvoSysBio (e.g. studies in the volume edited by Soyer (2012) and studies cited by Loewe (2009) and below). However, a more precise definition without loss of generality would be preferred, as it could facilitate creating powerful formal interfaces between EvoSysBio and other fields. These interfaces could help to interpret data collected elsewhere in its bigger evolutionary context. Conceptual advantages like these motivate a continued search for more precise definitions of EvoSysBio.

Formalizing EvoSysBio without Loss of Generality

Indeed, it turns out that EvoSysBio can be defined more precisely without loss of generality by building on the powers of abstraction offered by evolutionary theory. At the heart of this definition of EvoSysBio are evolving populations of *individual replicating organisms*, where evolution is governed by their *fitness landscapes*, a metaphor for mathematical causality statements about the traits of evolving populations in an abstract space with very many dimensions (see **Table 2** and Section Fitness Landscapes for more details).

To give EvoSysBio something to study, individual organisms have to exist and be near-perfect replicators of genotypic information, which affects phenotypic fitness traits like survival and reproduction in finite discrete populations. Hence, in these organisms all five fundamental factors of evolution can be active. This generic view includes viruses, microbes, and multicellular organisms (nested replicators), but also a fringe with exotic systems such as ribozymes (Martin *et al.*, 2015), prions (Li *et al.*, 2010), and data structures in computers (as exploited in biological individual-based simulations (Grimm and Railsback, 2005), evolutionary computation (De Jong, 2006), and artificial life studies (Adami, 1998)). Using the abstractions of evolutionary theory enables simulated populations with appropriately chosen parameters to closely mimic natural populations, facilitating their study. Organisms defined in this general way make it possible to see EvoSysBio as an integrating bridge between the great quantitative traditions of modern biology that complement each other through

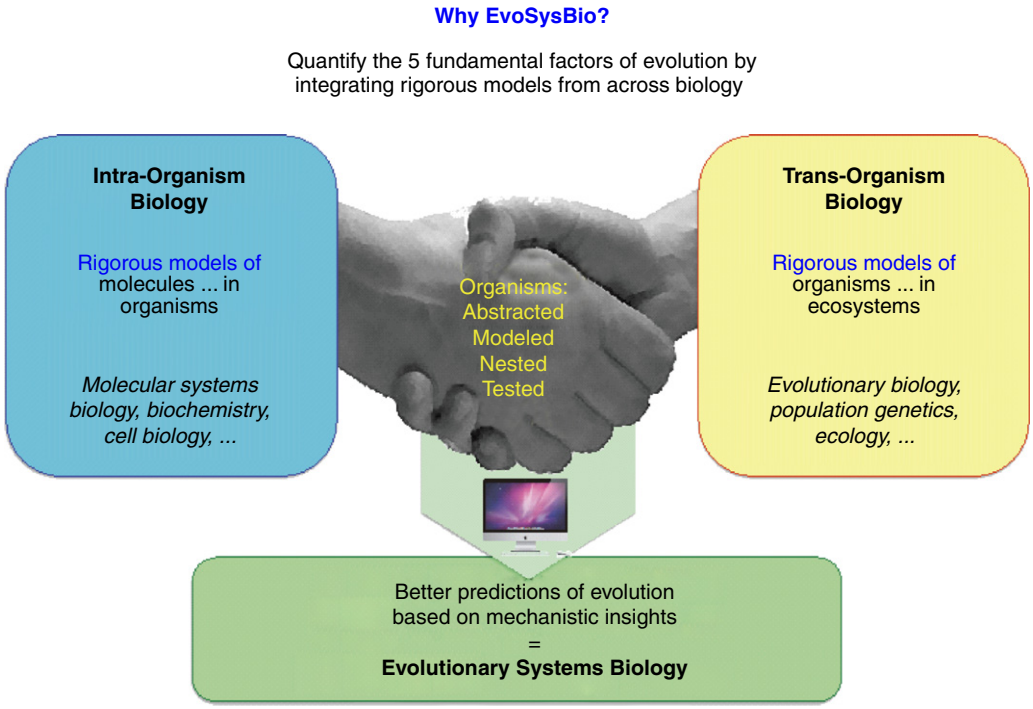

Figure 2 Overview of EvoSysBio integration. Traditionally, models from Intra-Organism Biology (IOB) and Trans-Organism Biology (TOB) have ignored each other to simplify rigorous modeling in their respective domains. This separation of concerns is facilitated by powerful abstractions such as the fitness of individual organisms and the five fundamental factors of evolution. When this separation fails at the cutting-edge of research, EvoSysBio offers a reversal that uses the same abstractions to bond branches of biology together. By measuring fitness traits of individual organisms and their contributions to the five factors, EvoSysBio enables asking new trans-disciplinary questions in addition to rigorously quantifying fitness landscapes. Fitness is either affected by – or itself affects – almost everything in biology, and 'nonquantitative' biology often breaks the conceptual ground for more precise studies; such observations have suggested: nothing in biology makes sense except *when properly quantified* in the light of evolution (Dobzhansky, 1973; Loewe, 2009). Picture credits: © Laurence Loewe (2015a), reuse under CC-BY 4.0, updated from previous versions (Loewe, 2009, 2012).

abstractions pioneered in population genetics (see **Table 1** and **Figure 2**):

(i) **I**ntra-**O**rganism **B**iology (IOB): biochemistry, molecular-, cell-, … -biology, … physiology; organisms may be nested (e.g., mitochondria in cells in humans; similarly gut bacteria, parasites, cancer cells, etc.). If nested, the IOB of the containing organism provides an environment for the smaller contained organisms, all of which come with their own IOB and TOB.

(ii) **T**rans-**O**rganism **B**iology (TOB): includes **E**xtra-**O**rganism **B**iology (EOB) disciplines such as ecology, **P**opulation **G**enetics **B**iology (PGB), the complementary genomics disciplines, and many diverse genotype-environment interactions (Pavlicev and Wagner, 2015). Nesting applies here too.

Over many decades, these traditions have produced outstanding empirical results and rigorous theory in their respective areas by using evolutionary fitness as a powerful abstraction for enabling an efficient separation of concerns. Both remain far from reaching their long-term goals: research in organisms ultimately aims to predict fitness and other phenotypic properties from genotypes (usually simplifying models of environmental changes), while population research ultimately aims to predict how populations evolve under a given set of environmental conditions (usually simplifying models of **IOB** to just a few numbers, such as fitness, epistasis, etc.). This mutual exclusion of broad areas of biology has enabled much progress, but the simplifications have started to break down for cutting-edge research. EvoSysBio can provide more detailed and explicit views of the relevant abstractions to facilitate more powerful evolutionary hypotheses that combine the strengths of both traditions. The elegance of these abstractions and the separation of concerns they enable between IOB and TOB are well illustrated by fitness landscapes, which provide a powerful intuition for explaining how evolution works (see Section Fitness Landscapes).

Based on these considerations and previous discussions (Loewe, 2009; National Research Council USA, 2009; Loewe, 2012; O'Malley and Soyer, 2012; Soyer, 2012; Calvert, 2012; Soyer and O'Malley, 2013; O'Malley, 2012; O'Malley *et al.*, 2015; Loewe *et al.*, 2015–2009), EvoSysBio can be defined in a way that leverages the integrative capabilities of fitness landscapes:

EvoSysBio is a trans-disciplinary **framework for constructing** reliable, testable, interactive **overviews of** nestable, dynamic, multi-dimensional **fitness landscapes**, **which mechanistically predict:** (i) **changes in fitness** of individual organisms when their states and environments change; (ii) **how populations evolve** when organisms traverse fitness landscapes.

This definition derives its formal power from IOB and TOB models of how individual organisms and populations change over time, respectively. These models can be efficient analytic summaries or approximations or computational models of much more complex underlying Continuous Time Markov Chains; they may require simulations to explore them. Reviewing relevant TOB theories (see e.g., Crow and Kimura, 1970, 2009; Charlesworth and Charlesworth, 2010; Kirkpatrick *et al.*, 2002; this Encyclopedia) and IOB theories (see e.g., Gillespie, 2007; Gillespie *et al.*, 2013; Klipp *et al.*, 2009; Gutenkunst *et al.*, 2007; Karr *et al.*, 2015; Kirk *et al.*, 2015; Section Systems Science: The Whole is Different from the Sum of Its Parts; **Table 3**) is beyond the scope of this article.

Informally, EvoSysBio and other work contributing to relevant computational models and appropriate visualizations might also be summarized as:

> EvoSysBio aims to improve the real-world reliability of 'flight-simulators for multidimensional fitness landscapes.'

This analogy is easily adapted to multi-dimensional data analyses, where researchers interactively explore interesting local features while relevant data is automatically selected, summarized, and presented by computers (as if 'flying' over the data; by comparison, manual coding for static snapshots feels 'pedestrian'). This requires a deep computational toolchain that integrates many diverse results extremely well; the approach thus enables EvoSysBio to leverage the full integrative power of *real-world* fitness landscapes, which are conceptually as powerful as evolution itself – known as the only theory general enough to unify biology.

In practice, this elegance comes at the steep price of requiring a deep mechanistic understanding of *real-world* fitness landscapes, an extraordinary challenge as relevant modeling expertise is often lacking in biology, statistics, formal modeling research, and more (see end of Section Systems Approaches to Genome Evolution).

This definition of EvoSysBio encourages developing more precise ways of defining fitness landscapes (see **Tables 2** and **3** and Loewe, 2009, 2012) to reduce the confusion surrounding the exact nature of these landscapes.

Fitness Landscapes

Sewall Wright (1932) introduced the now well-known concept of fitness landscapes to provide an intuitive understanding of how evolution works. These landscapes have been defined in various ways and labeled by different names, including adaptive landscapes, surfaces of selective value, and seascapes (Svensson and Calsbeek, 2012; Gavrilets, 2004; Loewe, 2009; Mustonen and Lassig, 2009). In statistics, similar landscapes have been described as response surfaces (Box and Draper, 2007).

Intuition behind Fitness Landscapes

Fitness landscapes provide a strong intuitive analogy for how evolution works. They are based on the familiar experience of walking in a hilly landscape. By representing potential genotypic states as points in a **plane** with fitness as the **height**, it becomes easy to see how evolution works:

> In a fitness landscape, *populations* evolve by moving toward higher (fitter) points whenever such points can be reached by the stochastic movement of *organisms* between positions in the plane (e.g., by mutation to other genotypic states); organisms tend to accumulate at these positions (genotypes) because there they are more likely to survive or reproduce (selection).

The strength of selection (selection coefficients) is governed by an organism's life history of survival and reproduction in comparison to other organisms in the same population. Selection and population size together govern the probabilities that the genotypes existing in one generation make it into the next, but these factors do not change genotypes (which is done by mutation and recombination). Thus, parents at higher points in the landscape have higher probabilities of generating offspring that also live at these higher points: the population evolves (see movie in **Figure 3**, Movie 1).

At a very high level, this intuition correctly captures the essence of what the complicated evolutionary mechanisms are all about: explaining why organisms are likely to exist in states of higher fitness. This intuitive nature of fitness landscapes has captured the imagination of many interested in evolution, from popular science writers to professional evolutionary biologists (Svensson and Calsbeek, 2012; Gavrilets, 2004; Loewe, 2009). Taken to one extreme, it leads to adaptationist interpretations of evolution that view populations *always* at the top of some adaptive peak and ignore the exceptions to the rule (e.g., quasi-species populations of RNA-viruses mutate at such high rates that they form 'bands around peaks' (Biebricher and Eigen, 2006); see **Figure 3**). Taken to another extreme, fitness landscapes become neutral networks, where all genotypes share the same fitness (Kimura, 1983).

Criticisms and Limitations

Critics of fitness landscapes as a concept highlight (1) the frequent lack of precise definitions, (2) its abstract nature, which does not easily lend itself to directly proposing new experiments, and (3) general human difficulties with navigating multi-dimensional spaces (Kaplan, 2008).

Little can be done about the limits of human imagination, which often fails when attempting to translate the multi-dimensional planes and heights of fitness landscapes into more familiar spatial dimensions. Different useful visualizations have been developed for highlighting various aspects of fitness landscapes (McCandlish, 2011), yet it is difficult to convey an undistorted picture. Visualizing landscapes from larger EvoSysBio projects may require interactivity, but appropriate 'flight simulators' are missing. These problems have not prevented numerous speculative drawings, which are prone to misinterpretation. It might be beneficial for the field to label them accordingly to contrast them with *attempts* that precisely define and measure the much less comprehensive real-world LIFTs (see below and **Figure 4**).

