

Package: rmdd (via r-universe)

May 17, 2026

Title Resolve Mammal Names with the Mammal Diversity Database

Version 0.0.0.1

Description Download, inspect, reconcile, and summarize mammal taxonomic names with the Mammal Diversity Database (MDD). Supports accepted names, synonyms, original combinations, distribution summaries, and mapped outputs derived from packaged MDD releases. Designed for reproducible mammal name resolution workflows in 'R'.

License MIT + file LICENSE

URL <https://github.com/PaulESantos/rmdd>,
<https://paulesantos.github.io/rmdd/>,
<https://www.mammaldiversity.org/>

BugReports <https://github.com/PaulESantos/rmdd/issues>

Depends R (>= 4.3.0)

Imports cli, dplyr, fs, ggplot2, fuzzyjoin, lifecycle, readr, naturalearth, sf, rlang, stringr, tibble, tools, utils

Suggests covr, knitr, lintr, pkgdown, rmarkdown, rnaturalearthdata, roxygen2, testthat (>= 3.0.0)

Config/testthat/edition 3

Encoding UTF-8

LazyData true

LazyDataCompression xz

Roxygen list(markdown = TRUE)

VignetteBuilder knitr

Config/roxygen2/version 8.0.0

Config/pak/sysreqs libabsl-dev cmake libgdal-dev gdal-bin libgeos-dev make libicu-dev libuv1-dev libssl-dev libproj-dev libsqlite3-dev libudunits2-dev libx11-dev

Repository <https://paulesantos.r-universe.dev>

Date/Publication 2026-05-17 04:06:27 UTC

RemoteUrl <https://github.com/PaulESantos/rmdd>

RemoteRef HEAD

RemoteSha 91ad5add22083c5851981bfa46c4f031f2cc4cb

Contents

as.list.mdd_taxon_info	2
build_mdd_match_backbone	3
classify_mammal_names	4
mdd_cache_dir	4
mdd_checklist	5
mdd_checklist_records	8
mdd_distribution_map	9
mdd_distribution_summary	11
mdd_distribution_summary_raw	12
mdd_download	13
mdd_load	14
mdd_matching	14
mdd_name_index	15
mdd_reference	16
mdd_release_records	17
mdd_synonym_records	17
mdd_synonyms	18
mdd_taxon_info	21
mdd_taxon_record	23
mdd_type_specimen_metadata	24
print.mdd_reference	25
print.mdd_taxon_info	26

Index 27

as.list.mdd_taxon_info

Convert mdd_taxon_info to list

Description

Convert mdd_taxon_info to list

Usage

```
## S3 method for class 'mdd_taxon_info'
as.list(x, ...)
```

Arguments

x An object of class `mdd_taxon_info`.
... Additional arguments.

Value

A list representation of the `mdd_taxon_info` object.

`build_mdd_match_backbone`

Build an MDD reconciliation backbone

Description**[Experimental]**

This function is experimental and may change in future releases.

Construct a species-level reconciliation backbone that combines accepted MDD checklist names with synonym names and links every matched name to its accepted species context.

Usage

```
build_mdd_match_backbone(checklist = NULL, synonyms = NULL)
```

Arguments

checklist A data frame like `mdd_checklist`.
synonyms Optional data frame like `mdd_synonyms`.

Value

A tibble used internally by `mdd_matching()`.

Examples

```
bb <- build_mdd_match_backbone(  
  checklist = dplyr::slice(mdd_checklist, 1:10),  
  synonyms = dplyr::slice(mdd_synonyms, 1:20)  
)  
bb
```

`classify_mammal_names` *Classify mammal scientific names into taxonomic components*

Description

[Experimental]

This function is experimental and may change in future releases.

Parse mammal scientific names into genus, optional subgenus, species, optional subspecies, and trailing author text. The parser is intentionally conservative and targets species-level MDD reconciliation workflows.

Usage

```
classify_mammal_names(splist)
```

Arguments

`splist` Character vector of scientific names.

Value

A tibble with one row per input name and standardized columns for species-level reconciliation.

Examples

```
classify_mammal_names(c(  
  "Puma concolor",  
  "Capromys (Pygmaeocapromys) angelcabrerai"  
))
```

`mdd_cache_dir` *Return the default rmdd cache directory*

Description

[Stable]

This function is stable and its interface is expected to remain compatible.

Usage

```
mdd_cache_dir(path = NULL)
```

Arguments

`path` Optional path. If NULL, a user cache directory is constructed.

Value

A scalar character path.

Examples

```
mdd_cache_dir(tempdir())
```

mdd_checklist	<i>Current Mammal Diversity Database checklist</i>
---------------	--

Description

Current species-level checklist from the Mammal Diversity Database (MDD).

