## Description for dataset genomics

## Data provided by: Renaud Berlemont

Strain: organism name

**N** Gene tot: total number of predicted gene (see bioinformatics BIOL474 if interested!)

**DNApol\_III\_a\_sub**: number of gene encoding the DNA polymerase 3 alpha subunit. This enzyme is essential for DNA replication. You would expect at least one copy of this gene per genome.

**tRNA\_synthetase**: number of gene encoding the tRNA-synthetase. These are the enzymes that attach the amino acid to the corresponding tRNA.

**X16rRNA**: number of gene encoding a non-translated RNA involved in the production of the ribosome (for protein production).

**cellulase**: number of accessory genes encoding enzymes that degrade cellulose. These are genes encoding enzymes that degrade cellulose. Cellulose from the plant cell wall is the most abundant source of organic carbon on Earth.

**Blactamase**: number of accessory genes encoding the microbial enzymes involved in the hydrolysis (inactivation) of the B-lactam antibiotics (e.g., penicillin, ampicillin, cephalotin, etc).