Combining transcriptional and post-transcriptional regulation to predict mutations altering the gene regulatory program in cancer

Jaime A Castro-Mondragón[1], Miriam R Aure[2], Vessela N Kristensen[2] and Anthony Mathelier[1,2]

1. Centre for Molecular Medicine Norway (NCMM), Nordic EMBL Partnership, University of Oslo, 0318 Oslo, Norway

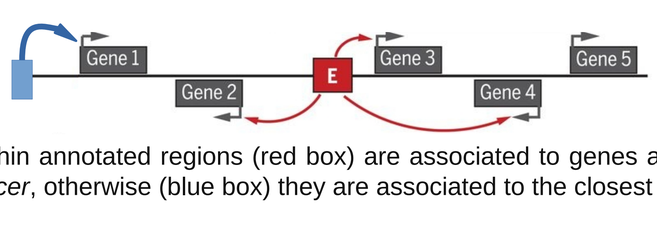
2. Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, 0310 Oslo, Norway

@jaimicore

j.a.c.mondragon@ncmm.uio.no

TFBS mutations with a cascading effect on miRNA-target expression

- Most somatic mutations are non-coding, a small fraction occurs at transcription factor binding sites (TFBSs).
- Interpretation of the effect of TFBS mutations can be eased by using expression data.
- Most of the known cancer drivers genes are protein-coding, but non-coding genes may also be cancer drivers.
- We mapped somatic mutations to TFBS with experimental and computational evidence derived from *UniBind*[1].
- Each TFBS was associated to its potential targets by combining the annotations from geneHancer[2] and associations to the closest TSS.

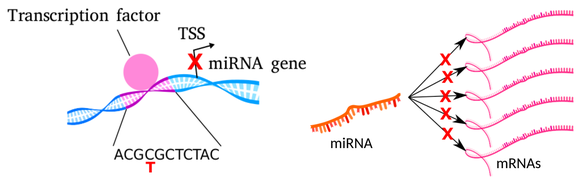


TFBS within annotated regions (red box) are associated to genes according to *geneHancer*, otherwise (blue box) they are associated to the closest TSS.

1. Gheorghe et al, NAR 2018
2. Fishilevich et al, Database 2017

- We adapted *xseq*[3], a bayesian probabilistic framework, to assess the likely association between mutated miRNA-associated TFBSs with dysregulation (cascading effect) in miRNA networks.

