

ILLUMINAHUMANMETHYLATION450K.DB

December 11, 2017

ILLUMINAHUMANMETHYLATION450KACCNUM

Map Manufacturer identifiers to Accession Numbers

Description

ILLUMINAHUMANMETHYLATION450KACCNUM is an R object that contains mappings between a manufacturer's identifiers and manufacturers accessions.

Details

For chip packages such as this, the ACCNUM mapping comes directly from the manufacturer. This is different from other mappings which are mapped onto the probes via an Entrez Gene identifier.

Each manufacturer identifier maps to a vector containing a GenBank accession number.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

Examples

```
x <- ILLUMINAHUMANMETHYLATION450KACCNUM
# Get the probe identifiers that are mapped to an ACCNUM
## mapped_probes <- mappedkeys(x)
# Convert to a list
## xx <- as.list(x[mapped_probes])
## if(length(xx) > 0) {
## # Get the ACCNUM for the first five probes
## xx[1:5]
## # Get the first one
## xx[[1]]
## }
```

IlluminaHumanMethylation450kALIAS2PROBE

Map between Common Gene Symbol Identifiers and Manufacturer Identifiers

Description

IlluminaHumanMethylation450kALIAS is an R object that provides mappings between common gene symbol identifiers and manufacturer identifiers.

Details

Each gene symbol is mapped to a named vector of manufacturer identifiers. The name represents the gene symbol and the vector contains all manufacturer identifiers that are found for that symbol. An NA is reported for any gene symbol that cannot be mapped to any manufacturer identifiers.

This mapping includes ALL gene symbols including those which are already listed in the SYMBOL map. The SYMBOL map is meant to only list official gene symbols, while the ALIAS maps are meant to store all used symbols.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

Examples

```
# Convert the object to a list
xx <- as.list(IlluminaHumanMethylation450kALIAS2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first two aliases
  xx[1:2]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450k.db

Bioconductor annotation data package

Description

Welcome to the IlluminaHumanMethylation450k.db annotation Package. The purpose of this package is to provide detailed information about the IlluminaHumanMethylation450k platform. This package is updated biannually.

You can learn what objects this package supports with the following command:

```
ls("package:IlluminaHumanMethylation450k.db")
```

Each of these objects has their own manual page detailing where relevant data was obtained along with some examples of how to use it.

Examples

```
ls("package:IlluminaHumanMethylation450k.db")
```

IlluminaHumanMethylation450kBLAME

Who checked the latest build of these annotations into version control?

Description

IlluminaHumanMethylation450kBLAME answers the important question of where to point the finger when it turns out that an annotation is out of date. It comes with a free money-back guarantee; your purchase price will be refunded.

Details

Typically this will be an overworked and underprepared grad student or RA.

Examples

```
IlluminaHumanMethylation450kBLAME
```

IlluminaHumanMethylation450kBUILD

Divulge the GoldenPath genome build ID for these annotations.

Description

This is a static map that returns the genome build for the probes in the DB.

Details

Usually this will be number between 1 and something resembling infinity.

Examples

```
IlluminaHumanMethylation450kBUILD
```

IlluminaHumanMethylation450kCHR

Map Manufacturer IDs to Chromosomes

Description

IlluminaHumanMethylation450kCHR is an R object that provides mappings between a manufacturer identifier and the chromosome that contains the gene of interest.

Details

Each manufacturer identifier maps to a vector of chromosomes. Due to inconsistencies that may exist at the time the object was built, the vector may contain more than one chromosome (e.g., the identifier may map to more than one chromosome). If the chromosomal location is unknown, the vector will contain an NA.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

Examples

```
x <- IlluminaHumanMethylation450kCHR
# Get the probe identifiers that are mapped to a chromosome
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the CHR for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kCHR36

Chromosome annotations against genome build 36

Description

What chromosome does the target sequence for a probe align to, in build 36?

Examples

```
x <- IlluminaHumanMethylation450kCHR36
xx <- as.list(x)[1:10]
xx
```

IlluminaHumanMethylation450kCHR37

Chromosome annotations against genome build 37

Description

What chromosome does the target sequence for a probe align to, in build 37?

Examples

```
x <- IlluminaHumanMethylation450kCHR37
xx <- as.list(x)[1:10]
xx
```

IlluminaHumanMethylation450kCHRENGTHS

A named vector for the length of each of the chromosomes

Description

IlluminaHumanMethylation450kCHRENGTHS provides the length measured in base pairs for each of the chromosomes.

Details

This is a named vector with chromosome numbers as the names and the corresponding lengths for chromosomes as the values.

Total lengths of chromosomes were derived by calculating the number of base pairs on the sequence string for each chromosome.

Examples

```
tt <- IlluminaHumanMethylation450kCHRENGTHS
# Length of chromosome 1
tt["1"]
```

IlluminaHumanMethylation450kCHRLOC

Map Manufacturer IDs to Chromosomal Location

Description

IlluminaHumanMethylation450kCHRLOC is an R object that maps manufacturer identifiers to the starting position of the gene. The position of a gene is measured as the number of base pairs.

The CHRLOCEND mapping is the same as the CHRLOC mapping except that it specifies the ending base of a gene instead of the start.

Details

Each manufacturer identifier maps to a named vector of chromosomal locations, where the name indicates the chromosome. Due to inconsistencies that may exist at the time the object was built, these vectors may contain more than one chromosome and/or location. If the chromosomal location is unknown, the vector will contain an NA.

Chromosomal locations on both the sense and antisense strands are measured as the number of base pairs from the p (5' end of the sense strand) to q (3' end of the sense strand) arms. Chromosomal locations on the antisense strand have a leading "-" sign (e. g. -1234567).

Since some genes have multiple start sites, this field can map to multiple locations.

Mappings were based on data provided by: UCSC Genome Bioinformatics (Homo sapiens) <ftp://hgdownload.cse.ucsc.edu>
With a date stamp from the source of: 2010-Mar22

Examples

```
x <- IlluminaHumanMethylation450kCHRLOC
# Get the probe identifiers that are mapped to chromosome locations
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the CHRLOC for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kCOLORCHANNEL

Color Channel for locus assay

Description

IlluminaHumanMethylation450kCOLORCHANNEL is an R object that provides mappings between a manufacturer identifier and the color channel for the assay for that locus

Details

Value is a simple Grn or Red for each probe, based simply on the Illumina manifest,

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 2010-Oct-10

Examples

```
x <- IlluminaHumanMethylation450kCOLORCHANNEL
# Get the probe identifiers that are mapped
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the values for the first five probes
  xx[1:5]
}
```

IlluminaHumanMethylation450kCPG36

CpG location annotations against genome build 36

Description

Two genome builds have been widely used in the annotation of Illumina's DNA methylation microarrays: GRC36 and GRC37. Both the chromosomal and CpG target locations for the 450k probes have been mapped against both.

Examples

```
as.list(IlluminaHumanMethylation450kCPG36)[10:20]
```

`IlluminaHumanMethylation450kCPG37`*CpG location annotations against genome build 37*

Description

Two genome builds have been widely used in the annotation of Illumina's DNA methylation microarrays: GRC36 and GRC37. Both the chromosomal and CpG target locations for the 450k probes have been mapped against both.

Examples

