BURNHAM INSTITUTE for MEDICAL RESEARCH

From Research, the Power to Cure

JAFA

a protein function prediction meta-server

Tim Harder[1,2], Iddo Friedberg[2] and Adam Godzik[2]

[1]University of Applied Sciences Bingen, Bioinformatics,
Berlinstrasse 109, D-55411 Bingen, Germany
[2]Burnham Institute for Medical Research, Program in Bioinformatics and Systems Biology,
10901 North Torrey Pines Road, La Jolla, CA 92037, USA

FH BINGEN

Life Sciences and Engineering

Start a new session

To start a new JAFA Session, enter a protein sequence in the text field below. The following servers will be queried using default values. To change these values, click on the **[+]** sign next to the server you want to change values. Have fun.

☑ Include **GOblet** using user defined values **[+]** < 5 min

☑ Include **GOfigure** using default values **[+]** < 5 min

☑ Include **GOtcha** using default values **[+]** < 5 min

☑ Include **InterProScan** using default values **[+]** < 5 min

☐ Include **Phydbac** using default values *(Sorry, nothing to configure here)* 5-15 min

Common:
* Email:
Title:
* Sequence (FASTA format, or sequence only): Only protein sequences for now
Upload FASTA File: Durchsuchen...
* mandatory

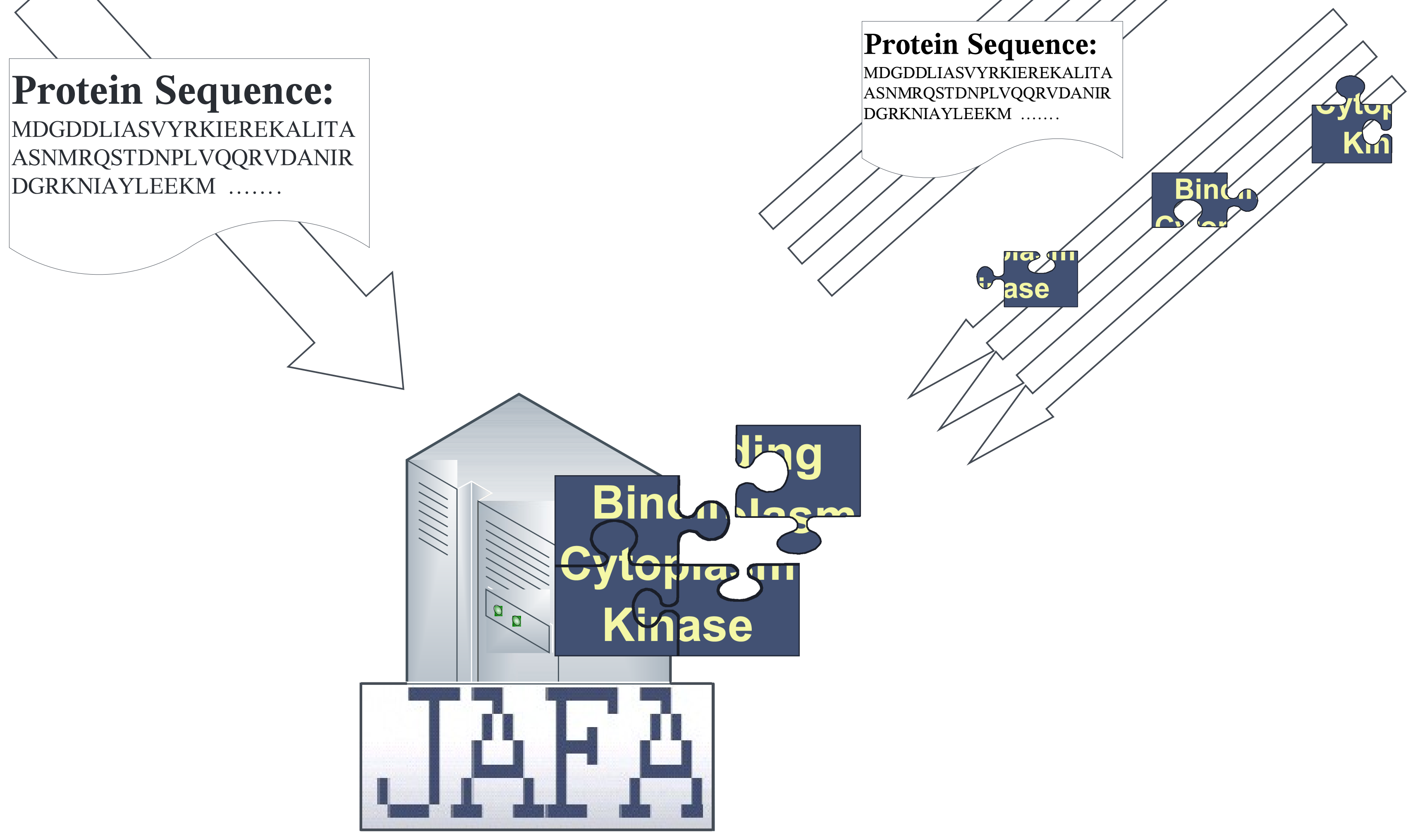
reset form | send request

JAFA Input:

- Enter a query sequence
- All parameters can be adjusted to the query

JAFA Run

- Query remote servers, wait for them to finish, receive results.
- Interim result are available, as soon as they are received
- Currently queried servers: Gotcha[3], Goblet[4], GoFigure[5], InterproScan[6], Phydbac[7]



Protein Sequence:
MDGDDLIASVYRKIEREKALITA
ASNMRQSTDNPLVQQRVDANIR
DGRKNIAYLEEKM …….

Protein Sequence:
MDGDDLIASVYRKIEREKALITA
ASNMRQSTDNPLVQQRVDANIR
DGRKNIAYLEEKM …….

Function Prediction is difficult …

- Biological function has different aspects: what happens, where in the cell, which process does this function belong to.
- Each method / program has it own strengths in predicting one or the other.

… but many different pieces build a picture.

- Combining the prediction of different methods / programs can provide a better understanding of the proteins function.
- Differences in each prediction lead to a broader over all picture
- Similar or identical predictions by different programs increase prediction confidence

Future Plans:

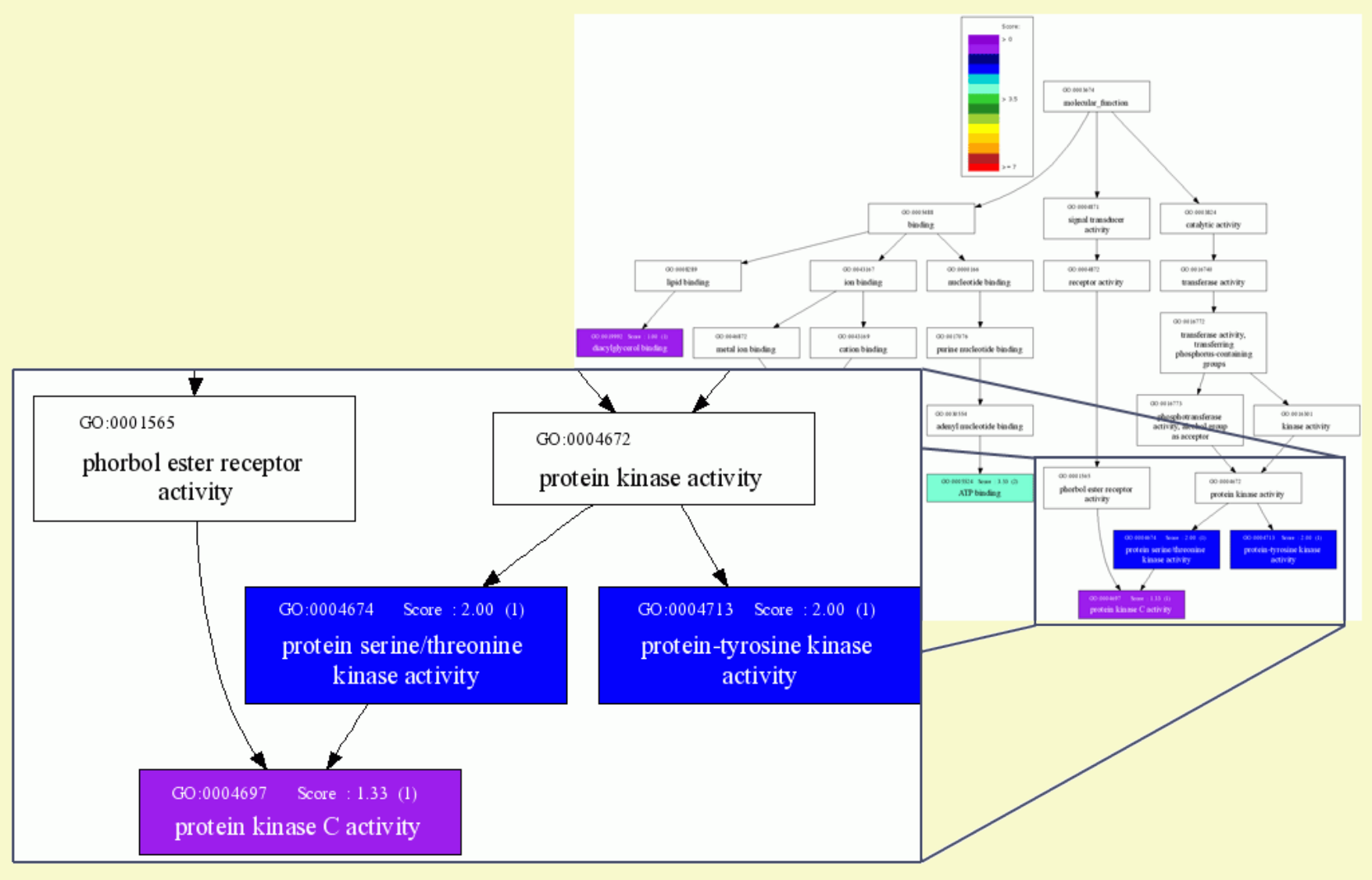
- A standalone version of JAFA, for use in high throughput research
- Create a standardized XML protocol for reporting function predictions
- Analysis of the different prediction programs capabilities will lead us to better overall predictions
- Improve scoring function to weight different programs based on known strengths and weaknesses

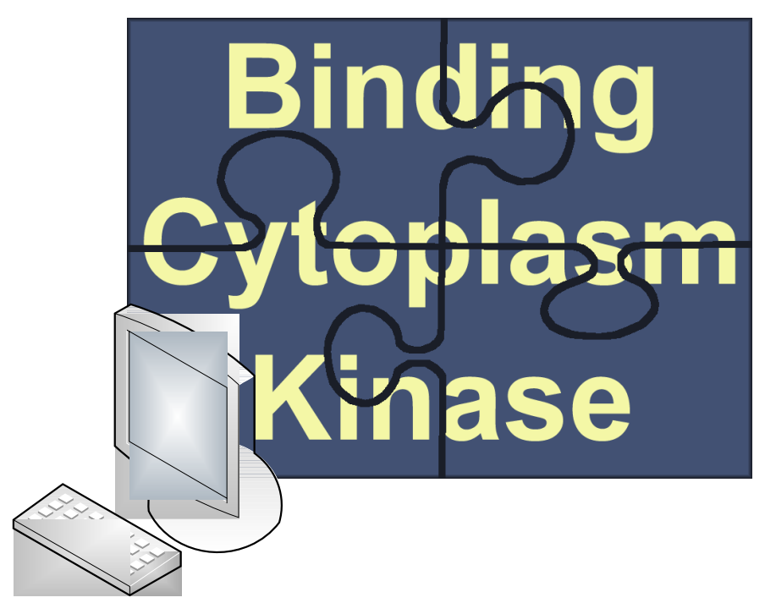
Gene Ontology (GO):