Usage

```
mdd_checklist
```

Format

A tibble with 6,871 rows and 52 variables in the MDD v2.4 release, with source column names normalized to snake_case during data import. Variables and their original camelCase equivalents:

sci_name (sciName) Unique genus–epithet key joined by an underscore; derived programmatically from genus and specific_epithet.

id (id) Unique integer identifier used for indexing and permlinking. Initial batch (14 Sep 2020) was numbered from 1,000,001 upward when sorted by phylosort; subsequent additions start at 1,006,485.

phylosort (phylosort) Numeric sort key ordering the 27 extant mammal orders according to the phylogenetic hierarchy in Figure 1 of the *Illustrated Checklist of the Mammals of the World* (2020).

main_common_name (mainCommonName) Primary vernacular name following the *Handbook of the Mammals of the World* style conventions: all words capitalised except pre-hyphen elements; "and" constructions use hyphens (e.g. *Black-and-white Ruffed Lemur*); directional modifiers are not hyphenated (e.g. *Southwestern Myotis*).

other_common_names (otherCommonNames) Pipe-separated (|) list of additional vernacular names documented in the literature, following the same formatting rules as main_common_name; primarily English, but widely used names from other languages are occasionally included.

subclass (subclass) Taxonomic subclass; NA when not applicable.

infraclass (infraclass) Taxonomic infraclass; NA when not applicable.

magnorder (magnorder) Taxonomic magnorder; NA when not applicable.

superorder (superorder) Taxonomic superorder; NA when not applicable.

order (order) Taxonomic order; present for all taxa.

- suborder** (suborder) Taxonomic suborder; NA when not applicable; "incertae sedis" for taxa unassigned within an order that uses suborders.
- infraorder** (infraorder) Taxonomic infraorder; NA when not applicable; "incertae sedis" when applicable but unassigned.
- parvorder** (parvorder) Taxonomic parvorder; NA when not applicable; "incertae sedis" when applicable but unassigned.
- superfamily** (superfamily) Taxonomic superfamily; NA when not applicable; "incertae sedis" when applicable but unassigned.
- family** (family) Taxonomic family; present for all taxa.
- subfamily** (subfamily) Taxonomic subfamily; NA when not applicable; "incertae sedis" when applicable but unassigned.
- tribe** (tribe) Taxonomic tribe; NA when not applicable; "incertae sedis" when applicable but unassigned.
- subtribe** (subtribe) Taxonomic subtribe; NA when not applicable; "incertae sedis" when applicable but unassigned.
- genus** (genus) Taxonomic genus; present for all taxa.
- subgenus** (subgenus) Taxonomic subgenus; NA when not applicable; "incertae sedis" when applicable but unassigned.
- specific_epithet** (specificEpithet) Taxonomic species epithet.
- authority_species_author** (authoritySpeciesAuthor) Author surname(s) of the original species description, sourced from Species_Syn_Current; all authors on the author line are included, with an Oxford comma before the last name when three or more authors are present; an "in" statement is added when the work appears in a volume with different editors; shared surnames are disambiguated by initials or full middle names; Chinese, Korean, and Indochinese names are written out in full with surname first and hyphens removed.
- authority_species_year** (authoritySpeciesYear) Year of the original species description, sourced from Species_Syn_Current.
- authority_parentheses** (authorityParentheses) Parenthesis flag: 0 = no parentheses; 1 = authority in parentheses (indicating the species was originally described under a different genus).
- original_name_combination** (originalNameCombination) Name combination exactly as it appears in the original description, sourced from Species_Syn_Current.
- authority_species_citation** (authoritySpeciesCitation) Full or abbreviated literature citation for the authority publication; APA format when verified by PDF or physical copy, abbreviated otherwise.
- authority_species_link** (authoritySpeciesLink) URL to the authority publication or abstract page; Biodiversity Heritage Library page-level links are preferred when available; DOIs are used for recent publications.
- type_voucher** (typeVoucher) Museum catalogue number(s) of the type series (holotype, syntypes, lectotype, or neotype); blank when type material has not been verified; multiple syntypes are listed when applicable.
- type_kind** (typeKind) Category of type specimen listed in type_voucher: one of "holotype", "syntypes", "lectotype", "neotype", or "nonexistent" (the last value is used when the MDD team has confirmed no type material exists); blank when existence of type material has not been verified.

- type_voucher_ur_is** (typeVoucherURIs) Pipe-separated links to type material records in external museum collection databases.
- type_locality** (typeLocality) Geographic locality where the holotype was collected or observed; edited place names follow CMW 2020 and are updated to current taxonomy; format may not match Species_Syn_Current as standardisation is ongoing.
- type_locality_latitude** (typeLocalityLatitude) Latitude of the type locality in decimal degrees; sourced from the original description or via georeferencing (web search or GeoLocate).
- type_locality_longitude** (typeLocalityLongitude) Longitude of the type locality in decimal degrees; sourced as for type_locality_latitude.
- nominal_names** (nominalNames) Pipe-separated list of all available and unavailable specific epithets subsumed under the current species concept, each with authority and year; names originally described in a different genus are shown in parentheses; reasons for unavailability are noted in brackets; gender changes are also noted.
- taxonomy_notes** (taxonomyNotes) Semicolon-separated notes by MDD staff documenting taxonomic changes.
- taxonomy_notes_citation** (taxonomyNotesCitation) Pipe-separated APA citations supporting the changes described in taxonomy_notes.
- distribution_notes** (distributionNotes) Detailed distributional narrative including notes on recently introduced populations; uses abbreviations: Mt/Mts, I/Is, N, S, E, W, C, and combinations (e.g. NW). **Not currently curated — reserved for future use.**
- distribution_notes_citation** (distributionNotesCitation) Pipe-separated citations supporting distribution_notes. **Not currently curated — reserved for future use.**
- subregion_distribution** (subregionDistribution) Pipe-separated (|) list of countries with native or reintroduced extant populations; subnational regions (currently US states only) are appended in parentheses and comma-separated; uncertain or possibly extinct occurrences are marked with ?.
- country_distribution** (countryDistribution) Pipe-separated list of countries where the species has a native or ancient (pre-1500 CE) introduced distribution; recent introductions are excluded; uncertain presences marked with ?; domesticated species are listed as "Domesticated"; extinct species reflect their post-1500 CE range; marine species list coastal or riverine countries.
- continent_distribution** (continentDistribution) Pipe-separated list of continents following the same rules as country_distribution; recognised values: Africa, Antarctica, Asia, Europe, North America, Oceania, South America, Domesticated; continental boundaries follow MDD-specific definitions (e.g. Asia–Europe split at Ural/Caucasus Mts and Ural River; Oceania east of Weber’s Line).
- biogeographic_realm** (biogeographicRealm) Pipe-separated list of biogeographic realms where the species occurs, following the same inclusion rules as country_distribution; realm boundaries follow the WWF schema (https://en.wikipedia.org/wiki/Biogeographic_realm).
- iucn_status** (iucnStatus) IUCN Red List status matched to MDD species and updated to current taxonomy (IUCN 2025-2 assessment); "NE" = not yet evaluated; standard IUCN acronyms otherwise; 19 domestic species are also not evaluated.
- extinct** (extinct) Extinction flag: 0 = extant; 1 = extinct after 1500 CE (following the IUCN criterion).

