RNA-Decoder

Development of a novel DNA-sequencing-based method to identify proteins in ribonucleoprotein complexes

Anttila Beda[1], PhD. Turtola Matti[1], PhD. Garcia Pichardo Desire[2] and Prof. Marquardt Sebastian[2]
[1]Department of Life Technologies, University of Turku, [2]Department of Plant and Environmental Sciences, University of Copenhagen

INTRODUCTION

Long non-coding RNAs (lncRNAs) constitute a diverse class of RNA-molecules that do not encode proteins. They are important regulators of various cellular processes including cell cycle, cellular differentation and metabolism. Typically, lncRNAs do not act by themselves in cell, but they interact with RNA-binding proteins. to form functional ribonucleoprotein complexes (RNPs). (Kung et al. 2013.) Together these complexes can affect gene expression by modelling chromatin structure or altering transcription or translation processes (Fig. 1). Although thousands of lncRNAs have been detected in cells, functional mechanisms have been characterized for only few of them.

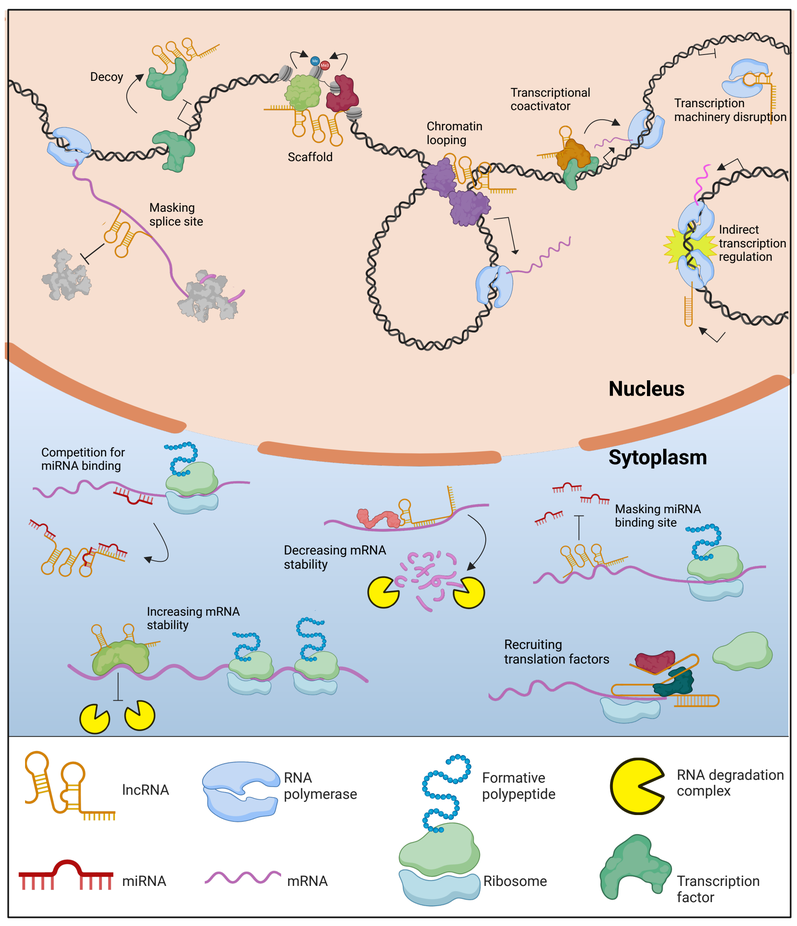

Fig 1. LncRNAs regulate gene expression through various mechanisms. Adapted from Kung et al. 2013. (Created with BioRender.com)

AIM

The aim of this study is to reveal differences between RNP complexes formed by coding mRNA transcripts and lncRNA transcripts, and to understand which RNP complexes mediate regulatory functions within the lncRNA-protein complexes.

To achieve this, the study concentrates on the development of a new method called RNA-Decoder, designed to identify proteins associating with specific RNA transcripts.

CONCLUSION

Preliminary results indicate successful immunoprecipitation step and acceptable PCR meditated barcode amplification. The next step involves sequencing the purified PCR-products and analyzing the results using bioinformatics to confirm the efficiency of the method and to detect differences between mRNA-protein and lncRNA-protein complexes.

References: *Kung, J. T. Y., Colognori, D. & Lee, J. T. (2013) Long Noncoding RNAs: Past, Present, and Future. Genetics* ***193****:651–669.*

METHODS

RNA-Decoder method focuses on identifying the RNA-binding proteins within RNPs that assemble onto RNA transcripts that originate from a locus of divergent transcription in *Saccharomyces cerevisiae*. The locus of interest includes protein coding GCG1 region and lncRNA coding region SUT098. Each TAP-tagged yeast strain is associated with unique barcodes at this region. (Fig. 2A).

