Project title: The role of adenosine demethylases in regulation of gene expression and cellular metabolism in human cells

Student: Praveenkumar Rengaraj
From: Stepanka Vanacova, CEITEC-MU
To: Vladimir Benes and bioinformatics team, EMBL

Summary

The role of m^6A RNA modification has been studied to great extent in the past 10 years. One of the major attractiveness was initially based on the existence of specific eraser(s) and thus its potential dynamic property. FTO and ALKBH5 have been linked to various regulatory roles on the cellular and organismal level, and to defects in development and disease. This project aims on tackling these questions by systems genetic screens, RNA binding and transcriptome wide modification analyses to uncover the involvement and role of the individual enzymes in the physiology of human cells, such as cellular signalling, proliferation or differentiation. For that we need to develop robust bioinformatics and statistical analyses which will be done in collaboration with Vladimir Benes and the bioinformatics team at EMBL. Results from the proposed project, will significantly contribute to a better understanding of the roles of m^6A and m^6A_m demethylation in cellular homeostasis and differentiation.

Scientific background

Post-transcriptional modifications of coding and non-coding RNAs are a critical part of the regulation of gene expression in eukaryotic cells. The most frequent mRNA marks are the internal N6-methyladenosine (m^6A) and the cap-adjacent m^6A_m that have progressively gained on importance in the field. Mammalian cells possess a dedicated cellular machinery to write, erase and read m^6A and m^6A_m marks. At least two distinct methyltransferases (MTs, writers) can deposit internal m^6As (METTL3/14 complex and the METTL16) and PCIF1 mediating cap-adjacent m^6Am (Liu 2014, Boulias 2019, Akichika 2019, He 2021). The levels of m^6A and m^6A_m in the cells are fine-tuned by the action of two demethylases (DMTs, erasers): FTO and ALKBH5 (Jia 2008, Zheng 2013). The activity of MTs and DMTs determines the *N6*-methylation status of mRNAs and ncRNAs in the cell and, in turn, their metabolism and function. Several pieces of evidence support a model of co-transcriptional m^6A and m^6A_m dynamics (Akichika 2019, Baltz 2012, Slobodin 2017). METTL3 and CAPAM are recruited to active chromatin by RNAPII (Akichika 2019, Slobodin 2017) and specific chromatin marks guide m^6A deposition (Huang 2019). In turn, m^6A modulates gene expression via regulation of histone modifications (Li 2020). Compared to MTs, recruitment and regulation of DMTs remains largely unknown. They both share a link to pre-mRNA splicing and nuclear localization, however, they differ in the expression pattern and KO mouse phenotypes (Li 2020, Tang 2018). Despite the fast developments in the field, the transcript and mark specificity, mechanism of their regulation and involvement in cellular physiology remains to be addressed for both DMTs.

Experimental approach and aims

1. Identification of direct RNA targets and their dependency on methylation marks

The depletion of ALKBH5 and FTO, respectively exhibits mild increase of m^6A in mRNAs (Mauer 2018) and our unpublished data). It thus remains unclear whether the m^6A demethylation activity is the primary role of these two proteins in vivo. It is not known to what extent is the RNA binding and demethylation regulated by additional cofactors. Identification of *in vivo* RNA targets is complicated by rather weak and transient RNA binding (Zhang 2019; Xu 2014). To overcome these obstacles and to address the above questions, we developed a protocol for robust and highly reproducible transcriptome-wide definition of ALKBH5 and FTO targets in mammalian cells which will address (i) the direct RNA targets, (ii) protein-binding sites and (iii) identify proteins that simultaneously occupy RNAs with each of the DMT. We call this approach with a working title TRAPOME (transcript-protein-interactome). This will help to unravel which ALKBH5 and FTO-protein interactions are direct and which are bridged by RNA.