domestic (domestic) Domestication flag: 0 = wild; 1 = domesticated (19 species including *Homo sapiens*); follows the nomenclatural guidelines of Gentry, Clutton-Brock & Groves (2004, *J. Archaeol. Sci.* 31(5): 645–651).

flagged (flagged) Taxonomic quality flag: 0 = valid; 1 = flagged as taxonomically questionable or actively debated in the literature.

cmw_sci_name (CMW_sciName) Genus–epithet key from the *Illustrated Checklist of the Mammals of the World* (CMW), formatted identically to sci_name.

diff_since_cmw (diffSinceCMW) CMW comparison flag: 0 = species present in CMW 2020 taxonomy; 1 = species new since CMW 2020.

msw3_matchtype (MSW3_matchtype) Method by which the taxon was linked to MSW3: one of "matched", "unmatched", or "manual".

msw3_sci_name (MSW3_sciName) Scientific name as matched to the MSW3 taxonomy.

diff_since_msw3 (diffSinceMSW3) MSW3 comparison flag: 0 = species present in MSW3 (~2004 cutoff); 1 = species new since MSW3.

Details

The checklist is the primary tabular output of the MDD and covers all extant and recently extinct (post-1500 CE) mammal species recognised by the database curators. Column names have been normalised from the original camelCase (sciName, mainCommonName, etc.) to snake_case during data import; the original names are shown in parentheses in each \item above for cross-reference with the upstream spreadsheet.

Distribution fields operate on three nested spatial scales: subregion_distribution (subnational units, currently US states), country_distribution, and continent_distribution. Multiple values within each field are pipe-separated (|). Only native ranges and ancient introductions (before 1500 CE) are included; ? marks uncertain or possibly extirpated occurrences.

The distribution_notes and distribution_notes_citation columns are present in the data but are **not currently curated** and are reserved for future use.

Cross-release comparison is provided by three column pairs: cmw_sci_name / diff_since_cmw (vs. CMW 2020) and msw3_sci_name / msw3_matchtype / diff_since_msw3 (vs. MSW3, ~2004 cutoff).

Source

Mammal Diversity Database release archive (<https://www.mammaldiversity.org>). Field definitions are derived from the META_v2.4.csv file and the column-level annotations in the MDD_Current sheet of the official MDD spreadsheet.

mdd_checklist_records *Return checklist records in a normalized tibble*

Description

[Stable]

This function is stable and its interface is expected to remain compatible.

Usage

```
mdd_checklist_records(checklist = NULL)
```

Arguments

checklist Optional checklist data frame. Defaults to mdd_checklist.

Value

A tibble.

Examples

```
mdd_checklist_records()
```

mdd_distribution_map *Generate a distribution map from MDD country distributions*

Description**[Experimental]**

This function is experimental and may change in future releases.

Resolve a mammal name with `mdd_matching()` and map its `country_distribution` values against an `rnaturlaearth` countries layer. The function preserves the package's exact vs partial input validation by reporting whether the taxon input was matched directly or through a fuzzy stage before plotting the accepted taxon distribution. Spatial polygons come from Natural Earth through `rnaturlaearth::ne_countries()`.

