SPARQL Queries Part 2 – Pascal Brunner – brunnpa7 – eHealth

Task 1:

Print the names of the genes along with chromosomes

```
select ?geneName ?chromosom where {
?entry :gene ?gen .
?gen :name ?geneName .
?gen :chromosome ?chromosom.
}
```

Query time is 1.833[s] for 20531 rows

geneName	chromosom
"ETFBKMT"^^xsd:string	"12"^^xsd:string
"TPH2"^^xsd:string	"12"^^xsd:string
"MSRB3"^^xsd:string	"12"^^xsd:string
"CCDC60"^^xsd:string	"12"^^xsd:string
"FBXW8"^^xsd:string	"12"^^xsd:string
"TMEM132C"^^xsd:string	"12"^^xsd:string
"GALNT4"^^xsd:string	"12"^^xsd:string
"GPRC5A"^^xsd:string	"12"^^xsd:string
"NUP37"^^xsd:string	"12"^^xsd:string
"PLBD2"^^xsd:string	"12"^^xsd:string
"DDX55"^^xsd:string	"12"^^xsd:string
"RASSF8"^^xsd:string	"12"^^xsd:string
"NEDD1"^^xsd:string	"12"^^xsd:string
"CHPT1"^^xsd:string	"12"^^xsd:string
"FRS2"^^xsd:string	"12"^^xsd:string
"FAM216A"^^xsd:string	"12"^^xsd:string
"OSBP2"^^xsd:string	"22"^^xsd:string
"L3MBTL2"^^xsd:string	"22"^^xsd:string
"GUCD1"^^xsd:string	"22"^^xsd:string
"AIFM3"^^xsd:string	"22"^^xsd:string
"PANX2"^^xsd:string	"22"^^xsd:string

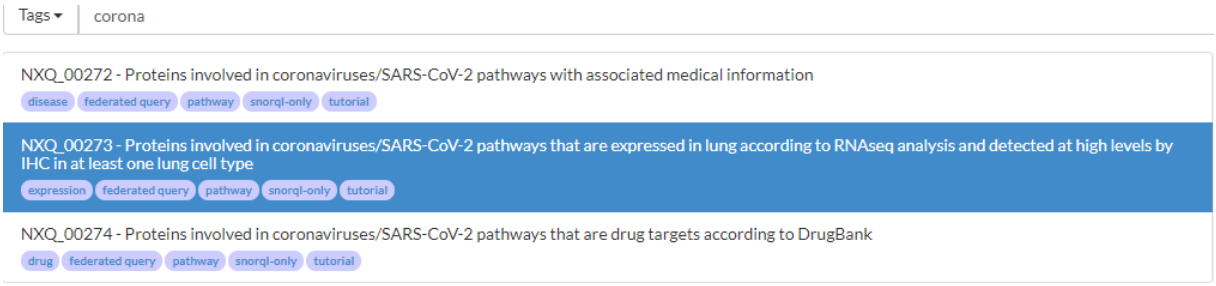
Name of proteins whose genes are on chromosome 11

```
select ?entry where {
?entry :gene / :chromosome "11"^^xsd:string .
}
```

Query time is 0.091[s] for 1332 rows

entry
entry:NX_A0AVK6 -- (neXtProt link)
entry:NX_A0JLT2 -- (neXtProt link)
entry:NX_A0PJW6 -- (neXtProt link)
entry:NX_A0A096LNP1 -- (neXtProt link)
entry:NX_A0A0U1RRN3 -- (neXtProt link)
entry:NX_A0A0U1RQE8 -- (neXtProt link)
entry:NX_A0A1B0GVV1 -- (neXtProt link)
entry:NX_A0A2R8Y4Y8 -- (neXtProt link)
entry:NX_A0A1B0GVM6 -- (neXtProt link)
entry:NX_A0A1B0GVG6 -- (neXtProt link)
entry:NX_A0A286YFK9 -- (neXtProt link)
entry:NX_A0A2R8YFL7 -- (neXtProt link)
entry:NX_A0A2R8YFM6 -- (neXtProt link)
entry:NX_A2RU14 -- (neXtProt link)
entry:NX_A5D8V6 -- (neXtProt link)
entry:NX_A5YM72 -- (neXtProt link)
entry:NX_A1A5B4 -- (neXtProt link)
entry:NX_A1L4Q6 -- (neXtProt link)
entry:NX_A6NEL3 -- (neXtProt link)
entry:NX_A6NDN8 -- (neXtProt link)
entry:NX_A6NDI0 -- (neXtProt link)

