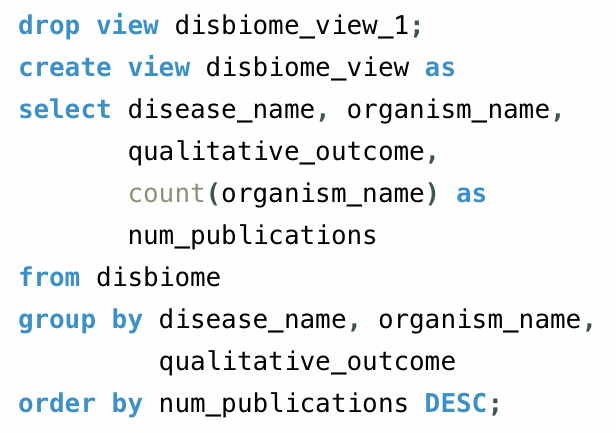
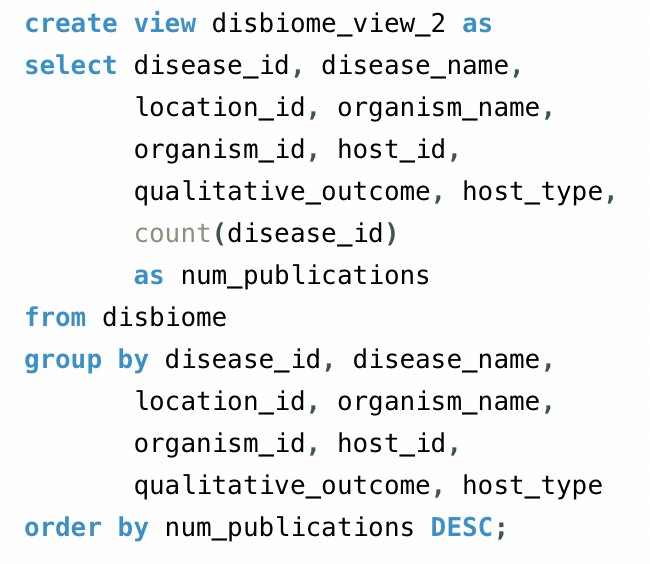
Project Code

**MySQL:**



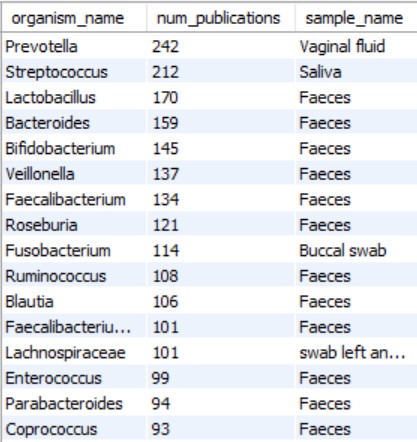
**Which diseases (or which bacteria) are the most heavily studied?**

select organism\_name, count(organism\_id) as num\_publications, sample\_name

from disbiome

group by organism\_name

order by num\_publications DESC;

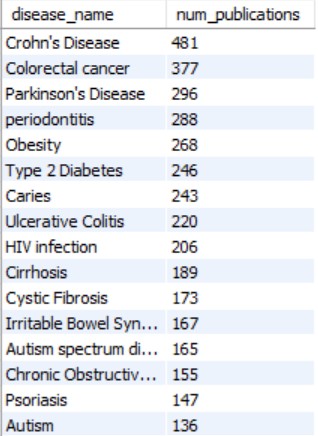


select disease\_name, count(disease\_id) as num\_publications, sample\_name

from disbiome

group by disease\_name

order by num\_publications DESC



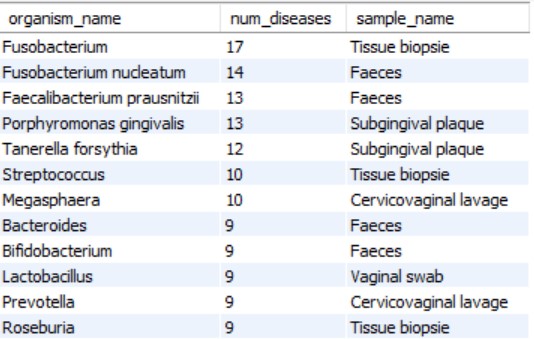
**Which bacterium is linked to the greatest number of diseases?**

select distinct organism\_name, count(disease\_id) as num\_diseases, sample\_name

from disbiome

group by organism\_id, disease\_id

order by count(organism\_id) desc;

****

**Neo4j:**

call apoc.load.jdbc("jdbc:mysql://localhost:3306/disbiome?serverTimezone=EST5EDT&user=ds4300user&password=ds4300password", "disbiome\_view\_1") yield row

merge (o:Organism {organism:row.organism\_name})

merge (d:Disease {disease:row.disease\_name})

merge (h:Host {host\_id:row.host\_id, host\_type:row.host\_type, control\_name:row.control\_name})

merge (o)-[:Assoc {publications:row.num\_publications, outcome:row.qualitative\_outcome}]->(d)

merge (o)-[:lives\_in]->(h)

return o,d,h

call apoc.load.jdbc("jdbc:mysql://localhost:3306/disbiome?serverTimezone=EST5EDT&user=disbiome&password=disbiome", "disbiome\_gad\_view") yield row

merge (o:Organism {organism\_id: row.organism\_id, organism:row.organism\_name,location:row.location\_id,sample\_name:row.sample\_name, organism\_ncbi\_id:row.organism\_ncbi\_id})

merge (d:Disease {disease:row.disease\_name, disease\_id:row.disease\_id, meddra\_level: row.meddra\_level, meddra\_id:row.meddra\_id,class:row.disease\_class})

merge (h:Host {host\_id:row.host\_id, host\_type:row.host\_type, control\_name:row.control\_name})

merge (o)-[:Assoc {publications:row.num\_publications, outcome:row.qualitative\_outcome}]->(d)

merge (o)-[:lives\_in]->(h)

