SHORT TERM SCIENTIFIC MISSION (STSM)
SCIENTIFIC REPORT

This report is submitted for approval by the STSM applicant to the STSM coordinator

Action number: CA17103 - Delivery of Antisense RNA Therapeutics
STSM title:SRp55 splicing landscape and AONs-based splicing modulation
STSM start and end date: 03/02/2019 to 24/02/2019
Grantee name: Maria Ines de Oliveira Alvelos

PURPOSE OF THE STSM:

The main goals of my PhD project are to: 1. Understand the role of alternative splicing in pancreatic β-cell dysfunction and death in the context of type 1 diabetes (T1D); 2. Identify key splicing factors and harmful mRNA splice variants that can be modulated using antisense oligonucleotides (ASOs) to prevent β-cell loss. I have already identified, by RNA-seq and biological validation experiments, that the splicing factor SRp55 (SRSF6) is a master regulator of β-cell function and survival. In collaboration with Dr. Kathi Zarnack from the Buchmann Institute for Molecular Life Sciences, Frankfurt, Germany, the aim of this STSM is to identify direct SRp55 RNA targets and its respective binding sites, using individual-nucleotide UV-crosslinking and immunoprecipitation (iCLIP). This approach will allow us to characterize the SRp55 regulatory map and to identify suitable targets for splicing modulation. The specific aims of the present STSM were: 1) To learn the required bioinformatics tools (R programming) and understand /interpret the downstream RNA-seq + iCLIP analysis, to identify splicing regulatory elements that will be targeted by ASOs. 2) To SRp55-splicing regulatory map and its binding motif(s).
The results obtained during this Short Term Scientific Mission (STSM) are expected to help the design of future AONs and this project therefore fits within the scope of the COST action CA17103 - Delivery of RNA therapeutics.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

During my visit to Dr. Kathi Zarnack's lab I used the R programming language combined with other specific tools to perform the downstream analysis of the SRp55 knockdown (KD) RNA-seq and SRp55 iCLIP-seq data that I had previously generated.
For the SRp55 iCLIP-seq analysis, the 'peak calling' was performed using the ASpeak software (version 2.0). The identified SRp55 binding sites were annotated using the Gencode annotation (Release 29, GRCh38.p12). Annotations were filtered by feature (only verified and/or manually annotated loci were included, while automatically annotated loci were excluded) and transcript annotation level (well supported transcripts based on Gencode definition).
The identification of the SRp55 consensus motif was conducted using the HOMER software (http://homer.ucsd.edu/homer/motif/) with the script developed for RNA motifs (findMotifs.pl/findMotifsGenome.pl).

In order to investigate the relationship between the observed splicing events after SRp55 KD and the genomic location of SRp55 binding sites, the overlap between the SRp55 KD RNA-seq and the SRp55 iCLIP-seq was conducted using the R/Bioconductor packages GenomicRanges, ChIPseeker, regioneR and seqPattern.
The pathway enrichment analysis was accomplished using DAVID Bioinformatics Resources 6.8 (https://david.ncifcrf.gov/) to assess the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. For the pathway enrichment analysis, all genes expressed in pancreatic β-cell genes (Reads per kilobase per million mapped reads, RPKM > 0.5) were used as background. Only the significant pathways (adjusted p-value < 0.05) were considered.

DESCRIPTION OF THE MAIN RESULTS OBTAINED

We identified more than 300,000 reproducible SRp55 binding sites in around 10,000 genes.
In agreement with what was observed for other SR protein members, we verified that SRp55 binding occurs preferentially in exons. We analyzed the sequence around the SRp55 binding sites, and found that the most highly enriched motif is a purine-enriched motif.
The overlap between the SRp55 KD RNA-seq and the iCLIP results showed a good correlation between both datasets with around 3,000 genes harboring splicing modifications and being direct SRp55 targets. These genes are functionally related and the pathway enrichment analysis showed that they are involved in crucial pathways for β-cell function and survival.
The correlation between the SRp55-regulated splicing events observed by RNA-seq and the positional information obtained by iCLIP-seq showed that the density of SRp55 crosslinking is higher across the SRp55-enhanced exons than across SRp55-repressed ones, suggesting that SRp55 binding promotes exon inclusion.
This extensive bioinformatic analysis allowed the identification of the transcriptome-wide binding sites of the SRp55 and also its direct tragets. The SRp55 RNA-map information is now being used to develop specific AON molecules targeting crucial *cis*-regulatory regions, placing this project in the scope of the COST action CA17103 - Delivery of RNA therapeutics.

FUTURE COLLABORATIONS (if applicable)

The eCOST STSM funding facilitated a very fruitful collaboration between our group at the Center of Diabetes Research (Brussels, Belgium) and the Computational RNA Biology group of Dr. Kathi Zarnack (Frankfurt, Germany), allowing the integration of complex bioinformatics analysis with specific biological questions. We intend to maintain this collaborative work with Dr. Zarnack's group. Additionally, concerning the major challenges ahead of us, we will pursue our ongoing collaboration with Dr Aartsma-Rus, Leiden University, for the AON design and aim to establish new collaborations in the field of AON delivery, since our goal is to design and target AONs specifically to pancreatic β-cells. We have already identified two promising biomarkers, expressed at the β-cell surface, that can be used for this purpose.