2. miCLIP analysis to dinstinguish between modification dependent and independent binding

To further evaluate the nature of DMT-RNA interactions, we will perform transcriptome-wide miCLIP assays (Grozhik 2017) in WT, FTO- and ALKBH5-deficient cells. These analyses will reveal whether bound RNAs is due to the methylation marks. Moreover, we will aim to find a bioinformatics tools to distinguish FTO targets modified by m^6A, m^6A_m and m^1A, respectively. Identification of specific m^6A and m^6A_m sites undergoes constant methodology improvements. The deep sequencing outcomes of miCLIP cDNA libraries from KO cells and WT will be subjected differential expression set analysis, allowing a confident definition of ALKBH5 and FTO-dependent m^6A/m^6A_m sites in the HEK293T transcriptome. We will compare our results to the previously published data on m^6A profiling in HEK293 and other human cell lines and datasets generated in the collaborating laboratories.

REFERENCES

Akichika, S. et al. *Science* 2019, *363* (6423).
Baltz, A. G. et al. *Mol Cell* 2012, *46* (5), 674-90.
Boulias, K. et al. *Mol Cell.* 2019 *75*(3): 631–643.e8
Grozhik, A. V. et al. *Methods Mol Biol* 2017, *1562*, 55-78.
He, P. C. & He, C. *EMBOJ* 2021, *40* (3), e105977.
Huang, H. et al.. *Nature* 2019, *567* (7748), 414-419.
Jia, G. et al. *FEBS Lett* 2008, *582* (23-24), 3313-9.
Li, Y. et al. *Nat Genet* 2020, *52* (9), 870-877.
Liu, J. et al., *Nat Chem Biol.* 2014 *10(2):* 93–95.
Mauer, J. et al. *FEBS Lett* 2018, *592* (12), 2012-2022.
Slobodin, B. et al. *Cell* 2017, *169* (2), 326-337.e12.
Tang, C. et al. *Proc Natl Acad Sci U S A* 2018, *115* (2), E325-e333.
Xu, C. et al. *J Biol Chem* 2014, *289* (25), 17299-311.
Zhang, X. et al. *Proc Natl Acad Sci U S A* 2019, *116* (8), 2919-2924.
Zheng, G. et al. *Mol Cell* 2013, *49* (1), 18-29.

Work programme:

To develop robust tools and approach for data analyses and for comparative analyses between the sequencing datasets that originate from different experimental approaches.
Conventional mRNA sequencing data tend to have larger number of reads. In CLIP-seq, we often have fairly low read densities (10 reads per site). Also, the technical variability relatively high compared to conventional RNA sequencing data. Due to various chemical processing in the clip seq sample preparation. It is very much likely to get Non-specific background binding signal and Crosslinking biases. So, these data demand special attention during normalization process. I will learn various normalization methods for the downstream analysis.

Another key aspect is to handle the multi-mapped reads against repetitive elements. It's widely accepted that the multi-mapped reads were discarded during the downstream analysis. But our interest is to focus where these RMDs are interacting in the genome. Is there any preference towards specific sub population of repetitive elements? If so, what is binding sites and does it follow any motifs to interact with because FTO and ALKBH5 targets are generally different. Another key function of FTO and ALKBH5 in splicing so, additionally I will learn to use the tools for differential exon usage and quantification of poly-A signals.

The proper analysis of such high-throughput data will often constitute a critical bottleneck for molecular biologists. Currently, Praveenkumar Rengaraj is familiar in computational biology from his basic training and analysis of RNAseq datasets. Statistical training appears necessary to be able to analyse these increasingly complex datasets. Therefore we plan a one month stay for PhD student Praveenkumar Rengaraj at the EMBL genomics facility where he will in close collaboration and under tight guidance develop bioinformatics tools and analyse the sequencing datasets generated by the experiments described above.

Budget:

One moth stay accommodation and diets 900 EUR
Travel costs: 400 EUR: includes return flight ticket (cca 200 EUR and airport-city transports cca 200 EUR)

Since the whole work is based on bioinformatics data analyses, we do not inquire any consumable costs.