

Funciones de PyBioGateway

Ontología	Uri	Prefijo
Simple Knowledge Organization System (SKOS)	http://www.w3.org/2004/02/skos/core#	skos
Resource Description Framework (RDF)	http://www.w3.org/1999/02/22-rdf-syntax-ns#	rdf
RDF Schema (RDFS)	http://www.w3.org/2000/01/rdf-schema#	rdfs
Semanticscience Integrated Ontology (SIO)	http://semanticscience.org/resource/	sio
Open Biomedical Ontologies (OBO)	http://purl.obolibrary.org/obo/	obo
Dublin core (DC)	http://purl.org/dc/terms/	dc
Schema (SCH)	http://schema.org/	sch
Open Biomedical Ontologies (OBO)	http://www.geneontology.org/formats/oboInOwl#	oboowl

Tabla 1. Tabla con la información sobre la ontología y los prefijos usados en las consultas.

Función `type_data(name)`:

Descripción: Esta función sirve para consultar el tipo de entidad biológica a la que pertenece una entidad específica introducida por el usuario.

Esta función tiene como parámetro el nombre (`prefLabel` o `altLabel`) de la entidad biológica específica que queremos consultar en el grafo de conocimiento, y devuelve el tipo de entidad biológica a la que pertenece, es decir si es un gen, proteína, dominio asociado topológicamente (TAD), módulo cis regulador (CRM), proceso biológico, componente celular o función molecular.

Parámetros:

-“**name**”: Nombre de la entidad biológica que el usuario quiere consultar. Siete tipos de entidades pueden consultarse: proteínas, genes, TADs, CRMs, procesos biológicos, funciones moleculares o componentes celulares. Los vocabularios empleados son los siguientes:

- Proteínas: Son válidos tanto el “entry name” de Uniprot como otros identificadores alternativos. Ejemplo: “INSR_HUMAN”.
- Genes: gene symbol HGNC. Ejemplo: “INSR”
- TADs y CRMs: identificadores propios de BioGateway. Ejemplo: “crm/CRMHS00000003515”, “tad/TADHS000000020654”.
- Procesos biológicos, funciones moleculares o componentes celulares: Identificadores propios de Gene Ontology (GO). Ejemplo: “GO:0000206”

Salida:

-“**bioentity_type**”: Tipo de entidad biológica a la que pertenece el valor de entrada. Valores posibles:

- Proteína: Ejemplo salida, “protein” (prefLabel).
- Gen: Ejemplo salida, “gene” (prefLabel).
- TAD: Ejemplo salida, “topologically_associated_domain (tad)” (prefLabel).
- CRM: Ejemplo salida, “cis_regulatory_module (crm)” (prefLabel).
- Proceso biológico: Ejemplo salida, “biological_process” (hasOBONamespace).
- Componente celular: Ejemplo salida, “cellular_component” (hasOBONamespace).
- Función molecular: Ejemplo salida, “molecular_function” (hasOBONamespace).

Ejemplo de ejecución:

Entrada:

type_data(“INSR_HUMAN”)

Salida:

“protein”

Entrada con nombre alternativo de la misma proteína:

type_data(“P06213”)

Salida:

“protein”

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología
“name”	entrada	string	Nombre de la entidad biológica	skos:prefLabel o skos:altLabel
“bioentity_type”	salida	string	Tipo de entidad biológica	rdfs:subClassof obo:hasOBONamespace

Función: **getGene_info(gene, taxon)**

Descripción: Función que proporciona la información asociada al gen que se desea consultar, disponible en el grafo de conocimiento: <http://rdf.biogateway.eu/graph/gene>. Esta función tiene dos parámetros: el nombre del gen que queremos consultar y el taxón.

Parámetros:

-**"gene"**. Insertamos el gen en formato gene symbol (propiedad: prefLabel). Ejemplo: "BRCA1"

-**"taxon"**. Insertamos o bien el número del taxon o bien el nombre del organismo (propiedad: label). Ejemplo: "Homo sapiens", "9606".

Salida:

Devuelve un diccionario con los siguientes campos:

-**"chromosome"**: Indica el cromosoma al que pertenece el gen (propiedad: BFO_0000050 (part of)).

-**"start"**: Indica la posición de inicio del gen en la secuencia genómica (propiedad: GENO_0000894 (start_position)).

-**"end"**: Indica la posición de finalización del gen en la secuencia genómica (propiedad: GENO_0000895 (end_position)).

-**"strand"**: Indica la orientación del gen en la secuencia genómica (propiedad: GENO_0000906 (on strand)).

-**"assembly"**: Indica el ensamblaje del genoma humano correspondiente a nuestro gen (propiedad: hasVersion).

-**"alt_gene_sources"**: Corresponde a otras fuentes o bases de datos que contienen información relacionada con el gen de interés (propiedad: closeMatch).

-**"definition"**: Proporciona una definición del gen (propiedad: definition).

Ejemplo de ejecución:

Entrada: **getGene_info("Brca1", "Mus musculus")**

Salida:

```
{'start': '101379587',
'end': '101442808',
'strand': 'ReverseStrandPosition',
'chr': 'NC_000077.7',
'assembly': 'GCF_000001635.27',
'alt_gene_sources': 'ensembl/ENSMUSG00000017146;
ensembl/ENSMUSG00000017146.13; ncbigene/12189',
'definition': 'gene 10090/Brca1 encoding [A0A087WP26_MOUSE A0A087WPE1_MOUSE
A0A087WPK5_MOUSE BRCA1_MOUSE]}'
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“gene”	entrada	string	Nombre del gen en formato symbol	skos:prefLabel	gene
“taxon”	entrada	string	Número de taxón o nombre del organismo	rdfs:label	taxon
“chromosome”	salida	string	Cromosoma al que pertenece el gen	obo:BFO_0000050 (part of)	gene
“start”	salida	integer	Posición de inicio del gen	obo:GENO_0000894 (start position)	gene
“end”	salida	integer	Posición de fin del gen	obo:GENO_0000895 (end position)	gene
“strand”	salida	string	Orientación del gen	obo:GENO_0000906 (on strand)	gene
“assembly”	salida	string	Ensamblaje del genoma humano al que correspond	dc:hasVersion	gene

			e el gen		
"alt_gene_sources"	salida	string	Fuentes alternativas con información sobre el gen	skos:closeMatch	gene
"definition"	salida	string	Descripción del gen	skos:definition	gene

Función `getGenes_by_coord(chr, start, end , strand)`:

Descripción: Función que devuelve los genes en formato gene symbol(prefLabel), que se encuentran dentro de las coordenadas genómicas especificadas. Usamos la información disponible en el grafo <http://rdf.biogateway.eu/graph/gene>.

Parámetros:

-"**chr**": Indica el cromosoma al que pertenece el segmento (identificador del cromosoma en el NCBI). Ejemplo: "NC_000074.7".

-"**start**": Indica la posición de inicio del segmento en la secuencia genómica. Ejemplo: "90973665"

-"**end**": Indica la posición de finalización del segmento en la secuencia genómica. Ejemplo: "91075654"

-"**strand**": Indica la cadena de ADN en la que se quiere realizar la búsqueda. Si se declara como none, buscará en ambas cadenas. Valores permitidos: "ReverseStrandPosition" y "ForwardStrandPosition".

Salida:

Lista de genes que se encuentran dentro de las coordenadas genómicas especificadas. A su vez, para cada gen se indica:

-"**gene_name**": Nombre del gen en formato gene symbol (propiedad: prefLabel).

-**"start"**: Indica la posición de inicio del gen en la secuencia genómica (propiedad: GENO_0000894).

-**"end"**: Indica la posición de finalización del gen en la secuencia genómica (propiedad: GENO_0000895).

-**"strand"**: Indica la orientación del gen en la secuencia genómica (propiedad: GENO_0000906). Este campo sólo se proporciona si no se ha especificado cadena de ADN en la función.

Ejemplo ejecución:

Entrada: `getGenes_by_coord("NC_051352.1", 52565276, 58596412, "ForwardStrandPosition")`

Salida:

```
[{'gene_name': 'Fzd8', 'start': '57312924', 'end': '57320551'},
 {'gene_name': 'Hist2h3c2', 'start': '183797721', 'end': '41378877'},
 {'gene_name': 'Hist2h3c2', 'start': '183837311', 'end': '41532577'},
 {'gene_name': 'Hist2h3c2', 'start': '183797721', 'end': '41532577'},
 {'gene_name': 'Hist2h3c2', 'start': '183837311', 'end': '41426735'},
 {'gene_name': 'Hist2h3c2', 'start': '183837311', 'end': '41378877'},
 {'gene_name': 'Hist2h3c2', 'start': '183797721', 'end': '41426735'},
 {'gene_name': 'Hnrnpk', 'start': '91756628', 'end': '6275001'},
 {'gene_name': 'Lgals8', 'start': '58024652', 'end': '58052764'},
 {'gene_name': 'Map3k8', 'start': '53382908', 'end': '53403216'},
 {'gene_name': 'Mtr', 'start': '58219998', 'end': '58308560'},
 {'gene_name': 'Actn2', 'start': '58143334', 'end': '58210622'},
 {'gene_name': 'Crem', 'start': '54238889', 'end': '54305989'}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"chr"	entrada	string	Cromosoma al que pertenece el segmento genómico	obo:BFO_0000050 (part of)	gene
"start"	entrada	integer	Posición de inicio del segmento genómico	obo:GENO_0000894 (start position)	gene

“end”	entrada	integer	Posición de fin del segmento genómico	obo:GENO_0000895 (end position)	gene
“strand”	entrada (opcional)	string	Cadena de DNA donde se quiere realizar la búsqueda	obo:GENO_0000906 (on strand)	gene
“gene_name”	salida	string	Nombre del gen en formato symbol	skos:prefLabel	gene
“start”	salida	integer	Posición de inicio del gen	obo:GENO_0000894 (start position)	gene
“end”	salida	integer	Posición de fin del gen	obo:GENO_0000895 (end position)	gene
“strand”	salida (opcional)	string	Orientación del gen	obo:GENO_0000906 (on strand)	gene

Función getProtein_info(protein)

Descripción: Función que facilita la información asociada a la proteína de interés, disponible en el grafo de conocimiento: <http://rdf.biogateway.eu/graph/protein>.

Esta función tiene un parámetro, el nombre de la proteína que queremos consultar en formato entry name de Uniprot.

