Model code for the EpiGen model used in Walworth et al. 2020 and example output

Website: https://www.bco-dmo.org/dataset/862458
Data Type: model results
Version: 1
Version Date: 2021-10-12

Project
» How does intensity and frequency of environmental variability affect phytoplankton growth? (Enviro variability and phytoplankton growth)

Contributors	Affiliation	Role
Levine, Naomi M.	University of Southern California (USC)	Principal Investigator, Contact
Soenen, Karen	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Model code and example model output for the EpiGen model used in Walworth et al. 2020. The EpiGen model is an individual-based model of adaptation modified from Fisher's model in which a simulated population moves between a "new" and "ancestral" environment following a step function with varying frequencies. This model calculates the rate of adaptation of the population where adaption proceeds through both fast variation, low transmission and slow variation, high transmission (HT) modifications.

Table of Contents

Dataset Description

The EpiGen model is an individual-based model of adaptation modified from Fisher's model in which a simulated population moves between a "new" and "ancestral" environment following a step function with varying frequencies. This model calculates the rate of adaptation of the population where adaption proceeds through both fast variation, low transmission and slow variation, high transmission (HT) modifications.

The model output is for the adaptation of a population of 1000 individuals to a fluctuating environment (50 generations selective pressure, 50 generations non-selective) over 15000 generations. Every row is a generation (time). Columns are different statistics characterizing the population of 1000 individuals.

- Column 1 is generation #,
- Column 2 is mean fitness at that time step,
- Column 3 is mean # of HT modifications at that time step,
- Column 4 is mean number of LT modifications,
- Column 5 is cumulative distance traveled towards optimum by HT modifications,
- Column 6 is cum. distance traveled towards optimum by LT modifications,
- Column 7 is standard dev of number of HT modifications,
- Column 8 is standard dev of number of LT modifications,
- Column 9 is mean distance traveled toward optimum by HT modifications,
- Column 10 is sd of distance traveled toward optimum by HT modifications,
- Column 11 is mean distance travelled toward optimum by LT modifications, and
- Column 12 is sd of distance travelled toward optimum by LT modifications.

The model code is available in a .zip file: (see section data files). These files are also available in the following GitHub repository: https://github.com/BCODMO/EpiGen

Methods & Sampling

We modeled an individual based adaptive walk using a modified version of Fisher's geometric adaptation model from Kronholm and Collins – the EpiGen model.

Fitness changes were driven by both LT (low transmission) modifications and HT (high transmission) modifications, where HT modifications were fixed and LT modifications reverted with probability urev (LT reversion rate). Model formulation: Phenotypic space was represented as an n-dimensional hypersphere where an individual's phenotype, z, was characterized by its distance from the hypersphere origin with radius r.

Fitness for each individual (w) was calculated as: w(z) = e(-z^2)/2 (Eq. 1)
such that an individual located at the origin had an optimal fitness of 1 and fitness declined as a Gaussian function as the phenotype moved away from the origin of the hypersphere. The simulations began with z = r = 1 for all individuals in the population such that w = 0.6065.

The phenotype was altered through both LT and HT modifications which were represented as mutational vectors with random directions and magnitudes in phenotypic space. A new phenotypic value, Zmut, was then calculated as:
zmut^2 = z^2 + m^2 + 2mzsin(sigma) (Eq. 2)
where m is the length of the mutational vector and sigma is the angle between the mutational vector and the vector running from the current phenotype to the origin (sigma is an element of [-pi/2, pi/2]).

For each new modification, sigma in n-dimensional space was drawn from the probability density (P):
P=Zcos(sigma)^(n-2) (Eq. 3)
where Z is a scaling constant and was calculated as:
integral[pi/2,−pi/2] Zcos(sigma)^(n-2) dsigma = 1 (Eq. 4)

All model parameters are given in Supplemental Table 1 in Walwort et al., 2020.

The model was initialized with a population of N uniform individuals: here N was varied from N = 103 to N = 105. HT modifications (NHT =10) and LT modifications (NLT =90) were then introduced into the population. The modification supply (population size x modification rate) remained constant in each generation and no more than one LT and one HT modification per generation was allowed to occur in a single individual. Eq. 2 was used to calculate new mutant 2 phenotypes. Isotropic modifications in phenotypic space were represented through the uniform distribution of HT modifications, mht, between 0 and 2r, mht ~ U(0, 2r), which generated nonuniform fitness effects (2, 4). While LT modifications (mlt) were introduced in the same manner as HT modifications, the effects of LT modifications were smaller than HT modifications with a uniform distribution of me between 0 and l. Hence, mlt ~ U(0, l) instead of mht ~ U(0, 2r), where 2r is the maximum effect of HT modifications and l £ 2r is the maximum effect of LT modifications. Fitness for each mutant phenotype was then calculated using Eq. 1, and the next generation was then created by sampling from the current population with replacement.

Selection: In the 'new' environment, selection was based on fitness in the 'new' environment so the sampling probability of an individual was weighted by its fitness until N offspring had been produced. Selection in the 'ancestral' environment occurred through the stochastic removal of organisms with relatively more HT modifications (i.e. higher HT modification abundance), which corresponds to stabilizing selection. We assume that all modifications have an equal chance of being conditionally deleterious (being neutral or adaptive in the 'selection' or 'new' environment, but deleterious in some other environment) so that individuals who have accumulated a high number of modifications in the selection environment
have a higher probability of decreased fitness in the 'ancestral' environment.

