## Example SPARQL query

*What are the human genes which have a known association to glioblastoma (a type of brain cancer) and which furthermore have an orthologous gene expressed in the rat's brain?*

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX obo: <http://purl.obolibrary.org/obo/>

PREFIX orth: <http://purl.org/net/orth#>

PREFIX sio: <http://semanticscience.org/resource/>

PREFIX taxon: <http://purl.uniprot.org/taxonomy/>

PREFIX up: <http://purl.uniprot.org/core/>

PREFIX lscr: <http://purl.org/lscr#>

PREFIX genex:<http://purl.org/genex#>

SELECT DISTINCT ?protein ?orthologous\_protein\_rat ?id WHERE {

SELECT \* {

SERVICE <http://sparql.uniprot.org/sparql> {

SELECT ?protein WHERE {

?protein a up:Protein;

up:organism taxon:9606 ;

up:annotation ?annotation .

?annotation rdfs:comment ?annotation\_text.

?annotation a up:Disease\_Annotation.

FILTER CONTAINS (?annotation\_text, "**glioblastoma**")

}}

SERVICE <https://sparql.omabrowser.org/sparql> {

SELECT ?orthologous\_protein\_rat ?protein ?id WHERE {

?protein\_OMA a orth:Protein.

?orthologous\_protein\_rat a orth:Protein.

?cluster a orth:OrthologsCluster.

?cluster orth:hasHomologousMember ?node1.

?cluster orth:hasHomologousMember ?node2.

?node2 orth:hasHomologousMember\* ?protein\_OMA.

?node1 orth:hasHomologousMember\* ?orthologous\_protein\_rat.

?orthologous\_protein\_rat orth:organism/obo:RO\_0002162 taxon:10116.

?orthologous\_protein\_rat sio:SIO\_010079/lscr:xrefEnsemblGene ?id.

?protein\_OMA lscr:xrefUniprot ?protein.

FILTER(?node1 != ?node2)

}}

SERVICE <http://biosoda.expasy.org:8080/rdf4j-server/repositories/bgeelight> {

?gene genex:isExpressedIn ?anatEntity .

?anatEntity rdfs:label "brain".

?gene orth:organism ?org .

?org obo:RO\_0002162 taxon:10116.

?gene lscr:xrefEnsemblGene ?id.

}}}

## RESULTS

The following table displays the results of executing the SPARQL query above, where:

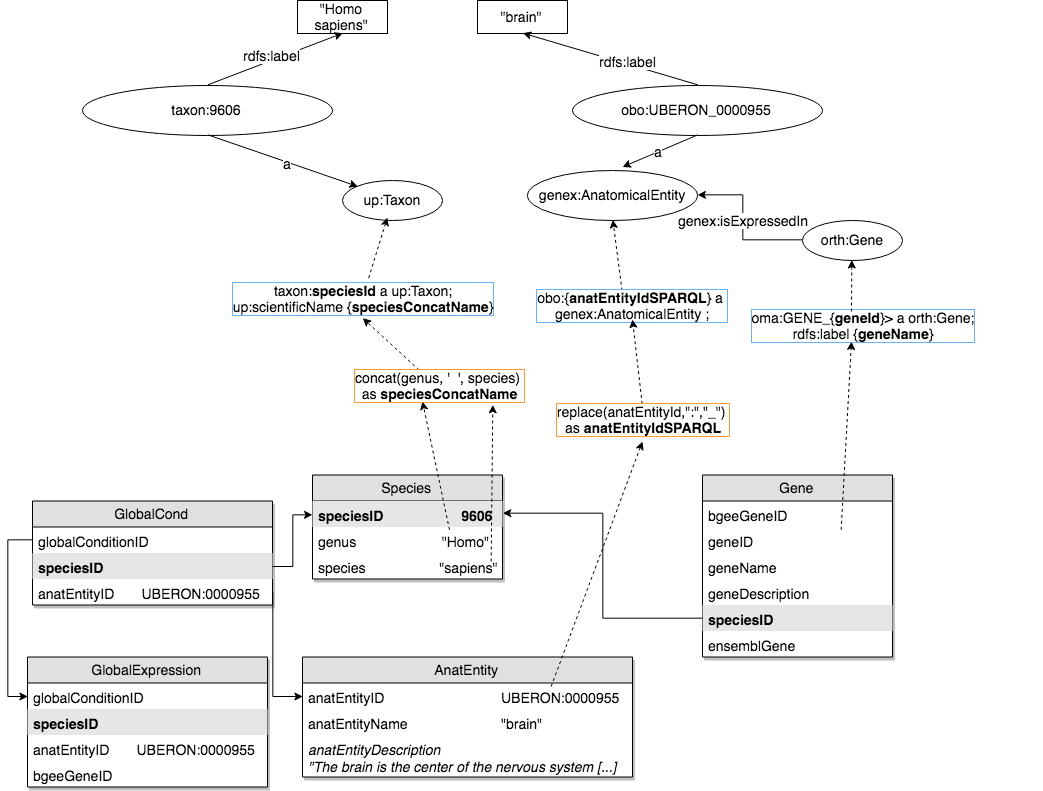
* The first column, “***protein***”, shows a UniProt human protein with a known association with glioblastoma for which there exists an orthologous protein expressed in the rat’s brain. Clicking on any of the links in this column will redirect to the corresponding UniProt entry online.
* The second column, “***orthologous\_protein\_rat***”, shows the orthologous rat protein (for which there exists known expression in the brain according to data from Bgee)
* The third column, *“****id****”*, shows the Ensembl ID of the gene encoded by the rat protein (from column 2). Note that the ensemble ID (e.g. *ENSRNOG00000008839*) can be used in the Bgee search interface at <https://bgee.org/> for validating the results.