Usage

```
mdd_distribution_map(  
  name,  
  checklist = NULL,  
  synonyms = NULL,  
  target_df = NULL,  
  atlas = NULL,  
  max_dist = 1,  
  method = "osa",  
  zoom = c("world", "auto", "manual"),  
  xlim = NULL,  
  ylim = NULL,  
  quiet = FALSE,  
  title = NULL,  
  base_fill = "#E8E4DB",  
  dist_fill = "#111111",  
  border_color = "#111111",
```

```

country_color = "#8B887F",
ocean_fill = "#F7F4EE",
plot_fill = "#F7F4EE",
graticule_color = "#BBB5A8",
graticule_linewidth = 0.18,
graticule_alpha = 0.55,
country_linewidth = 0.22,
dist_linewidth = 0.34,
title_color = "#111111",
title_size = 16,
title_face = "bold"
)

```

Arguments

name	A single scientific name.
checklist	Optional checklist data frame. Defaults to <code>mdd_checklist</code> .
synonyms	Optional synonym data frame. Defaults to <code>mdd_synonyms</code> .
target_df	Optional reconciliation backbone from <code>build_mdd_match_backbone()</code> .
atlas	Optional <code>sf</code> object with world country polygons. Defaults to <code>rnaturalearth::ne_countries(scale = "medium", returnclass = "sf")</code> .
max_dist	Maximum string distance used if the input name needs fuzzy reconciliation.
method	Distance method passed to <code>mdd_matching()</code> .
zoom	Zoom mode. Use "world" for the full map, "auto" to zoom to the mapped distribution extent, or "manual" to use <code>xlim</code> and <code>ylim</code> .
xlim	Optional longitude limits used when <code>zoom = "manual"</code> .
ylim	Optional latitude limits used when <code>zoom = "manual"</code> .
quiet	Logical. If TRUE, suppress cli progress messages.
title	Optional plot title. Defaults to the accepted taxon name.
base_fill	Fill color for non-selected countries.
dist_fill	Fill color for mapped distribution units.
border_color	Border color for mapped country outlines.
country_color	Border color for non-selected country outlines.
ocean_fill	Background color for the map panel.
plot_fill	Background color for the full plot area.
graticule_color	Color for graticule lines.
graticule_linewidth	Line width for graticule lines.
graticule_alpha	Alpha level for graticule lines.
country_linewidth	Line width for non-selected country outlines.

dist_linewidth Line width for mapped country outlines.
title_color Color for the plot title.
title_size Size of the plot title.
title_face Font face for the plot title.

Value

A ggplot2 map object. Additional metadata are attached in the "mdd_distribution_info" attribute.

Examples

```
if (interactive()) {
  mdd_distribution_map("Lama vicugna", quiet = TRUE)
}
```

mdd_distribution_summary

Summarize mammal diversity by country, continent, or subregion

Description**[Experimental]**

This function is experimental and may change in future releases.

Wrapper around `mdd_distribution_summary_raw()` that optionally excludes domesticated species and species considered widespread at the requested geographic level.

Usage

```
mdd_distribution_summary(
  level = c("country", "continent", "subregion"),
  checklist = NULL,
  exclude_domesticated = FALSE,
  exclude_widespread = FALSE,
  widespread_threshold = NULL
)
```

Arguments

level Geographic level to summarize. Use "country", "continent", or "subregion".
checklist Optional checklist data frame. Defaults to `mdd_checklist`.
exclude_domesticated Logical. If TRUE, drop rows where `domestic == 1`.
exclude_widespread Logical. If TRUE, drop species whose distribution spans more than `widespread_threshold` units at the chosen level.

widespread_threshold

Optional threshold used to define widespread species. If NULL, a level-specific default is used.

Value

A tibble with one row per geographic unit and the columns region, orders, families, genera, living_species, extinct_species, and total_species.

Examples

```
mdd_distribution_summary(
  level = "country",
  checklist = dplyr::slice(mdd_checklist, 1:50)
)
```

mdd_distribution_summary_raw

Summarize mammal diversity by country, continent, or subregion

Description

[Experimental]

This function is experimental and may change in future releases.

Expand MDD distribution fields and compute diversity summaries by geographic unit. The summary reports counts of distinct orders, families, genera, living species, and extinct species for each unit listed in the selected MDD distribution column.

Usage

```
mdd_distribution_summary_raw(
  level = c("country", "continent", "subregion"),
  checklist = NULL
)
```

Arguments

level Geographic level to summarize. Use "country", "continent", or "subregion".

checklist Optional checklist data frame. Defaults to mdd_checklist.

Value

A tibble with one row per geographic unit and the columns region, orders, families, genera, living_species, extinct_species, and total_species.

Examples

```
mdd_distribution_summary_raw(  
  level = "country",  
  checklist = dplyr::slice(mdd_checklist, 1:50)  
)
```

mdd_download

Download a Mammal Diversity Database file

Description

[Stable]

This function is stable and its interface is expected to remain compatible.

Usage

```
mdd_download(url, dest_dir = mdd_cache_dir(), dest_file = NULL, mode = "wb")
```

Arguments

url	URL to the source file.
dest_dir	Destination directory for the downloaded file.
dest_file	Optional output filename.
mode	Transfer mode passed to <code>utils::download.file()</code> .

Value

The downloaded file path, invisibly.

Examples

```
if (interactive()) {  
  mdd_download(  
    url = "https://www.mammaldiversity.org/robots.txt",  
    dest_dir = tempdir()  
  )  
}
```

mdd_load	<i>Load an MDD comma-separated export</i>
----------	---

Description**[Stable]**

This function is stable and its interface is expected to remain compatible.

Usage

```
mdd_load(path)
```

Arguments

path Path to a local CSV export from MDD.

Value

A tibble.

Examples

```
sample_path <- system.file("extdata", "mdd_sample.csv", package = "rmdd")
mdd_load(sample_path)
```

mdd_matching	<i>Reconcile mammal names against MDD</i>
--------------	---

Description**[Experimental]**

This function is experimental and may change in future releases.