```
as.list(IlluminaHumanMethylation450kCPG37)[10:20]
```

`IlluminaHumanMethylation450kCPGCOORDINATE`*Map Manufacturer IDs to genomic coordinate*

Description

`IlluminaHumanMethylation450kCPGCOORDINATE` is an R object that provides mappings between a manufacturer identifier and the CpG coordinate as defined by Illumina in the manifest.

Details

Simple probe mapping to CpG coordinate within the human genome, Build 37, as defined by Illumina.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 19-Sep-2009

Examples

```
x <- IlluminaHumanMethylation450kCPGCOORDINATE
# Get the probe identifiers that are mapped to a CpG coordinate
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the CHR for the first five probes
  xx[1:5]
}
```

IlluminaHumanMethylation450kCPGILOCATION

Map Illumina probeIDs to the the nearest UCSC CpG island, if any

Description

IlluminaHumanMethylation450kCPGILOCATION maps between Illumina probe IDs and the UCSC CpG island features, if any, with which they are associated.

Details

Map a probe ID to CpG island with which it is associated, and find what sort of relationship the probe has to said island (in it, shore, etc.)

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kCPGILOCATION)[150:200]
```

IlluminaHumanMethylation450kCPGNAME

Map Illumina probeIDs to the the nearest UCSC CpG island, if any

Description

IlluminaHumanMethylation450kCPGNAME maps between Illumina probe IDs and the UCSC CpG island, if any, with which they are associated.

Details

Map a probe ID to the CpG island with which it is associated.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kCPGNAME)[150:200]
```

`IlluminaHumanMethylation450kCPGIRELATION`*Map Illumina probeIDs to the the nearest UCSC CpG island, if any*

Description

IlluminaHumanMethylation450kCPGIRELATION maps between Illumina probe IDs and the UCSC CpG island, if any, with which they are associated.

Details

What relationship does a probe have to the nearest UCSC CpG island?

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kCPGIRELATION)[150:200]
```

`IlluminaHumanMethylation450kCPGS`*Number of CpG dinucleotides in target genomic sequence*

Description

IlluminaHumanMethylation450kCPGS is an R object that provides mappings between Illumina probe IDs and the number of CpGs in the source sequence targeted by Illumina, based on information in their manifest.

Details

Probe mapping to CpG density within the human genome, Build 36, as defined by Illumina.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kCPGS)[150:200]
```

IlluminaHumanMethylation450kDESIGN

Probe design type for HumanMethylation450k arrays

Description

IlluminaHumanMethylation450kDESIGN provides a mapping from probe ID to the design type of that probe. On 27k arrays, all probes had the same chemistry in the sense that paired allelic probes always fluoresced in the same channel. On the newer 450k arrays, there are probes of the 27k design (Design I), but there are also probes which are extended base-by-base with either Cy5 or Cy3 fluorophores, leading to a differential fluorescence in the two channels. This provides a dense method of assaying methylation at closely spaced CpG bases, but also requires some accommodation in preprocessing and detection calls. By using a mapped database table to keep track of which probes are Design I vs II, the 'methyumi' package is able to quickly preprocess and later normalize the two chemistries in a fashion most appropriate to each.

Details

The map returns either a 'I' or 'II' depending on probe chemistry.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 2010-Oct-10

Examples

```
x <- IlluminaHumanMethylation450kDESIGN
xx <- as.list(x)[1:10]
if(length(xx) > 0) {
  # Get the values for the first five probes
  xx[1:5]
}
```

IlluminaHumanMethylation450kDHS

Is the probe targeted at a DNase hypersensitive site in the genome?

Description

IlluminaHumanMethylation450kDHS provides mappings to Illumina's DHS annotations for the 450k probes, which were experimentally determined.

Details

Illumina/UCSC DNase hypersensitive site annotations.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kDHS)[10:20]
```

IlluminaHumanMethylation450kDMR

Is the probe targeted at a differentially methylated region? What sort?

Description

IlluminaHumanMethylation450kDMR provides mappings to Illumina's DMR annotations for the 450k probes, which indicate what sort of differentially methylated region is associated with the probe target sequence (if any).

Details

Illumina/UCSC DMR annotations.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kDMR)[10:20]
```

IlluminaHumanMethylation450kENHANCER

Is the probe associated with an enhancer binding site?

Description

IlluminaHumanMethylation450kENHANCER provides mappings to Illumina's ENHANCER annotations for the 450k probes (which in turn were suggested by Ben Berman).

Details

Illumina/UCSC enhancer binding site annotations.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kENHANCER)[10:20]
```

 IlluminaHumanMethylation450kENSEMBL

Map Ensembl gene accession numbers with Entrez Gene identifiers

Description

IlluminaHumanMethylation450kENSEMBL is an R object that contains mappings between Entrez Gene identifiers and Ensembl gene accession numbers.

Details

This object is a simple mapping of Entrez Gene identifiers <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene> to Ensembl gene Accession Numbers.

Mappings were based on data provided by BOTH of these sources: <http://www.ensembl.org/biomart/martview/> <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA>

This mapping is a combination of NCBI to ensembl IDs from BOTH NCBI and ensembl. Users who wish to only use mappings from NCBI are encouraged to see the `ncbi2ensembl` table in the appropriate organism package. Users who wish to only use mappings from ensembl are encouraged to see the `ensembl2ncbi` table which is also found in the appropriate organism packages. These mappings are based upon the `ensembl` table which contains data from BOTH of these sources in an effort to maximize the chances that you will find a match.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

Examples

```
x <- IlluminaHumanMethylation450kENSEMBL
# Get the entrez gene IDs that are mapped to an Ensembl ID
mapped_genes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_genes])
if(length(xx) > 0) {
  # Get the Ensembl gene IDs for the first five genes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
#For the reverse map ENSEMBL2PROBE:
# Convert to a list
xx <- as.list(IlluminaHumanMethylation450kENSEMBL2PROBE)
if(length(xx) > 0){
  # Gets the entrez gene IDs for the first five Ensembl IDs
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kENTREZID

Map between Manufacturer Identifiers and Entrez Gene

Description

IlluminaHumanMethylation450kENTREZID is an R object that provides mappings between manufacturer identifiers and Entrez Gene identifiers.

Details

Each manufacturer identifier is mapped to a vector of Entrez Gene identifiers. An NA is assigned to those manufacturer identifiers that can not be mapped to an Entrez Gene identifier at this time.

If a given manufacturer identifier can be mapped to different Entrez Gene identifiers from various sources, we attempt to select the common identifiers. If a consensus cannot be determined, we select the smallest identifier.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>

Examples

```
x <- IlluminaHumanMethylation450kENTREZID
# Get the probe identifiers that are mapped to an ENTREZ Gene ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the ENTREZID for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kENZYME

Map between Manufacturer IDs and Enzyme Commission (EC) Numbers

Description

IlluminaHumanMethylation450kENZYME is an R object that provides mappings between manufacturer identifiers and EC numbers.

Details

Each manufacturer identifier maps to a named vector containing the EC number that corresponds to the enzyme produced by that gene. The name corresponds to the manufacturer identifier. If this information is unknown, the vector will contain an NA.

Enzyme Commission numbers are assigned by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology <http://www.chem.qmw.ac.uk/iubmb/enzyme/> to allow enzymes to be identified.

