UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

Kvik: Interactive exploration of biological pathways

Bjørn Fjukstad
Ph. D. Student / Department of Computer Science
bjorn@cs.uit.no

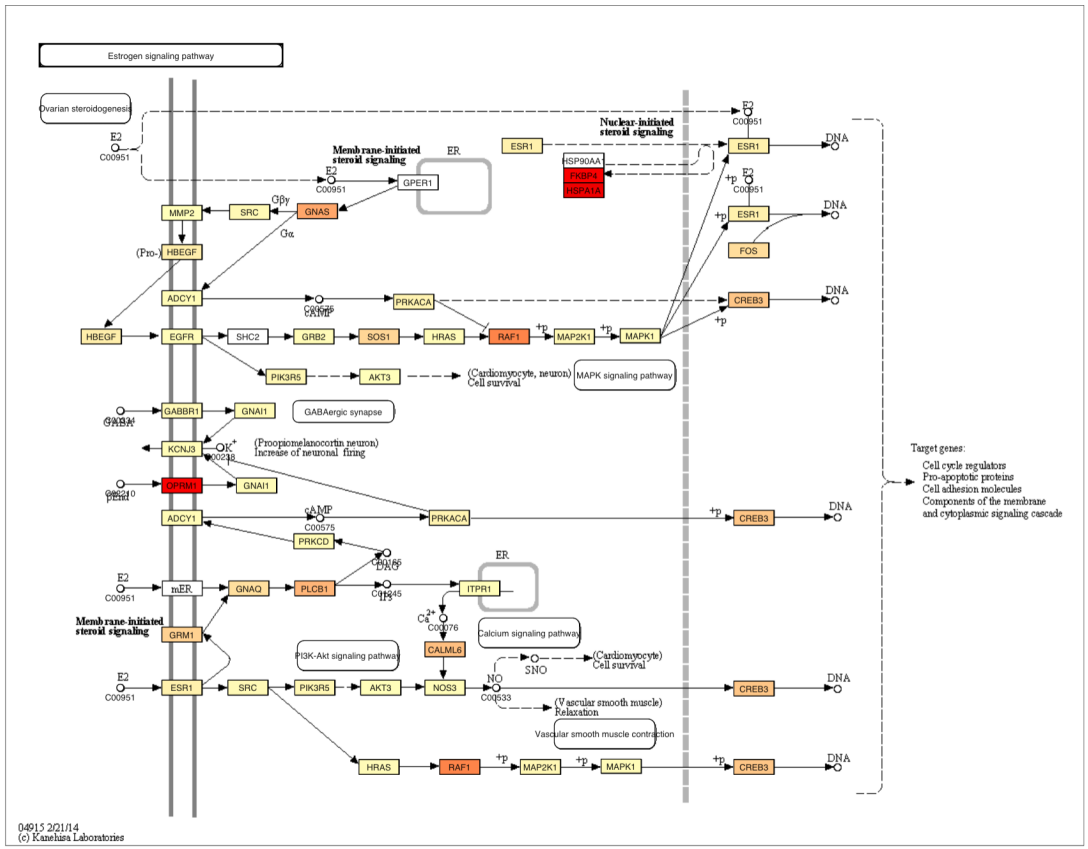
20. October 2014

UNIVERSITY OF TROMSØ

Pathway Visualizer

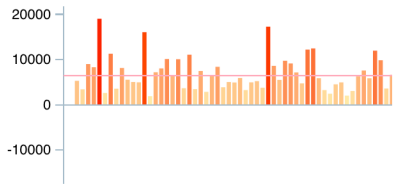
localhost:8000/browser/pathwaySelect=hsa04915

Kvik Home About Browser



PLCB1, EIEE12, PI-PLC, PLC-154, PLC-I, PLC154, PLCB1A, PLCB1B

Gene expression values for entire dataset



Mean: 6480.310

Standard deviation: 3437.448

Variance:11816047.351

Similar pathways

- Inositol phosphate metabolism
- Metabolic pathways
- Rap1 signaling pathway
- Calcium signaling pathway
- cGMP-PKG signaling pathway
- Chemokine signaling pathway
- Phosphatidylinositol signaling system
- Adrenergic signaling in cardiomyocytes
- Vascular smooth muscle contraction

Logarithmic Absolute

☑ Filtering disabled ☑ Smoker ☑ Hormone Therapy

me

- Finished my master's in June and became a Ph. D. student in July
- Collaborating with epidemiology researchers at the Dept. of Community Medicine
- I'm not a biologist. Sorry 😅

Norwegian Women and Cancer
NOWAC

- Identify relationships between lifestyle and risk of cancer
- Ongoing study since the early nineties
- Lots of data: Over 60 000 blood samples and 800 biopsies

Norwegian Women and Cancer
NOWAC

- Identify relationships between lifestyle and risk of cancer
- Ongoing study since the early nineties
- Lots of data: Over 60 000 blood samples and 800 biopsies

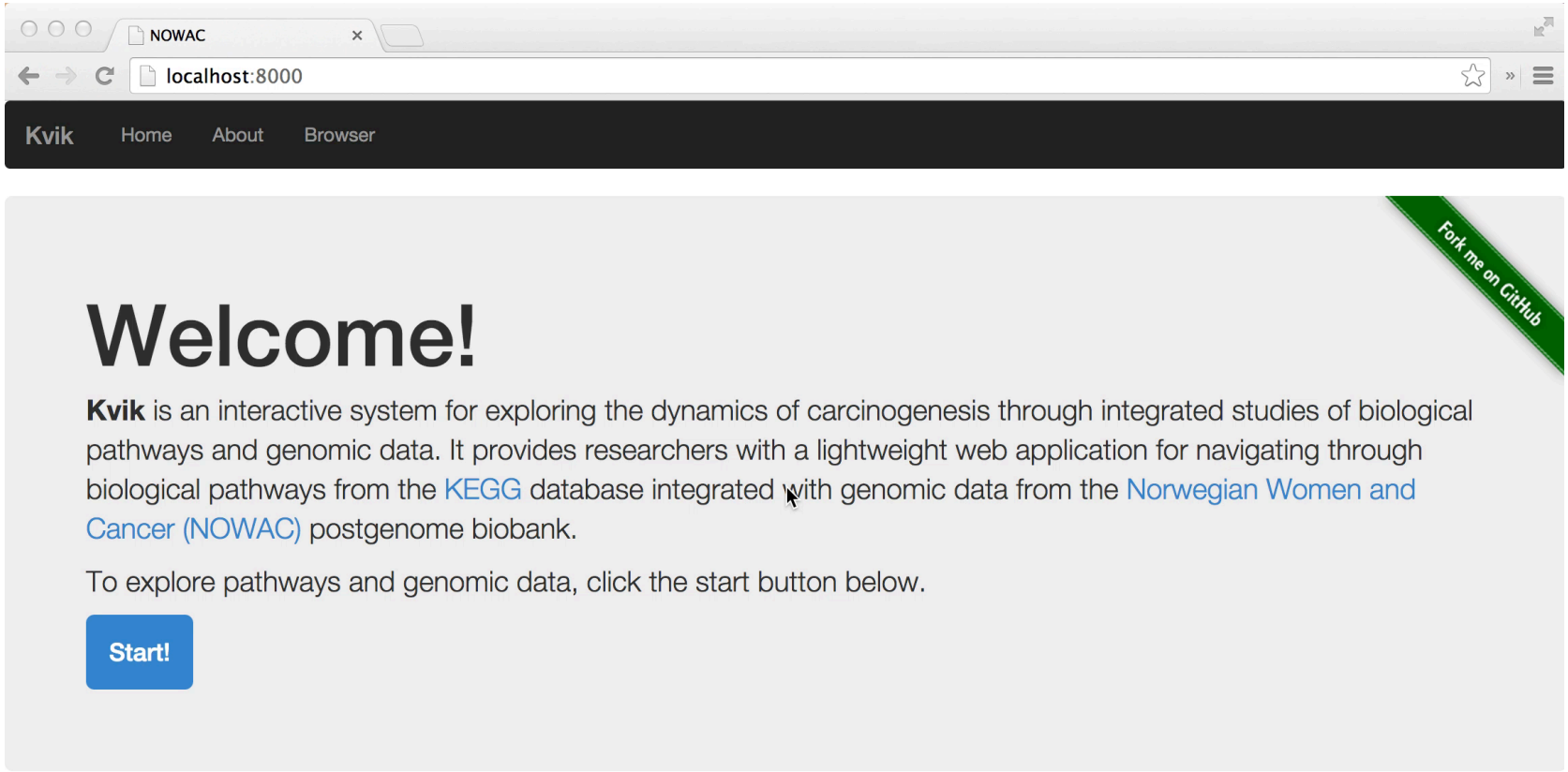
Background and Motivation

- Need a fast and interactive way of looking at many biological processes
- Manually collecting information from different databases is tedious and boring
- Researchers should focus on what they do best

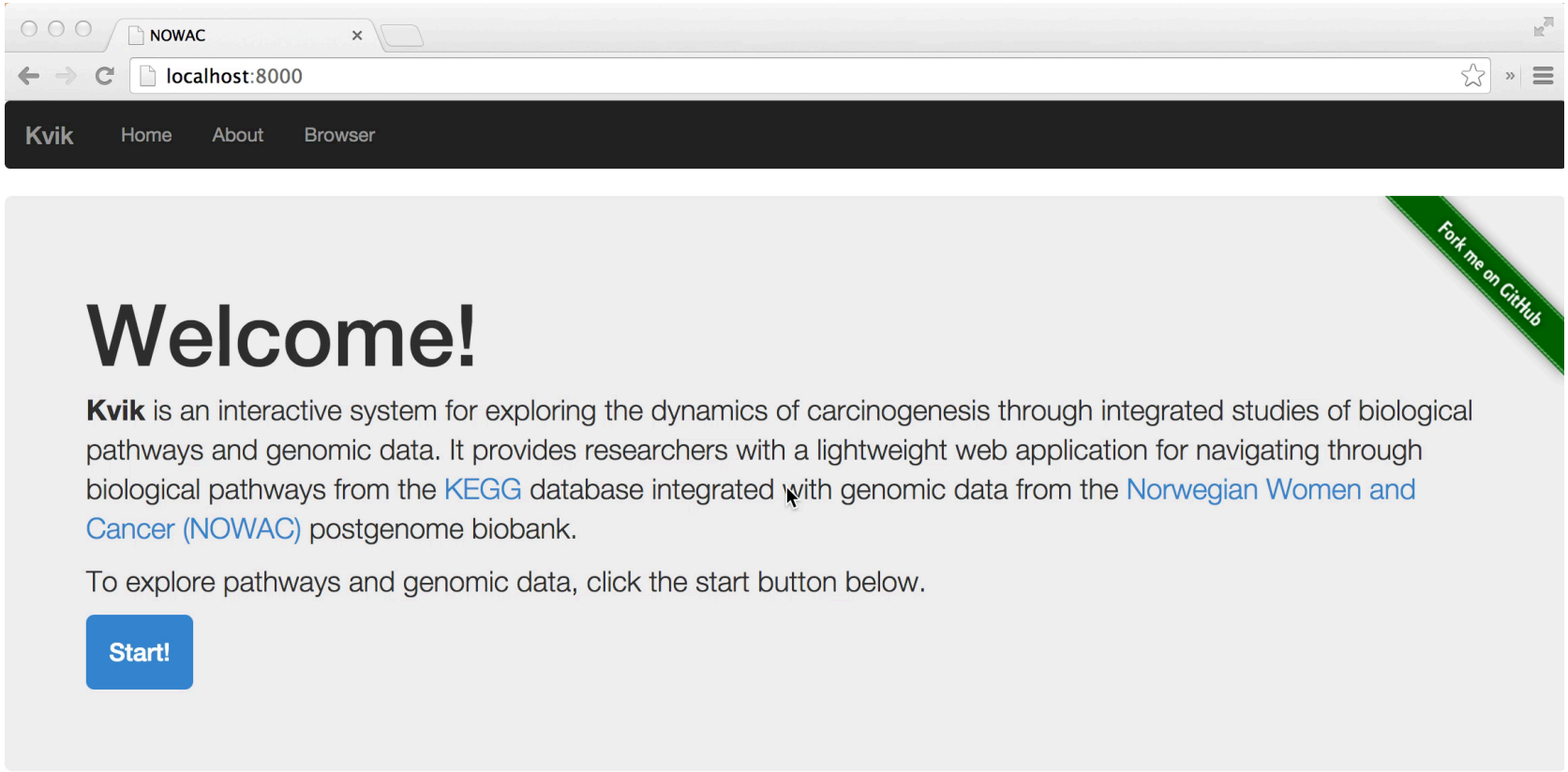
Kvik

- Integrates pathways and other information from KEGG together with data from the NOWAC study
- Makes it accessible through a web application that you can use on any device

