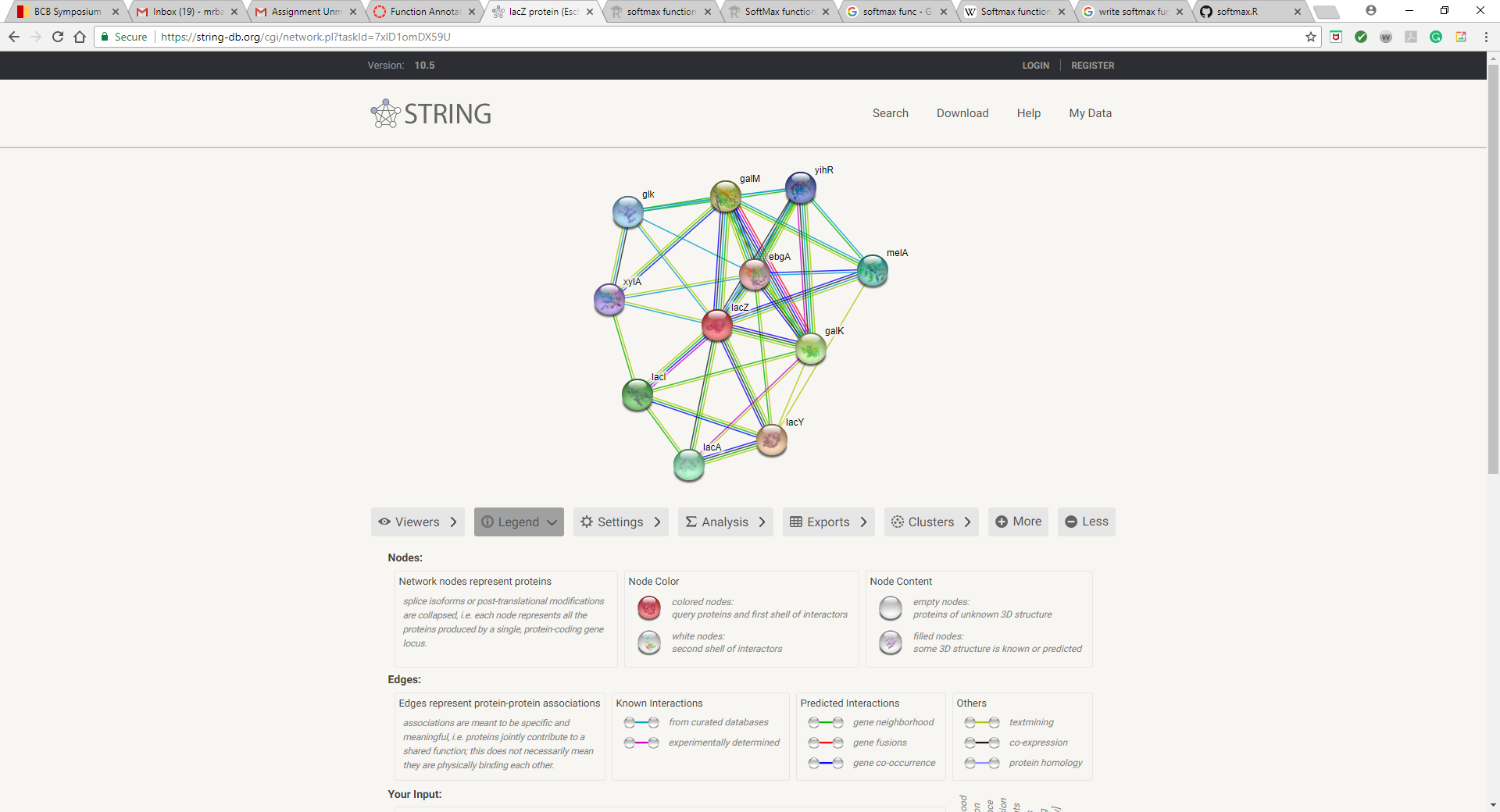
BCB 570 Assisgnment: Functional Annotation

-Urminder Singh

Sol1.



Sol2.

The gene lacZ encodes for β-galactosidase which is an enzyme which cleaves disaccharide lactose into glucose and galactose. lacZ is part of the lac operon which consists of other two genes as well. These genes are

lacY encodes for Beta-galactoside permease (LacY), a transmembrane symporter that pumps β-galactosides including lactose into the cell.

lacA encodes β-galactoside transacetylase (LacA), an enzyme that transfers an acetyl group from acetyl-CoA to β-galactosides.

As lacY and lacA are in the same operon as lacZ they are highly associated with each other.