Table 2 Concepts related to fitness landscapes

Concept	*Description*
Fitness Landscape Ultimate EvoSysBio goal	An abstract 'landscape' defined by causal 'positions in a plane' and their consequential 'heights' as defined by a corresponding FCNet, mapping genotypic traits to IFTs (see below). Each position in a plane describes a potential causal state of an individual organism and its given environment; both govern its height(s). Fitness landscapes are near-impossible to measure, compute, or visualize, since both planes and heights often have very many dimensions and their 'points' often turn into fuzzy distributions due to stochasticity. Nevertheless, Evolutionary biology studies them due to their importance (even if resorting to MOCA-LIFTs, see below), and EvoSysBio formalizes these studies to make them more efficient.
Fitness Causality Network FCNets define fitness Landscapes	A network of nodes (data) and links (functions), together describing causal influences on consequential IFTs over a given time period. Causal influences include genotypes, environments, and initial states (e.g., maternal methylation patterns of DNA). Nodes in FCNets are defined by IFTs and links between nodes by LIFTs. A complete FCNet of a single organism links all LIFTs (see **Table 3**) relevant for predicting the fitness consequences of mutations and environmental changes.
Landscape of Incomplete Fitness Traits LIFT, LIFTs are always context specific	An abstract representation of a fitness landscape that maps causal IFTs to consequential IFTs, thus providing a causality statement about how input governs output, comparable to probabilistic mathematical functions. Inputs are distributions of causal states of organisms ('positions in a plane,' e.g., mutations), and outputs are distributions of consequential IFTs ('local heights,' e.g., survival, reproduction); both only apply to a given context and time duration. IFT designations 'causal' and 'consequential' are relative to a LIFT, sometimes indicated by adding 'more' to distinguish 'more causal IFTs' from the 'most causal IFTs' (e.g., DNA). If complete and rigorous, these representations of fitness landscapes can be viewed as existence probability theorems that facilitate predictions of the probability that a given number of organisms will exist in a given set of locations of the plane. Such results are mathematically related to the analysis of Continuous Time Markov-Chains, which is extremely useful for defining and analyzing such fitness landscapes. This technical aspect makes LIFTs markedly different from the MOCA-LIFTs often used for illustration (see below). The strength of LIFTs is in their simplicity, which enables precise measurements and simulations. See **Figures 7** and **8** for examples.
Incomplete Fitness Trait IFTs never quantify fitness in full	A phenotypic trait that impacts aspects of the fitness of an organism, its offspring, or its genetic relatives. Traits are IFTs when they, at least occasionally, affect rates of survival, reproduction, merging, etc., or when they modify evolutionary factors (mutation, migration, etc.) in one or more environments, at least to a small degree.
Organism Single individual replicating system	A single, individual system that consists of different parts that builds a whole, which – in biology – must be able to replicate and may be nestable. Intra-Organism Biology (IOB) investigates how these parts work together. Extra-Organism Biology (EOB) studies how each individual organism interacts with its environment and other organisms of any type. If organisms (i) use hereditary information to grow on environmental resources and (ii) can produce identical or similar descendants, then their populations necessarily evolve as modeled in Population-Genetics Biology (PGB). Remaining indirect interactions are captured by Trans-Organism Biology (TOB), which also integrates EOB and PGB at its level of replication (but not IOB). Note the simplification when a huge nested stack of replicators is denoted as a 'single multi-cellular organism.'
Nesting (in Organisms) Organisms in organisms	A nest organism is a single, individual organism that contains one or more populations of different types of nested organisms, which replicate(d) by themselves. A nested organism is a single, individual organism encapsulated by a larger nest organism, which is the smaller nested organism's environment. Organisms can be both nest and nested organisms at the same time. Examples of nesting include mitochondria in cells in multi-cellular organisms, microbes in the gut, viruses in cells, parasites in hosts, cancer in patients, and more. In each case, the smaller contained ('nested') organism is nested in the larger containing ('nest') organism. The Trans-Organism Biology (TOB) of evolving populations of nested organisms and the Intra-Organism Biology (IOB) of the nest organism affect each other, but the IOB of a nest organism has to describe the TOB of its nested organisms well enough to exclude relevant surprises (e.g., evolution of tumors). Each population of nested organisms needs its own EvoSysBio analysis, where the IOB of the nest organism plays the role of the ecological environment in the TOB of the nested organism. The complex TOB of nested organisms can sometimes be simplified by assuming nested populations are homogeneous and incapable of mutation.
MOCA-LIFT 'Instant' LIFT cartoons	**M**assively **O**versimplified **C**artoonish **A**bstract **LIFT**s are essentially like LIFTs, but without much clarity on how their various dimensions map to reality. Their cartoon-like distortions can be caused by complex transformations, lack of data, imprecise definitions and other problems that are often difficult to resolve. The value of a MOCA-LIFT can range from useless speculation to useful stop-gap (capturing some aspects of reality) to ground-breaking 'Gedankenexperient.' While researchers put in the enormous effort to measure more rigorous LIFTs, they will likely construct increasingly realistic MOCA-LIFTs, since their cartoonish nature facilitates smooth transitions from fact-free figures to high-precision result repositories with varying degrees of usefulness. As a rule of thumb: LIFTs that are (i) not fully defined by measurements or simulations, (ii) are not quantified in defined units, and (iii) do not specify their real-world uncertainty are MOCA-LIFTs (see **Figures 3**, **4**, and **6** for examples).

Notes: A fitness landscape provides an intuitive understanding of how complicated evolutionary mechanisms impact the fitness of an organism. To comprehend the complexities of fitness landscapes and use them to predict the fitness consequences of mutational and environmental change, it is important to understand (i) the concept of an organism as it relates to biology, (ii) how researchers use Incomplete Fitness Traits (IFTs) to construct Landscapes of IFTs (LIFTs) that represent different aspects of the fitness landscape of an organism, and (iii) how IFTs and LIFTs form FCNets, which mediate finely tuned trade-offs in response to the environment. Only when complete FCNets are fully quantified is it possible to compute a complete fitness landscape, which can predict the distributions of phenotypes and fitness from an organism's distributions of genotypes and environments. See Table 3 for the types of LIFTs that can collectively bridge the whole gap from DNA to fitness.

Table 3 **Fitness predictions from genotypes** are, in principle, enabled by the types of **L**andscapes of **I**ncomplete **F**itness **T**raits (LIFTs) shown here

PT: Plane Type	→	***LT: Landscape Type***	→	***HT: Height Type***
Compute: Input type	→	Compute: Function type	→	Compute: Output type
FCNet: causal Node type	→	FCNet: Link type	→	FCNet: consequential Node
More causal IFT type	→	*LIFT: Landscape of IFT type*	→	*More consequential IFT type*
Type of multidimensional point in the causal plane of a FCNet node is given by its input data type that governs height as calculated by its LIFT type.	→	*Type of LIFT that represents a mapping function type that accepts a point in its causal plane as input and computes a consequential height as output. LIFTs define types of links between FCNet nodes in a larger FCNet.*	→	*Type of multidimensional point of consequential height that integrates all effects from replication (with all existing nested levels). It is computed as output type by its LIFT from a given input point in its plane type.*
PT7: Organism life history of fitness traits: survival, reproduction, etc.	→	LT7: **Summarizing** statistics of survival, reproduction, etc. for relevant groups of organisms over a given duration of time in a specified environment	→	HT7: Fitness summary statistics for organism genotypes in specified environments per time interval
PT6: Real-world Incomplete Fitness Traits (RIFTs)	→	LT6: **Balancing** trade-offs in IFT networks of organism life-history and physiology	→	HT6: Organism life history: survival, reproduction, etc.
PT5: Simulated Incomplete Fitness Traits (SIFTs)	→	LT5: **Mapping** simulations (*in silico* IFTs) to observed real-world IFTs	→	HT5: Real-world Incomplete Fitness Traits (RIFTs)
PT4: Time-series of phenotypic traits	→	LT4: **Extracting** fitness relevant traits (IFTs) from time-series analyses	→	HT4: Simulated Incomplete Fitness Traits (SIFTs)
PT3: Molecular functions network	→	LT3: **Simulating** dynamic time-series predictions (over multiple scales) for a given initial state in a given environment	→	HT3: Time-series of phenotypic traits
PT2: Molecular structures	→	LT2: **Abstracting** the structural biology of structure–function relationships	→	HT2: Molecular functions network
PT1: Hereditary information	→	LT1: **Folding** of expressed hereditary information	→	HT1: Molecular structures

Note: All LIFTs contribute some aspect to constructing the full **F**itness **C**ausality **Net**work (FCNet) that governs the dynamic distributions of various incomplete fitness traits of an organism. Each FCNet node connects at least two LIFTs (In → Out; see header row for terminology illustrative for different contexts). The fitness of an individual organism depends on the structure of its FCNet and on its genotype (broadly defined as any hereditary material), its environment (abiotic and biotic), its initial state (immediately after being 'produced'), and the time period over which the expected change in organism numbers is to be measured. Real-world examples exist for all LIFT types in this table, albeit spread across different organisms (see Loewe, 2009). Defining a full FCNet for any *single* organism will be a major milestone for EvoSysBio and requires the computational integration of all LIFT types in this table that are required for representing the organism. These LIFTs require many independent studies of Simulated IFTs (SIFTs) and Real-world IFTs measurements (RIFTs). The huge costs and high risks of pursuing this integrative vision are matched by the high rewards of enabling rational, *in silico* analyses of arbitrary mutations. Hence, such EvoSysBio work is at the core of personalizing medicine and reliably predicting evolution.

Fitness Landscapes Map to Real-World Concepts

Criticism of the notorious lack of quantitative and semantic precision in depictions of fitness landscapes does not imply that they are not real or cannot be defined rigorously.

(i) As introduced above, real-world fitness landscapes *are by definition* all possible states of real-world populations of replicating organisms, which *by definition* are subject to the five fundamental factors of evolution that describe how organisms traverse fitness landscapes.

(ii) *Models* of fitness landscapes *aim to approximate* real-world fitness landscapes by choosing features deemed relevant by modelers; thus, human reconstructions of fitness landscapes are *by definition not perfect*.

(iii) The quality of these approximations is determined by *testing* how useful their predictions and/or insights are for navigating their real-world counterparts.

(iv) Neither existence nor quality of these approximations influences real-world fitness landscapes, as long as they do not affect how humans shape their world.

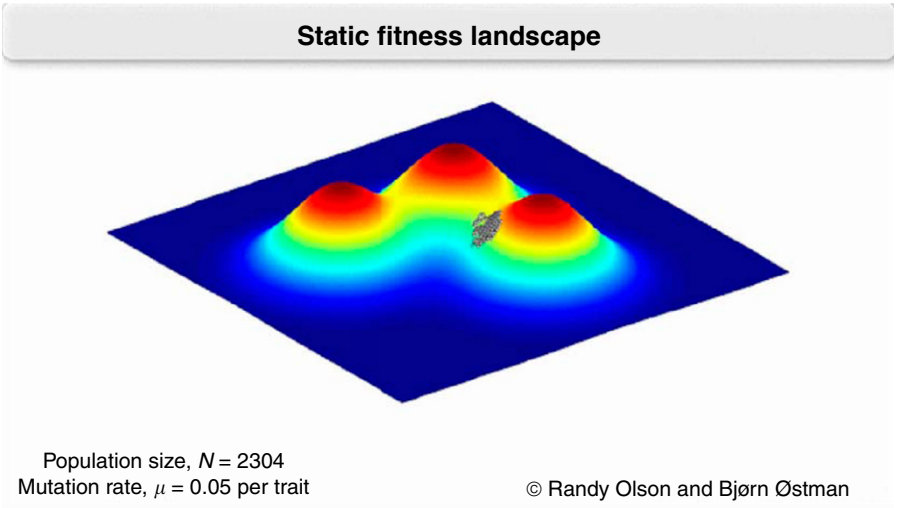

Figure 3 Future EvoSysBio 'flight simulators for fitness landscapes' might provide dynamic overviews of how populations evolve on measured data-rich Landscapes of Incomplete Fitness Traits. This MOCA-LIFT (see **Figure 4**) shows a snapshot from a cartoonish movie of a population that evolves on a static MOCA-LIFT (for more details, see also links in Section Relevant Websites). Picture credits: © Østman and Olson (2014a), reusable under CC-BY-SA 3.0.

