Effects of hypoxia on the urothelium transcriptome

Miguel Alburo
University of York

January 24, 2024

Abstract

This report investigates the transcriptomic alterations induced by hypoxic conditions in human bladder epithelium cells, performing a full transcriptomic pipeline from raw reads and alignments before DEA and GSEA. Starting with RNAseq data from four pairs of normoxic-hypoxic cultured urothelium samples, DEA identified 717 DEGs, including 367 upregulated and 350 downregulated genes. The analysis unveiled a surprising downregulation of HIF1A, a key regulator in hypoxia, and upregulation of six genes from a proposed 24-gene hypoxia signature. GSEA gave further insight into the higher-level transcriptional impacts, revealing 11 significantly enriched GO terms. Led by hypoxia, other GO terms were found relating to glycolysis, UV response and DNA repair. Consistencies and also inconsistencies present in our results show the conserved yet variable nature of the complex hypoxia response, emphasizing the potential of additional transcriptomic profiling approaches in further capturing response dynamics in urothelial cells.

1 Introduction

Oxygen, the critical substrate in aerobic respiration, is fundamental to life on Earth. A myriad of biological processes in our own cells depend on the availability of oxygen. Hypoxia is the condition where oxygen availability is insufficient for cells and tissues to perform effectively. The family of transcription factors called hypoxia-inducible factors (HIFs) play a central role in regulating the cellular response to hypoxic conditions. Consisting of HIF-α and HIF-β subunits, HIF is not stabilised at normal oxygen due to the degradation of HIF-α. HIFs regulate a panel of genes involved in various processes involved in reduced oxygen consumption, metabolic switch to glycolysis, cell death (Greijer and Wall 2004), and blood vessel formation (Hirota and Semenza 2006). However, the exact response is a complex system and dependent on factors such as the scale, intensity, duration of hypoxia, and the type of cell (Höckel and Vaupel 2001; Faller 1999).

Experimental and clinical work has consistently shown that hypoxic conditions are favoured by solid tumours, and this unfortunately comes with poorer prognosis and therapeutic resistance (Shannon et al. 2003). While helpful in combating low oxygen,

the processes regulated by HIFs simultaneously promote tumour growth, survival, and malignancy progression (Brahimi-Horn, Chiche, and Pouysségur 2007). Specifically, it is hypothesised that metabolic reprogramming in tumours, the Warburg effect, is a combination of HIF overexpression, tumour suppressor loss of function, and oncogene activation along with other factors (Vaupel and Multhoff 2021). The role of HIFs in cancer is a topic that is ever-growing with many new interesting developments such as implications for the use of HIF inhibitors as cancer therapeutics (Toustrup et al. 2011), highlighting the importance of research into hypoxia.

Transcriptome profiling of cells under normoxic and hypoxic conditions has previously been done using DNA microarrays (Chi et al. 2006; Faller 1999; Gracey, Troll, and Somero 2001). Recent improvements in genomic tools, notably the development of RNA sequencing (RNAseq), now illustrate the massive potential of transcriptomics techniques along with systems biology approaches. Many new analyses use RNAseq which offers further insight into the transcriptome by uncovering additional differentially expressed genes (DEGs), identifying genetic variants (Zhao et al. 2014) and also non-coding RNA (Rao et al. 2019). This immense amount of bioinformatic research into the effects of hypoxia has been key in describing gene expression programmes and identifying gene expression signatures, with experiments performed on different cell types (Chi et al. 2006; Faller 1999) and even fish (Gracey, Troll, and Somero 2001). Their conclusions have shown that the response is both conversed between organisms and yet displays a high degree of specificity between cell types (Chi et al. 2006). Efforts have also been made to develop a hypoxia gene expression classifiers, which can accurately predict hypoxia presence in tumour samples and also prognose (Toustrup et al. 2011; Puente-Santamaría et al. 2022; Ong, Prêle, and Dilley 2023). These classifiers were modelled with other cell types in the head and neck tissue, and are variable in their set of genes for classification. Since we are solely focusing on the effects of hypoxia in cells of the human bladder epithelium, these classifiers may not predict effectively.