Run a staged species-level reconciliation pipeline against an MDD backbone. The workflow prioritizes exact evidence first, then progressively relaxes criteria through exact genus, fuzzy genus, exact species-within-genus, and fuzzy species-within-genus matching.

Usage

```
mdd_matching(  
  x,  
  target_df = NULL,  
  prefilter_genus = TRUE,  
  allow_duplicates = FALSE,  
  max_dist = 1,  
  method = "osa"  
)
```

Arguments

<code>x</code>	Character vector of names or a data frame with parsed name columns.
<code>target_df</code>	Optional backbone produced by <code>build_mdd_match_backbone()</code> .
<code>prefilter_genus</code>	Logical. If TRUE, restrict the backbone to exact and fuzzy candidate genera before the full pipeline.
<code>allow_duplicates</code>	Logical. If TRUE, deduplicate internally and expand results back to the original rows.
<code>max_dist</code>	Maximum string distance used in fuzzy stages.
<code>method</code>	Distance method passed to <code>fuzzyjoin::stringdist*_join()</code> .

Value

A tibble with row-level traceability, pathway flags, matched name context, accepted-name context, and fuzzy distance columns.

Examples

```
checklist <- tibble::tibble(
  id = c("1", "2"),
  sci_name = c("Puma_concolor", "Vicugna_vicugna"),
  genus = c("Puma", "Vicugna"),
  specific_epithet = c("concolor", "vicugna"),
  authority_species_author = c("Linnaeus", "Molina")
)
synonyms <- tibble::tibble(
  mdd_syn_id = c("1001", "1002"),
  mdd_species_id = c("1", "2"),
  mdd_author = c("Linnaeus", "Molina"),
  mdd_original_combination = c("Felis concolor", "Auchenia vicugna")
)
backbone <- build_mdd_match_backbone(checklist, synonyms)
mdd_matching(
  c("Puma concolor", "Felis concolor", "Pumma concolor"),
  target_df = backbone
)
```

mdd_name_index

Build a normalized MDD name index

Description**[Experimental]**

This function is experimental and may change in future releases.

Create a reusable name index that combines accepted checklist names, checklist original combinations, and synonym names while preserving the accepted-species linkage for each row.

Usage

```
mdd_name_index(checklist = NULL, synonyms = NULL)
```

Arguments

`checklist` Optional checklist data frame. Defaults to `mdd_checklist`.
`synonyms` Optional synonym data frame. Defaults to `mdd_synonyms`.

Value

A tibble with one row per queryable accepted name or synonym.

Examples

```
idx <- mdd_name_index(
  checklist = dplyr::slice(mdd_checklist, 1:10),
  synonyms = dplyr::slice(mdd_synonyms, 1:20)
)
idx
```

mdd_reference

Reference citations for the Mammal Diversity Database

Description**[Stable]**

This function is stable and its interface is expected to remain compatible.

Return the recommended citation for the current MDD Zenodo release and, optionally, format a citation for a specific MDD taxon entry.

Usage

```
mdd_reference(taxon_id = NULL, taxon_name = NULL)
```

Arguments

`taxon_id` Optional MDD taxon identifier for a specific entry.
`taxon_name` Optional scientific name to use in the entry citation. If `NULL` and `taxon_id` is provided, the function tries to recover the accepted name from `mdd_checklist`.

Value

An object of class `mdd_reference` containing dataset-level and, when requested, entry-level citation strings.

Examples

```
mdd_reference()
mdd_reference(taxon_id = "1001892", taxon_name = "Dipodomys deserti")
```

mdd_release_records *Return release records used by rmdd*

Description

[Stable]

This function is stable and its interface is expected to remain compatible.

Usage

```
mdd_release_records(  
  checklist = NULL,  
  synonyms = NULL,  
  type_specimen_metadata = NULL  
)
```

Arguments

checklist Optional checklist data frame.
synonyms Optional synonym data frame.
type_specimen_metadata
 Optional type specimen metadata table.

Value

A named list of tibbles.

Examples

```
x <- mdd_release_records()  
names(x)
```

mdd_synonym_records *Return synonym records in a normalized tibble*

Description

[Stable]

This function is stable and its interface is expected to remain compatible.

Usage

```
mdd_synonym_records(synonyms = NULL)
```

Arguments

synonyms Optional synonym data frame. Defaults to mdd_synonyms.

Value

A tibble.

Examples

```
mdd_synonym_records()
```

mdd_synonyms	<i>Mammal Diversity Database synonym table</i>
--------------	--

Description

Synonymy and nomenclatural table from the Mammal Diversity Database (MDD), covering all names applicable to species- and subspecies-level mammal taxa within Class Mammalia. Designed to align with the online nomenclature database Hesperomys (<https://hesperomys.com>).

