MEK5-mediated gene expression changes in PC3 human prostate cancer cells

Retrieved from: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE156401

Background

- Mitogen/extracellular signal-regulated kinase kinase-5 (MEK5) belongs to the family of MAP kinases. It has been confirmed to play a pivotal role in tumor carcinogenesis and progression, including prostate cancer. MEK5 protein is overexpressed in prostate cancer cells compared with normal cells and MEK5 levels are associated with prostate cancer metastasis.
- Short hairpin RNA (shRNA) is a type of RNA molecule that can be used to silence or downregulate the expression of specific genes in a cell.
- PC3 is a human prostate cancer cell line. It is widely used in cancer research, particularly in studies related to prostate cancer.

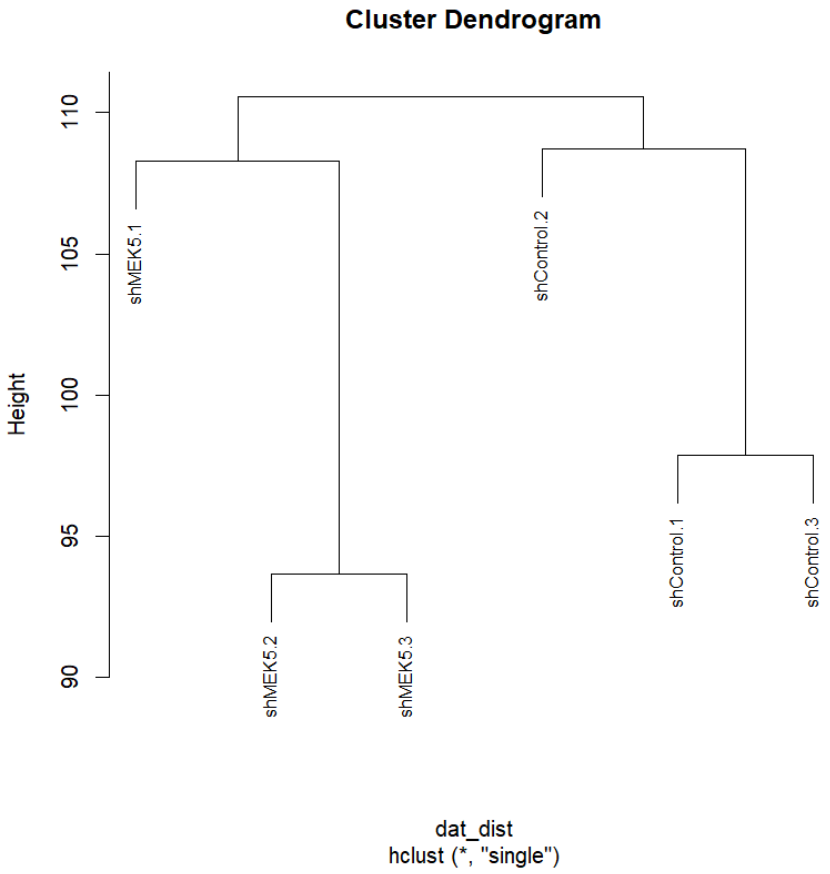
Introduction

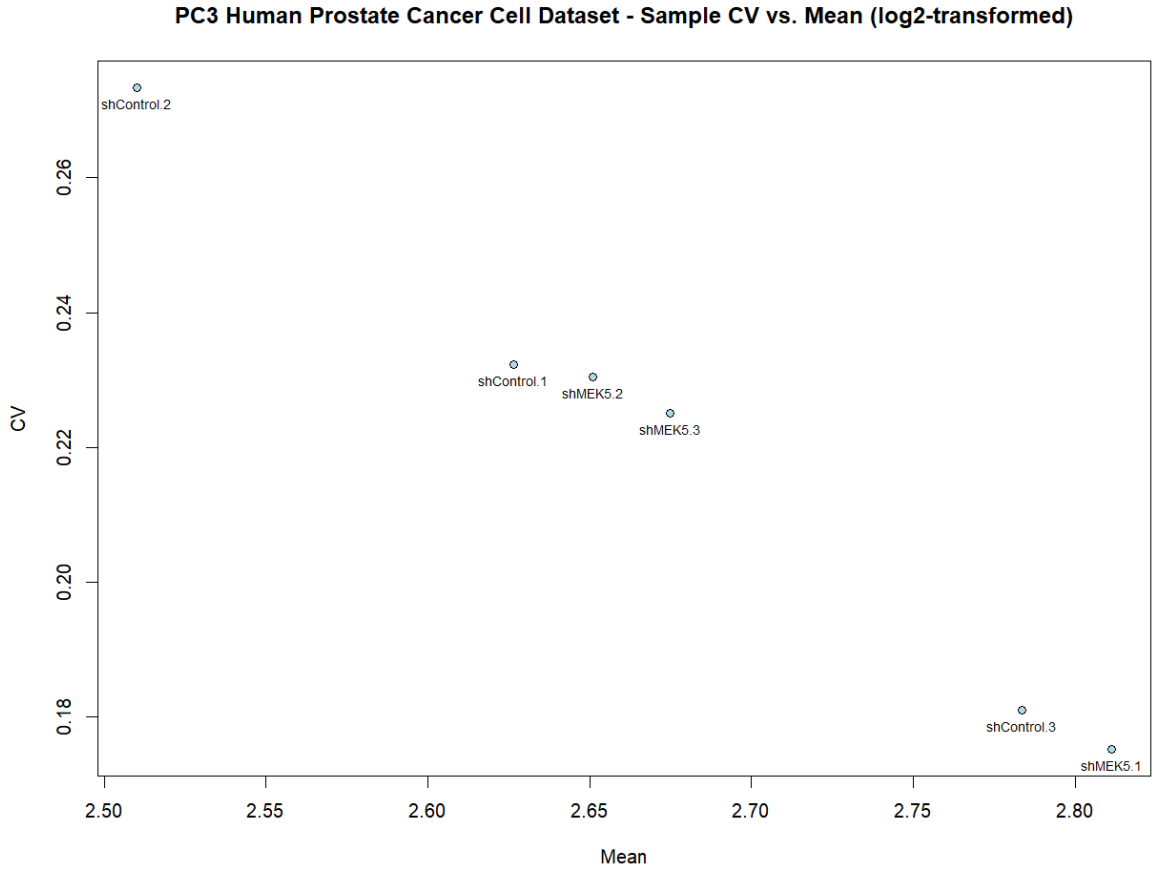
- GEO accession: GSE156401
- Experimental Design: Gene expression was measured in PC3 human prostate cancer cells stably expressing a scrambled shRNA (shControl) or MEK5 shRNA. Three replicates were used for each condition.
- Platform: Agilent-026652 Whole Human Genome Microarray 4x44K v2
- Data Dimensions: 10056 (probes) x 6 (samples)
- Annotation:

GSM4730204	shControl.1
GSM4730205	shControl.2
GSM4730206	shControl.3
GSM4730207	shMEK5.1
GSM4730208	shMEK5.2
GSM4730209	shMEK5.3

Outlier Sample Identification

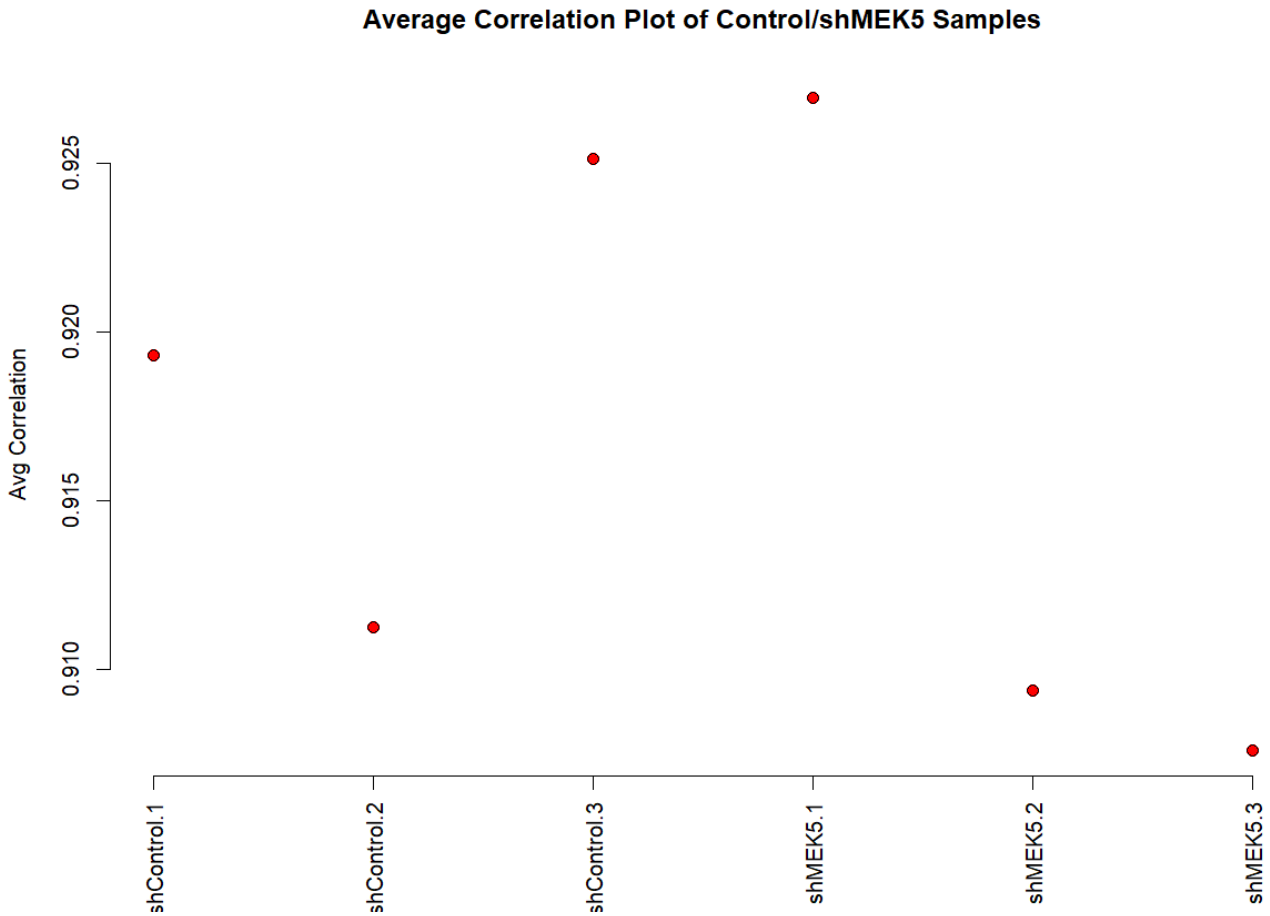
- To identify any outlier samples, the following graphs are plotted:





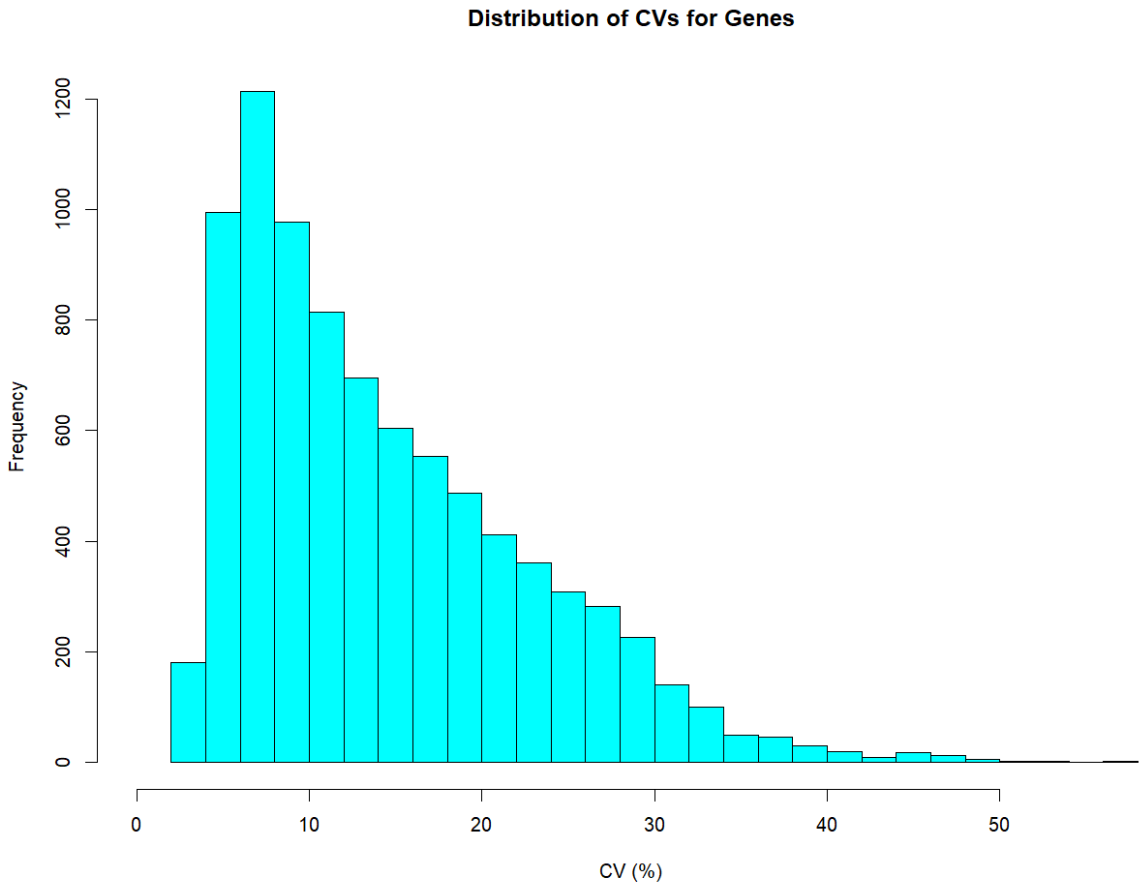
Outlier Sample Identification

- According to the three graphs presented, no clear outliers are observed. The cluster dendrogram reveals that samples are appropriately grouped based on experimental conditions without any distinct anomalies. In both the CV vs. Mean plot and the average correlation plot, the variations among samples are not substantial enough to indicate the presence of outliers. Therefore, the evidence suggests that the data points follow the expected patterns and do not exhibit any significant deviations or outliers.



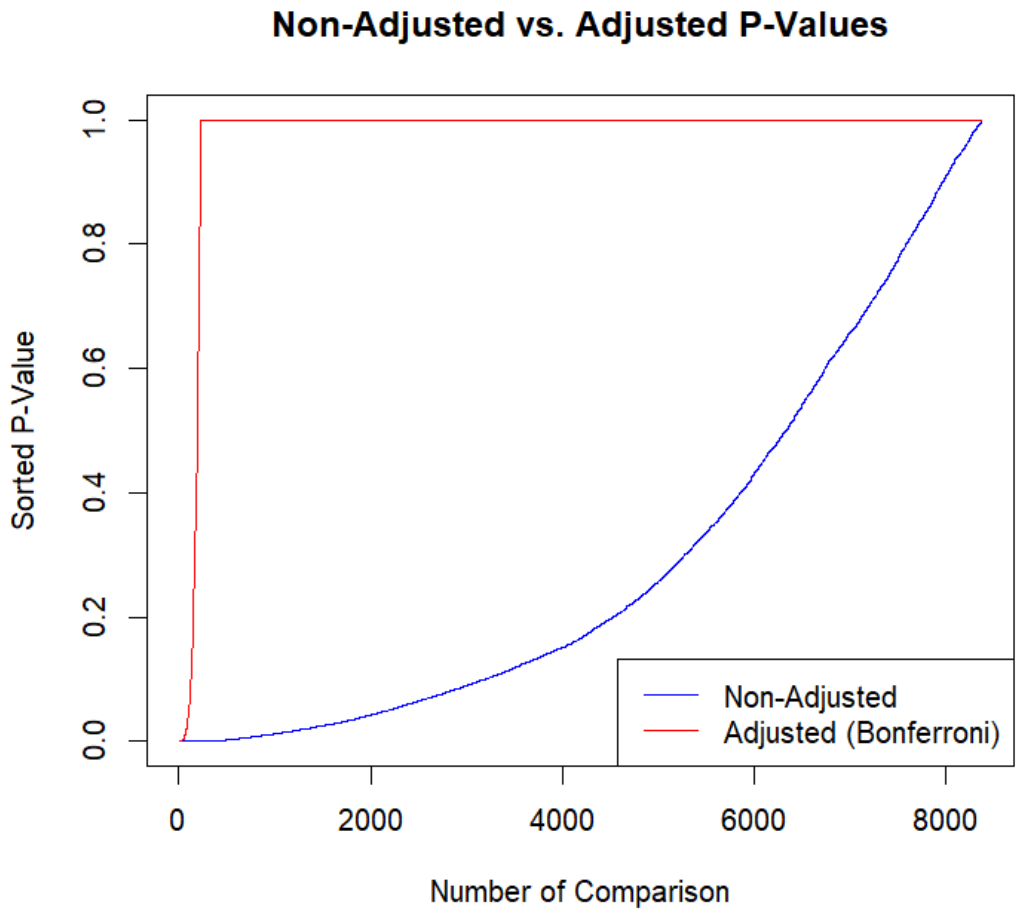
Gene Filtering: Removing Low-Expression Candidates