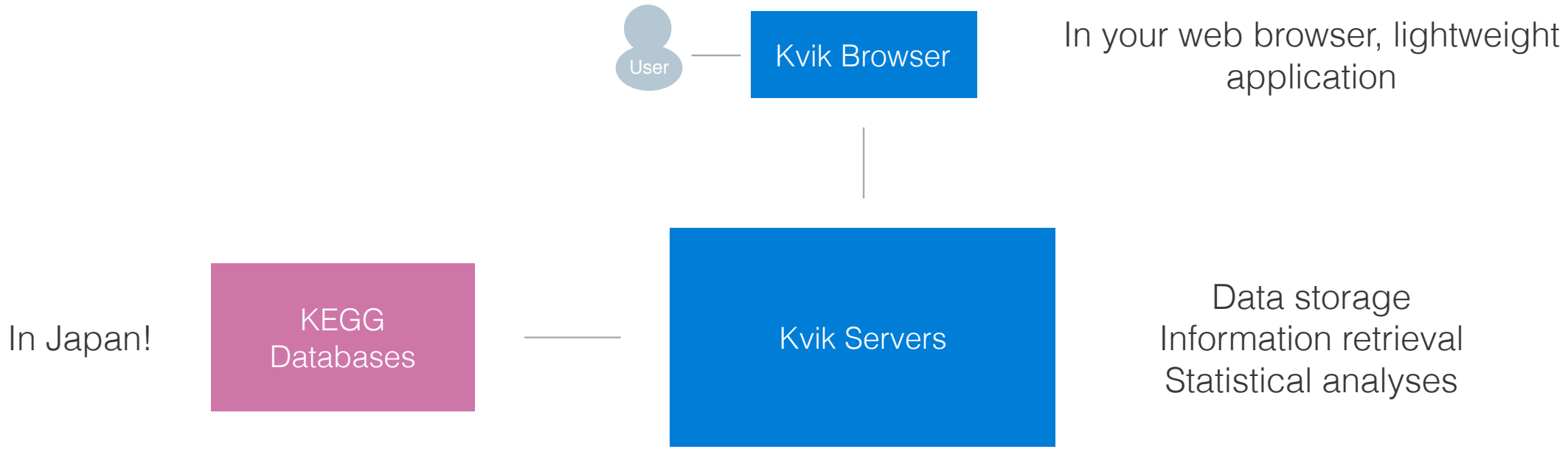
Demo


NOWAC
localhost:8000
Kvik
Home
About
Browser
Fork me on GitHub
Welcome!
Kvik is an interactive system for exploring the dynamics of carcinogenesis through integrated studies of biological pathways and genomic data. It provides researchers with a lightweight web application for navigating through biological pathways from the KEGG database integrated with genomic data from the Norwegian Women and Cancer (NOWAC) postgenome biobank.
To explore pathways and genomic data, click the start button below.
Start!

Demo


NOWAC
localhost:8000
Kvik
Home
About
Browser
Fork me on GitHub
Welcome!
Kvik is an interactive system for exploring the dynamics of carcinogenesis through integrated studies of biological pathways and genomic data. It provides researchers with a lightweight web application for navigating through biological pathways from the KEGG database integrated with genomic data from the Norwegian Women and Cancer (NOWAC) postgenome biobank.
To explore pathways and genomic data, click the start button below.
Start!

What's under the hood?



KEGG: Kyoto Encyclopedia of Genes and Genomes

Category	Entry Point	Search & Compute	DBGET Search
Systems information	**KEGG PATHWAY** **KEGG BRITE** **KEGG MODULE** **KEGG Mapper** **KEGG Atlas**	Search Pathway Search Brite Reconstruct Module Map Taxonomy	PATHWAY BRITE MODULE
Genomic information	**KEGG ORTHOLOGY** **KEGG Annotation**	BlastKOALA *New!* KO system	ORTHOLOGY
	KEGG GENOME **KEGG GENES** **KEGG Organisms** [Species | Genus]	SSDB search OC viewer† BLAST† / FASTA† KAAS†	GENOME GENES DGENES MGENOME† MGENES†
Chemical information	**KEGG LIGAND** **KEGG COMPOUND** **KEGG GLYCAN** **KEGG REACTION** **Reaction Modules**	SIMCOMP† / SUBCOMP† KCaM† PathSearch† PathComp† PathPred† E-zyme†	COMPOUND GLYCAN REACTION RPAIR RCLASS ENZYME
Health information	**KEGG DISEASE** **KEGG DRUG** **KEGG ENVIRON** **KEGG MEDICUS**	MEDICUS search Drug interaction checker Human diseases Infectious diseases ATC drug classification	DISEASE DRUG DGROUP ENVIRON

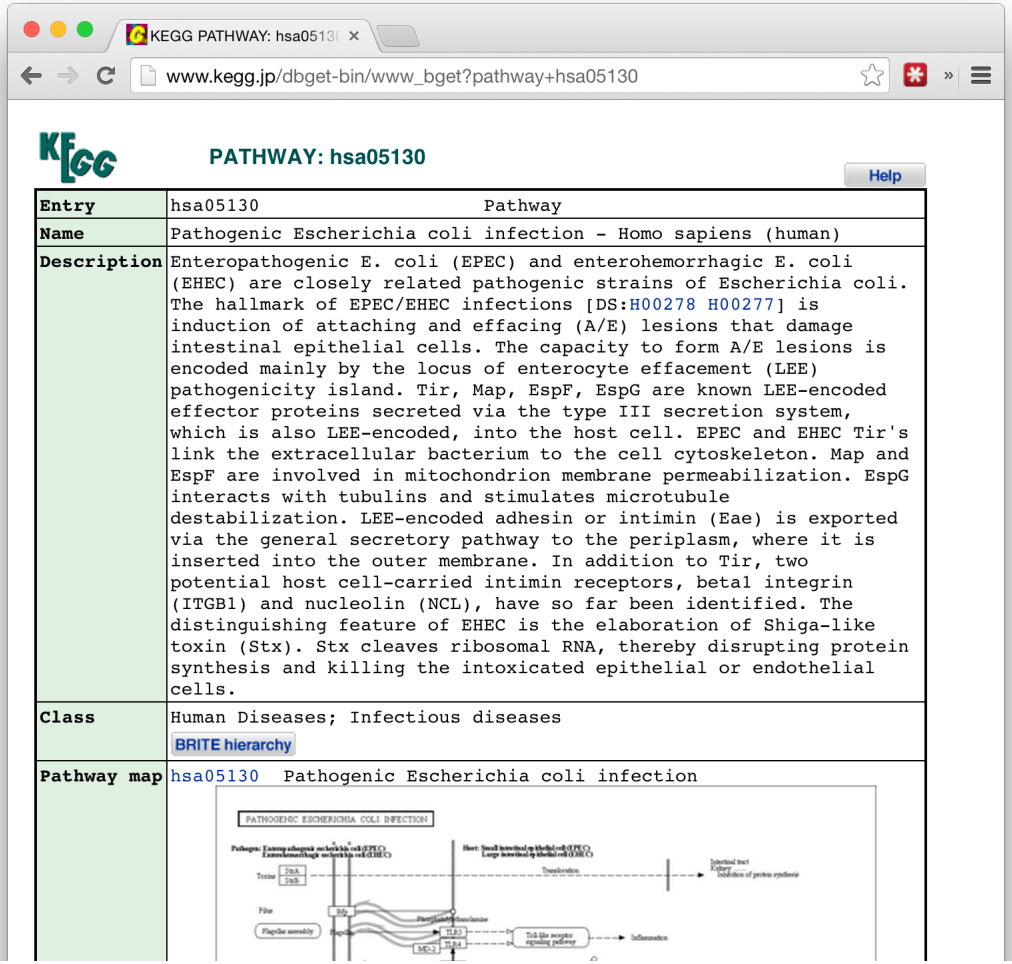
KEGG: Kyoto Encyclopedia of Genes and Genomes

Category	Entry Point	Search & Compute	DBGET Search
Systems information	**KEGG PATHWAY** **KEGG BRITE** **KEGG MODULE** **KEGG Mapper** **KEGG Atlas**	Search Pathway Search Brite Reconstruct Module Map Taxonomy	PATHWAY BRITE MODULE
Genomic information	**KEGG ORTHOLOGY** **KEGG Annotation**	BlastKOALA *New!* KO system	ORTHOLOGY
	KEGG GENOME **KEGG GENES** **KEGG Organisms** [Species | Genus]	SSDB search OC viewer† BLAST† / FASTA† KAAS†	GENOME GENES DGENES MGENOME† MGENES†
Chemical information	**KEGG LIGAND** **KEGG COMPOUND** **KEGG GLYCAN** **KEGG REACTION** **Reaction Modules**	SIMCOMP† / SUBCOMP† KCaM† PathSearch† PathComp† PathPred† E-zyme†	COMPOUND GLYCAN REACTION RPAIR RCLASS ENZYME
Health information	**KEGG DISEASE** **KEGG DRUG** **KEGG ENVIRON** **KEGG MEDICUS**	MEDICUS search Drug interaction checker Human diseases Infectious diseases ATC drug classification	DISEASE DRUG DGROUP ENVIRON