Data Processing Description

BCO-DMO Processing Notes:

* Originally submitted GitHub repository https://github.com/LevineLab/EpiGen forked to https://github.com/BCODMO/EpiGen and tagged with release v1.0 which corresponds with this dataset submission. This version is also attached in the data files section. The original repository may have continued

updates.

* Added model output example to section data files.

[table of contents | back to top]

Data Files

File
Epigen Model v1.0 filename: EpiGen-master.zip (ZIP Archive (ZIP), 9.00 KB) MD5:c4337f8e95cdd99a2ff09f6ca7ec1255 The model code is available in the following GitHub repository: https://github.com/BCODMO/EpiGen
Epigen Model Example Output filename: results_50n_all_15K_generations.txt (Octet Stream, 1.17 MB) MD5:c3cbc6b25d066d002677c8bdf846801d The model output is for the adaptation of a population of 1000 individuals to a fluctuating environment (50 generations selective pressure, 50 generations non-selective) over 15000 generations. Every row is a generation (time). Columns are different statistics characterizing the population of 1000 individuals. Column 1 is generation #, Column 2 is mean fitness at that time step, Column 3 is mean # of HT modifications at that time step, Column 4 is mean number of LT modifications, Column 5 is cumulative distance traveled towards optimum by HT modifications, Column 6 is cum. distance traveled towards optimum by LT modifications, Column 7 is standard dev of number of HT modifications, Column 8 is standard dev of number of LT modifications, Column 9 is mean distance traveled toward optimum by HT modifications, Column 10 is sd of distance traveled toward optimum by HT modifications, Column 11 is mean distance travelled toward optimum by LT modifications, and Column 12 is sd of distance travelled toward optimum by LT modifications. The model code is available in a .zip file: [link to our local zip of package] These files are also available in the following GitHub repository: https://github.com/BCODMO/EpiGen

[table of contents | back to top]

Related Publications

Fisher, R.A. (1930). The Genetical Theory of Natural Selection. Clarendon Press, Oxford, UK, est. 272 pp.
Methods

Hartl, D. L., & Taubes, C. H. (1998). Genetica, 102/103, 525–533. doi:10.1023/a:1017071901530 https://doi.org/10.1023/A:1017071901530
Methods

Kronholm, I., & Collins, S. (2015). Epigenetic mutations can both help and hinder adaptive evolution. Molecular Ecology, 25(8), 1856–1868. doi:10.1111/mec.13296
Methods

Orr, H. A. (1998). The Population Genetics of Adaptation: The Distribution of Factors Fixed during Adaptive Evolution. Evolution, 52(4), 935. doi:10.2307/2411226

Methods

Walworth, N. G., Zakem, E. J., Dunne, J. P., Collins, S., & Levine, N. M. (2020). Microbial evolutionary strategies in a dynamic ocean. Proceedings of the National Academy of Sciences, 117(11), 5943–5948. doi:10.1073/pnas.1919332117
Results

[table of contents | back to top]

Parameters

Parameters for this dataset have not yet been identified

[table of contents | back to top]

Project Information

How does intensity and frequency of environmental variability affect phytoplankton growth? (Enviro variability and phytoplankton growth)

Coverage: laboratory experiment

NSF Award Abstract:
Microscopic plants called phytoplankton are key members of global oceanic ecosystems, since their photosynthesis supports the majority of the marine food chain and produces about as much oxygen as land plants. Because of this, oceanographers have often carried out experiments examining how factors such as temperature and carbon dioxide levels may affect phytoplankton growth. Most previous experiments have used constant levels of temperature and carbon dioxide, but it is clear from looking at measurements from real ocean ecosystems that these two factors often vary greatly over timescales of days to weeks. Using field and laboratory experiments along with computer modeling, this project will test how the growth of several major groups of phytoplankton differs under constant conditions of temperature and carbon dioxide, compared to conditions in which these factors fluctuate in intensity and frequency. This research will give marine scientists a better picture of how phytoplankton may respond to a varying natural environment today and in the future, and therefore help us to understand how ocean food webs function to support critical living resources such as fisheries. The project will train graduate and undergraduate students and a postdoctoral researcher, and the lead scientists will be involved in an ocean science education program for largely minority high school students from a downtown Los Angeles school district.

The goal of this project is to use laboratory culture and natural community experiments to understand how realistically fluctuating temperature and pCO2 conditions may affect globally important phytoplankton groups in ways that differ from the artificial constant exposures used in previous work. Culture experiments will test how the intensity and frequency of short-term thermal and carbonate fluctuations affects the growth responses of diazotrophic and picoplanktonic cyanobacteria, coccolithophores, and diatoms under both current and projected future environmental conditions. These lab results will be supported and extended by parallel experiments using mixed natural assemblages from the California upwelling regime, allowing us to test these same questions using phytoplankton communities that experience large seasonal shifts between highly dynamic thermal and carbonate system conditions during the spring upwelling season, and relatively much more static conditions during fall stratification events. These results will be synthesized using a new generation of numerical models that employ novel approaches to incorporating realistic environmental variations to allow more accurate predictions of phytoplankton responses to a dynamic environment in today's marine ecosystems, and in the future changing ocean.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1538525

[table of contents | back to top]