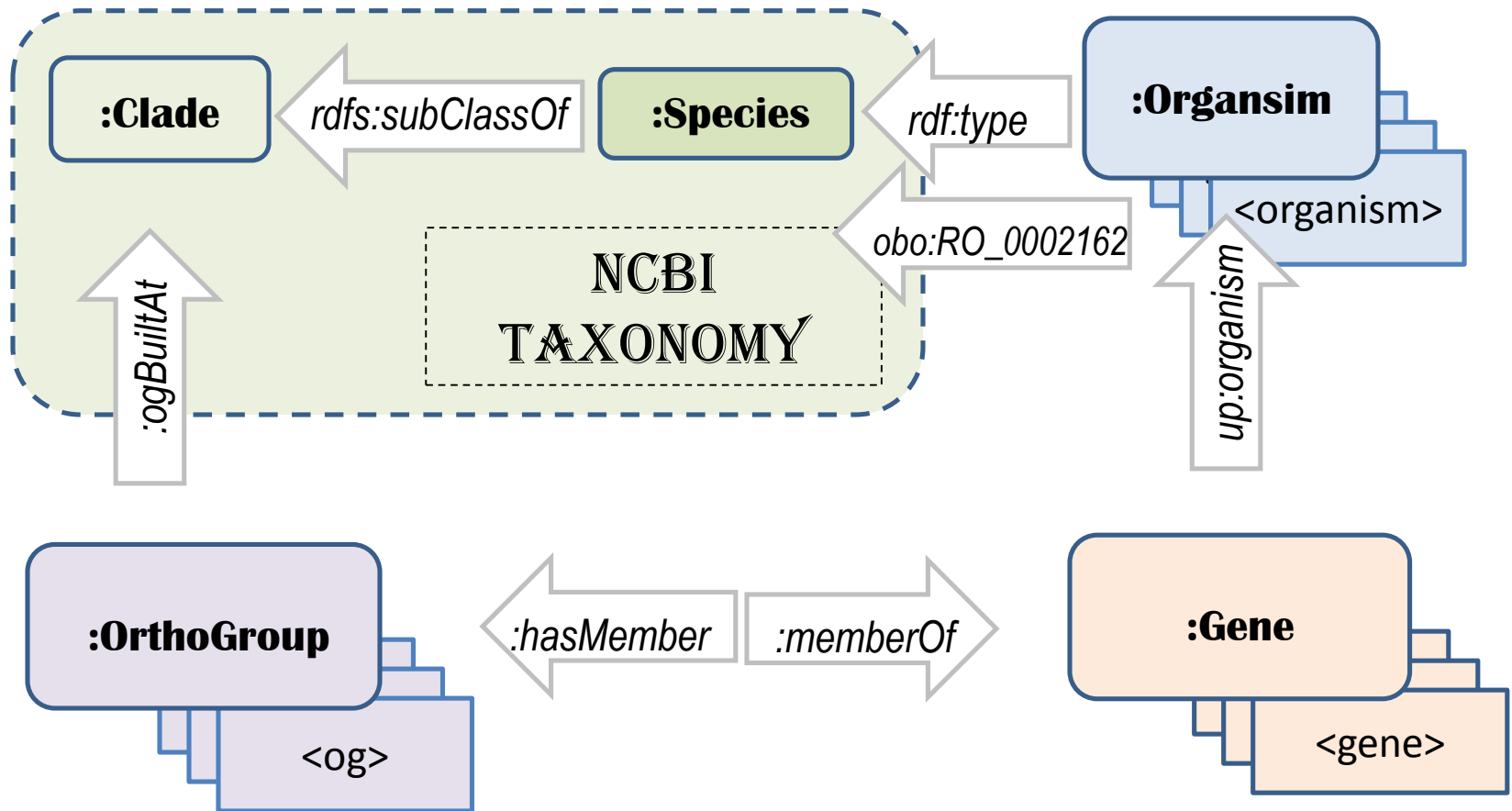


OrthoDB

Querying data using SPARQL

Genes and orthologous groups



Clades, Species and Organisms

Clades

?clade a :Clade; rdfs:subClassOf+ taxon:2; up:scientificName ?name.

Species, Organsims

?clade a :Clade; up:scientificName "Escherichia".

?taxon a :Species; up:scientificName ?**tx_name**; rdfs:subClassOf+ ?clade.