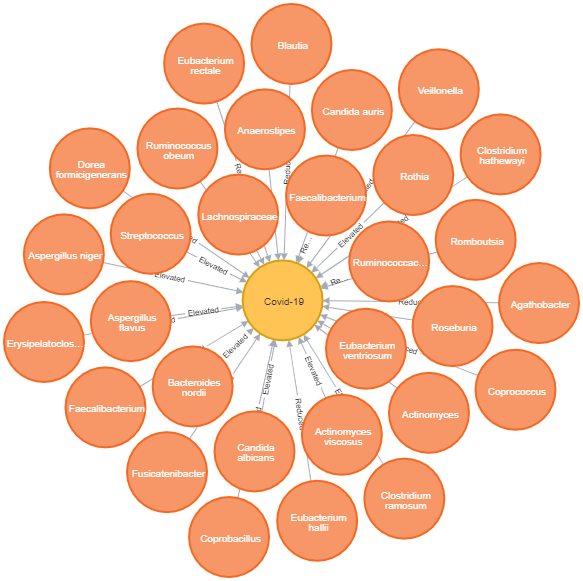
return o,d,h

Queries:

**Which bacteria are linked to Covid-19?**

match (o:Organism)-[a:Assoc]->(d:Disease {disease:"Covid-19"})

return o, a, d



**What diseases are associated with bacteria linked to Covid-19?**

call {

match (o:Organism)-[a:Assoc]->(d:Disease {disease:"Covid-19"})

