BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences

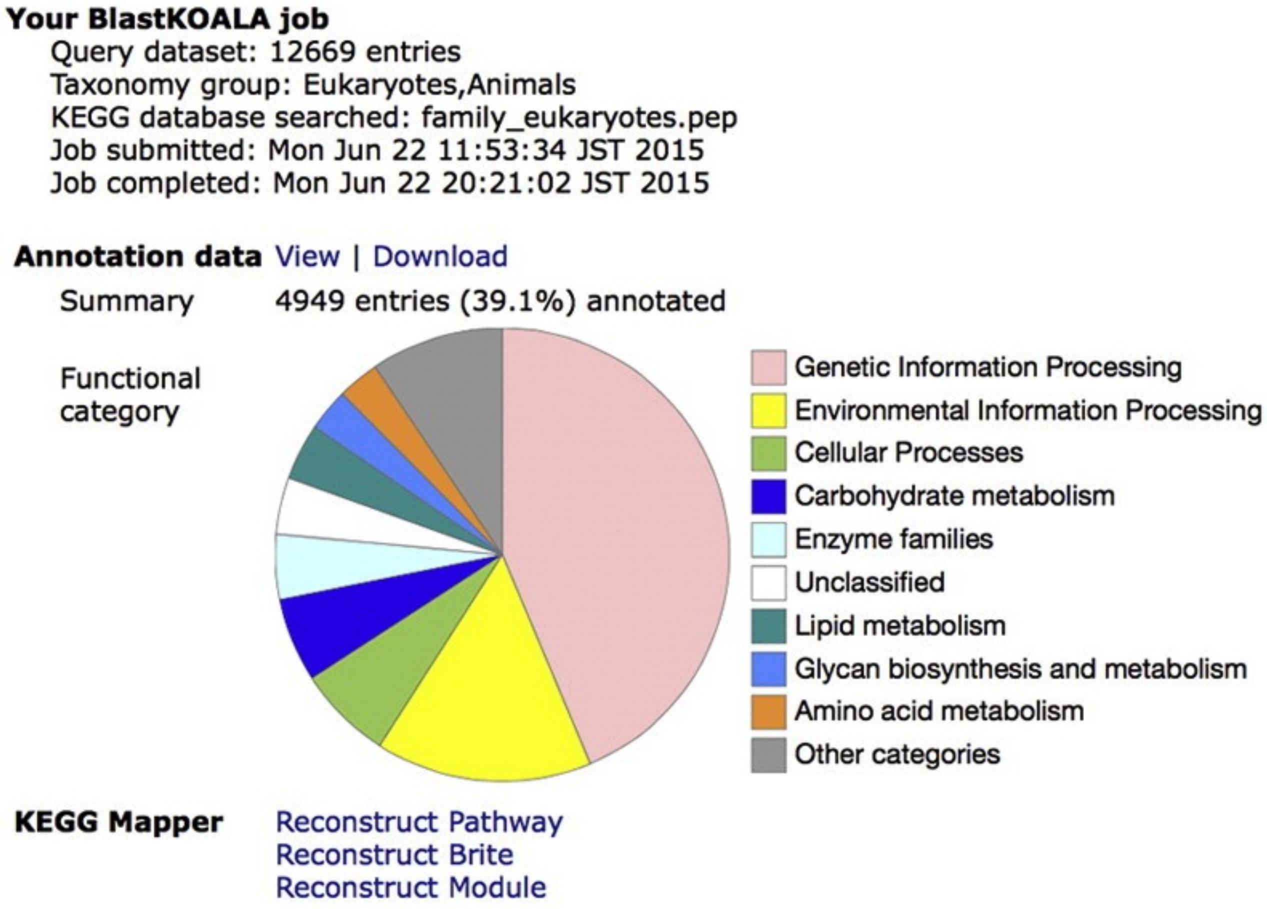
Minoru Kanehisa, Yoko Sato, Kanae Morishima

presented by
Bo Zheng, Franziska Hicking

Why do we need that

?