?org a :Organism,?taxon; up:scientificName ?**org_name**.

tx_name	org_name
"Escherichia albertii"	"Escherichia albertii, genome GCF_001549955.1"
"Escherichia coli Nissle 1917"	"Escherichia coli Nissle 1917, genome GCF_000714595.1"
"Escherichia coli"	"Escherichia coli, genome GCF_001617565.1"
"Escherichia coli"	"Escherichia coli, genome GCF_001900655.1"
"Escherichia coli"	"Escherichia coli, genome GCF_001900735.1"
"Escherichia coli"	"Escherichia coli, genome GCF_001900945.1"
"Escherichia fergusonii ATCC 35469"	"Escherichia fergusonii ATCC 35469, genome GCF_000026225.1"
"Escherichia coli O157:H16"	"Escherichia coli O157:H16, genome GCF_000827105.1"

Result IRIs for clade, taxon, org are functional URLs pointing to OrthoDB pages

Genes

Genes

?org a :Organism; up:scientificName "Escherichia coli, genome GCF_001617565.1".

?gene a :Gene; **up:organism** ?org; :name ?**gene_name**; :description ?**description**.

gene_name	description
"fucA"	"L-fuculose phosphate aldolase"
"alaS"	"Alanine--tRNA ligase"
"hemG"	"Molybdopterin-guanine dinucleotide biosynthesis protein B"
"SY51_RS02910"	"Lipoprotein"
"rbsA"	"Ribose import ATP-binding protein RbsA"
"WM90_RS12030"	"HCP oxidoreductase"
"gltB"	"Glutamate synthase"
"smpA"	"Outer membrane protein assembly factor BamE"
"WM90_RS24495"	"Addiction module toxin RelE"

Result IRIs for org, gene are functional URLs pointing to OrthoDB pages

Orthologous groups (OGs)

Orthogroups

```
select ?og ?og_name ?distance ?clade_name (count(1) as ?cnt)
where {
  ?org a :Organism; up:scientificName "Escherichia coli, genome GCF_001617565.1".
  ?gene a :Gene; up:organism ?org; :name "alaS"; :memberOf ?og.
  ?og :ogBuiltAt [up:scientificName ?clade_name; :taxTreeDistance ?distance];
  :name ?og_name; :hasMember ?gene2
} group by ?clade_name ?og ?og_name ?distance order by ?distance
```

og	og_name	distance	clade_name	cnt
http://purl.orthodb.org/odbgroup/91428at2	"Alanine--tRNA ligase"	0	"Bacteria"	5604
http://purl.orthodb.org/odbgroup/13856at1224	"Alanine--tRNA ligase"	1	"Proteobacteria"	2294
http://purl.orthodb.org/odbgroup/26372at1236	"Alanine--tRNA ligase"	2	"Gammaproteobacteria"	956
http://purl.orthodb.org/odbgroup/7453at91347	"Alanine--tRNA ligase"	3	"Enterobacteriales"	211
http://purl.orthodb.org/odbgroup/1091at543	"Alanine--tRNA ligase"	4	"Enterobacteriaceae"	98
http://purl.orthodb.org/odbgroup/132at561	"Alanine--tRNA ligase"	5	"Escherichia"	7

Result IRIs for org, gene are functional URLs pointing to OrthoDB pages

Hierarchy of OGs is a paraphrase of the gene in evolutionary retrospective

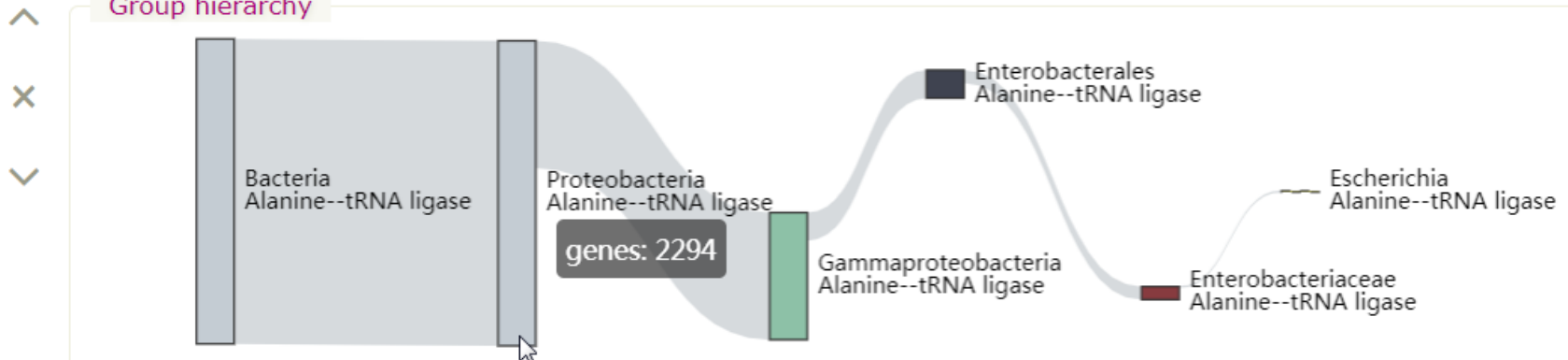
Group [132at561](#) at Escherichia level

