REvoSim v3: A fast evolutionary simulation tool with ecological processes

Euan N. Furness [1,2,¶], **Russell J. Garwood** [3,4], and **Mark D. Sutton** [1]

1 Department of Earth Sciences and Engineering, Imperial College London, United Kingdom **2** Grantham Institute, Imperial College London, United Kingdom **3** Department of Earth and Environmental Sciences, University of Manchester, United Kingdom **4** Natural History Museum, London, United Kingdom **¶** Corresponding author

Summary

Simulations are an important tool for understanding ecology and evolution because they allow greater precision and customisation than traditional experiments or observation (Barido-Sottani et al., 2020; Cabral et al., 2016; Zurrell et al., 2010). Additionally, they can model phenomena that occur over timescales too long to be easily empirically observed, but too short to be studied using the fossil record (Garwood et al., 2019). REvoSim v3 is a flexible, spatially explicit, individual-based eco-evolutionary simulation tool written in C++. It is an update of REvoSim v2 (Garwood et al., 2019) that increases the flexibility of genetic architecture within the simulation while also incorporating ecological processes such as inter-organism interactions.

Background

REvoSim v2 (Garwood et al., 2019) simulates the ecology of hundreds of thousands of sessile organisms on a spatially explicit environment grid. Each organism is a 64-bit binary genome. The first 32 genome bits determine the organism's ability to collect energy from the environment by interacting with three independent environmental variables, each represented by one colour channel of an RGB image of the same size as the environment grid. The simulation proceeds in timesteps during which organisms collect energy from their environment, breed, if they have sufficient energy and, after a fixed number of timesteps, die. Mutations may arise during breeding, and organisms are grouped into species based on a strict biological species concept (Mayr, 1942), with breeding compatibility based on genetic similarity.

New Features

Genome architecture

REvoSim v3 replaces the 64-bit genome architecture of REvoSim v2 with a configurable system of 32-bit genome "words"- organisms can comprise any number of these words, and different word combinations can be utilised by REvoSim's different systems.

Systems

All REvoSim processes (e.g. fitness, mutation, breeding, species ID) are now designated systems, and can be applied to specified genome words as required. Newly added systems include interactions and pathogens, introduced below. Further additions allow genome words to be assigned to modify mutation rate and/or reproductive mode (sexual or asexual).

Pathogen dynamics

REvoSim v3 incorporates an optional "pathogen layer" — a grid with identical dimensions to the environment. Pathogen genomes' architecture is defined by the system's options and, at a configurable rate, pathogen genome words are compared with those of organisms in the same position on the grid. These comparisons have a 50% chance of killing an organism if it has an identical genome string to the pathogen, and a 0% chance if every bit is different, with linear interpolation between these extremes.

Pathogens evolve over time, and may do so through drift or selection for virulence. When pathogens evolve through drift, each position on the grid contains a single pathogen which changes through mutation at a configurable rate. When pathogens evolve through selection, each position on the grid contains five pathogen genomes. In each timestep, the most virulent pathogen in each position in the pathogen grid, or positions adjacent to it, is applied, duplicated, and then returned, overwriting an existing pathogen.

Variable masks

REvoSim v3 adds mechanism that allows for the productivity and/or mutation rates of organisms to vary spatially and temporally within the simulation. This is achieved via variable masks - image files with the same dimensions as the simulated environment, where the magnitude of a masked variable at any given position in the environment is equal to that variable's global value, multiplied by the brightness (0-255) of the red channel of the pixel in the mask image at that position.

Interactions

Inter-organism interactions can be simulated in REvoSim v3 using one of two approaches, referred to as the 'genome blocks' and 'XOR' mechanisms. Both interaction mechanisms are designed such that small genome changes result in small changes to interaction scores. In both cases, the number of attempted interactions per-timestep per-individual is configurable. Individuals attempt to interact by searching for another individual at the same position as them in the environment. The probability of success of this search increases as organism-count in the cell increases. If another individual is found, an interaction occurs. Under 'genome blocks', an 'interaction score' is derived from a configurable table that describes how every possible combination of two-bit organism genome components interact Figure 1. Under 'XOR', one of the genomes is rotationally shifted by one bit both to the left and to the right Figure 2, and each shifted genome is combined with an exclusive-or function with the unshifted genome. The number of 1s in the right-shifted, XORed binary sequence subtracted from the number of 1s in the left-shifted, XORed binary sequence, creates an interaction score Figure 2.

Furness et al. (2023). REvoSim v3: A fast evolutionary simulation tool with ecological processes. *Journal of Open Source Software*, *8*(89), 5284. 2
https://doi.org/10.21105/joss.05284.

