Package ‘insectDisease’

January 7, 2023

Title Ecological Database of the World’s Insect Pathogens
Version 1.2.2
Description David Onstad provided us with this insect disease database, sometimes referred to as the ‘Ecological Database of the Worlds Insect Pathogens’ or ED-WIP. Files have been converted from 'SQL' to csv, and ported into 'R' for easy exploration and analysis. Thanks to the Macroeology of Infectious Disease Research Coordination Network (RCN) for funding and support. Data are also served online in a static format at <https://edwip.ecology.uga.edu/>.

Depends R (>= 3.5.0), plyr, taxize
Suggests corrplot, dplyr, knitr, rmarkdown, testthat
License GPL-3
LazyData true

BugReports https://github.com/viralemergence/insectDisease/issues/
RoxygenNote 7.2.3
VignetteBuilder knitr
Encoding UTF-8
NeedsCompilation no
Author Tad Dallas [aut, cre]
Maintainer Tad Dallas <tad.a.dallas@gmail.com>
Repository CRAN
Date/Publication 2023-01-06 23:00:02 UTC

R topics documented:

assocref .......................................................... 2
citation .......................................................... 3
getNCBI ........................................................... 4
hosts ............................................................... 4
hostTaxonomy .................................................... 5
InsectDisease ................................................... 6
### Description

Edgelist of known associations between insect host (‘Host’) and pathogen (‘Pathogen’), and associated reference and indexing values. The variables are as follows:

### Usage

```r
data(assocref)
```

### Format

A data.frame with 11005 observations

### Details

- **ERNnvp. EDWIP record number**
- **refCode. Reference code**
- **Reference. Actual citation**
- **HostSpecies. Host species**
- **PathogenSpecies. Pathogen species**
- **HostTaxID. Host NCBI ID number**
- **HostGenus. Host genus**
- **HostFamily. Host family**
- **HostOrder. Host order**
- **HostClass. Host class**
- **PathTaxID. Pathogen NCBI ID number**
- **PathGenus. Pathogen genus**
- **PathFamily. Pathogen family**
• PathOrder. Pathogen order
• PathClass. Pathogen class
• PathKingdom. Pathogen kingdom

Description

Edgelist of known associations between insect host ('Host') and pathogen ('Pathogen'), and associated references and indexing values. These data are smaller than assocref, which provides more data on host-parasite interactions. CreationDate and ModificationDate are incorrect. The variables are as follows:

Usage

data(citation)

Format

A data.frame with 1966 observations

Details

• CitationCode. Citation code
• Reference. Actual citation
• CreationDate. Record creation date
• ModificationDate. Record modification date
• ReadBy. Comments about the reading and identity of reader
• GetIt. Notes on article acquisition
• nvpCount. Number of pathogens reported in the citation.
getNCBI  
*Get NCBI taxonomy data*

**Description**
Process through a vector of species names to obtain taxonomic data

**Usage**
getNCBI(species, host = TRUE)

**Arguments**
- species: a vector of species names
- host: (boolean) affects column naming (nice to keep host and pathogen separate)

**Value**
a data.frame with nrow == length(species)

**Examples**
```r
## Not run:
mouse <- getNCBI("Peromyscus leucopus")
worm <- getNCBI("Ascaris lumbricoides", host=FALSE)
## End(Not run)
```

**hosts**  
*Detailed information on insect host species*

**Description**
Data on host species taxonomy, habitat, diet preferences, and distribution (in Canada)

**Usage**
data(hosts)

**Format**
A data.frame with 4392 observations
Details

- **RecordNo.** Sequence from 1:nrow(hosts)
- **DateEntered.** Date of initial data entry
- **Habitat.** Habitat of host
- **HostSpecies.** Host species
- **Synonyms.** Other names for the host species
- **Food.** What does the host eat?
- **genYr.** Number of generations per year. Can be <1, =1, >1, or some combination of these.
- **CommonName.** Host common name
- **ProvinceI.** Canadian provinces where host has been found.
- **InsectStatus.** Is the insect a pest, beneficial, endangered, unknown? Factor variable with 7 unique values
- **ModificationDate.** Modification date of entry
- **InCanada.** Citations for presence/absence of host in Canada. Numeric indices can be related to the citations in the ‘citations.rda’ data file. ‘Y’ and ‘N’ relate to presence and absence, respectively.
- **ChangeSpeciesTo.** Taxonomic verification column
- **CommonNameOther.** Other common names?
- **Complete.** Is this record complete?
- **AdditionalReferences.** Additional reference indices.
- **HostTaxID.** Host NCBI ID number
- **HostGenus.** Host genus
- **HostFamily.** Host family
- **HostOrder.** Host order
- **HostClass.** Host class