Parámetros:

-“**protein**”: Insertamos el nombre de la proteína en formato entry name de Uniprot (propiedad: prefLabel). Ejemplo: “BRCA1_HUMAN”.

Salida:

Devuelve un diccionario con los siguientes campos:

-“**protein_alt_names**”: Devuelve los nombres alternativos de la proteína (propiedad: altLabel), en concreto el nombre de la proteína en formato entry de Uniprot, así como el nombre del gen que la codifica y sinónimos de este gen.

-**"definition"**: Proporciona una definición de la proteína (propiedad: definition).

-**"evidence_level"**: Este término se refiere al nivel de evidencia que respalda la información asociada con la proteína (propiedad: evidenceLevel).

-**"alt_sources"**: Hace referencia a otras fuentes o bases de datos que contienen información relacionada con la proteína (propiedad: closeMatch).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la proteína de interés que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo ejecución:

Entrada: getProtein_info("TOX3_HUMAN")

Salida:

```
{'protein_alt_ids': 'O15405; TOX3; TNRC9; CAGF9', 'definition': 'TOX high mobility group box family member 3 (CAG trinucleotide repeat-containing gene F9 protein) (Trinucleotide repeat-containing gene 9 protein)', 'evidence_level': '5.0', 'alt_sources': 'ensembl/ENSP00000219746.9; ensembl/ENSP00000385705.3; refseq/NP_001073899.2; refseq/NP_001139660.1', 'articles': 'pubmed/15616553; pubmed/9225980; pubmed/21172805; pubmed/14702039'}
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"protein"	entrada	string	Nombre de la proteína en formato entry name de Uniprot	skos:prefLabel	protein
"protein_alt_names"	salida	string	Nombres alternativos de la proteína	skos:altLabel	protein
"definition"	salida	string	Definición de la proteína	skos:definition	protein

"evidence_level"	salida	string	Nivel de evidencia asociado a la proteína	sch:evidenceLevel	protein
"alt_sources"	salida	string	Fuentes alternativas con información sobre la proteína	skos:closeMatch1	protein
"articles"	salida	string	Artículos relacionados con la proteína	sio:SIO_000772 (has evidence)	protein

Función getPhenotype(phenotype):

Descripción: Esta función permitirá obtener el fenotipo a partir de un identificador OMIM o del nombre de una enfermedad, usa el grafo <http://rdf.biogateway.eu/graph/omim>:

-Por un lado si el valor introducido es un identificador OMIM de un fenotipo, nos devolverá su etiqueta preferente.

-Por otro lado si el valor introducido es el nombre de una enfermedad (por ejemplo "breast cancer"), nos devolverá la etiqueta preferente y el identificador omim de aquellos fenotipos cuya etiqueta, tanto preferente como alternativa, contenga la enfermedad introducida. Es decir, los fenotipos que contengan el nombre introducido.

Parámetros:

-**"phenotype"**: Los valores permitidos de este parámetro serán:

-Identificador OMIM de un fenotipo. Ejemplo "MTHU036782".

-Nombre de una enfermedad. Ejemplo "lung cancer".

Salida:

La salida dependerá del tipo de parámetro introducido:

-Si ha sido introducido un identificador OMIM de un fenotipo, la salida será simplemente la etiqueta preferente de este identificador (prefLabel).

-Si se ha introducido el nombre de una enfermedad, la salida estará compuesta por aquellos fenotipos que contienen en su etiqueta preferente esta enfermedad (prefLabel o altLabel), así como su identificador OMIM.

Ejemplo ejecución:

Entrada: getPhenotype("breast cancer")

Salida: [{ 'omim_id': '604704', 'label': 'BREAST CANCER ANTIESTROGEN RESISTANCE 3'},
{ 'omim_id': 'MTHU036782', 'label': 'Breast cancer, lobular'},
{ 'omim_id': 'MTHU068657', 'label': 'Breast cancer, early-onset'},
{ 'omim_id': '137215',
 'label': 'DIFFUSE GASTRIC AND LOBULAR BREAST CANCER SYNDROME'},
{ 'omim_id': 'MTHU015471',
 'label': 'Paraneoplastic SPS is associated with breast cancer and other malignancies'},

Entrada: getPhenotype("604704")

Salida: [{ 'phen_label': 'BREAST CANCER ANTIESTROGEN RESISTANCE 3'}]

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"phenotype"	entrada	string	Nombre de una enfermedad o identificador OMIM de un fenotipo	skos:prefLabel	omim
"omim_id"	salida	string	Identificador OMIM del fenotipo (opcional)	skos:prefLabel	omim
"phen_label"	salida	string	Nombre común del fenotipo	skos:prefLabel	omim

Función getCRM_info(crm):

Descripción: Función que devuelve la información asociada al módulo cis-regulador (crm) que queremos consultar, disponible en el grafo de conocimiento: <http://rdf.biogateway.eu/graph/crm>.

Esta función tiene un parámetro el cual es el nombre del crm que queremos consultar (propiedad: prefLabel).

Parámetros:

-**"crm"**: El parámetro introducido en la función será el nombre preferente del módulo cis-regulador (propiedad: prefLabel). Ejemplo "crm/CRMHS00003225754".

Salida:

Devuelve un diccionario con los siguientes campos:

-**"start"**: Indica la posición de inicio del crm en la secuencia genómica (propiedad: GENO_0000894 (start_position)).

-**"end"**: Indica la posición de finalización del crm en la secuencia genómica (propiedad: GENO_0000895 (end_position)).

-**"chromosome"**: Indica el cromosoma al que pertenece el crm, (identificador del cromosoma en el NCBI) (propiedad: BFO_0000050 (part_of)).

-**"assembly"**: Indica el ensamblaje del genoma humano correspondiente a nuestro crm (propiedad: hasVersion).

-**"taxon"**: Devuelve el taxón al que pertenece este crm (identificador taxonómico del NCBI) (propiedad: RO_0002162).

-**"definition"**: Definición del crm disponible en la propiedad definition del grafo de conocimiento.

Ejemplo ejecución:

Entrada: getCRM_info("crm/CRMHS00000005387")

Salida:

```
[{'start': '355447',  
  'end': '358949',  
  'chromosome': 'NC_000011.10',  
  'assembly': 'GCF_000001405.26',
```

```
'taxon': 'NCBITaxon_9606',
  'definition': 'Cis-regulatory module located in Homo sapiens chr11 between 355447 and 358949']}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“crm”	entrada	string	Nombre del módulo cis-regulador (crm)	skos:prefLabel	crm
“taxon”	salida	string	Número de taxón al que pertenece el crm	obo:RO_0002162 (in taxon)	crm
“chromosome”	salida	string	Cromosoma al que pertenece el crm	obo:BFO_0000050 (part of)	crm
“start”	salida	integer	Posición de inicio del crm	obo:GENO_0000894 (start position)	crm
“end”	salida	integer	Posición de fin del crm	obo:GENO_0000895 (end position)	crm
“strand”	salida	string	Orientación del crm	obo:GENO_0000906 (on strand)	crm
“assembly”	salida	string	Ensamblaje del genoma humano al que corresponde el crm	dc:has Version	crm
“definition”	salida	string	Descripción del crm	skos:definition	crm

Función getCRM_additional_info(crm):

Descripción: Esta función proporcionará información adicional sobre el crm de interés, disponible en el grafo <http://rdf.biogateway.eu/graph/crm> que no se obtiene de la función getCRM_info.

Parámetros:

-**"crm"**: El parámetro introducido en la función será el nombre preferente del módulo cis-regulador (propiedad: prefLabel). Ejemplo "crm/CRMHS00000005387".

Salida:

Devuelve un diccionario con los siguientes campos:

-**"evidence"**: Corresponde a la evidencia que respalda la información disponible sobre el módulo-cis regulador (propiedad: evidenceOrigin).

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre el crm de interés (propiedad: SIO_000253 (has source)).

-**"biological_samples"**: Se refiere a los diferentes tipos de muestras biológicas que están asociados con el estudio del módulo cis-regulador (propiedad: TXPO_0003500 (observed in)). En concreto, devolverá los identificadores en formato de términos ontológicos. Ejemplo: "CLO_0001601", "UBERON_0002113", "BTO_0000018".

-**"articles"**: Hace referencia a los artículos científicos que han sido publicados en Pubmed relacionados con el módulo cis-regulador introducido (propiedad: SIO_000772 (has evidence)).

Ejemplo ejecución:

Entrada: getCRM_add_info("crm/CRMHS00000005387")

Salida:

{'evidence':

'http://www.licpathway.net/ENdb/search/Detail.php?Species=Human&Enhancer_id=E_01_273',

'database': '<http://www.licpathway.net/ENdb/>',

'biological_samples': 'BTO_0000007; BTO_0000018; CLO_0001230; CLO_0001601; CL_0002518; CL_0000082; UBERON_0002048; UBERON_0002113',

'articles': 'pubmed/28511927'}

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"crm"	entrada	string	Nombre del módulo cis-regulador (crm)	skos:prefLabel	crm

"evidence"	salida	string	Nivel de evidencia asociado al crm	sch:evidenceOrigin	crm
"database"	salida	string	Base de datos donde encontramos la información del crm	sio:SIO_000253 (has source)	crm
"biological_samples"	salida	string	Muestras biológicas asociadas al crm de interés	obo:TXPO_0003500 (observed in)	crm
"articles"	salida	string	Artículos relacionados con el crm	sio:SIO_000772 (has evidence)	crm

Función `getCRMs_by_coord(chromosome, start, end)`:

Descripción: Función que devuelve el identificador de los crms (prefLabel) que se encuentran dentro de las coordenadas genómicas especificadas.

Parámetros:

-**"chromosome"**: Indica el cromosoma en el que queremos realizar la búsqueda, los valores permitidos **son chr-número de cromosoma**. Ejemplo: "chr-1". "chr-18", "mitochondrial" (para buscar en el ADN mitocondrial).

-**"start"**: Indica la posición de inicio del segmento en la secuencia genómica. Ejemplo: "90973665"

-**"end"**: Indica la posición de finalización del segmento en la secuencia genómica. Ejemplo: "91075654"

Salida:

Lista de módulos-cis reguladores que se encuentran dentro de las coordenadas genómicas especificadas. A su vez, para cada crm se indica:

-**"crm_name"**: Nombre del módulo cis regulador que se encuentra dentro de las coordenadas genómicas especificadas (propiedad: prefLabel).