RNA-Decoder method harnesses Epi-Decoder yeast library, which consists of thousands of yeast strains, each strain containing two unique DNA-barcodes at the locus of interest and a unique protein with a TAP-tag. (Fig. 2B). Barcodes and TAP-tagged proteins make it possible to compare which proteins bind to protein coding mRNA transcripts and which to lncRNA transcripts.

RNA-Decoder method is based on DNA barcoding, RNA immunoprecipitation and high-throughput sequencing (Fig. 2C-E.)

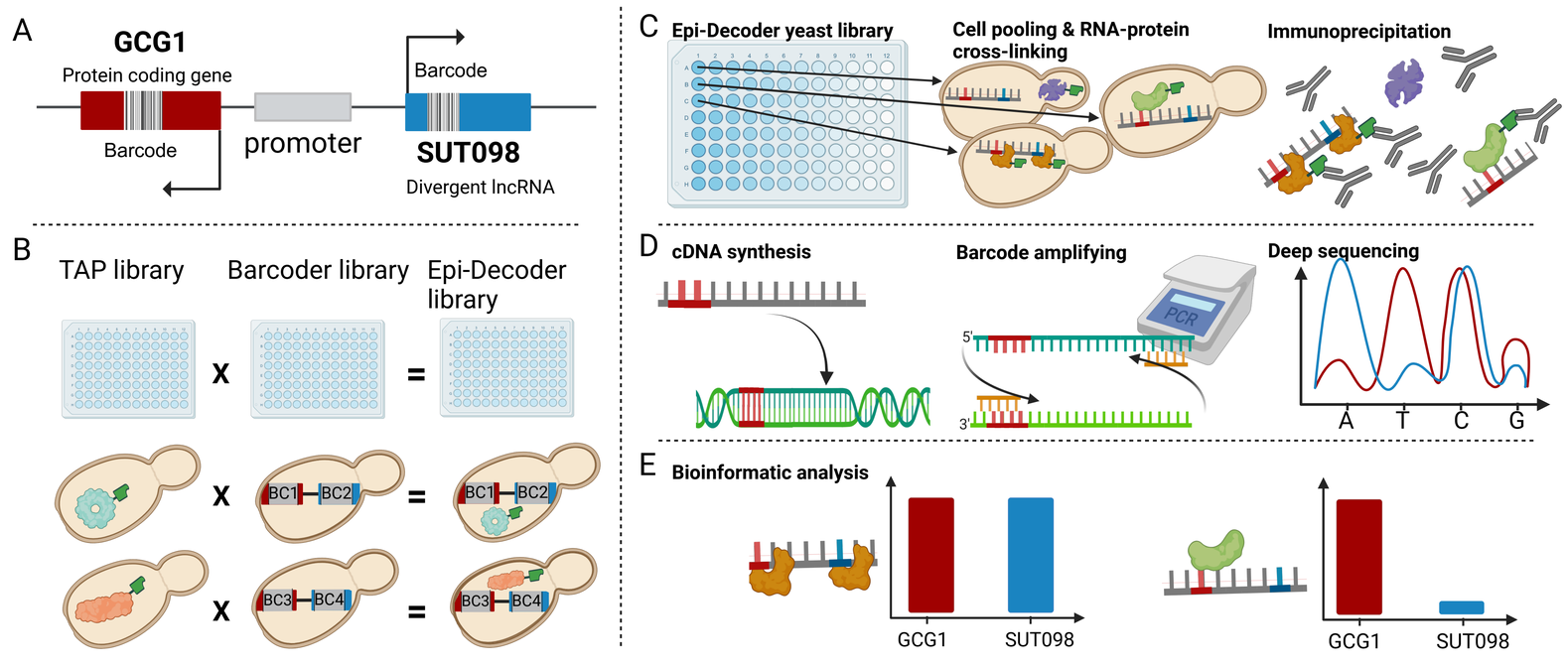

Fig 2. (A) The locus of interest. (B) Epi-Decoder library was created by PhD. Desire Garcia Pichardo by mating TAP-library with a custom made Barcoder library. (C-E) RNA-Decoder method: (C) All yeast strains from Epi-Decoder are pooled, RNPs formed within yeast cells are crosslinked by formaldehyde and then isolated by TAP-tag immunoprecipitation. (D) Barcode regions from immunoprecipitated RNA transcripts are converted into cDNA, amplified by PCR, and sequenced. (E) Bioinformatic analyses are used to identify which TAP-tagged proteins bound to the barcoded RNA transcripts (GCG1 mRNA and SUT098 lncRNA). (Created with BioRender.com)

RESULTS

The preliminary results of RNA-Decoder indicate that the method works, although the PCR step would require further optimization.

At the first, the immunoprecipitation step was tested using only four Epi-Decoder strains ADH2-TAP, and HEK2-TAP, NAB2-TAP and NAB3-TAP. After immunoprecipitation, binding of Adh2, Hek2, Nab2, and Nab3 proteins to RNA transcripts was checked by qPCR. (Fig 3.)

