JGI JOINT GENOME INSTITUTE
UNITED STATES DEPARTMENT OF ENERGY

U.S. DEPARTMENT OF ENERGY
Office of Science

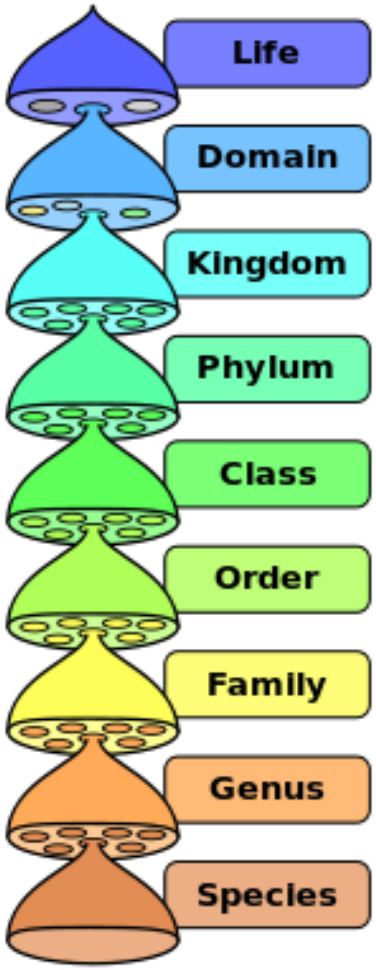
Taxonomy Quick Guide

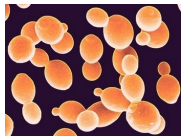
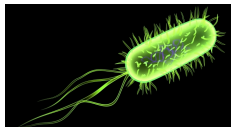
Objective

- Help ease the process of providing appropriate names to organisms to be sequenced by the JGI and identify potential errors in naming early.
- This guide provides an overview of organism taxonomy for isolates and single cells
- We will describe the process for identification of inconsistencies between sequencing project names and organism names and how to address these issues.

Taxonomy 101

Taxonomy is the science of defining groups of biological organisms on the basis of shared characteristics

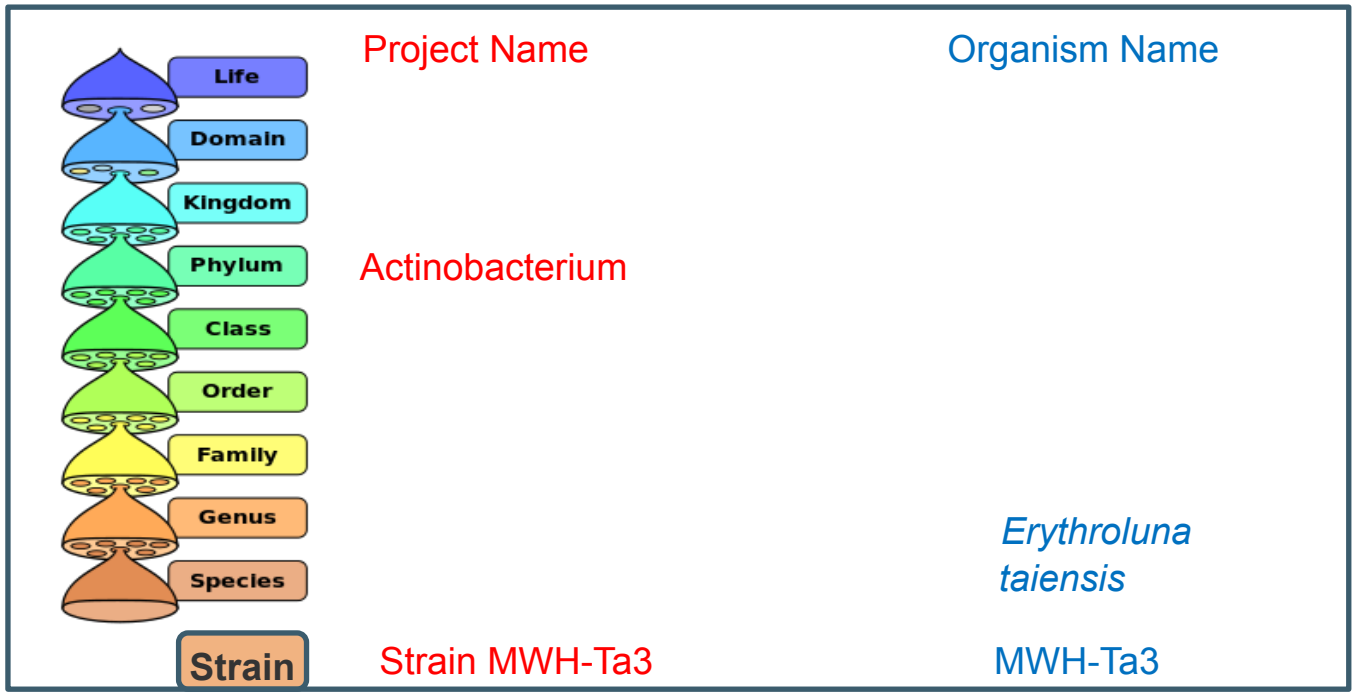


 **Yeast**	 ***E. coli***
Eukarya	Prokaryotes
Fungi	Bacteria
Ascomycota	Proteobacteria
Saccharomycetes	Gammaproteobacteria
Saccharomycetales	Enterobacteriales
Saccharomycetaceae	Enterobacteriaceae
Saccharomyces	*Escherichia*
S. cerevisiae	*E. coli*

Verifying Isolate Projects

- **Another example:**
  - The project name field should be the same as the organism name

ITS Project ID: 1104880:
Project Name: Actinobacterium Strain MWH-Ta3
Organism name: *Erythroluna alpina* MWH-Ta3



More things to look out for….