- Before filtering out the genes with low expression values, all rows with missing values were removed, which resulted in a dataset with 8552 genes/probes. The histogram shows the distribution of coefficient of variation (CV) values for these 8552 genes. Visual inspection suggested a potential cutoff at CV (%) < 4. Assuming genes with low variation will not be differentially expressed, 181 genes with CV (%) < 4 were removed, resulting in a dataset with 8371 genes.



Feature Selection: Empirical Bayes

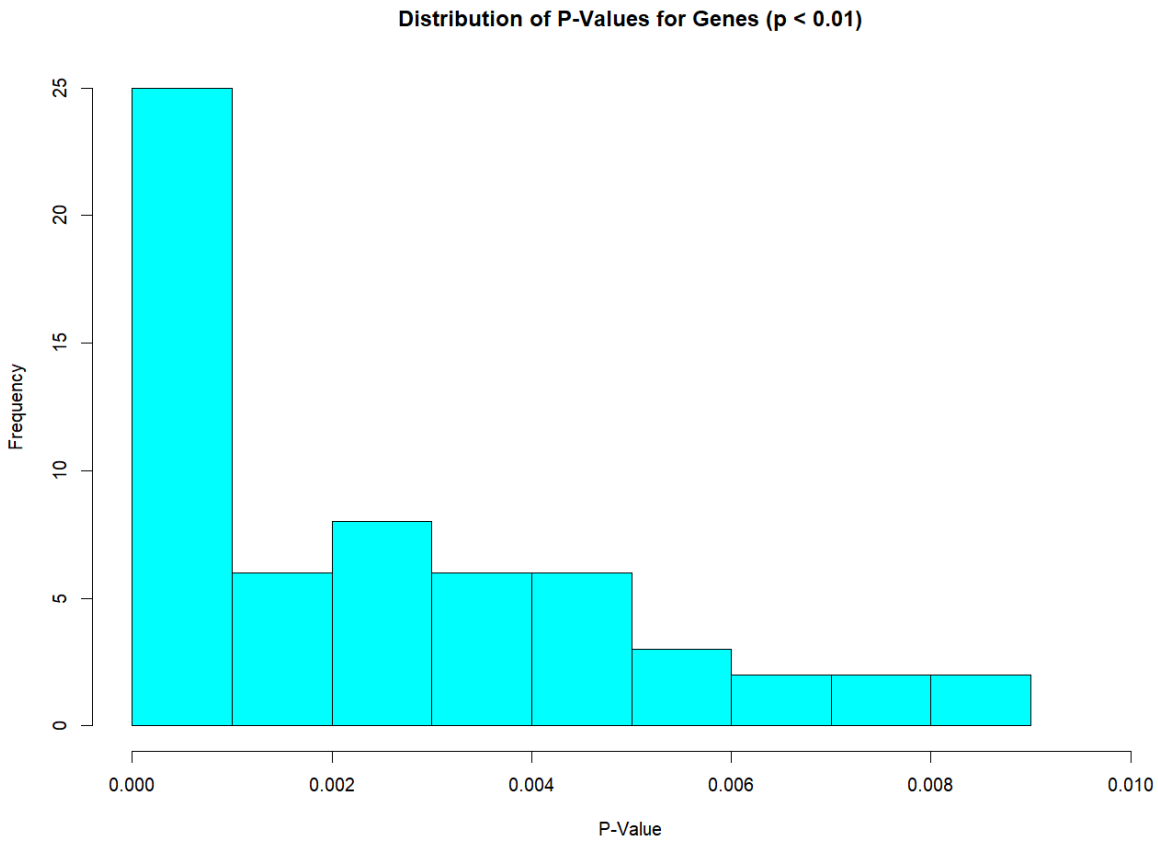
- The p-values were calculated for all genes using the eBayes() function in the limma package. Using mt.rawp2adjp in the multtest library, the p-values were adjusted for multiplicity with the Bonferroni method. The plot compares the non-adjusted and adjusted p-values against the number of comparisons. A clear distinction is observed between the two trends, emphasizing the effectiveness of the Bonferroni correction in minimizing false positives (non-differential genes with initially low p-values).