-**"start"**: Indica la posición de inicio del crm en la secuencia genómica (propiedad: GENO_0000894).

-**"end"**: Indica la posición de finalización del crm en la secuencia genómica (propiedad: GENO_0000895).

Ejemplo ejecución:

Entrada: getCRMs_by_coord("mitochondrial","1", "500")

Salida:

```
{'crm_name': 'crm/CRMHS00032244267', 'start': '175', 'end': '426'},  
{'crm_name': 'crm/CRMHS00032244371', 'start': '230', 'end': '389'},  
{'crm_name': 'crm/CRMHS00032244378', 'start': '32', 'end': '334'}}
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Graf o
"chromosome"	entrada	string	Cromosoma al que pertenece el segmento genómico	obo:BFO_0000050 (part of)	crm
"start"	entrada	integer	Posición de inicio del segmento genómico	obo:GENO_0000894 (start position)	crm
"end"	entrada	integer	Posición de fin del segmento genómico	obo:GENO_0000895 (end position)	crm
"crm_name"	salida	string	Nombre del crm que se encuentra	skos:prefLabel	crm

			dentro de las coordenadas genómicas		
“start”	salida	integer	Posición de inicio del crm	obo:GENO_0000894 (start position)	crm
“end”	salida	integer	Posición de fin del crm	obo:GENO_0000895 (end position)	crm

Función getTAD_info(tad):

Descripción: Función que proporciona la información asociada al dominio topológicamente asociado (tad) que queremos consultar, disponible en el grafo de conocimiento: <http://rdf.biogateway.eu/graph/tad>.

Esta función tiene un parámetro el cual es el identificador del tad que queremos consultar (propiedad: prefLabel).

Parámetros:

-”**tad**”: El parámetro introducido en la función será el nombre preferente del dominio topológicamente asociado (propiedad: prefLabel). Ejemplo "tad/TADHS00000038004"

Salida:

Devuelve un diccionario con los siguientes campos:

-”**chromosome**”: Indica el cromosoma al que pertenece el tad, (identificador del cromosoma en el NCBI) (propiedad: BFO_0000050).

-”**start**”: Indica la posición de inicio del tad en la secuencia genómica (propiedad: GENO_0000894).

-”**end**”: Indica la posición de finalización del tad en la secuencia genómica (propiedad: GENO_0000895).

-”**assembly**”: Indica el ensamblaje del genoma humano correspondiente a nuestro crm (propiedad: hasVersion).

-”**taxon**”: Devuelve el taxón al que pertenece este tad (identificador taxonómico del NCBI) (propiedad: RO_0002162).

-**"definition"**: Definición del tad disponible en la propiedad definition del grafo de conocimiento.

Ejemplo ejecución:

Entrada: getTAD_info("tad/TADHS00000038004")

Salida:

```
[{'start': '34120000',  
  'end': '35840000',  
  'chromosome': 'NC_000013.11',  
  'assembly': 'GCF_000001405.26',  
  'taxon': 'NCBITaxon_9606',  
  'definition': 'Topologically associated domain located in Homo sapiens chr13 between  
34120000 and 35840000'}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"tad"	entrada	string	Nombre del dominio topológicamente asociado	objeto de la relación prefLabel	tad
"chromosome"	salida	string	Cromosoma al que pertenece el tad	obo:BFO_0000050 (part of)	tad
"start"	salida	integer	Posición de inicio del tad	obo:GENO_0000894 (start position)	tad
"end"	salida	integer	Posición de fin del tad	obo:GENO_0000895 (end position)	tad
"assembly"	salida	string	Ensamblaje del genoma humano al que corresponde el tad	dc:hasVersion	tad
"taxon"	salida	string	Taxón al que pertenece este tad (identificador del NCBI)	obo:RO_0002162 (in taxon)	tad
"definition"	salida	string	Descripción del tad	skos:definition	tad

Función getTAD_additional_info(tad):

Descripción: Esta función proporcionará información adicional sobre el tad de interés, disponible en el grafo <http://rdf.biogateway.eu/graph/tad> que no se obtiene de la función getTAD_info.

Parámetros:

-**"tad"**: El parámetro introducido en la función será el nombre preferente del dominio topológicamente asociado (propiedad: prefLabel). Ejemplo "tad/TADHS00000038004".

Salida:

Devuelve un diccionario con los siguientes campos:

-**"evidence"**: Corresponde a la evidencia que respalda la información disponible sobre el tad (propiedad: evidenceOrigin).

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre el tad de interés (propiedad: SIO_000253 (has source)).

-**"biological_samples"**: Se refiere a los diferentes tipos de muestras biológicas que están asociados con el estudio del dominio topológicamente asociado (propiedad: TXPO_0003500 (observed in)). En concreto, devolverá los identificadores en formato de términos ontológicos. Ejemplo: "CLO_0001601", "UBERON_0002113", "BTO_0000018".

-**"articles"**: Hace referencia a los artículos científicos que han sido publicados en Pubmed relacionados con el dominio topológicamente asociado introducido (propiedad: SIO_000772 (has evidence)).

Ejemplo ejecución:

Entrada: getTAD_add_info("tad/TADHS00000038004")

Salida:

{'evidence':

'http://dna.cs.miami.edu/TADKB/domain.php?sp=hum&cl=HMEC&rg=hg19&chr=14&se=96700000_96850000&id=83&res=50kb&caller=IS',

'database': '<http://dna.cs.miami.edu/TADKB/>',

'biological_samples': 'BTO_0001229; CLO_0006951; CL_0002327; UBERON_0000310; UBERON_0002048; BTO_0004300; CL_0002553',

'articles': 'pubmed/30871473'}

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“tad”	entrada	string	Nombre del dominio topológicamente asociado (tad)	skos:prefLabel	tad
”evidence”	salida	string	Nivel de evidencia asociado al tad	sch:evidenceOrigin	tad
“database”	salida	string	Base de datos donde encontramos la información del tad	sio:SIO_000253 (has source)	tad
“biological_samples”	salida	string	Muestras biológicas asociadas al tad de interés	obo:TXPO_0003500 (observed in)	tad
“articles”	salida	string	Artículos relacionados con el tad	sio:SIO_000772 (has evidence)	tad

Función `getTAD_by_coord(chromosome,start,end)`:

Descripción: Función que devuelve el identificador de los tads (prefLabel) que se encuentran dentro de las coordenadas genómicas especificadas.

Parámetros:

-”**chromosome**”: Indica el cromosoma al que pertenece el segmento (identificador del cromosoma en el NCBI). Ejemplo: “chr-13”.

-”**start**”: Indica la posición de inicio del segmento en la secuencia genómica. Ejemplo: “90973665”

-”**end**”: Indica la posición de finalización del segmento en la secuencia genómica. Ejemplo: “91075654”

Salida:

-**"tad_name"**: Devuelve una lista de los ids de los dominios topológicamente asociados que se encuentran dentro de las coordenadas genómicas especificadas (propiedad: prefLabel).

-**"start"**: Indica la posición de inicio del tad en la secuencia genómica (propiedad: GENO_0000894).

-**"end"**: Indica la posición de finalización del tad en la secuencia genómica (propiedad: GENO_0000895).

Ejemplo ejecución:

Entrada: getTADs_by_coord("chr-13", "34120000", "35840000")

Salida:

```
[{'tad_id': 'tad/TADHS000000038004', 'start': '34120000', 'end': '35840000'},
{'tad_id': 'tad/TADHS000000029314', 'start': '35200000', 'end': '35840000'},
{'tad_id': 'tad/TADHS000000071459', 'start': '34125863', 'end': '35175863'},
{'tad_id': 'tad/TADHS000000071460', 'start': '34165863', 'end': '35825863'},
{'tad_id': 'tad/TADHS000000071461', 'start': '34185863', 'end': '35155863'},
{'tad_id': 'tad/TADHS000000071462', 'start': '34305863', 'end': '35455863'},
{'tad_id': 'tad/TADHS000000071463', 'start': '34325863', 'end': '35025863'},
{'tad_id': 'tad/TADHS000000071465', 'start': '35150863', 'end': '35800863'},
{'tad_id': 'tad/TADHS000000071468', 'start': '35170863', 'end': '35810863'},
{'tad_id': 'tad/TADHS000000071479', 'start': '35195863', 'end': '35815863'},
{'tad_id': 'tad/TADHS000000071484', 'start': '35215863', 'end': '35825863'},
{'tad_id': 'tad/TADHS000000071485', 'start': '35225863', 'end': '35825863'},
{'tad_id': 'tad/TADHS000000071489', 'start': '35375863', 'end': '35725863'},
{'tad_id': 'tad/TADHS000000071490', 'start': '35375863', 'end': '35775863'},
{'tad_id': 'tad/TADHS000000071491', 'start': '35375863', 'end': '35825863'},
{'tad_id': 'tad/TADHS000000071492', 'start': '35425863', 'end': '35725863'}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"chromosome"	entrada	string	Cromosoma al que pertenece el segmento genómico	obo:BFO_0000050 (part of)	tad
"start"	entrada	integer	Posición de	obo:GENO_0000894	tad

			inicio del segmento genómico	(start position)	
“end”	entrada	integer	Posición de fin del segmento genómico	obo:GENO_0000895 (end position)	tad
“tad_name”	salida	string	Nombre del tad que se encuentra dentro de las coordenadas genómicas	skos:prefLabel	tad
“start”	salida	integer	Posición de inicio del tad	obo:GENO_0000894 (start position)	tad
“end”	salida	integer	Posición de fin del tad	obo:GENO_0000895 (end position)	tad

Función gene2protein(gene,taxon)

Descripción: Esta función permite obtener las proteínas codificadas por el gen introducido en la consulta, y en el taxón seleccionado. Si el valor de taxón es **None**, devolverá las proteínas codificadas por el gen en los distintos taxones. Esta información se obtendrá del grafo <http://rdf.biogateway.eu/graph/gene>

Parámetros:

-”**gene**”: Este parámetro corresponde al nombre del gen en formato symbol (propiedad: prefLabel). Ejemplo: “BRCA1”

-”**taxon**”: Permite seleccionar el taxón en el que se quiere realizar la consulta. El valor puede ser el identificador taxonómico del NCBI o el nombre del organismo (propiedad: label). Ejemplo: “Homo sapiens”, “9606”.

Salida:

Devuelve una lista con las proteínas codificadas por el gen de interés:

-”**protein_name**”: La función devuelve los nombres en formato entry name de Uniprot (prefLabel) de las proteínas codificadas por el gen introducido (propiedad: SIO_010078(encodes)).

Ejemplo de ejecución:

Entrada: gene2protein("TOX3", "9606")

Salida:

```
[{'protein_name': 'H3BTZ9_HUMAN'},  
{protein_name: 'J3QQQ6_HUMAN'},  
{protein_name: 'TOX3_HUMAN'}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"gene"	entrada	string	Nombre del gen en formato symbol	skos:prefLabel	gene
"taxon"	entrada	string	Número de taxón o nombre del organismo	skos:label	taxon
"protein_name"	salida	string	Nombre de la proteína en formato entry name de Uniprot codificada por el gen introducido	sio:SIO_010078 (encodes)	gene

Función protein2gene(protein)

Descripción: Esta función permite obtener el gen que codifica la proteína introducida en la consulta. Esta información se obtendrá del grafo <http://rdf.biogateway.eu/graph/gene>

Parámetros:

-**"protein"**: Insertamos el nombre de la proteína tanto en formato entry name de Uniprot (propiedad: prefLabel) como en formato entry de Uniprot (propiedad: altLabel). Ejemplo: "BRCA1_HUMAN", "P38398".