**According to STRING**, lacZ seems to be associated with the following ten genes:

lacY: Lactose permease; Responsible for transport of beta-galactosides into the cell,

galM Aldose 1-epimerase; Mutarotase converts alpha-aldose to the beta-anomer.

galK Galactokinase; Catalyzes the transfer of the gamma-phosphate of ATP to D-galactose to form alpha-D-galactose-1-phosphate

lacl DNA-binding transcriptional repressor; Repressor of the lactose operon.

lacA Thiogalactoside acetyltransferase; May assist cellular detoxification by acetylating non- metabolizable pyranosides,

melA Alpha-galactosidase, NAD(P)-binding (451 aa)

glk Glucokinase; Not highly important in E. coli as glucose is transported into the cell by the PTS system

yihR Predicted aldose-1-epimerase (308 aa)

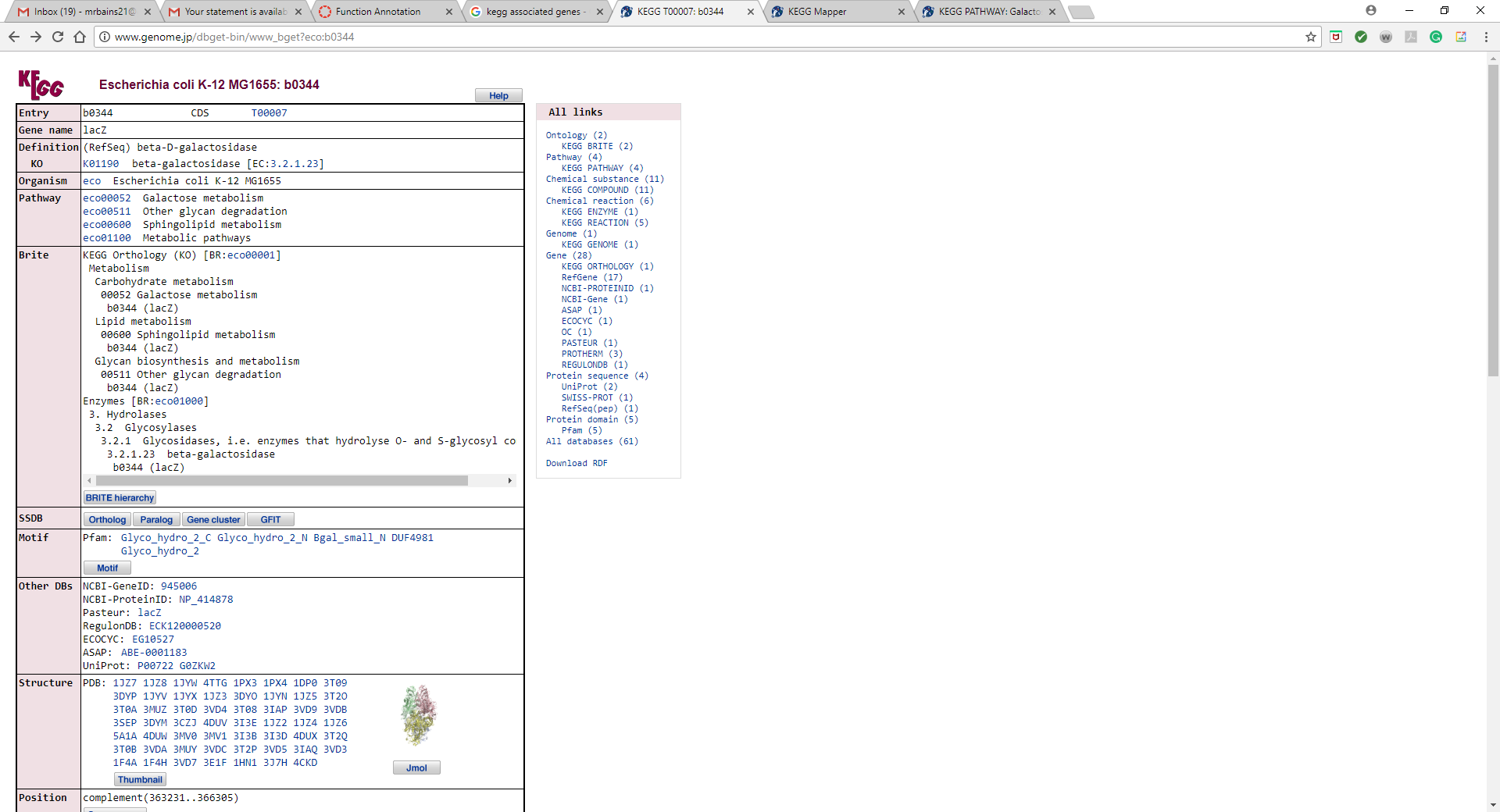
xylA D-xylose isomerase (440 aa)

It looks like STRING associates more than two genes to lacZ. This is because other genes are associated with lacZ via other pathways. E.g. galM and galK are associated with lacZ as these are, broadly, involved in galactose metabolism. Other criterion for association by STRING is by using text mining, co-expression, protein homology, gene fusions, experimentally determined etc. Therefore, using all these different criterion STRING is able to associate more genes to lacZ.