Organism 1 Genome Bit Pair / Organism 2 Genome Bit Pair	00	01	10	11
	Interaction Score			
00	0	+3	0	+1
01	-3	0	+2	+1
10	0	-2	0	-2
11	+1	-1	-2	0

Figure 1: An example interactions table for the 'genome blocks' mechanism. Adjacent bits in an interacting organism's genomes are paired (top) and compared with the corresponding pair in the other interacting organism (left) to determine their contribution to the interaction score. The total interaction score is the sum of the scores generated by every pair of bits.
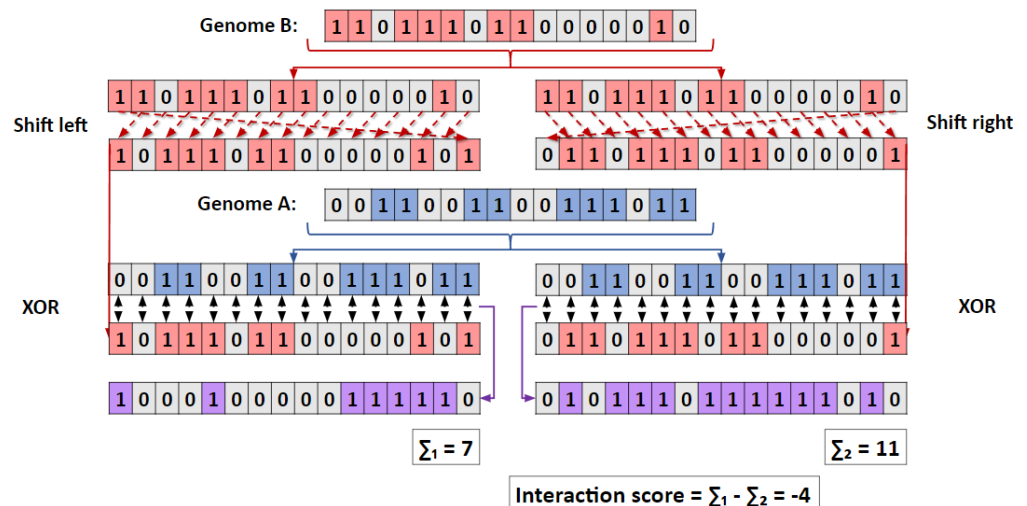

Figure 2: A flowchart depicting the XOR mechanism for two genomes (A: blue, and B: red), abbreviated to 16 bits in length for ease of interpretation. Swapping genomes A and B would reverse the sign of the interaction score.

Interactions may alter either organism fitnesses, modifying the rate at which they collect energy from the environment, or allow organisms to transfer energy from one another. Under the former, the interaction score (which may be negative) is added to the fitness of the instigating organism. Under the latter, the interaction score controls the proportion of the sought organism's current energy that is lost in the interaction Figure 3. Optionally, energy transfer can be disabled where this proportion would be below some threshold. Of the energy lost by the sought organism, a configurable fraction is transferred to the instigating organism, reflecting loss of energy to metabolic inefficiency (Anderson et al., 2009; Lindeman, 1942).
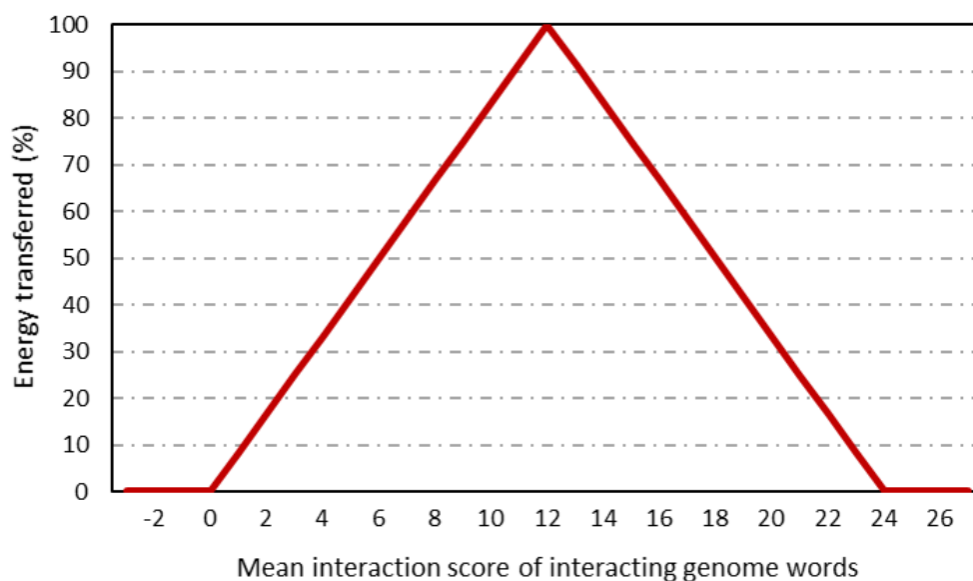

Figure 3: The proportion of energy lost by the sought organism, when direct energy transfer is occurring, for any given mean genome word interaction score.

