Neoplasia 79 (2026) 101327

Contents lists available at ScienceDirect

Neoplasia

journal homepage: www.elsevier.com/locate/neo

Short Communication

Inaccurate interpretation of hypoxia signatures from transcriptomic analyses in neuro-oncology

Rebecca K Simmons [a,b], Pratima Raut [a,b], Subhi Talal Younes [a,b], Marco Gallo [a,b,c,*]

[a] Department of Pediatrics, Baylor College of Medicine, Houston, TX, 77030, United States of America
[b] Cancer and Hematology Center, Texas Children's Hospital, Houston, TX, 77030, United States of America
[c] Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX 77030, United States of America

ARTICLE INFO

Keywords:
Hypoxia
Brain
Brain tumors
Bioinformatics

ABSTRACT

We aim to highlight problems with the biological interpretation of computationally-derived hypoxia signatures for neuro-oncology research. By reviewing important experimental work in the field, we discuss common misinterpretations of the concepts of normoxia and hypoxia. In fact, the concentration of oxygen in the adult brain is in the range of 1-7%. Unfortunately, this oxygen concentration range is often mistakenly considered hypoxic in neuro-oncology experiments, although it should be considered physioxic. Further, given the lack of a universal hypoxic signature, we stress the importance of deriving cell type-specific hypoxic signatures using carefully-executed experiments.

Introduction to the problem

The availability of large numbers of transcriptomic studies requires the development of approaches to make sense of large amounts of data. The basic question that most researchers ask is "What genes are differentially expressed in condition A compared to condition B?" and "What do these differentially expressed genes mean from a biological standpoint?" Such lists typically include hundreds if not thousands of genes. There is therefore a clear need to parse such a long list of genes into molecular processes and pathways that are biologically meaningful.

One of the most used computational approaches to extract biological meaning from transcriptomic data is Gene Set Enrichment Analysis (GSEA) [1]. GSEA allows the identification of transcriptional signatures, or gene sets, that are enriched in an experimentally-derived list of genes. One of the most widely used repositories of gene sets is the Molecular Signatures Database (MSigDB). MSigDB includes thousands of reference gene sets, and derived HALLMARK gene sets 10 years ago with the intent to reduce redundancies [2]. These signatures have been incredibly useful for understanding biological principles and the results are usually biologically meaningful, likely a reflection of their high quality and manual curation.

GSEA and MSigDB represent fundamental tools for meaningful discovery. However, the interpretation of the enrichment results needs to be carefully examined. We want to emphasize that it is the responsibility of the user, not of the developers of these tools, to critically evaluate the results of their own analyses.

We have noticed an increasing trend of taking the names of gene sets at face value without critically examining them for accuracy, biological relevance and the experimental conditions used to derive them. HALLMARK_HYPOXIA is one of the MSigDB gene sets that is regularly misinterpreted by users.

Hypoxia signatures

The HALLMARK_HYPOXIA signature includes 200 genes. Details on the gene list can be found at https://www.gsea-msigdb.org/gsea/msigdb/human/gene_families.jsp?geneSetName=HALLMARK_HYPOXIA. Anyone actually looking at the gene list will realize that most of these genes have nothing to do with hypoxia. We are not the first ones to bring up this problem. Puente-Santamaria et al [3] systematically addressed this issue, for instance. The authors performed a meta-analysis of 430 RNA-seq samples from 43 studies that included 34 cell types. They identified a list of 291 genes that are likely to be biologically and meaningfully affected by hypoxia. The overlap between their signature and the HALLMARK_HYPOXIA signature included only 64 genes. As the authors note in their paper, some of the discrepancies are due to the inclusion of genes in the HALLMARK_HYPOXIA gene set that are affected by hypoxia only in specific cell types, but not across different

* Corresponding author at: Department of Pediatrics, Baylor College of Medicine, Houston, TX, 77030, United States of America.
E-mail address: marco.gallo@bcm.edu (M. Gallo).

https://doi.org/10.1016/j.neo.2026.101327
Received 7 May 2026; Received in revised form 4 June 2026; Accepted 9 June 2026

cell types. GSEA users should therefore be aware that even the HALLMARK gene sets are not universally applicable to all cell types and experimental conditions.

We want to bring the attention to a few genes that are highly significant to the neuro oncology field and that are included in the HALLMARK_HYPOXIA signature.

- EGFR (epidermal growth factor receptor). This gene plays significant roles in brain tumors, including in glioblastoma, where it is commonly amplified [4,5].
- JUN and other genes encoding AP-1 family of transcription factors. This gene is expressed in response to a variety of cell stressors (not just hypoxia) and is a hallmark of the mesenchymal-like/injury response molecular subtype of glioblastoma [6,7].
- PDGFB (platelet-derived growth factor subunit beta) is an important mitogen with roles in development and cancer (reviewed in [8]).
- CXCR4 (C-X-C chemokine receptor 4) is expressed across brain tumor types, including glioblastoma and pediatric diffuse midline glioma, where it has a role in tumor cell invasion [9–11]. Its expression is enriched in mesenchymal-like glioblastoma cells [12].
- BCL2 (B-cell lymphoma 2) is known to antagonize cell death (reviewed extensively in [13]).

The dependence of expression of these genes on hypoxia in brain tumors (or other cell systems) is not currently established experimentally. These are just examples of genes with pleiotropic functions that are included in the HALLMARK_HYPOXIA signature. Expression of these genes is likely determined by a plethora of different molecular pathways, and the contribution of hypoxia to their transcription is unlikely to be major.