The specialised epithelium, i.e. urothelium, lines the bladder and functions as a barrier that prevents toxic substances and pathogens from entering the blood. Literature on hypoxia in the epithelium is sparse considering that, in the UK alone, over 10,000 cases of bladder cancer are diagnosed each year with a shocking 46% chance of survival and an expensive treatment (*Bladder cancer statistics* 2015). Furthermore, seemingly all the analysis is performed on samples from cancer cells. We found no references dealing with samples from healthy cells. Nonetheless, the results conform to the narrative in the literature, emphasising the predictive and prognostic value of hypoxia regulators in bladder cancer (Blick et al. 2015; Theodoropoulos et al. 2005; Yang et al. 2017). Though a classifier model is yet to be made, a 24-gene expression signature was identified which describes what a core set of DEGs should look like (Yang et al. 2017).

In the report, we have completed the full RNAseq analysis workflow. Starting with our 4x4 read data, we first check read quality before aligning reads to the human transcriptome. We then carry out differential expression analysis (DEA) to identify the DEGs on a gene-by-gene basis which we compare with the forementioned 24-gene derived hypoxia signature (Yang et al.). Finally, we use a relatively novel approach called gene set enrichment analysis (GSEA) to find system-level expression patterns

and related Gene Ontology (GO) terms using a predefined collection of gene sets (Subramanian et al. 2005).

2 Brief Methods

We used RNAseq data collected by the Mason lab, which included 4 pairs of normoxia-hypoxia urothelium samples. These samples originated from urothelium cells taken from 4 different individuals and then cultured. The cells underwent re-differentiation to create a biomimetic tissue, which was then divided. One half was maintained under standard normoxic conditions (20% O2), while the other half was cultivated under hypoxic conditions (1% O2). Each sample had 2 reads and, before any analysis, quality control was performed on each read with FastQC. We then used kallisto (Bray et al. 2016) to align the raw reads to the human transcriptome annotation by the Gencode v44 FASTA of protein-coding genes (Frankish et al. 2019). By mapping the transcripts to their genes, we constructed a matrix of gene expression values (TPM) from each kallisto output.

We then performed differential expression analysis (DEA) using sleuth (Pimentel et al. 2017) to obtain the significantly DEGs between the hypoxia and normoxia samples. Additionally, the log2 fold change (log2FC) was calculated to estimate the effect size, i.e. change in expression between hypoxia and normoxia. Prior transformation (+1 to TPM) was done to reduce high fold changes between low expression values. Gene set enrichment analysis (GSEA) was then performed in R using the 'fgsea' package along with MSigDB's hallmark gene set collection, and used pi, the product of log2FC with significance, as the metric for ranking (Sergushichev 2016; Liberzon et al. 2015).

3 Results

FastQC reports showed that all reads were of good quality and passed most criteria. However, all reads failed for 'Sequence Duplication Levels', and some reads failed for 'Per Base Sequence Content'; informing that reads had a small number of unique sequences that were overly represented in the sequencing data, and base composition was imbalanced. After some research, we found that these errors are not uncommon and may be due to techniques in library preparation or PCR.

DEA revealed 717 DEGs out of around 20,000 protein coding genes (Figure 1), which were selected if: significance $q < 0.05$ and log2FC > 1 or < -1. These included 367 upregulated (log2FC$>$1) and 350 downregulated (log2FC$<$-1) genes, displaying the magnitude of the hypoxia response on the transcriptome. HIF1A, the gene that encodes subunit HIF1α, was found to be significantly downregulated (-1.06 log2FC). When compared with the 24-gene hypoxia signature (Figure 2) selected only 6 DEGs, all upregulated, from the signature that passed our criteria above. One of these, SYD2, was ranked third in pi, inferring the biological significance of this gene within the signature. Finally, GSEA revealed 11 significantly enriched GO terms (Figure 3), including 9 upregulated and 2 downregulated gene sets. As expected, the top hit from the leading edge was the 'hallmark hypoxia' gene set. The following GO terms covered other hypoxia-related processes, notably glycolysis, but also DNA repair and UV response.

