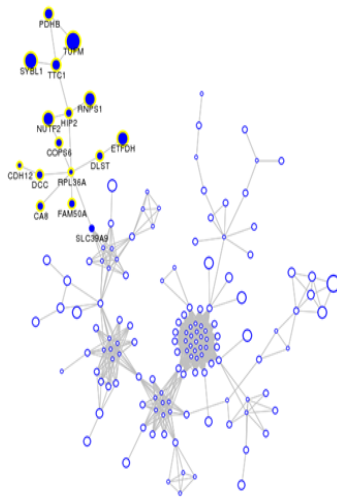




biolab
Bioinformatics Laboratory

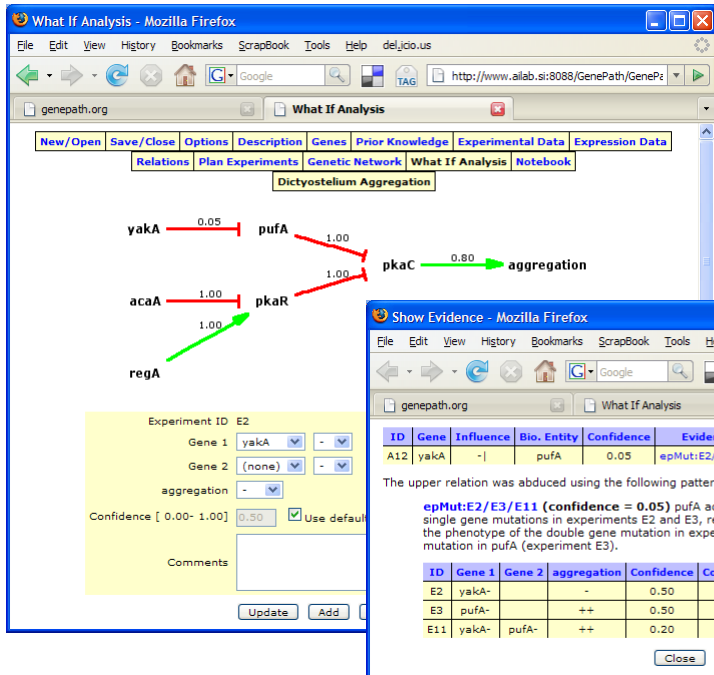
University of Ljubljana
Faculty of Computer and Information Science

Interests

- High-throughput genomics data analysis
- Data mining tools and methods
- Web-based data analytics
- Visualization

Principal Partners

- Baylor College of Medicine, Houston
 - Functional genomics of social amoeba *Dictyostelium*
- MRC Laboratory of Molecular Biology, Cambridge
 - Genome analytics of neurodegenerative diseases
- University of Pavia, Italy
 - Bioinformatics for tissue engineering
- AstraZeneca, Sweden & UK
 - Data mining tools in pharmacology



GenePath

The screenshot displays the GenePath software interface within a Mozilla Firefox browser window. The main window shows a metabolic pathway diagram with genes (yaka, pufa, pkaR, pkaC, aggregation) and their interactions (inhibition, activation) with associated confidence values. A 'Show Evidence' window is open, displaying a table of evidence for the relationship between yaka and pufa.

Pathway Diagram:

```

graph LR
    yaka -- 0.05 --| pufa
    pufa -- 1.00 --> pkaC
    pkaC -- 0.80 --> aggregation
    acaA -- 1.00 --| pkaR
    pkaR -- 1.00 --> pkaC
    regA -- 1.00 --> pkaR
  
```

Show Evidence Window:

ID	Gene	Influence	Bio. Entity	Confidence	Evidence
A12	yaka	-	pufa	0.05	epMut:E2/E3/E11

The upper relation was abducted using the following pattern:

epMut:E2/E3/E11 (confidence = 0.05) pufa acts after yaka because the phenotypes of the single gene mutations in experiments E2 and E3, respectively, are different from each other and the phenotype of the double gene mutation in experiment E11 is the same as for the single gene mutation in pufa (experiment E3).

ID	Gene 1	Gene 2	aggregation	Confidence	Comments
E2	yaka-		-	0.50	
E3	pufa-		++	0.50	
E11	yaka-	pufa-	++	0.20	

Inference of Gene Networks

genetic experiments

ID	Gene 1	Gene 2	aggregation
E1			+
E2	yakA-		-
E3	pufA-		++
E4	pkaR-		++
E5	pkaC-		-
E6	acaA-		-
E7	regA-		++
E8	acaA+		++
E9	pkaC+		++
E10	pkaC-	regA-	-
E11	yakA-	pufA-	++
E12	yakA-	pkaR-	±
E13	yakA-	pkaC-	-
E14	pkaC-	yakA+	-
E15	yakA-	pkaC+	++

pattern-based inference

ID	Gene	Influence	Bio. Entity
A3	yakA	->	aggregation
A4	pufA	-	aggregation
A5	pkaR	-	aggregation
A6	pkaC	->	aggregation
A7	acaA	->	aggregation
A8	regA	-	aggregation
A9	regA	-	pkaC
A10	yakA	-	pufA
A11	pkaR		yakA

