

PROCEED v6.0

Phenotypic Rates of Change Evolutionary and Ecological Database

RatesDB – Metadata

General

Release date

2022/08/12

Citation

Gotanda, Kiyoko M.; Gorné, Lucas D. (2022) "Phenotypic Rates of Change Evolutionary and Ecological Database (PROCEED)", version 6, Borealis, <https://doi.org/10.5683/SP3/NXSL3Q>

Subject

Ecology and Evolution

Keywords

Contemporary evolution, haldanes, darwins, quantitative traits, time series.

Current curator

Kiyoko M. Gotanda^{1, 2} <kgotanda@brocku.ca>

Lucas D. Gorné¹ <gorneld@gmail.com>

Affiliations

1. Department of Biological Sciences, Brock University, St. Catharines, Ontario, Canada.

2. Département de Biologie, Université de Sherbrooke, Sherbrooke, Quebec, Canada.

Data Entry

Carlos Arias, Marc-Olivier Beausoleil, Maryse Boisjoly, Cristian Correa, Erika Crispo, Daniel Berner, Luis De Leon, Joey DiBattista, Thomas Farrugia, Jason Gencher, Lucas D. Gorné, Kiyoko M. Gotanda, Grant E. Haines, Benjamin C. Haller, Andrew P. Hendry, Renaud Kauffer, Elspeth Kenny, Michael T. Kinnison, Shahin Muttalib, Ann McKellar, Rose E. O’Dea, Fanie Pelletier, Winer Daniel Reyes-Corral, Krista B. Oke, Sarah Sanderson, Amy Schwartz, Freedom Sorbara, Xavier Thibert-Plante, Zachary T. Wood.

Data Contributions

Carnegie Museum of Natural History, Maine Department of Marine Resources, NS-ITBS surveys (www.ices.dk/marine-data/dataportals/Pages/DATRAS.aspx), Powdermill Avian Research Station

Technical

Language: English

Decimal point="."

Not available data (NA)=" " (i.e. empty cell)

Description, Inclusion/Exclusion criteria

This database is an ongoing compilation of rates of phenotypic change, typically Haldanes and Darwins, published in peer reviewed manuscripts. We compile and add data regularly to the data set. This dataset is continually being updated as more people ask it include new variables.

This database includes studies that measure intraspecific change in quantitative (continuous or counting) traits and report the time elapsed from the onset of environmental novelty or refer to a historical or biological event reported in other sources (e.g., a mine opening, a well documented biological invasion). The maximum elapsed time between the environmental change and the sampling was no longer than 500 years (however, currently published papers had cut-off number of years at 200/300 years; it is up to the user to decide to keep those above that previous threshold). The included studies followed a single population through time, or compared two or more populations, diverging from an originally single population where (at least) one of them was a new condition of known age (i.e., allochronic and synchronic designs according to Hendry & Kinnison, 1999).

The database includes systems in natural or experimental conditions but in which reproduction was not manipulated (i.e. NO artificial selection is included). The experimental systems manipulate the environment of the populations being studied but do not select traits deliberately (e.g. an herbivory exclusion experiment, where the type and load of herbivory are manipulated and the populations evolve under this new condition).

The database does not include systems where the phenotypic change is due to interspecific hybridization, polyploidy, or other chromosomal mutations.

Variables:

es_ID

Unique identifier for each individual rate.

ref_ID

Unique identifier for each primary source of information (paper, thesis, report, etc).

sys_ID

Unique identifier for each study system (see "system" for the definition).

released_sys

sys_ID of the system originating the new system after removal/extinguishment/stop of the driver/disturbance. For example, the system s445 (Ovis canadensis in Ram Mountain, Alberta, Canada when a hunting moratorium was placed starting in 1996) is the same population of system s122 (Ovis canadensis in Ram Mountain, Alberta, Canada when hunting was allowed up to 1995). So, for the system s445 "released_sys"="s122". On the contrary, for the system s122 "released_sys"=NA (because s122 was not derived from any other system in the database). Future studies with similar situations (i.e. a single population in periods with different evolutionary drivers) should include, in "released_sys" of the new condition, the sys_ID of the former condition.

released_es

es_ID of the case originating the new case after removal/extinguishment/stop of the driver/disturbance in the system. A system being released from the original driver generates a new system. If more than one trait is measured in both systems (the original and the new one), "released_es" links the cases in the new system with the measurements of the same trait in the original systems. For example, in the systems s122 and s445, two traits were measured: mean weight (kg) and mean horn length (cm). In s122, the change in mean weight is the case "es_ID"="es2798" and the change in mean horn length is the case "es_ID"="es2799". So, for the case es9277 (change in mean weight in the system s445), "released_es"="es2798". For the case es9278 (change in mean horn length in

the system s445), “released_es”=“es2799”. On the contrary, for es2798 and es2799 (system s122), “released_es”=NA (because s122 was not derived from any other system in the database). Future studies with similar situations (i.e. a single population in periods with different evolutionary drivers) should include, in “released_es” of the new condition, the es_ID of the same traits in the former condition.

species

Species name (or subspecific taxon name) of the population(s) being assessed in each case, as was published in the primary source of information.

sp_ncbi

Species or subspecific taxon name of the population(s) being assessed in each case, according to the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Root>) in February 2022. Two species were not found in the NCBI database, *Stipa occidentalis* and *Diaptomus sanguineus*, these names were checked in The Plant List (<http://www.theplantlist.org/>) and WORMS (<https://www.marinespecies.org/index.php>) respectively. The supplementary table "PROCEEDv6_taxonomy.csv" contains the complete list of "species" and its "sp_ncbi".

taxa

This is a coarse classification, not necessarily a clade (monophyletic group). Current categories are: **Annelid, Mollusc, Arthropod, Fish, Amphibian, Mammal, Bird, Reptile, Plant.**

taxa_fine

This is an old variable kept from previous versions of the database. For future versions, there is no need to code new data entries according to this variable. Currently, this variable has been replaced by "class" and "order" for more precision. The last definition of "taxa_fine" was:

“Vertebrates classified by class. Invertebrates classified to phylum. Plants classified to family. This is a finer taxa classification. Current classifications are: Actinopterygii,

Amphibia, Annelid, Araliaceae, Asteraceae, Aves, Betulaceae, Boraginaceae, Brassicaceae, Caryophyllaceae, Crustacea, Fabaceae, Gastropod, Hypericaceae, Insecta, Lythraceae, Mammalia, Melastomataceae, Mollusca, Myrtaceae, Plantaginaceae, Poaceae, Reptilia, Additional finer level taxa can be added as necessary."

phylum

Taxonomic clade of rank "phylum" including the studied species, according to the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Root>) last retrieved May 2022. The supplementary table "PROCEEDv6_taxonomy.csv" contains the complete list of "species" and its "phylum".

class

Taxonomic clade of rank "class" including the studied species, according to the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Root>) last retrieved May 2022. The supplementary table "PROCEEDv6_taxonomy.csv" contains the complete list of "species" and its "class".

order

Taxonomic clade of rank "order" including the studied species, according to the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Root>) last retrieved May 2022. The supplementary table "PROCEEDv6_taxonomy.csv" contains the complete list of "species" and its "order".

system

It is a population or a set of populations, of a given species, experiencing a specific driver/disturbance. If the design is synchronic, the system includes both diverging populations (or sets of populations). Each system has its own unique combination of species, disturbance/driver and location. Within a given system, you can have multiple traits – for example, tarsus length and fledging date. We suggest keeping it general. Please note you can have both allochronic and synchronic data within the same system (e.g. Leaver 2012, Biol. J. Linnean. Soc. 107: 494-509). Example systems: "Salmo trutta (brown trout) monitored for 37 y following fish ladder

construction (1966) in River Gudbrandsdalslågen, Norway"; "Troglodytes troglodytes (Eurasian Wren) migration arrival date monitored in its native range in Europe"; "Oncorhynchus gorbuscha (pink salmon) long term trends on northwest coast of US".

experimental

States if the phenotypic change is driven by experimentally set conditions (**N**=no; **Y**=yes). It is not artificial selection (i.e. selection of specific attributes by choosing breeders and manipulating mating) but a process of natural selection in an experimental environment.

exp_type

When "experimental"="N", "exp_type"=NA (empty cell). When "experimental"="Y", the following categories were applied:

- **01**: A new environmental condition is imposed on a natural population, in the field, in situ.
- **03**: A new environmental condition is imposed on a natural population, in the field, ex situ (transplant).
- **05**: A new environmental condition is imposed on a natural population, in a common garden/greenhouse experiment.
- **07**: A new environmental condition is imposed on an artificially generated population in the field. Here, the artificial population is a set of genotypes from the natural population, evenly represented by (full or half) sibling families.
- **09**: A new environmental condition is imposed on an artificially generated population in a common garden experiment. Here, the artificial population is a set of genotypes from the natural population, evenly represented by (full or half) sibling families.
- **11**: A new environmental condition is imposed on an artificial population (it was obtained by randomly crossing F2 individuals from a set of interbreeds between contrasting genotypes selected from a natural population) in a common garden experiment.

comments__on__system

Any comments about the system should be stated here.

c__value

The amount (in picograms) of DNA contained within a haploid nucleus (e.g. a gamete) or one half the amount in a diploid somatic cell. Most of the values were taken from:

- Gregory TR. 2021. Animal Genome Size Database. <http://www.genomesize.com>.
- Leitch IJ, Johnston E, Pellicer J, Hidalgo O, Bennett MD. 2019. Plant DNA C-values Database (Release 7.1). <https://cvalues.science.kew.org/>

Detailed information of the variable, sources, ploidy level, analytical methods, etc in the supplementary table "PROCEEDv6_species_genome.csv" and "PROCEEDv6_species_genome_README.txt".

WHEN MORE THAN ONE C-VALUE IS AVAILABLE FOR ONE SPECIES

- If there is no difference in ploidy level or chromosome number among records, c-values were averaged.
- Records using Feulgen densitometry (Fe), flow cytometry (FC__any type), DAPI microdensitometry (MDAPI), Gallocyanin chrom alum densitometry (GDC), or Fulgen image analysis densitometry (FIA) methods are preferred over other methods. Thus, records using these methods were chosen for averaging when available.
- In general, if the genome size is available from several subspecies, the value for each subspecies was previously averaged by LG, then the average among them was imputed for the necessary species.
- In general, but especially in plants, when a species has more than one genome size due to differences in ploidy level or chromosome number, information for the specific populations within the system were researched. If no information was found, the available c-values were averaged.

SPECIAL CASES in C-value

- in birds (low c-value variability), if the genome size is available from several congeners (but not for the necessary species), the values of the multiple measurements of c-value (if more than one) for each congener were previously averaged by LG, then the average among congeners was imputed for the necessary species.

sex

Sex of the individuals measured (male, female, asexual, mixed=male+female, NA(empty cell)=sex not reported in the study).

sample1

Identification of Sample 1 (e.g. year, population, identifier given by author(s)). In the case of allochronic studies, Sample 1 should always be the earlier sample. In the case of synchronic studies, sample 1 should always be the source population (e.g. in the case of introductions or range expansions, Sample 1 should be the site “before introduction or expansion”) or the population(s) in the original/ancestral environment condition (e.g. in the case of pollution, sample 1 should be the population in the non-polluted area; in the case of response to introduction, sample 1 should be the population in the area where the introduced species is not present). When the traits are measured in a common garden experiment with different conditions, in more than one common garden location, or in reciprocal transplant experiments, these conditions are noted here (e.g. sample1:"Control origin, control treatment", sample2:"Polluted origin, Control treatment"; sample1: "Control origin, Drought treatment", sample2: "Polluted origin, Drought treatment"; sample1:"Native populations, Jokers Hill common garden", sample2:"Introduced populations, Jokers Hill common garden"; sample1:"Native populations, Montpellier common garden", sample2:"Introduced populations, Montpellier common garden").