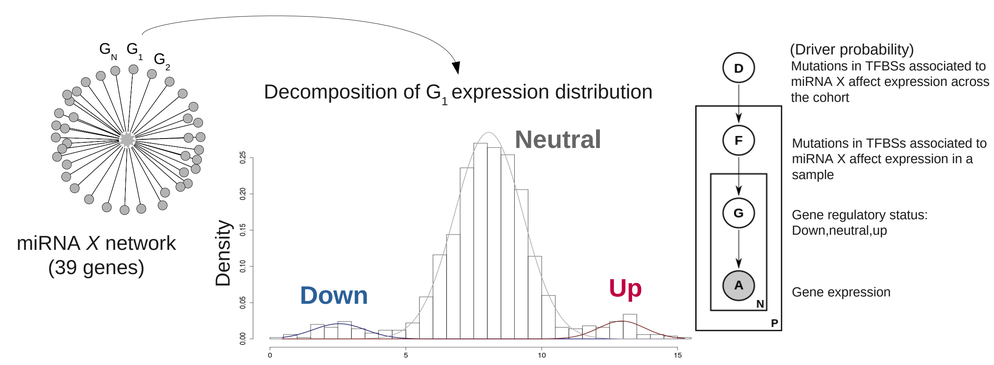
3. Ding et al, Nat Commun 2015



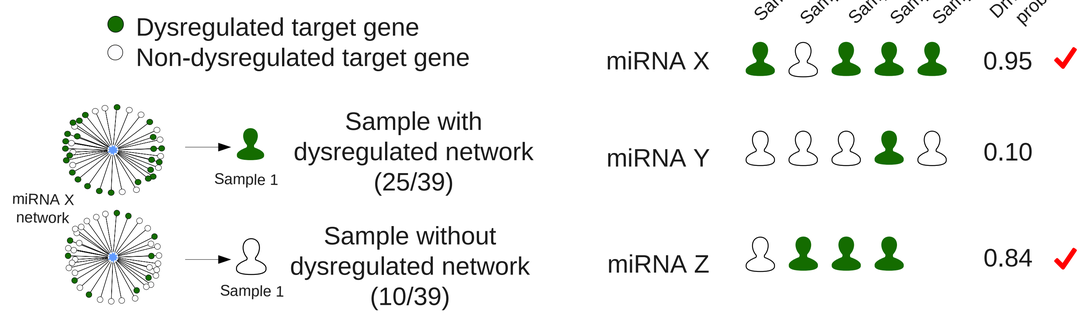
- We combined transcriptional (TFBS mutations) and post-transcriptional (miRNA networks) information to highlight cancer driver miRNAs across 7 TCGA[4] cohorts .

4. Zhang et al, Database 2011

Associating mutations with gene dysregulation



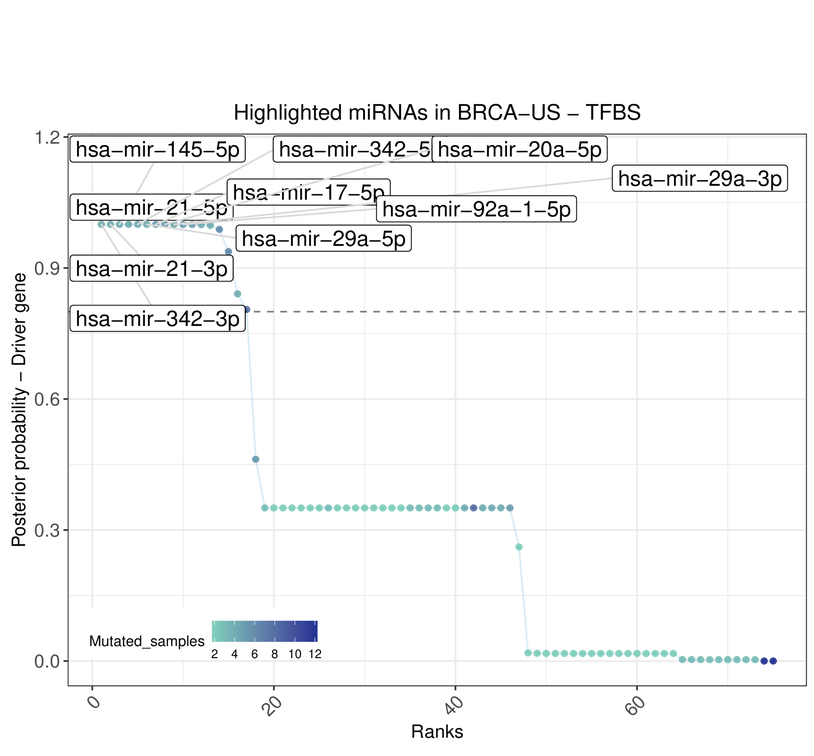
On each sample, we infer the regulatory status (down,neutral,up) of each gene within a miRNA network.



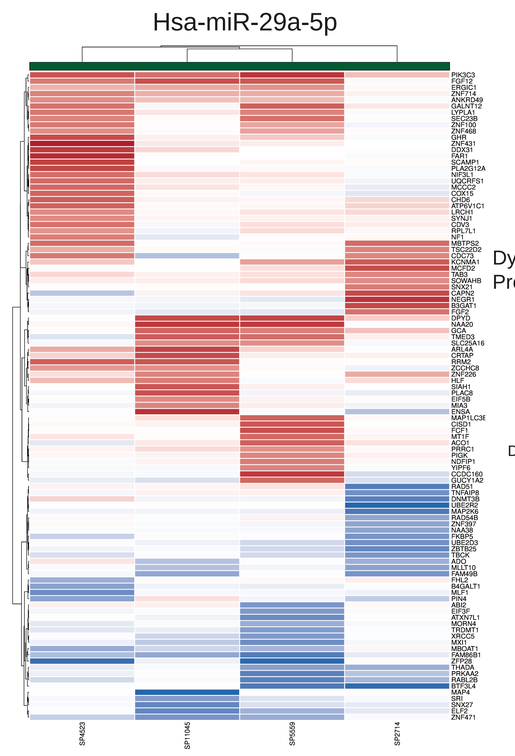
miRNA X mutations are associated with dysregulation in the network (Sample 1).

