

PROCEED v6.0

Phenotypic Rates of Change Evolutionary and Ecological Database

Species Genome — 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.

Technical

Language: English

Decimal point="."

Not available data (NA)="NA"

Description, Inclusion/Exclusion criteria

This is a complementary table of PROCEEDv6.0 (Phenotypic Rates of Change Evolutionary and Ecological Database (PROCEED). Version 6.0. Usar repositorio estable Brock y citar segun sus reglas. DOI: XXXXXXXXXXXX)

Variables:

group

A grouping factor coarser than "taxa". Current categories are: **invert** (invertebrates), **vert** (vertebrates), and **plant** (plants).

taxa

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

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".

GS_source

Source of information (citation) for the C-value of a given species.

GS_obs

Any relevant information related to the obtaining of the C-value.

Cvalue_1C_pg

C-value, i.e. amount of DNA, in picograms (pg), in a haplontic (gamete) nucleus.

ploidy

Ploidy level of the species or population, i.e number of chromosome sets in somatic cells of the diplophase (2n).

N_chromosomes_2n

Number of chromosomes in somatic cells of the diplophase (2n).

method

Determination method of the C-value.

- **Fe**: Feulgen microdensitometry
- **FC_X**: flow cytometry
- **FC_PI**: flow cytometry with propidium iodide
- **FC_DAPI**: flow cytometry with 4',6'-diamidinophenylindole
- **FC_EB**: flow cytometry with ethidium bromide
- **FC_MI**: flow cytometry with mithramycin
- **FC_EBO**: flow cytometry with ethidium bromide and olivomycin
- **MDAPI**: microdensitometry with 4',6'-diamidinophenylindole
- **RK**: reassociation kinetics
- **Ch**: Chemical extraction
- **CIA**: unknown
- **FIA**: Feulgen Image Analysis Densitometry
- **SCF**: Static cell fluorometry
- **BFA**: Bulk fluorometric assay
- **BCA**: Biochemical analysis
- **UVM**: Ultraviolet microscopy
- **GCD**: Gallocyanin chrom alum densitometry
- **seq**: full sequencing

congener

Point out if the genome size information belongs to the needed species (i.e. the species in "sp_ncbi"; so, "congener"="no") or if it belongs to a different species from the same genus (so, "congener"="yes").

PLANTS

The main source is:

Leitch IJ, Johnston E, Pellicer J, Hidalgo O, Bennett MD. 2019. Plant DNA C-values Database (Release 7.1). <https://cvalues.science.kew.org/>

Search options and criteria:

Search for: All Plant C-values

Show estimates: All estimates

Show fields: Genus, Species, Subspecies, Chromosome number, Ploidy level, Estimation method, Prime estimate

C-value: 1C(pg)

I choose the prime genome size estimate (i.e. the most consistent value obtained under best-practice methods, as defined by Bennett MD, Smith JB. 1976. Nuclear DNA amounts in angiosperms. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 274:227–274)

If the prime estimation is not linked to ploidy and Chromosome number, I choose the estimations with these data.

For the not found species, I repeated the search looking for C-value: 1C(Mbp)

When a species name is not found in this database, all synonyms listed in The Plant List (<http://www.theplantlist.org/>) were checked in the database.

Search of species not found in C-value database

Scholar Google

Search string:

"Species name"+"DNA content"

ANIMALS

The main source is:

Gregory TR. 2021. Animal Genome Size Database. <http://www.genomesize.com>.

Search options and criteria:

When multiple estimations are available for a given species, those with not specified (NS) method, standard species or cell type were ignored.

All, estimations with fitting method, standard species and cell type were recorded.

In birds variability in genome size is narrow (see Chandler Bruce Andrews Thesis, 2009).

So, when the genome size for a given species is not available, but it is available for some congeners I took these values to average as a proxy of the actual genome size. (only for birds). Belonging to the same genus was checked according taxonomic nomenclature in the NCBI taxonomic database.

To get a preliminary table to load c value in RatesDB

```
setwd("")
datos <- read.csv("PROCEEDv6_species_genome.csv", na.string="NA")
names <- paste(datos$species, datos$sp_ncbi, sep="_")

dup <- duplicated(names)
dup2 <- duplicated(names, fromLast=T)
dupp <- dup+dup2
duppm <- dup-dup2

datos2.1 <- datos[which(dupp==0),
c("group","taxa","species","sp_ncbi","Cvalue_1C_pg")]

datos2.2 <- datos[which(duppm==
1),c("group","taxa","species","sp_ncbi","Cvalue_1C_pg")]
datos2.2$Cvalue_1C_pg <- "REVISAR"
```

```
datos2 <- rbind(datos2.1, datos2.2)

write.csv(datos2, "species_Cval.csv")
```

Then, check out in “species_Cval.csv” file the cases where Cvalue_1C_pg=”REVISAR” to make a decision according to the criteria described below (When more than one value is available for one species).

When more than one 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 citometry (FC_any type), DAPI microdensitometry (MDAPI), Gallocyanin chrom alum densitometry (GDC), or Fulgen image analysis densitometry (FIA) methods were preferred over other methods. So, when available, records using these methods were chosen for averaging.

In birds, if the genome size was estimated from several congeners, the values of the multiple measurements of c-value (if more than one) for each congener were averaged previously, then the average among congeners was computed.

In general, if the genome size was estimated from several subspecies, the value for each subspecies was averaged previously of averaging among them.

In general, but specially in plants, when a species has more than one genome size due to differences in ploidy level or chromosome number, the system was revisited looking for information about particular ploidy levels of the involved populations for each case. If no information was found, the available c-values were averaged.