

miRBaseVersions.db-vignette

Stefan Haunsberger

2018-03-26

Contents

1. Introduction	1
2. Use Cases	2
Additional information	4
References	5

The miRBase database (Griffiths-Jones, 2004; Griffiths-Jones, Grocock, Dongen, Bateman, & Enright, 2006; Griffiths-Jones, Saini, Dongen, & Enright, 2008; Kozomara & Griffiths-Jones, 2011, 2014) is the official repository for miRNAs and includes a miRNA naming convention (AMBROS et al., 2003; Meyers et al., 2008). Over the years of development miRNAs have been added to, or deleted from the database, while some miRNA names have been changed. As a result, each version of the miRBase database can differ substantially from previous versions.

The *miRBaseVersions.db* R package has been developed to provide an easy accessible repository for several different miRBase release versions.

1. Introduction

The *miRBaseVersions.db* package is an annotation package which includes mature miRNA names from 22 miRBase release versions. Due to ongoing growth and changes with each release miRNA names can have different names in different versions or even are not listed as valid miRNAs anymore. This annotation package serves as a repository and can be used for quick lookup for mature miRNA names. The *miRBaseVersions.db* package has implemented the AnnotationDbi-select interface. By implementing this `select` interface the user is able to use the same methods as for any other annotation package.

The main four implemented methods are

- `columns`, presents the values one can retrieve in the final result,
- `keytypes`, which presents the tables that can be used in this package,
- `keys`, is used to get viable keys of a particular `keytype` and
- `select`, which is used to extract data from the annotation package by using values provided by the other three methods.

To load the package and gain access to the functions just run the following command:

```
library(miRBaseVersions.db)
```

Vignette Info

This vignette has been generated using an R Markdown file with `knitr:rmarkdown` as vignette engine (Boettiger, 2015; Francois, 2014; Xie, 2014, 2015b, 2015a).

Database information

The data is the *miRNANameConverter* package is stored in an SQLite database. All entries contained in the database were downloaded from the miRBase ftp-site. The following versions are available:

miRBase Version	Release Date	# Mature entries
6.0	04/05	1650
7.1	10/05	3424
8.0	02/06	3518
8.1	05/06	3963
8.2	07/06	4039
9.0	10/06	4361
9.1	02/07	4449
9.2	05/07	4584
10.0	08/07	5071
10.1	12/07	5395
11.0	04/08	6396
12.0	09/08	8619
13.0	03/09	9539
14.0	09/09	10883
15.0	04/10	14197
16.0	08/10	15172
17.0	04/11	16772
18.0	11/11	18226
19.0	08/12	21264
20.0	06/13	24521
21.0	06/14	28645
22.0	03/18	38589

from 271 organisms.

2. Use Cases

2.1 Function keytypes

Use this function to receive table names from where data can be retrieved:

```
keytypes(miRBaseVersions.db);
```

```
## [1] "MIMAT"          "VW-MIMAT-10.0" "VW-MIMAT-10.1" "VW-MIMAT-11.0"
## [5] "VW-MIMAT-12.0" "VW-MIMAT-13.0" "VW-MIMAT-14.0" "VW-MIMAT-15.0"
## [9] "VW-MIMAT-16.0" "VW-MIMAT-17.0" "VW-MIMAT-18.0" "VW-MIMAT-19.0"
## [13] "VW-MIMAT-20.0" "VW-MIMAT-21.0" "VW-MIMAT-22.0" "VW-MIMAT-6.0"
## [17] "VW-MIMAT-7.1"  "VW-MIMAT-8.0"  "VW-MIMAT-8.1"  "VW-MIMAT-8.2"
## [21] "VW-MIMAT-9.0"  "VW-MIMAT-9.1"  "VW-MIMAT-9.2"
```

The output lists 23 tables where each one of them can be queried. The keytype “MIMAT” is the main table containing all records from all supported miRBase release versions. Keytypes starting with the prefix “VW-MIMAT” are so called SQL views. For example the keytype “VW-MIMAT-22.0” is an SQL view from the “MIMAT” table which only holds records from miRBase version 22.0.

2.2 Function columns

Use the `columns` function to retrieve information about the kind of variables you can retrieve in the final output:

```
columns(miRBaseVersions.db);
```

```
## [1] "ACCESSION" "NAME" "ORGANISM" "SEQUENCE" "VERSION"
```

All 5 columns are available for all 23 keytypes.