Mutations in miRNAs X and Z, are recurrently associated with dysregulation across the cohort.

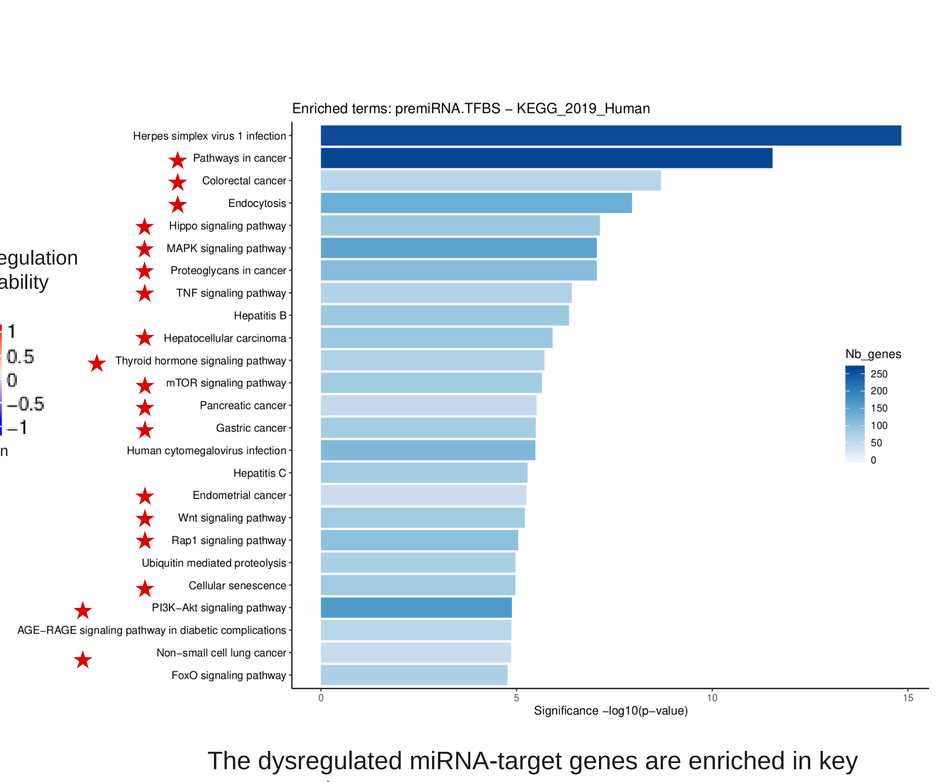
Dysregulated miRNA-target genes are enriched in key cancer pathways



For each cohort, we selected the miRNAs with high *Driver posterior probability.*