Sol3.

According to KEGG lacZ appears in following pathways:

|  |  |
| --- | --- |
| [eco00052](http://www.genome.jp/kegg-bin/show_pathway?eco00052+b0344) | Galactose metabolism |
| [eco00511](http://www.genome.jp/kegg-bin/show_pathway?eco00511+b0344) | Other glycan degradation |
| [eco00600](http://www.genome.jp/kegg-bin/show_pathway?eco00600+b0344) | Sphingolipid metabolism |
| [eco01100](http://www.genome.jp/kegg-bin/show_pathway?eco01100+b0344) | Metabolic pathways |



After searching the associated genes in these pathways, I found that there are many other genes present than listed by STRING in the lacZ associated pathways except lacA and lacY were not present in the pathways. This may be because lacA and lacY may have an indirect role in the lacZ associated pathways.

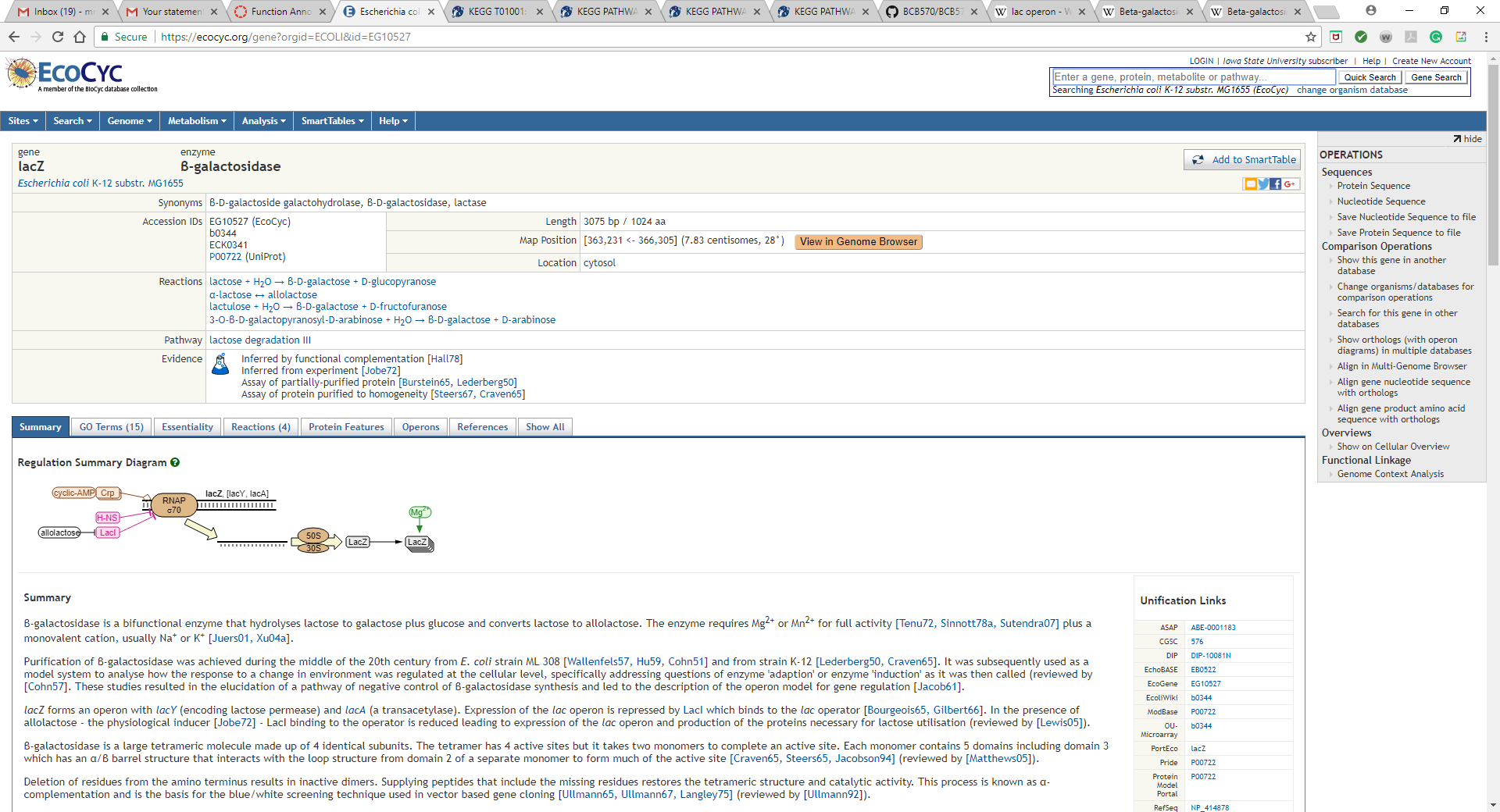
Sol 4

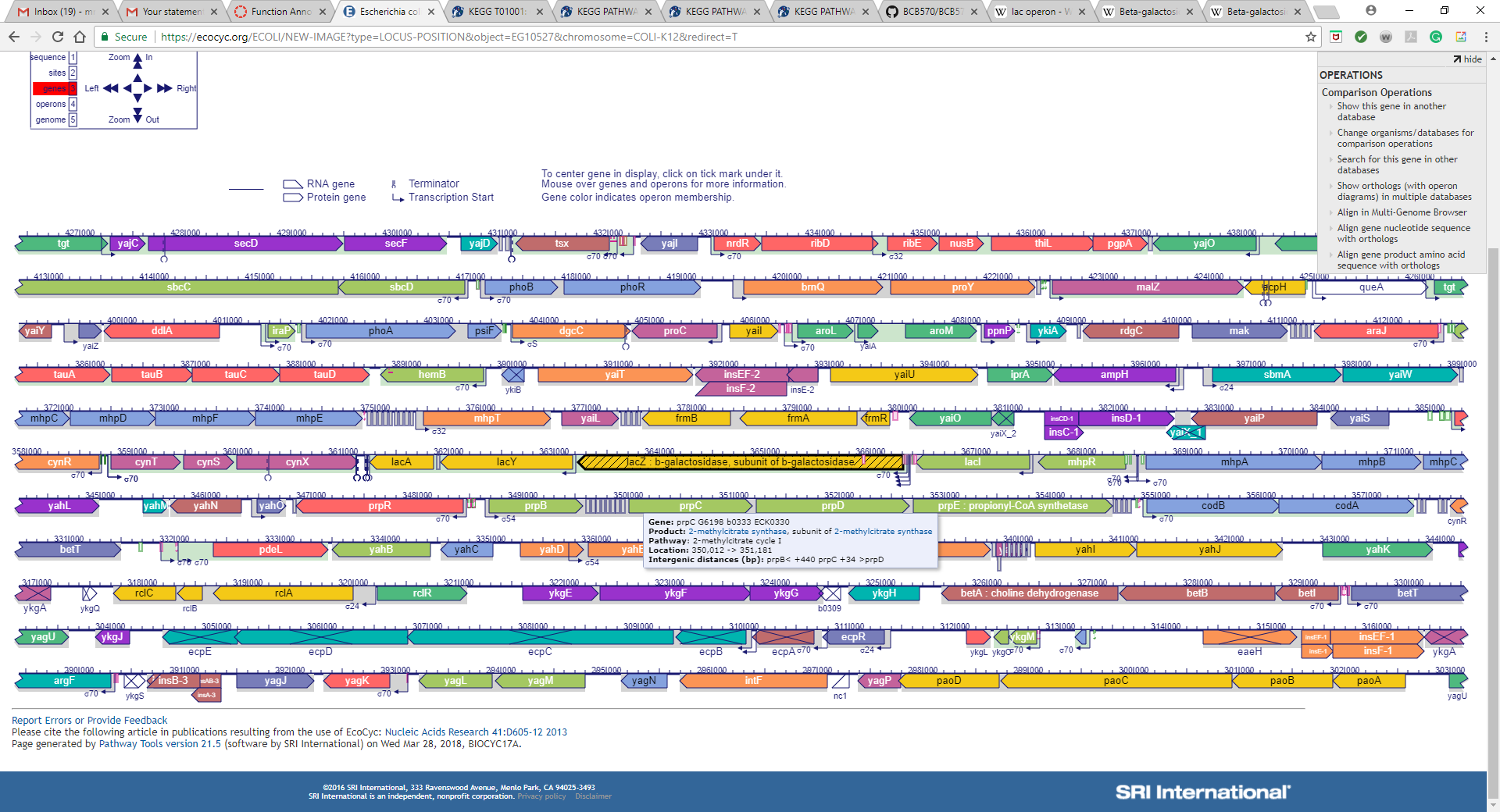
I repeated the above experiment for lacY and lacA.