sample2

Identification of Sample 2 (e.g. year, population, identifier given by author(s)). In the case of allochronic studies, Sample 2 should always be the later sample. In the case of synchronic studies, Sample 2 should always be the later populations (e.g. in the case of introductions or range expansions, Sample 2 should be the site “after introduction or range expansion”) or the population in the new environmental condition (e.g. in the case of pollution, sample 2 should be the population in the

polluted area; in the case of response to introduction, sample 2 should be the population in the area where the introduced species is present). When the traits are measured in a common garden experiment with different conditions, or in more than one common garden location, or in reciprocal transplant experiments, these conditions should be log here (e.g. sample1:"Control origin, control treatment", sample2:"Polluted origin, Control treatment"; sample1: "Control origin, Drought treatment", sample2: "Polluted origin, Drought treatment"; sample1:"Native populations, Jokers Hill common garden", sample2:"Introduced populations, Jokers Hill common garden"; sample1:"Native populations, Montpellier common garden", sample2:"Introduced populations, Montpellier common garden").

samp1__anc

(yes/no) If sample 1 was from the ancestral population or the population from the ancestral place/source/condition, "samp1__anc"="yes". If there is no ancestral-to-derived relationship between samples, then "samp1__anc"="no".

sample1.latitude

Latitude in decimals (to a minimum of three decimal places if possible) of the location of the "sample1" population. Note that South is negative. If the lat/longs are not given please look up via other studies of the same system, Wikipedia, Google Maps, or OpenStreetMap. You can use these websites to convert from degrees or UTM:

- For URL: http://andrew.hedges.name/experiments/convert_lat_long/
- For UTM: http://www.engineeringtoolbox.com/utm-latitude-longitude-d_1370.html

sample1.longitude

Longitude in decimals (to a minimum of three decimal places if possible) of the location of the "sample1" population. Note that West is negative. If the lat/longs are not given please look up via other studies of the same system, Wikipedia, Google Maps, or OpenStreetMap. You can use the above websites to convert from degrees or UTM.

sample2.latitude

Same as "sample1.latitude" but for the location of "sample2" populations.

sample2.longitude

Same as "sample1.longitude" but for the location of "sample2" populations.

coordinate.means

If the coordinates for sample1 and sample2 are the average of coordinates of more than one population (**Y**) or not (**N**).

coordinate.notes

Source of coordinate information and any other relevant information about the coordinates.

design

- **Allochronic**: same population/different time points, longitudinal design.
- **Synchronic**: populations with known divergence time.

genphen

It distinguishes studies measuring individuals that developed in nature (**Phenotypic**) from those based on measurements of individuals bred in common-garden conditions or applying other quantitative-genetic methods (**Genetic**). Phenotypic does not mean it is not genetic; we just do not have the information and it does not automatically imply plasticity. Quantitative genetic methods (heritability does not count) are usually based on animal model methods and calculate breeding values which is the genetic contribution to a trait (Wilson et al. 2010). Common garden experiments performed from vegetative (clonal) propagules collected in the field (e.g. tiller, an anatomic unit in grasses) are still "Phenotypic".

years

If the study is "design"="Allochronic", the number of years from the start time point to the end time point. If the study is "design"="Synchronic", the number of years from the divergence point to the date of sampling. If it is less than one year, enter a decimal (number days of study/365).

gtime

Generation time expressed in years. There are several definitions of the generation time. For our purposes, we define the generation time as the time (in years) in which a population replaces itself.

- For species with non-overlapping generations or species whose individuals reproduce only once in life (e.g. monocarpic plants), the average age at maturity is appropriate.
- For species with non-overlapping generations or species whose individuals reproduce only once in life, with the average age at maturity shorter than one year but experiencing a period of non reproduction in the year (e.g. diapause in mosquitoes, plants with dormant seeds), the generation time was computed as the reciprocal of the number of generations per year (1/GPY). If the generation time is less than one year, the number will be a decimal (# days/365 or # months/12).
- For species with overlapping generations, individuals reproducing multiple times in life and age-structured populations, the age at which members of a given cohort are expected to reproduce is an accurate proxy. It can be computed from a life table as:

$$GT = \frac{\sum_{i=1}^n x_i \cdot l_{(x_i)} \cdot m_{(x_i)}}{\sum_{i=1}^n l_{(x_i)} \cdot m_{(x_i)}}$$

Where x_i is the age (or the mean for an age class), $l_{(x_i)}$ is the proportion of survivors (from the age=0) to a given age (x_i), and $m_{(x_i)}$ is the average fecundity of the individuals at a given age (x_i).

A similar measurement is the average difference in age between parent and offspring, but it requires a stable age structure in the population.

- For species with overlapping generations, individuals reproducing multiple times in life and stage-structured populations, the time it takes for the population to grow by a factor of its net reproductive rate (R_0) is an accurate proxy. It can be computed as:

$$GT = \frac{\ln(R_0)}{r}$$

such as:

$$e^{r \cdot GT} = R_0$$

Where $r = \ln(\lambda)$, and λ is the finite population growth rate. These parameters can be computed from the transition matrix.