Salida:

-**"gene_name"**: La función devuelve el nombre en formato symbol (prefLabel) del gen que codifica la proteína introducida (propiedad: SIO_010078).

Ejemplo de ejecución:

Entrada: `protein2gene("BRCA1_MOUSE")`

Salida: `[{'gene_name': 'Brca1'}]`

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"protein"	entrada	string	Nombre de la proteína en formato entry name de Uniprot codificada por el gen introducido	skos:prefLabel	protein
"gene_name"	salida	string	Nombre del gen en formato symbol	sio:SIO_010078 (encodes)	gene

Función gene2phen(gene)

Descripción: Esta función posibilita la obtención de los fenotipos asociados a un gen previamente introducido como parámetro de esta función. Para ello explotaremos el grafo <http://rdf.biogateway.eu/graph/gene2phen>.

Parámetros:

-**"gene"**: Este parámetro corresponde al nombre del gen en formato symbol (propiedad: prefLabel). Ejemplo: "BRCA1".

Salida:

La función devuelve una lista de los fenotipos relacionados con el gen de interés. A subes de cada fenotipo obtenemos los siguientes datos:

-**"omim_id"**: Corresponde al identificador Omim del fenotipo que está asociado al gen introducido (propiedad: RO_0002331 (involved in)).

-**"phen_label"**: Se refiere a la etiqueta preferente del fenotipo al que corresponde el identificador omim (pref_label).

Ejemplo de ejecución:

Entrada: gene2phen("BRCA1")

Salida:

```
[{'omim_id': '114480', 'phen_label': 'Breast cancer (BC)'},  
{ 'omim_id': '167000', 'phen_label': 'Ovarian cancer (OC)'},  
{ 'omim_id': '604370',  
  'phen_label': 'Breast-ovarian cancer, familial, 1 (BROVCA1)'},  
{ 'omim_id': '617883',  
  'phen_label': 'Fanconi anemia, complementation group S (FANCS)'}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"gene"	entrada	string	Nombre de una enfermedad o identificador OMIM de un fenotipo	obo:RO_0002331 (involved in)	gene2phen
"omim_id"	salida	string	Identificador OMIM del fenotipo (opcional)	obo:RO_0002331 (involved in)	gene2phen
"phen_label"	salida	string	Nombre común del fenotipo	skos:prefLabel	gene2phen

Función phen2gene(phenotype)

Descripción: Esta función permite la obtención de los genes asociados a un fenotipo previamente introducido como parámetro de esta función. Para ello explotaremos el grafo <http://rdf.biogateway.eu/graph/gene2phen>.

Parámetros:

-**“phenotype”**: Este parámetro corresponde al fenotipo de interés. Permite tanto su identificador OMIM (“211980”), como el nombre de una enfermedad(“lung cancer”).

Salida:

-**“gene_name”**: La función devuelve una lista con los nombres en formato symbol (propiedad: prefLabel) de los genes relacionados con el fenotipo de interés (propiedad: RO_0002331).

Ejemplo de ejecución:

Entrada: phen2gene("lung cancer")

Salida:

[{'gene_name': 'MXRA5'}, {'gene_name': 'BRAF'}, {'gene_name': 'ERBB2'}, {'gene_name': 'SLC22A18'}]

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“phenotype”	entrada	string	Nombre de una enfermedad o identificador OMIM de un fenotipo	obo:RO_0002331 (involved in)	gene2phen
“gene_name”	salida	string	Nombre del gen en formato symbol	obo:RO_0002331 (involved in)	gene2phen

Función prot2bp(protein)

Descripción: Esta función permite estudiar los procesos biológicos en los que participa la proteína de interés, explotando la información del grafo <http://rdf.biogateway.eu/graph/prot2bp>.

Parámetros:

-**“protein”**: Insertamos el nombre de la proteína tanto en formato entry name de Uniprot (propiedad: prefLabel) como en formato entry de Uniprot (propiedad: altLabel). Ejemplo: “BRCA1_HUMAN”, “P38398”.

Salida:

La función devuelve una lista con los procesos biológicos relacionados con la proteína seleccionada. Presenta los siguientes campos:

-**“bp_id”**: Corresponde al identificador del proceso biológico en Gene Ontology (propiedad: obol:Owl#id). Ejemplo: “GO:0035349”

-**“bp_label”**: Se refiere a la etiqueta del proceso biológico en el grafo de conocimiento <http://rdf.biogateway.eu/graph/go> (propiedad: label). Ejemplo: “coenzyme A transmembrane transport”

-**“relation_label”**: Hace referencia a la etiqueta que presenta la relación entre la proteína y el proceso biológico de interés (propiedad: prefLabel) disponible en el grafo <http://rdf.biogateway.eu/graph/prot2mf>.

-**“database”**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre la proteína y el proceso biológico de interés (propiedad: SIO_000253 (has source)).

-**“articles”**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre la proteína y el proceso biológico de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: `prot2bp("BRCA1_HUMAN")`

Salida:

```
[{'bp_id': 'GO:0035066',  
  'bp_label': 'positive regulation of histone acetylation',  
  'relation_label': 'P38398--GO:0035066',  
  'database': 'goa',  
  'articles': 'pubmed/21873635'},  
 {'bp_id': 'GO:0045786',  
  'bp_label': 'negative regulation of cell cycle',
```

```

'relation_label': 'P38398--GO:0045786',
'database': 'goa/',
'articles': 'pubmed/15159397'},
{'bp_id': 'GO:0007095',
'bp_label': 'mitotic G2 DNA damage checkpoint signaling',
'relation_label': 'P38398--GO:0007095',
'database': 'goa/',
'articles': 'pubmed/19261748; pubmed/17643121; pubmed/17525340'}}

```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“protein”	entrada	string	Nombre de la proteína en formato entry name	obo:RO_0002331 (involved in)	prot2bp
“bp_id”	salida	string	Identificador en Gene Ontology del proceso biológico relacionado con la proteína de interés	obo:RO_0002331 (involved in)	prot2bp
“bp_label”	salida	string	Etiqueta del proceso biológico en el grafo	skos:label	go
“relation_label”	salida	string	Etiqueta que presenta la relación entre la proteína y el proceso biológico de interés	skos:prefLabel	prot2bp
“database”	salida	string	Base de datos con la información sobre la relación	sio:SIO_0000253 (has source)	prot2bp

“articles”	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	prot2bp
-------------------	---------------	---------------	--	---------------------------------------	----------------

Función bp2prot(biological_process,taxon)

Descripción: Esta función devuelve las proteínas relacionadas con un proceso biológico específico explotando la información del grafo <http://rdf.biogateway.eu/graph/prot2bp>.

Parámetros:

-**“biological_process”**: Este parámetro es el proceso biológico de interés, los valores permitidos son: su identificador en Gene Ontology (Ejemplo: “GO:0035349”) o el nombre de un proceso biológico (Ejemplo: “coenzyme A transmembrane transport”).

-**“taxon”**: Permite seleccionar el taxón en el que se quiere realizar la consulta. El valor puede ser el identificador taxonómico del NCBI o el nombre del organismo (propiedad: label). Ejemplo: “Homo sapiens”, “9606”. Si el valor del taxon es **None**, aplicará la búsqueda con todos los taxones disponibles en el grafo de conocimiento.

Salida:

La función devuelve una lista con las proteínas relacionadas con el proceso biológico especificado. Presenta los siguientes campos:

-**“protein_name”**: La función devuelve una lista con los nombres en formato entry name de Uniprot (prefLabel) de las proteínas relacionadas con el proceso biológico de interés (propiedad: RO_0002331 (involved in)).

-**“relation_label”**: Hace referencia a la etiqueta que presenta la relación entre la proteína y el proceso biológico de interés (propiedad: prefLabel) disponible en el grafo <http://rdf.biogateway.eu/graph/prot2mf>.

-**“database”**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre la proteína y el proceso biológico de interés (propiedad: SIO_000253 (has source)).

-**“articles”**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre la proteína y el proceso biológico de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: bp2prot("GO:0043524","Homo sapiens")

Salida:

```
[{'protein_name': 'FZD9_HUMAN',  
  'relation_label': 'O00144--GO:0043524',  
  'database': 'goa/',  
  'articles': 'pubmed/27509850'},  
{'protein_name': 'HTRA2_HUMAN',  
  'relation_label': 'O43464--GO:0043524',  
  'database': 'goa/',  
  'articles': 'pubmed/18221368'},  
{'protein_name': 'NGF_HUMAN',  
  'relation_label': 'P01138--GO:0043524',  
  'database': 'goa/',  
  'articles': 'pubmed/21873635'},  
{'protein_name': 'GDNF_HUMAN',  
  'relation_label': 'P39905--GO:0043524',  
  'database': 'goa/',  
  'articles': 'pubmed/8493557; pubmed/21873635'}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“biological_process”	entrada	string	Identificador en Gene Ontology o nombre del proceso biológico	obo:RO_0002331 (involved in)	prot2bp
“taxon”	entrada (opcional)	string	Número de taxón o nombre del organismo	obo:RO_0002162 (in taxon)	prot
“protein_name”	salida	string	Nombre de la proteína relacionada con el proceso biológico en formato entry name de Uniprot	obo:RO_0002331 (involved in)	prot2bp
“relation_label”	salida	string	Etiqueta que presenta la	skos:prefLabel	prot2bp

			relación entre la proteína y el proceso biológico de interés		
"database"	salida	string	Base de datos con la información sobre la relación	sio:SIO_0000253 (has source)	prot2bp
"articles"	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	prot2bp

Función prot2cc(protein)

Descripción: Esta función facilita el estudio de los componentes celulares en los que participa la proteína de interés, explotando la información del grafo
<http://rdf.biogateway.eu/graph/prot2cc>

Parámetros:

-**"protein"**: Insertamos el nombre de la proteína tanto en formato entry name de Uniprot (propiedad: prefLabel) como en formato entry de Uniprot (propiedad: altLabel). Ejemplo: "BRCA1_HUMAN", "P38398".