2.3 Function keys

The `keys` function returns all viable keys of a particular keytype. The following example retrieves all possible keys for miRBase release version 6.0.

```
k = head(keys(miRBaseVersions.db, keytype = "VW-MIMAT-6.0"));
k;
```

```
## [1] "MIMAT0000001" "MIMAT0000002" "MIMAT0000003" "MIMAT0000004"
```

```
## [5] "MIMAT0000005" "MIMAT0000006"
```

2.4 Function select

The `select` function is used to extract data. As input values the function takes outputs received from the other three functions `keys`, `columns` and `keytypes`.

For example to extract all information about the mature accession 'MIMAT0000092' we can run the following command:

```
result = select(miRBaseVersions.db,
               keys = "MIMAT0000092",
               keytype = "MIMAT",
               columns = "*")
result;
```

##	ACCESSION	NAME	SEQUENCE	VERSION	ORGANISM
## 1	MIMAT0000092	hsa-miR-92a-3p	UAUUGCACUUGUCCCGGCCUGU	22.0	hsa
## 2	MIMAT0000092	hsa-miR-92a-3p	UAUUGCACUUGUCCCGGCCUGU	21.0	hsa
## 3	MIMAT0000092	hsa-miR-92a-3p	UAUUGCACUUGUCCCGGCCUGU	20.0	hsa
## 4	MIMAT0000092	hsa-miR-92a-3p	UAUUGCACUUGUCCCGGCCUGU	19.0	hsa
## 5	MIMAT0000092	hsa-miR-92a-3p	UAUUGCACUUGUCCCGGCCUGU	18.0	hsa
## 6	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	17.0	hsa
## 7	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	16.0	hsa
## 8	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	15.0	hsa
## 9	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	14.0	hsa
## 10	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	13.0	hsa
## 11	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	12.0	hsa
## 12	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	11.0	hsa
## 13	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	10.1	hsa
## 14	MIMAT0000092	hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	10.0	hsa
## 15	MIMAT0000092	hsa-miR-92	UAUUGCACUUGUCCCGGCCUG	9.2	hsa
## 16	MIMAT0000092	hsa-miR-92	UAUUGCACUUGUCCCGGCCUG	9.1	hsa
## 17	MIMAT0000092	hsa-miR-92	UAUUGCACUUGUCCCGGCCUG	9.0	hsa
## 18	MIMAT0000092	hsa-miR-92	UAUUGCACUUGUCCCGGCCUG	8.2	hsa
## 19	MIMAT0000092	hsa-miR-92	UAUUGCACUUGUCCCGGCCUG	8.1	hsa
## 20	MIMAT0000092	hsa-miR-92	UAUUGCACUUGUCCCGGCCUG	8.0	hsa
## 21	MIMAT0000092	hsa-miR-92	UAUUGCACUUGUCCCGGCCUG	7.1	hsa
## 22	MIMAT0000092	hsa-miR-92	UAUUGCACUUGUCCCGGCCUG	6.0	hsa

As we can see the result returns all miRNA names the accession had among the different miRBase releases. If we for example only want to extract the fields for 'accession', 'name' and 'version' we simply run the following command:

```
result = select(miRBaseVersions.db,
               keys = "MIMAT0000092",
               keytype = "MIMAT",
               columns = c("ACCESSION", "NAME", "VERSION"))
result;
```

##	ACCESSION	NAME	VERSION
## 1	MIMAT0000092	hsa-miR-92a-3p	22.0
## 2	MIMAT0000092	hsa-miR-92a-3p	21.0
## 3	MIMAT0000092	hsa-miR-92a-3p	20.0
## 4	MIMAT0000092	hsa-miR-92a-3p	19.0
## 5	MIMAT0000092	hsa-miR-92a-3p	18.0
## 6	MIMAT0000092	hsa-miR-92a	17.0
## 7	MIMAT0000092	hsa-miR-92a	16.0
## 8	MIMAT0000092	hsa-miR-92a	15.0
## 9	MIMAT0000092	hsa-miR-92a	14.0
## 10	MIMAT0000092	hsa-miR-92a	13.0
## 11	MIMAT0000092	hsa-miR-92a	12.0
## 12	MIMAT0000092	hsa-miR-92a	11.0
## 13	MIMAT0000092	hsa-miR-92a	10.1
## 14	MIMAT0000092	hsa-miR-92a	10.0
## 15	MIMAT0000092	hsa-miR-92	9.2
## 16	MIMAT0000092	hsa-miR-92	9.1
## 17	MIMAT0000092	hsa-miR-92	9.0
## 18	MIMAT0000092	hsa-miR-92	8.2
## 19	MIMAT0000092	hsa-miR-92	8.1
## 20	MIMAT0000092	hsa-miR-92	8.0
## 21	MIMAT0000092	hsa-miR-92	7.1
## 22	MIMAT0000092	hsa-miR-92	6.0