Ultimately, the generation time is used to compute the number of elapsed generations between Sample 1 and Sample 2. Ideally, we want the specific generation time for the population (because the generation time can vary among populations). When the authors from the primary source of information provide an estimation of the generation time, or number of generations per year or total number of generations, this information was used. If this is unavailable, the following steps can be taken:

- Query the author directly (see contacting authors).
- Web of Science or Google Scholar search looking for the generation time.
- For marine species, consult the "The Marine Life Information Network" (<https://www.marlin.ac.uk/>). Then, search for your species, in the "Life history" table you might find the "generation time".
- An option for fish is to consult the FishBase (<https://www.fishbase.se/search.php>).
- The databases COMPADRE (<https://compadre-db.org/>) and COMADRE (<https://compadre-db.org/Data/Comadre>) provide population transition matrices for plants and animal species respectively. From these matrices, the generation time can be computed as the time it takes for the population to grow by a factor of its net reproductive rate by using the R package "popbio" (Stubben C & Milligan B. 2007. Estimating and analyzing demographic models using the popbio package in R. Journal of Statistical Software 22(11): 1-23). Generation time from Caswell eq. 5.73:

$$GT = \frac{\ln(R_0)}{\ln(\lambda)}$$

Example using "popbio" in R:

```
#Fmtr: fecundity matrix
#Umtr: survival matrix
```

```

#Amtr: (full) transition matrix

#R0:
Rmtr <- Fmtr%%fundamental.matrix(Umtr)$N
R0 <- eigen(Rmtr)$value[1]

#λ:
lmd <- pop.projection(Amtr, rep(100, nrow(Amtr)), 100)$lambda

#GT:
log(R0)/log(lmd)

```

- If the generation time, a table of life, a transition matrix, or the demographic information to build a table of life are not available, you can look for detailed information on the life cycle and the time and probability for each stage and transition. Based on this information you can produce an educated guess of the actual generation time.

- Look at the Encyclopedia of Life website (<http://eol.org/>). Then, (i) search for your species, (ii) click on read full entry, (iii) look for the box on reproduction – there might be information on age at maturity. For animals with a long gestation period (e.g. bears), sum this gestation period to the age at maturity. Then, the generation time is the earliest age when they give birth, not when they become sexually mature. Even this is (most likely) an underestimation of the actual generation time.

- The supplementary table "PROCEEDv6_species_GT.csv" summarizes all the information about generation time for each species, compiled in the current version of PROCEED.

gtime.notes

Source of the "gtime" and any other comment related to "gtime" and "generations".

generations

The number of generations that have elapsed given the number of years that have elapsed for a given system. It can be calculated as: years/gtime.

trait_description

Trait being quantified in the words of the authors of the paper. If the name of the trait is not self descriptive, please include the definition from the paper (e.g. shell slenderness ($\log_{10}[\text{width}] / \log_{10}[\text{length}]$)). If units are given, please include. Each trait has its own row.

Roff_Mousseau

This is an old variable kept from previous versions of the database. For future versions, there is no need to code new data entries according to this variable. The current version of this variable is "trait_type". The last definition of "Roff_Mousseau" was:

Class of traits as determined by Roff and Mosseau 1987.

- *Morphological: e.g. body size, wing size, other metric characters.*
- *Behavioural: e.g. alarm reaction, activity level, sensitive to conditioning.*
- *Physiological: e.g. oxygen consumption, resistance to heat stress, body temperature.*
- *Life-history: Fecundity, phenology, viability, survival, development rate, "directly related to fitness"*

Kingsolver_Diamond

This is an old variable kept from previous versions of the database. For future versions, there is no need to code new data entries according to this variable. The current version of this variable is "trait_type". The last definition of "Kingsolver_Diamond" was:

"Class of traits as determined by Kingsolver and Diamond 2011. Current classifications are:

- *size: Overall aspect of body size (body size, length, mass, PCs for body mass). Size proxies are NOT included here (e.g. tarsus length, intercranial-length, etc.)*
- *othermorphology: all remaining morphological traits. This includes morphological components that are used as proxies for size (e.g. tarsus length, intercranial-length, etc.)*

- *phenology*: the timing of life-history events (e.g., photoperiod, initial date, peak date, or duration of events such as germination, flowering, laying date, and hatching date)

- *otherLH*: all remaining life history traits (growth rate, offspring size, # of offspring, survival), if biotic interactions (e.g. competition), then classified as *otherLH*, if abiotic stressor, classify as *physio*. For herbivory, see *physio* – will depend on if it is resistance or tolerance to herbivory.

- *behaviour*

- *physio*: metabolism, bite force, growth in stress (abiotic factor), resistance to herbivory (e.g. a chemical defense; tolerance will be based on trait), diet: trophic level, diet, food type, if both abiotic and biotic interactions, classify as *physio*.”

trait__type

Types of traits based on Kingsolver and Diamond 2011 but with some additional categories. Current categories are:

- **phenology**: the timing of cyclic and seasonal life-history events (e.g. flowering date, laying date, and hatching date). Do not confound with "otherLH". For example, the time spent in a given stage of the life cycle (e.g. larval stage) or the time to reach a given stage of the life cycle (e.g. time to maturity, time to first flower) are life-history traits, not phenology. "new.data.type" for "phenology" traits is typically a date while for "otherLH" it is time.

- **behaviour**: "Broadly speaking, animal behavior includes all the ways animals interact with other members of their species, with organisms of other species, and with their environment. Behavior can also be defined more narrowly as a change in the activity of an organism in response to a stimulus, an external or internal cue or combination of cues." (<https://www.khanacademy.org/science/ap-biology/ecology-ap/responses-to-the-environment/a/intro-to-animal-behavior>)

- **physio**: traits related to physiology or a biochemical parameter (e.g. enzymatic activity). Example physiology traits include metabolism, bite force, leaf toughness, decomposition rate, stoichiometry (e.g. C:N), concentration or amount (yield) of a given chemical compound in the body (or part of the body), diet (trophic level, diet, food type), severity or level of affectation by a

disease/pathogen/parasite/herbivore (e.g. Fungal infection % of plants, % leaf consumed), dose/concentration/time exposed to a given substance/condition (e.g. pesticide, drought, light, temperature, etc) needed to produce some effect (e.g. death, flowering, etc.), performance of a standard parasite/herbivore/competitor feeding/competing on/with the focal population (e.g. Pupal weight of the herbivorous *Trichoplusia ni* feeding on native vs introduced populations of *Eschscholzia californica*; or Leaf feeder (*Galerucella pusilla*, L1-Imago) survival, allelopathic effects of *Centaurea maculosa*, *Achillea* biomass ratio with-competition to without-competition).