Salida:

La función devuelve una lista con los componentes celulares relacionados con la proteína seleccionada. Presenta los siguientes campos:

-**"cc_id"**: Corresponde al identificador del componente celular en Gene Ontology (propiedad: obol:Owl#id) relacionado con nuestra proteína (propiedad: BFO_0000050). Ejemplo: "GO:0005634"

-**"cc_label"**: Se refiere a la etiqueta del componente celular en el grafo de conocimiento <http://rdf.biogateway.eu/graph/go> (propiedad: label). Ejemplo: "nucleus".

-**"relation_label"**: Hace referencia a la etiqueta que presenta la relación entre la proteína y el componente celular de interés (propiedad: prefLabel) disponible en el grafo <http://rdf.biogateway.eu/graph/prot2mf> .

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre la proteína y el componente celular de interés (propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre la proteína y el componente celular de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: prot2cc("BRCA1_HUMAN")

Salida:

```
[{'cc_id': 'GO:0000151',  
  'cc_label': 'ubiquitin ligase complex',  
  'relation_label': 'P38398--GO:0000151',  
  'database': 'goa/',  
  'articles': 'pubmed/14976165'},  
{'cc_id': 'GO:0000152',  
  'cc_label': 'nuclear ubiquitin ligase complex',  
  'relation_label': 'P38398--GO:0000152',  
  'database': 'goa/',  
  'articles': 'pubmed/14636569'},  
{'cc_id': 'GO:0000800',  
  'cc_label': 'lateral element',  
  'relation_label': 'P38398--GO:0000800',  
  'database': 'goa/',  
  'articles': 'pubmed/9774970'},  
{'cc_id': 'GO:0000931',  
  'cc_label': 'gamma-tubulin ring complex',  
  'relation_label': 'P38398--GO:0000931',  
  'database': 'goa/',  
  'articles': 'pubmed/12214252'},  
{'cc_id': 'GO:0005634',  
  'cc_label': 'nucleus',  
  'relation_label': 'P38398--GO:0005634',  
  'database': 'goa/',  
  'articles': 'pubmed/17643121; pubmed/20656689; pubmed/26833090; pubmed/17525340;  
pubmed/14636569; pubmed/9342365; pubmed/22369660; pubmed/19369211;  
pubmed/20160719; pubmed/23855721; pubmed/18171670'}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“protein”	entrada	string	Nombre de la proteína en formato entry name de Uniprot	obo:RO_0002331 (involved in)	prot2cc
“cc_id”	salida	string	Identificador en Gene Ontology del componente celular relacionado con la proteína de interés	obo:RO_0002331 (involved in)	prot2cc
“cc_label”	salida	string	Etiqueta del proceso biológico componente celular en el grafo	skos:label	go
“relation_label”	salida	string	Etiqueta que presenta la relación entre la proteína y el componente celular de interés	skos:prefLabel	prot2cc
“database”	salida	string	Base de datos con la información sobre la relación	sio:SIO_0000253 (has source)	prot2cc
“articles”	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	prot2cc

Función cc2prot(cellular_component,taxon)

Descripción: Esta función devuelve las proteínas relacionadas con un componente celular específico explotando la información del grafo <http://rdf.biogateway.eu/graph/prot2cc>.

Parámetros:

-**"cellular_component"**: Este parámetro es el componente celular de interés, los valores permitidos son: su identificador en Gene Ontology (Ejemplo: "GO:0034703") o el nombre del componente celular (Ejemplo: "cation channel complex").

-**"taxon"**: Permite seleccionar el taxón en el que se quiere realizar la consulta. El valor puede ser el identificador taxonómico del NCBI o el nombre del organismo (propiedad: label). Ejemplo: "Homo sapiens", "9606". Si el valor del taxon es **None**, aplicará la búsqueda con todos los taxones disponibles en el grafo de conocimiento.

Salida:

La función devuelve una lista con las proteínas relacionadas con el componente celular especificado. Presenta los siguientes campos:

-**"protein_name"**: La función devuelve una lista con los nombres en formato entry name de Uniprot (prefLabel) de las proteínas relacionadas con el componente celular de interés (propiedad: BFO_0000050(part of)).

-**"relation_label"**: Hace referencia a la etiqueta que presenta la relación entre la proteína y el componente celular de interés (propiedad: prefLabel) disponible en el grafo <http://rdf.biogateway.eu/graph/prot2mf>.

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre la proteína y el componente celular de interés (propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre la proteína y el componente celular de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: cc2prot("GO:0034703","9606")

Salida:

```
[{'protein_name': 'TRPC3_HUMAN',  
  'relation_label': 'Q13507--GO:0034703',  
  'database': 'goa',
```

```

'articles': 'pubmed/21873635'},
{'protein_name': 'PKD2_HUMAN',
'relation_label': 'Q13563--GO:0034703',
'database': 'goa/',
'articles': 'pubmed/30093605'},
{'protein_name': 'UNC80_HUMAN',
'relation_label': 'Q8N2C7--GO:0034703',
'database': 'goa/',
'articles': 'pubmed/21873635'},
{'protein_name': 'TRPC6_HUMAN',
'relation_label': 'Q9Y210--GO:0034703',
'database': 'goa/',
'articles': 'pubmed/21873635'},
{'protein_name': 'PKD1_HUMAN',
'relation_label': 'P98161--GO:0034703',
'database': 'goa/',
'articles': 'pubmed/30093605'}}]

```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“cellular_component”	entrada	string	Identificador en Gene Ontology o nombre del componente	obo:RO_0002331 (involved in)	prot2cc
“taxon”	entrada (opcional)	string	Número de taxón o nombre del organismo	obo:RO_0002162 (in taxon)	prot
“protein_name”	salida	string	Nombre de la proteína relacionada con el componente celular en formato entry name de Uniprot	obo:RO_0002331 (involved in)	prot2cc
“relation_label”	salida	string	Etiqueta que presenta la relación entre la proteína y el	skos:prefLabel	prot2cc

			componente celular de interés		
"database"	salida	string	Base de datos con la información sobre la relación	sio:SIO_0000253 (has source)	prot2cc
"articles"	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	prot2cc

Función prot2mf(protein)

Descripción: Esta función permite estudiar las funciones moleculares en las que participa la proteína de interés, explotando la información del grafo <http://rdf.biogateway.eu/graph/prot2mf>

Parámetros:

-**"protein"**: Insertamos el nombre de la proteína tanto en formato entry name de Uniprot (propiedad: prefLabel) como en formato entry de Uniprot (propiedad: altLabel). Ejemplo: "BRCA1_HUMAN", "P38398".

Salida:

La función devuelve una lista con las funciones moleculares relacionadas con la proteína seleccionada. Presenta los siguientes campos:

-**"mf_id"**: Corresponde al identificador de la función molecular en Gene Ontology (propiedad: obol:Owl#id) de las funciones moleculares relacionadas con la proteína de interés (propiedad: RO_0002327 (enables)) . Ejemplo: "GO:0004672"

-**"mf_label"**: Se refiere a la etiqueta de la función molecular en el grafo de conocimiento <http://rdf.biogateway.eu/graph/go> (propiedad: label). Ejemplo: "protein kinase activity"

-**"relation_label"**: Hace referencia a la etiqueta que presenta la relación entre la función molecular y la proteína de interés (propiedad: prefLabel) disponible en el grafo <http://rdf.biogateway.eu/graph/prot2mf> .

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre la función molecular y la proteína de interés (propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre la función molecular y la proteína de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: prot2mf("BRCA1_HUMAN")

Salida:

```
{'mf_id': 'GO:0002039',  
  'mf_label': 'p53 binding',  
  'relation_label': 'P38398--GO:0002039',  
  'database': 'goa/',  
  'articles': 'pubmed/15571721'},  
{'mf_id': 'GO:0003677',  
  'mf_label': 'DNA binding',  
  'relation_label': 'P38398--GO:0003677',  
  'database': 'goa/',  
  'articles': 'pubmed/9662397'},  
{'mf_id': 'GO:0003713',  
  'mf_label': 'transcription coactivator activity',  
  'relation_label': 'P38398--GO:0003713',  
  'database': 'goa/',  
  'articles': 'pubmed/9662397'},  
{'mf_id': 'GO:0003723',  
  'mf_label': 'RNA binding',  
  'relation_label': 'P38398--GO:0003723',  
  'database': 'goa/',  
  'articles': 'pubmed/12419249'},  
{'mf_id': 'GO:0004842',  
  'mf_label': 'ubiquitin-protein transferase activity',  
  'relation_label': 'P38398--GO:0004842',  
  'database': 'goa/',  
  'articles': 'pubmed/20351172; pubmed/21873635; pubmed/17349954; pubmed/19117993;  
pubmed/12890688'}}
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"protein"	entrada	string	Nombre de la proteína en formato entry name	obo:RO_0002331 (involved in)	prot2mf

“mf_id”	salida	string	Identificador en Gene Ontology de la función molecular relacionada con la proteína de interés	obo:RO_0002331 (involved in)	prot2mf
“mf_label”	salida	string	Etiqueta la función molecular en el grafo	skos:label	go
“relation_label”	salida	string	Etiqueta que presenta la relación entre la proteína y la función molecular de interés	skos:prefLabel	prot2mf
“database”	salida	string	Base de datos con la información sobre la relación	sio:SIO_0000253 (has source)	prot2mf
“articles”	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	prot2mf

Función mf2prot(molecular_function,taxon)

Descripción: Esta función devuelve las proteínas relacionadas con la función molecular especificada explotando la información del grafo <http://rdf.biogateway.eu/graph/prot2mf>

Parámetros:

-**"molecular_function"**: Este parámetro es la función molecular de interés, los valores permitidos son: su identificador en Gene Ontology (Ejemplo: "GO:0004672") o el nombre del componente celular (Ejemplo: "protein kinase activity").