- **Genus/species field is left blank**
  - Possible cause: Investigator is sequencing a new organism and is unsure of the identity
  - Resolution:
    - Investigator must provide some information for higher taxonomic levels (e.g. Gammaproteobacteria sp. 123)
    - ➢ In this case, the organism and project name will be “Gammaproteobacteria sp. 123”. The GOLD convention is to leave the “genus” field blank since this is not well defined, populate the entire name in the “species” field and populate the “strain” field
    - If the investigator is proposing a **new genus**, we accept it without a publication with the understanding that if/when the project is registered with NCBI, they may not accept the new genus without an associated publication and suggest alternative names until a publication is available describing the genus
- **Strain field is left blank (this is a RED FLAG)**
  - Possible cause: Investigator is sequencing a new organism and is unsure of the identity
  - Resolution:
    - For bacterial and fungal projects, the strain is used to distinguish between two organisms within the same genus/species and cannot be left blank. The investigator must specify a strain ID.
    - If the organism has a repository ID (culture collection ID), that can be used as the strain ID

Culture Collections

Repository ID (or Culture Collection ID): Culture collection centers are repositories for organisms & reference materials.
Examples include:

DSMZ

ATCC®

NCTC
National Collection of Type Cultures
Operated by Public Health England

NRRL

THE CHINESE GENERAL MICROBIOLOGICAL CULTURE COLLECTION CENTER

Verifying Isolate Names

- **If you have doubts about a name, you can consult StrainInfo**

http://www.straininfo.net/

Taxon Passport
Escherichia coli

overview

species	*Escherichia coli*
parent taxon	*Escherichia sp.*
type strain	ACM 1803 T, AJ 2617 T, ATCC 11775 T, BCRC 10675 T, BTCC U5/41 T, CAPM 6101 T, CCM 5172 T, CCRC 10675 T, CCT 0547 T, CCT 1357 T, CCTM 2067 T, CCTM La 2067 T, CCUG 24 T, CCUG 29300 T, CDBB 964 T, CECT 515 T, CGMCC 1.2389 T, CGMCC 1.2993 T, CIP 54.8 T, CIP 54.8T[T], CN 4382 T, CNCTC 6859 T, CNCTC Eck 206/59 T, CNCTC Eck 58/59 T, DSM 30083 T, DSMZ 30083 T, ELI 50 T, F. Oerskov T, FIRDI 675 T, GISK 240001 T, IAM 12119 T, ICMP 15663 T, JCM 1649 T, K. Lincoln T, Kauffmann U5/41 T, KCTC 2441 T, LMD 54.8 T, LMG 2092 T, LRA 73.08.009 T, NBIMCC 3398 T, NBRC 102203 T, NCAIM B.01874 T, NCCB 54008 T, NCDO 1989 T, NCFB 1989 T, NCIB 11943 T, NCIMB 11943 T, NCTC 9001 T, NCTC 9001 U5/41[T], NCTC 9001.[T], NZRM 3309 T, PCM 172 T, PCM 321 T, R. Sakazaki T, SSIC U 5/41 T, strain U 5/41 T, U 5-41 T, U5/41 T, USCC 2054 T, USCC 2520 T, VTT E-94564 T, WDCM 00090 T, WDCM 00155 T, Y. Kosako 82039 T
16S rRNA gene	AB681728 (LTP: X80725)
external links	Catalogue of Life, LPSN; J.P. Euzéby, NCBI, WikiSpecies
genome projects	Escherichia coli APEC O1 Escherichia coli O157:H7 str. Sakai Escherichia coli UTI89 Escherichia coli CFT073 Escherichia coli str. K-12 substr. MG1655 Escherichia coli 536 Escherichia coli str. K-12 substr. W3110

Verifying Isolate Names

- **You can also consult the NCBI Taxonomy Browser**

http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/index.cgi

Escherichia coli DSM 30083 = JCM 1649 = ATCC 11775

Taxonomy ID: 866789
Inherited blast name: **enterobacteria**
Rank: no rank
Genetic code: Translation table 11 (Bacterial, Archaeal and Plant Plastid)
Other names:
synonym: **Escherichia coli JCM 1649**
synonym: **Escherichia coli DSM 30083**
synonym: **Escherichia coli CIP 54.8**
synonym: **Escherichia coli ATCC 11775**

Lineage(full)
cellular organisms; Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia; Escherichia coli

Verifying Isolate Names

- **Another resource is DSMZ**

https://www.dsmz.de/bacterial-diversity/helpful-links-regarding-taxonomy.html

BACTERIA	How to read the following data (Example)
Name:	***Escherichia coli*** (Migula 1895) Castellani and Chalmers 1919
DSM No.:	**30083, Type strain**
Strain designation:	U5/41
Other collection no. or WDCM no.:	ATCC 11775, CCM 5172, CIP 54.8, IAM 12119, JCM 1649, NCDO 1989, NCTC 9001, WDCM 00090
Serovar:	O1:K1:H7
Isolated from:	urine
History:	<- ATCC <- NCTC <- F. Kauffmann, U5/41
Cultivation conditions:	Medium 1, 37°C Complete DSMZ Media List

Key Points

- The organism name should be a part of the project name (at least the stem)
- Both genus and species fields cannot be empty. The investigator must provide some taxonomic identifier for the sample to be sequenced.
- Strain fields for isolates cannot be empty
- If you have questions about naming, consult your JGI project manager for assistance prior to submission of project metadata

JGI JOINT GENOME INSTITUTE
UNITED STATES DEPARTMENT OF ENERGY

Sample and metadata submission information can be found at:

http://jgi.doe.gov/user-program-info/pmo-overview/project-materials-submission-overview/