Epigenome-wide DNA methylation and transcriptome profiling of localized and locally advanced prostate cancer: uncovering new molecular markers

Madison Reed[1], Haiying Zhu[2], Qian Liu[1], Ricardo Ribeiro[3] and Pingzhao Hu[1]
[1]Department of Biochemistry and Medical Genetics, University of Manitoba, Canada, [2]Dalla Lana School of Public Health, University of Toronto, Canada, [3] i3S/INEB, Instituto de Investigação e Inovação em Saúde Universidade do Porto, Portugal

University of Manitoba

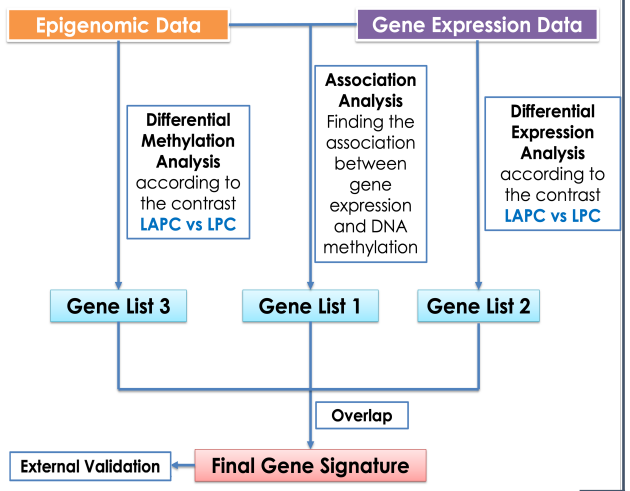
The Hu Lab
Bioinformatics and Statistical Genetics

Background: Prostate Cancer (PC) is one of the most common neoplasms and a leading cause of cancer-related deaths among men. PC can be further distinguished as localized PC (LPC), being confined to the prostate gland, and locally advanced PC (LAPC), having metastasized past the primary site, but not to distant sites. Reliable markers are needed to distinguish LAPC from LPC for a better estimation of patients' prognosis and further treatment plan.

Hypothesize: We hypothesize that a gene signature can be built to distinguish LAPC from LPC and illustrate its molecular mechanism based on the genes' differential expression profiles and the CpGs' differential methylation profiles in the genes' promoter regions. We also hypothesize that the expression profiles of these genes is down-regulated while their methylation profiles in their prompter regions are hypermethylated.

1

Overall study workflow



2

Methods

Discovery dataset:
- Collected in 2009 by the Portuguese Institute of Oncology, Porto Centre
- We generated the gene pression profiles [1] and the DNA methylation profiles [2] of the 10 patients (4 LPC patients and 6 LAPC patients)

External validation dataset:

Study-specific sample sizes

Study	No. LAPC	No. LPC
MSKCC/DFCI, Nature Genetics 2018	333	680
MSKCC, Cancer Cell 2010	37	181
Fred Hutchinson CRC, Nat Med 2016	154	22
Broad/Cornell, Cell 2013	2	55

- Genetic data for the LPC and LAPC samples from 4 studies was obtained from cBioPortal.
- Mutation frequencies of the genes of interest were downloaded

3

Methods

Association analysis: LASSO model

$$Y = \beta 0 + \beta 1 * X_1 + \cdots + \beta k * X_K + \epsilon$$

- Y is each gene expression
- Explanatory variables ($X_1,\ldots,Xk$): methylation level of the kth CpG in this gene's promoter region.

$$\hat{\beta}^{LASSO} = argmin_{\beta} \parallel Y - X\beta \parallel_2^2 + \lambda \parallel \beta \parallel_1$$

- λ chosen by 10-fold cross-validation

Gene list 1:
- Statistically significant (p-value < 5×10^{-8})
- Anti-correlation between the gene expression and the methylation level ($\beta < 0$)

4

Methods

Differential expression analysis:
Linear model based differential expression analysis between LAPC and LPC was performed for each gene.
Gene list 2:
- Identify the genes with down-regulated gene expression (fold change < -1)

Differential methylation analysis:
Linear model based differential methylation analysis between LAPC and LPC was performed for CpG.
Gene list 3 (differential methylation analysis):
- Identify the genes with hyper-methylated CpGs in the promoter region (fold change > 1)

5

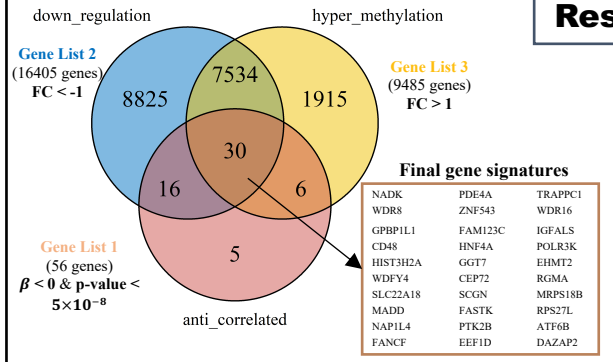
Methods

6

External validation

- The number of mutations for each gene of interest across 4 studies were counted in LAPC and LPC group, respectively.
- Fisher's exact test was used to compare the number of mutations between the LPC and LAPC samples for each gene to select a list of significant genes after multiple testing.
- OncoPrint was used to visualize the genomic alterations in the genes of interest
  - Top panel: 4 studies
  - Second panel: LAPC (Red) and LPC (Blue)
  - Gene-specific panels: Amplification (red) and deletion (Blue)

Results



Intersection of downregulated, hypermethylated, and anti-correlated genes.

7

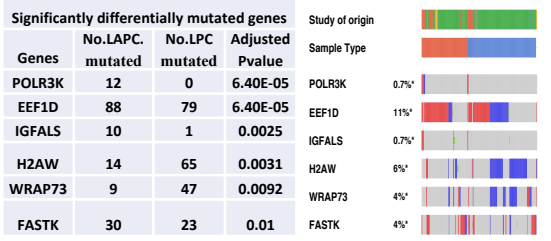
Results

8

External validation

Significantly differentially mutated genes

Genes	No.LAPC. mutated	No.LPC mutated	Adjusted Pvalue
POLR3K	12	0	6.40E-05
EEF1D	88	79	6.40E-05
IGFALS	10	1	0.0025
H2AW	14	65	0.0031
WRAP73	9	47	0.0092
FASTK	30	23	0.01



Most significantly affected genes from the potential gene signature list.

Conclusion: 30 downregulated, hypermethylated genes were identified as gene signatures for LAPC. From these 30 genes we further determined that 6 are the most statistically significant in distinguishing LAPC from LPC at DNA level. Further experimental validation can be used to refine the gene list.

9

References:

[1] Chen Y, et al. Epigenome-wide DNA methylation profiling of periprostatic adipose tissue in prostate cancer patients with excess adiposity-a pilot study. Clinical epigenetics vol. 10 54. 17 Apr. 2018.
[2] Ribeiro R, et al. Obesity and prostate cancer: gene expression signature of human periprostatic adipose tissue. BMC medicine vol. 10 108. 25 Sep. 2012.

10