- **growth**: a measurement of the change in size (or a proxy of size) in a time period (where the starting point is not the birth, germination or equivalent), or a growth rate (irrespective of the starting point). It can be a relative or an absolute measurement of growth or growth rate. Irrespective of the environmental condition (pollution, stress, competition, herbivory/parasites, fertigated/non-fertigated, etc.), the measurements of growth or growth rate are classified as "growth". If the trait is a change in growth (or another performance measurement) see "response".

- **size**: Total body mass, or an overall aspect/measurement of body size as a whole (e.g. total length, plant height, thorax width, plant diameter, basal area, number of tillers (grasses)). Size of parts of the body (e.g. tarsus length, intracranial length, length of the longest leaf, etc.) are not included here but in "othermorphology". In plants, measurements of the whole aboveground size (e.g. aboveground biomass, plant height, plant diameter or basal area, etc.), as well as measurements of the whole belowground size (i.e. root mass) are included as "size". Other morphological features of roots (e.g. length, diameter, etc.) are classified as "othermorphology". The "stem diameter" of tree species is a measurement of "size" but in herbaceous plants, it is "othermorphology". "size" is not necessarily the final or average body size in a population, the measurements of size at a given age/stage (but not at birth/hatch or equivalent) are also classified as "size" (e.g. Body mass lambs (kg)). Note that, similar situations can be classified differently according to the operational decision of the authors of the primary source of information, e.g.: in a common garden experiment for a short period of time if the authors report final biomass, the trait is classified as "size", but if the authors report growth rate (biomass over time) it is classified as "growth". If the trait is a change in size (or another performance measurement) see "response".

- **othermorphology**: all remaining morphological traits. This includes the size of parts of the body (e.g. tarsus length, intercranial-length, etc.) or plant (e.g. leaf size). The size of a part of the body at a given age/stage is also "othermorphology". It is independent of the environmental condition of the measurement, i.e. a measurement of tarsus length (or leaf size, etc) is always "othermorphology" independently of if it was measured in a controlled environment or in a polluted environment (or in another condition), and so on.

- **otherLH**: all remaining life history traits that are not classified as phenology or growth (e.g., offspring size, # of offspring, fecundity, survival, parental investment, reproductive effort/allocation, time to reach maturity or any stage in the life cycle). It is independent on the environmental condition of the measurement, i.e. a measurement of survival (or fecundity, etc.) is always "otherLH" independently of if it was measured in a controlled environment or in a polluted environment (or in another condition), and so on. If the trait is a change in survival (or another performance measurement) see "response".

- **response**: response to a number of abiotic and biotic factors. It is the change, relative change, or ratio of a measurement of performance (size, growth, survival, fecundity, etc) between two conditions (e.g. Size reduction—difference in size between pH 7 and 4, relative to size in pH 7—; Leaf Cpi—[no.leaves_woc - no.leaves_wc]/no.leaves_woc—; Mycorrhizal response [inoculated_biomass-Uninoculated_biomass]/Uninoculated_biomass).

data.type

This is an old variable kept from previous versions of the database. For future versions, there is no need to code new data entries according to this variable. The current version of this variable is "new.data.type". The last definition of "data.type" was:

“Old classification where categories are: Area, Concentration, Count, Count Residuals, Date, Index, Length, Mass, Other, Proportion, Rate, Speed, Temperature, Time, Volume, Weight.”

new.data.type

The current categories are:

- **linear**: a measurement of length with a dimensionality of 1. A ratio between lengths (or a length over another variable) is not a "linear" variable because its dimensionality is not 1. A ratio between linear dimensions (e.g. length over width, or height over diameter), or a linear dimension over another variable (e.g. specific root length (SRL) [mm/mg]) is not "linear" because its dimensionality is not 1 (e.g. $SRL = [1D]/[3D]$, dimensionality -2).
- **area (2D)**: a measurement of area with a dimensionality of 2. A ratio between areas, or an area over another variable (e.g. specific leaf area (SLA) is not "area (2D)" because its dimensionality is not 2 (e.g. $SLA = [2D]/[3D]$, dimensionality -1).
- **cube (3D)**: a measurement of volume or mass or amount of matter with a dimensionality of 3. A ratio between volumes/mass, or a volume/mass over another variable (e.g. "Stem mass ratio (stem mass / whole plant mass)" or "Rhizome mass over length") is not "cube (3D)" because its dimensionality is not 3 (e.g. stem mass / whole plant mass = $[3D]/[3D]$ is a dimensionless proportion; Rhizome mass over length = $[3D]/[D]$, dimensionality is 2 but it is not an "area (2D)" because it depends on the shape of the organ and the density of the tissue, so it is "other").
- **count**: an amount of discrete units (e.g. number of leaves, clutch size, fecundity when number of offspring or propagules, etc.).
- **proportion**: the ratio between a part over the total (including that part). It can be expressed as a proportion (0-1), as a percentage (0-100). Some concentrations are also proportions, concentrations expressed as % M/M, % V/V, ppm, mg/g, ug/g (or similar). When the ratio is not a part over the total, this is not a "proportion" (e.g. root to shoot biomass, reproductive over somatic/vegetative biomass), this is a dimensionless ratio (i.e. "ad_ratio").

SPECIAL CASES OF PROPORTION:

Molar concentration or $\mu\text{mol/g}$ were also classified as "proportion". These are not exactly proportions because the numerator and denominator are not in the same units but are still a measurement of the ratio between a part over the total.

A ratio between a given amount in a stressing/disturbing condition (e.g. biomass in a toxic environment or in presence of parasite/herbivore or competitors) over a control condition expected to express the maximum potential (e.g. biomass in a non-toxic environment or in absence of parasite/herbivore or competitors), was also

considered a "proportion". Similarly, the relative difference between these conditions was also classified as "proportion", because it is a proportional change (e.g. tolerance to herbivory: (damaged plant biomass - undamaged plant biomass) / undamaged plant biomass). The difference is that in the first situation "transf_data"="raw" and in the last one "transf_data"="resid".

