KEGG API

This is a brief specification document for the REST-style KEGG API.

General form of the URL

URL form

```
http://rest.kegg.jp/<operation>/<argument>[/<argument2>][/<option>]

<operation> = info | list | find | get | conv | link
<argument> = <database> | <dbentries>
```

Database

```
<database> = KEGG database including KEGG organism (see Table 1)
```

Table 1. KEGG database names and abbreviations

Database	Name	Abbrev	kid	Remark
KEGG PATHWAY	pathway	path	map number	
KEGG BRITE	brite	br	br number	
KEGG MODULE	module	md	M number	
KEGG ORTHOLOGY	orthology	ko	K number	
KEGG GENOME	genome	genome	T number	
KEGG GENOMES	genomes	gn	T number	Composite database: genome + egenome + mgenome
KEGG GENES	genes	-	-	Composite database: consisting of KEGG organisms
KEGG LIGAND	ligand	ligand	-	Composite database: compound + glycan + reaction + rpair + rclass + enzyme
KEGG COMPOUND	compound	cpd	C number	Japanese version: compound_ja cpd_ja
KEGG GLYCAN	glycan	gl	G number	
KEGG REACTION	reaction	rn	R number	
KEGG RPAIR	rpair	rp	RP number	
KEGG RCLASS	rclass	rc	RC number	
KEGG ENZYME	enzyme	ec	-	
KEGG DISEASE	disease	ds	H number	Japanese version: disease_ja ds_ja
KEGG DRUG	drug	dr	D number	Japanese version: drug_ja dr_ja
KEGG DGOUP	dgroup	dg	DG number	Japanese version: dgroup_ja dg_ja
KEGG ENVIRON	environ	ev	E number	Japanese version: environ_ja ev_ja

Auxiliary databases, dgenes, egenes, mgenes, egenome and mgenome, may also be used.

Database entry

```
<dbentries> = <dbentry>1[+<dbentry>2...]
<dbentry> = <db:entry> | <kid> | <org:gene>
```

Each database entry is identified by:

db:entry

where

"db" is the database name or its abbreviation shown above and
"entry" is the entry name or the accession number that is uniquely assigned within the database.

In reality "db" may be omitted, for the entry name called the KEGG object identifier (kid) is unique across KEGG.

kid = database-dependent prefix + five-digit number

In the KEGG GENES database the db:entry combination must be specified. This is more specifically written as:

org:gene

where

"org" is the three- or four-letter KEGG organism code or the T number genome identifier and
"gene" is the gene identifier, usually locus_tag or ncbi GeneID, or the primary gene name.

Output

The output of all operations is in a text format:

- tab-delimited text returned from list, find, conv and link
- flat file database format returned from get
- text message returned from info

Status code

The HTTP status code can be used to check if the operation was successful.

Code	Meaning
200	Success
400	Bad request (syntax error, wrong database name, etc.)
404	Not found (e.g., requesting amino acid sequence for RNA)

Operations

Database information

Operation

info – displays the current statistics of a given database

URL form

```
http://rest.kegg.jp/info/<database>

<database> = pathway | brite | module | ko | genome | <org> | compound | glycan |
             reaction | rpair | rclass | enzyme | disease | drug | dgroup | environ |
             genomes | genes | ligand | kegg
<org> = KEGG organism code or T number
```

Examples

/info/kegg	displays the current statistics of the KEGG database
/info/pathway	displays the number pathway entries including both the reference and organism-specific pathways
/info/hsa	displays the number of gene entries for the KEGG organism *Homo sapiens*

Entry list

Operation

list – returns a list of entry identifiers and associated definition for a given database or a given set of database entries

URL form

```
http://rest.kegg.jp/list/<database>
```

```
<database> = pathway | brite | module | ko | genome | <org> | compound | glycan |
             reaction | rpair | rclass | enzyme | disease | drug | dgroup | environ |
             organism
<org> = KEGG organism code or T number
```

```
http://rest.kegg.jp/list/<database>/<org>

<database> = pathway | module
<org> = KEGG organism code
```

```
http://rest.kegg.jp/list/<dbentries>

<dbentries> = KEGG database entries involving the following <database>
<database> = pathway | brite | module | ko | genome | <org> | compound | glycan |
             reaction | rpair | rclass | enzyme | disease | drug | dgroup | environ |
<org> = KEGG organism code or T number
```

Notes

The special keyword "organism" for the database name is applicable only to this operation.
In the second form an organism-specific version of pathway/module list is returned.
In the third form the maximun number of database entries that can be given is 100.

Examples

/list/pathway	returns the list of reference pathways
/list/pathway/hsa	returns the list of human pathways
/list/organism	returns the list of KEGG organisms with taxonomic classification
/list/hsa	returns the entire list of human genes
/list/T01001	same as above
/list/hsa:10458+ece:Z5100	returns the list of a human gene and an E.coli O157 gene
/list/cpd:C01290+gl:G00092	returns the list of a compound entry and a glycan entry
/list/C01290+G00092	same as above

Data search

Operation

find – finds entries with matching query keywords or other query data in a given database

URL form

```
http://rest.kegg.jp/find/<database>/<query>

<database> = pathway | module | ko | genome | <org> | compound | glycan |
             reaction | rpair | rclass | enzyme | disease | drug | dgroup | environ |
             genes | ligand
<org> = KEGG organism code or T number
```

```
http://rest.kegg.jp/find/<database>/<query>/<option>

<database> = compound | drug
<option> = formula | exact_mass | mol_weight
```

Notes

The first form searches entry identifier and associated fields shown below for matching keywords.