An Enzyme Commission number is of the format EC x.y.z.w, where x, y, z, and w are numeric numbers. In IlluminaHumanMethylation450kENZYME2PROBE, EC is dropped from the Enzyme Commission numbers.

Enzyme Commission numbers have corresponding names that describe the functions of enzymes in such a way that EC x is a more general description than EC x.y that in turn is a more general description than EC x.y.z. The top level EC numbers and names are listed below:

EC 1 oxidoreductases

EC 2 transferases

EC 3 hydrolases

EC 4 lyases

EC 5 isomerases

EC 6 ligases

The EC name for a given EC number can be viewed at <http://www.chem.qmul.ac.uk/iupac/jcbn/index.html#6>

Mappings between probe identifiers and enzyme identifiers were obtained using files provided by: KEGG GENOME <ftp://ftp.genome.jp/pub/kegg/genomes> With a date stamp from the source of: 2010-Sep7

References

<ftp://ftp.genome.ad.jp/pub/kegg/pathways>

Examples

```
x <- IlluminaHumanMethylation450kENZYME
# Get the probe identifiers that are mapped to an EC number
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the ENZYME for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kENZYME2PROBE

Map between Enzyme Commission Numbers and Manufacturer Identifiers

Description

IlluminaHumanMethylation450kENZYME2PROBE is an R object that maps Enzyme Commission (EC) numbers to manufacturer identifiers.

Details

Each EC number maps to a named vector containing all of the manufacturer identifiers that correspond to the gene that produces that enzyme. The name of the vector corresponds to the EC number.

Enzyme Commission numbers are assigned by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology <http://www.chem.qmw.ac.uk/iubmb/enzyme/> to allow enzymes to be identified.

An Enzyme Commission number is of the format EC x.y.z.w, where x, y, z, and w are numeric numbers. In IlluminaHumanMethylation450kENZYME2PROBE, EC is dropped from the Enzyme Commission numbers.

Enzyme Commission numbers have corresponding names that describe the functions of enzymes in such a way that EC x is a more general description than EC x.y that in turn is a more general description than EC x.y.z. The top level EC numbers and names are listed below:

EC 1 oxidoreductases

EC 2 transferases

EC 3 hydrolases

EC 4 lyases

EC 5 isomerases

EC 6 ligases

The EC name for a given EC number can be viewed at <http://www.chem.qmul.ac.uk/iupac/jcbn/index.html#6>

Mappings were based on data provided by: KEGG GENOME <ftp://ftp.genome.jp/pub/kegg/genomes>
With a date stamp from the source of: 2010-Sep7

References

<ftp://ftp.genome.ad.jp/pub/kegg/pathways>

Examples

```
# Convert to a list
xx <- as.list(IlluminaHumanMethylation450kENZYME2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first five enzyme
  #commission numbers
  xx[1:5]
  # Get the first one
```

```
xx[[1]]
}
```

IlluminaHumanMethylation450kFANTOM

Is the probe associated with an enhancer binding site?

Description

IlluminaHumanMethylation450kFANTOM provides mappings to Illumina's FANTOM annotations for the 450k probes (which in turn were suggested by Ben Berman).

Details

Illumina/FANTOM annotations.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kFANTOM)[10:20]
```

IlluminaHumanMethylation450kGENEBODY

Determine whether a probe maps into a gene body

Description

IlluminaHumanMethylation450kGENEBODY is an R object that maps onto a Boolean column in the probe design table indicating whether the target sequence maps onto the gene bodies of any of the associated transcripts.

Details

Value is a simple TRUE or FALSE for each probe.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 19-Sep-2009

Examples

```
x <- IlluminaHumanMethylation450kGENEBODY
# Get the probe identifiers that are mapped
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the values for the first five probes
  xx[1:5]
}
```

IlluminaHumanMethylation450kGENENAME

Map between Manufacturer IDs and Genes

Description

IlluminaHumanMethylation450kGENENAME is an R object that maps manufacturer identifiers to the corresponding gene name.

Details

Each manufacturer identifier maps to a named vector containing the gene name. The vector name corresponds to the manufacturer identifier. If the gene name is unknown, the vector will contain an NA.

Gene names currently include both the official (validated by a nomenclature committee) and preferred names (interim selected for display) for genes. Efforts are being made to differentiate the two by adding a name to the vector.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

Examples

```
x <- IlluminaHumanMethylation450kGENENAME
# Get the probe identifiers that are mapped to a gene name
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the GENENAME for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kGO

Map between Manufacturer IDs and Gene Ontology (GO)

Description

IlluminaHumanMethylation450kGO is an R object that provides mappings between manufacturer identifiers and the GO identifiers that they are directly associated with.

Details

Each Entrez Gene identifier is mapped to a list of lists. The names on the outer list are GO identifiers. Each inner list consists of three named elements: GOID, Ontology, and Evidence.

The GOID element matches the GO identifier named in the outer list and is included for convenience when processing the data using 'lapply'.

The Ontology element indicates which of the three Gene Ontology categories this identifier belongs to. The categories are biological process (BP), cellular component (CC), and molecular function (MF).

The Evidence element contains a code indicating what kind of evidence supports the association of the GO identifier to the Entrez Gene id. The evidence codes in use include:

IMP: inferred from mutant phenotype

IGI: inferred from genetic interaction

IPI: inferred from physical interaction

ISS: inferred from sequence similarity

IDA: inferred from direct assay

IEP: inferred from expression pattern

IEA: inferred from electronic annotation

TAS: traceable author statement

NAS: non-traceable author statement

ND: no biological data available

IC: inferred by curator

Mappings between probe identifiers and GO information were obtained through their mappings to Entrez Gene identifiers. NAs are assigned to probe identifiers that can not be mapped to any Gene Ontology information. Mappings between Gene Ontology identifiers and Gene Ontology terms and other information are available in a separate data package named GO.

Mappings were based on data provided by: Gene Ontology <ftp://ftp.geneontology.org/pub/go/godatabase/archive/latest-lite/> With a date stamp from the source of: 20100904

Examples

```
x <- IlluminaHumanMethylation450kGO
# Get the probe identifiers that are mapped to a GO ID
mapped_probes <- mappedkeys(x)[1:100]
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Try the first one
  got <- xx[[1]]
  got[[1]][["GOID"]]
  got[[1]][["Ontology"]]
  got[[1]][["Evidence"]]
}
```

IlluminaHumanMethylation450kGO2ALLPROBES

Map between Gene Ontology (GO) Identifiers and all Manufacturer Identifiers in the subtree

Description

IlluminaHumanMethylation450kGO2ALLPROBES is an R object that provides mappings between a given GO identifier and all manufacturer identifiers annotated at that GO term or one of its children in the GO ontology.

Details

GO consists of three ontologies—molecular function (MF), biological process (BP), and cellular component (CC). All ontologies are structured as directed acyclic graphs (DAGs). Each node in each DAG (tree) is a GO term (id) associated with a named vector of manufacturer identifiers. The name associated with each manufacturer id corresponds to the evidence code for that GO identifier. This object *IlluminaHumanMethylation450kGO2ALLPROBES* maps between a given GO identifier and all manufacturer identifiers annotated at that GO term or one of its children in the GO ontology.

The evidence code indicates what kind of evidence supports the association between the GO and Entrez Gene identifiers. Evidence codes currently in use include:

IMP - inferred from mutant phenotype

IGI - inferred from genetic interaction

IPI - inferred from physical interaction

ISS - inferred from sequence similarity

IDA - inferred from direct assay

IEP - inferred from expression pattern

IEA - inferred from electronic annotation

TAS - traceable author statement

NAS - non-traceable author statement

ND - no biological data available

IC - inferred by curator

A GO identifier may be mapped to the same manufacturer identifier more than once but the evidence code can be different. Mappings between Gene Ontology identifiers and Gene Ontology terms and other information are available in a separate data package named GO.

Mappings were based on data provided by:

Gene Ontology <ftp://ftp.geneontology.org/pub/go/godatabase/archive/latest-lite/> With a date stamp from the source of: 20100904 and Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA/> With a date stamp from the source of: 2010-Sep7

References

<ftp://ftp.ncbi.nlm.nih.gov/gene/DATA/>

Examples