- **time**: the time needed for some process or event.
- **date**: a temporal landmark for a given event in a reference system with an arbitrary starting point (e.g. number of days from January 1st). A time in the day (e.g. hours from midnight) is the same type of trait.
- **temperature**: temperature in the scale and units reported in the primary source of information (e.g. °C, °K, °F).
- **rate**: a measurement of speed, an amount over time (e.g., growth rate, metabolic rate (cal/s), etc.). "rate" is a diverse category because the numerator can be "count" (e.g. eggs/day), "linear" (e.g. Growth (mm/day)), "proportion" (Grow(%/d)), "cube (3D)" (e.g. growth rate (pupal mass divided by preadult developmental period) OR Resting Metabolic Rate (mg O₂ h⁻¹)), "ad_ratio" (e.g. $RGR = \ln[m_2/m_1]/\text{time}$), "other" ratios (e.g. Stomatal Conductance (mmol m⁻² s⁻¹)), etc. However, even when a trait is called "rate" in the primary source of information is not necessarily a rate according to the current classification (e.g. "pregnancy rate", "survival rate", "germination rate" are "proportion"; "Unit leaf rate ($\ln [\text{mass_t1} / \text{mass_t0}]$)" is an "ad_ratio"; etc.).
- **ad_ratio**: the ratio between two independent amounts with the same dimensionality is a dimensionless ratio. It is not a proportion because the numerator is not a component or fraction of the denominator (e.g. shoot:root ratio, C:N ratio, reproductive over vegetative biomass).
- **index**: ordinal variables and their derivations. However, when a trait is called "index" in the primary source of information it is not necessarily an index according to the current classification. For example, "index of ear size" ($[(EL \times EW)^{0.5}]$) is a called index but it is quantitative variable proportional to the ear size, its dimensionality is 1. We classified this trait as "other". Another example is "Gonadosomatic index (ova mass over body mass)" this is a "proportion" even when called index.

- **other**: other ratios and other quantitative variables not meeting the previous descriptions.

"new.data.type" is particularly relevant when computing and comparing darwins (D) among different data types. The rate of change darwins is the relative (to mean) rate of change of a given variable (X) at a given time (dt) (Haldane 1949 Suggestions as to quantitative measurement of rates of evolution. *Evolution* 3(1): 51–56). Darwins is:

$$(1) \quad D = \frac{d\left(\frac{dX}{X}\right)}{d(t)}$$

by integrating $1/X$ on dX we obtain:

$$(2) \quad D = \frac{d(\ln(X))}{d(t)}$$

In a discrete interval of time (t), we get:

$$(3) \quad D = \frac{\ln(x_2) - \ln(x_1)}{t}$$

The dimensionality of the trait affects darwins. As an example, if we have three traits total length (l), body mass (m), and mass of gonads over total body mass (p). The relation between these three traits will depend on the allometric relationships, but it would be approximately as follows:

$$m = a \cdot l^3$$

$$p = \frac{b \cdot l^3}{a \cdot l^3} = \frac{b}{a}$$

So, the rate of change in darwins for these traits would be:

$$D(l) = \frac{\ln(l_2) - \ln(l_1)}{t}$$

$$\begin{aligned}
D(m) &= \frac{\ln(m_2) - \ln(m_1)}{t} \\
&= \frac{\ln(a \cdot l_2^3) - \ln(a \cdot l_1^3)}{t} \\
&= \frac{\ln(a) + \ln(l_2^3) - \ln(a) - \ln(l_1^3)}{t} \\
&= \frac{3 \cdot \ln(l_2) - 3 \cdot \ln(l_1)}{t} \\
&= \frac{3 \cdot (\ln(l_2) - \ln(l_1))}{t} \\
&= 3 \cdot D(l)
\end{aligned}$$

If the allometric relationships remain constant:

$$\begin{aligned}
D(p) &= \frac{\ln(p_2) - \ln(p_1)}{t} \\
&= \frac{\ln\left(\frac{b}{a}\right) - \ln\left(\frac{b}{a}\right)}{t} \\
&= 0
\end{aligned}$$

As was shown in the previous example, the type of variable, by itself, strongly affects the estimation of evolutionary change in darwins.

transf_data

If the trait measurement has been transformed and the means and standard deviations were computed on basis of the transformed data (e.g. body mass expressed as log(g) instead of grams, so the mean is the average of the logarithm of the body mass). The current categories are:

- **raw**: mean and sd computed from raw data (no transformation).
- **ord**: the trait is the score in the axis of an ordination analysis (i.e. PCA, CVA).
- **arcsin**: arc sin (\sin^{-1}) transformation.
- **arcsin.sqr**: angular transformation, i.e. $2 \cdot \sin^{-1}(\text{sqrt}(x))$.
- **resid**: residuals of regression or other statistical model ($y - \hat{y}$), centered variables ($y - \bar{y}$), absolute ($Y_t - Y_c$) or relative difference ($(Y_t - Y_c)/Y_c = Y_t/Y_c - 1$) between a treatment and a control.

- **ln**: mean and sd computed from natural log-transformed data. We also include data like $\ln(x)/\ln(y)$ and $\ln[x/y]/t$
- **log10**: mean and sd computed from log10-transformed data. We also include data like $\log_{10}(x)/\log_{10}(y)$ and $\log_{10}[x/y]$.

data.scale

- **ratio**: constant interval with a precise zero.
- **interval**: constant interval with an arbitrary zero (e.g. temperature, time of day, rescaled data: Principal components or discriminate functions).
- NA (**empty cells**): undefined scales and ordinal scales (i.e. "new.data.type"="index") because ordinal scales do not have a constant interval.