-**"taxon"**: Permite seleccionar el taxón en el que se quiere realizar la consulta. El valor puede ser el identificador taxonómico del NCBI o el nombre del organismo (propiedad: label). Ejemplo: "Homo sapiens", "9606". Si el valor del taxon es **None**, aplicará la búsqueda con todos los taxones disponibles en el grafo de conocimiento.

Salida:

La función devuelve una lista con las proteínas relacionadas con la función molecular de interés. Presenta los siguientes campos:

-**"protein_name"**: Son los nombres en formato entry name de Uniprot (prefLabel) de las proteínas relacionadas con el componente celular de interés (propiedad: RO_0002327).

-**"relation_label"**: Hace referencia a la etiqueta que presenta la relación entre la proteína y la función molecular de interés (propiedad: prefLabel) disponible en el grafo **<http://rdf.biogateway.eu/graph/prot2mf>**.

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre la proteína y la función molecular de interés (propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre la proteína y la función molecular de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: mf2prot("protein binding","9606")

Salida:

```
[{'protein_name': 'ZN217_HUMAN',  
  'relation_label': 'O75362--GO:0005515',  
  'database': 'goa',  
  'articles': 'pubmed/16940172'},  
{'protein_name': 'SIR6_HUMAN',  
  'relation_label': 'Q8N6T7--GO:0005515',  
  'database': 'goa',  
  'articles': 'pubmed/19135889; pubmed/23217706; pubmed/23911928'},  
{'protein_name': 'NELFE_HUMAN',  
  'relation_label': 'P18615--GO:0005515',  
  'database': 'goa',  
  'articles': 'pubmed/32296183; pubmed/26496610; pubmed/14667819; pubmed/28514442;  
pubmed/25416956; pubmed/24981860; pubmed/12612062; pubmed/20211142'}]
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“molecular_function”	entrada	string	Identificador en Gene Ontology o nombre de la función molecular	obo:RO_0002331 (involved in)	prot2mf
“taxon”	entrada (opcional)	string	Número de taxón o nombre del organismo	obo:RO_0002162 (in taxon)	prot
“protein_name”	salida	string	Nombre de la proteína relacionada con la función molecular en formato entry name de Uniprot	obo:RO_0002331 (involved in)	prot2mf
“relation_label”	salida	string	Etiqueta que presenta la relación entre la proteína y la función molecular de interés	skos:prefLabel	prot2mf
“database”	salida	string	Base de datos con la información sobre la relación	sio:SIO_0000253 (has source)	prot2mf
“articles”	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	prot2mf

Función gene2crm(gene)

Descripción: Esta función nos permite obtener los crms asociados al gen introducido como parámetro. Para ello usaremos la información del grafo <http://rdf.biogateway.eu/graph/crm2gene>.

Parámetros:

-**"gene"**: Este parámetro corresponde al nombre del gen en formato symbol (propiedad: prefLabel). Ejemplo: "TOX3".

Salida:

La función devuelve un diccionario con los crms que están relacionados con el gen introducido. Presenta los siguientes campos:

-**"crm_name"**: Nombre del módulo cis-regulador (propiedad: prefLabel) que está relacionado con el gen introducido (propiedad: RO_0002429).

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre el crm y el gen de interés (propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre el crm y el gen de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: gene2crm("TOX3")

Salida:

```
{'crm_name': 'crm/CRMHS00000005857',  
  'database': 'http://lcbi.swjtu.edu.cn/EnhFFL'; http://218.8.241.248:8080/SEA3/  
http://bioinfo.vanderbilt.edu/AE/HACER;  
https://genome.ucsc.edu/cgi-bin/hgTrackUi?db=hg38&g=geneHancer;  
http://enhanceratlas.net/scenhancer; https://webs.iiitd.edu.in/raghava/cancerend/  
https://enhancer.lbl.gov; http://www.licpathway.net/sedb/  
http://health.tsinghua.edu.cn/jianglab/endisease/  
http://biocc.hrbmu.edu.cn/DiseaseEnhancer/; https://fantom.gsc.riken.jp/5/  
http://acgt.cs.tau.ac.il/focs; https://asntech.org/dbsuper/  
http://yiplab.cse.cuhk.edu.hk/jeme/  
  'articles': 'pubmed/24119843; pubmed/28869592; pubmed/30247654; pubmed/23374354;  
pubmed/29716618; pubmed/23001124; pubmed/35694152; pubmed/24670763;  
pubmed/32360910; pubmed/17130149; pubmed/30371817; pubmed/28605766;  
pubmed/34761274; pubmed/31667506'}
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"gene"	entrada	string	Nombre del gen en formato symbol	obo:RO_0002429 (involved in positive regulation of)	crm2gene
"crm_name"	salida	string	Nombre del módulo cis-regulador relacionado con el gen introducido	obo:RO_0002429 (involved in positive regulation of)	crm2gene
"database"	salida	string	Base de datos con la información sobre la relación entre el crm y el gen.	sio:SIO_0000253 (has source)	crm2gene
"articles"	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	crm2gene

Función crm2gene(crm)

Descripción: Esta función facilita la obtención de los genes a los que un crm específico afecta. Para ello usaremos la información del grafo <http://rdf.biogateway.eu/graph/crm2gene> .

Parámetros:

-"**crm**": El parámetro introducido en la función será el identificador preferente del módulo cis-regulador (propiedad: prefLabel). Ejemplo "crm/CRMHS00003225754".

Salida:

La función devuelve un diccionario con los genes que están relacionados con el crm de interés. Presenta los siguientes campos:

-**"gene_name"**: La función devuelve una lista con los nombres en formato symbol (propiedad: prefLabel) de los genes relacionados con el crm de interés (propiedad: RO_0002429).

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre el gen y el crm de interés (propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre el gen y el crm de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: crm2gene("crm/CRMHS000000137026")

Salida:

```
[{'gene_name': 'CRAT',  
  'database': 'http://bioinfo.vanderbilt.edu/AE/HACER; https://fantom.gsc.riken.jp/5;  
https://webs.iitd.edu.in/raghava/cancerend/; http://yiplab.cse.cuhk.edu.hk/jeme/',  
  'articles': 'pubmed/24670763; pubmed/32360910; pubmed/28869592; pubmed/30247654'},  
 {'gene_name': 'NTMT1',  
  'database': 'http://bioinfo.vanderbilt.edu/AE/HACER; https://fantom.gsc.riken.jp/5;  
https://webs.iitd.edu.in/raghava/cancerend/; http://yiplab.cse.cuhk.edu.hk/jeme/',  
  'articles': 'pubmed/24670763; pubmed/32360910; pubmed/28869592; pubmed/30247654'},  
 {'gene_name': 'PTPA',  
  'database': 'http://bioinfo.vanderbilt.edu/AE/HACER; https://fantom.gsc.riken.jp/5;  
https://webs.iitd.edu.in/raghava/cancerend/; http://yiplab.cse.cuhk.edu.hk/jeme/',  
  'articles': 'pubmed/24670763; pubmed/32360910; pubmed/28869592; pubmed/30247654'}
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"crm"	entrada	string	Nombre del módulo cis-regulador Nombre del gen en	obo:RO_0002429 (involved in positive regulation of)	crm2gene

			formato symbol		
"gene_name"	salida	string	Nombre del gen en formato symbol relacionado con el crm	obo:RO_0002429 (involved in positive regulation of)	crm2gene
"database"	salida	string	Base de datos con la información sobre la relación entre el crm y el gen	sio:SIO_0000253 (has source)	crm2gene
"articles"	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	crm2gene

Función tfac2crm(tfac)

Descripción: La función posibilita la obtención de los crms asociados a un factor de transcripción específico, información que encontramos en el grafo <http://rdf.biogateway.eu/graph/crm2tfac>

Parámetros:

-**"tfac"**: Este parámetro corresponde al nombre en formato entry name de Uniprot del factor de transcripción del cual queremos conocer con qué crms interactúa.

Salida:

La función devuelve un diccionario con los crms que están relacionados con el factor de transcripción introducido. Presenta los siguientes campos:

-**"crm_name"**: Nombre de los módulos cis-reguladores ((propiedad: prefLabel)) con los que interactúa el factor de transcripción especificado (propiedad: RO_0002436).

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre el crm y el factor de transcripción(propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre el crm y el factor de transcripción de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

-**"evidence"**: Corresponde a la evidencia que respalda la información disponible sobre la relación entre el factor de transcripción y el crm (propiedad: evidenceOrigin).

-**"biological_samples"**: Se refiere a los diferentes tipos de muestras biológicas que están asociados con el estudio de la relación entre el módulo cis-regulador y el factor de transcripción (propiedad: TXPO_0003500 (observed in)). En concreto, devolverá los identificadores en formato de términos ontológicos. Ejemplo: "CLO_0001601", "UBERON_0002113", "BTO_0000018".

Ejemplo de ejecución:

Entrada: tfac2crm("FOSL2_HUMAN")

Salida:

```
{'crm_name': 'crm/CRMHS00000005425',  
  'database': 'http://www.licpathway.net/ENdb/',  
  'articles': 'pubmed/29149598',  
  'evidence':  
    'http://www.licpathway.net/ENdb/search/Detail.php?Species=Human&Enhancer_id=E_01_292',  
  'biological_samples': 'UBERON_0001003; CL_0000148'},  
{'crm_name': 'crm/CRMHS00000006865',  
  'database': 'http://lcbbswjtuedu.cn/EnhFFL/',  
  'articles': 'pubmed/35694152',  
  'evidence':  
    'http://lcbbswjtuedu.cn/EnhFFL/details/?term=enH639113&subtype=enhancer&species=human',  
  'biological_samples': 'CL_0002319; UBERON_0002421; BTO_0000601'},  
{'crm_name': 'crm/CRMHS00000006866',  
  'database': 'http://lcbbswjtuedu.cn/EnhFFL/',  
  'articles': 'pubmed/35694152',  
  'evidence':  
    'http://lcbbswjtuedu.cn/EnhFFL/details/?term=enH594653&subtype=enhancer&species=human',  
  'biological_samples': 'CL_0002319; UBERON_0002421; BTO_0000601'}}
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"tfac"	entrada	string	Nombre del del factor de	obo:RO_0002436 (molecularly	crm2tfac

			transcripción en formato entry name de Uniprot	interacts with)	
"crm_name"	salida	string	Nombre del módulo cis-regulador relacionado con el factor de transcripción introducido	obo:RO_0002436 (molecularly interacts with)	crm2tfac
"database"	salida	string	Base de datos con la información sobre la relación entre el crm y el tfac.	sio:SIO_0000253 (has source)	crm2tfac
"articles"	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	crm2tfac
"evidence"	salida	string	Nivel de evidencia asociado la relación	sch:evidenceOrigin	crm2tfac
"biological_samples"	salida	string	Muestras biológicas asociadas a la relación	obo:TXPO_0003500 (observed in)	crm2tfac

Función crm2tfac(crm)

Descripción: Esta función permite estudiar los factores de transcripción que interactúan con un crm determinado, usando la información del grafo <http://rdf.biogateway.eu/graph/crm2tfac>.