Database	Search fields (see flat file format)
pathway	ENTRY and NAME
module	ENTRY and NAME
disease	ENTRY and NAME
drug	ENTRY and NAME
environ	ENTRY and NAME
ko	ENTRY, NAME and DEFINITION
genome	ENTRY, NAME and DEFINITION
<org>	ENTRY, NAME, DEFINITION and ORTHOLOGY

compound	ENTRY and NAME
glycan	ENTRY, NAME, COMPOSITION and CLASS
reaction	ENTRY, NAME and DEFINITION
rpair	ENTRY and NAME
rclass	ENTRY, NAME and DEFINITION
enzyme	ENTRY and NAME

Keyword search against brite is not supported. Use /list/brite to retrieve a short list.

In the second form the chemical formula search is a partial match irrespective of the order of atoms given.
The exact mass (or molecular weight) is checked by rounding off to the same decimal place as the query data.
A range of values may also be specified with the minus(-) sign.

Examples

/find/genes/shiga+toxin	for keywords "shiga" and "toxin"
/find/genes/"shiga toxin"	for keywords "shiga toxin"
/find/compound/C7H10O5/formula	for chemical formula "C7H10O5"
/find/compound/O5C7/formula	for chemical formula containing "O5" and "C7"
/find/compound/174.05/exact_mass	for 174.045 =< exact mass < 174.055
/find/compound/300-310/mol_weight	for 300 =< molecular weight =< 310

Data retrieval

Operation

get – retrieves given database entries

URL form

```
http://rest.kegg.jp/get/<dbentries>[/<option>]

<dbentries> = KEGG database entries involving the following <database>
<database> = pathway | brite | module | ko | genome | <org> | compound | glycan |
             reaction | rpair | rclass | enzyme | disease | drug | dgroup | environ
<org> = KEGG organism code or T number

<option> = aaseq | ntseq | mol | kcf | image | kgml
```

Notes

The input is limited up to 10 entries.
The input is limited to one pathway entry with the image or kgml option.
The input is limited to one compound/glycan/drug entry with the image option.

Examples

/get/cpd:C01290+gl:G00092	retrieves a compound entry and a glycan entry
/get/C01290+G00092	same as above
/get/hsa:10458+ece:Z5100	retrieves a human gene entry and an E.coli O157 gene entry
/get/hsa:10458+ece:Z5100 /aaseq	retrieves amino acid sequences of a human gene and an E.coli O157 gene
/get/hsa05130/image	retrieves the png image file of a pathway map
/get/C00002/image	retrieves the gif image file of a compound
/get/hsa05130/kgml	retrieves the kgml file of a pathway map

See also

KEGG web links to retrieve entries in HTML.

ID conversion

Operation

conv – convert KEGG identifiers to/from outside identifiers

URL form

```
http://rest.kegg.jp/conv/<target_db>/<source_db>

(<target_db> <source_db>) = (<kegg_db> <outside_db>) | (<outside_db> <kegg_db>)

For gene identifiers:
<kegg_db> = <org>
<org> = KEGG organism code or T number
<outside_db> = ncbi-proteinid | ncbi-geneid | uniprot

For chemical substance identifiers:
<kegg_db> = drug | compound | glycan
<outside_db> = pubchem | chebi
```

```
http://rest.kegg.jp/conv/<target_db>/<dbentries>

For gene identifiers:
<dbentries> = database entries involving the following <database>
<database> = <org> | genes | ncbi-proteinid | ncbi-geneid | uniprot
<org> = KEGG organism code or T number

For chemical substance identifiers:
<dbentries> = database entries involving the following <database>
<database> = drug | compound | glycan | pubchem | chebi
```

Note

The database name "genes" may be used only in the second form.

Examples

/conv/eco/ncbi-geneid	conversion from NCBI GeneID to KEGG ID for E. coli genes
/conv/ncbi-geneid/eco	opposite direction
/conv/ncbi-proteinid /hsa:10458+ece:Z5100	conversion from KEGG ID to NCBI GI
/conv/genes/ncbi-geneid:3113320	conversion from NCBI GeneID to KEGG ID when the organism code is not known

Linked entries

Operation

link – find related entries by using database cross-references

URL form

```
http://rest.kegg.jp/link/<target_db>/<source_db>

<target_db> = <database>
<source_db> = <database>

<database> = pathway | brite | module | ko | genome | <org> | compound | glycan |
             reaction | rpair | rclass | enzyme | disease | drug | dgroup | environ
```

```
http://rest.kegg.jp/link/<target_db>/<dbentries>

<dbentries> = KEGG database entries involving the following <database>
<database> = pathway | brite | module | ko | genome | <org> | compound | glycan |
             reaction | rpair | rclass | enzyme | disease | drug | dgroup | environ |
             genes
```

Note

The database name "genes" may be used only in the second form.

Examples

/link/pathway/hsa	KEGG pathways linked from each of the human genes
/link/hsa/pathway	human genes linked from each of the KEGG pathways

/link/pathway/hsa:10458+ece:Z5100	KEGG pathways linked from a human gene and an E. coli O157 gene
/link/genes/K00500	List of genes with the KO assignment of K00500
/link/genes/hsa00010 or /link/hsa/hsa00010	List of human genes in pathway hsa00010
/link/ko/map00010 or /link/ko/ko00010	List of KO entries in pathway map00010 or ko00010
/link/rn/map00010 or /link/rn/rn00010	List of reaction entries in pathway map00010 or rn00010
/link/ec/map00010 or /link/ec/ec00010	List of EC number entries in pathway map00010 or ec00010
/link/cpd/map00010	List of compound entries in pathway map00010

Last updated: January 15, 2016