Feature Selection: Empirical Bayes

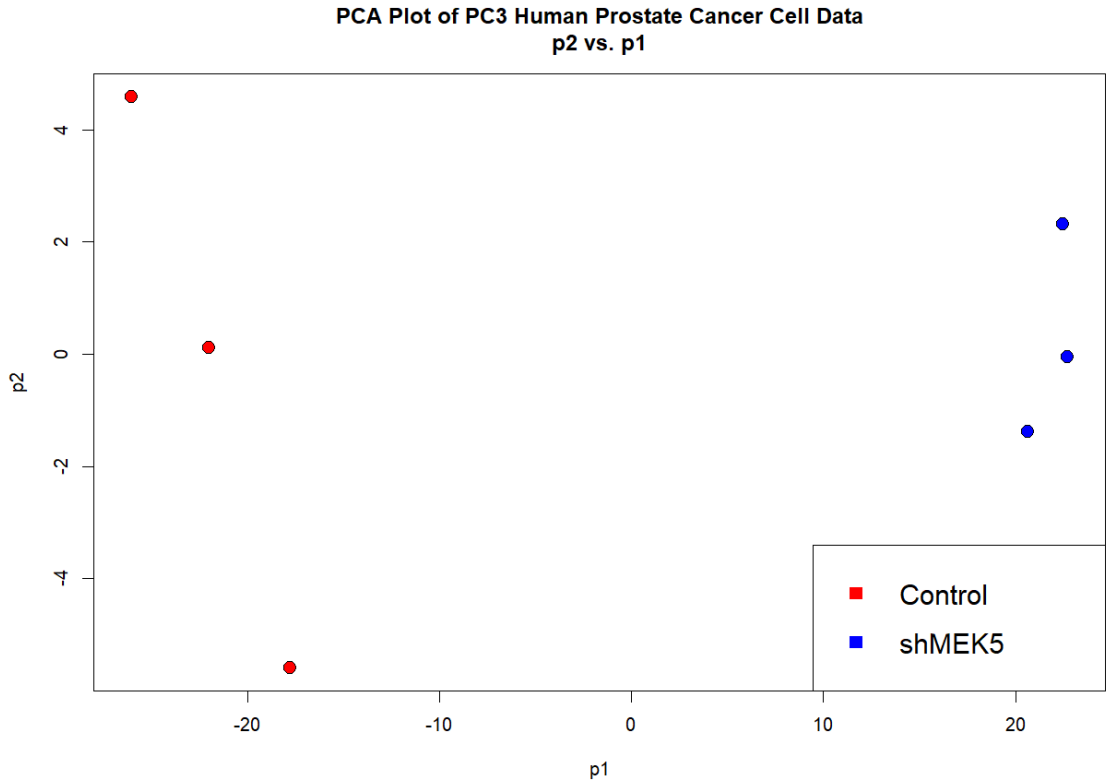
- Only genes with an adjusted p-value less than 0.01 were extracted. The 60 genes retained, along with their corresponding p-values, are presented below:

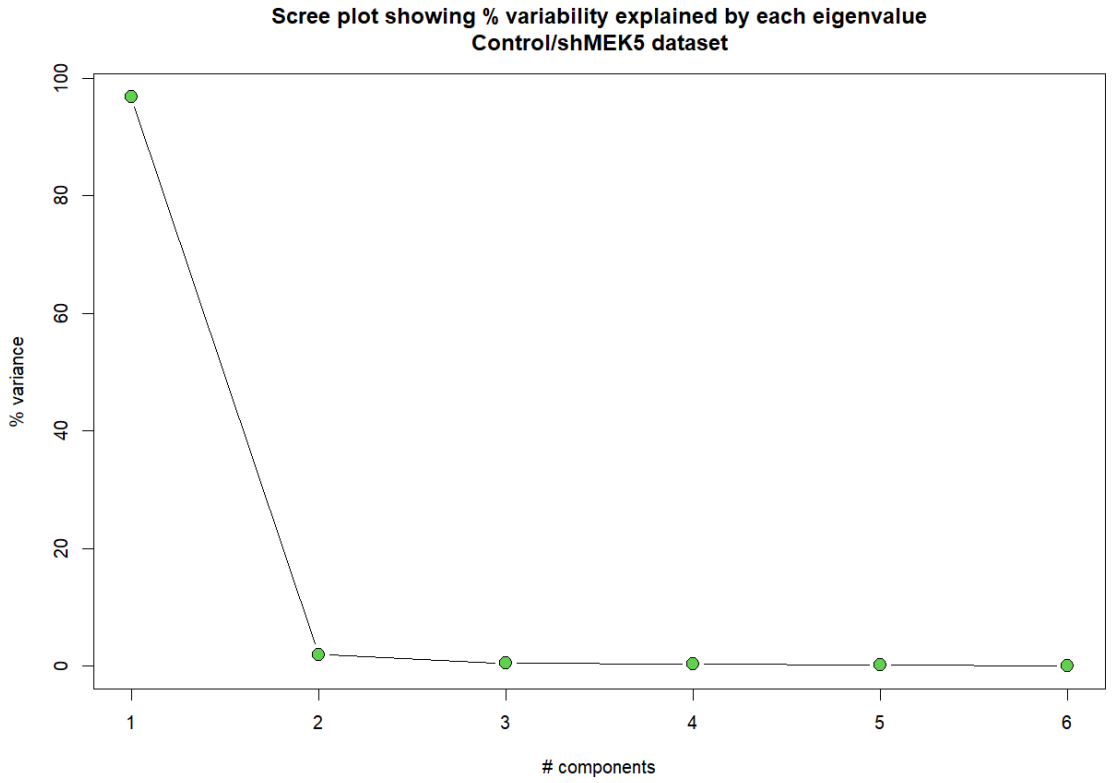
A_23_P124619	A_23_P103877	A_23_P4536	A_33_P3303810
1.352838e-06	7.178993e-06	8.957189e-06	1.119911e-05
A_23_P169437	A_33_P3220698	A_23_P115022	A_23_P206359
5.547899e-05	6.015597e-05	9.337653e-05	1.221593e-04
A_23_P11800	A_23_P393051	A_23_P49155	A_33_P3364741
1.411652e-04	1.637088e-04	1.666161e-04	1.821159e-04
A_23_P47282	A_32_P129752	A_33_P3374723	A_33_P3243449
2.054799e-04	2.971868e-04	2.988188e-04	3.080995e-04
A_23_P67847	A_33_P3507542	A_33_P3256391	A_33_P3335590
3.673980e-04	3.685185e-04	5.226974e-04	7.521148e-04
A_33_P3303649	A_23_P130027	A_23_P2181	A_33_P3281191
7.696162e-04	8.111979e-04	9.019212e-04	9.346272e-04
A_23_P9255	A_23_P49448	A_32_P23125	A_23_P255444
9.560981e-04	1.182082e-03	1.278194e-03	1.488887e-03
A_33_P3325704	A_23_P161190	A_23_P143817	A_32_P192692
1.682492e-03	1.979193e-03	1.990184e-03	2.262023e-03
A_33_P3246833	A_33_P3210099	A_23_P259127	A_24_P270728
2.286564e-03	2.314081e-03	2.590961e-03	2.871526e-03
A_23_P29067	A_23_P10591	A_23_P119943	A_33_P3738458
2.886627e-03	2.896264e-03	2.934402e-03	3.254113e-03
A_23_P42397	A_33_P3329088	A_23_P132910	A_33_P3228460
3.309901e-03	3.508940e-03	3.674942e-03	3.823868e-03
A_23_P26024	A_23_P57268	A_23_P143348	A_23_P23048
3.832026e-03	4.239941e-03	4.345871e-03	4.547803e-03
A_33_P3258061	A_33_P3389827	A_23_P149529	A_32_P164246
4.707844e-03	4.758516e-03	4.871068e-03	5.544282e-03
A_32_P29806	A_23_P97990	A_33_P3313055	A_23_P501933
5.706528e-03	5.727085e-03	6.018284e-03	6.695041e-03
A_33_P3725227	A_33_P3307113	A_23_P122906	A_33_P3813128
7.546632e-03	7.958958e-03	8.126630e-03	8.244654e-03