```
# Convert to a list
# xx <- as.list(IlluminaHumanMethylation450kGO2ALLPROBES)
# if(length(xx) > 0){
#   # Get the probe identifiers for the top 2nd and 3rd GO identifiers
#   goids <- xx[2:3]
#   # Get all the probe identifiers for the first element of goids
#   goids[[1]]
#   # Evidence code for the mappings
#   names(goids[[1]])
# }
```

 IlluminaHumanMethylation450kGO2PROBE

Map between Gene Ontology (GO) and Manufacturer Identifiers

Description

IlluminaHumanMethylation450kGO2PROBE is an R object that provides mappings between GO identifiers and manufacturer identifiers.

Details

Each GO term maps to a named vector of manufacturer identifiers. The name associated with each manufacturer identifier corresponds to the evidence code for that GO identifier. The evidence code indicates what kind of evidence supports the association between the GO and Entrez Gene identifiers. Evidence codes currently in use include:

IMP - inferred from mutant phenotype
 IGI - inferred from genetic interaction
 IPI - inferred from physical interaction
 ISS - inferred from sequence similarity
 IDA - inferred from direct assay
 IEP - inferred from expression pattern
 IEA - inferred from electronic annotation
 TAS - traceable author statement
 NAS - non-traceable author statement
 ND - no biological data available
 IC - inferred by curator

A GO identifier may be mapped to the same probe identifier more than once but the evidence code can be different. Mappings between Gene Ontology identifiers and Gene Ontology terms and other information are available in a separate data package named GO.

Mappings were based on data provided by:

Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

References

<ftp://ftp.ncbi.nlm.nih.gov/gene/DATA/>

Examples

```
# Convert to a list
# xx <- as.list(IlluminaHumanMethylation450kGO2PROBE)
# if(length(xx) > 0){
#   # Get the probe identifiers for the top 2nd and 3rd GO identifiers
#   goids <- xx[2:3]
#   # Get the probe identifiers for the first element of goids
#   goids[[1]]
#   # Evidence code for the mappings
#   names(goids[[1]])
#}
```

`IlluminaHumanMethylation450kISCPGISLAND`*Determine whether a probe is within a CpG island*

Description

`IlluminaHumanMethylation450kISCPGISLAND` is an R object that provides mappings between a manufacturer identifier and whether or not the CpG site being assayed meets Illumina's criteria for being in a CpG island as defined by Illumina in the manifest.

Details

Value is a simple TRUE or FALSE for each probe, based simply on the Illumina manifest, TRUE meaning that the CpG site is within a CpG island.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 19-Sep-2009

Examples

```
x <- IlluminaHumanMethylation450kISCPGISLAND
# Get the probe identifiers that are mapped
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the values for the first five probes
  xx[1:5]
}
```

`IlluminaHumanMethylation450kMAP`*Map between Manufacturer Identifiers and cytogenetic maps/bands*

Description

`IlluminaHumanMethylation450kMAP` is an R object that provides mappings between manufacturer identifiers and cytoband locations.

Details

Each manufacturer identifier is mapped to a vector of cytoband locations. The vector length may be one or longer, if there are multiple reported chromosomal locations for a given gene. An NA is reported for any manufacturer identifiers that cannot be mapped to a cytoband at this time.

Cytogenetic bands for most higher organisms are labeled p1, p2, p3, q1, q2, q3 (p and q are the p and q arms), etc., counting from the centromere out toward the telomeres. At higher resolutions, sub-bands can be seen within the bands. The sub-bands are also numbered from the centromere out toward the telomere. Thus, a label of 7q31.2 indicates that the band is on chromosome 7, q arm, band 3, sub-band 1, and sub-sub-band 2.

The physical location of each band on a chromosome can be obtained from another object named "organism"CYTOLOC in a separate data package for human(`humanCHRLOC`), mouse(`mouseCHRLOC`), and rat(`ratCHRLOC`).

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

References

<http://www.ncbi.nlm.nih.gov>

Examples

```
x <- IlluminaHumanMethylation450kMAP
# Get the probe identifiers that are mapped to any cytoband
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the MAP for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kMAPCOUNTS

Number of mapped keys for the maps in package IlluminaHumanMethylation450k.db

Description

IlluminaHumanMethylation450kMAPCOUNTS provides the "map count" (i.e. the count of mapped keys) for each map in package IlluminaHumanMethylation450k.db.

Details

This "map count" information is precalculated and stored in the package annotation DB. This allows some quality control and is used by the `checkMAPCOUNTS` function defined in AnnotationDbi to compare and validate different methods (like `count.mappedkeys(x)` or `sum(!is.na(as.list(x)))`) for getting the "map count" of a given map.

See Also

[mappedkeys](#), [count.mappedkeys](#), [checkMAPCOUNTS](#)

Examples

```
## IlluminaHumanMethylation450kMAPCOUNTS
## mapnames <- names(IlluminaHumanMethylation450kMAPCOUNTS)
## IlluminaHumanMethylation450kMAPCOUNTS[mapnames[1]]
## x <- get(mapnames[1])
## sum(!is.na(as.list(x)))
## count.mappedkeys(x) # much faster!

## Check the map count of all maps in package IlluminaHumanMethylation450k.db
## checkMAPCOUNTS("IlluminaHumanMethylation450k.db")
```

`IlluminaHumanMethylation450kMETHYL27`

Determine whether a probe is also found on the 27k chip

Description

`IlluminaHumanMethylation450kMETHYL27` is a mapping to a boolean column in the annotation database that indicates whether a 450k probe was on the 27k chips.

Details

Value is a simple TRUE or FALSE for each probe. Note: the 27k probes on the 450k chips do not all retain 27k probe chemistry!

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 19-Sep-2009

Examples

```
x <- IlluminaHumanMethylation450kMETHYL27
xx <- as.list(x)
if(length(xx) > 0) {
  # Get the values for the first five probes
  xx[1:5]
}
```

`IlluminaHumanMethylation450kNUID`

Map Illumina probe IDs to the NuID representation of the target sequence

Description

`IlluminaHumanMethylation450kNUID` is an R object that provides mappings between Illumina probe IDs and the NuID string generated by `lumi::seq2id(sourceSeq)` based on the target sequences provided by Illumina in the 450k manifest.