---

**hostTaxonomy**

Cached version of host taxonomy (see vignette for code to build from scratch)

---

Description

Host taxonomic cached data from NCBI (see vignette for code to re-make)

- **HostSpecies original host name from EDWIP**
- **HostTaxID NCBI taxonomic ID**
- **HostGenus host genus**
- **HostFamily host family**
- **HostOrder host order**
- **HostClass host class**
### Usage

data(hostTaxonomy)

### Format

A data.frame

---

#### InsectDisease

*InsectDisease.*

---

#### Description

This package provides a ‘R’ workspace with insect disease database data already present. Data are visible by using the R command `ls("package:InsectDisease")`, and individual data frames can be imported to your local workspace by using the command ‘data(NAMEOFDATA)’

Data frames can be grouped by parasite type, as ‘nemaref’, ‘nematode’ and ‘newnema’ provide information on nematode infections, as ‘nvpassoc’ provides information on protozoans, and ‘viraref’ and ‘viruses’ provide details of viral infections. ‘pathogen’ provides general taxonomic information on pathogens in the database. Citations are available in a number of places, including but not limited to ‘assocref’, ‘citation’, ‘nemaref’, and ‘viraref’.

*When using this database, cite this reference*


Also, this database was originally created by the following people, to whom we are indebted:

- David W. Onstad, EDWIP Director. Center for Economic Entomology, Illinois Natural History Survey
- Ellen Brewer, Research Programmer Center for Economic Entomology, Illinois Natural History Survey
- Susan Braxton, Science & Technology Librarian Milner Library, Illinois State University

---

#### negative

*Information on negative host-parasite interactions*

---

#### Description

Hosts were challenged with pathogen, and did not become infected. These are data on what pathogens do not infect certain hosts.

#### Usage

data(negative)


Format

A data.frame with 529 observations

Details

- ERNntr. EDWIP record number
- PathogenSpecies. Pathogen species
- DateEntered. Date of initial data entry
- DateModified. Modification date of entry
- LogMaxDose. Dosage, in many different units
- HostStageTested. Host stage exposed to pathogen (e.g. Larvae, Nymph, Adult)
- HostSpecies. Host species examined
- Group. Pathogen group (e.g. viruses)
- HighTaxon. General classification of pathogen (e.g. DNA virus)
- LowTaxon. More specific classification of pathogen (e.g. Baculoviridae)
- HostTaxID. Host NCBI ID number
- HostGenus. Host genus
- HostFamily. Host family
- HostOrder. Host order
- HostClass. Host class
- PathTaxID. Pathogen NCBI ID number
- PathGenus. Pathogen genus
- PathFamily. Pathogen family
- PathOrder. Pathogen order
- PathClass. Pathogen class
- PathKingdom. Pathogen kingdom

---

**nemaref**

Citation information for host-nematode interactions

---

Description

Edgelist of known associations between insect host (‘Host’) and nematode parasite (‘Nematode’), and associated references and indexing values. These data are smaller than ‘asscref’, and ‘citation’. The variables are as follows:

Usage

data(nemaref)
Format

A data.frame with 338 observations

Details

- ERNnem. EDWIP record number
- refCode. Index of reference obtained from 'nematode' data frame
- Reference. Citation for host-nematode record
- HostSpecies. Host species
- PathogenSpecies. Nematode parasite species
- HostTaxID. Host NCBI ID number
- HostGenus. Host genus
- HostFamily. Host family
- HostOrder. Host order
- HostClass. Host class
- PathTaxID. Pathogen NCBI ID number
- PathGenus. Pathogen genus
- PathFamily. Pathogen family
- PathOrder. Pathogen order
- PathClass. Pathogen class
- PathKingdom. Pathogen kingdom

---

**nematode**

Information on nematode parasite occurrences

---

Description

Host associations for nematode parasites

Usage

data(nematode)