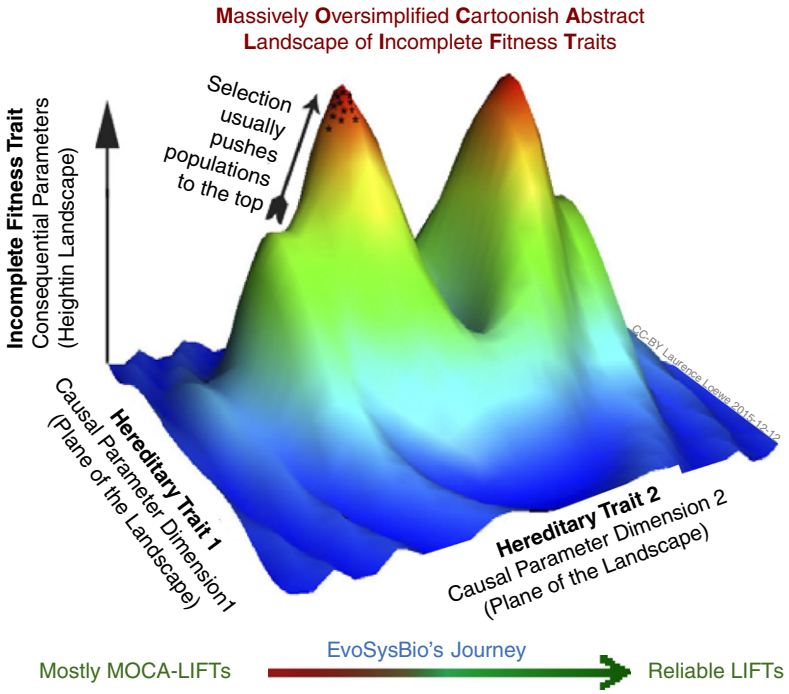

Figure 4 **M**assively **O**versimplified **C**artoonish **A**bstract **L**andscape of **I**ncomplete **F**itness **T**raits. Such a MOCA-LIFT captures the central causality statement of a fitness landscape: each position in the causal plane determines its consequential height. Its value usually ends here, as real-world LIFTs tend to differ substantially. Picture credits: © Laurence Loewe (2015a), reuse under CC-BY 4.0, updated from previous versions (Loewe, 2009, 2012).

(v) Human measuring or managing can severely limit the ability of EvoSysBio to improve model quality if these actions substantially perturb real-world fitness landscapes.

Fitness landscapes play a central role in the EvoSysBio definition because they integrate all fundamental factors of evolution. Quantifying fitness landscapes accurately is a key long-term goal for EvoSysBio (and other systems approaches to genome evolution, even if they have other priorities).

Though compelling, the mechanistic goals of EvoSysBio can be extraordinarily difficult to achieve. In such cases, less-than-ideal fitness landscape definitions can enable important progress. These may (1) *only describe* observed phenomena without requiring explicit rational hypotheses on real-world

causal mechanisms, or (2) are *only system-specific* and so enmeshed with their particular context that abstraction obscures their meaning, or (3) are so *highly abstracted* from the real world that they are essentially void of interpretable biological facts.

How to Avoid Misreading Fitness Landscapes

The following simple principles might help to keep the useful aspects of **fitness** landscapes while reducing the confusion:

1. **Fitness landscapes exist** for all organisms that produce offspring at all nested levels of replication, even if the landscapes are impossibly difficult to define.
2. **Pictures mislead.** A single picture of a fitness landscape is usually misleading, as it can rarely be created without distorting vital views of the underlying model when distilling the potentially many dimensions of a plane and height down to the three dimensions humans are used to.
3. **Uncertainty quantifies accuracy.** As elsewhere, quantifications of fitness landscapes are at most as reliable as the uncertainty quantifications performed for them from diverse perspectives to obtain reliable error bounds. Hence, the qualities of uncertainty measures of a landscape indicate the accuracy with which it was measured.
4. **Fitness has many measures**. The definition of fitness landscapes is intricately linked to the definition of fitness, which has its own challenges. While it first appears easy enough to measure (survival, reproduction), crucial complications appear when the subtle timing and indirect effects of real life are accounted for as well. Life history theory has compiled a substantial body of theory on how multiple traits combine in order to contribute to fitness (Stearns, 1992, 2000; Brommer, 2000; Houle *et al.*, 2011; Gardner, 2015; Gardner and West, 2014).
5. ***Incomplete F*itness *T*raits (IFTs) are easily defined**. The current inability to define or compute *complete* fitness measures does not prevent the definition and measurement of IFTs, which are defined as having at least some probability to impact survival and/or reproduction in at least one known environment (irrespective of whether the traits are continuous or not). Defining such IFTs is comparatively easy, and predicting them computationally, as well as observing them in the lab, is already possible for some IFTs. Thus, researchers can accumulate now the observations and prediction capabilities required for later EvoSysBio analyses.

From MOCA-LIFTs to Reasonable LIFTs

Intuitive appeal with difficulties of measurement and visualization have generated many speculative depictions of 'fitness landscapes'; most are probably misleading and deserve a name that distinguishes them from real-world measurements. Figure 4 shows such an entirely fact-free ***M*assively *O*versimplified *C*artoonish *A*bstract *L*andscape of *I*ncomplete *F*itness *T*raits**, or in short, a **MOCA-LIFT**:

- *Massively Oversimplified*: the *many* dimensions of plane and height are arbitrarily distilled to one or two dimensions, often without defined mapping or appropriate justification;
- *Cartoonish*: the landscape only captures a causality statement; a lack of real-world data precludes any further statements;
- *Abstract*: plane and height do not represent any biological systems in the real world.

Unless spoiled by 'MOCA properties,' LIFTs represent real-world progress in EvoSysBio, as LIFTs measure, model, simulate, and/or summarize real-world biology in a reproducible way:

- *Landscape*: a function that maps a multi-dimensional point in a plane of *more causal* IFTs (input) to its computable multidimensional height, which is a *more consequential* IFT (output);
- *Incomplete*: the recognition that IFTs are not complete and do not even attempt to be; IFTs may or may not affect other traits or be affected by other traits; all LIFT statements are conditional on 'all else being equal';
- *Fitness*: the statement that some aspect of this trait affects survival, replication, or some other evolutionary factor directly or indirectly at least with a small probability in some environments;
- *Trait*: a type of property of an organism; if traits do not affect fitness, they may be only phenotypic (e.g., gene expression without impact on fitness is irrelevant), or neutral (e.g., DNA sequences never expressed and without any impact on fitness).

It often makes more biological sense to investigate LIFTs than to attempt direct predictions of fitness (which often require abstracting too many complex processes at once, increasing the likelihood of failure). At their core, LIFTs can be thought of as the basic building blocks of fitness landscapes; they provide the most direct link between real-world biology and EvoSysBio simulations, either by interpolating real-world measurements or by simulating known processes. Deliberately ignoring the bigger picture of fitness in favor of simplicity makes basic LIFTs much easier to use and provides two additional conceptual advantages, one experimental and one conceptual.

Experimentally, the artificial nature of many LIFTs can remove them far enough from the finely tuned fitness trade-offs that dominate in the wild; it might thus be easier to find IFT mutants that are measurably different from wild-types, enabling tests of *in silico* prediction quality (see Loewe, 2009; Figure 3 there, read IFT for Candidate Fitness Correlate).

Conceptually, large collections of basic building-block LIFTs may facilitate the construction of **F**itness **C**ausality **Net**works (FCNets, see Table 2), which capture all causal influences that govern the expected distribution of fitness values for an individual organism over a given time interval in a given environment. Table 3 gives an overview of the diverse LIFT types and fitness causality node types that enable, in principle, the construction of a full causality chain, which ranges from DNA to fitness. Biological examples for each particular LIFT type have been given elsewhere (see discussion of Table 2 in Loewe, 2009) and details of linking them into full FCNets are beyond the scope of this article.

An alternative to mechanistic predictions of LIFT networks is to observe evolution by empirically measuring the speed at which different genotypes grow in a given environment. Such direct IFT measurements can be conducted for each individual genotype, as has been done for bacterial genotypes with respect to their ability to survive antibiotics (see below, EvoSysBio Milestone 5: Antibiotics Resistance Evolution). Several empirically observed LIFTs (Weinreich *et al.*, 2013) have been compiled for visualization with the MAGELLAN tool (Brouillet *et al.*, 2015). These measurements can provide important high-resolution views of very small local LIFT areas, assuming the context of otherwise constant fitness landscapes.

It is also possible to mix empirical observations of IFTs with statistical modeling in order to predict fitness landscapes with very many points from measurements of much fewer genotypes. As more big data sets become available, this approach becomes increasingly powerful, fueling a revival of the fitness landscape paradigm (Schuster, 2012). It has been used to computationally explore the complexity of HIV for *in vitro* LIFTs (Kouyos *et al.*, 2012). Statistical methods and quasi-species theory have also been employed to infer fitness landscapes *in vivo* using sequenced HIV samples from patients (Seifert *et al.*, 2015).

As evolutionary geneticists have struggled with the question of how to best capture relevant glimpses of fitness landscapes, they have developed a number of useful abstractions. Some of these quantify particular aspects of fitness landscapes and build on formalisms, which produce general complex landscapes of different types (Orr, 2005; Gravner *et al.*, 2007). Others analyze fitness landscapes from particular angles, including speciation (Gavrilets, 1997, 2004), game theory (Nowak and Sigmund, 2004), and more (e.g., Svensson and Calsbeek, 2012; Richter and Engelbrecht, 2014). Also, evolutionary geneticists have defined evolutionary parameters that, in principle, could be measured in the real world or be computed from fully known fitness landscapes. These provide excellent summaries of particular aspects of fitness landscapes.

Abstractions for Aspects of Fitness Landscapes

The complexity of multi-dimensional fitness landscapes and the notorious difficulties of exploring them have long been motivating evolutionary biologists to develop concepts that quantify more limited aspects of fitness landscapes, sometimes empirically or without requiring a full understanding. Such incomplete empirical summary statistics of fitness landscapes include:

- *distributions of mutational effects on fitness*: pick a point on the landscape as wildtype starting point, jump into all directions that represent genotype changes from naturally occurring mutations, then compare fitness to observe selection coefficients (e.g., Schenk *et al.*, 2012; Eyre-Walker and Keightley, 2007; Loewe and Charlesworth, 2006; Loewe and Hillston, 2008);
- *epistasis*: interactions between mutations that increase or decrease the effects of additional mutations play a major role in evolution, but nomenclature can be confusing (Wolf *et al.*, 2000; Phillips, 2008; Loewe and Hill, 2010a); epistasis captures the gene-regulatory and biochemical reaction network complexity of IOB (Phillips, 2008), so it is no surprise that higher-order epistasis can result in surprising changes of fitness (Weinreich *et al.*, 2013); it can sometimes be measured (Schenk *et al.*, 2012; Schenk and de Visser, 2013); it can determine the accessibility of certain evolutionary paths (Poelwijk *et al.*, 2007; Weinreich *et al.*, 2006) and affect robustness, even within a protein (Bershtein *et al.*, 2006); epistasis is also important for understanding the evolution of antibiotic resistance (e.g., MacLean *et al.*, 2010; Hall and MacLean, 2011; Schenk *et al.*, 2012; Schenk and de Visser, 2013);
- *robustness*: helps developmental or simpler processes produce patterns that reduce observable changes in phenotypes (e.g., Wagner, 2014, 2012; Payne and Wagner, 2014);
- *fragility* or *capacitance*: is the opposite of *robustness*; increases variability beyond usual amounts (e.g., Bergman and Siegal, 2003).

These measures can be used for investigating the evolvability of a system (e.g., Wagner, 2005) or its mechanisms of adaptation (e.g., Wagner, 2011). More on summary statistics of fitness landscapes can be found elsewhere (see Loewe, 2009 and in this Encyclopedia).