Sample Visualization: Dimensionality Reduction

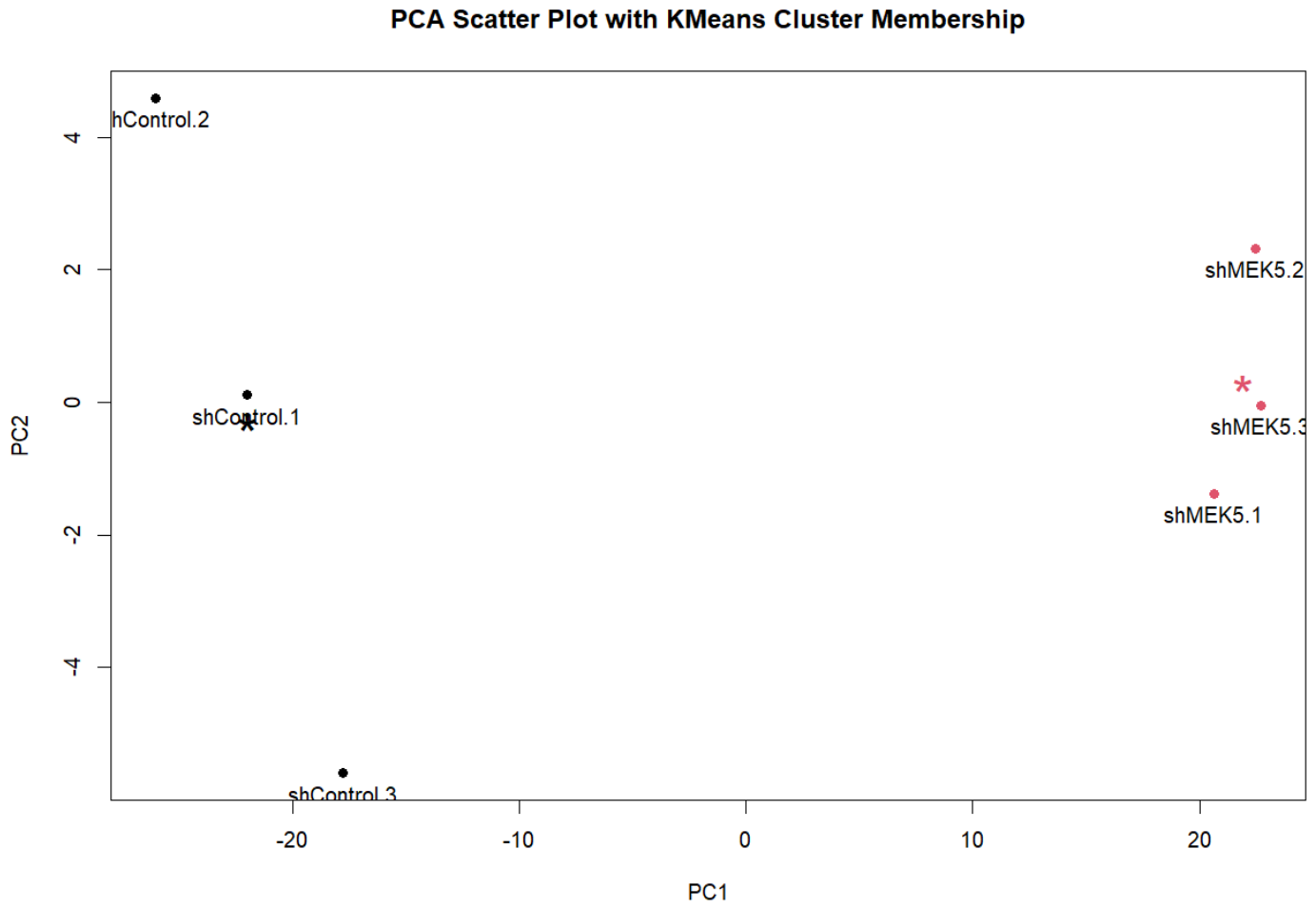
- The dataset was narrowed down to the 60 chosen genes. The Principal Components Analysis (PCA) plot on the left illustrates overall patterns among samples in two-dimensional space, using the first two eigenvalues. Notably, the control and shMEK5 groups exhibit clear distinguishability in the plot. The scree plot on the right shows % variability explained by each eigenvalue. Calculation shows that approximately 98.9% variability in the data is explained by the first two eigenvalues.