|  |  |  |
| --- | --- | --- |
| **protein** | **orthologous\_protein\_rat** | **id** |
| <http://purl.uniprot.org/uniprot/P37231> | <https://omabrowser.org/oma/info/RATNO15188> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000008839> |
| <http://purl.uniprot.org/uniprot/P08922> | <https://omabrowser.org/oma/info/RATNO12308> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000000406> |
| <http://purl.uniprot.org/uniprot/P68431> | <https://omabrowser.org/oma/info/RATNO09038> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000053155> |
| <http://purl.uniprot.org/uniprot/P68431> | <https://omabrowser.org/oma/info/RATNO09042> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000056281> |
| <http://purl.uniprot.org/uniprot/P68431> | <https://omabrowser.org/oma/info/RATNO11352> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000060366> |
| <http://purl.uniprot.org/uniprot/Q14956> | <https://omabrowser.org/oma/info/RATNO14717> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000008816> |
| <http://purl.uniprot.org/uniprot/P84243> | <https://omabrowser.org/oma/info/RATNO18582> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000032401> |
| <http://purl.uniprot.org/uniprot/P84243> | <https://omabrowser.org/oma/info/RATNO06508> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000003220> |
| <http://purl.uniprot.org/uniprot/O75140> | <https://omabrowser.org/oma/info/RATNO07117> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000018144> |
| <http://purl.uniprot.org/uniprot/Q12980> | <https://omabrowser.org/oma/info/RATNO03263> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000020541> |
| <http://purl.uniprot.org/uniprot/Q8WTW4> | <https://omabrowser.org/oma/info/RATNO20234> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000021660> |
| <http://purl.uniprot.org/uniprot/Q9BZH6> | <https://omabrowser.org/oma/info/RATNO02052> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000020430> |
| <http://purl.uniprot.org/uniprot/Q9HD26> | <https://omabrowser.org/oma/info/RATNO12311> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000000408> |
| <http://purl.uniprot.org/uniprot/Q9Y243> | <https://omabrowser.org/oma/info/RATNO06482> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000021497> |
| <http://purl.uniprot.org/uniprot/Q9UM73> | <https://omabrowser.org/oma/info/RATNO17047> | <http://rdf.ebi.ac.uk/resource/ensembl/ENSRNOG00000008683> |

## Example relational-to-RDF mappings

The following figure illustrates graphically an example of exposing relational data from Bgee (shown in the bottom half of the figure, the relational tables) in a virtual RDF graph (a fraction of the ontology and 2 instances shown in the upper half of the figure: the triple representing “Homo sapiens” and the triple describing the anatomical entity “brain”).

The orange rectangles show the “source” fragment (a simplified SQL statement) of a relational-to-RDF mapping, while the blue rectangles illustrate the “target” (resulting RDF triples). The full set of mappings employed with the Bgee relational data is available in our github repository [REF].



The Bgee data is now available in a SPARQL endpoint at <http://biosoda.expasy.org:8080/rdf4j-workbench/repositories/bgeelight/> (to be used in SERVICE blocks in federated queries) or through a graphical user interface at <http://biosoda.expasy.org:8080/rdf4j-workbench/repositories/bgeelight/query> . The complete list of federated queries is available in our template-based search interface at <http://biosoda.expasy.org:8080/build_biosodafrontend/> .