Darwins cannot use interval data, but Haldanes can (because it's rescaled units, essentially removing an absolute zero point). Ratio data are actually interval data, but with a clearly defined zero.

n1

Sample size for sample 1.

n2

Sample size for sample 2.

mean1

Mean of the phenotypic trait for a given system for sample 1.

mean2

Mean of the phenotypic trait for a given system for sample 2.

sd1

The standard deviation around the mean of the phenotypic trait for sample 1. Can be calculated manually if necessary.

sd2

The standard deviation around the mean of the phenotypic trait for sample 2. Can be calculated manually if necessary.

out_in

Population(s) outside or inside its native range. Of your two data points (samples) to calculate the rates – if one of the datapoints is outside the native range, it is considered “out”. Natural range expansions are considered “in”. You can have an introduction within the native range which would be “in”. Invasive species, with range expansions, would be “out”.

urban_disturbance

Category used in Alberti et al. 2017 PNAS that classified urban disturbance category (decision tree in Figure S2):

- **Biotic:** "biotic interactions stem from introductions, and are subcategorized depending on the study organism's ecological role: introduced species, or species in its native range responding to an introduction. Introduced species were further subdivided into instances where the introduced species adapts immediately following introduction, vs adaptation after range expansion (post introduction)".
- **Habitat Mod:** "changes due to climate change in general, modification of the landscape, or pollution".
- **Heterogeneity:** "heterogeneity may refer to micro-habitat or micro-climate, and can refer to heterogeneity in space or time".
- **Novel:** "Novel disturbances were defined as those disturbances to which organisms respond to with novel adaptations, such as the rapid evolution of zinc tolerance".
- **Social:** "Social interactions refers to disturbances with are direct and intentional result of human agency".

driver

This is an old variable kept from previous versions of the database. For future versions, there is no need to code new data entries according to this variable. The

current version of this variable is "new_disturbance". The last definition of "driver" was:

“This is the classification used for the Evo. Applications world map (Palkovacs EP, Kinnison MT, Correa C, Dalton CM & Hendry AP. 2012. Fates beyond traits: Ecological consequences of human-induced trait change. Evolutionary Applications 5(2): 183–191). The categories are: (1) Climate change, (2) Harvest, (3) In situ natural variation, (4) Introduction, (5) Introduction of predator/prey/host/competitor, (6) Landscape change, (7) Pollution, (8) Range expansion after introduction, (9) Self induced range expansion.”

old_disturbance

This is an old variable kept from previous versions of the database. For future versions, there is no need to code new data entries according to this variable. The current version of this variable is "new_disturbance". The last definition of "old_disturbance" was:

It is a classification of the environmental change (or driver/disturbance) driving the phenotypic change. Definitions from Hendry et al. 2008 Mol. Ecol.:

- *Introduction: “when humans transferred a species to a new geographical location, and comparisons were then made between introduced and ancestral populations.”*

- *In situ natural variation: “when established populations were not subject to an obvious human impact. Most studies in this category involve the long-term monitoring of natural populations, such as Darwin’s finches of the Galápagos (Grant & Grant 2002, 2006).”*

- *In situ anthropogenic disturbance: “when established populations were subject to an obvious human impact. Examples here include harvesting, anthropogenic acidification, localized thermal inputs, and point-source chemical pollution.”*

- *Self induced range or host expansion: “when new populations were established without any direct human influence, and comparisons were then made among the new populations or among the new and ancestral populations. Examples include birds colonizing new habitats within their native range (Yeh & Price 2004) or new islands outside their native range (Clegg et al. 2002).”*

- *Introduction of a new host: “when humans transferred an exotic host species into the range of a native species, with comparisons then made between populations on the native and introduced hosts. Examples in our database are currently limited to native phytophagous insects colonizing introduced host plants (e.g. Carroll et al. 2005).”*
- *Range expansion after introduction: “when humans introduced a species to a new geographical location, and the species then spread on its own accord to occupy multiple sites. Comparisons were here made among the self-colonized populations.”*

new__disturbance

It is a classification of the environmental change (or driver/disturbance) driving the phenotypic change, based on the categories used by Hendry 2008 (Hendry AP, Farrugia TJ & Kinnison MT. 2008. Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology* 17(1): 20–29) and Palkovacs 2012 (Palkovacs EP, Kinnison MT, Correa C, Dalton CM & Hendry AP. 2012. Fates beyond traits: Ecological consequences of human-induced trait change. *Evolutionary Applications* 5(2): 183–191). The current categories are:

- **Climate change:** this disturbance can often be difficult to clearly identify. A disturbance is classified as "Climate change" when the primary source of information explicitly considers climate change as a potential effect, or when the driver of the system is a climatic event (e.g. *Brassica rapa* (field mustard) changed flowering time after drought in Southern California, USA). Experimental systems manipulating CO2 levels were considered "Pollution" (See "Pollution" definition).
- **Hunt__harv:** any case of hunting or harvesting of a species by humans.
- **Introduction:** “when humans transferred a species to a new geographical location, and comparisons were then made between introduced and ancestral populations”. Also, range expansion after introduction (i.e. “when humans introduced a species to a new geographical location, and the species then spread on its own accord to occupy multiple sites. Comparisons were here made among the self-colonized populations”).
- **Pollution:** an environmental change is classified as "Pollution" when chemical substances (e.g. heavy metals, pesticides, etc) are added or released into the environment contaminating, dirtying, and/or making harmful an area (usually air, water, or soil) where the focal species occur. The environmental change resulting

from changing the natural concentration of substances present in the environment (e.g. a change in the level of CO₂ in the air or a change in the level of nutrients in the soil) or a physical parameter (e.g. temperature, pH, etc) is also categorized as "Pollution" even if not harmful to the species being studied.

- **Landscape change:** It is a habitat modification (different from "Pollution" and the introduction of a new species) in the patches occupied by the population(s) being studied. It can be the creation/restoration of new patches/habitats for the population(s) (e.g. a newly formed freshwater pond for an aquatic species, newly restored habitat), changes in the connectivity among patches/habitats (e.g. the construction of a dam in a river or the construction of a fish ladder, level of fragmentation of the habitat), or a change in the land-use/quality of the patches (e.g. forest clearing, livestock/herbivores inclusion/exclusion, newly irrigation system).

