# Mining structural patterns in non-coding RNAs related to virus infections

BSc project supervisor Sacha Gultyaev (LIACS/ErasmusMC)

- Recent data indicate that a number of non-coding RNAs (ncRNAs) could regulate virus replication.
- Such functions may depend on ncRNA structural features.
- The project goal is to identify potential structures and/or consensus motifs in ncRNAs that play a role in replication of viruses.
- Methods: sequence database similarity search, multiple sequence alignment, thermodynamics-based RNA structure predictions.

An example of secondary structure model for a ncRNA that play a role in influenza virus replication [Winterling et al., 2014].

Colors indicate the estimated probabilities of base pairs in the structure.
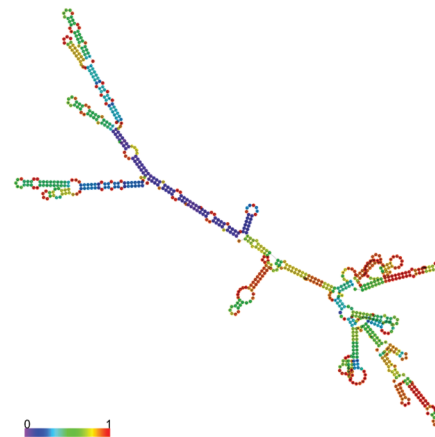
Identification of conserved RNA structures encoded in DNA sequences used for barcoding in metagenomic projects

BSc project supervisor Sacha Gultyaev (LIACS)
(together with the group Understanding Evolution, Naturalis Biodiversity Center)

- Specific sequences (amplicons) are used for identification of species in metagenomic projects, e.g. fungal communities.
- Detection and classification of species is crucially dependent on alignment of amplicons.
- Frequently used amplicons: so-called internal transcribed spacers (ITS) in the clusters of ribosomal RNA genes.
- Alignment of these sequences is difficult due to considerable sequence diversity and the lack of information on conserved motifs in encoded RNA structures.
- The project goal is to identify conserved structures in ITS sequences.
- Methods: sequence database similarity search, multiple sequence alignment, thermodynamics-based RNA structure predictions.

Organization of the ribosomal RNA gene locus [from Woolford & Baserga, 2013]. ITS1 and ITS2 are important amplicon sequences in the metagenomic barcoding.

Predicted structure of an ITS2 sequence from a fungus species. Length: 264 nucleotides.

5'ETS 18S ITS1 5.8S ITS2 25S 3' ETS