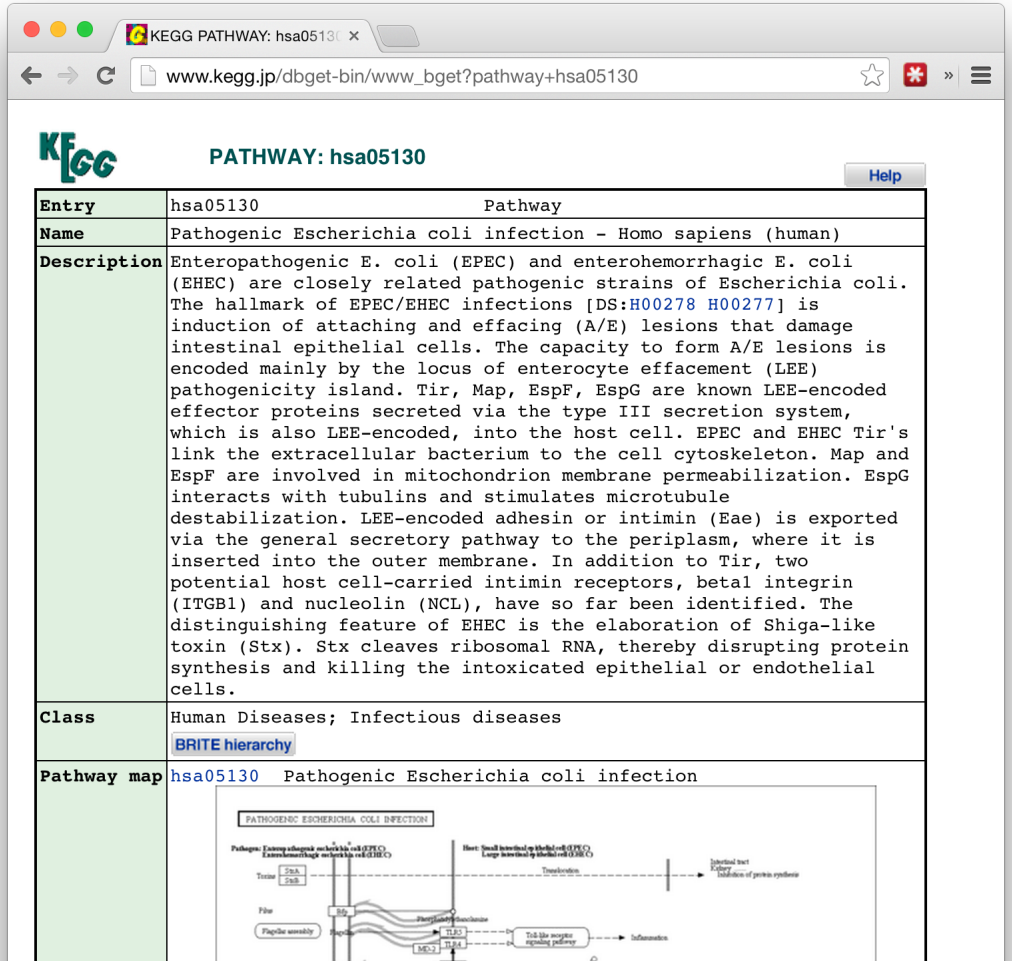
KEGG

- Two options for retrieving data
  - Expensive FTP-licence
  - Free REST API

KEGG REST API

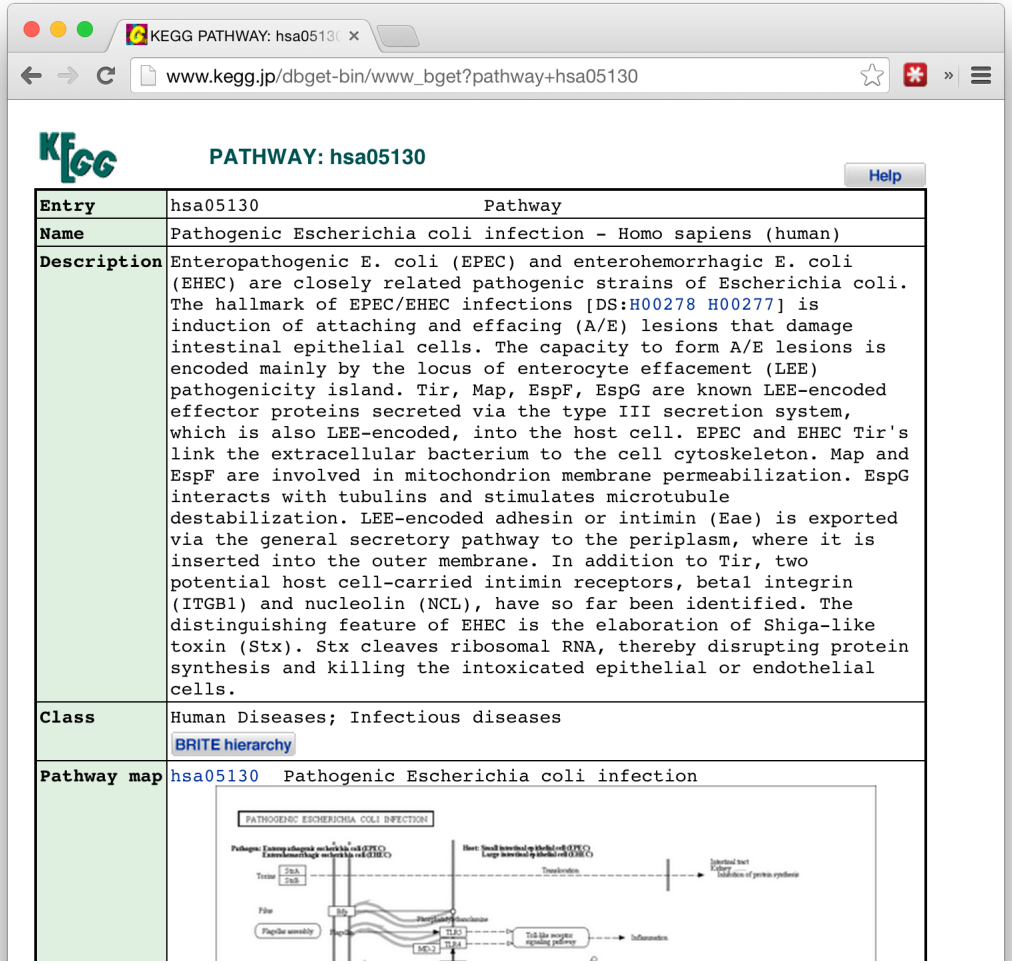


KEGG REST API



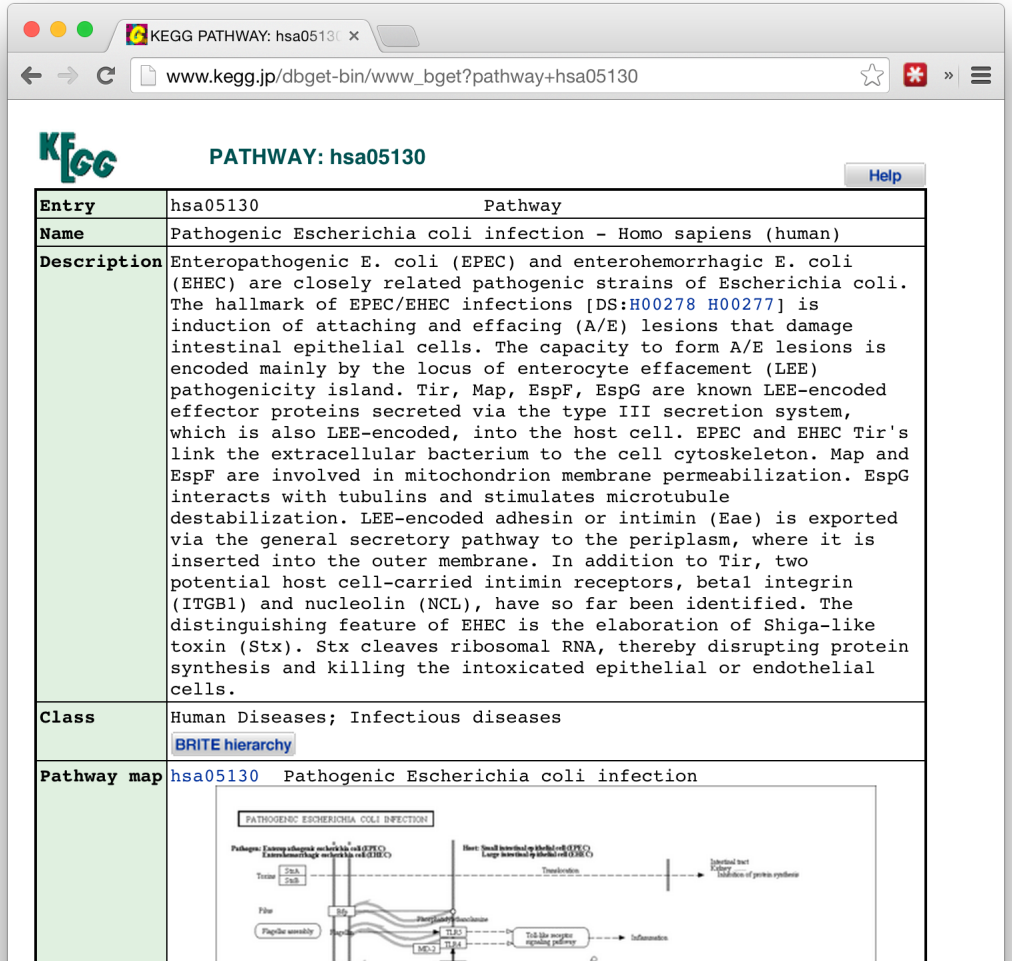
rest.kegg.jp/get/hsa05130

KEGG REST API



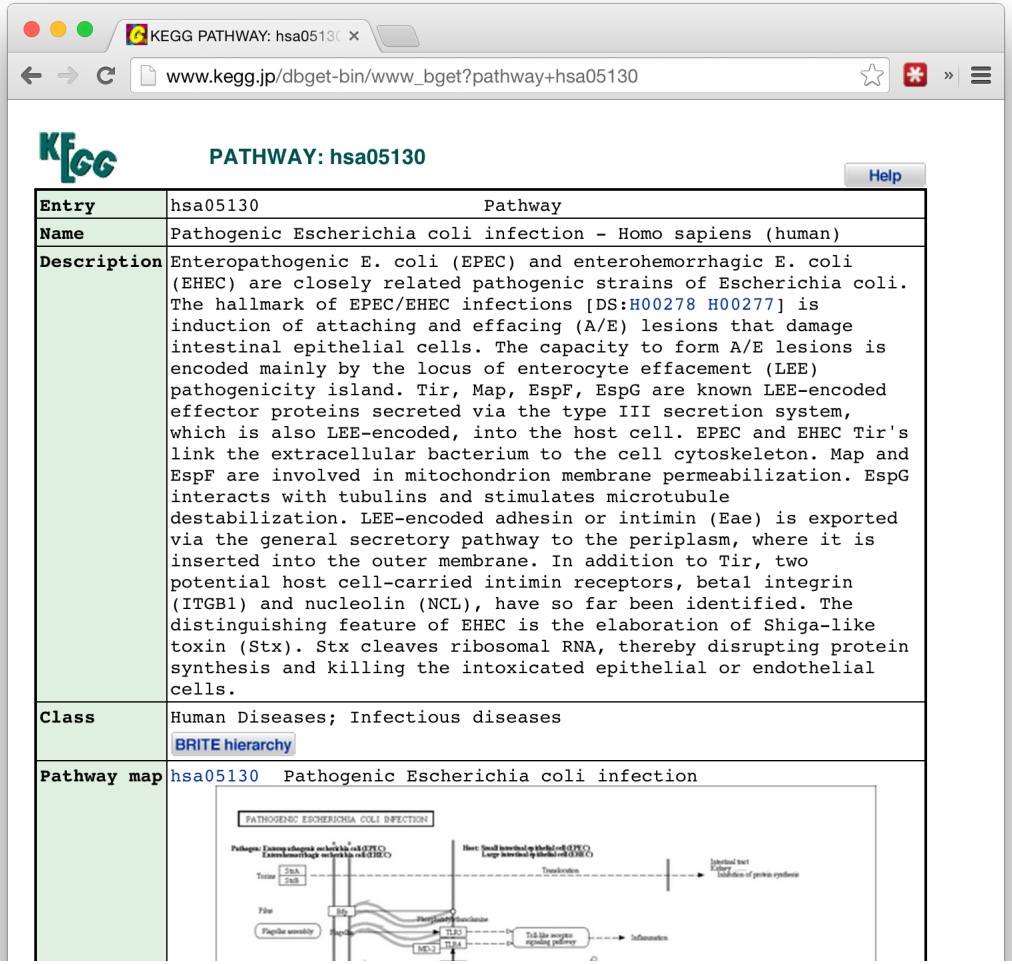
rest.kegg.jp/get/hsa05130

KEGG REST API



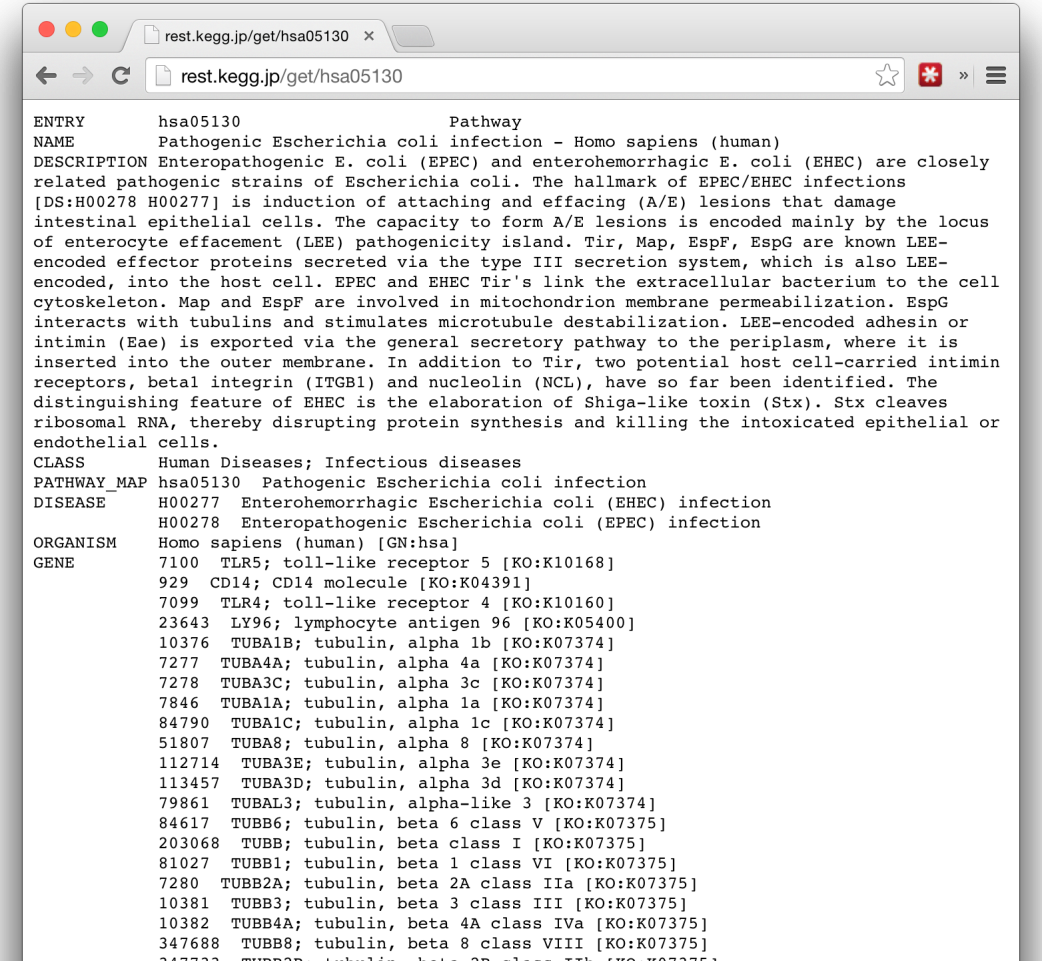
rest.kegg.jp/get/hsa05130

KEGG REST API



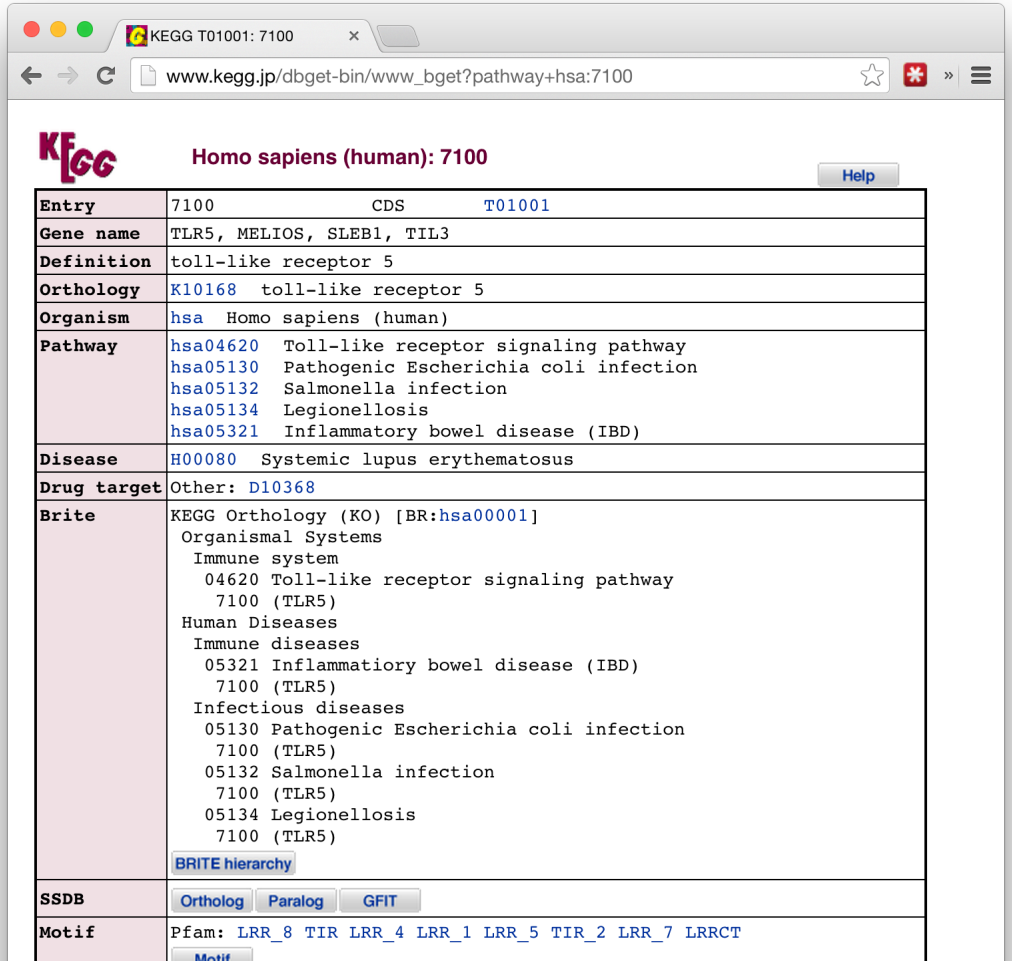
rest.kegg.jp/get/hsa05130

KEGG REST API

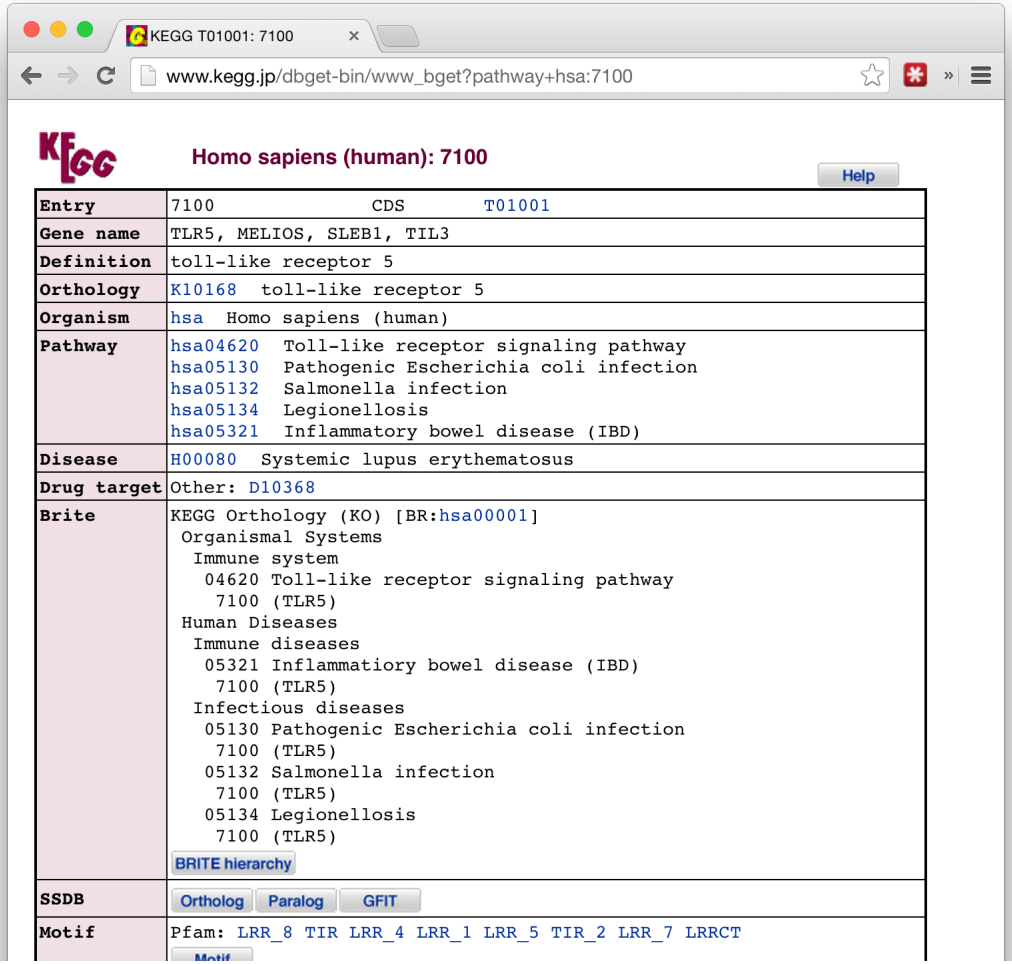



```
ENTRY       hsa05130                    Pathway
NAME        Pathogenic Escherichia coli infection - Homo sapiens (human)
DESCRIPTION Enteropathogenic E. coli (EPEC) and enterohemorrhagic E. coli (EHEC) are closely
related pathogenic strains of Escherichia coli. The hallmark of EPEC/EHEC infections
[DS:H00278 H00277] is induction of attaching and effacing (A/E) lesions that damage
intestinal epithelial cells. The capacity to form A/E lesions is encoded mainly by the locus
of enterocyte effacement (LEE) pathogenicity island. Tir, Map, EspF, EspG are known LEE-
encoded effector proteins secreted via the type III secretion system, which is also LEE-
encoded, into the host cell. EPEC and EHEC Tir's link the extracellular bacterium to the cell
cytoskeleton. Map and EspF are involved in mitochondrion membrane permeabilization. EspG
interacts with tubulins and stimulates microtubule destabilization. LEE-encoded adhesin or
intimin (Eae) is exported via the general secretory pathway to the periplasm, where it is
inserted into the outer membrane. In addition to Tir, two potential host cell-carried intimin