Practical Relevance of EvoSysBio

Ideas for applying EvoSysBio to solve practical problems are easy to conceive. From agriculture to medicine to zoos, replicating organisms are everywhere. For some, replication is desirable (e.g., rare species in zoos), for others not (e.g., cancer cells, agricultural pests, superbugs). Humans can sometimes shape the impact of these organisms as planned by increasing growth of desired organisms and blocking growth of undesired ones. Success usually requires the ability to predict growth with some reliability, which in turn often requires a deeper understanding of evolution, as also needed for computationally exploring potential management decisions and their side effects. EvoSysBio models can help by providing a rich framework that facilitates the integration of all important aspects of IOB, EOB, PGB, and TOB (see **Table 1**). If not overwhelmed by statistical prediction errors or numerical rounding errors, such mechanistic models might reliably predict important practical aspects of very rare, high-impact events, such as the extinction of endangered species, catastrophic virus epidemics, the evolution of superbugs resistant to all known antibiotics, or the origin of tumors.

The next milestones on the long journey to formal EvoSysBio analyses mirror many models in mixing mechanistic and descriptive aspects. Using their motivating questions and chosen levels of abstractions, most models combine known cause and effect mechanisms with a coarse-grained phenomenological basis that merely describes statistical estimates of empirical data. For example, modeling metabolic regulation neither requires a simulation of the full quantum mechanics of ribosomes (too fine-grained) nor empirically observed cell-division rates (too coarse-grained). Summarizing individuals as fitness values is the ultimate coarse-graining in IOB; however, to merge IOB and TOB requires the inclusion of substantially more details than currently possible. The following milestones mark important points in building the capabilities for such broad integration.

EvoSysBio Milestone 1: Observed LIFTs

This marks the definition and observation of a new consequential IFT for a set of diverse causal traits that mechanistically specify how consequential IFTs are computed from causal traits, irrespective of how many LIFT types from **Table 3** are implicitly integrated. For example, consider the low probability that light-colored **mice** will be caught by birds of prey on sandy hills (Linnen *et al.*, 2009). It is easy to map genotypes to phenotypes if coat color is controlled by few known genes and mutations of coat color can be recognized in DNA sequences. Thus, a LIFT for survival in such an environment can be trivial to predict if good measurements of predation risks are available for different coat colors. Full measurements of fitness are much more complicated, as mice can die from many causes and reproductive success requires essentially a prediction of everything a mouse can do in that environment. The complexities of such measurements are beyond what most researchers would be willing to contemplate, let alone adding similarly complex models for the corresponding birds of prey, whose survival may depend on their ability to learn how to detect the mice (adding complex neuronal and evolutionary feedback loops). This example illustrates key reasons behind the 'Incomplete' in IFTs: measurements of mouse survival are clearly fitness relevant; yet they are also clearly incomplete and need to be complemented by additional studies that might never be conducted if 'fitness has been measured in this mouse.'

A simpler example uses flux-balance analysis to predict *in silico* how *Escherichia coli* evolves by adapting to a certain environment (Edwards *et al.*, 2001). It only needs a network of relevant metabolic reaction stoichiometries and the assumption of flux-balance equilibrium (influx = efflux everywhere). Such models are easily coupled with genomic datasets that indicate the presence of genes for particular enzymes that catalyze certain metabolic reactions; thus, inferring relevant metabolic networks becomes much easier. Additional empirically observed LIFTs are listed elsewhere (Brouillet *et al.*, 2015).

EvoSysBio Milestone 2: Fragmented LIFTs

Biology has now conceptually mapped much of the most causal LIFTs in the FCNet of model organisms (DNA sequences available). It has also followed diverse LIFTs in the network, spanning all the way to direct fitness contributions. At least since 2009, it has become possible to provide specific realistic examples for *all* LIFT types specified in **Table 3** (see discussion of **Table 2** in Loewe, 2009). Thus, each critical step on the full path from DNA to fitness can, in principle, be modeled; such LIFT-type models have been developed independently in different model-organisms, greatly complicating a potential integration. As biological research continues, the addition of increasing numbers of LIFTs will simply connect different types of LIFTs in the same organism.

EvoSysBio Milestone 3: Simulate a Whole Cell

While the first two milestones were passed some time ago, somewhere on the EvoSysBio journey toward integrating IOB and TOB it must become possible to mechanistically simulate a whole cell in a simple lab environment. This milestone is well aligned with the goal of Evolutionary Cell Biology (Lynch *et al.*, 2014), which requires the quantitative integration of a very large number of diverse quantitative models of cellular subsystems. It has recently been demonstrated that a very simple (but complete) single cell in a simple growth medium can be simulated at the level of biochemical interaction over a whole cell cycle (Karr *et al.*, 2012), though much work remains before such simulations become robust enough and available for more complicated cells. While organizing and curating the thousands of parameters required for simulating the complete biochemical reaction network of a single cell remains challenging (Macklin *et al.*, 2014), obtaining reasonable values for them can be even harder (Karr *et al.*, 2015). New experimental methods provide a wealth of diverse information about cellular processes that was previously difficult to conceive: genomic ontologies are being constructed to list all genes, types of RNAs and proteins (many with functional annotations); advances in microscopy are approaching imaging at atomic resolution in cells (Kuhlbrandt, 2014); fluorescent proteins allow the recording of live single-cell time series of intracellular amounts or tracking the movement of individual molecules. In fact, for some types of cells, it is no longer clear if EvoSysBio is more limited by the need to observe relevant data or the need to organize, interpret, and integrate data that has already been collected (Macklin *et al.*, 2014). This growing need for more efficient data analysis can also be seen in sequencing, where the cost of analysis can substantially exceed the cost of obtaining raw data.

EvoSysBio Milestone 4: Predict Mouse Cancer

Successfully predicting the growth of diverse, complex cells on controlled growth media might encourage addressing the challenges posed by more complex ecologies. The need to understand the full-scale ecology of an animal or plant can be postponed by focusing on the small-scale 'internal ecology' of the body of a mouse from the perspective of cancer cells. In cancer cell biology, cells cannot grow outside their 'mouse-body-ecosystem'; from an experimentalist perspective, mice with known cancer mutations are about as well-controlled and well-studied as ecosystems can be. While predicting the full IOB of a mouse will remain very unlikely for some time, it is easily replicated with high accuracy by growing more mice, all of which can be analyzed with the tools of modern biology. Thus, in comparison to the full-scale EOB of wild mice, it is relatively easy to quantify the IOB of mice for the purposes of describing the environment that controls much of the possibilities of mouse cancer cells. Such mice facilitate addressing a number of interesting challenges in EvoSysBio and cancer research, simply because they can be so extraordinarily well characterized. Given the big interest in mouse cancer research, important breakthroughs are more likely to occur here first.

Cancer therapy resistance remains the most difficult challenge in the diagnosis and treatment of cancer. It is enabled by diverse populations of cancer cells, some of which keep surviving therapy to grow back. Populations of cancer cells are, in principle, governed by the same complex TOB processes studied in ecology and evolutionary biology. Measurements of survival and reproduction rates of cells enable

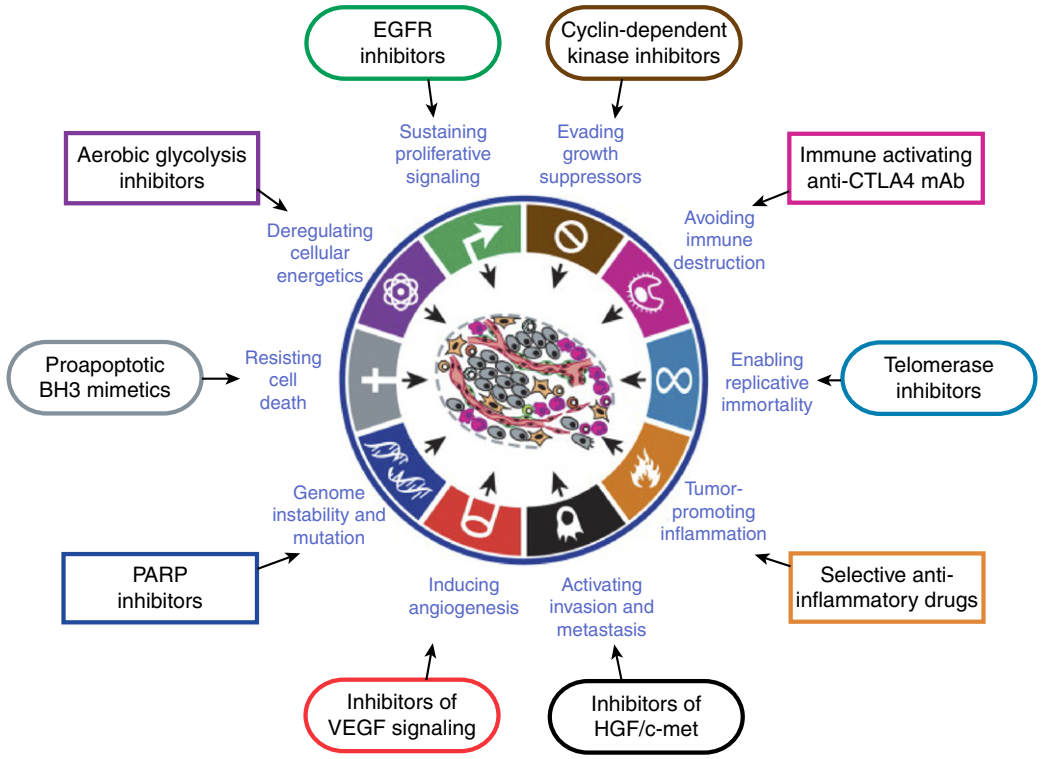

Figure 5 The hallmarks of cancer have strong links to either survival incomplete fitness traits (IFTs) or reproductive IFTs, hence the evolutionary processes triggered by cancer cell biology are interesting test cases for EvoSysBio. Picture credits: Hanahan and Weinberg (2011) © Elsevier, reuse now allowed under CC-BY-SA 4.0 (http://creativecommons.org/licenses/by-sa/4.0/).

predicting evolutionary outcomes of a population in a given environment. Survival of cancer cells in therapy depends on the dynamics of intracellular reaction networks, which can, in principle, be studied with the modeling tools of CSysBio. EvoSysBio aims to combine the theoretical modeling with the real-world observations needed to simulate evolutionary outcomes for cancer populations with increasing precision.

Cancer is an evolutionary process of a population of cells that replicate too much and can grow outside of their normal boundaries. A substantial number of survival and reproductive traits in cancer cells provide a target for natural selection among them (see **Figure 5**): Evolutionary factors like surviving attacks from the immune system, mutating, migrating to new tissues, and other such traits have vexed cancer biologists for a long time (Hanahan and Weinberg, 2011). It is not difficult to define corresponding IFTs for cancer cells and all fundamental factors of evolution are active (Gerlinger *et al.*, 2014).

Mouse cancer as a milestone for EvoSysBio is reasonable due to high independent interest, a wealth of existing information, and readily applicable methodologies. Once it becomes routinely possible to simulate the biochemistry of whole cells with reasonable accuracy, these techniques could be used to quantify relevant mouse IOB and define the TOB for a type of mouse cancer that is comparatively well understood. It is relatively easy to compare predicted numbers and sizes of a given tumor type with actual observations. This could even be done in many replicate mice to explore the impact of chance and necessity in cancer evolution at a very fine-grained level. Such analyses are impossible in large-scale ecology, since 're-running' the world is impossible. Mouse cancer provides unique opportunities for learning about the dynamic aspects of LIFTs in a context where real-world checks are conceivable, as cancer cells are affected by dynamics in their environment (rhythms of mouse life; attempts to cure cancer, etc.) and in return affect their environment (eventually killing the mouse). This might eventually allow turning the dynamic MOCA-LIFT in **Figure 6** into a LIFT (Movie 2).

Ongoing discussions between evolutionary biology and cancer biology (Horne *et al.*, 2015) include work in evolutionary modeling (Nagy, 2005), ecology (Korolev *et al.*, 2014), investigations of life-history trade-offs (Aktipis *et al.*, 2013), evidence for positive selection (Crespi and Summers, 2006), evidence for the role of mutations during development (Frank, 2010), interactions between cancer and viruses (Brandon Ogbunugafor *et al.*, 2013), and epigenetics (Swanton and Beck, 2014).