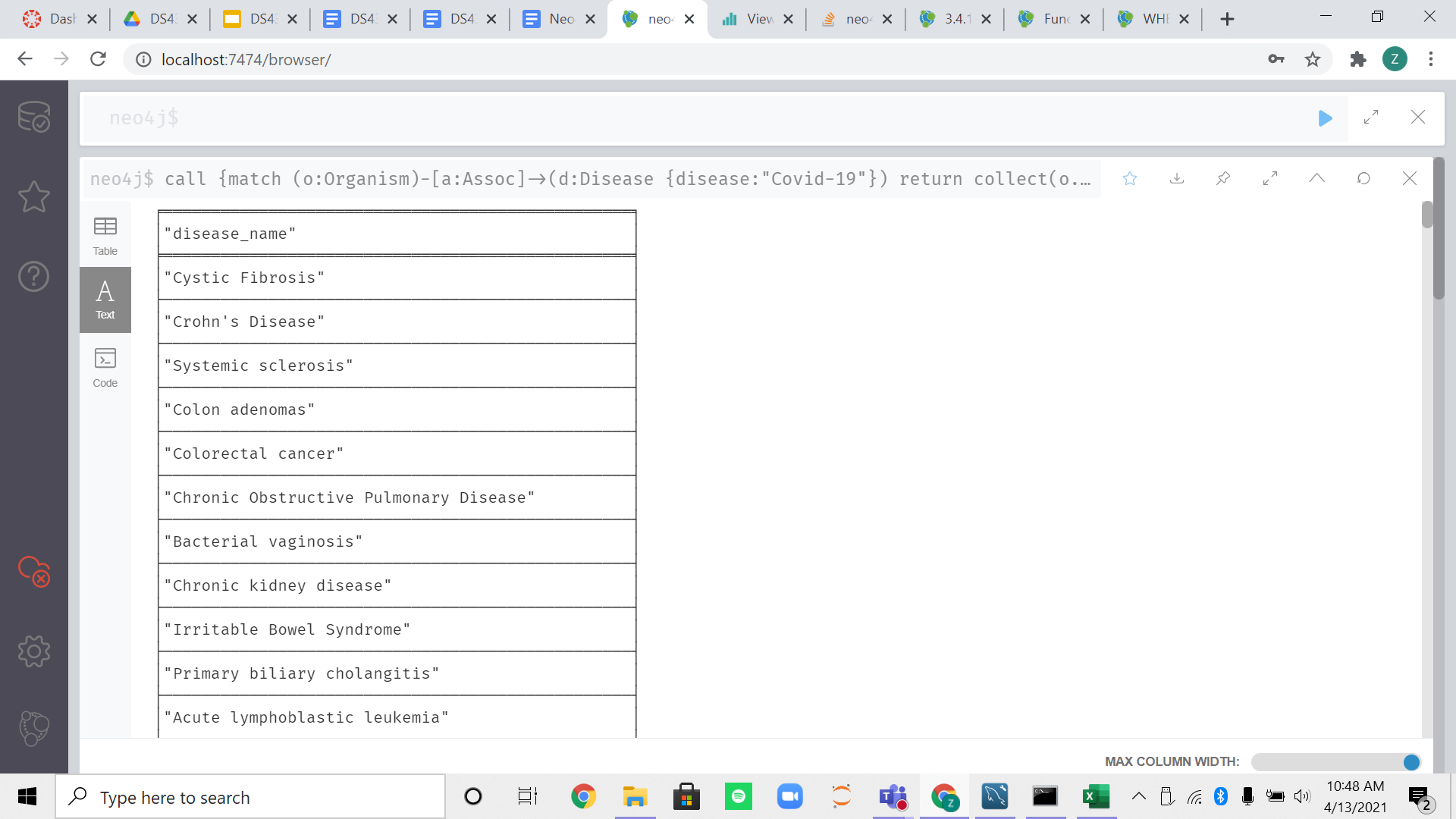
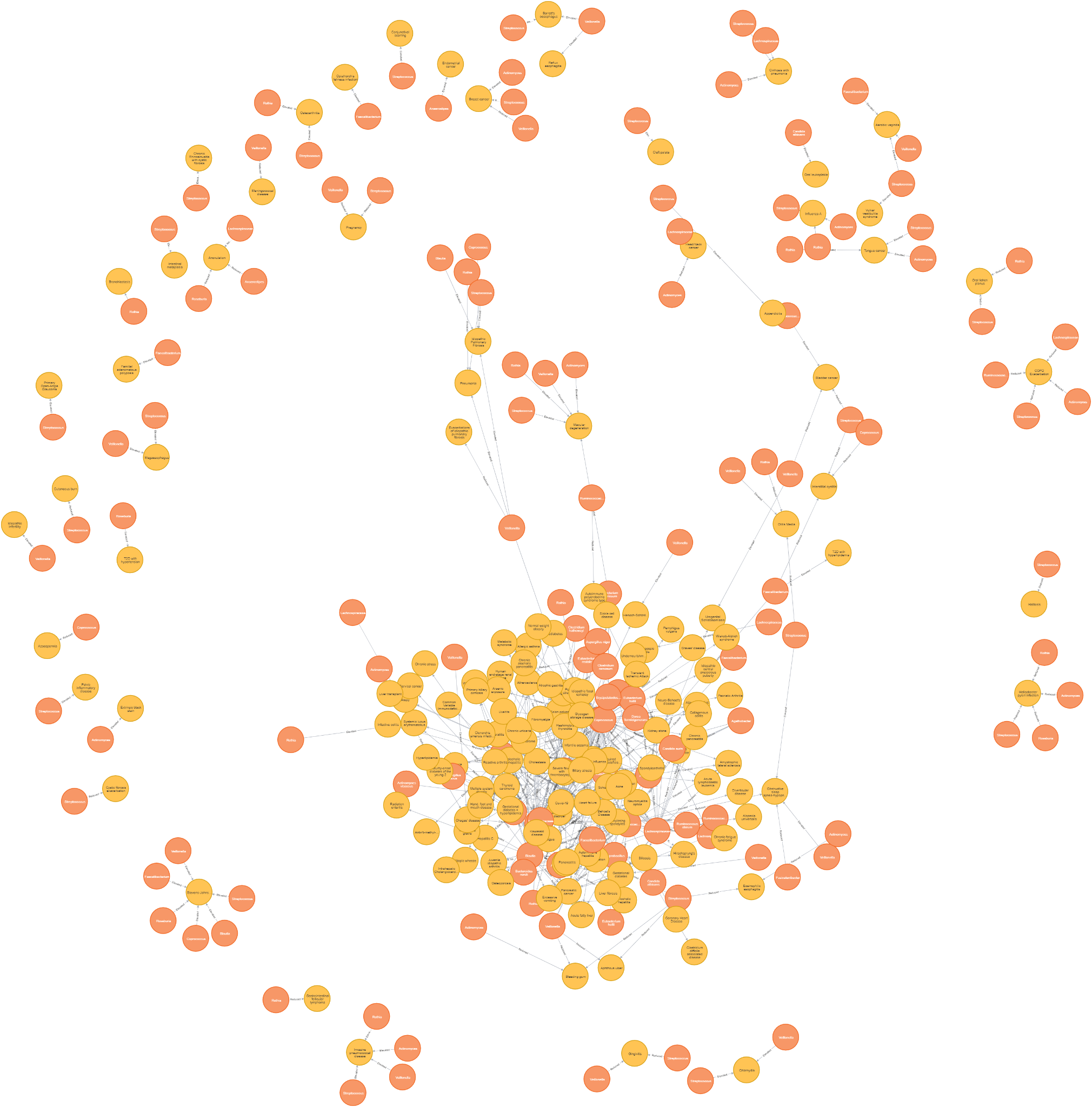
return collect(o.organism) as bacteria