Task 2:




```
PREFIX wp: <http://vocabularies.wikipathways.org/wp#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX dcterms: <http://purl.org/dc/terms/>

select distinct ?entry str(?gen) ?pathwayname  group_concat(distinct str(?rnatissname), ',') as
?RNAseq_positive group_concat(distinct str(?ihctissname), ',') as ?IHC_high   where {
   service <http://sparql.wikipathways.org/sparql> {

      {?geneProduct a wp:Protein}
         union
      {?geneProduct a wp:GeneProduct}

      ?geneProduct rdfs:label ?gen .
      filter(!regex(?gen,"[ a-z-]")). # ensures official gene names for subsequent neXtprot matching

      ?geneProduct dcterms:isPartOf ?pathway .
      ?pathway a wp:Pathway .
      ?pathway wp:organism ?organism .
      filter(contains(str(?organism),"9606"))
      ?pathway dc:title ?pathwayname .
      filter(regex(?pathwayname,"sars-cov-2","i")|| regex(?pathwayname,"corona","i") ).
   }

   ?entry a :Entry .
   ?entry :gene / :name ?gen .
   ?entry :isoform ?iso.
   ?iso :expression ?exprna, ?expihc.
   ?exprna :term ?rnatiss .
   ?exprna :evidence ?evirna.
   ?exprna :quality :GOLD .
   ?evirna :expressionLevel :Positive.
   ?evirna :evidenceCode cv:ECO_0000295. # RNA-seq
   ?rnatiss :childOf cv:TS-0568. #lung
   ?rnatiss rdfs:label ?rnatissname.
   ?expihc :quality :GOLD .
   ?expihc :term ?ihctiss .
   ?expihc :evidence ?eviihc.
   ?eviihc :evidenceCode cv:ECO_0001055. # Immunohistochemistry (IHC)
   ?eviihc :expressionLevel :High.
   ?ihctiss :childOf ?rnatiss .
   ?ihctiss rdfs:label ?ihctissname.
}
order by ?entry
```

entry	callret-1	pathwayname	callret-3	medsource
entry:NX_O14920 -- (neXtProt link)	"IKBKB"	"Type I Interferon Induction and Signaling During SARS-CoV-2 Infection"@en	"Immunodeficiency 15B (IMD15B) [MIM:615592]: An autosomal recessive primary immunodeficiency disorder characterized by onset in infancy of life-threatening bacterial, fungal, and viral infections and failure to thrive. Laboratory studies show hypo- or agammaglobulinemia with relatively normal numbers of B and T-cells, and impaired differentiation and activation of immune cells. The disease is caused by mutations affecting the gene represented in this entry."	source:Uniprot
entry:NX_O14920 -- (neXtProt link)	"IKBKB"	"Host-pathogen interaction of human corona viruses - Interferon induction"@en	"Immunodeficiency 15B (IMD15B) [MIM:615592]: An autosomal recessive primary immunodeficiency disorder characterized by onset in infancy of life-threatening bacterial, fungal, and viral infections and failure to thrive. Laboratory studies show hypo- or agammaglobulinemia with relatively normal numbers of B and T-cells, and impaired differentiation and activation of immune cells. The disease is caused by mutations affecting the gene represented in this entry."	source:Uniprot
entry:NX_O14920 -- (neXtProt link)	"IKBKB"	"Type I Interferon Induction and Signaling During SARS-CoV-2 Infection"@en	"Immunodeficiency 15A (IMD15A) [MIM:618204]: An autosomal dominant primary immunodeficiency disorder characterized by lymphopenia, inflammation and immune activation of both CD4+ and CD8+ T cells. Patients suffer from recurrent respiratory tract infections, oral candidiasis, and otitis media. The disease is caused by mutations affecting the gene represented in this entry."	source:Uniprot
entry:NX_O14920 -- (neXtProt link)	"IKBKB"	"Host-pathogen interaction of human corona viruses - Interferon induction"@en	"Immunodeficiency 15A (IMD15A) [MIM:618204]: An autosomal dominant primary immunodeficiency disorder characterized by lymphopenia, inflammation and immune activation of both CD4+ and CD8+ T cells. Patients suffer from recurrent respiratory tract infections, oral candidiasis, and otitis media. The disease is caused by mutations affecting the gene represented in this entry."	source:Uniprot
entry:NX_O14920 -- (neXtProt link)	"IKBKB"	"Type I	"Severe combined immunodeficiency due to IKK2 deficiency"	source:Orphanet

Count of publication in 2019

```
SELECT DISTINCT Count(?public) WHERE {
?public :author ?author .
?public :year "2019"^^xsd:string .
}
```

callret-0
87192