Sample Visualization: K-Means Clustering

- K-Means clustering (k = 2) was performed on the first two eigenvalues. In the PCA scatter plot on the right, the samples are colored according to the predicted cluster membership, while the text labels indicate the actual sample identification. The results show that the K-Means clustering aligns well with the actual grouping of the samples.

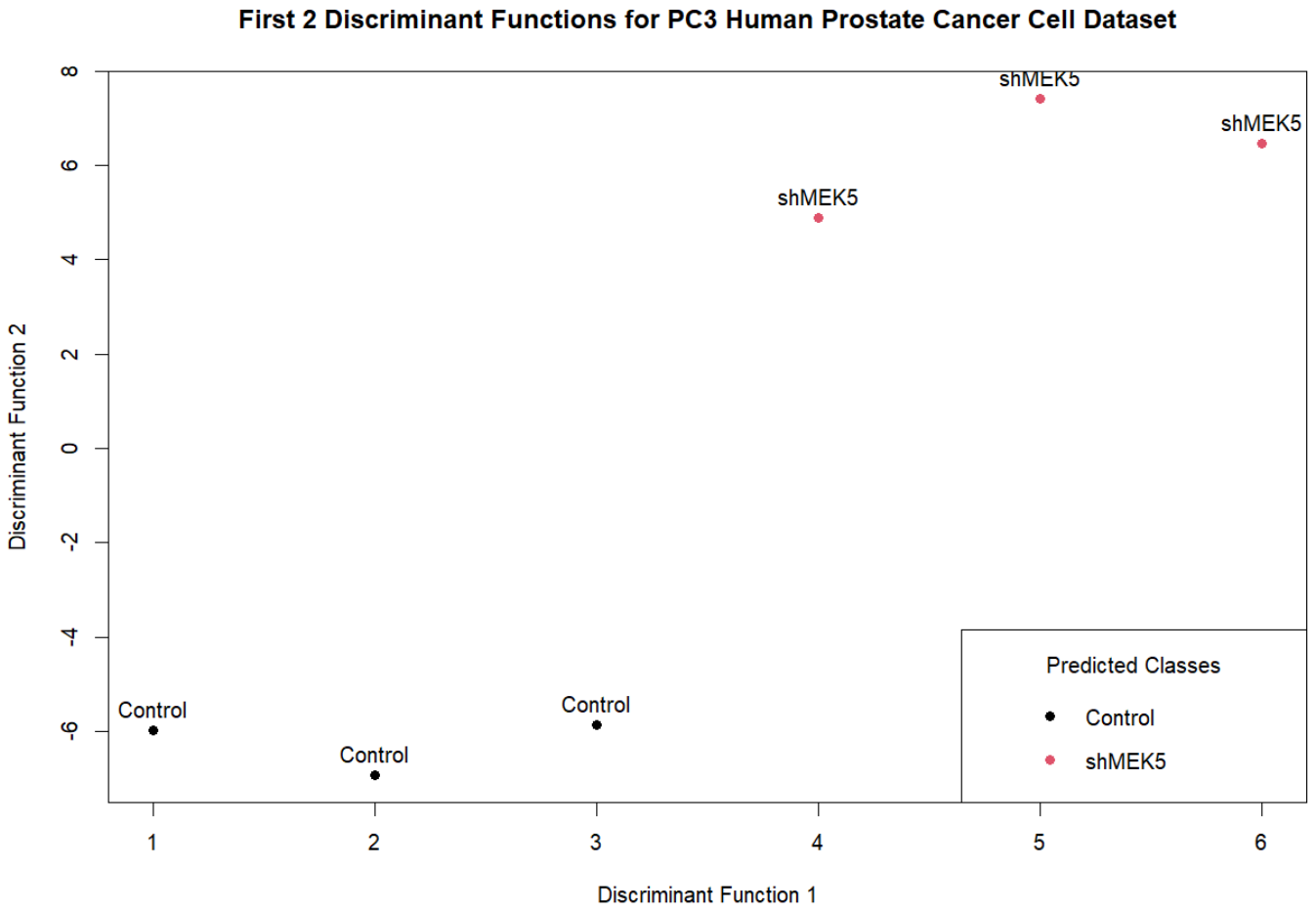


Classification: Linear Discriminant Analysis

- Linear Discriminant Analysis (LDA) was performed using all 60 genes to classify the samples into their respective classes. The confusion matrix is as follows:

	classes	
	Control	shMEK5
Control	3	0
shMEK5	0	3

- The right side features a plot of the first 2 discriminant functions against each other, displayed in an xy plot. Sample points are colored based on their predicted class, while text labels indicate their actual class membership.
- All 6 samples were correctly classified.