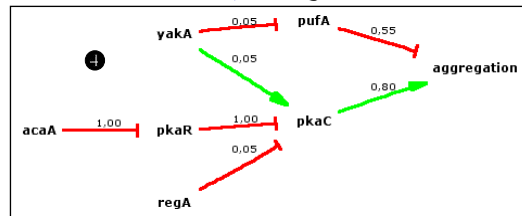
relations btw genes and outcomes

constraint satisfaction

prior knowledge

ID	Gene	Influence	Bio. Entity
P1	pkaR	-	pkaC
P2	acaA	-	pkaR

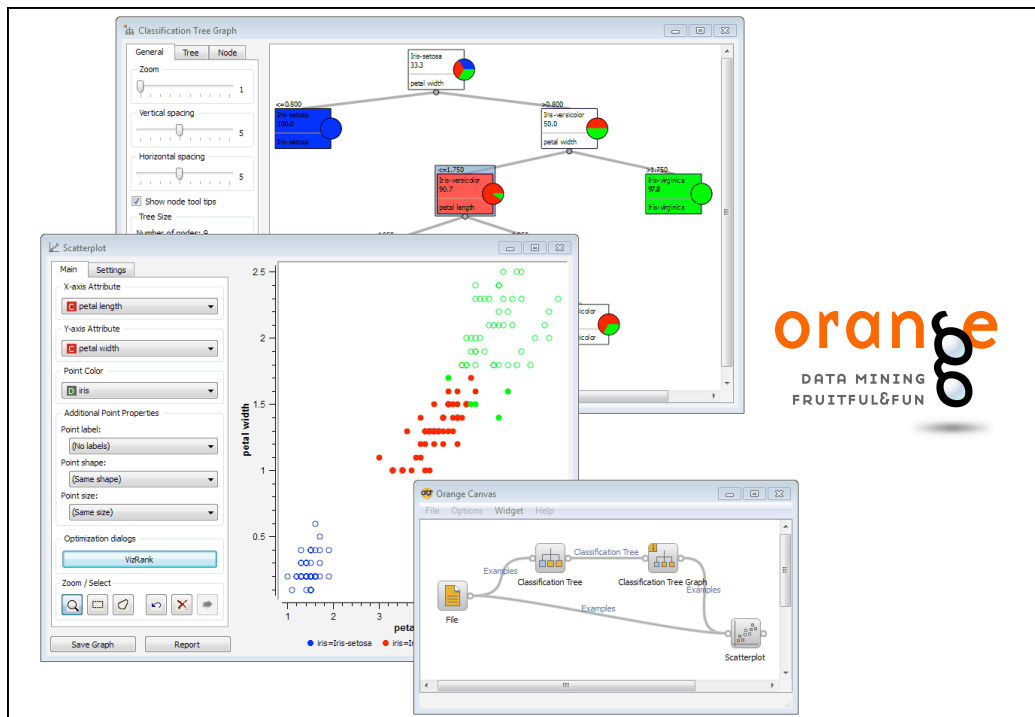
genetic network



Zupan et al. *Bioinformatics* 2003: 19(3).

Juvan et al. *Nucl Acids Res* 2005: 33.

Van Driessche et al. *Nature Genetics* 2005: 37(5).



Orange Scripting

```
import obiKEGG
import obiGeneMatch

organism = "hsa"
keggorg = obiKEGG.KEGGorganism(organism)
pathways = keggorg.list_pathways()
geneids = keggorg.get_genes()

print "Organism: %s, Pathways: %d, Genes: %d" % (organism, len(pathways), len(geneids))

genes = ["CCR5", "Bcl2", "BRCA1", "LEP"]
matcher = obiGeneMatch.GeneMatch(genes, organism=organism, caseSensitive=False)
matched = matcher.match(geneids)

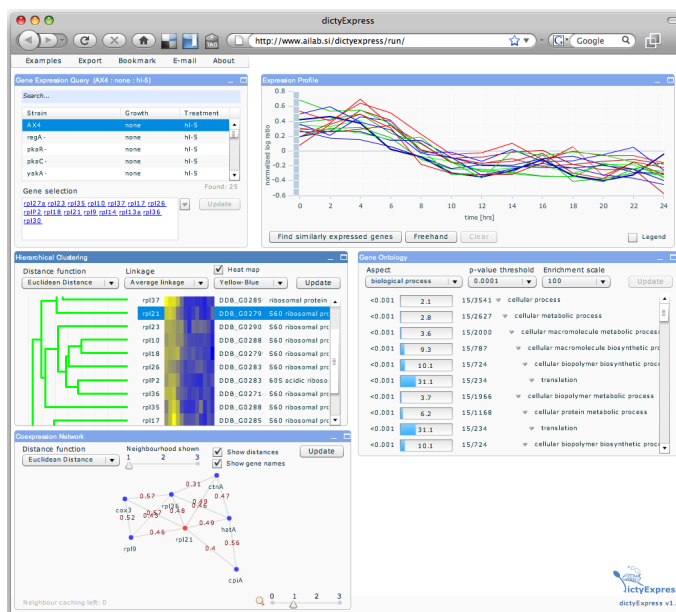
for geneid, genename in matched:
    matchedpathways = keggorg.get_pathways_by_genes([geneid])
    print "Gene: %s (%s)" % (genename, geneid)
    print "\n".join(["%s (%s)" % (pathways[p], p) for p in matchedpathways])
```