}

match (o:Organism)-[a:Assoc]->(d:Disease)

where o.organism in bacteria

return o, a, d

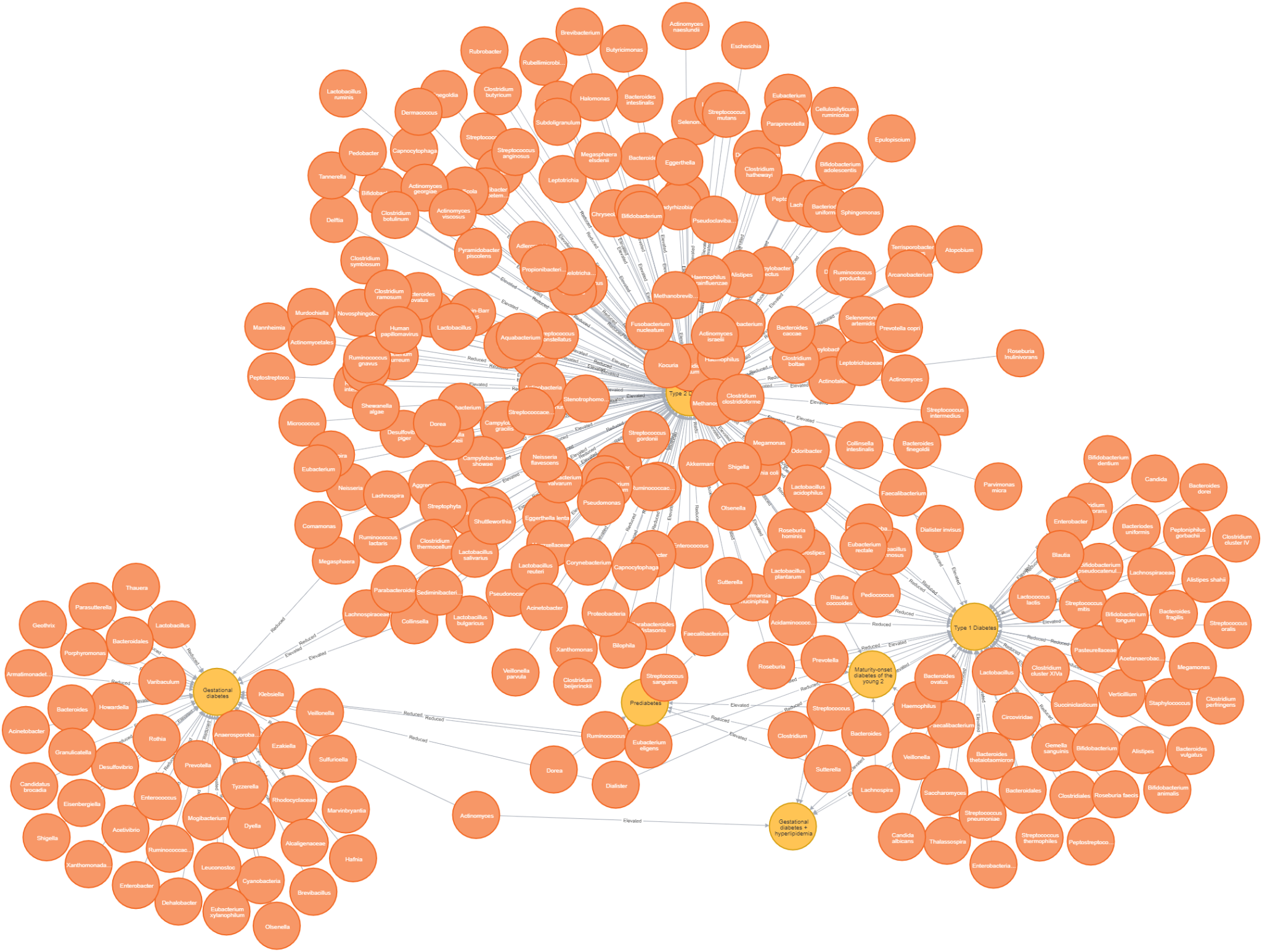
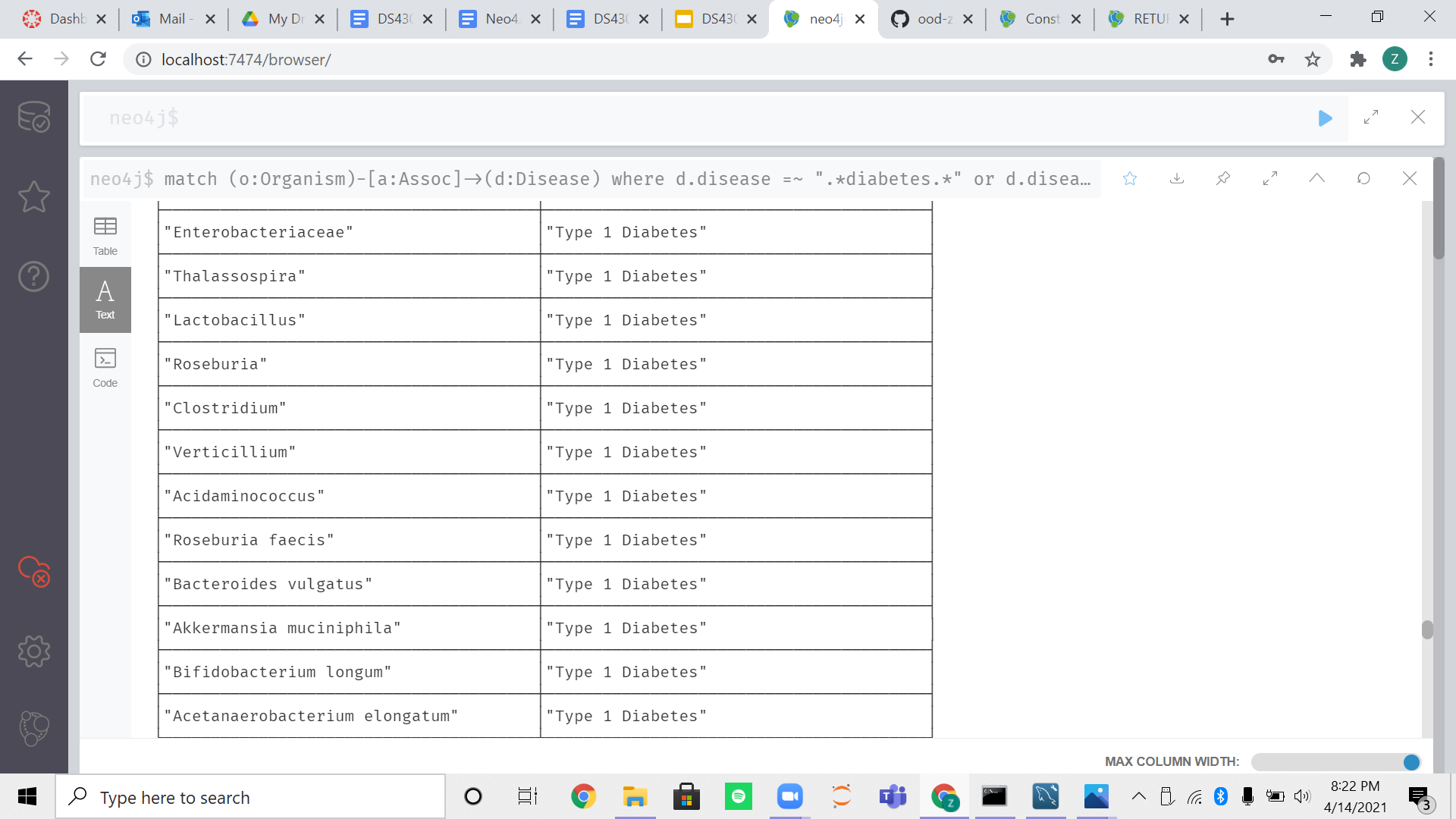
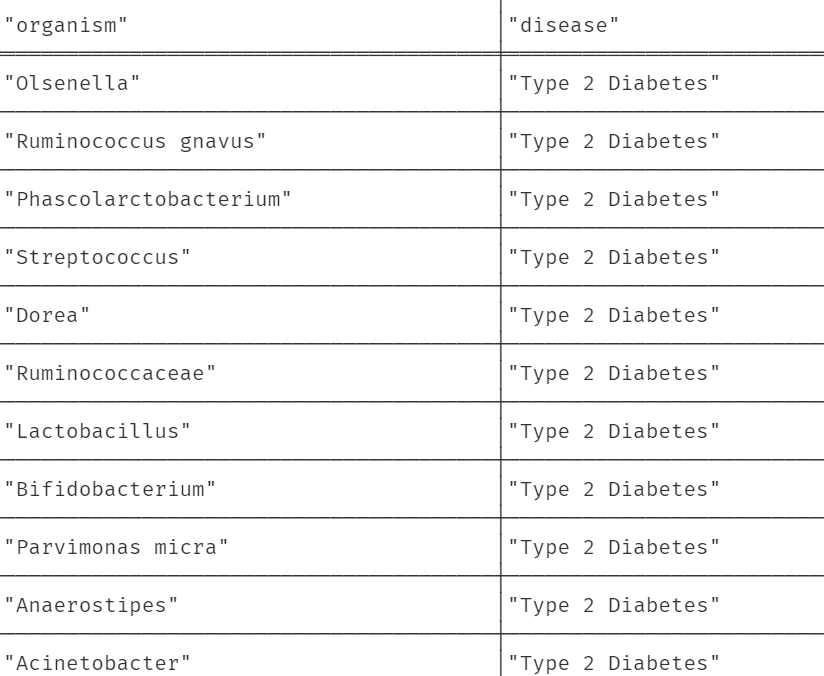