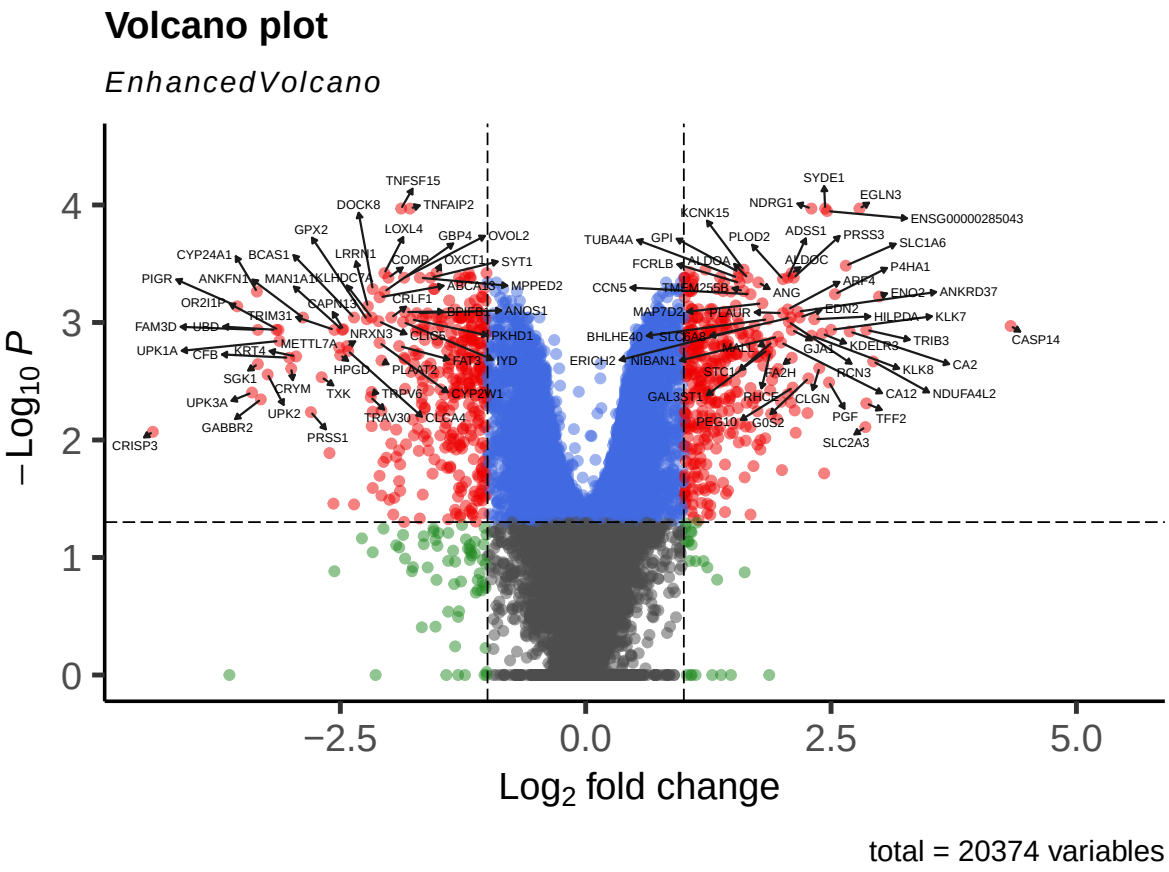

Figure 1: Volcano plot showing results of DEA. Labels indicate top 100 in pi metric.

