

## Standardized Metagenome-Assembled Genome (MAG) naming in GOLD

To create Metagenome-Assembled Genome (MAG) analysis project correctly, you need to populate all the required fields in the MAG creation form properly. The correct naming is important because MAGs are treated as organisms. That is why we require you to provide taxonomy information and populate fields such as Species, Strain, and, in some cases, Genus, if known.

The species and strain fields are mandatory; if the species is not known then you can use either the known **Family** or **Order** or **Class** or **Phylum** name to populate the Species field such as:

*Enterobacteriaceae bacterium ABC123 (family is known)*

*Enterobacteriales bacterium ABC123 (order is known)*

*Gammaproteobacteria bacterium ABC123 (class is known)*

*Proteobacteria bacterium ABC123 (phylum is known)*

We accept MAGs with the taxonomy known at least down to the Phylum level. Phylum cannot be undefined. The NCBI Taxonomy ID must be consistent with the name. Finally, each MAG must have a distinct identifier to differentiate it from other MAGs. As an identifier, you may use a bin number or anything unique that would separate one MAG from another.

### Examples of how to name a MAG AP in GOLD

#### Case 1: You know the taxonomy of your MAG down to the genus level:

Name your MAG using the following format:

**[Genus] sp. [identifier]**

Examples:

Olsenella sp. RUG778

Acidipila sp. ABC123

Catabacter sp. bin 77

If you know the genus of your MAG, then the field Genus MUST be populated, and NCBI Taxonomy ID should point to the Genus.

Example of the correctly populated fields:

**Analysis Project Name:** Olsenella sp. RUG778

**Genus:** Olsenella

**Species:** Olsenella sp. RUG778

**Strain:** RUG778

**NCBI Taxon ID:** 133925

**Case 2: You DON'T know the genus:**

Name your MAG using one of the formats below:

***Unclassified [Rank] [identifier]***

***[Rank] bacterium [identifier]***

***[Rank] archaeon [identifier]***

Examples:

*Proteobacteria bacterium FL.bin\_25*

*Unclassified Clostridiales ABC123*

*Unclassified Bacteroidales bin 12*

*Prevotellaceae bacterium RUG849*

*Thaumarchaeota archaeon SP3992*

Please note that in this case, the field Genus should be left blank, and the NCBI Taxonomy ID points to the known taxon.

Example of correctly populated fields:

**Analysis Project Name:** Unclassified Clostridiales ABC123

**Genus:**

**Species:** Unclassified Clostridiales ABC123

**Strain:** ABC123

**NCBI Taxon ID:** 186802

**Examples of unacceptable and incorrect MAGs' names:**

JalPR\_LakeR bin 3

Bacteroidales\_bin12

Acidipila sp.ABC123 [no space between 'sp.' and an identifier]

Acidipila sp ABC123 [no period after 'sp']