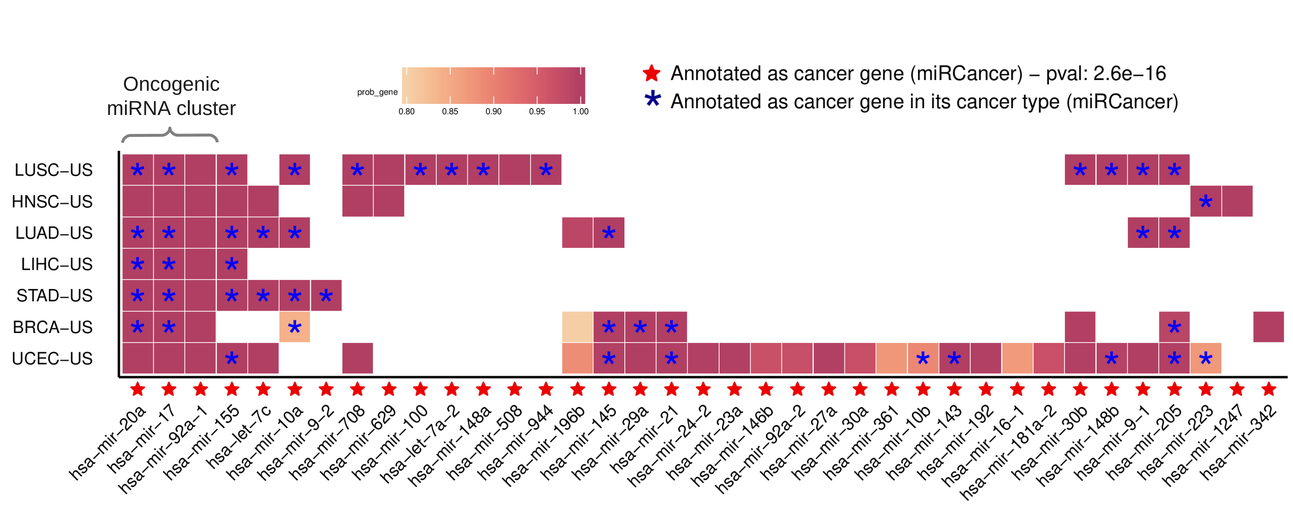


For each mutated sample (columns), we selected the dysregulated target genes (rows) of the highlighted miRNAs.



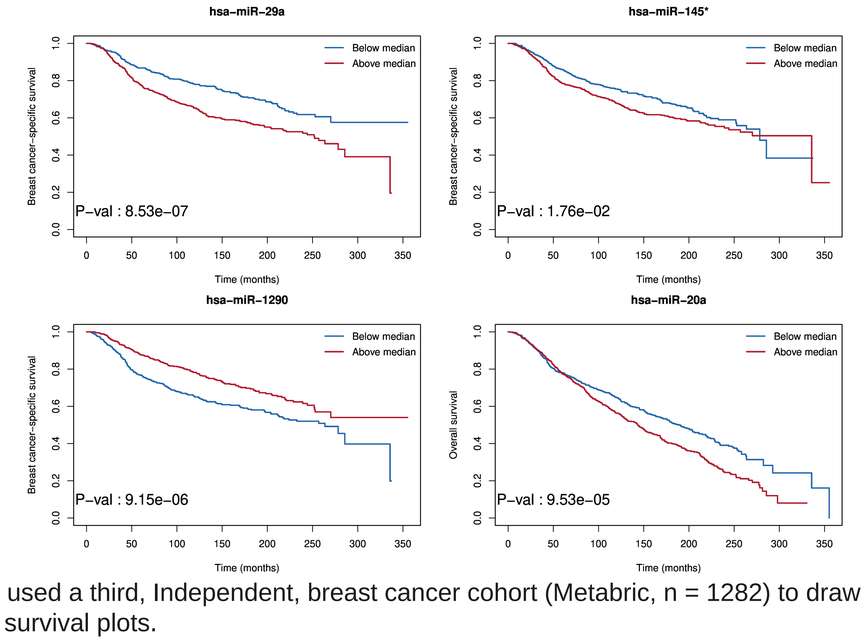
The dysregulated miRNA-target genes are enriched in key cancer pathways.

Pan-cancer predicted cancer driver miRNAs



We predicted 38 miRNAs in 7 TCGA cohorts. Three well known oncogenic miRNAs (miR-20a, miR-17, miR-92a) were predicted independently in the 7 cohorts. All the predicted miRNAs are annotated as cancer miRNAs in miRCancer.

Predicted cancer driver miRNAs are associated to prognosis



We used a third, Independent, breast cancer cohort (Metabric, n = 1282) to draw the survival plots.

Conclusions

- By combining transcriptional and post-transcriptional information we highlighted potential cancer driver miRNAs (with mutations at their TFBSs) likely associated to a cascading effect on the miRNA networks.
- Non-coding mutations coupled with gene expression can be explored to highlight cancer driver genes.
- The same methodology also works in TFBS mutations associated to protein coding genes, and could be adapted for other genes such as lncRNAs.