receptors, beta1 integrin (ITGB1) and nucleolin (NCL), have so far been identified. The
distinguishing feature of EHEC is the elaboration of Shiga-like toxin (Stx). Stx cleaves
ribosomal RNA, thereby disrupting protein synthesis and killing the intoxicated epithelial or
endothelial cells.
CLASS       Human Diseases; Infectious diseases
PATHWAY_MAP hsa05130  Pathogenic Escherichia coli infection
DISEASE     H00277  Enterohemorrhagic Escherichia coli (EHEC) infection
            H00278  Enteropathogenic Escherichia coli (EPEC) infection
ORGANISM    Homo sapiens (human) [GN:hsa]
GENE        7100  TLR5; toll-like receptor 5 [KO:K10168]
            929  CD14; CD14 molecule [KO:K04391]
            7099  TLR4; toll-like receptor 4 [KO:K10160]
            23643  LY96; lymphocyte antigen 96 [KO:K05400]
            10376  TUBA1B; tubulin, alpha 1b [KO:K07374]
            7277  TUBA4A; tubulin, alpha 4a [KO:K07374]
            7278  TUBA3C; tubulin, alpha 3c [KO:K07374]
            7846  TUBA1A; tubulin, alpha 1a [KO:K07374]
            84790  TUBA1C; tubulin, alpha 1c [KO:K07374]
            51807  TUBA8; tubulin, alpha 8 [KO:K07374]
            112714  TUBA3E; tubulin, alpha 3e [KO:K07374]
            113457  TUBA3D; tubulin, alpha 3d [KO:K07374]
            79861  TUBAL3; tubulin, alpha-like 3 [KO:K07374]
            84617  TUBB6; tubulin, beta 6 class V [KO:K07375]
            203068  TUBB; tubulin, beta class I [KO:K07375]
            81027  TUBB1; tubulin, beta 1 class VI [KO:K07375]
            7280  TUBB2A; tubulin, beta 2A class IIa [KO:K07375]
            10381  TUBB3; tubulin, beta 3 class III [KO:K07375]
            10382  TUBB4A; tubulin, beta 4A class IVa [KO:K07375]
            347688  TUBB8; tubulin, beta 8 class VIII [KO:K07375]
```

