Package 'PCAWG7'

October 20, 2020

Title Repository of data from 'Repertoire of Mutational Signatures in Human Cancer'

Version 0.0.3

Description Contains data from Alexandrov, Kim, Haradhvala, Huang et al., 'Repertoire of Mutational Signatures in Human Cancer'. Please see ?PCAWG7. The reference for the data is Alexandrov, L.B., Kim, J., Haradhvala, N.J. et al. The repertoire of mutational signatures in human cancer. Nature 578, 94-101 (2020). https://doi.org/10.1038/s41586-020-1943-3. The funny name comes from the fact that this paper was generated by Working Group 7 of the Pan Cancer Analysis of Whole Genomes (PCAWG) consortium.

License GPL-3

Language en-US

Encoding UTF-8

LazyData true

Depends R (>= 3.5),

RoxygenNote 7.1.1

URL https://github.com/steverozen/PCAWG7

BugReports https://github.com/steverozen/PCAWG7/issues

R topics documented:

exposure 2
PCAWG.sample.id 2
PCAWG.WGS.DBS 3
PCAWG.WGS.SBS.96 3
PCAWG7 4
signature 4
spectra 5

Index **6**

exposure *PCAWG7 SigProfiler assignments (exposures).*

Description

PCAWG7 SigProfiler assignments (exposures).

Usage

```
exposure
```

Format

A list with the elements:

PCAWG A list with the elements:

- **SBS96** Strand-agnostic single-base substitutions in trinucleotide context.
- **DBS78** Strand-agnostic doublet-base substitutions.
- **ID** Strand-agnostic indels. These are signature assignments for the PCAWG platinum genomes.

TCGA A list with the elements:

- **SBS96** As above.
- **ID** As above. These are signature assignments for the TCGA exomes.

other.genome A list with the element:

- **SBS96** As above. This contains signature assignments for non-TCGA genomes.

other.exome A list with the element:

- **SBS96** As above. This contains signature assignments for non-TCGA exomes.

Source

Files of https://www.synapse.org/#!Synapse:syn12009743, 2019 Oct 09, populated by data-raw/sig.profiler.e

PCAWG.sample.id *Vectors of the PCAWG* tumor_wgs_icgc_specimin_id*s.*

Description

Note that the PCAWG7 spectra catalogs have 2 sample ids that were blacklisted after the mutational signature analysis was underway. The blacklisted samples are SP116419 and SP116883, which are in PCAWG.sample.id$black.

Usage

```
PCAWG.sample.id
```

Format

A list with the elements:

white Whitelisted IDs

grey Greylisted IDs

black Blacklisted IDs

Source

https://dcc.icgc.org/api/v1/download?fn=/PCAWG/data_releases/latest/release_may2016.v1.4.with_consensus_calls.tsv, 2019 Oct 09

`PCAWG.WGS.DBS` *Doublet Base Substitution (SBS) spectra (deprecated). Use* `spectra$PCAWG$DBS78` *instead.*

Description

Doublet Base Substitution (SBS) spectra (deprecated). Use `spectra$PCAWG$DBS78` instead.

Usage

```
PCAWG.WGS.DBS
```

Format

An object of class `matrix` (inherits from `array`) with 78 rows and 2780 columns.

`PCAWG.WGS.SBS.96` *Single Base Substitution (SBS) spectra in trinucleotide context (deprecated). Use* `spectra$PCAWG$SBS96` *instead.*

Description

Single Base Substitution (SBS) spectra in trinucleotide context (deprecated). Use `spectra$PCAWG$SBS96` instead.

Usage

```
PCAWG.WGS.SBS.96
```

Format

An object of class `matrix` (inherits from `array`) with 96 rows and 2780 columns.

PCAWG7 *PCAWG7: A package of data from 'Repertoire of Mutational Signatures in Human Cancer'*

Description

This is a data package with 3 main package variables: `exposure`, `signature`, and `spectra`.

Details

There are also PDF plots of the signatures in `data-raw/plots/`.

The reference for the data is

Alexandrov, L.B., Kim, J., Haradhvala, N.J. et al. The repertoire of mutational signatures in human cancer. Nature 578, 94-101 (2020). https://doi.org/10.1038/s41586-020-1943-3.

`signature` *PCAWG7 SigProfiler reference signatures.*

Description

PCAWG7 SigProfiler reference signatures.

Usage

```
signature
```

Format

A list with the elements:

genome A list with the elements:

- **SBS96** Strand-agnostic single-base substitutions in trinucleotide context.
- **SBS192** Transcriptionally stranded single-base substitutions in trinucleotide context.
- **DBS78** Strand-agnostic doublet-base substitutions.
- **ID** Strand-agnostic indels.

exome A list with the elements:

- **SBS96** As above, for exome count signatures, which look different than genome count signatures, because of differences in trinucleotide frequencies in exomes versus whole genomes.

Source

Subdirectories of https://www.synapse.org/#!Synapse:syn12009743, 2019 Oct 09, populated by `data-raw/populate.variable.siganture.R`.

`spectra` *PCAWG7 mutational spectra (catalogs).*

Description

PCAWG7 mutational spectra (catalogs).

Usage

```
spectra
```

Format

A list with the elements:

SBS96 Deprecated.

DBS78 Deprecated.

PCAWG A list with the elements:

- **SBS96** Strand-agnostic single-base substitutions in trinucleotide context.
- **SBS192** Single-base substitutions in transcripts based on the sense strand.
- **SBS1536** Strand-agnostic single-base substitutions in pentanucleotide context.
- **DBS78** Strand-agnostic doublet-base substitutions.
- **ID** Strand-agnostic indels.

TCGA A list with the same elements as the `PCAWG` element.

other.genome A list with the same elements as the `PCAWG` element but with `ID` omitted.

other.exome A list with the same elements as the `PCAWG` element but with `ID` omitted.

Source

Files below https://www.synapse.org/#!Synapse:syn11801889, 2019 Oct 09. Populated by `data-raw/spectra/load.package.variable.specra.R`.

Index

∗ **datasets**
exposure, 2
PCAWG.sample.id, 2
PCAWG.WGS.DBS, 3
PCAWG.WGS.SBS.96, 3
signature, 4
spectra, 5

exposure, 2, *4*

PCAWG.sample.id, 2
PCAWG.WGS.DBS, 3
PCAWG.WGS.SBS.96, 3
PCAWG7, 4

signature, *4*, 4
spectra, *3*, *4*, 5