Format

A data.frame with 234 observations
Details

- ERNnem. EDWIP record identifier
- HostSpecies. Host species
- PathogenSpecies. Nematode parasite species
- PathogenStrain. Nematode parasite strain
- StageInfected. Host stage infected
- TissueInfected. Host tissue infected
- FieldOrLab. Was this a field or lab tested association?
- Country. What country did the interaction occur in?
- SoilType. Type of soil where interaction was observed
- AssociatedBacterium. Associated bacterium.
- IntermediateHost. Is there an intermediate host present?
- CreationDate. Date of initial data entry
- ModificationDate. Modification date of entry
- HostTaxID. Host NCBI ID number
- HostGenus. Host genus
- HostFamily. Host family
- HostOrder. Host order
- HostClass. Host class
- PathTaxID. Pathogen NCBI ID number
- PathGenus. Pathogen genus
- PathFamily. Pathogen family
- PathOrder. Pathogen order
- PathClass. Pathogen class
- PathKingdom. Pathogen kingdom

Description

These data are included because they were originally part of the EDWIP data. We caution the user to not use these data, as we do not believe the records are correct.

These data are included because they were originally part of the EDWIP data. These data are identical to the ‘nematode’ data, as far as we can tell.

Usage

data(new_asso)
data(newnema)
Format

A data.frame of not good records
A data.frame

Details

- ERNnew. EDWIP record number
- HostSpecies. Host species
- HostOrder. Host order
- HostFamily. Host family
- HostHabitat. Habitat type of host
- HostFood. What does the host eat?
- HostGenYr. Number of generations of hosts per year
- PathSpecies. Nematode species
- PathGroup. Pathogen group (fungi, protozoa, nematode, etc.)
- PathHighTaxon. Pathogen taxonomic information (mostly ‘NA’)
- PathLowTaxon. Pathogen taxonomic information (mostly ‘NA’)
- StageInf. Host life stage infected
- TissueInfected. Host tissue infected
- Field. Was this a field or lab tested association?
- Country. Country of host-pathogen association
- IntermediateHost. Information on intermediate hosts
- Citation. Citation for host-pathogen record
- MoreInfo. Additional comments or notes
- Who. Identity of researcher who entered data
- CreationDate. Record creation date
- ModificationDate. Record modification date
- StainFCB. takes values: Adult, egg, larvae, pupa
- ERNnem. EDWIP record identifier
- Host. Host species
- Nematode. Nematode parasite species
- NemaOrder. Pathogen order
- NemaFamily. Pathogen family
- NemaStrain. Pathogen strain
- StageInfected. Host stage infected
- TissueInfected. Host tissue infected
- FieldOrLab. Was this a field or lab tested association?
- Country. What country did the interaction occur in?
• SoilType. Type of soil where interaction was observed
• AssociatedBacterium. Associated bacterium
• IntermediateHost. Is there an intermediate host present?
• CreationDate. Date of initial data entry
• ModificationDate. Modification date of entry
• Group. all just say ‘nematode’

Description
These data are included because they were originally part of the EDWIP data. We caution the user in using these data, as they do not have associated metadata that the other records have.

Usage
data(noassref)

Format
A data.frame

Details
• RefCode. Index of reference
• ERNntr. EDWIP record number
• HostSpecies. Host species
• PathogenSpecies. Pathogen species
• Citation. Citation for host-pathogen record
• HostTaxID. Host NCBI ID number
• HostGenus. Host genus
• HostFamily. Host family
• HostOrder. Host order
• HostClass. Host class
• PathTaxID. Pathogen NCBI ID number
• PathGenus. Pathogen genus
• PathFamily. Pathogen family
• PathOrder. Pathogen order
• PathClass. Pathogen class
• PathKingdom. Pathogen kingdom
Information on non-viral pathogens of insect hosts

Description
Non-viral parasite (protozoan, fungi, bacteria) occurrences on host species. The variables are as follows:

Usage
data(nvpassoc)

Format
A data.frame with 7164 observations

Details
- ERNvp. EDWIP record number
- PathogenSpecies. Pathogen species
- Group. Pathogen group (e.g. Protozoa)
- HostSpecies. Host species examined
- HostStageTested. Host stage exposed to pathogen (e.g. Larvae, Nymph, Adult)
- HostTissueInfected. Host tissue infected
- FieldOrLab. Was this a field or lab tested association?
- Country. What country did the interaction occur in?
- IntermediateHost. Is there an intermediate host present?
- DateEntered. Date of initial data entry
- DateModified. Modification date of entry
- BiogeographicRegion. Biogeographic region (or some combination thereof)
- HostTaxID. Host NCBI ID number
- HostGenus. Host genus
- HostFamily. Host family
- HostOrder. Host order
- HostClass. Host class
- PathTaxID. Pathogen NCBI ID number
- PathGenus. Pathogen genus
- PathFamily. Pathogen family
- PathOrder. Pathogen order
- PathClass. Pathogen class
- PathKingdom. Pathogen kingdom
Information on pathogen species in the database

