Construction of regulatory networks by integrating gene expression, promoter methylation and copy number alteration data for prostate cancer

Rinat Sultanov[1,2,3], G. Arapidi [1,2], K. Babalyan[1,3], V. Gordeeva[1,2,3], E. Generozov[3], V. Govorun[1,2,3]

Affiliations:
1) Moscow Institute of Physics and Technology (State University), Moscow, Russia
2) Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia
3) Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia

Key words:
Prostate cancer, EWAS, DNA methylation, Transcriptomics, Network analysis

Prostate cancer (PCa) is one of the most commonly diagnosed malignancy in men. It is a very heterogeneous disease, with a phenotype ranging from indolent behavior lasting decades to highly aggressive metastatic cancer which can be lethal in just a few years.

To identify novel drivers of PCa, we constructed a network by integrating gene expression, promoter methylation, miRNA expression and copy number alteration (CNA) data from TCGA (The Cancer Genome Atlas). We used the linear mixed model package pyLMM to test for association among the samples. We applied three regression models: i) Ei = Ej + error, ii) Ei = Ej+ Mi + error, iii) Ei = Ej + CNVi + error, where Ei - i-th gene expression, Mi - i-th gene's promoter methylation, CNVi - copy number variation of i-th gene and error - residual expression. The same models were used for miRNA expression data. We also considered model Ei = CpGi + error to identify associations between genes and CpG-sites. Bonferroni threshold was used as correction for multiple comparisons. Reconstructed network had four big connected components (clusters). We selected the second one for further analysis.

We identified that selected sub-network related to cell migration and metastatic PCa. 95% genes of this cluster was downregulated. We found that one of the corresponding master-regulators of cluster was TP63. Interestingly, that TP63 belonged to selected sub-network and its expression was downregulated in tumor samples. Promoter regions of cluster were enriched with ChIP-seq peaks of TP63 in cell line RWPE1 (normal prostate). Sub-network's gene expression was highly correlated with hypermethylated enhancer CpG-sites (eCpG) (spearman correlation coefficient < -0.8). Furthermore, we found that correlated sites were enriched with

ChIP-seq peaks of TP63 and H3K27ac in normal prostate cell lines (RWPE1, PrEC) but not in the tumour lines (PC3, LNCaP, VCaP). To validate enhancer-promoter interaction we used Hi-C data from three prostate cell lines (PrEC, PC3, LNCaP) and found 199 differential interactions between correlated eCpG and cluster' genes. We also observed that eCpG were enriched with DNMT3A peaks in epithelial stem cells.

In this study we reconstructed network based on EWAS (Epigenome Wide Association Study) and TWAS (Transcriptome Wide Association Study). We constructed subnetwork associated with invasion and metastatic PCa and identified its master-regulator - TP63. We found that TP63 bound with active enhancers in normal prostate cell line but not in the cancer one. Chromatin interaction networks based on Hi-C data confirmed that eCpG interacted with cluster's genes. Thus, cluster's genes were regulated by TP63 via distal enhancers. Based on our findings and recently published data we speculate that TP63 binds with active enhancer through DNMT3A and maintain expression level of cluster' genes and methylation level of correlated eCpG.