I found that lacA and lacY are not associated with any pathway according to kegg. We can try to answer why, based on the function of these genes. lacY is involved in coding for a protein galactoside permease which controls flow of lactose molecules into the cell. Thus, it is not directly involved in breaking lactose into glucose and hence is not associated with the lacZ pathway.

lacA codes for a protein β-galactoside transacetylase which is an enzyme that transfers an acetyl group from acetyl-CoA to galactosides but not to the beta-galactosidase which is coded by lacZ. Further, the precise role of lacA’s product is not clearly understood in the lac operon. This shows that lacA is not directly involved in any pathway and thus is absent from lacZ’s pathways.

Sol 5.





Sol6.

I repeated 3 to 5 with GalK

According to kegg, GalK is involved in following metabolic pathways

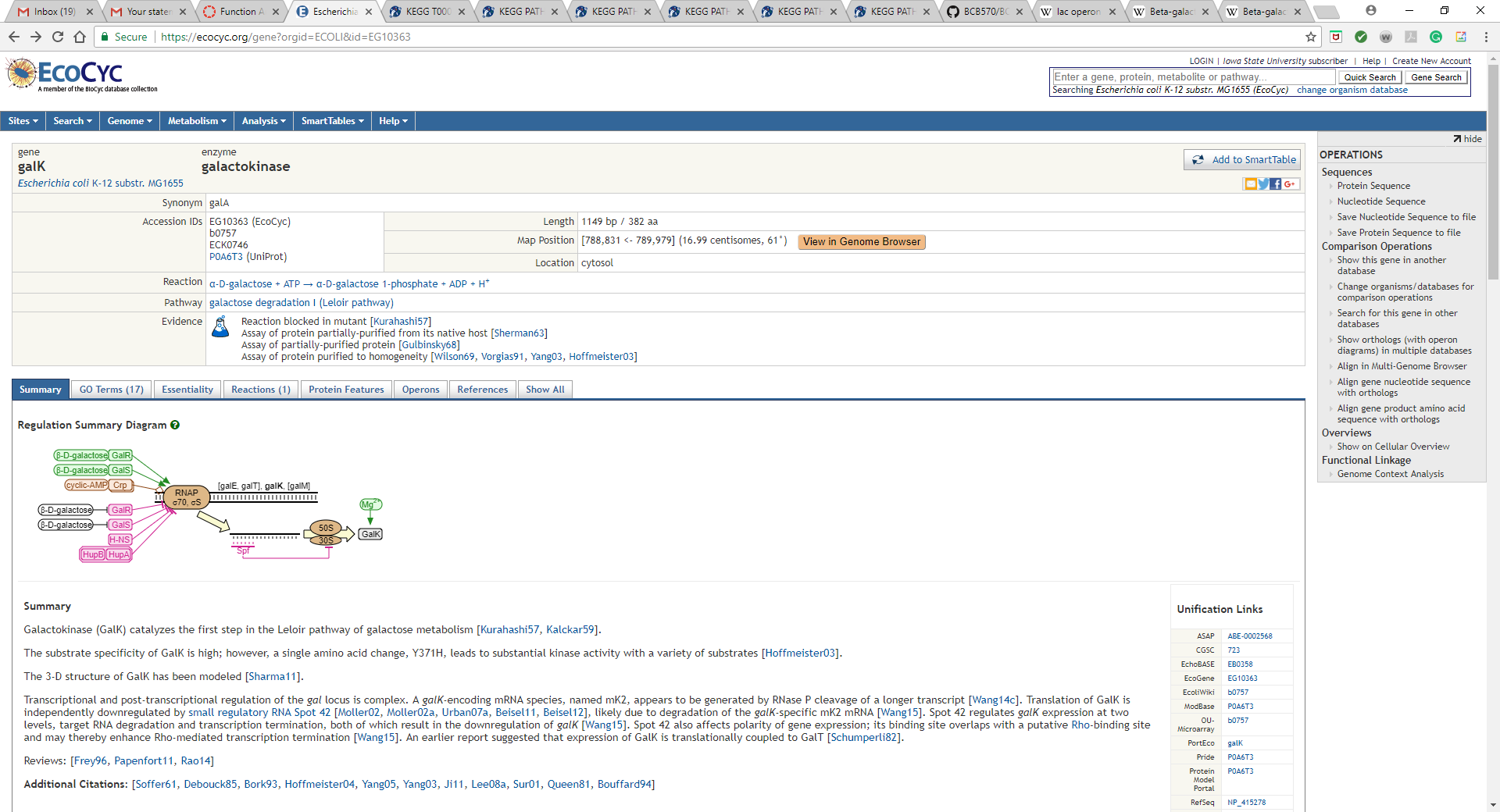
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | [eco00052](http://www.genome.jp/kegg-bin/show_pathway?eco00052+b0757) | Galactose metabolism | | [eco00520](http://www.genome.jp/kegg-bin/show_pathway?eco00520+b0757) | Amino sugar and nucleotide sugar metabolism | | [eco01100](http://www.genome.jp/kegg-bin/show_pathway?eco01100+b0757) | Metabolic pathways | |