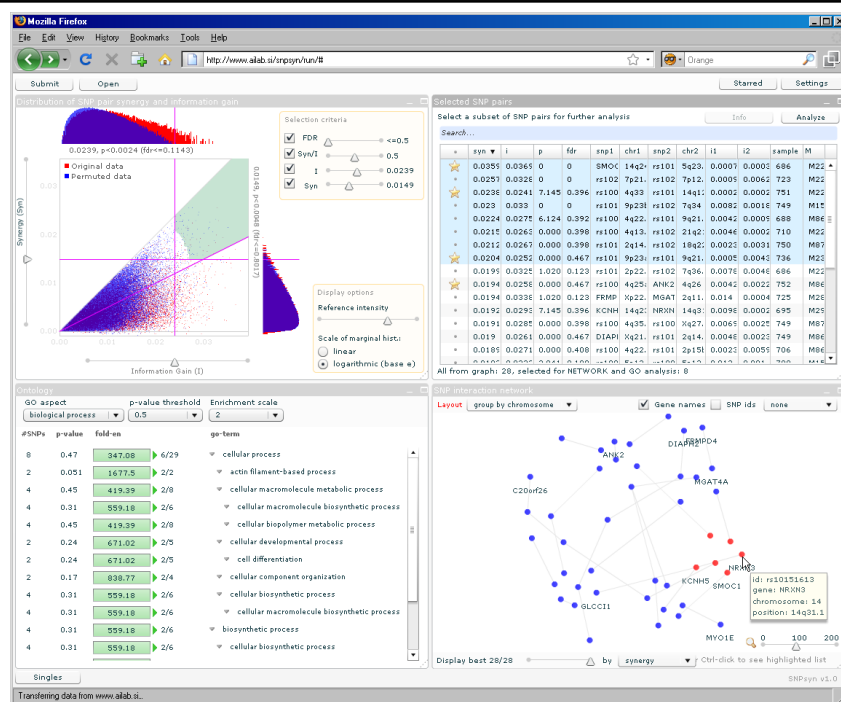
script in
Python

For each gene in a list, the script
finds which
KEGG pathways include them.

```
In [1]: run L4-kegg.py
Organism: hsa, Pathways: 213, Genes: 24293
Gene: CCR5 (hsa:1234):
  Cytokine-cytokine receptor interaction (path:hsa04060)
Gene: LEP (hsa:3952):
  Neuroactive ligand-receptor interaction (path:hsa04080)
  Jak-STAT signaling pathway (path:hsa04630)
  Cytokine-cytokine receptor interaction (path:hsa04060)
  Adipocytokine signaling pathway (path:hsa04920)
Gene: Bcl2 (hsa:596):
  Apoptosis (path:hsa04210)
  Focal adhesion (path:hsa04510)
  Amyotrophic lateral sclerosis (ALS) (path:hsa05030)
  Neurodegenerative Diseases (path:hsa01510)
  Small cell lung cancer (path:hsa05222)
  Prostate cancer (path:hsa05215)
  Prion disease (path:hsa05060)
  Colorectal cancer (path:hsa05210)
Gene: BRCA1 (hsa:672):
  Ubiquitin mediated proteolysis (path:hsa04120)
```

dictyexpress.biolab.si





Selected Publications

Žagar et al. (2011) Stage prediction of embryonic stem cell differentiation from genome-wide expression data, *Bioinformatics*.

Tollervy & Curk et al. (2011) Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nature Neuroscience*.

Toplak et al. (2010) Does replication groups scoring reduce false positive rate in SNP interaction discovery? *BMC Genomics*.

Štajdohar et al. (2010) FragViz: visualization of fragmented networks, *BMC Bioinformatics*.

Parikh et al. (2010) Conserved developmental transcriptomes in evolutionarily divergent species. *Genome Biology*.

König et al. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Structural and Molecular Biology*.

Demšar (2006) Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*.

Van Driessche et al. (2005) Epistasis analysis with global transcriptional phenotypes. *Nature Genetics*.

Bioinformatics Lab

Janez Demšar
vice chair

Blaž Zupan, chair

