Package: bedrockbio (via r-universe)

July 2, 2026

Title Open-Access Computational Biology Datasets

Version 2.0.0

Description Efficiently access the 'Bedrock Bio' library of open-access computational biology datasets. Lazily query datasets backed by 'DuckDB' and 'Apache Iceberg', with support for predicate pushdown and column projection to the cloud storage backend. This enables quick, iterative access to otherwise massive, unwieldy datasets without downloading them in full. See <`https://bedrock.bio`> for available datasets and documentation.

Language en-US

License GPL (>= 3)

URL `https://bedrock.bio`, `https://github.com/bedrock-bio/bedrock-bio-client`

BugReports `https://github.com/bedrock-bio/bedrock-bio-client/issues`

Depends R (>= 4.2)

Imports curl, DBI, dbplyr, dplyr, duckdb, jsonlite

Suggests testthat (>= 3.0.0), withr

Config/testthat/edition 3

OS_type unix

Encoding UTF-8

RoxygenNote 7.3.3

NeedsCompilation no

Author Liam Abbott [aut, cre, cph]

Maintainer Liam Abbott <`liam@bedrock.bio`>

Repository https://cran.r-universe.dev

Date/Publication 2026-07-02 13:20:07 UTC

RemoteUrl https://github.com/cran/bedrockbio

RemoteRef HEAD

RemoteSha 7e3cddfebf72c3466744b6a6e6620b2e358c38bd

Contents

`describe_namespace` *Describe a namespace: its name, citation, license, context, and tables*

Description

Describe a namespace: its name, citation, license, context, and tables

Usage

```
describe_namespace(name)
```

Arguments

`name`	Namespace identifier.

Value

A named list with `name`, `citation`, `license`, `context`, and `tables` (fully-qualified table identifiers). Use `describe_table()` for per-table details.

Examples

```
## Not run:
library(bedrockbio)
describe_namespace("ukb_ppp")$tables

## End(Not run)
```

`describe_table` *Describe a table: its context, columns, and partitions*

Description

Describe a table: its context, columns, and partitions

Usage

```
describe_table(name)
```

Arguments

`name`	Table identifier.

Value

A named list with `name`, `context`, `columns` (each with `name`, `type`, `description`, `nullable`), and `partitions` (a named list of partition column to `values` and `default`). Filter on partition columns for fastest reads.

Examples

```
## Not run:
library(bedrockbio)
describe_table("ukb_ppp.pqtls")$name

## End(Not run)
```

`list_namespaces` *List available namespaces (data sources)*

Description

List available namespaces (data sources)

Usage

```
list_namespaces()
```

Value

A character vector of namespace identifiers.

Examples

```
## Not run:
library(bedrockbio)
list_namespaces()

## End(Not run)
```

`list_tables` *List available tables, optionally filtered to one namespace*

Description

List available tables, optionally filtered to one namespace

Usage

```
list_tables(namespace = NULL)
```

Arguments

`namespace`	If given, return only that namespace's tables; otherwise all tables.

Value

A character vector of fully-qualified table identifiers.

Examples

```
## Not run:
library(bedrockbio)
list_tables("ukb_ppp")

## End(Not run)
```

`load_table` *Lazily query a table*

Description

Lazily query a table

Usage

```
load_table(name)
```

Arguments

`name`	Table identifier.

Value

A lazy `tbl` backed by DuckDB, compatible with dplyr verbs. Filter on partition columns (see `describe_table()`) for fastest reads.

Examples

```
## Not run:
library(bedrockbio)
library(dplyr)

load_table("dbsnp.vcf") |>
  filter(assembly == "GRCh38", chromosome == "22") |>
  select(rsid, position, ref_allele, alt_allele) |>
  head(5) |>
  collect()

## End(Not run)
```

Index