Parámetros:

-**"crm"**: El parámetro introducido en la función será el identificador preferente del módulo cis-regulador (propiedad: prefLabel). Ejemplo "crm/CRMHS00000007832".

Salida:

La función devuelve un diccionario con los factores de transcripción que están relacionados con el crm de interés. Presenta los siguientes campos:

-**“tfac_name”**: Nombres en formato entry name de Uniprot de los factores de transcripción que interactúan con el crm de interés. Ejemplo: “TF7L2_HUMAN”.

-**“database”**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre el factor de transcripción y el crm de interés (propiedad: SIO_000253 (has source)).

-**“articles”**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre el factor de transcripción y el crm de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

-**“evidence”**: Corresponde a la evidencia que respalda la información disponible sobre la relación entre el factor de transcripción y el crm de interés(propiedad: evidenceOrigin).

-**“biological_samples”**: Se refiere a los diferentes tipos de muestras biológicas que están asociados con el estudio de la relación entre el factor de transcripción y el crm introducido (propiedad: TXPO_0003500 (observed in)). En concreto, devolverá los identificadores en formato de términos ontológicos. Ejemplo: “CLO_0001601”, “UBERON_0002113”, “BTO_0000018”.

Ejemplo de ejecución:

Entrada: crm2tfac("crm/CRMHS00000007832")

Salida:

```
{'tfac_name': 'RAD21_HUMAN',  
  'database': 'http://lcbbswjtuedu.cn/EnhFFL',  
  'articles': 'pubmed/35694152',  
  'evidence':  
    'http://lcbbswjtuedu.cn/EnhFFL/details/?term=enH58066&subtype=enhancer&species=human',  
  'biological_samples': 'UBERON_0005090; CL_0000187; BTO_0000887'},  
{'tfac_name': 'SRF_HUMAN',  
  'database': 'http://lcbbswjtuedu.cn/EnhFFL',  
  'articles': 'pubmed/35694152',  
  'evidence':  
    'http://lcbbswjtuedu.cn/EnhFFL/details/?term=enH58066&subtype=enhancer&species=human',  
  'biological_samples': 'UBERON_0005090; CL_0000187; BTO_0000887'},  
{'tfac_name': 'TAF1_HUMAN',  
  'database': 'http://lcbbswjtuedu.cn/EnhFFL',  
  'articles': 'pubmed/35694152',
```

```

'evidence':
'http://lcbb.swjtu.edu.cn/EnhFFL/details/?term=enH58066&subtype=enhancer&species=human',
'biological_samples': 'UBERON_0005090; CL_0000187; BTO_0000887'},
{'tfac_name': 'RPB1_HUMAN',
'database': 'http://lcbb.swjtu.edu.cn/EnhFFL/',
'articles': 'pubmed/35694152',
'evidence':
'http://lcbb.swjtu.edu.cn/EnhFFL/details/?term=enH58066&subtype=enhancer&species=human',
'biological_samples': 'UBERON_0005090; CL_0000187; BTO_0000887']}

```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
"crm"	entrada	string	Nombre del módulo cis-regulador	obo:RO_0002436 (molecularly interacts with)	crm2tfac
"tfac_name"	salida	string	Nombre del del factor de transcripción en formato entry name de Uniprot relacionado con el crm introducido	obo:RO_0002436 (molecularly interacts with)	crm2tfac
"database"	salida	string	Base de datos con la información sobre la relación entre el crm y el tfac.	sio:SIO_0000253 (has source)	crm2tfac
"articles"	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	crm2tfac
"evidence"	salida	string	Nivel de evidencia asociado la relación	sch:evidenceOrigin	crm2tfac
"biological_samples"	salida	string	Muestras biológicas asociadas a la	obo:TXPO_0003500 (observed in)	crm2tfac

			relación		
--	--	--	----------	--	--

Función `crm2phen(crm)`

Descripción: Esta función posibilita la obtención de los fenotipos asociados a un `crm` previamente introducido como parámetro de esta función. Para ello explotaremos el grafo <http://rdf.biogateway.eu/graph/crm2phen>.

Parámetros:

-"**crm**": El parámetro introducido en la función será el identificador preferente del módulo `cis-regulador` (propiedad: `prefLabel`). Ejemplo `""crm/CRMHS00000005764""`.

Salida:

La función devuelve un diccionario con los fenotipos relacionados con el `crm` seleccionado. Asimismo, presenta los siguientes campos:

-"**phen_id**": Corresponde a los identificadores de los fenotipos asociados al `crm` de interés, en concreto puede devolver tanto el identificador OMIM, como el identificador DOID como el identificador MeSH.

-"**database**": Indica la base de datos donde se encuentra registrada la información sobre la relación entre el fenotipo y el `crm` de interés (propiedad: `SIO_000253` (has source)).

-"**articles**": Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre el fenotipo y el `crm` de interés, que se encuentran en la base de datos Pubmed (propiedad: `SIO_000772` (has evidence)).

Ejemplo de ejecución:

Entrada: `crm2phen("crm/CRMHS00000005764")`

Salida:

```
[{'phen_id': 'OMIM/181500',
  'database': 'http://biocc.hrbmu.edu.cn/DiseaseEnhancer/';
  http://health.tsinghua.edu.cn/jianglab/endisease/',
  'articles': 'pubmed/25453756'},
 {'phen_id': 'MESH/D012559',
  'database': 'http://biocc.hrbmu.edu.cn/DiseaseEnhancer/';
  http://health.tsinghua.edu.cn/jianglab/endisease/',
  'articles': 'pubmed/25453756'},
 {'phen_id': 'DOID/DOID_5419',
  'database': 'http://biocc.hrbmu.edu.cn/DiseaseEnhancer/';
  http://health.tsinghua.edu.cn/jianglab/endisease/'}
```


'articles': 'pubmed/25453756']}]

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“crm”	entrada	string	Nombre del módulo cis-regulador Nombre del gen en formato symbol	obo:RO_0002331 (involved in)	crm2phen
“phen_id”	salida	string	Identificador de los fenotipos asociados al crm de interés	obo:RO_0002331 (involved in)	crm2phen
“database”	salida	string	Base de datos con la información sobre la relación entre el crm y el fenotipo	sio:SIO_0000253 (has source)	crm2phen
“articles”	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	crm2phen

Función phen2crm(phenotype)

Descripción: Esta función devuelve los crms que están asociados a un fenotipo determinado. Explota la información disponible en el grafo <http://rdf.biogateway.eu/graph/crm2phen>.

Parámetros:

-”**phenotype**”: Este parámetro corresponde al fenotipo de interés. Se permite tanto su identificador OMIM (“181500”), como el nombre de una enfermedad(“schizophrenia”).

Salida:

La función devuelve un diccionario con los crms que están relacionados con el fenotipo especificado. Presenta los siguientes campos:

-**"crm_name"**: Nombre de los módulos cis-reguladores ((propiedad: prefLabel)) asociados al fenotipo introducido (propiedad: RO_0002331 (involved in)).

-**"omim_id"**: Corresponde al identificador OMIM del fenotipo que está asociado al crm (solo si se ha introducido el nombre de un fenotipo).

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre el crm y el fenotipo especificado (propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre el crm y el fenotipo de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

Ejemplo de ejecución:

Entrada: phen2crm("schizophrenia")

Salida:

```
{'crm_name': 'crm/CRMHS00000005764',  
  'omim_id': 'OMIM/181500',  
  'database': 'http://biocc.hrbmu.edu.cn/DiseaseEnhancer/;  
http://health.tsinghua.edu.cn/jianglab/endisease/',  
  'articles': 'pubmed/25453756'},  
{'crm_name': 'crm/CRMHS00000005770',  
  'omim_id': 'OMIM/181500',  
  'database': 'http://biocc.hrbmu.edu.cn/DiseaseEnhancer/;  
http://health.tsinghua.edu.cn/jianglab/endisease/',  
  'articles': 'pubmed/25453756'},  
{'crm_name': 'crm/CRMHS00000005771',  
  'omim_id': 'OMIM/181500',  
  'database': 'http://biocc.hrbmu.edu.cn/DiseaseEnhancer/;  
http://health.tsinghua.edu.cn/jianglab/endisease/',  
  'articles': 'pubmed/25434007'},  
{'crm_name': 'crm/CRMHS00000005773',  
  'omim_id': 'OMIM/181500',  
  'database': 'http://biocc.hrbmu.edu.cn/DiseaseEnhancer/;  
http://health.tsinghua.edu.cn/jianglab/endisease/',  
  'articles': 'pubmed/25453756'},  
{'crm_name': 'crm/CRMHS00000005816',  
  'omim_id': 'OMIM/181500',  
  'database': 'http://biocc.hrbmu.edu.cn/DiseaseEnhancer/;  
http://health.tsinghua.edu.cn/jianglab/endisease/',  
  'articles': 'pubmed/27276213; pubmed/25453756']}
```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“phenotype”	entrada	string	Identificador OMIM o nombre del fenotipo de interés	obo:RO_0002331 (involved in)	crm2phen
“crm_name”	salida	string	Nombre del módulo cis-regulador relacionado con el fenotipo introducido	obo:RO_0002331 (involved in)	crm2phen
“omim_id”	salida (opcional)	string	Identificador OMIM del fenotipo asociado al crm	obo:RO_0002331 (involved in)	crm2phen
“database”	salida	string	Base de datos con la información sobre la relación entre el crm y el fenotipo	sio:SIO_0000253 (has source)	crm2phen
“articles”	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	crm2phen

Función tfac2gene(tfac)

Descripción: Esta función permite obtener los genes que son regulados por un factor de transcripción determinado, gracias a la información proporcionada por el grafo <http://rdf.biogateway.eu/graph/tfac2gene>.