Annotation



Your BlastKOALA job

Query dataset: 12669 entries
Taxonomy group: Eukaryotes,Animals
KEGG database searched: family_eukaryotes.pep
Job submitted: Mon Jun 22 11:53:34 JST 2015
Job completed: Mon Jun 22 20:21:02 JST 2015

Annotation data View | Download

Summary 4949 entries (39.1%) annotated

Functional category

KEGG Mapper Reconstruct Pathway
Reconstruct Brite
Reconstruct Module

Figure 1: result by BlastKOALA

Minoru Kanehisa

KEGG

Kyoto Encyclopedia of Genes and Genomes

https://www.genome.jp/kegg/

BlastKOALA

- species, genus and family
- BLAST algorithm
- Dataset from KEGG GENES
- KOALA algorithm

GhostKOALA

- metagenome
- GhostX algorithm
- Dataset with Cd-Hit clusters of viral genes
- KOALA algorithm

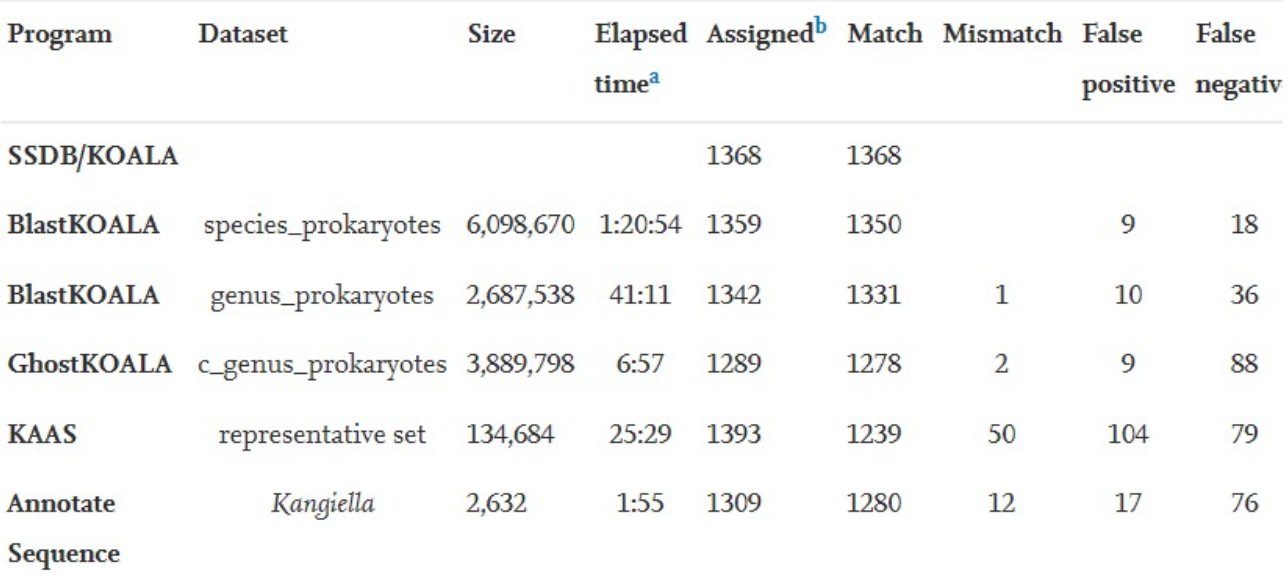
https://www.kegg.jp/blastkoala/

https://www.kegg.jp/ghostkoala/

Really works?

Any good?

Table 1: A comparison of K number assignments for the genome of *K. geojedonensis*.



Program	Dataset	Size	Elapsed time[a]	Assigned[b]	Match	Mismatch	False positive	False negativ
SSDB/KOALA				1368	1368			
BlastKOALA	species_prokaryotes	6,098,670	1:20:54	1359	1350		9	18
BlastKOALA	genus_prokaryotes	2,687,538	41:11	1342	1331	1	10	36
GhostKOALA	c_genus_prokaryotes	3,889,798	6:57	1289	1278	2	9	88
KAAS	representative set	134,684	25:29	1393	1239	50	104	79
Annotate Sequence	*Kangiella*	2,632	1:55	1309	1280	12	17	76

Thank you!

Bo Zheng   2021

Poster

SCAN
ME