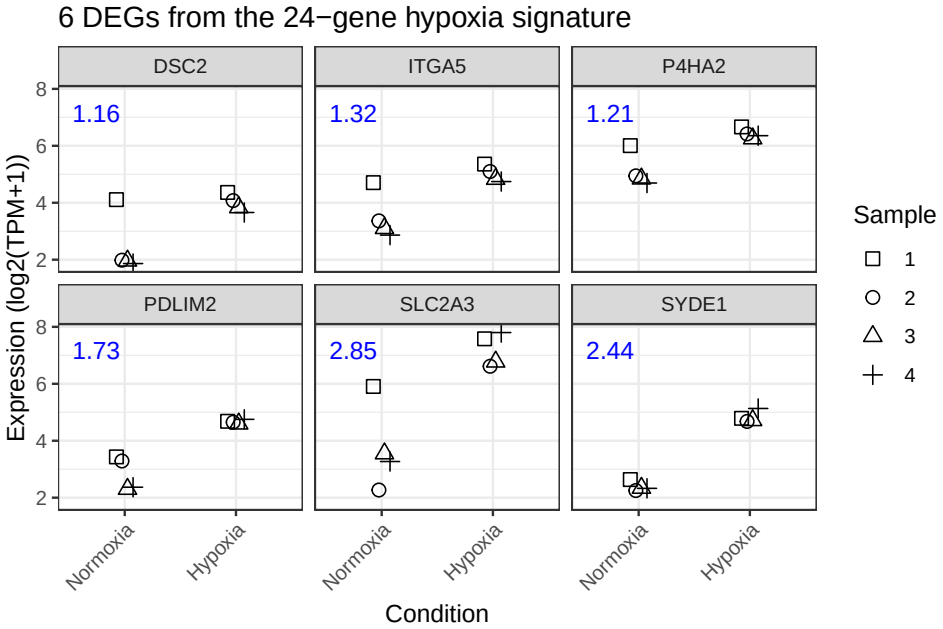

Figure 2: Plot showing the expression of 6 genes from the 24-gene hypoxia signature that were found to be differentially expressed from our data. log2FC values are located on the top-left in blue

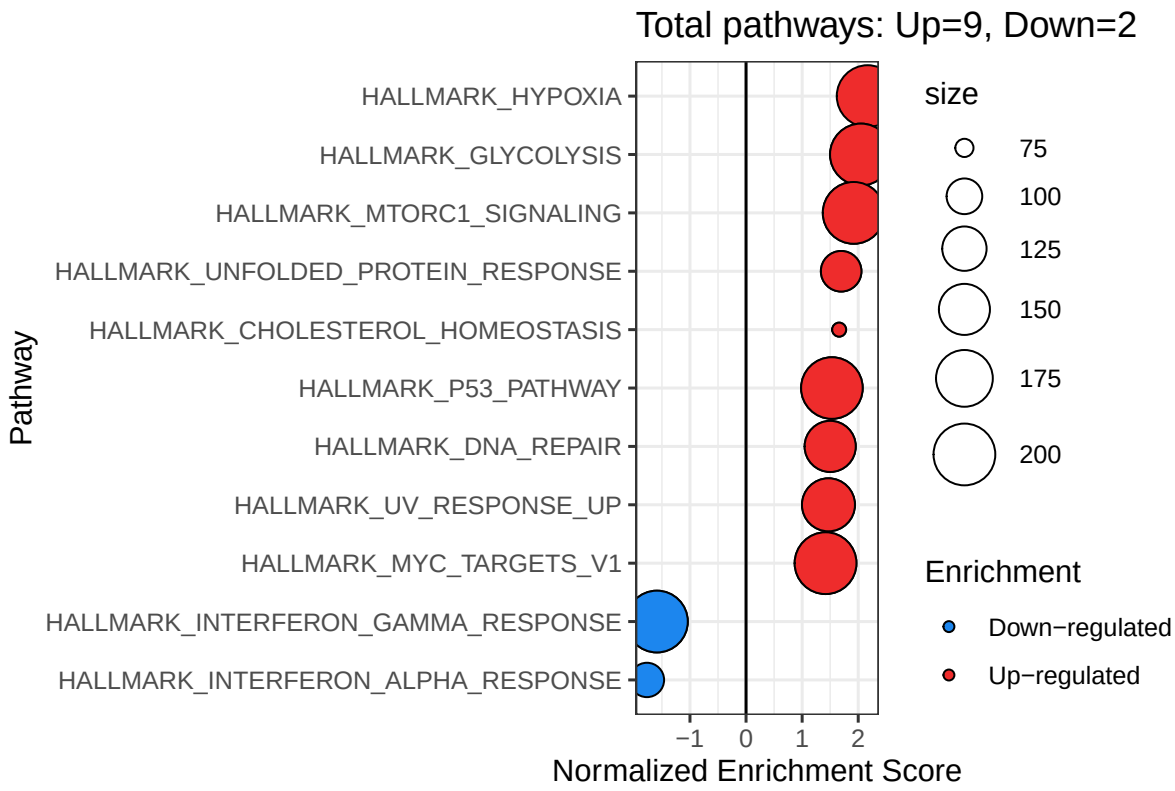

Figure 3: Plot of GO terms and their normalised enrichment score (NES) along with the size of the gene set (*How to do Gene Set Enrichment Analysis (GSEA) in R* 2024).