EvoSysBio Milestone 5: Antibiotics Resistance Evolution

The first part of this milestone predicts how fast bacterial model-organism cultures evolve resistance to well-understood antibiotics under defined laboratory conditions without nested organisms, a challenge comparable to Milestone 3. Adding nested organisms increases difficulties to **the level of** Milestone 4 by requiring the specification of a TOB for the bacteria. However, in simple applied real-world contexts, it is no longer possible to repeat the 'world history' for improving the model

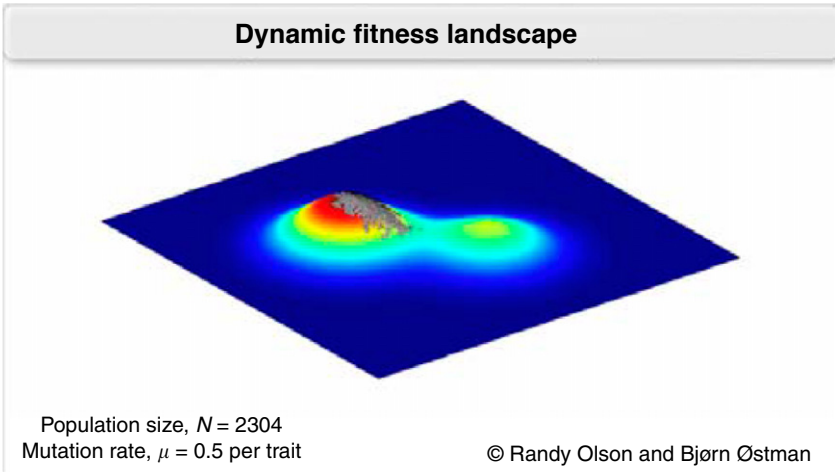

Figure 6 Fitness always depends on the environment. Frequent environmental changes can trigger perpetual changes in fitness, resulting in a constant need for adaptation. Dynamically changing Landscapes of Incomplete Fitness Traits pose additional challenges to EvoSysBio 'flight simulators,' which now also have to predict environmental properties and how they might be affected by populations of organisms over time. Here a snapshot is shown from a cartoonish movie of a population that evolves on a dynamic MOCA-LIFT (for more details, see also links in Section Relevant Websites). Picture credits: © Østman and Olson (2014c), reuse under CC-BY-SA 3.0.

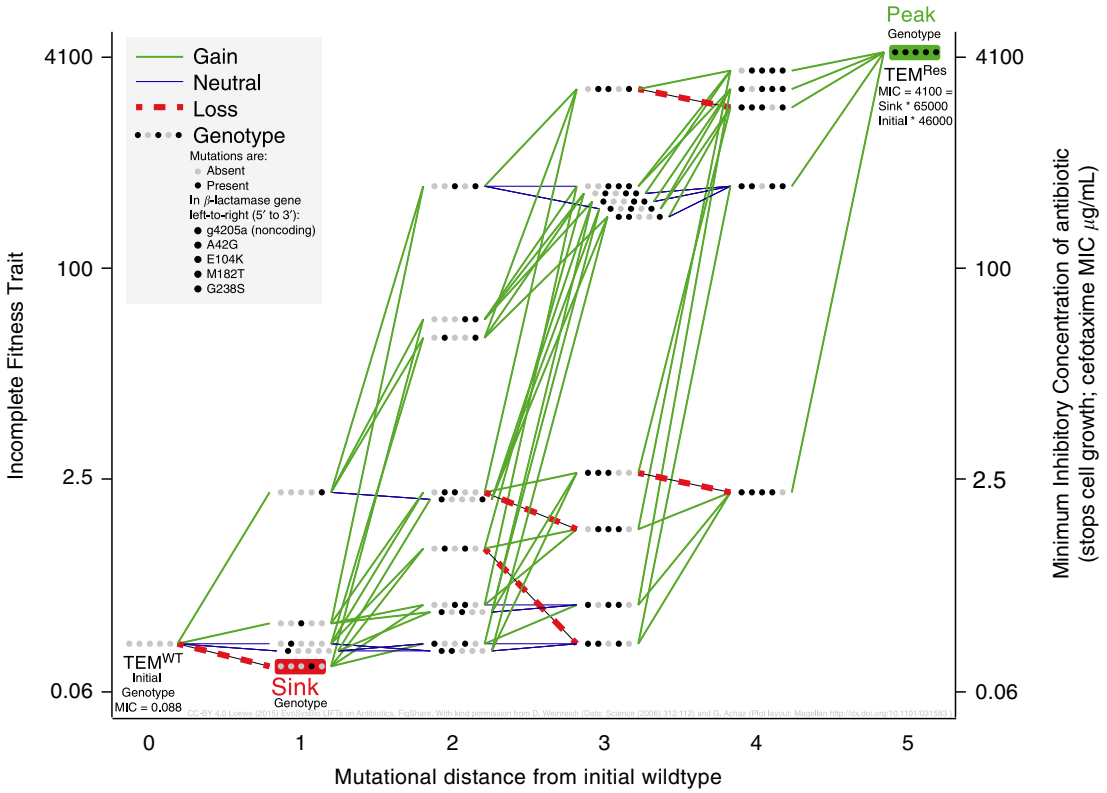

Figure 7 Evolution can take different paths to increase an IFT, such as in this experimentally observed LIFT snapshot of antibiotics resistance from five point mutations in a bacterial gene, where they can accumulate from the original TEM^{WT} genotype to the TEM^{Res} with different trajectories of fitness increases, which are followed with different probabilities as shown elsewhere (Weinreich *et al.*, 2006; DePristo *et al.*, 2007). This LIFT uses MIC, a cellular IFT that depends on a corresponding protein based LIFT where IFTs are determined by molecular features, which can be measured independently (Meini *et al.*, 2015). Picture credits: © Laurence Loewe (2015b), reusable under CC-BY 4.0. Data (Weinreich *et al.*, 2006) was plotted by the MAGELLAN fitness landscape mapping tool (Brouillet *et al.*, 2015), both shared with kind permission of Daniel Weinreich and the MAGELLAN authors.

(tumor evolution is repeatable by using new mice; but hospitals are difficult to sterilize, turning their resistance evolution into a historic process). Nested organisms, reservoirs, and very complex ecologies further complicate the picture by introducing many unknowns of potential importance. Unfortunately, these difficulties are matched by the urgency of the problem.

Prescribed doses of antibiotics have been climbing and now some bacterial strains are resistant to all known antibiotics. The evolution of antibiotic resistant bacteria has been discussed since antibiotics were first used (Neu, 1992; Normark and Normark, 2002; Choffnes *et al.*, 2010; Perros, 2015; Baker, 2015): How can we use these antimicrobial 'super-drugs' in a way that reduces the ability of 'superbugs' to evolve complete resistance? If we do not succeed, bacterial infections will become more deadly until one of the biggest medical advances of the twentieth century will have lost practical relevance. Predicting antibiotic resistance evolution (Martinez *et al.*, 2007) might allow us to find ways to use antibiotics that minimize the evolution of resistance.

Challenges of dynamic, multi-level simulations of evolution need to be mastered for truly understanding antibiotics resistance evolution. At the molecular level, questions include how many mutations a bacterial protein needs to accumulate to confer a higher level of resistance to the cell that produced it. The molecular features of such a protein can be interpreted as IFTs that define a protein-based LIFT (Meini *et al.*, 2015). The interplay of new mutations with the other content in crowded cells leads to an IFT at the cellular level: the Minimum Inhibitory Concentration (MIC) of an antibiotic above which a given bacterial cell is not longer able to grow. **Figure 7** shows the results of experimentally measuring a small, but combinatorially complete LIFT for an antibiotic-based IFT defined by a corresponding MIC (Weinreich *et al.*, 2006). At intermediate levels, events in organisms need to be modeled, since bacteria evolve during an infection. On geographic scales, the diverse use of antibiotics contributes to resistance evolution in many unexpected places (e.g., soil in agriculture, biofilms in hospitals). Despite progress, it is not clear how to optimize the use of antibiotics overall in hospitals and agriculture (where infected humans can carry resistant super-bugs between both, unwittingly exposing others). International travel adds further complications by moving infections over large distances. The challenges of both dynamic and static fitness landscapes (**Figures 6** and **3**) apply to antibiotics resistance evolution, as changes in antibiotics usage policies can easily generate either.

Complex evolutionary phenomena, such as epistasis, mutation rates in the stationary phase of bacteria and many other effects are important for understanding the evolution of antibiotics resistance (Loewe *et al.*, 2003; MacLean *et al.*, 2010; Hall and MacLean, 2011; Schenk and de Visser, 2013; DePristo *et al.*, 2007; Poelwijk *et al.*, 2007; Weinreich and Knies, 2013). For example, resistant bacteria often pay a fitness cost unless it is mitigated by compensatory mutations (Andersson, 2006). Unfortunately, antibiotics are often used in a way that facilitates the evolution of antibiotics resistance by creating environments where sublethal concentrations of antibiotics can select for resistant bacteria over longer periods of time (e.g., in sewers or agriculture; see **Figure 8**, (Gullberg *et al.*, 2011)). To preserve one of the biggest medical success stories of the twentieth century requires continually solving the riddles that bacteria pose by evolving antibiotics resistance (Choffnes *et al.*, 2010; Perros, 2015; Baker, 2015). Given that prokaryotes rule the world (by numbers and by speed of growth), it will require the best, most precise, and most integrated understanding of evolution to outsmart them – with the help of many computers.

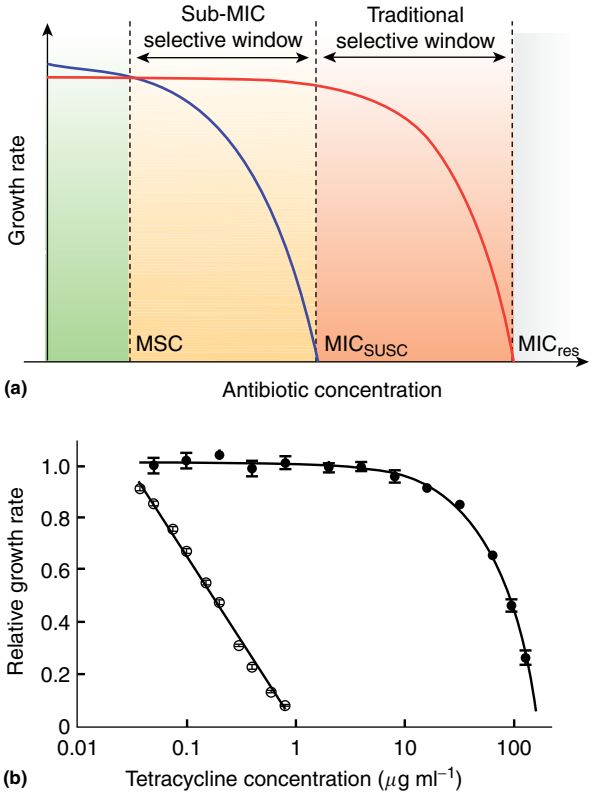

Figure 8 Small selective advantages can combine over a longer period of time to substantially accelerate antibiotics resistance evolution (Gullberg *et al.*, 2011). (a) Schematic overview, *MIC*: Minimal Inhibitory Concentration for susceptible (*susc*) and resistant (*res*) bacteria, *MSC*: Minimal Selective Concentration. (b) Experimental measurements for the antibiotic tetracycline. Picture credits: © Gullberg *et al.* (2011), reusable under CC-BY.

Conclusions

Even though the New Evolutionary Synthesis started in the 1920s and biological progress has continued at a phenomenal pace, much biology remains to be discovered and integrated. The New Synthesis is being renewed each day when researchers integrate new results into its framework through relentless synthesis (Wray *et al.*, 2014). The explosion of expertise on evolution and the complexity of systems science require reliable computational tools to give researchers a chance to keep up with the pace of the New Synthesis. Success critically depends on the strength of underpinning abstractions, which may be measured by intangibles such as conceptual clarity, completeness, simplicity, and elegance. The power of good abstractions is the best defense against the swamp of

complexity that otherwise mires researchers in endlessly redundant repeats of researching, reinventing, and rediscovering. There is nothing more practical than a good theory.

The ambitious aims of EvoSysBio for understanding and reliably predicting evolution depend on accelerating the pace of the relentless New Evolutionary Synthesis to which EvoSysBio is ultimately contributing. This can be done by making it easier to connect the five fundamental factors of evolution to results from models of complex Intra-Organism and Trans-Organism Biology. Toward this end, this article provides a high-level overview of FCNets and LIFTs, but many details remain to be worked out.