Usage

```
mdd_synonyms
```

Format

A tibble with 44 variables in the MDD v2.4 release, with source column names normalized to snake_case during data import. Variables and their original naming equivalents:

mdd_syn_id (MDD_syn_ID) Unique MDD synonym identification number assigned to each synonym, starting from 100000001 and incrementing as new synonyms are added.

mdd_species (MDD_species) Genus and specific epithet of the accepted species to which this synonym is assigned, with a space between the two elements; listed as "genus incertae_sedis" or "incertae_sedis incertae_sedis" for nomina dubia, nomina inquirenda, and names based on composite or hybrid type material.

mdd_species_id (MDD_species_id) Identification number of the accepted species as listed in mdd_checklist (the MDD_Current sheet).

hesp_id (Hesp_id) Identification number of the matching entry in the Hesperomys nomenclature database (<https://hesperomys.com>), used to link and synchronise data between MDD and Hesperomys.

mdd_root_name (MDD_root_name) Specific epithet used as the root to form a valid taxon name; spelling is adjusted to match generic gender when the epithet is an adjective, or Latinised when originally written in non-Latin characters.

- mdd_author** (MDD_author) Author surname(s) of the original description, following the same formatting conventions as `authority_species_author` in `mdd_checklist`: all authors on the author line are included; Oxford comma before the last name when three or more authors are present; an "in" statement is added when the work appears in a volume with different editors; shared surnames are disambiguated by initials or full middle names; Chinese, Korean, and Indochinese names are written with surname first and hyphens removed.
- mdd_year** (MDD_year) Year of the original description as given in the original publication.
- mdd_authority_parentheses** (MDD_authority_parentheses) Parenthesis flag: 0 = no parentheses; 1 = authority in parentheses (indicating the species was originally described under a different genus).
- mdd_nomenclature_status** (MDD_nomenclature_status) Nomenclatural status of the name under the ICZN Code; indicates whether the name is available, a spelling variant, a name combination, or unavailable (with the specific reason given); see `Nomenclature_Taxonomy_Metadata` in the MDD spreadsheet for full value definitions.
- mdd_validity** (MDD_validity) Taxonomic status of the name under the ICZN Code and primary literature; indicates whether the name is a valid species or subspecies, or is a nomen dubium, nomen inquirendum, or based on composite or hybrid type material; see `Nomenclature_Taxonomy_Metadata` for full value definitions.
- mdd_original_combination** (MDD_original_combination) Name combination exactly as it appears in the original description; all parts of the scientific name are written in full even if abbreviated in the source.
- mdd_original_rank** (MDD_original_rank) Taxonomic rank at which the name was first described; one of "species", "subspecies", "form", "variety", "infrasubspecific", "unranked", "synonym", or "other".
- mdd_authority_citation** (MDD_authority_citation) Full APA citation of the original description, including specific month and day of publication when known; populated only when the MDD team has verified the citation by obtaining a PDF or physical copy.
- mdd_unchecked_authority_citation** (MDD_unchecked_authority_citation) Unverified original description citation in non-standardised format, often abbreviated from regional or global taxonomic compendia; set to NA once `mdd_authority_citation` is filled.
- mdd_sourced_unverified_citations** (MDD_sourced_unverified_citations) Same as `mdd_unchecked_authority_citation` but includes the sources from which the citation was obtained; retained as an internal reference for the MDD team.
- mdd_citation_group** (MDD_citation_group) Journal name or city of publication (for books) of the original description; used internally by the MDD team when locating citation materials.
- mdd_citation_kind** (MDD_citation_kind) Whether the MDD team holds a physical, electronic, or no copy of the original description; used internally by the MDD team.
- mdd_authority_page** (MDD_authority_page) Page number where the scientific name first appears in the original description; special cases (footnotes, unnumbered pages, figures, plates) are noted in parentheses.
- mdd_authority_link** (MDD_authority_link) URL to an online copy of the original description; older publications link to the Biodiversity Heritage Library or other archives; newer publications use a DOI.

- mdd_authority_page_link** (MDD_authority_page_link) URL to the exact first page where the name appears in an online copy; primarily applicable to records in the Biodiversity Heritage Library, HathiTrust, Gallica, or the Internet Archive.
- mdd_unchecked_authority_page_link** (MDD_unchecked_authority_page_link) Candidate page links identified programmatically from the Biodiversity Heritage Library but not yet manually verified for inclusion of the name; multiple links separated by |.
- mdd_old_type_locality** (MDD_old_type_locality) Type locality as it appeared in the MDD synonym sheet before harmonisation with *Hesperomys*; retained for internal reference only — not for general use.
- mdd_original_type_locality** (MDD_original_type_locality) Verbatim type locality transcribed from the original description and verified by the MDD team; blank when a type locality is not applicable.
- mdd_unchecked_type_locality** (MDD_unchecked_type_locality) Verbatim type locality from non-original sources, with source attribution appended; multiple entries from different sources separated by |; blank when not applicable.
- mdd_emended_type_locality** (MDD_emended_type_locality) Verbatim emended type locality from sources making a restriction, emendation, or declaration, followed by the reference; **currently empty — reserved for future curation**; blank when not applicable.
- mdd_type_latitude** (MDD_type_latitude) Latitude of the type locality in decimal degrees; geo-referenced from the original description or via web search / GeoLocate; blank when not applicable.
- mdd_type_longitude** (MDD_type_longitude) Longitude of the type locality in decimal degrees; sourced as for mdd_type_latitude; blank when not applicable.
- mdd_type_country** (MDD_type_country) Modern country containing the type locality; see *Distribution_List* in the MDD spreadsheet for the full list of country names used; blank when not applicable.
- mdd_type_subregion** (MDD_type_subregion) First-level subnational unit containing the type locality (state, province, territory, or island group for larger countries); blank when not applicable.
- mdd_type_subregion2** (MDD_type_subregion2) Second-level subnational unit within mdd_type_subregion; primarily US counties and individual islands within archipelagoes; blank when not applicable.
- mdd_holotype** (MDD_holotype) Museum catalogue number(s) of the type series (holotype, syntypes, lectotype, or neotype); multiple syntypes are comma-separated; blank when type material has not been verified.
- mdd_type_kind** (MDD_type_kind) Category of type specimen listed in mdd_holotype: one of "holotype", "syntypes", "lectotype", "neotype", or "nonexistent" (confirmed absence of type material); blank when existence of type material has not been verified.
- mdd_type_specimen_link** (MDD_type_specimen_link) URL(s) to type material records in external museum collection databases.
- mdd_order** (MDD_order) Taxonomic order; "incertae_sedis" when not assigned to an order.
- mdd_family** (MDD_family) Taxonomic family; "incertae_sedis" when not assigned to a family.
- mdd_genus** (MDD_genus) Taxonomic genus; "incertae_sedis" when not assigned to a genus.
- mdd_specific_epithet** (MDD_specificEpithet) Specific epithet of the accepted species; "incertae_sedis" when not assigned to a species.