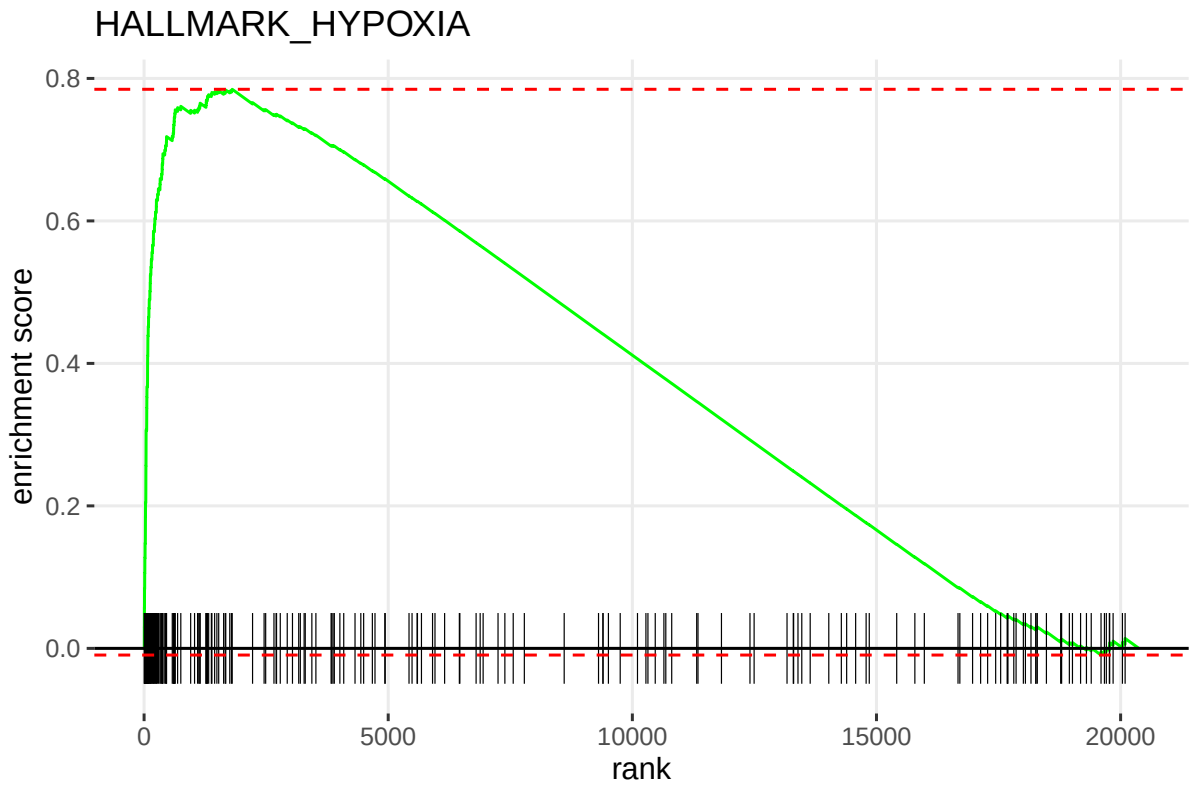

Figure 4: Enrichment plot of the hallmark hypoxia gene set.

4 Discussion

We found the downregulation of HIF1A a big surprise. The product, HIF1α, stabilises HIF and regulates the hypoxia response causing us to initially hypothesise upregulation. Downregulated HIF1A could result from the period of exposure to hypoxic conditions, though we were not given any details about hypoxia duration from the data. We also expected the 24-gene hypoxia signature to identify more DEGs than the 6 that were found, but this is likely due to our DEG selection criteria's log2FC threshold of 1. Namely CAV1, a known major hypoxia regulator that is present in the 24-gene signature, was not significantly differentially expressed. The results from GSEA gave us a better interpretation of the hypoxia response than DEA, and correctly evaluated the leading GO terms as hypoxia and glycolysis. Though we had initially expected angiogenesis as a significantly enriched GO term, we realised in hindsight that the RNAseq data came from in vitro samples which would have likely prevented any angiogenesis response.

5 Conclusion

In summary, we have given an overview of the hypoxic response in the urothelial transcriptome, concluding that hypoxia significantly affects the expression of a huge number of genes, including only 6 from a 24-gene signature. Furthermore, we show that hypoxia affects sets of genes related to glycolysis, DNA repair, UV response and more. These findings also show the ease and power of this experiment workflow to obtain biological interpretations from RNAseq data, and highlight the importance of these bioinformatics techniques such as DEA and GSEA along with their convenient R packages.