[View Fasta](#) | [View Tab Delimited](#)

Alanine--tRNA ligase



Group hierarchy



Functional descriptions

Functional Category

J: Translation, ribosomal structure and biogenesis
L: Replication, recombination and repair
F: Nucleotide transport and metabolism
T: Signal transduction mechanisms



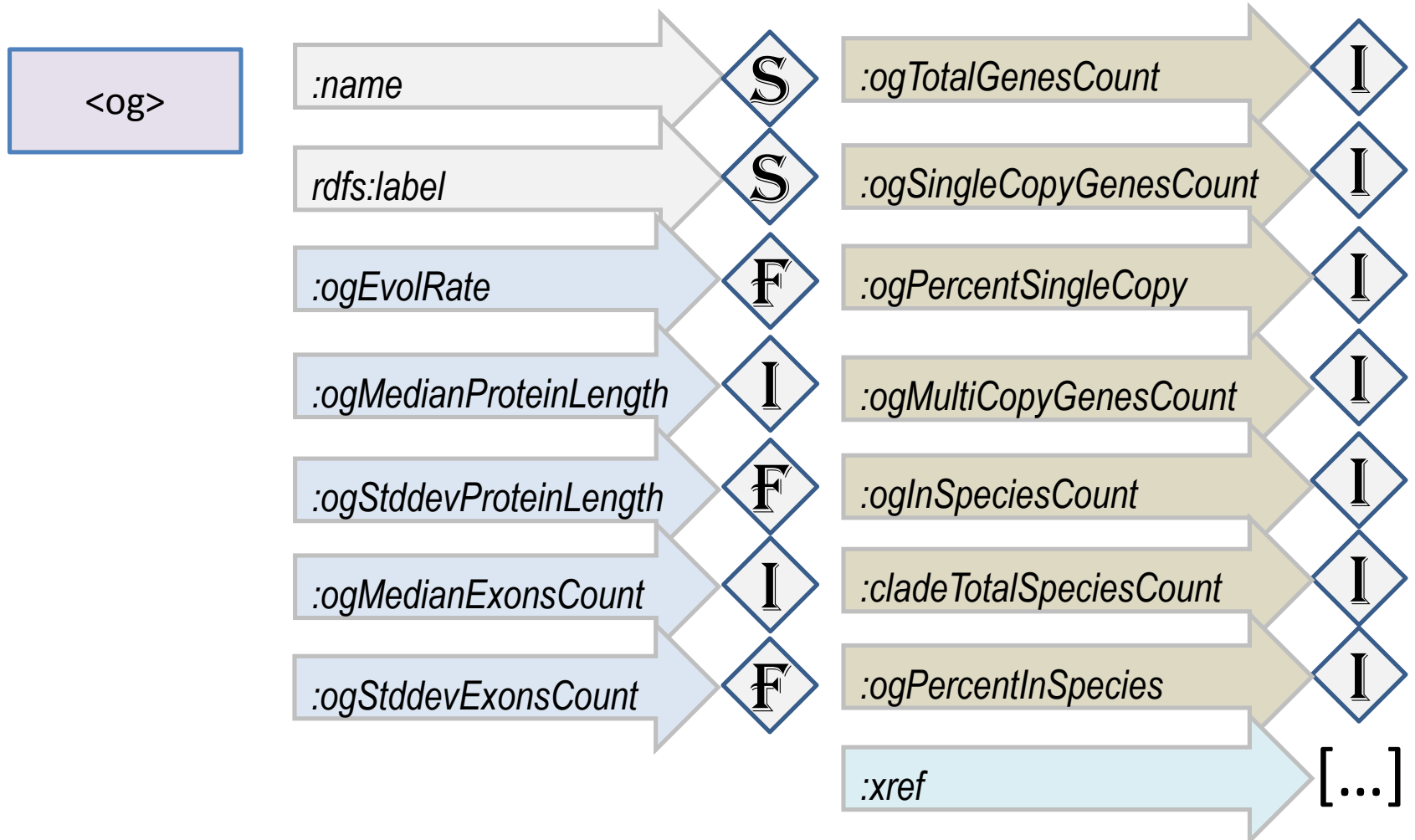
EC number

5 genes with [6.1.1.7](#): alanine--tRNA ligase; ATP + L-alanine + tRNAAla = AMP + diphosph

InterPro Domains

6 genes with [IPR002318](#): Alanine-tRNA ligase, class IIc
7 genes with [IPR003156](#): DHHA1 domain
7 genes with [IPR012947](#): Threonyl/alanyl tRNA synthetase, SAD

Orthologous group predicates



External references, aka xrefs