Details

Probe ID to NuID conversion, without the agonizing wait.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kNUID)[150:200]
```

IlluminaHumanMethylation450kOMIM

Map between Manufacturer Identifiers and Mendelian Inheritance in Man (MIM) identifiers

Description

IlluminaHumanMethylation450kOMIM is an R object that provides mappings between manufacturer identifiers and OMIM identifiers.

Details

Each manufacturer identifier is mapped to a vector of OMIM identifiers. The vector length may be one or longer, depending on how many OMIM identifiers the manufacturer identifier maps to. An NA is reported for any manufacturer identifier that cannot be mapped to an OMIM identifier at this time.

OMIM is based upon the book Mendelian Inheritance in Man (V. A. McKusick) and focuses primarily on inherited or heritable genetic diseases. It contains textual information, pictures, and reference information that can be searched using various terms, among which the MIM number is one.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene> <http://www3.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>

Examples

```
x <- IlluminaHumanMethylation450kOMIM
# Get the probe identifiers that are mapped to a OMIM ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the OMIM for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kORGANISM

The Organism information for IlluminaHumanMethylation450k

Description

IlluminaHumanMethylation450kORGANISM is an R object that contains a single item: a character string that names the organism for which IlluminaHumanMethylation450k was built. IlluminaHumanMethylation450kORGPkg is an R object that contains a character vector with the name of the organism package that a chip package depends on for its gene-centric annotation.

Details

Although the package name is suggestive of the organism for which it was built, IlluminaHumanMethylation450kORGANISM provides a simple way to programmatically extract the organism name. IlluminaHumanMethylation450kORGPKG provides a simple way to programmatically extract the name of the parent organism package. The parent organism package is a strict dependency for chip packages as this is where the gene centric information is ultimately extracted from. The full package name will always be this string plus the extension ".db". But most programatic acces will not require this extension, so its more convenient to leave it out.

Examples

```
IlluminaHumanMethylation450kORGANISM
IlluminaHumanMethylation450kORGPKG
```

```
IlluminaHumanMethylation450kPATH
```

Mappings between probe identifiers and KEGG pathway identifiers

Description

KEGG (Kyoto Encyclopedia of Genes and Genomes) maintains pathway data for various organisms. IlluminaHumanMethylation450kPATH maps probe identifiers to the identifiers used by KEGG for pathways in which the genes represented by the probe identifiers are involved

Details

Each KEGG pathway has a name and identifier. Pathway name for a given pathway identifier can be obtained using the KEGG data package that can either be built using AnnBuilder or downloaded from Bioconductor <http://www.bioconductor.org>.

Graphic presentations of pathways are searchable at url <http://www.genome.ad.jp/kegg/pathway.html> by using pathway identifiers as keys.

Mappings were based on data provided by: KEGG GENOME <ftp://ftp.genome.jp/pub/kegg/genomes> With a date stamp from the source of: 2010-Sep7

References

<http://www.genome.ad.jp/kegg/>

Examples

```
x <- IlluminaHumanMethylation450kPATH
# Get the probe identifiers that are mapped to a KEGG pathway ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the PATH for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kPATH2PROBE

Map between Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway identifiers and Manufacturer Identifiers

Description

IlluminaHumanMethylation450kPATH2PROBE is an R object that provides mappings between KEGG identifiers and manufacturer identifiers.

Details

Each KEGG identifier is mapped to a named vector of manufacturer identifiers. The name represents the KEGG identifier and the vector contains all manufacturer identifiers that are found in that particular pathway. An NA is reported for any KEGG identifier that cannot be mapped to any manufacturer identifiers.

Pathway name for a given pathway identifier can be obtained using the KEGG data package that can either be built using AnnBuilder or downloaded from Bioconductor <http://www.bioconductor.org>.

Graphic presentations of pathways are searchable at <http://www.genome.ad.jp/kegg/pathway.html> using pathway identifiers as keys.

Mappings were based on data provided by: KEGG GENOME <ftp://ftp.genome.jp/pub/kegg/genomes> With a date stamp from the source of: 2010-Sep7

References

<http://www.genome.ad.jp/kegg/>

Examples

```
# Convert the object to a list
xx <- as.list(IlluminaHumanMethylation450kPATH2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first two pathway identifiers
  xx[1:2]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kPFAM

Map Manufacturer IDs to Pfam IDs

Description

IlluminaHumanMethylation450kPFAM is an R object that provides mappings between a manufacturer identifier and the associated Pfam identifiers.

Details

Each manufacturer identifier maps to a named vector of Pfam identifiers. The name for each Pfam identifier is the IPI accession number where this Pfam identifier is found.

If the Pfam is a named NA, it means that the associated Entrez Gene id of this manufacturer identifier is found in an IPI entry of the IPI database, but there is no Pfam identifier in the entry.

If the Pfam is a non-named NA, it means that the associated Entrez Gene id of this manufacturer identifier is not found in any IPI entry of the IPI database.

Mappings were based on data provided by: The International Protein Index <ftp://ftp.ebi.ac.uk/pub/databases/IPI/current> With a date stamp from the source of: 2010-Aug19

References

Kersey P. J., Duarte J., Williams A., Karavidopoulou Y., Birney E., Apweiler R. The International Protein Index: An integrated database for proteomics experiments. *Proteomics* 4(7): 1985-1988 (2004) <http://www.ebi.ac.uk/IPI/IPIhelp.html>

Examples

```
x <- IlluminaHumanMethylation450kPFAM
# Get the probe identifiers that are mapped to any Pfam ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
# randomly display 10 probes
sample(xx, 10)
```

IlluminaHumanMethylation450kPMID

Map between Manufacturer Identifiers and PubMed Identifiers

Description

IlluminaHumanMethylation450kPMID is an R object that provides mappings between manufacturer identifiers and PubMed identifiers.

Details

Each manufacturer identifier is mapped to a named vector of PubMed identifiers. The name associated with each vector corresponds to the manufacturer identifier. The length of the vector may be one or greater, depending on how many PubMed identifiers a given manufacturer identifier is mapped to. An NA is reported for any manufacturer identifier that cannot be mapped to a PubMed identifier.

Titles, abstracts, and possibly full texts of articles can be obtained from PubMed by providing a valid PubMed identifier. The `pubmed` function of `annotate` can also be used for the same purpose.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed>

Examples

```

x <- IlluminaHumanMethylation450kPMID
# Get the probe identifiers that are mapped to any PubMed ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0){
  # Get the PubMed identifiers for the first two probe identifiers
  xx[1:2]
  # Get the first one
  xx[[1]]
  if(interactive() && !is.null(xx[[1]]) && !is.na(xx[[1]]))
    && require(annotate)){
    # Get article information as XML files
    xmls <- pubmed(xx[[1]], disp = "data")
    # View article information using a browser
    pubmed(xx[[1]], disp = "browser")
  }
}

```

IlluminaHumanMethylation450kPMID2PROBE

Map between PubMed Identifiers and Manufacturer Identifiers

Description

IlluminaHumanMethylation450kPMID2PROBE is an R object that provides mappings between PubMed identifiers and manufacturer identifiers.