Genes involved in Galactose metabolism are: GalM, lacZ, galT, galU, melA, yihR etc

Genes involved in Amino sugar and nucleotide sugar metabolism are: galT, glf, galT, arnC, arnD etc.

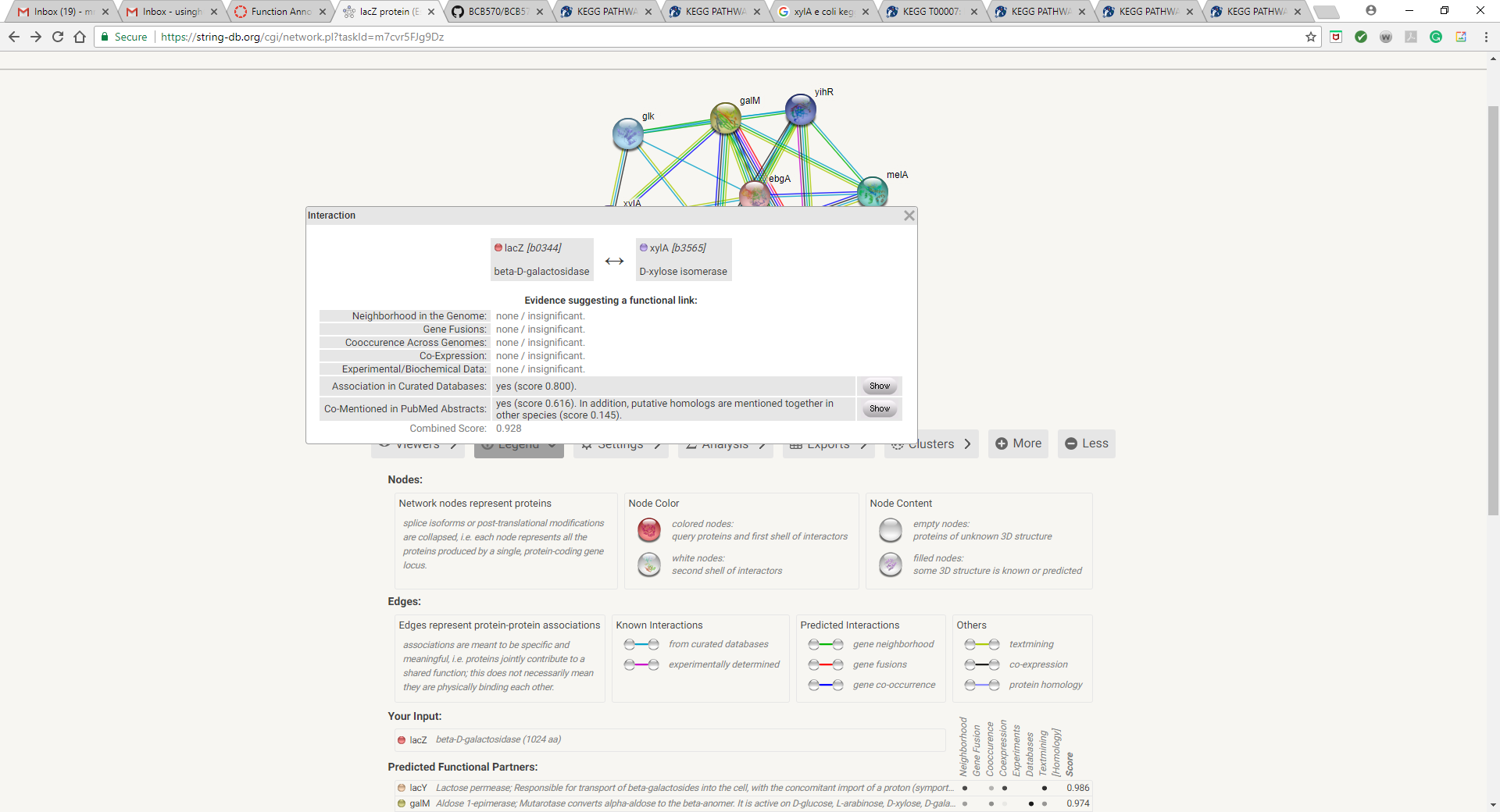
It is interesting that galM, galK and galT all are in the gal operon. Again, this has other genes associate which are not present in STRING. And genes which are present only in STRING. This is because STRING has different criterion to associate genes one of which is looking ate the biological pathways but it only reports strong associations.

