Project title: **Identification and analysis of tRNA-like sequences from *Arabidopsis thaliana***

Transfer RNAs (tRNA) are one of the best recognized essential classes of non-coding RNAs (ncRNAs) present in all living organisms. Their involvement in translation as a carrier of amino acids for the ribosomal biosynthesis of proteins and a link between the trinucleotides of the genetic code and protein sequences suggests that it is also one of the oldest molecular components of the cell.

However, the functions of tRNA molecules are not limited to the process of translation. There is an increasing body of evidence that suggests that both canonical tRNAs as well as molecules possessing certain structural features of tRNA (tRNA-lookalike or tRNA-like molecules) may also be involved in other functions in the cell including signalling, degradation, transport, modifications of biomolecules and as substrates in various biosynthetic pathways. Currently, one of the most intensively studied alternative function of tRNAs is their ability to be processed into smaller fragments (tRNA-derived small RNAs; tRFs). For many years these short RNAs were considered to be merely degradation products, but in recent years they were recognized as factors involved in important regulatory and biogenesis processes.

RNA species that resemble canonical tRNA structure can be recognized by cellular proteins involved in tRNA biogenesis, including ribonucleases responsible for pre-tRNA maturation. In several cases it has been demonstrated that the tRNA-like sequences within genomic DNA can provide control elements for transcription by RNA polymerase III. On the other hand, tRNA-like structures within the primary transcripts may supply signals for maturation, involving hydrolysis by RNase P or RNase Z associated with processing of 'true' tRNAs. However, the roles of these tRNA-like molecules and the extent of such phenomena is not yet fully characterized.

Based on currently available data, we postulate that the methods used for tRNA identification and analysis cannot recover the full spectrum of tRNA-like molecules in *Arabidopsis thaliana*. Taking into account a plethora of known as well as putative biological functions of tRNA and tRNA-like molecules, this leaves uncharted an important part of molecular machinery of the cell. Therefore, to understand the function of these molecules, it is essential to delineate and characterize them on a genome-wide scale. Thus, we propose to comprehensively identify and characterize the full spectrum of tRNA and tRNA-related molecules in a model plant *Arabidopsis thaliana*.

To accomplish this goal, we will deploy two strategies of tRNA identification: computational and experimental. The computational approach will involve whole-genome scanning for structural properties of well-defined tRNAs with specially designed algorithm trained to allow high flexibility in detection parameters. The experimental methods will employ high-throughput sequencing technology (NGS) with RNA preparation techniques specifically aimed to enrich the sequence space with functional (amino-acid-caring) tRNAs and / or possessing characteristic sequence modifications (i.e. the CCA 3'-tail). Additionally, we will explore data coming from specific enrichment experiments for detection of ribosomal associations, previously obtained results (small RNAs from *A. thaliana* RNA biogenesis mutants) as well as sequences from public databases (SRA).

The whole-genome identification of the tRNA and tRNA-like sequences will trigger additional level of analyses that will involve their genomic context evaluation and comparative genomics studies.

By characterizing the tRNA-space of the model plant *Arabidopsis thaliana* we will supplement the knowledge about its non-coding RNA component on the identification and functional characterization level. Moreover, all the results will be incorporated into redesigned tRex database (http://combio.pl/trex), primarily dedicated to tRNA-derived fragment analysis, that will be now transformed into public resource for tRNA biology in *Arabidopsis thaliana*.