mdd_subspecific_epithet (MDD_subspecificEpithet) Subspecific epithet of the accepted subspecies; blank when no subspecies are recognised within the species. **Note:** this field currently contains a tentative, unvetted subspecies list — assignments are incomplete and subject to revision.

mdd_variant_of (MDD_variant_of) For spelling variants and name combinations, the name they are a variant of, given with its original name combination, authority, and MDD synonym ID.

mdd_senior_homonym (MDD_senior_homonym) For preoccupied names (primary or secondary homonymy), the senior homonym that preoccupies the name, given with its original name combination, authority, and MDD synonym ID.

mdd_name_usages (MDD_name_usages) Later citations using the exact spelling of the name as given in mdd_original_combination; each citation is abbreviated to authors, year, and page number.

mdd_comments (MDD_comments) Free-text comments on the nomenclature or taxonomy of the name.

Details

This table covers every name — valid, synonymised, or otherwise — that has been applied to a species- or subspecies-level mammal taxon and is recognised by the MDD. It serves as the primary nomenclatural backbone for the checklist in mdd_checklist and is designed to interoperate with the Hesperomys database (<https://hesperomys.com>) via hesp_id.

Column names have been normalised from the original mixed-case convention (MDD_syn_ID, MDD_specificEpithet, etc.) to snake_case during data import; the original names are shown in parentheses in each \item above for cross-reference with the upstream spreadsheet.

Citation verification follows a two-tier structure: mdd_authority_citation holds APA-formatted, MDD-verified citations, while mdd_unchecked_authority_citation holds unverified or abbreviated citations from secondary sources and is set to NA once the verified field is populated.

Several fields are retained for internal MDD team reference and should not be used for analysis: mdd_old_type_locality, mdd_sourced_unverified_citations, mdd_citation_group, and mdd_citation_kind. The mdd_emended_type_locality and mdd_subspecific_epithet fields are present but **not yet fully curated**.

Source

Mammal Diversity Database release archive (<https://www.mammaldiversity.org>). Field definitions are derived from the META_v2.4.csv file and the column-level annotations in the Species_Syn_Current sheet of the official MDD spreadsheet.

mdd_taxon_info

Retrieve structured MDD taxon information by name

Description**[Experimental]**

This function is experimental and may change in future releases.

Resolve a mammal name against the local MDD backbone and return a structured object built from `mdd_checklist` and `mdd_synonyms`, including the accepted taxon record, synonym records, and grouped sections that mirror the main information shown in MDD species pages.

Usage

```
mdd_taxon_info(
  name,
  checklist = NULL,
  synonyms = NULL,
  target_df = NULL,
  max_dist = 1,
  method = "osa"
)
```

Arguments

<code>name</code>	A single scientific name.
<code>checklist</code>	Optional checklist data frame. Defaults to <code>mdd_checklist</code> .
<code>synonyms</code>	Optional synonym data frame. Defaults to <code>mdd_synonyms</code> .
<code>target_df</code>	Optional reconciliation backbone from <code>build_mdd_match_backbone()</code> .
<code>max_dist</code>	Maximum string distance used if the input name needs fuzzy reconciliation.
<code>method</code>	Distance method passed to <code>mdd_matching()</code> .

Value

An object of class `mdd_taxon_info`.

Examples

```
checklist <- tibble::tibble(
  id = "1",
  sci_name = "Puma_concolor",
  genus = "Puma",
  specific_epithet = "concolor",
  authority_species_author = "Linnaeus",
  authority_species_year = "1758",
  order = "Carnivora",
  family = "Felidae",
  main_common_name = "Puma"
)
synonyms <- tibble::tibble(
  mdd_syn_id = "1001",
  mdd_species_id = "1",
  mdd_author = "Linnaeus",
```

```
mdd_original_combination = "Felis concolor",
mdd_validity = "synonym",
mdd_nomenclature_status = "available"
)
x <- mdd_taxon_info(
  "Felis concolor",
  checklist = checklist,
  synonyms = synonyms
)
x
as.list(x)
```

mdd_taxon_record

Retrieve a normalized MDD taxon record

Description

[Experimental]

This function is experimental and may change in future releases.