Then, I found galK in ecocyc too



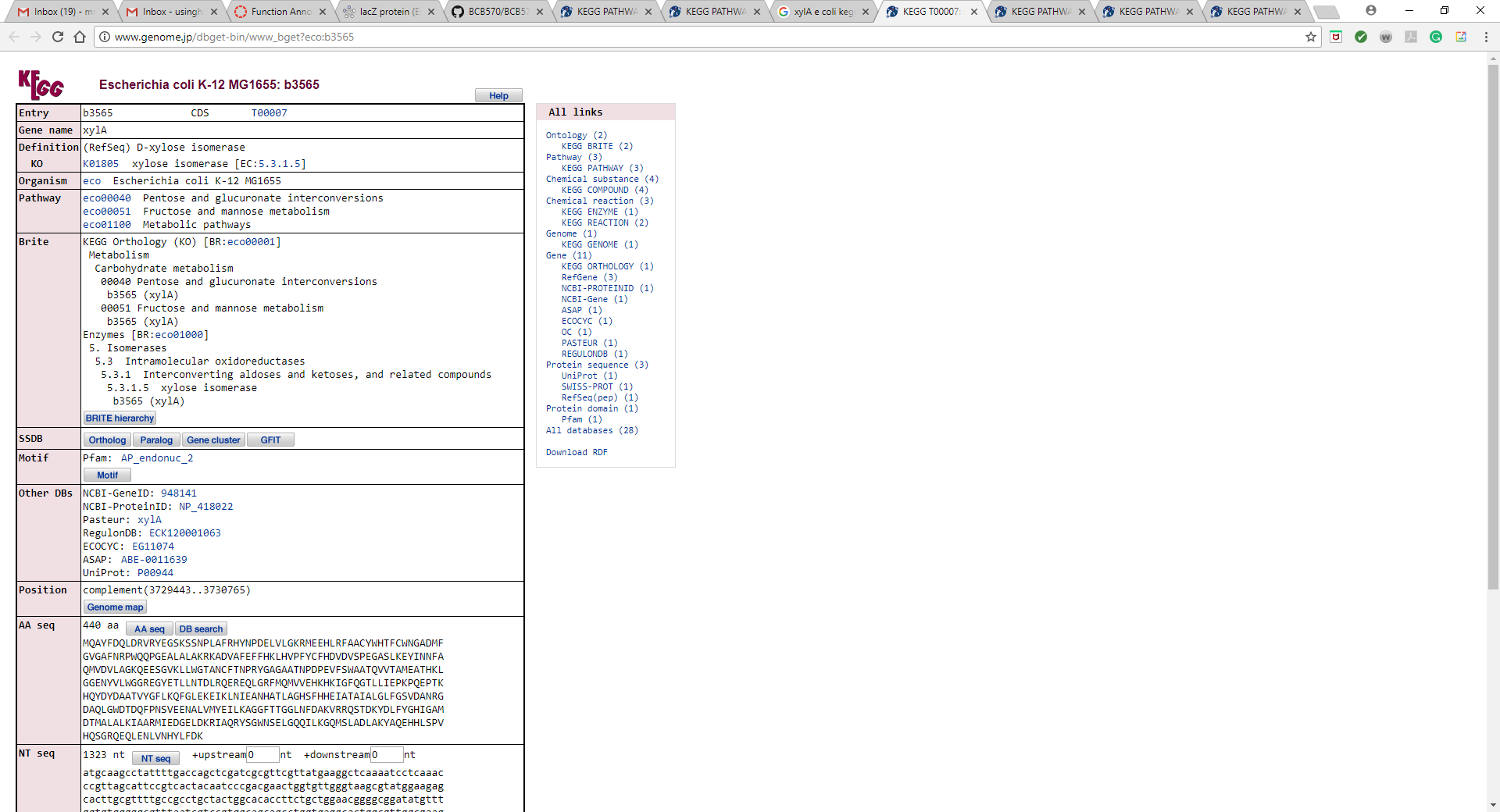
Sol 7.

From STRING we can see that xylA is associated to lacZ by two edges. These two edges are represent the evidence of association which are “curated database” and “from text mining”.