The relentless New Synthesis can benefit from lessons learned by programmers in their struggles in the swamp of complexity: aim for 'as simple as possible, but not simpler' (Raymond, 2003). To make EvoSysBio efficient, this view of Occam's Razor needs to inspire new computational approaches for knowledge organization, provenance, modeling, reliability, precise uncertainty quantification, ease of automation, abstraction management, automatic testing, efficient debugging, and reproducibility. Without such innovations, EvoSysBio researchers will either get stuck in the complexity swamp of constructing FCNets (and debugging their dependencies) or get lost in the fog of finding signals drowned by noise (due to missing uncertainty quantification).

Biology's complexity is so pervasive that efficient approaches for solving these problems often require semantic architectures with the strength of general programming languages that can integrate most biological research, if not all. Developing such general semantics is usually slow and challenging, but enables great leaps forward once the right concepts are ready for automation. Given the slow pace of developing semantics and the complexity of evolution, it may take ≥ 30 years until integrated computational pipelines can enable the use of all state-of-the-art expertise to mechanistically predict the effects of unknown mutations in well-studied model organisms. Analyzing resulting fitness landscapes with interactive 'multidimensional flight simulators' may sound like science fiction today – perhaps like 'population genomics' would have sounded in the 1970s. Yet it took only about 30 years to go from a method of DNA sequencing to population genomics. Maybe it is possible to start transitioning from a formal definition of EvoSysBio to usable flight simulators for fitness landscapes as the New Evolutionary Synthesis approaches its 100th year of relentless integration service. A bit of strategic long-term thinking might help to avoid swamps and fogs to make evolutionary research more efficient in the long run.

See also: Adaptive Landscapes. Evolutionary Medicine I. An Overview and Applications to Cancer. Gene Interactions in Evolution. Modularity and Integration. Modularity and Integration in Evo-Devo

References

Adami, C., 1998. Introduction to Artificial Life. New York, NY: Springer.

Adams, D., 1979. The Hitchhiker's Guide to the Galaxy. London: Pan Macmillan Ltd.

Aktipis, C.A., Boddy, A.M., Gatenby, R.A., *et al.*, 2013. Life history trade-offs in cancer evolution. Nature Reviews Cancer 13 (12), 883–892.

Andersson, D.I., 2006. The biological cost of mutational antibiotic resistance: Any practical conclusions? Current Opinion in Microbiology 9 (5), 461–465.

Anholt, R.R., Mackay, T.F., 2015. Dissecting the genetic architecture of behavior in *Drosophila*. Current Opinion in Behavioral Sciences 2, 1–7.

Aristotle (− 344 + 22). Metaphysics, Book 8, part 6 (W.D. Ross, Trans). The Internet Classics Archive: MIT. Available at: http://classics.mit.edu/Aristotle/metaphysics.8.viii.html (accessed 22.01.16).

Aronson, S.J., Rehm, H.L., 2015. Building the foundation for genomics in precision medicine. Nature 526 (7573), 336–342.

Ayroles, J.F., Carbone, M.A., Stone, E.A., *et al.*, 2009. Systems genetics of complex traits in *Drosophila melanogaster*. Nature Genetics 41 (3), 299–307.

Baker, S., 2015. Infectious disease. A return to the pre-antimicrobial era? Science 347 (6226), 1064–1066.

Bergman, A., Siegal, M.L., 2003. Evolutionary capacitance as a general feature of complex gene networks. Nature 424 (6948), 549–552.

Bershtein, S., Segal, M., Bekerman, R., *et al.*, 2006. Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. Nature 444 (7121), 929–932.

Biebricher, C.K., Eigen, M., 2006. What is a quasispecies? Current Topics in Microbiology and Immunology 299, 1–31.

Bohl, K., Hummert, S., Werner, S., *et al.*, 2014. Evolutionary game theory: Molecules as players. Molecular BioSystems 10 (12), 3066–3074.

Box, G.E.P., Draper, N.R., 2007. Response Surfaces, Mixtures, and Ridge Analyses, second ed. Hoboken, NJ: John Wiley.

Brandon Ogbunugafor, C., Alto, B.W., Overton, T.M., *et al.*, 2013. Evolution of increased survival in RNA viruses specialized on cancer-derived cells. American Naturalist 181 (5), 585–595.

Brommer, J.E., 2000. The evolution of fitness in life-history theory. Biological Reviews of the Cambridge Philosophical Society 75 (3), 377–404.

Brouillet, S., Annoni, H., Ferretti, L., Achaz, G., 2015. Magellan: A tool to explore small fitness landscapes. BioRxiv. Available at: http://wwwabi.snv.jussieu.fr/public/Magellan/ (accessed 27.01.16). doi:http://dx.doi.org/10.1101/031583.

Caetano-Anolles, G., 2010. Evolutionary Genomics and Systems Biology. Hoboken, NJ: Wiley-Blackwell.

Calvert, J., 2012. Systems biology, synthetic biology and data-driven research: A commentary on Krohs, Callebaut, and O'Malley and Soyer. Studies in History and Philosophy of Biological and Biomedical Sciences 43 (1), 81–84.

Carroll, S.B., 2008. Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. Cell 134 (1), 25–36.

Charlesworth, B., Betancourt, A.J., Kaiser, V.B., Gordo, I., 2009. Genetic recombination and molecular evolution. Cold Spring Harbor Symposia on Quantitative Biology 74, 177–186.

Charlesworth, B., Charlesworth, D., 2010. Elements of Evolutionary Genetics. Greenwood Village, CO: Roberts and Company Publishers.

Choffnes, E.R., Relman, D.A., Mack, A., 2010. Antibiotic Resistance. Implications for Global Health and Novel Intervention Strategies: Workshop Summary. Washington, DC: National Academies Press. Available at: http://www.ncbi.nlm.nih.gov/books/NBK54255 (accessed 22.01.16).

Cornish-Bowden, A., 2006. Putting the systems back into systems biology. Perspectives in Biology and Medicine 49 (4), 475–489.

Crespi, B.J., Summers, K., 2006. Positive selection in the evolution of cancer. Biological Reviews of the Cambridge Philosophical Society 81 (3), 407–424.

Crow, J.F., Kimura, M., 1970, 2009. An Introduction to Population Genetics Theory. Caldwell, NJ: Blackburn Press.

Dean, A.M., Thornton, J.W., 2007. Mechanistic approaches to the study of evolution: The functional synthesis. Nature Reviews Genetics 8 (9), 675–688.

De Jong, K.A., 2006. Evolutionary Computation: A Unified Approach. Cambridge, MA: MIT Press.

DePristo, M.A., Hartl, D.L., Weinreich, D.M., 2007. Mutational reversions during adaptive protein evolution. Molecular Biology and Evolution 24 (8), 1608–1610.

Dietze, M.C., Lebauer, D.S., Kooper, R., 2013. On improving the communication between models and data. Plant, Cell & Environment 36 (9), 1575–1585.

Dobzhansky, T., 1973. Nothing in biology makes sense except in the light of evolution. American Biology Teacher 35, 125–129. Available at: http://people.delphiforums.com/lordorman/light.htm (accessed 17.11.15).

Dodd, A.N., Salathia, N., Hall, A., *et al.*, 2005. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. Science 309 (5734), 630–633.

E, W., 2011. Principles of Multiscale Modeling. Cambridge; New York, NY: Cambridge University Press.

Editorial, 2015. Big data meets mechanism. 2015 Nature Medicine 21(7), 673.

Edwards, J.S., Ibarra, R.U., Palsson, B.O., 2001. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. Nature Biotechnology 19 (2), 125–130.

von Ehrenfels, C., 1890. Über gestaltqualitäten. Vierteljahrsschrift für wissenschaftliche Philosophie 17, 249–292.

Endy, D., You, L., Yin, J., Molineux, I.J., 2000. Computation, prediction, and experimental tests of fitness for bacteriophage T7 mutants with permuted genomes. Proceedings of the National Academy of Sciences of the United States of America 97 (10), 5375–5380.

Eyre-Walker, A., Keightley, P.D., 2007. The distribution of fitness effects of new mutations. Nature Reviews Genetics 8 (8), 610–618.

Felsenstein, J., 2004. Inferring Phylogenies. Sunderland, MA: Sinauer Associates.

Fisher, R.A., 1918. The correlation between relatives on the supposition of mendelian inheritance. Transactions of the Royal Society of Edinburgh 52, 399–433.

Frank, S.A., 2010. Evolution in health and medicine sackler colloquium: Somatic evolutionary genomics: Mutations during development cause highly variable genetic mosaicism with risk of cancer and neurodegeneration. Proceedings of the National Academy of Sciences of the United States of America 107 (Suppl. 1), 1725–1730.

Gardner, A., 2015. The genetical theory of multilevel selection. Journal of Evolutionary Biology 28 (2), 305–319.

Gardner, A., West, S.A., 2014. Inclusive fitness: 50 years on. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 369 (1642), 20130356.

Gavrilets, S., 1997. Evolution and speciation on holey adaptive landscapes. Trends in Ecology & Evolution 12 (8), 307–312.

Gavrilets, S., 2004. Fitness Landscapes and the Origin of Species. Monographs in Population Biology, vol. 41, 2004 ed. Princeton, NJ: Princeton University Press.

Gerlinger, M., McGranahan, N., Dewhurst, S.M., *et al*., 2014. Cancer: Evolution within a lifetime. Annual Review of Genetics 48, 215–236.

Gillespie, D.T., 2007. Stochastic simulation of chemical kinetics. Annual Review of Physical Chemistry 58, 35–55.

Gillespie, D.T., Hellander, A., Petzold, L.R., 2013. Perspective: Stochastic algorithms for chemical kinetics. Journal of Chemical Physics 138 (17), 170901.

Gravner, J., Pitman, D., Gavrilets, S., 2007. Percolation on fitness landscapes: Effects of correlation, phenotype, and incompatibilities. Journal of Theoretical Biology 248 (4), 627–645.

Grimm, V., Railsback, S.F., 2005. Individual-based modeling and ecology. Princeton Series in Theoretical and Computational Biology. Princeton, NJ: Princeton University Press.

Gullberg, E., Cao, S., Berg, O.G., *et al*., 2011. Selection of resistant bacteria at very low antibiotic concentrations. PLOS Pathogens 7 (7), e1002158.

Gutenkunst, R.N., Waterfall, J.J., Casey, F.P., *et al*., 2007. Universally sloppy parameter sensitivities in systems biology models. PLOS Computational Biology 3 (10), 1871–1878.

Haldane, J.B., 1964. A defense of beanbag genetics. Perspectives in Biology and Medicine 19, 343–359.

Hall, A.R., MacLean, R.C., 2011. Epistasis buffers the fitness effects of rifampicin-resistance mutations in *Pseudomonas aeruginosa*. Evolution 65 (8), 2370–2379.

Hallgrimsson, B., Hall, B.K., 2011. Epigenetics: Linking Genotype and Phenotype in Development and Evolution. Berkeley, CA: University of California Press.

Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: The next generation. Cell 144 (5), 646–674.

Harbison, S.T., Mackay, T.F., Anholt, R.R., 2009. Understanding the neurogenetics of sleep: Progress from *Drosophila*. Trends in Genetics 25 (6), 262–269.

Hood, L., Heath, J.R., Phelps, M.E., Lin, B., 2004. Systems biology and new technologies enable predictive and preventative medicine. Science 306 (5696), 640–643.

Horne, S.D., Pollick, S.A., Heng, H.H.Q., 2015. Evolutionary mechanism unifies the hallmarks of cancer. International Journal of Cancer 136 (9), 2012–2021.

Houle, D., Pelabon, C., Wagner, G.P., Hansen, T.F., 2011. Measurement and meaning in biology. Quarterly Review of Biology 86 (1), 3–34.

Huang, W., Massouras, A., Inoue, Y., *et al*., 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* genetic reference panel lines. Genome Research 24 (7), 1193–1208.

Huang, W., Perez-Garcia, P., Pokhilko, A., *et al*., 2012. Mapping the core of the *Arabidopsis* circadian clock defines the network structure of the oscillator. Science 336 (6077), 75–79.

Hughes, K.A., 2010. Mutation and the evolution of ageing: From biometrics to systems genetics. Philosophical Transactions of the Royal Society B: Biological Sciences 365 (1544), 1273–1279.