KEGG REST API

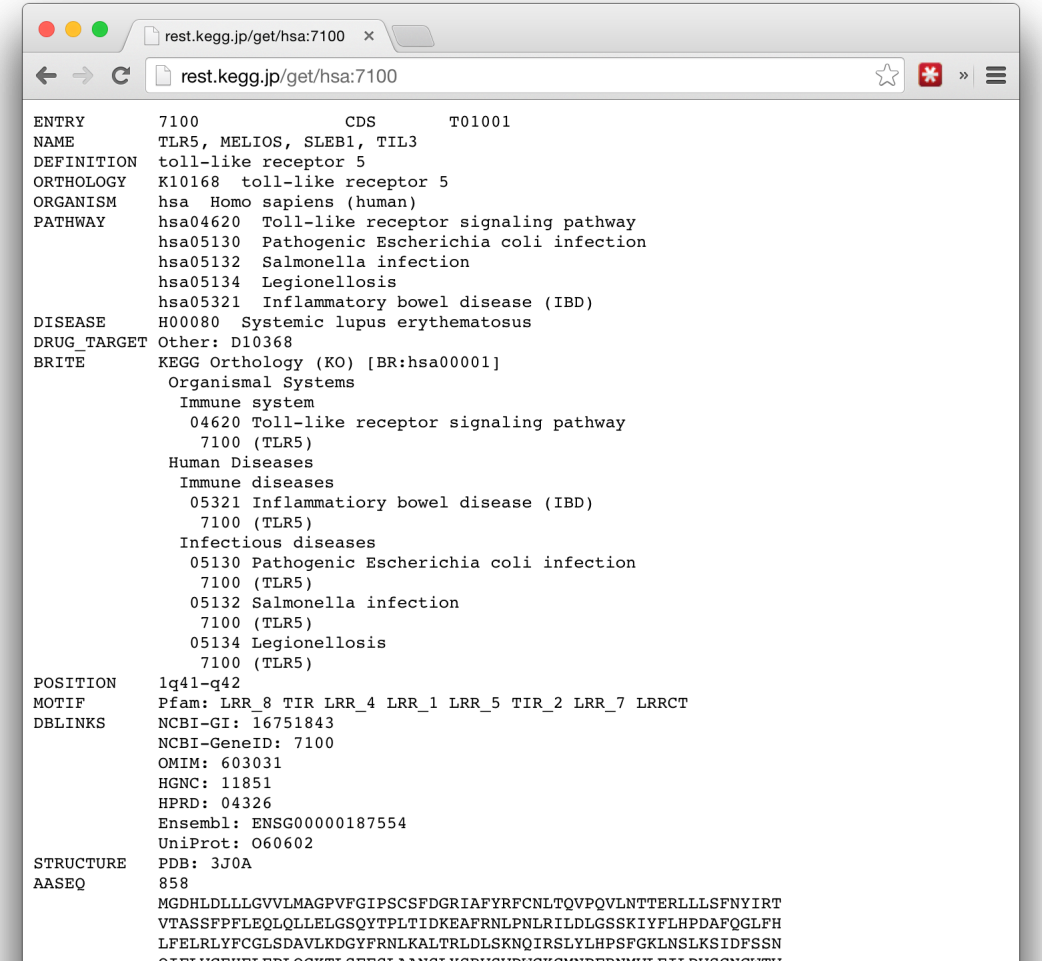


KEGG REST API



rest.kegg.jp/get/hsa:7100

KEGG REST API

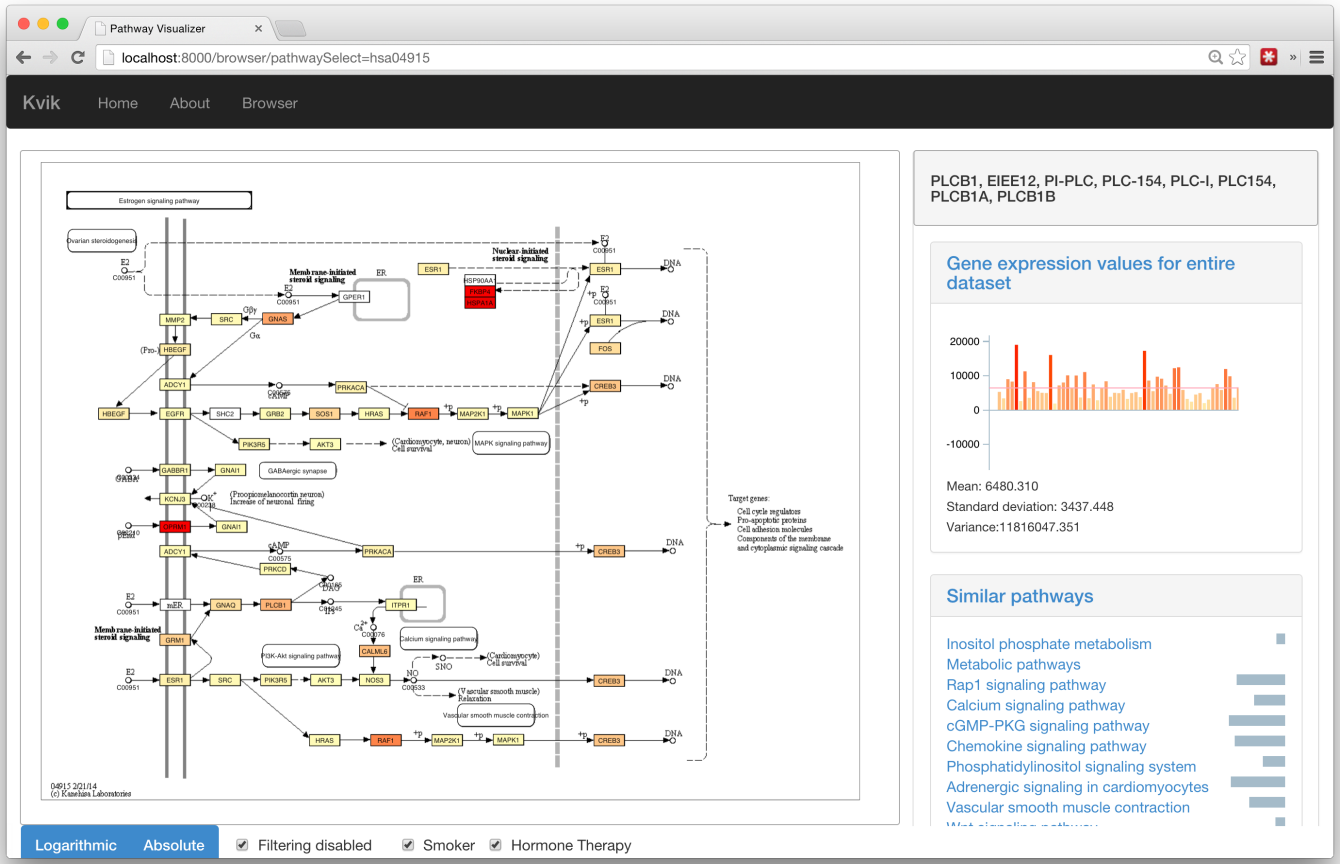


rest.kegg.jp/get/hsa:7100

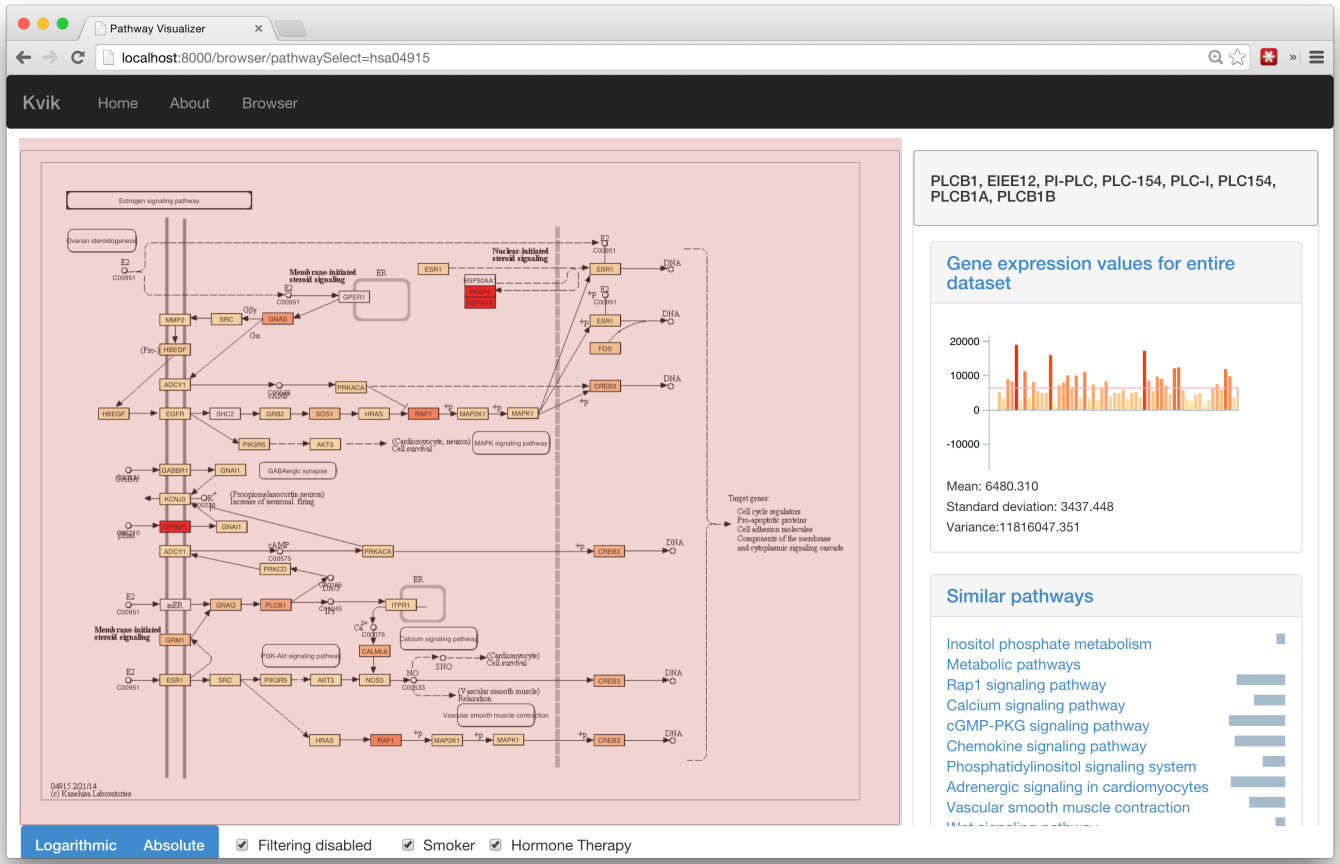
KEGG REST API

- There is a lot you can do with it!