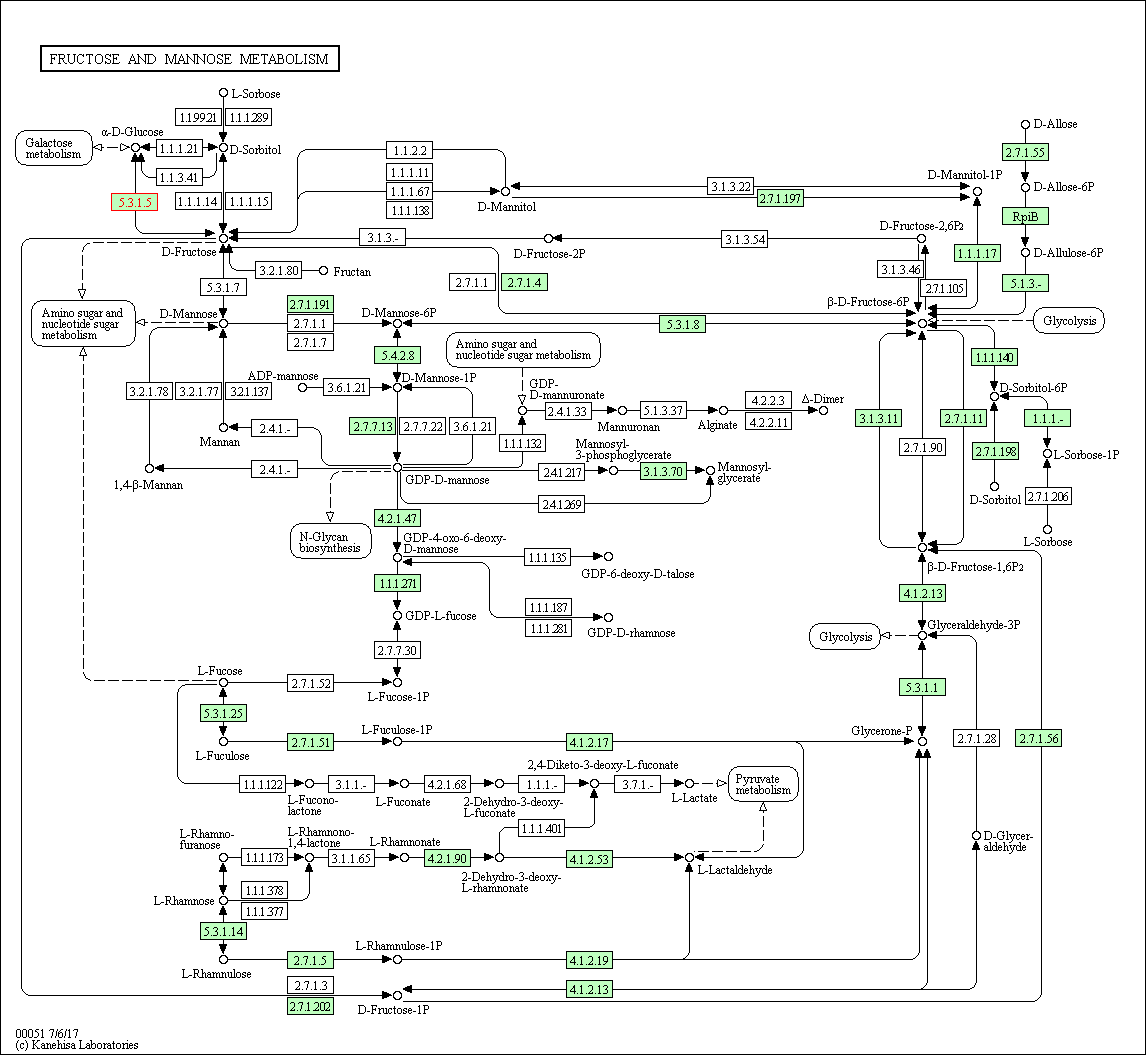


To verify this, I first found xylA in Kegg and found its present in the following pathways

|  |  |
| --- | --- |
| [eco00040](http://www.genome.jp/kegg-bin/show_pathway?eco00040+b3565) | Pentose and glucuronate interconversions |
| [eco00051](http://www.genome.jp/kegg-bin/show_pathway?eco00051+b3565) | Fructose and mannose metabolism |
| [eco01100](http://www.genome.jp/kegg-bin/show_pathway?eco01100+b3565) | Metabolic pathways |



In the Fructose and mannose metabolism pathway, eco00051, I found that the Galactose metabolism pathway is contained in it, which is associated with lacZ. This is the reason STRING shows the xylA is associated with lacZ via a “curated from database edge”.



Next, From STRING I found that there are 7 publications in PubMed which co-mentions lacZ and xylA. Thus, using text mining STRING associates lacZ and xylA.