Resolve a taxon by accepted MDD identifier or by scientific name against the local MDD name index. The returned object contains the match result, accepted taxon row, linked synonym rows, and the canonical MDD taxon URL.

Usage

```
mdd_taxon_record(
  name = NULL,
  taxon_id = NULL,
  checklist = NULL,
  synonyms = NULL,
  target_df = NULL,
  max_dist = 1,
  method = "osa"
)
```

Arguments

name	Optional scientific name.
taxon_id	Optional accepted MDD taxon identifier.
checklist	Optional checklist data frame.
synonyms	Optional synonym data frame.
target_df	Optional reconciliation backbone or name index.
max_dist	Maximum string distance used if name needs fuzzy matching.
method	Distance method passed to <code>mdd_matching()</code> .

Value

A list of class `mdd_taxon_record`.

Examples

```
checklist <- tibble::tibble(
  id = "1",
  sci_name = "Puma_concolor",
  genus = "Puma",
  specific_epithet = "concolor",
  authority_species_author = "Linnaeus"
)
synonyms <- tibble::tibble(
  mdd_syn_id = "1001",
  mdd_species_id = "1",
  mdd_author = "Linnaeus",
  mdd_original_combination = "Felis concolor"
)
mdd_taxon_record(
  name = "Felis concolor",
  checklist = checklist,
  synonyms = synonyms
)
```

`mdd_type_specimen_metadata`

Mammal Diversity Database type specimen metadata

Description

Auxiliary reference table of natural history museum collection metadata distributed with an MDD release. Each row corresponds to one institution that holds or has held type material cited in the MDD synonym table (`mdd_synonyms`).

Usage

```
mdd_type_specimen_metadata
```

Format

A tibble with 138 rows and 5 variables, with source column names normalized to `snake_case` during data import:

abbreviation Standard acronym or abbreviation used to identify the collection in the MDD and broader taxonomic literature (e.g. "AMNH", "MNHN", "USNM"); corresponds to the collection codes cited in the `mdd_holotype` and `mdd_type_specimen_link` fields of `mdd_synonyms`.

full_name Full official name of the institution or collection (e.g. "American Museum of Natural History", "Museum National d'Histoire Naturelle").

city_and_country City and country where the institution is located, formatted as "City, Country" (e.g. "New York, United States of America").

synonyms_notes Alternative or historical names by which the collection has been known in the synonymic literature (e.g. "British Museum (Natural History)" for BM, "United States National Museum" for USNM); NA when no alternative name is recorded.

online_website_database_if_available URL to the institution's online specimen database or collection portal, when publicly available; NA when no online resource has been identified.

Details

This table serves as a lookup reference linking collection abbreviations used throughout `mdd_synonyms` to their full institutional names, geographic locations, and online portals. It covers 138 institutions spanning all major zoogeographic regions, from large international collections such as AMNH, MNHN, and USNM to regional and national natural history museums in Latin America, Asia, Africa, and Oceania.

The `synonyms_notes` column documents historical or colloquial collection names that may appear in older taxonomic literature, aiding users who encounter non-standard abbreviations in pre-MDD sources. Notable examples include BM (formerly *British Museum (Natural History)*), MCZ (*Museum of Comparative Zoology, Harvard*), and ZMA (*Zoölogisch Museum, Amsterdam*, merged with RMNH in 2010).

URLs in `online_website_database_if_available` link directly to mammalogy or specimen-search pages where possible, and to general institutional pages otherwise. Link availability and validity may change over time.

Source

Mammal Diversity Database release archive (<https://www.mammaldiversity.org>), distributed as `TypeSpecimenMetadata_...csv` within each versioned release.

print.mdd_reference *Print mdd_reference object*

Description

Print `mdd_reference` object

Usage

```
## S3 method for class 'mdd_reference'
print(x, ...)
```

Arguments

`x` An object of class `mdd_reference`.
`...` Additional arguments.

Value

The mdd_reference object, invisibly.

```
print.mdd_taxon_info  Print mdd_taxon_info object
```

Description

Print mdd_taxon_info object

Usage

```
## S3 method for class 'mdd_taxon_info'  
print(x, ...)
```

Arguments

x	An object of class mdd_taxon_info.
...	Additional arguments.

Value

The mdd_taxon_info object, invisibly.

Index

* datasets

- mdd_checklist, 5
- mdd_synonyms, 18
- mdd_type_specimen_metadata, 24

as.list.mdd_taxon_info, 2

build_mdd_match_backbone, 3

build_mdd_match_backbone(), 10, 15, 22

classify_mammal_names, 4

mdd_cache_dir, 4

mdd_checklist, 5

mdd_checklist_records, 8

mdd_distribution_map, 9

mdd_distribution_summary, 11

mdd_distribution_summary_raw, 12

mdd_distribution_summary_raw(), 11

mdd_download, 13

mdd_load, 14

mdd_matching, 14

mdd_matching(), 3, 9, 10, 22, 23

mdd_name_index, 15

mdd_reference, 16

mdd_release_records, 17

mdd_synonym_records, 17

mdd_synonyms, 18

mdd_taxon_info, 21

mdd_taxon_record, 23

mdd_type_specimen_metadata, 24

print.mdd_reference, 25

print.mdd_taxon_info, 26

utils::download.file(), 13