- See rest.kegg.jp for more information

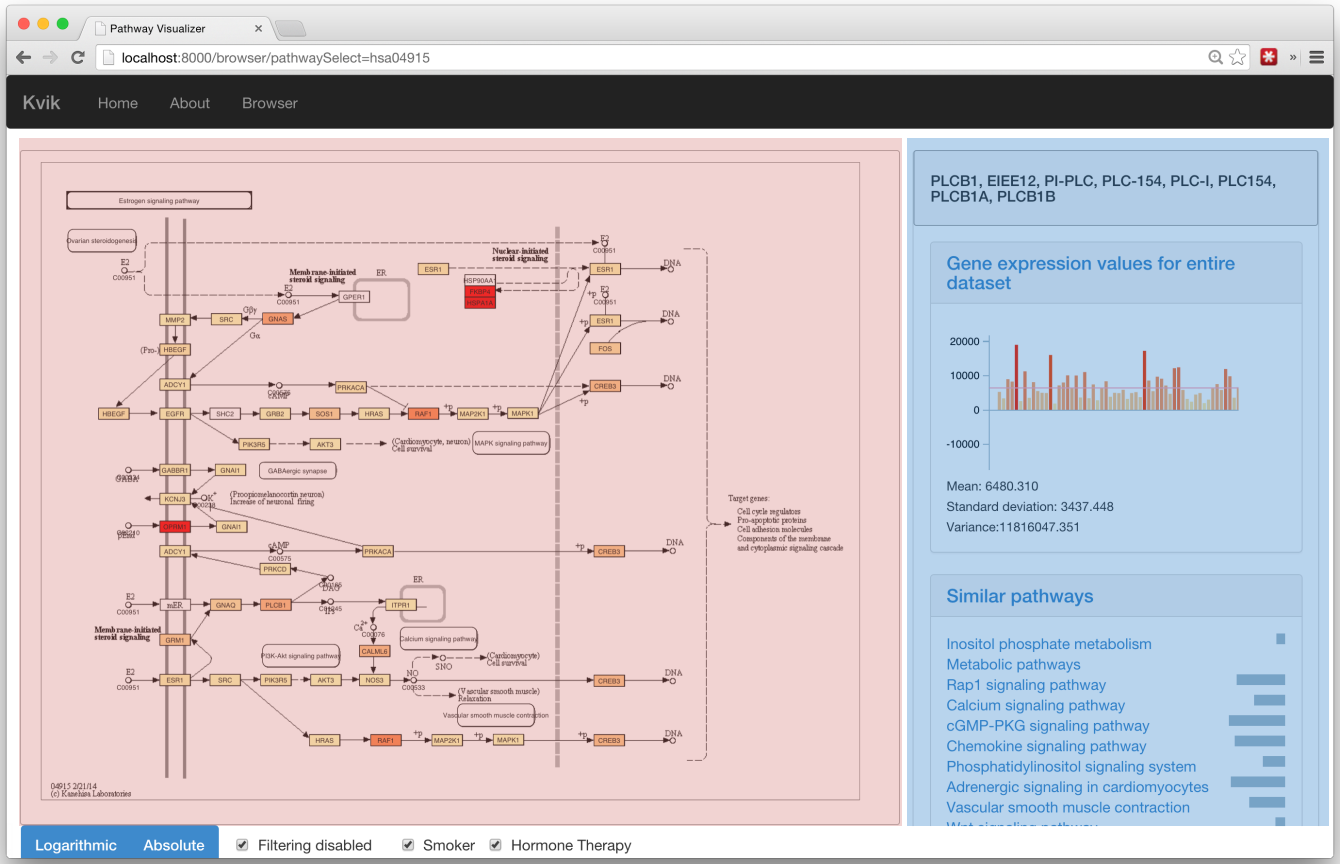
KEGG usage in Kvik



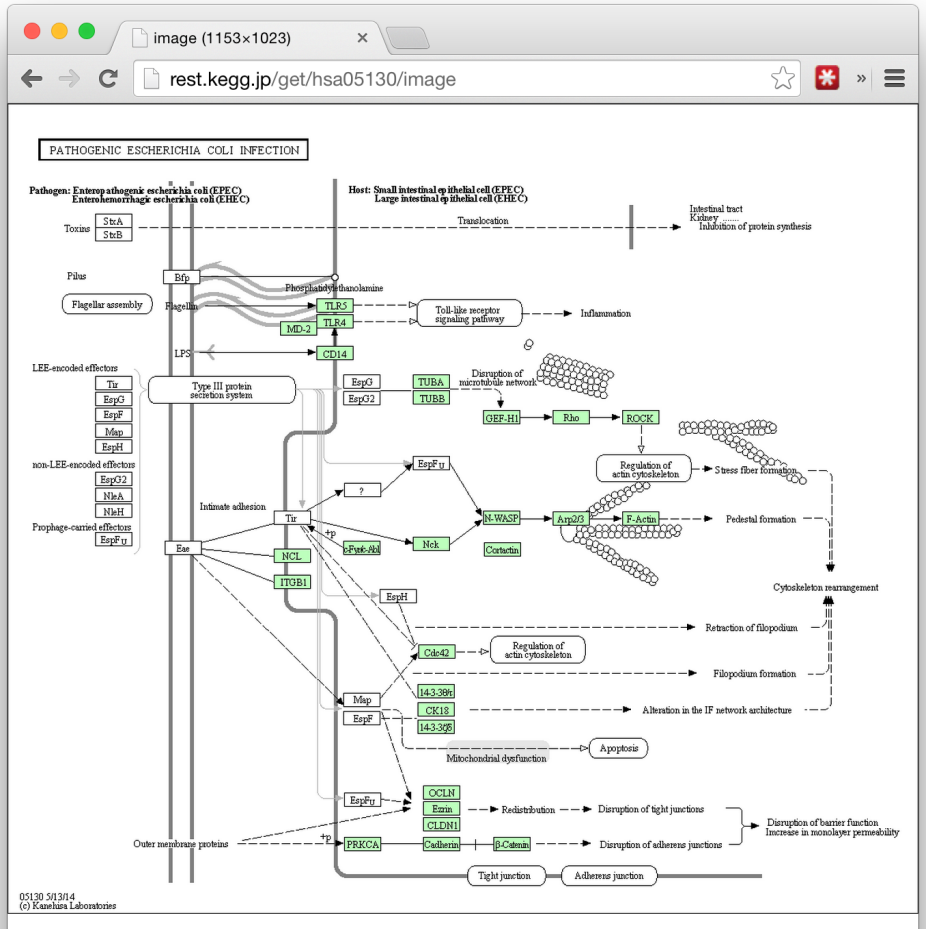
KEGG usage in Kvik

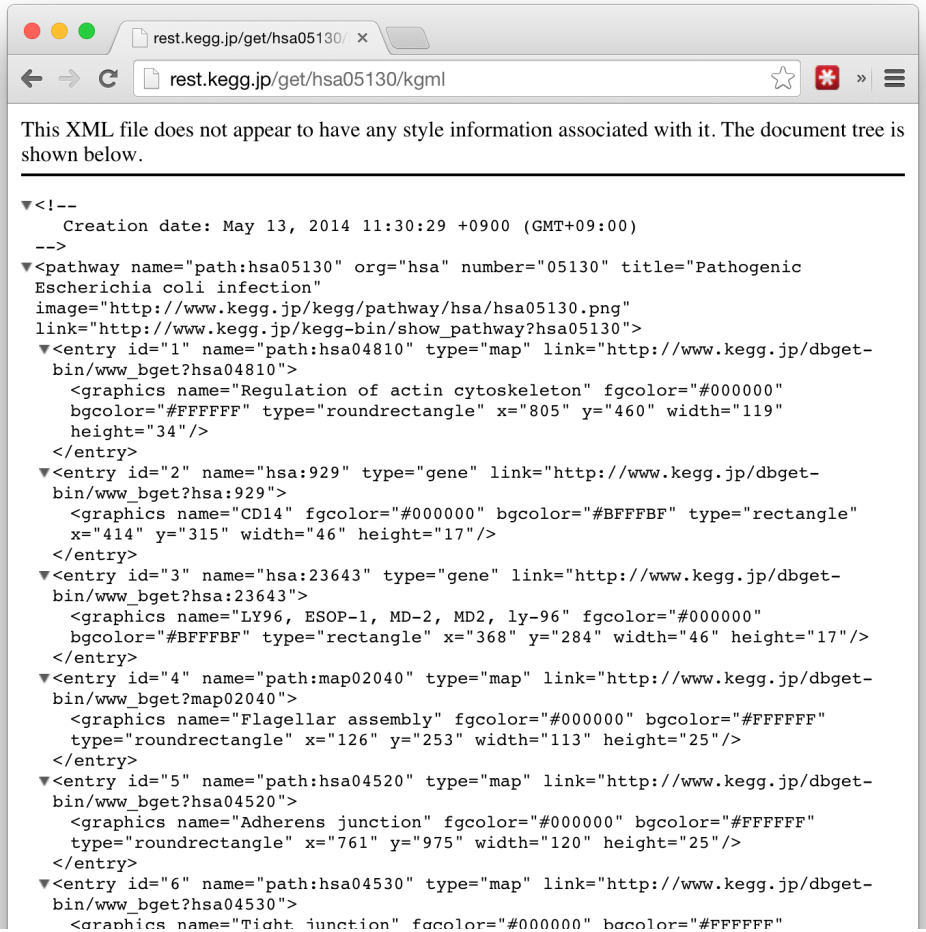


KEGG usage in Kvik



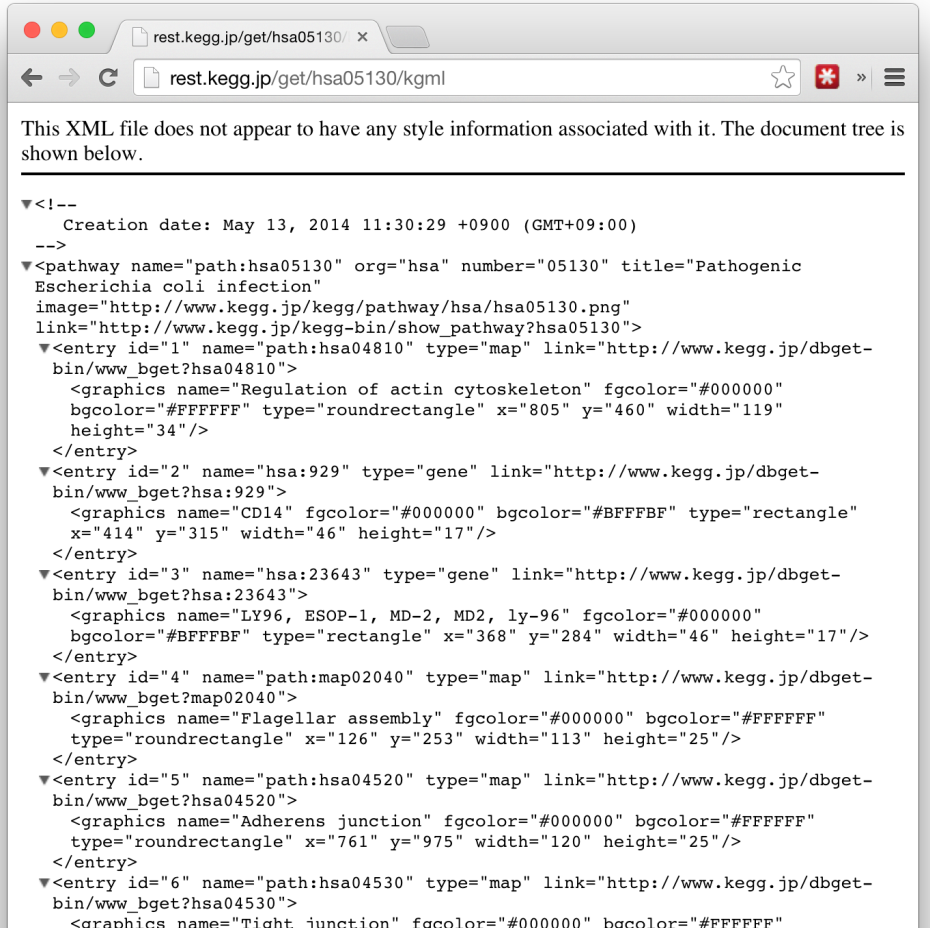
KEGG Pathways