Hummert, S., Bohl, K., Basanta, D., *et al*., 2014. Evolutionary game theory: Cells as players. Molecular BioSystems 10 (12), 3044–3065.

Ibarra, R.U., Edwards, J.S., Palsson, B.O., 2002. *Escherichia coli* k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. Nature 420 (6912), 186–189.

Ideker, T., Galitski, T., Hood, L., 2001a. A new approach to decoding life: Systems biology. Annual Review of Genomics and Human Genetics 2, 343–372.

Ideker, T., Thorsson, V., Ranish, J.A., *et al*., 2001b. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 292 (5518), 929–934.

Iyengar, R., Altman, R.B., Troyanskya, O., FitzGerald, G.A., 2015. Medicine: Personalization in practice. Science 350 (6258), 282–283.

James, D., Wilkins-Diehr, N., Stodden, V., *et al*., 2015. Standing together for reproducibility in large-scale computing: Report on reproducibility@xsede. *Arxiv*: Distributed, Parallel, and Cluster Computing (cs.DC). Available at: http://arxiv.org/abs/1412.5557 (accessed 17.11.15).

Kacser, H., Burns, J.A., 1981. The molecular basis of dominance. Genetics 97, 639–666.

Kaplan, J., 2008. The end of the adaptive landscape metaphor? Biology and Philosophy 23, 625–638.

Karr, J.R., Sanghvi, J.C., Macklin, D.N., *et al*., 2012. A whole-cell computational model predicts phenotype from genotype. Cell 150 (2), 389–401.

Karr, J.R., Williams, A.H., Zucker, J.D., *et al*., 2015. Summary of the dream8 parameter estimation challenge: Toward parameter identification for whole-cell models. PLOS Computational Biology 11 (5), e1004096.

Keightley, P.D., 1996. Metabolic models of selection response. Journal of Theoretical Biology 182 (3), 311–316.

Kimura, M., 1983. The Neutral Theory of Molecular Evolution. Cambridge: Cambridge University Press.

King, Z.A., Lloyd, C.J., Feist, A.M., Palsson, B.O., 2015. Next-generation genome-scale models for metabolic engineering. Current Opinion in Biotechnology 35C, 23–29.

Kirk, P.D., Babtie, A.C., Stumpf, M.P., 2015. Systems biology (un)certainties. Science 350 (6259), 386–388.

Kirkpatrick, M., Johnson, T., Barton, N., 2002. General models of multilocus evolution. Genetics 161 (4), 1727–1750.

Kitano, H., 2002a. Computational systems biology. Nature 420 (6912), 206–210.

Kitano, H., 2002b. Systems biology: A brief overview. Science 295 (5560), 1662–1664.

Klipp, E., Liebermeister, W., Wierling, C., *et al*., 2009. Systems Biology: A Textbook. Weinheim: Wiley-Blackwell-VCH.

Korolev, K.S., Xavier, J.B., Gore, J., 2014. Turning ecology and evolution against cancer. Nature Reviews Cancer 14 (5), 371–380.

Kouyos, R.D., Leventhal, G.E., Hinkley, T., *et al*., 2012. Exploring the complexity of the HIV-1 fitness landscape. PLOS Genetics 8 (3), e1002551.

Kuhlbrandt, W., 2014. Biochemistry. The resolution revolution. Science 343 (6178), 1443–1444.

Laland, K., Uller, T., Feldman, M., *et al*., 2014. Does evolutionary theory need a rethink? Nature 514 (7521), 161–164.

Law, A.M., Kelton, W.D., 2000. Simulation Modeling and Analysis. Mcgraw-Hill Series in Industrial Engineering and Management Science, third ed. Boston, MA: McGraw-Hill.

Li, J., Browning, S., Mahal, S.P., *et al*., 2010. Darwinian evolution of prions in cell culture. Science 327 (5967), 869–872.

Lim, K.I., Yin, J., 2009. Computational fitness landscape for all gene-order permutations of an RNA virus. PLOS Computational Biology 5 (2), e1000283.

Linnen, C.R., Kingsley, E.P., Jensen, J.D., Hoekstra, H.E., 2009. On the origin and spread of an adaptive allele in deer mice. Science 325 (5944), 1095–1098.

Loewe, L., 2009. A framework for evolutionary systems biology. BMC Systems Biology 3, 27.

Loewe, L., 2012. How evolutionary systems biology will help understand adaptive landscapes and distributions of mutational effects. In: Soyer, O. (Ed.), Evolutionary Systems Biology. Advances in Experimental Medicine and Biology, vol. 751. New York, NY: Springer, pp. 399–410.

Loewe, L., 2015a. EvoSysBio in 10 slides, FigShare 1427128. Available at: http://dx.doi.org/10.6084/m9.figshare.1427128 (accessed 22.01.16), usable under Creative Commons Attribution 4.0 International (https://creativecommons.org/licenses/by/4.0/).

Loewe, L., 2015b. EvoSysBio lifts: Microbes and antibiotics, FigShare 2056428. Available at: https://dx.doi.org/10.6084/m9.figshare.2056428 (accessed 22.01.16), usable under Creative Commons Attribution 4.0 International (https://creativecommons.org/licenses/by/4.0/).

Loewe, L., Charlesworth, B., 2006. Inferring the distribution of mutational effects on fitness in *Drosophila*. Biology Letters 2 (3), 426–430.

Loewe, L., Hill, W.G., 2010a. Introduction: The population genetics of mutations: Good, bad and indifferent. Philosophical Transactions of the Royal Society of London B: Biological Sciences 365 (1544), 1153–1167.
Loewe, L., Hill, W.G., 2010b. The population genetics of mutations: Good, bad and indifferent – Papers of a theme issue in honor of Brian Charlesworth. Philosophical Transactions of the Royal Society of London B: Biological Sciences 365 (1544), 1149–1294.
Loewe, L., Hillston, J., 2008. The distribution of mutational effects on fitness in a simple circadian clock. Lecture Notes in Bioinformatics 5307, 156–175.
Loewe, L., Papp, B., Lercher, M., *et al.*, 2015–2009. Symposia and workshops in evolutionary systems biology. Available at: http://evolutionarysystemsbiology.org/meeting/ (accessed 22.01.16).
Loewe, L., Textor, V., Scherer, S., 2003. High deleterious genomic mutation rate in stationary phase of *Escherichia coli*. Science 302 (5650), 1558–1560.
Lynch, M., 2007. The Origins of Genome Architecture. Sunderland, MA: Sinauer Associates.
Lynch, M., Field, M.C., Goodson, H.V., *et al.*, 2014. Evolutionary cell biology: Two origins, one objective. Proceedings of the National Academy of Sciences of the United States of America 111 (48), 16990–16994.
Mackay, T.F., Stone, E.A., Ayroles, J.F., 2009. The genetics of quantitative traits: Challenges and prospects. Nature Reviews Genetics 10 (8), 565–577.
Mackechnie, C., Maskell, L., Norton, L., Roy, D., 2011. The role of 'big society' in monitoring the state of the natural environment. Journal of Environmental Monitoring 13 (10), 2687–2691.
Macklin, D.N., Ruggero, N.A., Covert, M.W., 2014. The future of whole-cell modeling. Current Opinion in Biotechnology 28, 111–115.
MacLean, R.C., Perron, G.G., Gardner, A., 2010. Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in *Pseudomonas aeruginosa*. Genetics 186 (4), 1345–1354.
Markowetz, F., Boutros, M., 2015. Systems Genetics : Linking Genotypes and Phenotypes. New York, NY: Cambridge University Press.
Martin, L.L., Unrau, P.J., Muller, U.F., 2015. RNA synthesis by in vitro selected ribozymes for recreating an RNA world. Life (Basel) 5 (1), 247–268.
Martinez, J.L., Baquero, F., Andersson, D.I., 2007. Predicting antibiotic resistance. Nature Reviews Microbiology 5 (12), 958–965.
McCandlish, D.M., 2011. Visualizing fitness landscapes. Evolution 65 (6), 1544–1558.
Medina, M., 2005. Genomes, phylogeny, and evolutionary systems biology. Proceedings of the National Academy of Sciences of the United States of America 102 (Suppl. 1), 6630–6635.
Meini, M.R., Tomatis, P.E., Weinreich, D.M., Vila, A.J., 2015. Quantitative description of a protein fitness landscape based on molecular features. Molecular Biology and Evolution 32 (7), 1774–1787.
Michel, J.B., Shen, Y.K., Aiden, A.P., *et al.*, 2011. Quantitative analysis of culture using millions of digitized books. Science 331 (6014), 176–182.
Mustonen, V., Lassig, M., 2009. From fitness landscapes to seascapes: Non-equilibrium dynamics of selection and adaptation. Trends in Genetics 25 (3), 111–119.
Nagy, J.D., 2005. The ecology and evolutionary biology of cancer: A review of mathematical models of necrosis and tumor cell diversity. Mathematical Biosciences and Engineering 2 (2), 381–418.
National Research Council (USA), 2009. A New Biology for the 21st Century: Ensuring the United States Leads the Coming Biology Revolution. Washington, DC: National Academies Press. Available at: http://www.ncbi.nlm.nih.gov/books/NBK32509 (accessed 22.01.16).
Neu, H.C., 1992. The crisis in antibiotic resistance. Science 257, 1064–1073.
Noble, D., 2011. Successes and failures in modeling heart cell electrophysiology. Heart Rhythm 8 (11), 1798–1803.
Normark, B.H., Normark, S., 2002. Evolution and spread of antibiotic resistance. Journal of Internal Medicine 252 (2), 91–106.
Nowak, M.A., Sigmund, K., 2004. Evolutionary dynamics of biological games. Science 303 (5659), 793–799.
O'Malley, M.A., 2012. Evolutionary systems biology: Historical and philosophical perspectives on an emerging synthesis. Advances in Experimental Medicine and Biology 751, 1–28.
O'Malley, M.A., Soyer, O.S., 2012. The roles of integration in molecular systems biology. Studies in History and Philosophy of Biological and Biomedical Sciences 43 (1), 58–68.
O'Malley, M.A., Soyer, O.S., Siegal, M.L., 2015. A philosophical perspective on evolutionary systems biology. Biological Theory. doi:10.1007/s13752-015-0202-6.
Okasha, S., 2006. Evolution and the Levels of Selection. Oxford: Oxford University Press.
Oreskes, N., Shrader-Frechette, K., Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science 263 (5147), 641–646.
Orr, H.A., 2005. The genetic theory of adaptation: A brief history. Nature Review Genetics 6 (2), 119–127.
Østman, B., Olson, R., 2014a. Using fitness landscapes to visualize evolution in action. Available at: https://en.wikipedia.org/wiki/File:Visualization_of_a_population_evolving_in_a_static_fitness_landscape.gif (usable under Creative Commons Attribution-ShareAlike 3.0 International (https://creativecommons.org/licenses/by-sa/3.0/)); http://beacon-center.org/blog/2014/2004/2009/using-fitness-landscapes-to-visualize-evolution-in-action/ (accessed 22.01.16).
Østman, B., Olson, R., 2014c. Visualizing coevolution in dynamic fitness landscapes. Available at: https://commons.wikimedia.org/wiki/File:Visualization_of_a_population_evolving_in_a_dynamic_fitness_landscape.gif (usable under Creative Commons Attribution-ShareAlike 3.0 International (https://creativecommons.org/licenses/by-sa/3.0/)); http://beacon-center.org/blog/2014/2006/2011/visualizing-coevolution-in-dynamic-fitness-landscapes/ (accessed 22.01.16).
Pacheco, J.M., Santos, F.C., Dingli, D., 2014. The ecology of cancer from an evolutionary game theory perspective. Interface Focus 4 (4), 20140019.
Papp, B., Szappanos, B., Notebaart, R.A., 2011. Use of genome-scale metabolic models in evolutionary systems biology. Methods in Molecular Biology 759, 483–497.
Parysow, P., Gertner, G., Westervelt, J., 2000. Efficient approximation for building error budgets for process models. Ecological Modelling 135 (2–3), 111–125.
Pavlicev, M., Wagner, G.P., 2015. Evolutionary systems biology: Shifting focus to the context-dependency of genetic effects. In: Martin, L.B., Ghalambor, C.K., Woods, H.A. (Eds.), Integrative Organismal Biology. Hoboken, NJ: Wiley Blackwell, pp. 91–108.
Payne, J.L., Wagner, A., 2014. The robustness and evolvability of transcription factor binding sites. Science 343 (6173), 875–877.
Pelletier, F., Garant, D., Hendry, A.P., 2009. Eco-evolutionary dynamics. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 364 (1523), 1483–1489.
Perros, M., 2015. Infectious disease: A sustainable model for antibiotics. Science 347 (6226), 1062–1064.
Phillips, P.C., 2008. Epistasis – The essential role of gene interactions in the structure and evolution of genetic systems. Nature Reviews Genetics 9 (11), 855–867.
Poelwijk, F.J., Kiviet, D.J., Weinreich, D.M., Tans, S.J., 2007. Empirical fitness landscapes reveal accessible evolutionary paths. Nature 445 (7126), 383–386.
Raymond, E.S., 2003. The Art of Unix Programming. Addison-Wesley. Available at: http://www.catb.org/~esr/writings/taoup/ (accessed 16.11.15).
Richter, H., Engelbrecht, A.P. (Eds.), 2014. Recent Advances in the Theory and Application of Fitness Landscapes. Berlin: Springer.
Schenk, M.F., de Visser, J.A., 2013. Predicting the evolution of antibiotic resistance. BMC Biology 11, 14.
Schenk, M.F., Szendro, I.G., Krug, J., de Visser, J.A., 2012. Quantifying the adaptive potential of an antibiotic resistance enzyme. PLOS Genetics 8 (6), e1002783.
Schuster, P., 2012. A revival of the landscape paradigm: Large scale data harvesting provides access to fitness landscapes. Complexity 17 (5), 6–10.
Seifert, D., Di Giallonardo, F., Metzner, K.J., *et al.*, 2015. A framework for inferring fitness landscapes of patient-derived viruses using quasispecies theory. Genetics 199 (1), 191–203.
Seksik, P., Landman, C., 2015. Understanding microbiome data: A primer for clinicians. Digestive Diseases 33 (Suppl. 1), 11–16.
Singh, R.S., Krimbas, C.B., 2000. Evolutionary Genetics : From Molecules to Morphology. Cambridge; New York, NY: Cambridge University Press.
Sleeman, J.M., 2013. Has the time come for big science in wildlife health? Ecohealth 10 (4), 335–338.
Soyer, O.S. (Ed.), 2012. Evolutionary Systems Biology. Advances in Experimental Medicine and Biology, vol. 751. New York, NY: Springer.
Soyer, O.S., O'Malley, M.A., 2013. Evolutionary systems biology: What it is and why it matters. Bioessays 35 (8), 696–705.
Stearns, S.C., 1992. The Evolution of Life Histories. Oxford: Oxford University Press.
Stearns, S.C., 2000. Life history evolution: Successes, limitations, and prospects. Naturwissenschaften 87 (11), 476–486.
Stearns, S.C., Koella, J.C., 2008. Evolution in Health and Disease. Oxford biology, second ed. Oxford; New York, NY: Oxford University Press.
Stearns, S.C., Magwene, P., American Society of Naturalists, 2003. The naturalist in a world of genomics. American Naturalist 161 (2), 171–180.
Stodden, V., 2015. Reproducing statistical results. Annual Review of Statistics and Its Application 2, 1–19.