References

Greijer, A E and E van der Wall (Oct. 2004). "The role of hypoxia inducible factor 1 (HIF-1) in hypoxia induced apoptosis". In: *Journal of Clinical Pathology* 57.10, pp. 1009–1014. ISSN: 0021-9746. DOI: 10.1136/jcp.2003.015032. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1770458/ (visited on 01/17/2024).

Hirota, Kiichi and Gregg L. Semenza (July 1, 2006). "Regulation of angiogenesis by hypoxia-inducible factor 1". In: *Critical Reviews in Oncology/Hematology* 59.1, pp. 15–26. ISSN: 1040-8428. DOI: 10.1016/j.critrevonc.2005.12.003. URL: https://www.sciencedirect.com/science/article/pii/S1040842806000138 (visited on 01/17/2024).

Höckel, Michael and Peter Vaupel (Feb. 21, 2001). "Tumor Hypoxia: Definitions and Current Clinical, Biologic, and Molecular Aspects". In: *JNCI: Journal of the National Cancer Institute* 93.4, pp. 266–276. ISSN: 0027-8874. DOI: 10.1093/jnci/93.4.266. URL: https://doi.org/10.1093/jnci/93.4.266 (visited on 01/16/2024).

Faller, D. V. (Jan. 1999). "Endothelial cell responses to hypoxic stress". In: *Clinical and Experimental Pharmacology & Physiology* 26.1, pp. 74–84. ISSN: 0305-1870. DOI: 10.1046/j.1440-1681.1999.02992.x.

Shannon, Aoife M et al. (Aug. 1, 2003). “Tumour hypoxia, chemotherapeutic resistance and hypoxia-related therapies”. In: *Cancer Treatment Reviews* 29.4, pp. 297–307. ISSN: 0305-7372. DOI: 10.1016/S0305-7372(03)00003-3. URL: https://www.sciencedirect.com/science/article/pii/S0305737203000033 (visited on 01/16/2024).

Brahimi-Horn, M. Christiane, Johanna Chiche, and Jacques Pouysségur (Dec. 2007). “Hypoxia and cancer”. In: *Journal of Molecular Medicine (Berlin, Germany)* 85.12, pp. 1301–1307. ISSN: 0946-2716. DOI: 10.1007/s00109-007-0281-3.

Vaupel, Peter and Gabriele Multhoff (2021). “Revisiting the Warburg effect: historical dogma versus current understanding”. In: *The Journal of Physiology* 599.6. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1113/JP278810, pp. 1745–1757. ISSN: 1469-7793. DOI: 10.1113/JP278810. URL: https://onlinelibrary.wiley.com/doi/abs/10.1113/JP278810 (visited on 01/17/2024).

Toustrup, Kasper et al. (Sept. 1, 2011). “Development of a hypoxia gene expression classifier with predictive impact for hypoxic modification of radiotherapy in head and neck cancer”. In: *Cancer Research* 71.17, pp. 5923–5931. ISSN: 1538-7445. DOI: 10.1158/0008-5472.CAN-11-1182.

Chi, Jen-Tsan et al. (Mar. 2006). “Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers”. In: *PLoS medicine* 3.3, e47. ISSN: 1549-1676. DOI: 10.1371/journal.pmed.0030047.

Gracey, Andrew Y., Joshua V. Troll, and George N. Somero (Feb. 13, 2001). “Hypoxia-induced gene expression profiling in the euryoxic fish Gillichthys mirabilis”. In: *Proceedings of the National Academy of Sciences* 98.4. Publisher: Proceedings of the National Academy of Sciences, pp. 1993–1998. DOI: 10.1073/pnas.98.4.1993. URL: https://www.pnas.org/doi/full/10.1073/pnas.98.4.1993 (visited on 01/16/2024).

Zhao, Shanrong et al. (Jan. 16, 2014). “Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells”. In: *PLOS ONE* 9.1. Publisher: Public Library of Science, e78644. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0078644. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0078644 (visited on 01/17/2024).