**Which bacteria are linked to diabetes (any type)?**

match (o:Organism)-[a:Assoc]->(d:Disease)

where d.disease =~ ".\*diabetes.\*" or d.disease =~".\*Diabetes.\*"

return o, a, d



**Bacteria in common among the diabetes types**

**\*nothing for gestational diabetes or prediabetes**

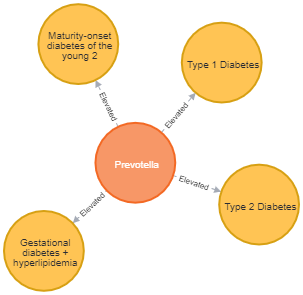
match (o:Organism)-[a]->(d:Disease{disease:"Type 2 Diabetes"})

match(o)-[a2]->(d2:Disease{disease:"Type 1 Diabetes"})

match(o)-[a3:Assoc]->(d3:Disease{disease:"Maturity-onset diabetes of the young 2"})

match(o)-[a4:Assoc]->(d4:Disease{disease:"Gestational diabetes + hyperlipidemia"})

return o, a, d, d2, a2, d3, a3, d4, a4



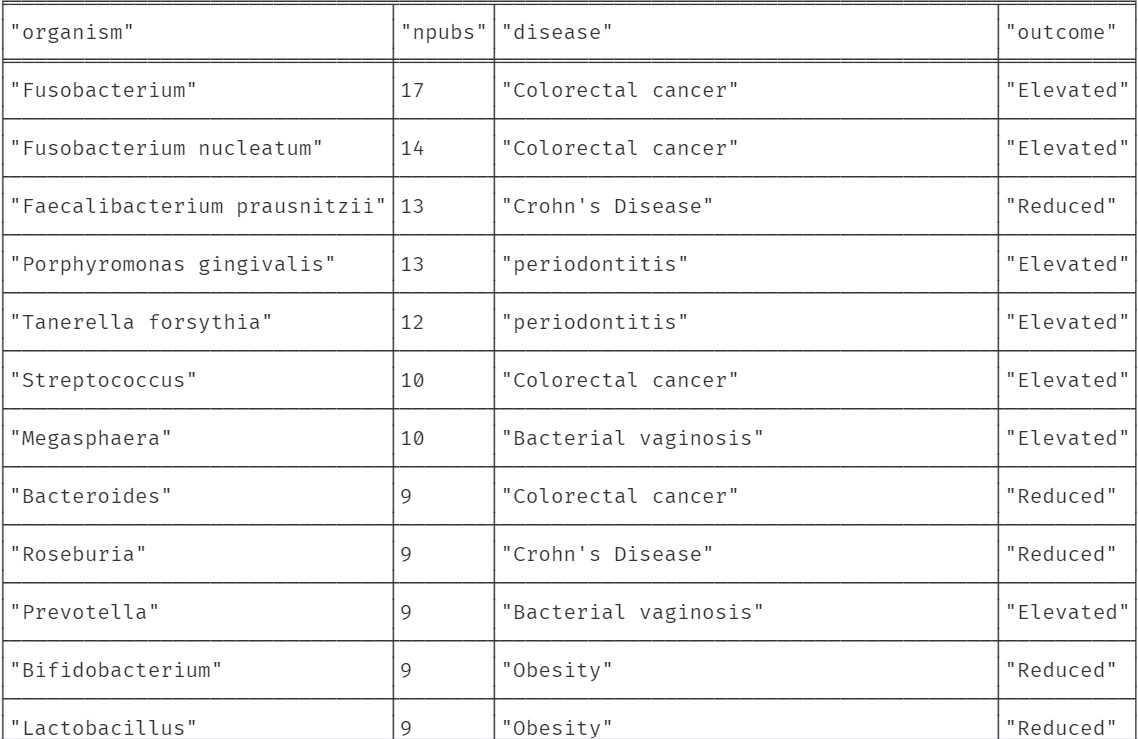
**Which diseases are most linked to specific gut bacteria?**

match (o:Organism)-[a:Assoc]->(d:Disease)

where a.publications > 1

return o.organism as organism, sum(a.publications) as npubs, d.disease as disease

order by npubs desc



**Which bacteria do neurodegenerative diseases have a common link to?**

**Alzheimer’s Disease, Parkinson’s Disease, Amyotrophic lateral sclerosis, Multiple Sclerosis**

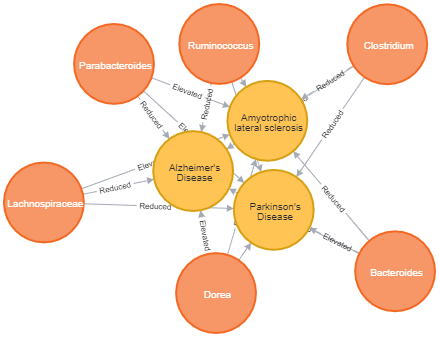
match (o:Organism)-[a:Assoc]->(d:Disease{disease:"Alzheimer's Disease"})

match(o)-[a2:Assoc]->(d2:Disease{disease:"Parkinson's Disease"})

match(o)-[a3:Assoc]->(d3:Disease{disease:"Amyotrophic lateral sclerosis"})

match(o)-[a4:Assoc]->(d4:Disease{disease:"Multiple Sclerosis"})

return o, a, d, d2, a2, a3, d3, a4, d4



**Which diseases are connected to the bacteria linked to the 3 neurodegenerative diseases listed above?**

call {

match (o:Organism)-[a:Assoc]->(d:Disease{disease:"Alzheimer's Disease"})

match(o)-[a2:Assoc]->(d2:Disease{disease:"Parkinson's Disease"})

match(o)-[a3:Assoc]->(d3:Disease{disease:"Amyotrophic lateral sclerosis"})

match(o)-[a4:Assoc]->(d4:Disease{disease:"Multiple Sclerosis"})