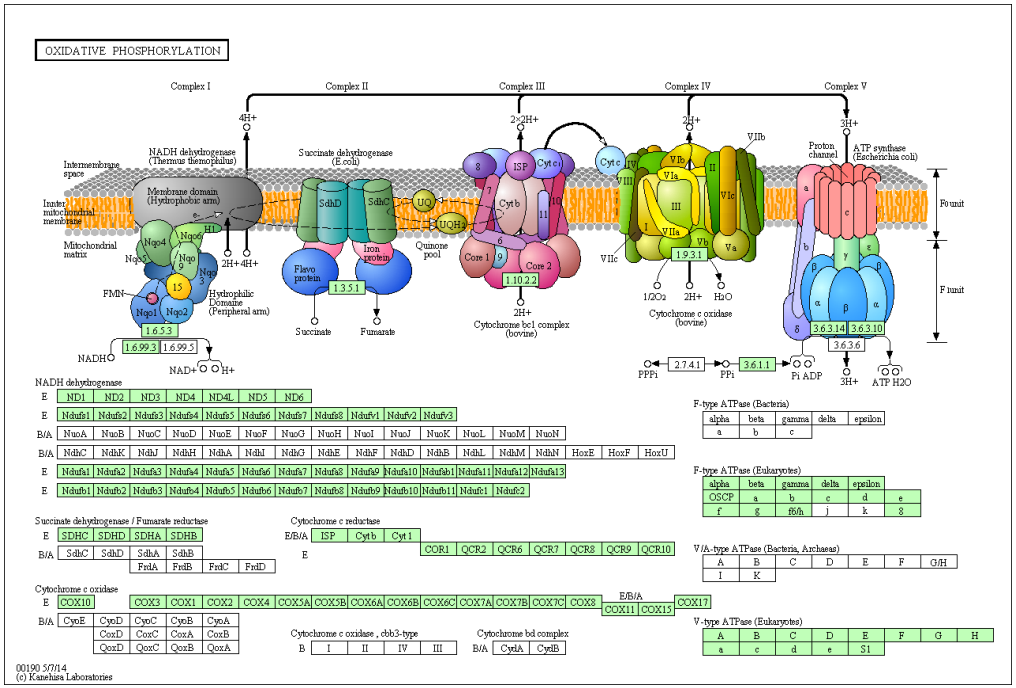


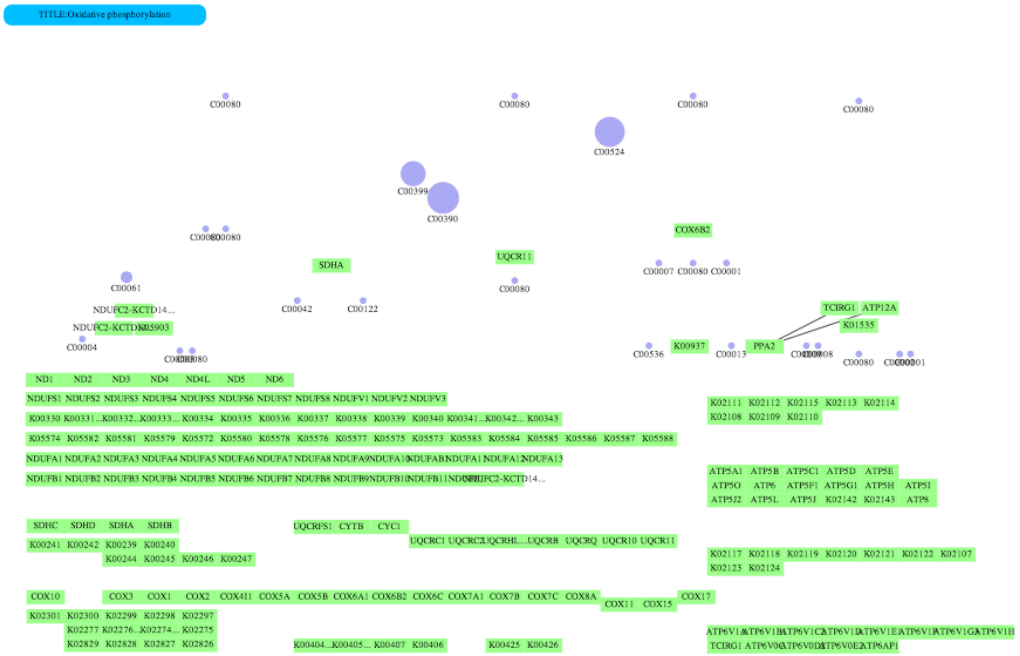
KEGG Pathways

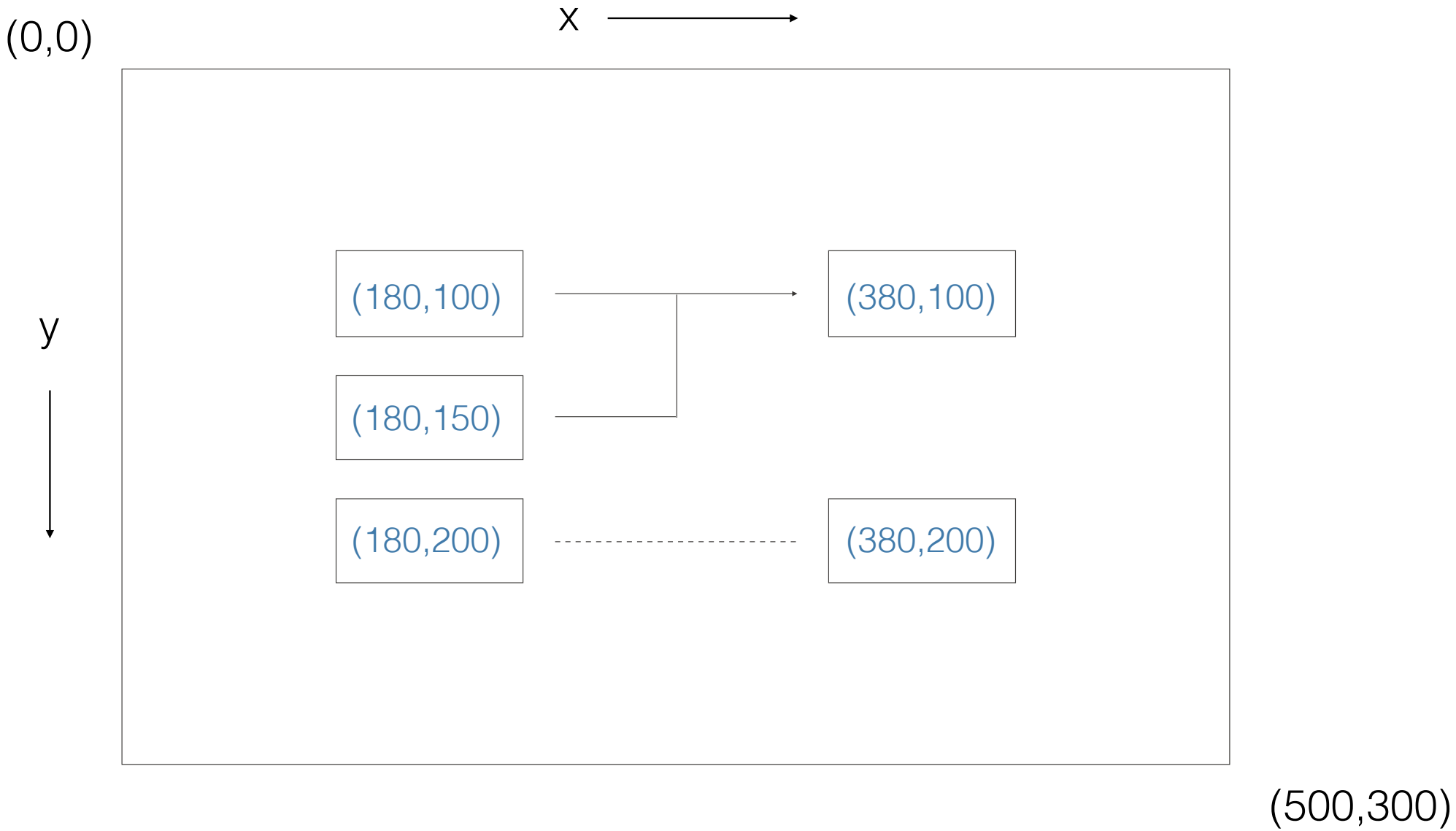
```
<pathway name="path:hsa05130" … >
  <entry id=1 name="path:hsa048140" … >
    .
    .
    .
  <relation entryid1=1 entryid2=3 … >
    .
    .
    .
</pathway>
```