To accompany the interactions system, REvoSim v3 includes the option to seed either three or five (in addition to the previous default of one) different genomes at the start of the simulation. When three or five genomes are seeded, preset genomes are used that are known to form a single, unbranching food-chain under the XOR mechanism (Barbier & Loreau, 2019).

To track the effects of energy-transfer interactions, REvoSim v3 records a 'trophic level' for each organism over the course of its lifetime (Anderson et al., 2009). Trophic level is initially inherited from the parent, but subsequently becomes one greater than an average of the trophic levels of all of the sources from which the organism has gained energy during its life, weighted by the amount gained from each source. This results in calculated trophic levels that mirror those in real world ecosystems (i.e. zero for producers, one for herbivores, two for primary carnivores, etc.).

Logging

The logging system within the software has been redesigned to improve flexibility. The user can now define a header, a string output for every polling iteration with grid-level statistics, and an output for every species every polling iteration. These allow any combination of outputs required for an experiment.

Species identification

The species identification algorithm has been recoded to improve performance.

Statement of need

Despite simulations becoming progressively more common in evolutionary ecology (Barido-Sottani et al., 2020; Cabral et al., 2016; Zurrell et al., 2010), many function at the level of the population, rather than the individual (Hagen et al., 2021; Rangel et al., 2018; Reddin et al., 2022). Simulations that do function at the level of the individual often focus on ecology, rather than evolution (Harfoot et al., 2014; Smith & Lundholm, 2010), although other individual-based evolutionary simulation tools do exist (Dolson & Ofria, 2021). All digital evolution and simulation tools experience a tradeoff between complexity and efficiency (Dolson & Ofria, 2021), and REvoSim is no exception. REvoSim minimises system complexity through abstraction of organism and environment traits: an approach that improves efficiency, while still allowing for the investigation of complex phenomena as emergent properties of the model. This approach allows REvoSim to simulate long time spans and large population sizes on modest hardware. However, this approach abstracts out any effect of genetic complexity or ontogeny on organisms. REvoSim is therefore complementary with other simulation and digital evolution tools that incorporate greater explicit complexity (particularly genetic complexity; (Batut et al., 2013)) at the cost of efficiency.

The new features added in REvoSim v3 allow for the study of ecological processes (predation, symbioses, and pathogens) that have been implicated in the evolution of diversity gradients, such as the latitudinal biodiversity gradient (Johnson et al., 2012; Roslin et al., 2017). Studies could directly modify the intensity or operation of these processes within REvoSim in order to determine their impact on the formation of such gradients. Furthermore, these processes allow REvoSim to be used as a tool for complex ecological simulation in the absence of evolutionary processes, and have also increased the versatility of the tool throughout.

REvoSim v3's mask tool allows it to simulate geographical areas of varying size, shape, and configuration and, therefore, study the impacts of these properties on evolutionary processes and biodiversity. For example, this allows REvoSim to be used to simulate habitat fragmentation (Matthews et al., 2014) and the island species-area relationship (Matthews et al., 2021), both of which are important topics in conservation science (Hanski et al., 2007).

Current associated projects

RJG - Pathogens and environmental disturbance as drivers of reproductive mode.

ENF and MDS - Effects of evolutionary processes on trophic level abundance equilibria.

ENF, MDS and RJG - Rates of evolution during speciation.

Availability

REvoSim v3.0.1 is freely available from Zenodo and GitHub. Newer releases of the REvoSim software will be available on GitHub.