Details

Each PubMed identifier is mapped to a named vector of manufacturer identifiers. The name represents the PubMed identifier and the vector contains all manufacturer identifiers that are represented by that PubMed identifier. The length of the vector may be one or longer, depending on how many manufacturer identifiers are mapped to a given PubMed identifier.

Titles, abstracts, and possibly full texts of articles can be obtained from PubMed by providing a valid PubMed identifier. The `pubmed` function of `annotate` can also be used for the same purpose

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

Examples

```

# Convert the object to a list
xx <- as.list(IlluminaHumanMethylation450kPMID2PROBE)
if(length(xx) > 0){
  # Get the probe identifiers for the first two PubMed identifiers
  xx[1:2]
  # Get the first one
  xx[[1]]
  if(interactive() && require(annotate)){
    # Get article information as XML files for a PubMed id
    xmls <- pubmed(names(xx)[1], disp = "data")
  }
}

```

```
    # View article information using a browser
    pubmed(names(xx)[1], disp = "browser")
  }
}
```

IlluminaHumanMethylation450kPROBELOCATION

Map Illumina probe IDs to the location of the target sequence

Description

IlluminaHumanMethylation450kPROBELOCATION is an R object that maps between Illumina probe IDs and the transcript(s) they target, along with a description of where in the associated gene the interrogated CpG site happens to be.

Details

Map a probe ID to all of the transcripts it is associated with, and where in the corresponding gene (body, 5' UTR, 3' UTR, within 1500bp of the TSS, etc.) the probe target sequence aligns.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kPROBELOCATION)[1:10]
```

IlluminaHumanMethylation450kPROSITE

Map Manufacturer IDs to PROSITE ID

Description

IlluminaHumanMethylation450kPROSITE is an R object that provides mappings between a manufacturer identifier and the associated PROSITE identifiers.

Details

Each manufacturer identifier maps to a named vector of PROSITE identifiers. The name for each PROSITE identifier is the IPI accession number where this PROSITE identifier is found.

If the PROSITE is a named NA, it means that the associated Entrez Gene id of this manufacturer identifier is found in an IPI entry of the IPI database, but there is no PROSITE identifier in the entry.

If the PROSITE is a non-named NA, it means that the associated Entrez Gene id of this manufacturer identifier is not found in any IPI entry of the IPI database.

Mappings were based on data provided by: The International Protein Index <ftp://ftp.ebi.ac.uk/pub/databases/IPI/current>
With a date stamp from the source of: 2010-Aug19

References

Kersey P. J., Duarte J., Williams A., Karavidopoulou Y., Birney E., Apweiler R. The International Protein Index: An integrated database for proteomics experiments. *Proteomics* 4(7): 1985-1988 (2004) <http://www.ebi.ac.uk/IPI/IPIhelp.html>

Examples

```
x <- IlluminaHumanMethylation450kPROSITE
# Get the probe identifiers that are mapped to any PROSITE ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xxx <- as.list(x[mapped_probes])
# randomly display 10 probes
xxx[sample(1:length(xxx), 10)]
```

IlluminaHumanMethylation450kRANDOM

Determine whether a probe was randomly selected in the design phase.

Description

Approximately 4000 probes on the 450k array were randomly targeted by selecting unique 50mer matches out of the GRC37 genome assembly. These all return TRUE from this mapping function. Other probes return FALSE.

Details

Value is a simple TRUE or FALSE for each probe.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 19-Sep-2009

Examples

```
x <- IlluminaHumanMethylation450kRANDOM
xx <- as.list(x)
if(length(xx) > 0) {
  # Get the values for the first five probes
  xx[1:5]
}
```

IlluminaHumanMethylation450kREFSEQ

Map between Manufacturer Identifiers and RefSeq Identifiers

Description

IlluminaHumanMethylation450kREFSEQ is an R object that provides mappings between manufacturer identifiers and RefSeq identifiers.

Details

Each manufacturer identifier is mapped to a named vector of RefSeq identifiers. The name represents the manufacturer identifier and the vector contains all RefSeq identifiers that can be mapped to that manufacturer identifier. The length of the vector may be one or greater, depending on how many RefSeq identifiers a given manufacturer identifier can be mapped to. An NA is reported for any manufacturer identifier that cannot be mapped to a RefSeq identifier at this time.

RefSeq identifiers differ in format according to the type of record the identifiers are for as shown below:

NG_XXXXX: RefSeq accessions for genomic region (nucleotide) records

NM_XXXXX: RefSeq accessions for mRNA records

NC_XXXXX: RefSeq accessions for chromosome records

NP_XXXXX: RefSeq accessions for protein records

XR_XXXXX: RefSeq accessions for model RNAs that are not associated with protein products

XM_XXXXX: RefSeq accessions for model mRNA records

XP_XXXXX: RefSeq accessions for model protein records

Where XXXXX is a sequence of integers.

NCBI <http://www.ncbi.nlm.nih.gov/RefSeq/> allows users to query the RefSeq database using RefSeq identifiers.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

References

<http://www.ncbi.nlm.nih.gov> <http://www.ncbi.nlm.nih.gov/RefSeq/>

Examples

```
x <- IlluminaHumanMethylation450kREFSEQ
# Get the probe identifiers that are mapped to any RefSeq ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the REFSEQ for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kREGULATORYGROUP

Map Illumina probeIDs to the the nearest UCSC regulatory feature, if any

Description

IlluminaHumanMethylation450kREGULATORYGROUP maps between Illumina probe IDs and the UCSC regulatory features, if any, with which they are associated.

Details

Map a probe ID to regulatory feature with which it is associated, and find what sort of regulatory region the probe is annotated to.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kREGULATORYGROUP)[30:40]
```

IlluminaHumanMethylation450kREGULATORYLOCATION

Map Illumina probeIDs to the the nearest UCSC regulatory feature, if any

Description

IlluminaHumanMethylation450kREGULATORYLOCATION maps between Illumina probe IDs and the UCSC regulatory features, if any, with which they are associated.

Details

Map a probe ID to regulatory feature with which it is associated, and find what sort of regulatory region the probe is annotated to.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kREGULATORYLOCATION)[10:20]
unique(na.omit(unlist(mget(IlluminaHumanMethylation450kALIAS2PROBE[['EVI1']], IlluminaHumanMethylati
```

IlluminaHumanMethylation450kREGULATORYNAME

Map Illumina probeIDs to the the nearest UCSC regulatory feature, if any

Description

IlluminaHumanMethylation450kREGULATORYNAME maps between Illumina probe IDs and the UCSC regulatory features, if any, with which they are associated.

Details

Map a probe ID to regulatory feature with which it is associated, and find what sort of regulatory region the probe is annotated to.

Mappings were based on data provided by: Illumina <http://Illumina.com> Downloaded 11-Jan-2011

Examples

```
as.list(IlluminaHumanMethylation450kREGULATORYNAME)[30:40]
```

IlluminaHumanMethylation450kSVNID

What is the in-house Subversion ID of this annotation package?

Description

IlluminaHumanMethylation450kSVNID answers the important question of where to point the finger when it turns out an annotation is out of date. Naturally, this particular feature will also indicate if the user has ended up with an out-of-date copy of the package, so that both maintainers and users can track any such discontinuities.

Details

The revision ID should include a numeric change identifier and its author.