Parámetros:

-“**tfac**”: Este parámetro corresponde al nombre en formato entry name de Uniprot del factor de transcripción del cual queremos conocer a qué genes regula. Ejemplo: "NKX31_HUMAN".

Salida:

La función devuelve dos diccionarios con los genes que son regulados positiva y negativamente por el factor de transcripción introducido. Presenta los siguientes campos:

-**"gene_name"**: Nombre en formato symbol de los genes ((propiedad: prefLabel)) que son regulados por el factor de transcripción especificado (propiedad: RO_0002429 (involved in positive regulation of) y propiedad: RO_0002430 (involved in negative regulation of)).

-**"database"**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre el gen y el factor de transcripción(propiedad: SIO_000253 (has source)).

-**"articles"**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre el gen y el factor de transcripción de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

-**"evidence_level"**: Corresponde a la evidencia que respalda la información disponible sobre la relación entre el factor de transcripción y el gen (propiedad: evidenceOrigin).

-**"definition"**: Proporciona la definición de la relación que se produce entre cada gen y el factor de transcripción especificado (propiedad: definition), disponible en el propio grafo (<http://rdf.biogateway.eu/graph/tfac2gene>).

Ejemplo de ejecución:

Entrada: tfac2gene("NKX31_HUMAN")

Salida:

```
'Positive regulation results:',  
[{'gene_name': 'CLIC4',  
  'database': 'https://github.com/saezlab/CollecTRI',  
  'articles': 'pubmed/16316993',  
  'evidence_level': '1',  
  'definition': 'Q99801 involved in positive regulation of 9606/CLIC4'},  
{'gene_name': 'DKK3',  
  'database': 'https://github.com/saezlab/CollecTRI',  
  'articles': 'pubmed/23975733',  
  'evidence_level': '1',  
  'definition': 'Q99801 involved in positive regulation of 9606/DKK3'},  
{'gene_name': 'MAP3K5',  
  'database': 'https://github.com/saezlab/CollecTRI',  
  'articles': 'pubmed/21594902',  
  'evidence_level': '1',  
  'definition': 'Q99801 involved in positive regulation of 9606/MAP3K5'},  
{'gene_name': 'NKX3-1',  
  'database': 'https://github.com/saezlab/CollecTRI'}]
```

```

    'articles': 'pubmed/15880262; pubmed/20855495; pubmed/19886863; pubmed/20195545;
pubmed/23368843; pubmed/16270157; pubmed/19263243; pubmed/21730289;
pubmed/16763719; pubmed/20716579; pubmed/16845664',
    'evidence_level': '11',
    'definition': 'Q99801 involved in positive regulation of 9606/NKX3-1']]

'Negative regulation results:',
[{'gene_name': 'AR',
  'database': 'https://github.com/saezlab/CollecTRI',
  'articles': 'pubmed/23492366; pubmed/17202838; pubmed/20363913; pubmed/18360715;
pubmed/16697957',
  'evidence_level': '5',
  'definition': 'Q99801 involved in regulation of 9606/AR'},
{'gene_name': 'BCL2',
  'database': 'https://github.com/saezlab/CollecTRI',
  'articles': 'pubmed/24273454; pubmed/17191317; pubmed/22869582; pubmed/16316993;
pubmed/22266868; pubmed/9581775; pubmed/15817464; pubmed/19137013;
pubmed/22331597; pubmed/23313858; pubmed/12679484; pubmed/24098340;
pubmed/17486276; pubmed/14684736; pubmed/8183578; pubmed/19266349;
pubmed/21940310',
  'evidence_level': '17',
  'definition': 'Q99801 involved in regulation of 9606/BCL2'},
{'gene_name': 'CCND1',
  'database': 'https://github.com/saezlab/CollecTRI',
  'articles': 'pubmed/17639064; pubmed/22179513',
  'evidence_level': '2',
  'definition': 'Q99801 involved in regulation of 9606/CCND1'}]]

```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“tfac”	entrada	string	Nombre del del factor de transcripción en formato entry name de Uniprot	obo:RO_0002429(involved in positive regulation of) y obo:RO_0002430(involved in negative regulation of)	tfac2gene
“gene_name”	salida	string	Nombre del gen en formato symbol relacionado	obo:RO_0002429(involved in positive regulation of) y obo:RO_0002430(involved in	tfac2gene

			con el factor de transcripción	negative regulation of)	
“database”	salida	string	Base de datos con la información sobre la relación entre el factor de transcripción y el gen	sio:SIO_0000253 (has source)	tfac2gene
“articles”	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	tfac2gene
“evidence_level”	salida	integer	Evidencia que respalda la información disponible sobre la relación entre el factor de transcripción y el gen	sch:evidenceLevel	tfac2gene
“definition”	salida	string	Definición de la relación	skos:definition	tfac2gene

Función gene2tfac(gene)

Descripción: Esta función devolverá los factores de transcripción que regulan el gen introducido, usando la información disponible en el grafo <http://rdf.biogateway.eu/graph/tfac2gene>.

Parámetros:

-”**gene**”: Este parámetro corresponde al nombre del gen en formato symbol (propiedad: prefLabel). Ejemplo: “TOX3”.

Salida:

La función tiene como salida dos diccionarios con los factores de transcripción que regulan de manera positiva al gen en uno de ellos, y en el otro aquellos que lo regulan de manera negativa. Presenta los siguientes campos:

-**“tfac_name”**: Nombres en formato entry name de Uniprot de los factores de transcripción que interactúan con el gen introducido (propiedad: RO_0002429 (involved in positive regulation of) y propiedad: RO_0002430 (involved in negative regulation of)).

-**“database”**: Indica la base de datos donde se encuentra registrada la información sobre la relación entre el gen y el factor de transcripción(propiedad: SIO_000253 (has source)).

-**“articles”**: Corresponde a artículos científicos o publicaciones que están asociadas con la relación entre el factor de transcripción y el gen de interés, que se encuentran en la base de datos Pubmed (propiedad: SIO_000772 (has evidence)).

-**“evidence_level”**: Corresponde a la evidencia que respalda la información disponible sobre la relación entre el factor de transcripción y el gen (propiedad: evidenceOrigin).

-**“definition”**: Proporciona la definición de la relación que se produce entre cada factor de transcripción y el gen especificado (propiedad: definition), disponible en el grafo <http://rdf.biogateway.eu/graph/tfac2gene>.

Ejemplo de ejecución:

Entrada: gene2tfac("BRCA1")

Salida:

```
('Positive regulation results:',  
[{'tfac_name': 'BHE41_HUMAN',  
  'database': 'https://github.com/saezlab/CollecTRI',  
  'articles': 'pubmed/20006609',  
  'evidence_level': '1',  
  'definition': 'Q9C0J9 involved in positive regulation of 9606/BRCA1'},  
 {'tfac_name': 'P63_HUMAN',  
  'database': 'https://github.com/saezlab/CollecTRI',  
  'articles': 'pubmed/24556685',  
  'evidence_level': '1',  
  'definition': 'Q9H3D4 involved in positive regulation of 9606/BRCA1'},  
 {'tfac_name': 'MBD2_HUMAN',  
  'database': 'https://github.com/saezlab/CollecTRI',  
  'articles': 'pubmed/16052033; pubmed/23011797',  
  'evidence_level': '2',  
  'definition': 'Q9UBB5 involved in positive regulation of 9606/BRCA1'}]
```

```
'Negative regulation results:',  
[{'tfac_name': 'HMGA1_HUMAN',  
  'database': 'https://github.com/saezlab/CollecTRI',  
  'articles': 'pubmed/16007157; pubmed/12640109',  
  'evidence_level': '2',  
  'definition': 'P17096 involved in regulation of 9606/BRCA1'},  
 {'tfac_name': 'HMGA1_HUMAN',  
  'database': 'https://github.com/saezlab/CollecTRI',  
  'articles': 'pubmed/16007157; pubmed/12640109',
```

```

'evidence_level': '2',
'definition': 'P17096 involved in negative regulation of 9606/BRCA1'},
{'tfac_name': 'ID4_HUMAN',
'database': 'https://github.com/saezlab/CollecTRI',
'articles': 'pubmed/24475217; pubmed/12032322; pubmed/17016441; pubmed/11136250;
pubmed/21194482; pubmed/16582598',
'evidence_level': '6',
'definition': 'P47928 involved in negative regulation of 9606/BRCA1'},
{'tfac_name': 'LMO4_HUMAN',
'database': 'https://github.com/saezlab/CollecTRI',
'articles': 'pubmed/12925972; pubmed/11751867',
'evidence_level': '2',
'definition': 'P61968 involved in negative regulation of 9606/BRCA1'},

```

Tabla resumen:

Variable	Rol	Tipo	Descripción	Propiedad de la Ontología	Grafo
“gene”	entrada	string	Nombre del gen en formato symbol relacionado con el factor de transcripción	obo:RO_0002429(involved in positive regulation of) y obo:RO_0002430(involved in negative regulation of)	tfac2gene
“tfac_name”	salida	string	Nombre del del factor de transcripción en formato entry name de Uniprot	obo:RO_0002429(involved in positive regulation of) y obo:RO_0002430(involved in negative regulation of)	tfac2gene
“database”	salida	string	Base de datos con la información sobre la relación entre el factor de transcripción y el gen	sio:SIO_0000253 (has source)	tfac2gene
“articles”	salida	string	Artículos asociados a la relación	sio:SIO_0000772 (has evidence)	tfac2gene

“evidence_level”	salida	integer	Evidencia que respalda la información disponible sobre la relación entre el factor de transcripción y el gen	sch:evidenceLevel	tfac2gene
“definition”	salida	string	Definición de la relación	skos:definition	tfac2gene

Funciones auxiliares

-data_processing: Facilita el procesamiento de los datos obtenidos en las consultas SPARQL, de manera que el usuario obtenga resultados fácilmente legibles y entendibles.

-translate_chr: Permite traducir un cromosoma humano en su identificador en el NCBI. Por ejemplo, chr-1 lo traduce a NC_000001.11.