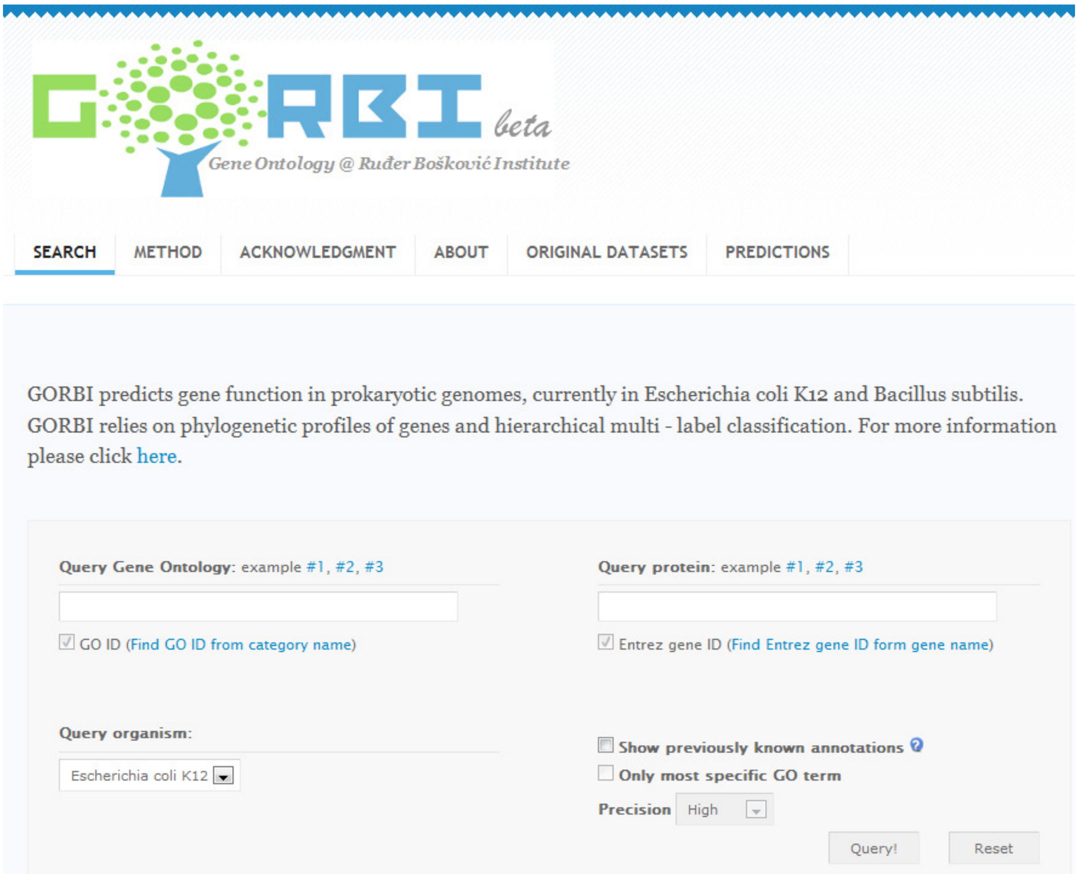
GORBI: Web application for the prediction of a protein's functional context

GORBI is an online database offering the results of computational gene function prediction in prokaryotic genomes. The analysis was done via the method of correlating gene occurrence patterns in selected organisms, termed phylogenetic profiling [1]. A machine learning algorithm based on decision trees for Hierarchical Multi-label Classification (HMC) [2] was used, and the annotations are represented using the Gene Ontology (GO) vocabulary.

When predicting GO assignments for genes, the HMC algorithm shows convincing predictive power. The web interface will be used to present the results in a clear and easily accessible manner, thereby covering two angles in wet lab research: focusing on 1) *function* (needing prediction for genes having that function in any organism), and 2) *organism* (needing prediction of any function in this particular organism). The layout of the interface is shown in the figure below. An early version of GORBI is freely available from http://gorbi.irb.hr/

The listing of function predictions together with the corresponding precision values, that is available, can be scanned, thus providing a condensed outline of target genes that are to be examined in the wet lab.



GORBI predicts gene function in prokaryotic genomes, currently in Escherichia coli K12 and Bacillus subtilis. GORBI relies on phylogenetic profiles of genes and hierarchical multi - label classification. For more information please click here.

Query Gene Ontology: example #1, #2, #3

☑ GO ID (Find GO ID from category name)

Query protein: example #1, #2, #3

☑ Entrez gene ID (Find Entrez gene ID form gene name)

Query organism: Escherichia coli K12

☐ **Show previously known annotations**

☐ **Only most specific GO term**

Precision High

1. Kensche, P.R. et al., Practical and theoretical advances in predicting the function of a protein by its phylogenetic distribution. *Journal of the Royal Society Interface*, 2008. **5**(19): p. 151-170.

2. Vens, C., et al. Decision trees for hierarchical multi-label classification. *Machine Learning*, 2008. **73**(2): p. 185-214.