Rao, Mohan S. et al. (Jan. 22, 2019). “Comparison of RNA-Seq and Microarray Gene Expression Platforms for the Toxicogenomic Evaluation of Liver From Short-Term Rat Toxicity Studies”. In: *Frontiers in Genetics* 9, p. 636. ISSN: 1664-8021. DOI: 10.3389/fgene.2018.00636. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6349826/ (visited on 01/16/2024).

Puente-Santamaría, Laura et al. (May 31, 2022). “Hypoxia classifier for transcriptome datasets”. In: *BMC Bioinformatics* 23.1, p. 204. ISSN: 1471-2105. DOI: 10.1186/s12859-022-04741-8. URL: https://doi.org/10.1186/s12859-022-04741-8 (visited on 01/16/2024).

Ong, Huan Ting, Cecilia M. Prêle, and Rodney J. Dilley (Apr. 17, 2023). “Using RNA-seq to identify suitable housekeeping genes for hypoxia studies in human adipose-derived stem cells”. In: *BMC Molecular and Cell Biology* 24.1, p. 16. ISSN: 2661-8850. DOI: 10.1186/s12860-023-00475-4. URL: https://doi.org/10.1186/s12860-023-00475-4 (visited on 01/16/2024).

Bladder cancer statistics (May 14, 2015). Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer (visited on 01/18/2024).

Blick, C et al. (Aug. 11, 2015). “Identification of a hypoxia-regulated miRNA signature in bladder cancer and a role for miR-145 in hypoxia-dependent apoptosis”. In: *British Journal of Cancer* 113.4, pp. 634–644. ISSN: 0007-0920. DOI: 10.1038/bjc.2015.203. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4647685/ (visited on 01/16/2024).

Theodoropoulos, Vasilios. E. et al. (2005). “Evaluation of hypoxia-inducible factor 1α overexpression as a predictor of tumour recurrence and progression in superficial urothelial bladder carcinoma”. In: *BJU International* 95.3, pp. 425–431. ISSN: 1464-410X. DOI: 10.1111/j.1464-410X.2005.05314.x. (Visited on 01/18/2024).

Yang, Lingjian et al. (Aug. 14, 2017). “A Gene Signature for Selecting Benefit from Hypoxia Modification of Radiotherapy for High-Risk Bladder Cancer Patients”. In: *Clinical Cancer Research* 23.16, pp. 4761–4768. ISSN: 1078-0432. DOI: 10.1158/1078-0432.CCR-17-0038. URL: https://doi.org/10.1158/1078-0432.CCR-17-0038 (visited on 01/18/2024).

Subramanian, Aravind et al. (Oct. 25, 2005). “Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles”. In: *Proceedings of the National Academy of Sciences* 102.43. Publisher: Proceedings of the National Academy of Sciences, pp. 15545–15550. DOI: 10.1073/pnas.0506580102. URL: https://www.pnas.org/doi/10.1073/pnas.0506580102 (visited on 01/22/2024).

Bray, Nicolas L. et al. (May 2016). “Near-optimal probabilistic RNA-seq quantification”. In: *Nature Biotechnology* 34.5, pp. 525–527. ISSN: 1546-1696. DOI: 10.1038/nbt.3519.

Frankish, Adam et al. (Jan. 8, 2019). “GENCODE reference annotation for the human and mouse genomes”. In: *Nucleic Acids Research* 47 (D1), pp. D766–D773. ISSN: 1362-4962. DOI: 10.1093/nar/gky955.

Pimentel, Harold et al. (July 2017). “Differential analysis of RNA-seq incorporating quantification uncertainty”. In: *Nature Methods* 14.7, pp. 687–690. ISSN: 1548-7105. DOI: 10.1038/nmeth.4324.

Sergushichev, Alexey A. (June 20, 2016). *An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation*. Pages: 060012 Section: New Results. DOI: 10.1101/060012. URL: https://www.biorxiv.org/content/10.1101/060012v1 (visited on 01/23/2024).

Liberzon, Arthur et al. (Dec. 23, 2015). “The Molecular Signatures Database (MSigDB) hallmark gene set collection”. In: *Cell Systems* 1.6, pp. 417–425. ISSN: 2405-4712. DOI: 10.1016/j.cels.2015.12.004.

How to do Gene Set Enrichment Analysis (GSEA) in R (Jan. 18, 2024). Bioinformatics Breakdown. URL: https://bioinformaticsbreakdown.com/how-to-gsea/ (visited on 01/18/2024).