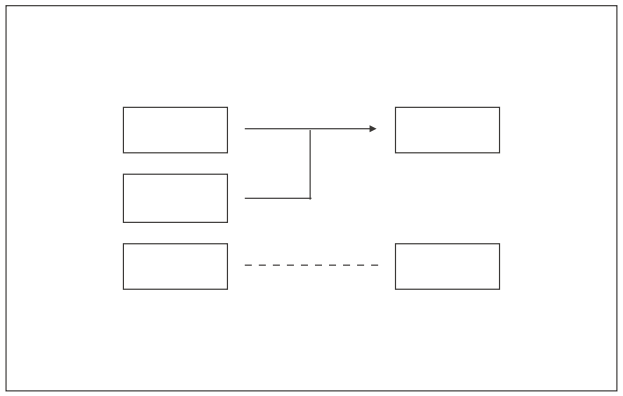
Combining image and KGML



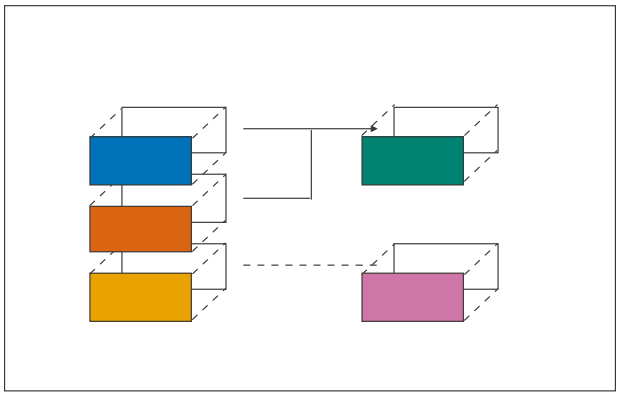



(0,0)
x
y
(180,100)
(380,100)
(180,150)
(180,200)
(380,200)
(500,300)

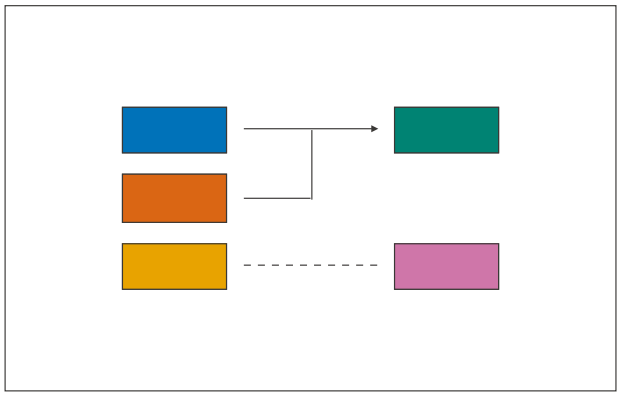
Pathways in Kvik



Static image from KEGG

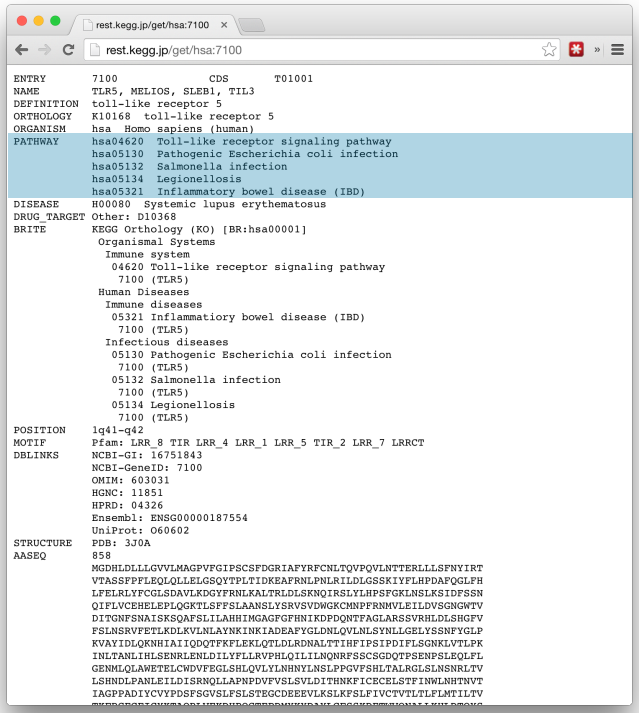


Get location from KGML file, and color the gene according to data from NOWAC

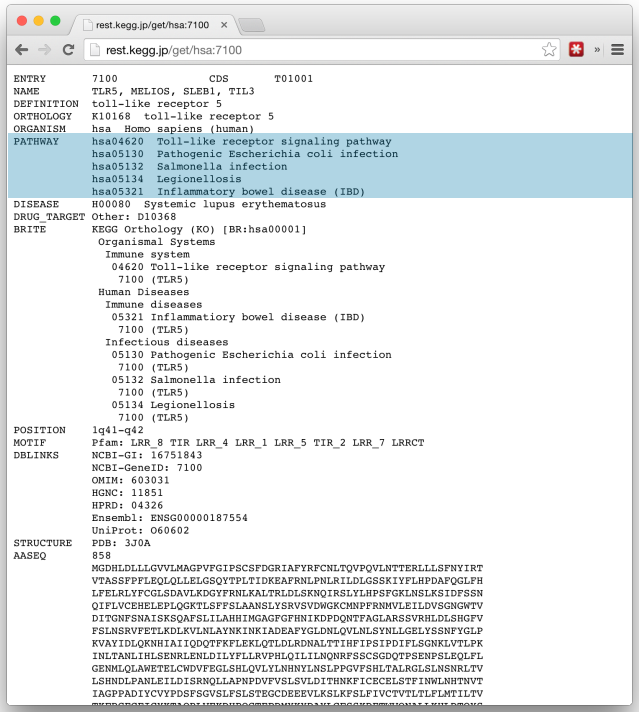


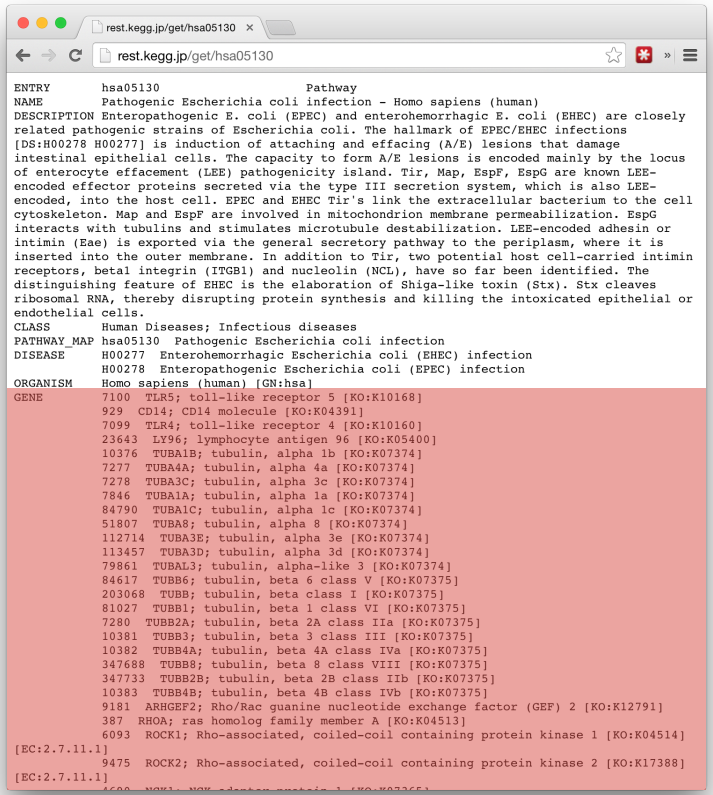
Final visualiation

Pathway *similarity*

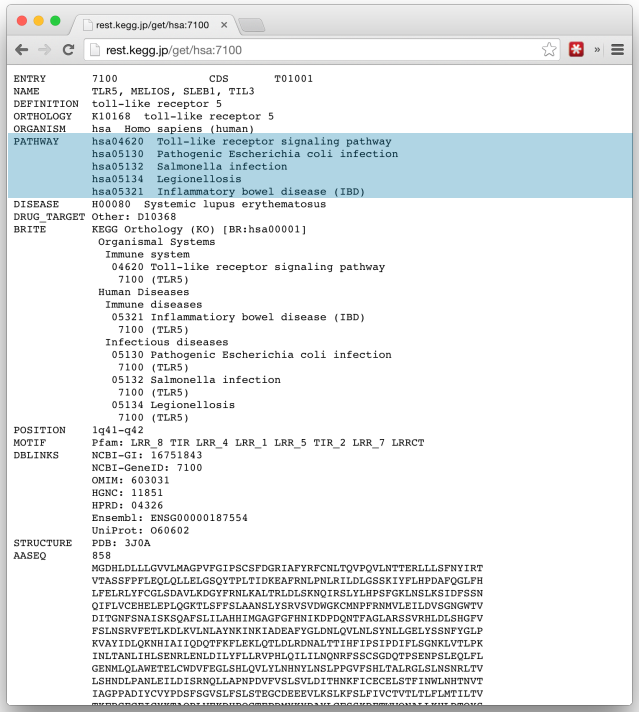


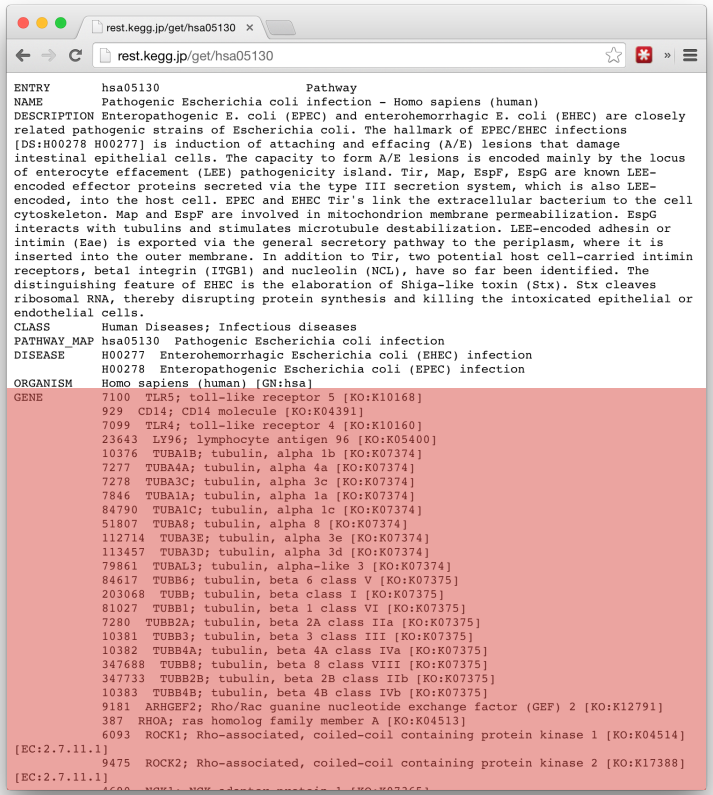
Pathway *similarity*





Pathway *similarity*





Performance

- Long latency (time to retrieve something) when using the REST API
- Kvik uses *caching* to reduce this latency

Future work

- Visualize more than one pathway at a time
- Make the web-application capable of handling more than one user
- Making it publicly available at kvik.cs.uit.no

Concluding Remarks

- Kvik provides interactive exploration of biological pathways from KEGG and gene expression data from the NOWAC study
- Without any installation or plug-ins, researchers can start exploring the data instantly
- Kvik uses KEGG to provide pathway maps and other information
- Open-sourced at github.com/fjukstad/kvik

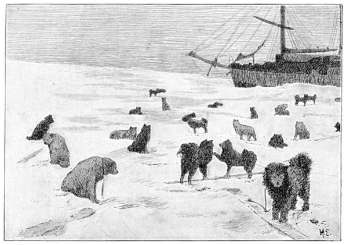
Pathways on the Display Wall

- At the dept. of Comp. Sci. we got a wall-sized display with a 7168x3072 resolution!
- Visualize all the pathways!

UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

Thank you!

Bjørn Fjukstad
bjorn@cs.uit.no



Kvik — Lead sled dog on Fridtjof Nansen's expedition to the North Pole

UNIVERSITY OF TROMSØ