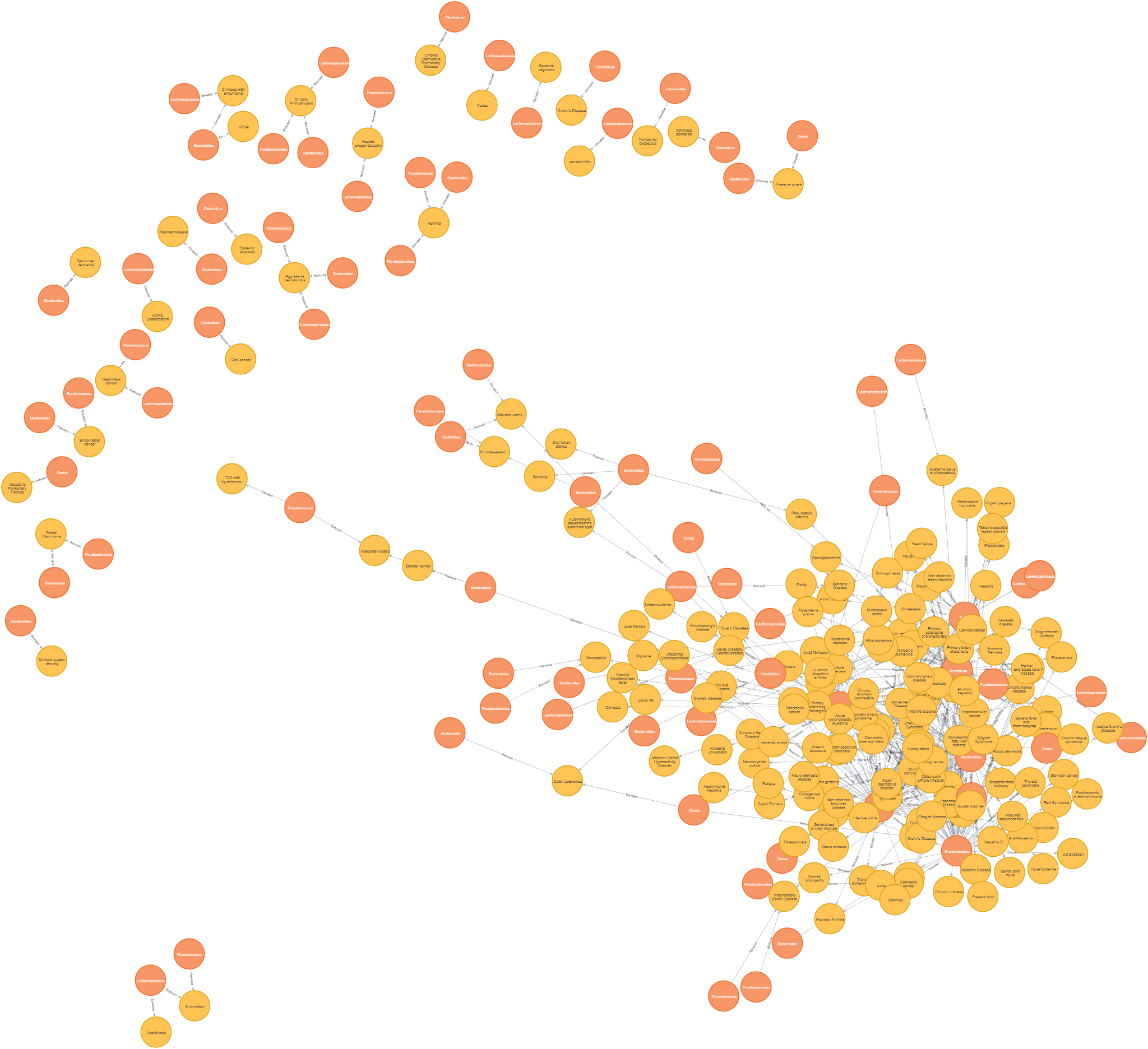
return collect(o.organism) as bacteria

}

match (o:Organism)-[a:Assoc]->(d:Disease)

where o.organism in bacteria

return o, a, d

****

**What diseases are associated with bacteria linked to cancer?**

call {

match (o:Organism)-[a:Assoc]->(d:Disease)

where d.disease =~ ".\*cancer.\*" or d.disease =~".Cancer.\*"