Top 5 Discriminant Genes (Positive Direction)

- The identification of the top 5 discriminant genes in the positive direction was determined by their low p-values and a positive log2 fold change. Specifically, these genes exhibit the 5 lowest p-values among the selected genes with a positive log2 fold change. The corresponding probesets are A_23_P124619, A_33_P3303810, A_23_P169437, A_33_P3220698, and A_23_P115022. The gene name and functional information for these 5 genes are as follows:

Probe ID	Gene Symbol	Gene Name	Functional Information (Reference: NCBI Gene)
A_23_P124619	S100A14	S100 calcium binding protein A14	Levels of the encoded protein have been found to be lower in cancerous tissue and associated with metastasis suggesting a tumor suppressor function
A_33_P3303810	LAD1	ladinin 1	The protein encoded by this gene may be an anchoring filament that is a component of basement membranes. It may contribute to the stability of the association of the epithelial layers with the underlying mesenchyme
A_23_P169437	LCN2	lipocalin 2	The protein encoded by this gene plays a role in innate immunity by limiting bacterial growth as a result of sequestering iron-containing siderophores. This protein is thought to be involved in suppression of invasiveness and metastasis.
A_33_P3220698	EPS8L1	EPS8-like 1	This gene encodes a protein that is related to epidermal growth factor receptor pathway substrate 8 (EPS8).
A_23_P115022	TMEM125	transmembrane protein 125	An integral component of membrane

Top 5 Discriminant Genes (Negative Direction)

- The identification of the top 5 discriminant genes in the negative direction was determined by their low p-values and a negative log2 fold change. Specifically, these genes exhibit the 5 lowest p-values among the selected genes with a negative log2 fold change. The corresponding probesets are A_23_P103877, A_23_P4536, A_33_P3364741, A_33_P3374723, and A_33_P3243449. The gene name and functional information for these 5 genes are as follows:

Probe ID	Gene Symbol	Gene Name	Functional Information (Reference: NCBI Gene)
A_23_P103877	LRRC38	leucine rich repeat containing 38	Involved in positive regulation of voltage-gated potassium channel activity and potassium ion transmembrane transport.
A_23_P4536	EPB41L3	erythrocyte membrane protein band 4.1-like 3	Predicted to be involved in several processes, including nervous system development, paranodal junction maintenance, and protein localization to paranode region of axon. Located in cell-cell junction and plasma membrane.
A_33_P3364741	MRC2	mannose receptor, C type 2	The encoded protein mediates the internalization and lysosomal degradation of collagen ligands. Expression of this gene may play a role in the tumorigenesis and metastasis of several malignancies.
A_33_P3374723	ZEB1	zinc finger E-box binding homeobox 1	This gene encodes a zinc finger transcription factor. The encoded protein likely plays a role in transcriptional repression of interleukin 2.
A_33_P3243449	CD70	CD70 molecule	The protein encoded by this gene is a cytokine that belongs to the tumor necrosis factor (TNF) ligand family. It induces proliferation of costimulated T cells, enhances the generation of cytolytic T cells, and contributes to T cell activation. This cytokine is also reported to play a role in regulating B-cell activation, cytotoxic function of natural killer cells, and immunoglobulin synthesis.

Conclusion

- After reviewing the functional information of the top 5 discriminant genes in both positive and negative directions, it appears that PC3 human prostate cancer cells with stable expression of MEK5 shRNA do exhibit a hindrance to tumor progression. This is supported by the increased expression of S100A14 and LCN2, known to suppress tumor growth and metastasis, in MEK5 knockout cells. The reduced expression of MRC2 and the TNF member CD70 in these cells further reinforces this observation. This conclusion is also consistent with previous studies, affirming the impact of MEK5 in prostate tumor development.

Reference

- Broustas, C. (2023). *MEK5-mediated gene expression changes in PC3 human prostate cancer cells*. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE156401
- Mehta, P. B., Jenkins, B. L., McCarthy, L., Thilak, L., Robson, C. N., Neal, D. E., & Leung, H. Y. (2003). MEK5 overexpression is associated with metastatic prostate cancer, and stimulates proliferation, MMP-9 expression and invasion. *Oncogene*, *22*(9), 1381–1389. https://doi.org/10.1038/sj.onc.1206154
- Moore, C. B., Guthrie, E. H., Huang, M. T., & Taxman, D. J. (2010). Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown. *Methods in molecular biology (Clifton, N.J.)*, *629*, 141–158. https://doi.org/10.1007/978-1-60761-657-3_10
- Tai, S., Sun, Y., Squires, J. M., Zhang, H., Oh, W. K., Liang, C. Z., & Huang, J. (2011). PC3 is a cell line characteristic of prostatic small cell carcinoma. *The Prostate*, *71*(15), 1668–1679. https://doi.org/10.1002/pros.21383
- https://www.ncbi.nlm.nih.gov/gene/57402
- https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=3898
- https://www.ncbi.nlm.nih.gov/gene/3934
- https://www.ncbi.nlm.nih.gov/gene/54869
- https://www.ncbi.nlm.nih.gov/gene/128218
- https://www.ncbi.nlm.nih.gov/gene/126755
- https://www.ncbi.nlm.nih.gov/gene/23136
- https://www.ncbi.nlm.nih.gov/gene/9902
- https://www.ncbi.nlm.nih.gov/gene/6935
- https://www.ncbi.nlm.nih.gov/gene/970