Acknowledgements

References

Anderson, S., Beyer, J. E., & Lundberg, K. V. (2009). Trophic and individual efficiencies of size-structured communities. *Proceedings of the Royal Society B - Biological Sciences*.

https://doi.org/10.1098/rspb.2008.0951

Barbier, M., & Loreau, M. (2019). Pyramids and cascades: a synthesis of food chain functioning and stability. *Ecology Letters*. https://doi.org/10.1111/ele.13196

Barido-Sottani, J., Saupe, E. E., Smiley, T. M., Soul, L. C., Wright, A. M., & Warnock, R. C. M. (2020). Seven rules for simulations in paleobiology. *Paleobiology*. https://doi.org/10.1017/pab.2020.30

Batut, B., Parsons, D. P., Fischer, S., Beslon, G., & Knibbe, C. (2013). In silico experimental evolution: a tool to test evolutionary scenarios. *BMC Bioinformatics*. https://doi.org/10.1186/1471-2105-14-S15-S11

Cabral, J. S., Valente, L., & Hartig, F. (2016). Mechanistic simulation models in macroecology and biogeography: state-of-art and prospects. *Ecography*. https://doi.org/10.1111/ecog.02480

Dolson, E., & Ofria, C. (2021). Digital evolution for ecology research: a review. *Frontiers in Ecology and Evolution*. https://doi.org/10.3389/fevo.2021.750779

Garwood, R. J., Spencer, A. R. T., & Sutton, M. D. (2019). REvoSim: organism-level simulation of macro and microevolution. *Palaeontology*. https://doi.org/10.1111/pala.12420

Hagen, O., Flück, B., Fopp, F., Cabral, J. S., Hartig, F., Pontarp, M., Rangel, T. F., & Pellissier, L. (2021). gen3sis: a general engine for eco-evolutionary simulations of the processes that shape Earth's biodiversity. *PLoS Biology*. https://doi.org/10.1371/journal.pbio.3001340

Hanski, I., Koivulehto, H., Cameron, A., & Rahagalala, P. (2007). Deforestation and apparent extinctions of endemic forest beetles in Madagascar. *Biology Letters*. https://doi.org/10.1098/rsbl.2007.0043

Harfoot, M. B. J., Newbold, T., Tittensor, D. P., Emmott, S., Hutton, J., Lyutsarev, V., Smith, M. J., Scharlemann, J. P. W., & Purves, D. W. (2014). Emergent global patterns of ecosystem structure and function from a mechanistic general ecosystem model. *PLoS Biology*. https://doi.org/10.1371/journal.pbio.1001841

Johnson, D. J., Beaulieu, W. T., Bever, J. D., & Clay, K. (2012). Conspecific negative density dependence and forest diversity. *Science*. https://doi.org/10.1126/science.1220269

Lindeman, R. L. (1942). The trophic-dynamic aspect of ecology. *Ecology*. https://doi.org/10.2307/1930126

Matthews, T. J., Cottee-Jones, H. E., & Whittaker, R. J. (2014). Habitat fragmentation and the species–area relationship: a focus on total species richness obscures the impact of habitat loss on habitat specialists. *Diversity and Distributions*. https://doi.org/10.1111/ddi.12227

Matthews, T. J., Triantis, K. A., & Whittaker, R. J. (2021). *The species-area relationship: theory and application*. Cambridge University Press. https://doi.org/10.1017/9781108569422

Mayr, E. (1942). *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. Columbia University Press.

Rangel, T. F., Edwards, N. R., Holden, P. B., Diniz-Filho, J. A. F., Gosling, W. D., Coelho, M. T. P., Cassemiro, F. A. S., Rahbek, C., & Colwell, R. K. (2018). Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums, and graves. *Science*. https://doi.org/10.1126/science.aar5452

Reddin, C. J., Abraham, M., Raja, N. B., & Kocsis, Á. T. (2022). Global warming generates predictable extinctions of warm- and cold-water marine benthic invertebrates via thermal habitat loss. *Global Change Biology*. https://doi.org/10.1111/gcb.16333

Roslin, T., Hardwick, B., Novotny, V., Petry, W. K., Andrew, N. R., Asmus, A., Barrio, I. C., Basset, Y., Boesing, A. L., Bonebrake, T. C., Cameron, E. K., Dáttilo, W., Donoso, D. A., Drozd, P., Gray, C. L., Hik, D. S., Hill, S. J., Hopkins, T., Huang, S., … Slade, E. M. (2017). Higher predation risk for insect prey at low latitudes and elevations. *Science*. https://doi.org/10.1126/science.aaj1631

Smith, T. W., & Lundholm, J. T. (2010). Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography*. https://doi.org/10.1111/j.1600-0587.2009.06105.x

Zurrell, D., Berger, U., Cabral, J. S., Jeltsch, F., Meynard, C. N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B., & Grimm, V. (2010). The virtual ecologist approach: simulating data and observers. *Oikos*. https://doi.org/10.1111/j.1600-0706.2009.18284.x