Svensson, E.I., Calsbeek, R., 2012. The Adaptive Landscape in Evolutionary Biology, first ed. Oxford; New York, NY: Oxford University Press.
Swanton, C., Beck, S., 2014. Epigenetic noise fuels cancer evolution. Cancer Cell 26 (6), 775–776.
Thiele, I., Fleming, R.M., Que, R., *et al.*, 2012. Multiscale modeling of metabolism and macromolecular synthesis in *E. coli* and its application to the evolution of codon usage. PLOS One 7 (9), e45635.
Voit, E.O., 2013. A First Course in Systems Biology. New York, NY: Garland Science.
Wagemans, J., Elder, J.H., Kubovy, M., *et al.*, 2012a. A century of gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. Psychological Bulletin 138 (6), 1172–1217.
Wagemans, J., Feldman, J., Gepshtein, S., *et al.*, 2012b. A century of gestalt psychology in visual perception: II. Conceptual and theoretical foundations. Psychological Bulletin 138 (6), 1218–1252.
Wagner, A., 2005. Robustness and Evolvability in Living Systems. Princeton Studies in Complexity. Princeton, NJ: Princeton University Press.
Wagner, A., 2011. The molecular origins of evolutionary innovations. Trends in Genetics 27 (10), 397–410.
Wagner, A., 2012. The role of robustness in phenotypic adaptation and innovation. Proceedings of the Royal Society B: Biological Sciences 279 (1732), 1249–1258.
Wagner, A., 2014. Mutational robustness accelerates the origin of novel RNA phenotypes through phenotypic plasticity. Biophysical Journal 106 (4), 955–965.
Wagner, G.P., Pavlicev, M., Cheverud, J.M., 2007. The road to modularity. Nature Reviews Genetics 8 (12), 921–931.
Walhout, A.J.M., Vidal, M., Dekker, J., 2013. Handbook of Systems Biology: Concepts and Insights, first ed. London: Waltham Academic Press.
Watt, W.B., 1985. Bioenergetics and evolutionary genetics — Opportunities for new synthesis. American Naturalist 125 (1), 118–143.
Weinreich, D.M., Delaney, N.F., Depristo, M.A., Hartl, D.L., 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science 312 (5770), 111–114.
Weinreich, D.M., Knies, J.L., 2013. Fisher's geometric model of adaptation meets the functional synthesis: Data on pairwise epistasis for fitness yields insights into the shape and size of phenotype space. Evolution 67 (10), 2957–2972.
Weinreich, D.M., Lan, Y., Wylie, C.S., Heckendorn, R.B., 2013. Should evolutionary geneticists worry about higher-order epistasis? Current Opinion in Genetics & Development 23 (6), 700–707.
Westerhoff, H.V., Palsson, B.O., 2004. The evolution of molecular biology into systems biology. Nature Biotechnology 22 (10), 1249–1252.
Westervelt, J.D., Cohen, G.L., 2012. Ecologist-Developed Spatially-Explicit Dynamic Landscape Models. Modeling Dynamic Systems. New York, NY: Springer.
Wolf, J.B., Brodie III, E.D., Wade, M.J. (Eds.), 2000. Epistasis and the Evolutionary Process. New York, NY: Oxford University Press.
Wolkenhauer, O., 2001. Systems biology: The reincarnation of systems theory applied in biology? Briefings in Bioinformatics 2 (3), 258–270.
Wooley, J.C., Godzik, A., Friedberg, I., 2010. A primer on metagenomics. PLOS Computational Biology 6 (2), e1000667.
Wray, G.A., Hoekstra, H.E., Futuyma, D.J., *et al.*, 2014. Does evolutionary theory need a rethink? Nature 514 (7521), 161–164.
Wright, S., 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proceedings of the Sixth International Congress of Genetics 1, 356–366.
Xu, M., Rhee, S.Y., 2014. Becoming data-savvy in a big-data world. Trends in Plant Sciences 19 (10), 619–622.
Zeigler, B.P., 2012. Guide to Modeling and Simulation of Systems of Systems. Simulation Foundations, Methods and Applications, first ed. New York, NY: Springer.
Zeigler, B.P., Praehofer, H., Kim, T.G., 2000. Theory of Modeling and Simulation : Integrating Discrete Event and Continuous Complex Dynamic Systems, second ed. San Diego, CA: Academic Press.

Further Reading

Klipp, E., Liebermeister, W., Wierling, C., *et al.*, 2009. Systems Biology: A Textbook. Weinheim: Wiley-Blackwell-VCH.
Loewe, L., 2009. A framework for evolutionary systems biology. BMC Systems Biology 3, 27.
Soyer, O.S. (Ed.), 2012. Evolutionary Systems Biology. Advances in Experimental Medicine and Biology, vol. 751. New York, NY: Springer.
Østman, B., Olson, R., 2014b. Various movies visualizing evolution in action on abstract fitness landscapes and related blog posts. Available at: http://pleiotropy.fieldofscience.com/2014/06/the-simulations-behind-fitness.html; http://pleiotropy.fieldofscience.com/2014/03/evolutionary-dynamics-in-holey-fitness.html; http://pleiotropy.fieldofscience.com/2014/06/video-visualizing-coevolution-in.html; http://spartanideas.msu.edu/2013/11/13/smooth-and-rugged-fitness-landscapes/; http://bjornostman.com/#landscapes; http://www.randalolson.com/2014/04/17/visualizing-evolution-in-action-dynamic-fitness-landscapes/; https://en.wikipedia.org/wiki/File:Visualization_of_a_population_evolving_in_a_static_fitness_landscape.gif; https://commons.wikimedia.org/wiki/File:Visualization_of_a_population_evolving_in_a_dynamic_fitness_landscape.gif (accessed 22.01.16).

Relevant Websites

http://bionumbers.hms.harvard.edu
B10NUMB3R5: The Database of Useful Biological Numbers.
http://beacon-center.org/blog/tag/fitness-landscapes/
BEACON, an NSF Center for the Study of Evolution in Action: Fitness Landscapes.
http://beacon-center.org/blog/2014/04/09/using-fitness-landscapes-to-visualize-evolution-in-action/
BEACON, an NSF Center for the Study of Evolution in Action: Using Fitness Landscapes to Visualize Evolution in Action.
http://beacon-center.org/blog/2014/06/11/visualizing-coevolution-in-dynamic-fitness-landscapes/
BEACON, an NSF Center for the Study of Evolution in Action: Visualizing Coevolution in Dynamic Fitness Landscapes.
http://brenda-enzymes.org
BRENDA: The Comprehensive Enzyme Information System.
http://dx.doi.org/10.1101/031583
Cold Spring Harbor Laboratory: bioRxiv.
http://EvolutionarySystemsBiology.org
Evolutionary Systems Biology Home.
https://dx.doi.org/10.6084/m9.figshare.1427128
EvoSysBio in 10 Slides.
http://bjornostman.com/#landscapes
Fitness Landscape Visualization.
http://www.hpcuniversity.org
HPC University: High-Performance Computing.
http://wwwabi.snv.jussieu.fr/public/magellan/Magellan.main.html
Magellan: Maps of Genetical Landscapes.
http://spartanideas.msu.edu/2013/11/13/smooth-and-rugged-fitness-landscapes/
Michigian State University: Spartan Ideas.
http://neos-server.org
NEOS Server: State-of-the-Art Solvers for Numerical Optimization.
http://pleiotropy.fieldofscience.com/2014/03/evolutionary-dynamics-in-holey-fitness.html
Pleiotropy: Evolutionary Dynamics in Holey Fitness Landscapes.
http://pleiotropy.fieldofscience.com/2014/06/the-simulations-behind-fitness.html
Pleiotropy: The Simulations Behind the Fitness Landscape Visualizations.
http://pleiotropy.fieldofscience.com/2014/06/video-visualizing-coevolution-in.html
Pleiotropy: Visualizing Coevolution in Dynamic Fitness Landscapes.
http://www.randalolson.com/2014/2004/2017/visualizing-evolution-in-action-dynamic-fitness-landscapes/
Randal Olson: Visualizing Evolution in Action: Dynamic Fitness Landscapes.
http://www.milosjohnson.com/Scapes/Weinreich/
Scapes: Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins.
http://virtualrat.org
The Virtual Physiological Rat Project.
http://floating-point-gui.de
What Every Programmer Should Know About Floating-Point Arithmetic.

Acknowledgments: It is impossible to appropriately thank the great many people who have contributed to the many discussions that inspired this paper over the years, either by working with me on relevant aspects, or by drawing my attention to important work, or pointing out weaknesses, or inspiring ideas improving this work, or by contributing to the EvoSysBio meetings, or patiently waiting for me to finish this chapter: thank you to all. In particular to Anthony Pietsch and Christine Javid for numerous suggestions that greatly improved the clarity of this text, and to the editors of this Encyclopedia for their patience and support. Thanks to the National Science Foundation for supporting this work under Career Award 1149123, helping to simplify modeling in biology. Any opinions, findings, and conclusions or recommendations expressed in this material are those of LL and do not necessarily reflect the views of the NSF or any person contributing to the discussions that have influenced this work.