In comparison to the previous output with parameter `columns = "*"` this time only the selected columns were returned.

Additional information

Packages loaded via namespace

The following packages are used in the `miRBaseVersions.db` package:

- AnnotationDbi_1.32.3 (Pages, Carlson, Falcon, & Li, n.d.)
- DBI_0.3.1 (Databases, 2014)
- RSQLite_1.0.0 (Wickham, James, & Falcon, 2014)
- gtools_3.5.0 (Warnes, Bolker, & Lumley, 2015)

Future Aspects

This database can only be of good use if it will be kept up to date. Therefore we plan to include new miRBase releases as soon as possible.

References

- AMBROS, V., BARTEL, B., BARTEL, D. P., BURGE, C. B., CARRINGTON, J. C., CHEN, X., ... TUSCHL, T. (2003). A uniform system for microRNA annotation. *RNA*, 9(3), 277–279. <http://doi.org/10.1261/rna.2183803>
- Boettiger, C. (2015). *Knitcitations: Citations for 'knitr' markdown files*. Retrieved from <http://CRAN.R-project.org/package=knitcitations>
- Databases, R. S. I. G. on. (2014). *DBI: R database interface*. Retrieved from <http://CRAN.R-project.org/package=DBI>
- Francois, R. (2014). *Bibtex: Bibtex parser*. Retrieved from <http://CRAN.R-project.org/package=bibtex>
- Griffiths-Jones, S. (2004). The microRNA registry. *Nucleic Acids Research*, 32(suppl 1), D109–D111. <http://doi.org/10.1093/nar/gkh023>
- Griffiths-Jones, S., Grocock, R. J., Dongen, S. van, Bateman, A., & Enright, A. J. (2006). MiRBase: MicroRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*, 34(suppl 1), D140–D144. <http://doi.org/10.1093/nar/gkj112>
- Griffiths-Jones, S., Saini, H. K., Dongen, S. van, & Enright, A. J. (2008). MiRBase: Tools for microRNA genomics. *Nucleic Acids Research*, 36(suppl 1), D154–D158. <http://doi.org/10.1093/nar/gkm952>
- Kozomara, A., & Griffiths-Jones, S. (2011). MiRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research*, 39(suppl 1), D152–D157. <http://doi.org/10.1093/nar/gkq1027>
- Kozomara, A., & Griffiths-Jones, S. (2014). MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research*, 42(D1), D68–D73. <http://doi.org/10.1093/nar/gkt1181>
- Meyers, B. C., Axtell, M. J., Bartel, B., Bartel, D. P., Baulcombe, D., Bowman, J. L., ... others. (2008). Criteria for annotation of plant micromRNAs. *The Plant Cell*, 20(12), 3186–3190.
- Pages, H., Carlson, M., Falcon, S., & Li, N. (n.d.). *AnnotationDbi: Annotation database interface*.
- Warnes, G. R., Bolker, B., & Lumley, T. (2015). *Gtools: Various r programming tools*. Retrieved from <https://CRAN.R-project.org/package=gtools>
- Wickham, H., James, D. A., & Falcon, S. (2014). *RSQLite: SQLite interface for r*. Retrieved from <http://CRAN.R-project.org/package=RSQLite>
- Xie, Y. (2014). Knitr: A comprehensive tool for reproducible research in R. In V. Stodden, F. Leisch, & R. D. Peng (Eds.), *Implementing reproducible computational research*. Chapman; Hall/CRC. Retrieved from <http://www.crcpress.com/product/isbn/9781466561595>
- Xie, Y. (2015a). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, Florida: Chapman; Hall/CRC. Retrieved from <http://yihui.name/knitr/>
- Xie, Y. (2015b). *Knitr: A general-purpose package for dynamic report generation in r*. Retrieved from <http://yihui.name/knitr/>