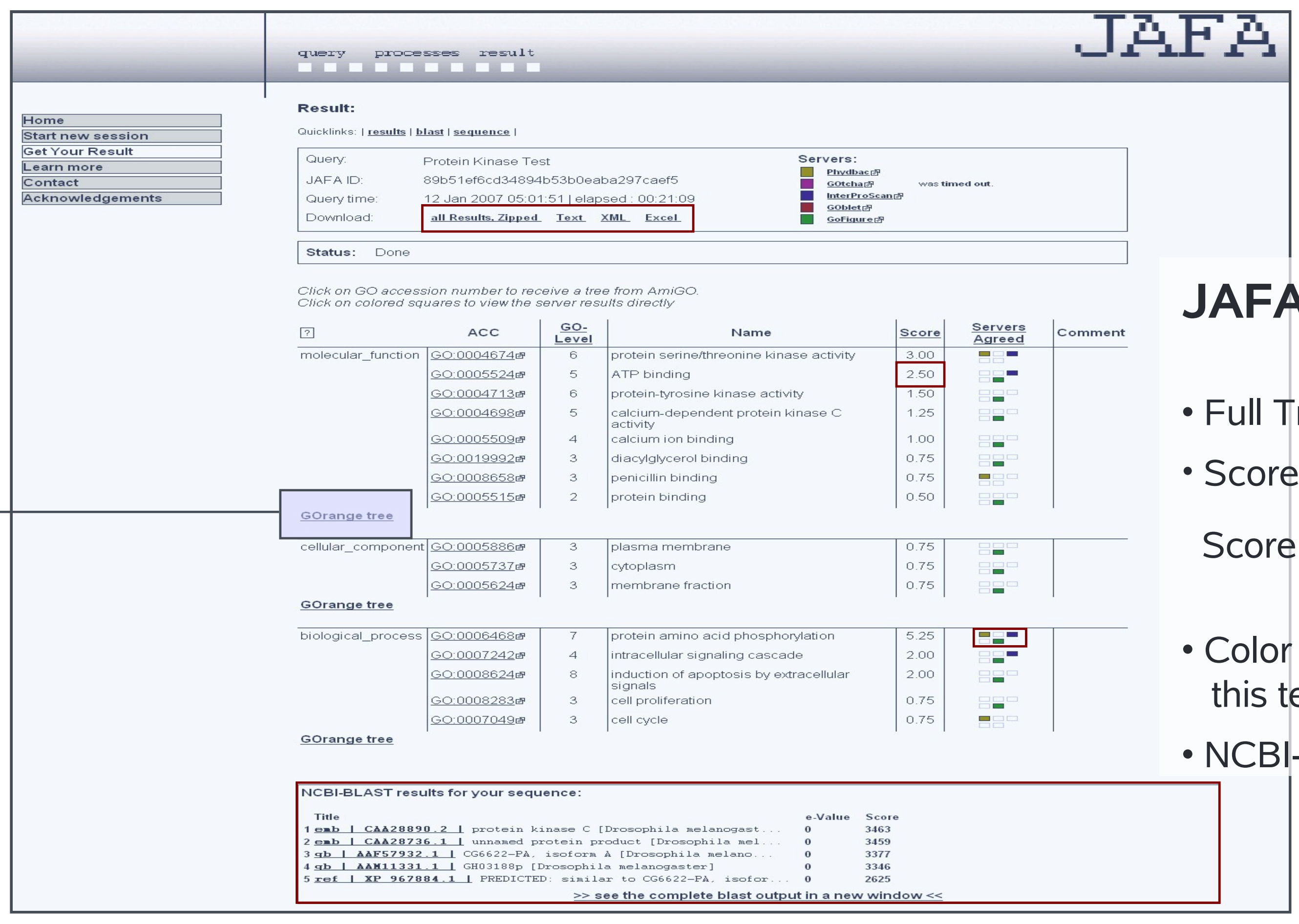
- Controlled vocabulary for the description of genes and gene product attributes
- Hierarchical: general to specific
- Different functional aspects: Biological Process, Molecular Function and Cellular Component
- http://www.geneontology.org [2]

JAFA - GOrange Tree:

- Predicted GO terms displayed hierarchically. Prediction confidence is color coded







Result:

Quicklinks: | results | blast | sequence |

Query: Protein Kinase Test
JAFA ID: 89b51ef6cd34894b53b0eaba297caef5
Query time: 12 Jan 2007 05:01:51 | elapsed : 00:21:09
Download: all Results, Zipped Text XML Excel

Servers:
🟨 Phydbac
🟪 GOtcha was timed out.
🟦 InterProScan
🟥 GOblet
🟩 GoFigure

Status: Done

Click on GO accession number to receive a tree from AmiGO.
Click on colored squares to view the server results directly

?	ACC	GO-Level	Name	Score	Servers Agreed	Comment
molecular_function	GO:0004674	6	protein serine/threonine kinase activity	3.00		
	GO:0005524	5	ATP binding	2.50		
	GO:0004713	6	protein-tyrosine kinase activity	1.50		
	GO:0004698	5	calcium-dependent protein kinase C activity	1.25		
	GO:0005509	4	calcium ion binding	1.00		
	GO:0019992	3	diacylglycerol binding	0.75		
	GO:0008658	3	penicillin binding	0.75		
	GO:0005515	2	protein binding	0.50		
cellular_component	GO:0005886	3	plasma membrane	0.75		
	GO:0005737	3	cytoplasm	0.75		
	GO:0005624	3	membrane fraction	0.75		
biological_process	GO:0006468	7	protein amino acid phosphorylation	5.25		
	GO:0007242	4	intracellular signaling cascade	2.00		
	GO:0008624	8	induction of apoptosis by extracellular signals	2.00		
	GO:0008283	3	cell proliferation	0.75		
	GO:0007049	3	cell cycle	0.75		

NCBI-BLAST results for your sequence:

```
  Title                                                          e-Value  Score
1 emb | CAA28890.2 |  protein kinase C [Drosophila melanogast...   0       3463
2 emb | CAA28736.1 |  unnamed protein product [Drosophila mel...  0       3459
3 gb | AAF57932.1 |  CG6622-PA, isoform A [Drosophila melano...    0       3377
4 gb | AAM11331.1 |  GH03188p [Drosophila melanogaster]           0       3346
5 ref | XP 967884.1 |  PREDICTED: similar to CG6622-PA, isofor... 0       2625
```

>> see the complete blast output in a new window <<

JAFA Results

- Full Traceback Information downloadable
- Scored GO Terms

$$\text{Score} = \frac{\text{number of predicting programs} \times \text{GO level}}{\text{total number of programs}}$$

- Color indication which server predicted this term
- NCBI-Blast [8] Run Results

Online:

JAFA is available online at **http://jafa.burnham.org**
Contact: **jafateam@gmail.com**

Acknowledgements:

JAFA was funded by NIH grant P01-GM63208.
A travel fellowship was generously granted by the International Society for Computational Biology, ISCB

ISCB INTERNATIONAL SOCIETY FOR COMPUTATIONAL BIOLOGY

References:

1. Friedberg I., Harder T. and Godzik A. (2006) JAFA: a Protein Function Annotation Meta-Server. Nucleic Acids Res. 2006 Jul 1;34(Web Server issue):W379-81.
2. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T. et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium Nat Genet, 25, 25-29.
3. Martin, D.M., Berriman, M. and Barton, G.J. (2004) GOtcha: a new method for prediction of protein function assessed by the annotation of seven genomes BMC Bioinformatics, 5, 178.
4. Groth, D., Lehrach, H. and Hennig, S. (2004) GOblet: a platform for Gene Ontology annotation of anonymous sequence data Nucleic Acids Res, 32, W313-317.
5. Khan, S., Situ, G., Decker, K. and Schmidt, C.J. (2003) GoFigure: automated Gene Ontology annotation Bioinformatics, 19, 2484-2485.
6. Zdobnov, E.M. and Apweiler, R. (2001) InterProScan--an integration platform for the signature-recognition methods in InterPro Bioinformatics, 17, 847-848.
7. Enault, F., Suhre, K. and Claverie, J.-M. (2005) Phydbac "Gene Function Predictor" : a gene annotation tool based on genomic context analysis BMC Bioinformatics, 6.
8. Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs Nucl. Acids Res., 25, 3389-3402.15

http://jafa.burnham.org