**Description**

Information on taxonomy of parasites present in this database

**Usage**

data(pathogen)

**Format**

A data.frame with 2041 observations

**Details**

- PathogenSpecies. Pathogen species.
- Group. Pathogen group (e.g. Protozoa)
- AdditionalNotes. Some additional notes.
- PathTaxID. Pathogen NCBI ID number
- PathGenus. Pathogen genus
- PathFamily. Pathogen family
- PathOrder. Pathogen order
- PathClass. Pathogen class
- PathKingdom. Pathogen kingdom

**pathTaxonomy**

Cached version of pathogen taxonomy (see vignette for code to build from scratch)

**Description**

- PathTaxID NCBI taxonomic ID
- PathNCBIResolved is the pathogen found in NCBI
- PathGenus pathogen genus
- PathFamily pathogen family
- PathOrder pathogen order
- PathClass pathogen class
- PathKingdom pathogen kingdom
Usage

data(pathTaxonomy)

Format

A data.frame

---

**viraref**  
*Citation information for host-virus interactions*

**Description**

Edgelist of known associations between insect host (‘Host’) and viral pathogen (‘Virus’), and associated references and indexing values. These data are smaller than ‘assocref’, and ‘citation’. The variables are as follows:

**Usage**

data(viraref)

**Format**

A data.frame with 2124 observations

**Details**

- RefCode. Index of reference obtained from ‘nematode’ data frame
- Citation. Reference.
- ERNv. EDWIP record number
- HostSpecies. Host species
- PathogenSpecies. Virus name
- HostTaxID. Host NCBI ID number
- HostGenus. Host genus
- HostFamily. Host family
- HostOrder. Host order
- HostClass. Host class
- PathTaxID. Pathogen NCBI ID number
- PathGenus. Pathogen genus
- PathFamily. Pathogen family
- PathOrder. Pathogen order
- PathClass. Pathogen class
- PathKingdom. Pathogen kingdom
viruses

Information on viral pathogen occurrences

Description

Host associations for viruses

Usage

data(viruses)

Format

A data.frame with 1659 observations

Details

- ERNv. EDWIP record number
- HostSpecies. Host species
- VirusType. DNA or RNA virus
- PathogenSpecies. Viral family
- Virus. Virus identity
- HostStageInfected. Host stage infected
- HostTissueInfected. Host tissue infected
- FieldOrLab. Was this a field or lab tested association?
- Country. What country did the interaction occur in?
- IntermediateHost. Is there an intermediate host present?
- CreationDate. Date of initial data entry (wrong)
- ModificationDate. Modification date of entry (wrong)
- ProvinceA. Canadian provinces where host-virus interaction occurs
- PathogenValue. Is there value to the pathogen (can it be used as a control agent?)
- HostTaxID. Host NCBI ID number
- HostGenus. Host genus
- HostFamily. Host family
- HostOrder. Host order
- HostClass. Host class
- PathTaxID. Pathogen NCBI ID number
- PathGenus. Pathogen genus
- PathFamily. Pathogen family
- PathOrder. Pathogen order
- PathClass. Pathogen class
- PathKingdom. Pathogen kingdom
Index

* datasets
  assocref, 2
  citation, 3
  hosts, 4
  hostTaxonomy, 5
  negative, 6
  nemaref, 7
  nematode, 8
  new_asso, 9
  noassref, 11
  nvpassoc, 12
  pathogen, 13
  pathTaxonomy, 13
  viraref, 14
  viruses, 15

assocref, 2

citation, 3

getNCBI, 4

hosts, 4
hostTaxonomy, 5

InsectDisease, 6

negative, 6
nemaref, 7
nematode, 8
new_asso, 9
newnema (new_asso), 9
noassref, 11
nvpassoc, 12

pathogen, 13
pathTaxonomy, 13

viraref, 14
viruses, 15