Examples

```
x <- IlluminaHumanMethylation450kSVNID
show(x)
```

IlluminaHumanMethylation450kSYMBOL

Map between Manufacturer Identifiers and Gene Symbols

Description

IlluminaHumanMethylation450kSYMBOL is an R object that provides mappings between manufacturer identifiers and gene abbreviations.

Details

Each manufacturer identifier is mapped to an abbreviation for the corresponding gene. An NA is reported if there is no known abbreviation for a given gene.

Symbols typically consist of 3 letters that define either a single gene (ABC) or multiple genes (ABC1, ABC2, ABC3). Gene symbols can be used as key words to query public databases such as Entrez Gene.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>

Examples

```
x <- IlluminaHumanMethylation450kSYMBOL
# Get the probe identifiers that are mapped to a gene symbol
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the SYMBOL for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

IlluminaHumanMethylation450kUNIGENE

Map between Manufacturer Identifiers and UniGene cluster identifiers

Description

IlluminaHumanMethylation450kUNIGENE is an R object that provides mappings between manufacturer identifiers and UniGene identifiers.

Details

Each manufacturer identifier is mapped to a UniGene identifier. An NA is reported if the manufacturer identifier cannot be mapped to UniGene at this time.

A UniGene identifier represents a cluster of sequences of a gene. Using UniGene identifiers one can query the UniGene database for information about the sequences or the Entrez Gene database for information about the genes.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

References

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>

Examples

```
x <- IlluminaHumanMethylation450kUNIGENE
# Get the probe identifiers that are mapped to an UNIGENE ID
mapped_probes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_probes])
if(length(xx) > 0) {
  # Get the UNIGENE for the first five probes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

 IlluminaHumanMethylation450kUNIPROT

Map Uniprot accession numbers with Entrez Gene identifiers

Description

IlluminaHumanMethylation450kUNIPROT is an R object that contains mappings between Entrez Gene identifiers and Uniprot accession numbers.

Details

This object is a simple mapping of Entrez Gene identifiers <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene> to Uniprot Accessions.

Mappings were based on data provided by: Entrez Gene <ftp://ftp.ncbi.nlm.nih.gov/gene/DATA> With a date stamp from the source of: 2010-Sep7

Examples

```
x <- IlluminaHumanMethylation450kUNIPROT
# Get the entrez gene IDs that are mapped to an Uniprot ID
mapped_genes <- mappedkeys(x)
# Convert to a list
xx <- as.list(x[mapped_genes])
if(length(xx) > 0) {
  # Get the Uniprot IDs for the first five genes
  xx[1:5]
  # Get the first one
  xx[[1]]
}
```

 IlluminaHumanMethylation450k_dbconn

Collect information about the package annotation DB

Description

Some convenience functions for getting a connection object to (or collecting information about) the package annotation DB.

Usage

```
IlluminaHumanMethylation450k_dbconn()
IlluminaHumanMethylation450k_dbfile()
IlluminaHumanMethylation450k_dbschema(file="", show.indices=FALSE)
IlluminaHumanMethylation450k_dbInfo()
```

Arguments

file	A connection, or a character string naming the file to print to (see the file argument of the <code>cat</code> function for the details).
show.indices	The CREATE INDEX statements are not shown by default. Use <code>show.indices=TRUE</code> to get them.

Details

IlluminaHumanMethylation450k_dbconn returns a connection object to the package annotation DB. **IMPORTANT:** Don't call `dbDisconnect` on the connection object returned by `IlluminaHumanMethylation450k_dbconn` or you will break all the `AnnDbObj` objects defined in this package!

IlluminaHumanMethylation450k_dbfile returns the path (character string) to the package annotation DB (this is an SQLite file).

IlluminaHumanMethylation450k_dbschema prints the schema definition of the package annotation DB.

IlluminaHumanMethylation450k_dbInfo prints other information about the package annotation DB.

Value

IlluminaHumanMethylation450k_dbconn: a DBIConnection object representing an open connection to the package annotation DB.

IlluminaHumanMethylation450k_dbfile: a character string with the path to the package annotation DB.

IlluminaHumanMethylation450k_dbschema: none (invisible NULL).

IlluminaHumanMethylation450k_dbInfo: none (invisible NULL).

See Also

[dbGetQuery](#), [dbConnect](#), [dbconn](#), [dbfile](#), [dbschema](#), [dbInfo](#)

Examples

```
## Count the number of rows in the "probes" table:
dbGetQuery(IlluminaHumanMethylation450k_dbconn(), "SELECT COUNT(*) FROM probes")

## The connection object returned by IlluminaHumanMethylation450k_dbconn() was
## created with:
dbConnect(SQLite(), dbname=IlluminaHumanMethylation450k_dbfile(), cache_size=64000,
synchronous=0)

IlluminaHumanMethylation450k_dbschema()

IlluminaHumanMethylation450k_dbInfo()
```

IlluminaHumanMethylation450k_get27k

Retrieve 27k probe IDs by design and/or color channel for processing

Description

Retrieve Illumina IDs of probes on the 450k chip which were on the 27k chip

Usage

```
IlluminaHumanMethylation450k_get27k(design=NULL, color=NULL)
```

Arguments

design Optional probe design type. If NULL, both will be returned in a list.
color Optional color for design I. If NULL, both will be returned as a list.

Details

IlluminaHumanMethylation450k_get27k simply queries a VIEW in the SQLite database accompanying this package to obtain probe addresses en masse.

Value

If no arguments are provided, a list containing a data.frame for design II probes along with a list of two data frames for design I probes, one for each color channel. If arguments are provided, the value depends on their status.

See Also

[dbconn](#), [dbfile](#), [dbschema](#), [dbInfo](#)

Examples

```
## Get both channels and both designs worth of 27k probes from the 450k chip
IlluminaHumanMethylation450k_get27k()

## Get only the red channel's worth of design I probes common to the 27k array
IlluminaHumanMethylation450k_get27k('I', 'Red')
```

IlluminaHumanMethylation450k_getControls

Retrieve control probe addresses, optionally by control probe type

Description

A convenience function to retrieve addresses of Illumina control probes.

Usage

```
IlluminaHumanMethylation450k_getControls(type=NULL)
```

Arguments

type Optional control probe type. If NULL, all control probes will be returned.

Details

IlluminaHumanMethylation450k_getControls simply queries a VIEW in the SQLite database accompanying this package to obtain control probe addresses.

Value

A data.frame with fields Address and Type for each probe returned.

See Also

[dbconn](#), [dbfile](#), [dbschema](#), [dbInfo](#)

Examples

```
## Get all the control probe addresses
IlluminaHumanMethylation450k_getControls()

## Get only the bisulfite conversion controls
IlluminaHumanMethylation450k_getControls('Bisulfite')
```

```
IlluminaHumanMethylation450k_getProbeOrdering
Retrieve probe IDs in the order loaded from the manifest
```

Description

A convenience function to provide consistent probe ordering from IDATs.

Usage

```
IlluminaHumanMethylation450k_getProbeOrdering()
```

Value

A vector of Probe_ID strings in the order found in the original manifest.

Examples

```
ordering <- IlluminaHumanMethylation450k_getProbeOrdering()
```

```
IlluminaHumanMethylation450k_getProbes
Retrieve probe addresses by design and/or color channel for processing
```

Description

A convenience function to retrieve probe addresses by design or color channel.