- **Response to introductions:** when a species (native or alien) is responding to the introduction of a new species. The newly introduced species can be a competitor, a predator, a parasite, a new host or prey, etc. The comparison is established among populations (or times, according to the design) of the responding species (not the newly introduced species). When the responding species is native, "out_in"="In"; when it is alien, "out_in"="Out".

- **Other:** It is "in situ natural variation" and "Self induced range or host expansion", i.e. when native populations were not subject to an obvious human impact, most studies in this category involve the long-term monitoring of natural populations, such as Darwin's finches of the Galápagos (Grant & Grant 2002, 2006); also, when new populations were established without any direct human influence, and comparisons were then made among the new populations or among the new and ancestral populations (Examples include birds colonizing new habitats within their native range—Yeh & Price 2004—or new islands outside their native range—Clegg et al. 2002—).

env_change

- **o:** ongoing environmental change. The environmental change has been occurring through time and the phenotypic change was measured in a temporal window within this process (but not from the starting point).

- **n**: novel environmental change. The environmental change has a defined starting point and phenotypic change was measured from this point.

This category applies only to allochronic studies, synchronic are always "env_change"="n" (because the proposed ancestral state is placed previous to the drivers happening). When "new_disturbance"="Other", if the driver it is unknown, "env_change"="NA" (empty cell). But in some cases the driver of change is known, like "self induced range expansion", also experimental systems can represent "in situ natural variation" but from a known starting point. In these cases, the value of "env_change" can be chosen.

Introductions

This is an arbitrary way of ranking systems with more than one disturbance/driver. It ranks from 0 to 5, with 0 being the disturbance not pertaining to the system at ALL, and 5 being very important to the system. The current method for ranking is:

- 1 - Range expansions without invasions
- 2 - Range expansions with invasions (but invasions not specified)
- 3 - paper specifically investigates response to an invader or introduced species
- 5 - paper specifically investigates an invader or introduced species

Other

This is an arbitrary way of ranking systems with more than one disturbance/driver. It ranks from 0 to 5, with 0 being the disturbance not pertaining to the system at ALL, and 5 being very important to the system. The current method for ranking is:

- 1 - Paper specifically investigates climate change
- 5 - Long term monitoring, even for commercially exploited species (e.g. salmonids), but not specifying hunting/harvesting
- 5 - Paper was in situ natural variation

resp_introduced

This is an arbitrary way of ranking systems with more than one disturbance/driver. It ranks from 0 to 5, with 0 being the disturbance not pertaining to the system at ALL, and 5 being very important to the system. The current method for ranking is:

1 - authors note that invasions happened and could be a source of change, but do not explicitly consider it (do not add number to invasions, but do add to landscape change)

4.1 - response to removal of invaders

5 - response to introduced organisms

Hunting_Harvesting

This is an arbitrary way of ranking systems with more than one disturbance/driver. It ranks from 0 to 5, with 0 being the disturbance not pertaining to the system at ALL, and 5 being very important to the system. The current method for ranking is:

1 - Looking at herbivory in plants

1 - predation on organisms was specifically considered, but natural predation

2 - predation on organisms was specifically considered, but introduced predation

2.1 predation on organisms specifically considered but REMOVAL of predators due to natural causes

3.1 predation on organisms specifically considered but REMOVAL of predators due to human influences

4 - Species is a commercially or recreationally exploited species (e.g. salmonids), but the study is not specifically considering hunting/harvesting

5 - Paper looks directly at targeted hunting or harvesting by humans

landscape_change

This is an arbitrary way of ranking systems with more than one disturbance/driver. It ranks from 0 to 5, with 0 being the disturbance not

pertaining to the system at ALL, and 5 being very important to the system. The current method for ranking is:

- 1 - Paper investigates range expansions w/o invasions
- 2 - Paper investigates invasions, but not range expansion
- 2.1 - paper investigates removal of organisms, but not range expansion
- 3 - Paper investigates range expansions w/ invasions (this includes range expansion into more lakes (for fish) via introduction)
- 4 - Paper investigates pollution/biocontrol/urbanization
- 5 - Paper specifically investigates landscape change

Pollution

This is an arbitrary way of ranking systems with more than one disturbance/driver. It ranks from 0 to 5, with 0 being the disturbance not pertaining to the system at ALL, and 5 being very important to the system. The current method for ranking is:

- 1 - pollution is considered as potential cause of observed phenotypic changes
- 2 - historical fluxes in a pollutant are mentioned in article (e. g. copper flux due to historical copper mining, eutrophication)
- 5 - Paper specifically investigates pollution/biocontrol

climate_change

This is an arbitrary way of ranking systems with more than one disturbance/driver. It ranks from 0 to 5, with 0 being the disturbance not pertaining to the system at ALL, and 5 being very important to the system. The current method for ranking is:

- 1 - Paper was in situ natural variation
- 2 - Paper does not mention climate change but talks about environmental variation that can be attributed to climate change (e.g. fluctuation in annual precipitation)
- 2.1 - time series for commercially exploited species (e.g. salmonids)

- 3 - Paper specifically investigates range expansion (but not invasions specifically, but if invaders expanded, goes here)
- 4 - Paper looks at thermal tolerance (natural or induced, e.g. warm water effluent from power plants)
- 5 - Paper specifically investigates climate change

time.series

Point out if this case is related to a time series that can be included in the TimeSeriesDB of PROCEED.

tserieshandling

It states how the time series was handled to take two samples to add to the RatesDB.

comments

Specify the source of information for each rate (e.g. data source within the paper: figure 3a). Other notes that might be useful to someone looking at the database.

DOI

Digital identifier number.

reference

This is the reference of the primary source of information, which is given in the following format (all one line): Last name of the first author and year, Journal (standard abbreviations) Volume: Page numbers or article identifier (e.g. Cox 1997, Can. J. Fish. Aquat. Sci. 54: 1159-1165).

reference_comments

Enter any comments related to the primary source of information.

copyright

If the authors of the manuscript do not want us to share their data, indicate it here. Or any other information with regards to copyright.

who

Who has entered the data (your name).

pending_issues

Any unresolved issues that need to be addressed.