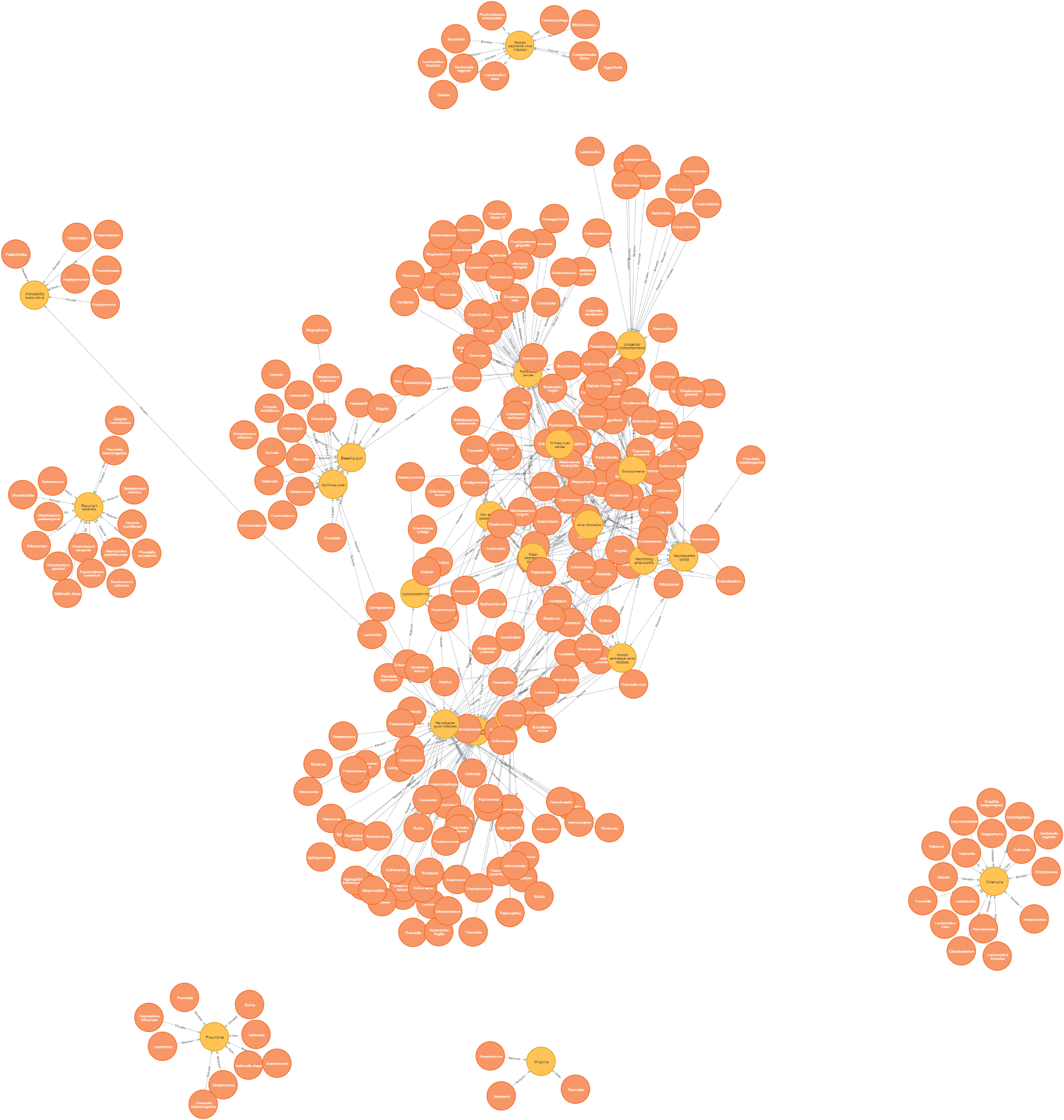
return collect(o.organism) as bacteria

}

match (o:Organism)-[a:Assoc]->(d:Disease)

where o.organism in bacteria

return o, a, d



**Which bacteria are associated with cardiovascular diseases?**

match (o:Organism)-[a:Assoc]->(d:Disease{class:"CARDIOVASCULAR"})

return o, a, d



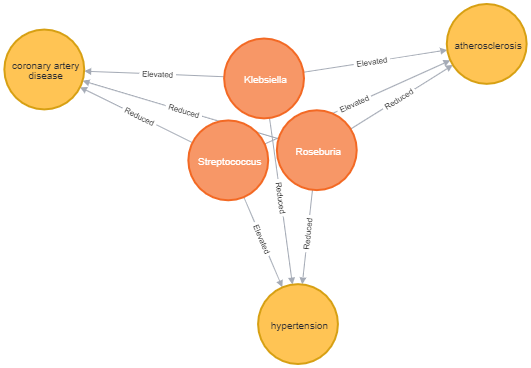
**Common bacteria among cardiovascular diseases?**

match (o:Organism)-[a:Assoc]->(d:Disease{disease:"hypertension"})

match(o)-[a2:Assoc]->(d2:Disease{disease:"atherosclerosis"})

match(o)-[a3:Assoc]->(d3:Disease{disease:"coronary artery disease"})

return o, a, d, d2, a2, d3, a3



match (o:Organism)-[a:Assoc]->(d:Disease{disease:"hypertension"})

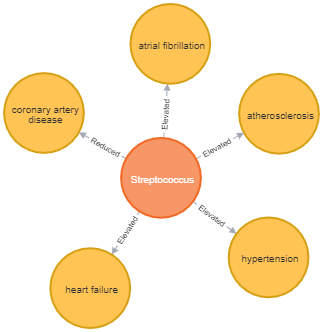
match(o)-[a2:Assoc]->(d2:Disease{disease:"atherosclerosis"})

match(o)-[a3:Assoc]->(d3:Disease{disease:"coronary artery disease"})

match(o)-[a4:Assoc]->(d4:Disease{disease:"heart failure"})

match(o)-[a5:Assoc]->(d5:Disease{disease:"atrial fibrillation"})

return o, a, d, d2, a2, d3, a3, a4, d4, a5, d5



**Which bacteria are most highly linked to mental health disorders?**

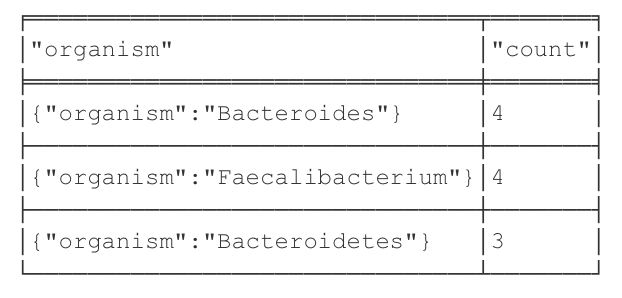
match (o:Organism)-[a:Assoc]->(d:Disease)

where d.disease =~ ".\*anxiety.\*" or d.disease =~".\*depressive.\*"

or d.disease =~ ".\*Bipolar.\*"

return o, count(a)

order by count(a) DESC



**What diseases are associated with the organisms most highly linked to mental health disorders?**

call {match (o:Organism)-[a:Assoc]->(d:Disease)

where d.disease =~ ".\*anxiety.\*" or d.disease =~".\*depressive.\*"

or d.disease =~ ".\*Bipolar.\*"

with o, count(a) as count

order by count DESC limit 3

return collect(o.organism) as bacteria}

match (o:Organism)-[a:Assoc]->(d:Disease)

where o.organism in bacteria and a.publications > 1

return o,a,d

limit 5

