Conducting bioinformatics workflow to produce experimental Ribo-Seq density maps for Neurospora crassa.

Ribo-Seq density maps on transcripts for ADAT2-silenced versus wild type Neurospora crassa

**The reference paper is :**

Lyu X, Yang Q, Li L, Dang Y, Zhou Z, Chen S, Liu Y. Adaptation of codon usage to tRNA I34 modification controls translation kinetics and proteome landscape. PLoS Genet. 2020 Jun 1;16(6):e1008836. doi: 10.1371/journal.pgen.1008836. PMID: 32479508; PMCID: PMC7289440.

**Dataset availability :**

**NCBI Gene Expression Omnibus accession number GSE130155**

The aim of the workflow is to get the detailed ribosome profiles (at codon resolution) for the ORFs of a number of transcripts (suggested list provides hereafter). It would be helpful to get for each of the transcript (ORF) a tab delimited text file including absolute codon occupancy values in 2 cases:

The 2 cases are : control/treatment samples:

* wild type WT=CONTROL --- https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM3732951
* ADAT2 silenced (Siadat2)=TREATMENT

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM3732952>

The complete SRA reference files are provided on the GEO accession number in NCBI.

**Suggested minimal list of transcripts/gene cds /ORFs for which the ribosome footprint codon occupancy maps are to be produced:**

* If available : codon manipulated sequences of Firefly Luciferase (in vitro luciferase reporter assay)
* Beta tubulin : NCU04054
* CPC-1 cross-pathway control protein (GCN4 yeast ortholog): NCU04050
* Ribosomal proteins of the large ribosome subunit (RPL):
* Selected transcripts from the gene list given in table S1 of the reference paper:

Ribosome: (downregulated proteins upon adat2 silencing):

NCU06047, NCU03393, NCU08389, NCU04331, NCU04552, NCU08620, NCU07562, NCU00618, NCU03757, NCU00475, NCU06661, NCU03565, NCU01948, NCU01949, NCU06431, NCU06432, NCU03150, NCU08990, NCU08627, NCU09109, NCU06843, NCU00489, NCU03703, NCU01776, NCU06743, NCU05804, NCU08960, NCU08963, NCU03988, NCU08964, NCU01221, NCU07014, NCU09089, NCU08344, NCU06892, NCU03806, NCU00413, NCU00258, NCU02905, NCU06210, NCU05599, NCU01827, NCU00634, NCU03738, NCU03635, NCU01317, NCU00315, NCU01966, NCU04779, NCU00294, NCU02181, NCU05554, NCU02509, NCU03302, NCU08502, NCU07182, NCU02707, NCU07830, NCU08500, NCU09476, NCU00971, NCU07826, NCU09475, NCU00979, NCU06469, NCU02744, NCU10498, NCU07829, NCU03102, NCU03038, NCU00706, NCU07857, NCU01552, NCU07408, NCU01452, NCU06226

Biosynthesis of amino acids: (upregulated proteins upon adat2 silencing):

NCU04754, NCU04277, NCU01667, NCU01666, NCU03425, NCU00409, NCU01429, NCU06974, NCU00567, NCU09116, NCU07725, NCU04411, NCU02004, NCU03608, NCU05420, NCU07987, NCU08411, NCU03010, NCU04579, NCU07941, NCU00200, NCU00106, NCU01439, NCU07982, NCU05410, NCU00554, NCU05805, NCU07608, NCU05129, NCU10468, NCU05526, NCU03118, NCU07156, NCU03139, NCU06023, NCU04118, NCU08409, NCU04292, NCU02505, NCU02010, NCU09230, NCU08162, NCU02785, NCU01300, NCU02430, NCU05548, NCU09320, NCU01412, NCU02639, NCU09817, NCU06360, NCU05683

**A fasta file of the cds (ORF) (mRNA) sequences of the Neurospora crassa transcripts could be useful (downloaded from Ensembl):**

<https://ftp.ebi.ac.uk/ensemblgenomes/pub/release-58/fungi/fasta/neurospora_crassa/cds/Neurospora_crassa.NC12.cds.all.fa>

A correspondence can be done between the NCU gene ids and the cds transcripts of Neurospora crassa from the above fasta file.

The tRNA-seq and mRNA-seq reads that mapped uniquely to the reference with at most two mismatches were considered. The tRNA and mRNA references were from GtRNAdb(http://gtrnadb.ucsc. edu/GtRNAdb2/genomes/eukaryota/Neur\_cras\_OR74A/) and GenBank (https://www.ncbi. nlm.nih.gov/bioproject/13841), respectively.

## Neurospora crassa OR74A

### Neurospora crassa OR74A RefSeq Genome

See [Genome](https://www.ncbi.nlm.nih.gov/genome/19) Information for Neurospora crassa

Navigate Across

[3175 additional projects](https://www.ncbi.nlm.nih.gov/bioproject?term=9522%5btop+bioproject%5d+NOT+132%5buid%5d) are related by organism.

The reference sequence (RefSeq) genome assembly is derived from the submitted GenBank assembly (see linked project PRJNA13841). Annotation on the RefSeq genomic records is propagated from the INSDC submission. RefSeq records may differ from the submitted annotation in format details, some protein names, and occasionally in feature location details.