Next, the pooled strains in RNA-Decoder were used to test the immunoprecipitation step. Detecting TAP-tagged proteins by western blot analysis demostrates the success of the immunoprecipitation. PCR-amplified barcode regions were reviewed by gel electrophoresis. (Fig 4 - 5).

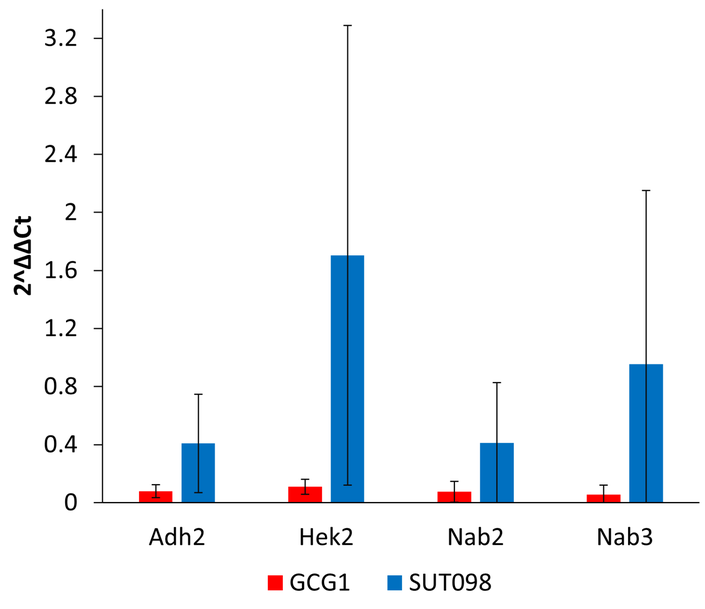

Fig 3. Binding of Adh2, Hek2, Nab2 and Nab3 proteins against GCG1 and SUT098 transcripts. Adh2 should not bind on RNA transcripts in yeast cells at all, Nab2 and Nab3 should bind on RNA transcripts in nucleosome and Hek2 should bind also in cytoplasm.

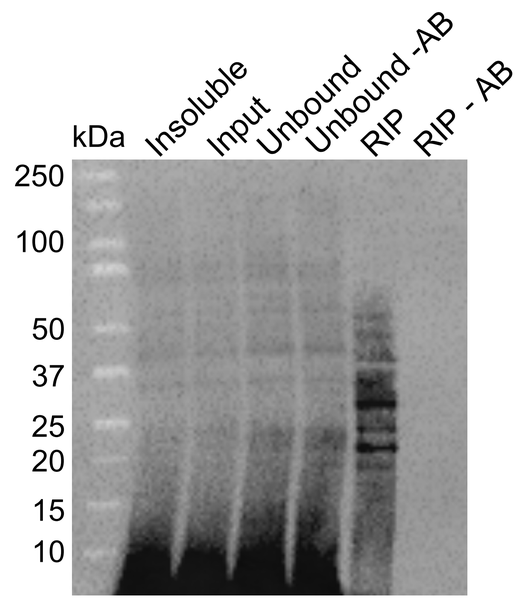

Fig 4. Immunoprecipitated sample (RIP) represents the proteins pulled down by the anti-TAP-tag antibody, which exhibit clearer and more visible bands than the control samples before immunoprecipitation (insoluble, input and unbound samples). A negative control was immunoprecipitated without the anti-TAP-tag antibodies (RIP -AB), an empty lane for this control, indicates that only TAP-tagged proteins have been pulled down.

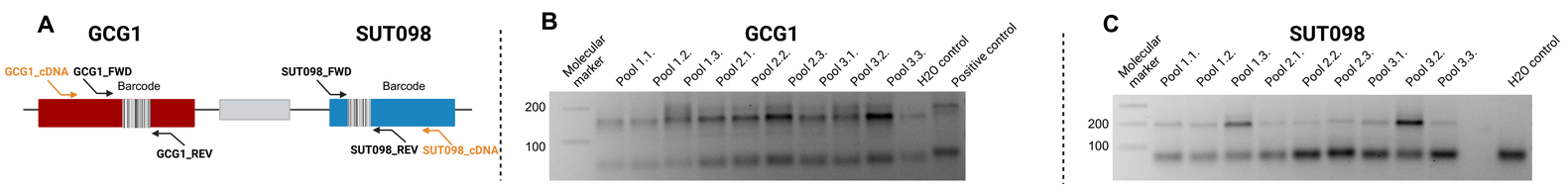

Fig 5. (A) A schematic representation of the locus of interest and the primers used to convert immunoprecipitated RNA transcripts into cDNAs (orange) and amplify barcode regions from cDNAs (black). (B-C) PCR-amplified barcode regions from SUT098 and GCG1 areas. The size of the PCR products seems to be approximately 200 nucleotides, same than the positive control samples, as expected. However, further optimization of the PCR is necessarily, because there are still some extra bands on the GCG1 gel caused by primer-dimers (C).