Usage

```
IlluminaHumanMethylation450k_getProbes(design=NULL, color=NULL)
```

Arguments

design	Optional probe design type. If NULL, both will be returned in a list.
color	Optional color for design I. If NULL, both will be returned as a list.

Details

`IlluminaHumanMethylation450k_getProbes` simply queries a VIEW in the SQLite database accompanying this package to obtain probe addresses en masse.

Value

If no arguments are provided, a list containing a data.frame for design II probes along with a list of two data frames for design I probes, one for each color channel. If arguments are provided, the value depends on their status.

See Also

[dbconn](#), [dbfile](#), [dbschema](#), [dbInfo](#)

Examples

```
## Get both channels' worth of probes for processing IDAT files
IlluminaHumanMethylation450k_getProbes()

## Get only the red channel's worth of design I probes
IlluminaHumanMethylation450k_getProbes('I', 'Red')
```

Index

*Topic **datasets**

 IlluminaHumanMethylation450k.db, [2](#)
 IlluminaHumanMethylation450k_dbconn, [35](#)
 IlluminaHumanMethylation450k_get27k, [36](#)
 IlluminaHumanMethylation450k_getControls, [37](#)
 IlluminaHumanMethylation450k_getProbeOrdering, [38](#)
 IlluminaHumanMethylation450k_getProbes, [38](#)
 IlluminaHumanMethylation450kACCNUM, [1](#)
 IlluminaHumanMethylation450kALIAS2PROBE, [2](#)
 IlluminaHumanMethylation450kBLAME, [3](#)
 IlluminaHumanMethylation450kBUILD, [3](#)
 IlluminaHumanMethylation450kCHR, [3](#)
 IlluminaHumanMethylation450kCHR36, [4](#)
 IlluminaHumanMethylation450kCHR37, [4](#)
 IlluminaHumanMethylation450kCHRENGTHS, [5](#)
 IlluminaHumanMethylation450kCHRLOC, [5](#)
 IlluminaHumanMethylation450kCOLORCHANNEL, [6](#)
 IlluminaHumanMethylation450kCPG36, [6](#)
 IlluminaHumanMethylation450kCPG37, [7](#)
 IlluminaHumanMethylation450kCPGCOORDINATE, [7](#)
 IlluminaHumanMethylation450kCPGILOCATION, [8](#)
 IlluminaHumanMethylation450kCPGINAME, [8](#)
 IlluminaHumanMethylation450kCPGIRELATION, [9](#)
 IlluminaHumanMethylation450kCPGS, [9](#)
 IlluminaHumanMethylation450kDESIGN, [10](#)
 IlluminaHumanMethylation450kDHS, [10](#)
 IlluminaHumanMethylation450kDMR, [11](#)
 IlluminaHumanMethylation450kENHANCER, [11](#)
 IlluminaHumanMethylation450kENSEMBL, [12](#)
 IlluminaHumanMethylation450kENTREZID, [13](#)
 IlluminaHumanMethylation450kENZYME, [13](#)
 IlluminaHumanMethylation450kENZYME2PROBE, [15](#)
 IlluminaHumanMethylation450kFANTOM, [16](#)
 IlluminaHumanMethylation450kGENEBODY, [16](#)
 IlluminaHumanMethylation450kGENENAME, [17](#)
 IlluminaHumanMethylation450kGO, [17](#)
 IlluminaHumanMethylation450kGO2ALLPROBES, [18](#)
 IlluminaHumanMethylation450kGO2PROBE, [20](#)
 IlluminaHumanMethylation450kISCPGISLAND, [21](#)
 IlluminaHumanMethylation450kMAP, [21](#)
 IlluminaHumanMethylation450kMAPCOUNTS, [22](#)
 IlluminaHumanMethylation450kMETHYL27, [23](#)
 IlluminaHumanMethylation450kNUID, [23](#)
 IlluminaHumanMethylation450kOMIM, [24](#)
 IlluminaHumanMethylation450kORGANISM, [24](#)

- IlluminaHumanMethylation450kCPGIRELATION, [9](#)
 IlluminaHumanMethylation450kCPGS, [9](#)
 IlluminaHumanMethylation450kDESIGN, [10](#)
 IlluminaHumanMethylation450kDHS, [10](#)
 IlluminaHumanMethylation450kDMR, [11](#)
 IlluminaHumanMethylation450kENHANCER, [11](#)
 IlluminaHumanMethylation450kENSEMBL, [12](#)
 IlluminaHumanMethylation450kENSEMBL2PROBE
 (IlluminaHumanMethylation450kENSEMBL), [12](#)
 IlluminaHumanMethylation450kENTREZID, [13](#)
 IlluminaHumanMethylation450kENZYME, [13](#)
 IlluminaHumanMethylation450kENZYME2PROBE, [15](#)
 IlluminaHumanMethylation450kFANTOM, [16](#)
 IlluminaHumanMethylation450kGENEBODY, [16](#)
 IlluminaHumanMethylation450kGENENAME, [17](#)
 IlluminaHumanMethylation450kGO, [17](#)
 IlluminaHumanMethylation450kGO2ALLPROBES, [18](#)
 IlluminaHumanMethylation450kGO2PROBE, [20](#)
 IlluminaHumanMethylation450kISCPGISLAND, [21](#)
 IlluminaHumanMethylation450kLOCUSID
 (IlluminaHumanMethylation450kENTREZID), [13](#)
 IlluminaHumanMethylation450kMAP, [21](#)
 IlluminaHumanMethylation450kMAPCOUNTS, [22](#)
 IlluminaHumanMethylation450kMETHYL27, [23](#)
 IlluminaHumanMethylation450kNUID, [23](#)
 IlluminaHumanMethylation450kOMIM, [24](#)
 IlluminaHumanMethylation450kORGANISM, [24](#)
 IlluminaHumanMethylation450kORGPKG
 (IlluminaHumanMethylation450kORGANISM), [24](#)
 IlluminaHumanMethylation450kPATH, [25](#)
 IlluminaHumanMethylation450kPATH2PROBE, [26](#)
 IlluminaHumanMethylation450kPFAM, [26](#)
 IlluminaHumanMethylation450kPMID, [27](#)
 IlluminaHumanMethylation450kPMID2PROBE, [28](#)
 IlluminaHumanMethylation450kPROBELOCATION, [29](#)
 IlluminaHumanMethylation450kPROSITE, [29](#)
 IlluminaHumanMethylation450kRANDOM, [30](#)
 IlluminaHumanMethylation450kREFSEQ, [30](#)
 IlluminaHumanMethylation450kREGULATORYGROUP, [31](#)
 IlluminaHumanMethylation450kREGULATORYLOCATION, [32](#)
 IlluminaHumanMethylation450kREGULATORYNAME, [32](#)
 IlluminaHumanMethylation450kSVNID, [33](#)
 IlluminaHumanMethylation450kSYMBOL, [33](#)
 IlluminaHumanMethylation450kUNIGENE, [34](#)
 IlluminaHumanMethylation450kUNIPROT, [35](#)
 IlluminaHumanMethylation450kUNIPROT2PROBE
 (IlluminaHumanMethylation450kUNIPROT), [35](#)
 mappedkeys, [22](#)