We want to emphasize that hypoxia has pleiotropic effects on cells, and that GSEA is probabilistic in nature and does not imply mechanism. In fact, we think that GSEA is a great first step toward making sense of large transcriptional datasets. However, in this commentary, we also want to emphasize that GSEA results need to be interpreted thoughtfully in light of the biology of any specific cell type being investigated.

What is 'hypoxia'?

An important consideration that is largely missing from the conversation in neuro oncology is the definition of hypoxia in the context of the brain. There is a general assumption that hypoxia means "not normoxia," or the oxygen concentration of 21% we live in. However, it is well known that the actual concentration of oxygen in the brain is well below 21%, with most estimates converging on the 2-7% range in the adult brain, and lower (<1%) in the fetal brain (reviewed extensively in [14]).

In general, tumors have lower oxygen pressures than their non-malignant tissue of origin. This concept is extensively reviewed in [15].

In brain tumors, oxygen concentrations have been estimated using several approaches. In one study, oxygen pressure was measured directly in WHO grade 3 and 4 gliomas in patients using electrodes implanted in the peritumoral and intratumoral regions [16]. Under awake conditions, oxygen pressure was 17.9 ± 9.3 mmHg in the peritumoral regions and 9.2 ± 5.8 mmHg in the intratumoral regions, representing statistically significant differences. These values correspond approximately to 2.4% oxygen concentration in the peritumoral regions and 1.2% in the intratumoral regions. An alternative approach indirectly measures oxygen pressure by quantifying the cellular uptake of 3-[18 F]-fluoro-1-(2-nitro-1-imidazolyl)-2-propanol ([^{18}F]-FMISO) in hypoxic cells by positron emission tomography. Using this approach, the intratumoral oxygen pressure in glioblastoma was estimated to be 4.8 ± 1.9 mmHg [17], or approximately 0.6%. Based on the complementary measurements, intratumoral oxygen concentration for high-grade glioma (including glioblastoma) should therefore be below 2%.

Because of this, Calvo Tardon et al [18] introduced the concept of physiologically relevant oxygen concentrations, or physioxia, in the context of brain tumors. Culturing a brain tumor cell line at 5% oxygen concentration therefore does not correspond to 'hypoxic conditions', but physioxic. Oxygen concentrations would need to be 1-2% or below to be truly hypoxic. This concept, and the definitions of normoxia, hypoxia and physioxia, are extensively reviewed in [15].

Calvo Tardon and colleagues proved this conceptual problem experimentally. They cultured glioblastoma cell lines in 1%, 5% and 21% oxygen concentrations. Transcriptome analyses showed that 1,040 genes were differentially expressed between 5% and 21% oxygen concentrations, whereas only 36 genes were differentially expressed between the 5% and 1% oxygen concentrations. 11 genes were unique to the 5% vs 1% comparison. Finally, the authors correctly pointed out the need to distinguish between genes that are differentially expressed in hypoxia and the ones that are both differentially expressed and have hypoxia response elements (HREs) in their putative cis regulatory regions to distinguish between likely direct and indirect effects of hypoxia on transcription. In light of these results, we suggest that experiments performed with brain tumor cell lines at oxygen concentrations between 1% and 5% to be defined as physioxic, not hypoxic.

In a systematic review of 70 hypoxia signatures, Di Giovannantonio et al [19] found that no gene was included in all of them, showcasing differences in biological responses to hypoxia in different cell types, and experimental and/or technical variability between experimental conditions. These results clearly indicate that the definition of a hypoxia signature is far from consensus, and cell types and experimental conditions need to be carefully considered before reaching conclusions about the involvement of hypoxia in a biological process.

In a practical way, many genes in the HALLMARK_HYPOXIA gene set are associated with metabolic functions. Gene ontology (GO) analysis of the 200 members of this gene set reveals that, in fact, most genes are defined by metabolic processes (Fig. 1). Only 26 genes are recognized as members of a HIF1 pathway (PID_HIF1_TFPATHWAY). Undisputably, oxygen levels not only regulate HIF1 and its downstream targets, but also other important aspects of cell biology, including metabolic and chromatin dynamics. However, many other factors also do as well. It is very important therefore to distinguish between (a) genes directly regulated by hypoxia, (b) genes indirectly regulated by hypoxia, ie genes regulated by targets of HIF-family transcription factors, and (c) metabolic changes that have nothing to do with hypoxia. These distinctions are not easy to make, but carefully controlled experiments should be able to at least refine the hypoxia signature for any given cell type.

Metabolism and oxygen levels are intertwined and both converge on epigenetic regulation of transcription (Fig. 2). A class of enzymes known as alpha-ketoglutarate dependent dioxygenases are well-studied and provide the most straightforward example of this. TET enzymes (TET1-3) and JmjC-domain histone demethylases are the best studied members of this enzymatic class [20–23]. They act as nodes to link oxygen and metabolic state of a cell to chromatin changes that eventually lead to transcriptional changes, by requiring oxygen and alpha-ketoglutarate (an intermediate of the TCA cycle) to perform their enzymatic functions. If oxygen and alpha-ketoglutarate levels are sufficient, TET enzymes catalyze a stepwise series of chemical reactions that result in removal of DNA methylation at CpG sites. Similarly, under appropriate conditions, JmjC-domain histone demethylases remove methyl marks from histones. However, these enzymatic functions cannot proceed if oxygen or alpha-ketoglutarate levels are too low. Please note that a decrease in alpha-ketoglutarate caused by changes in glucose levels, for instance, but without any changes in oxygen concentration, could lead to a decrease in enzymatic function similar to what would be expected for hypoxia. The effects of hypoxia on metabolism and transcription are therefore so wide-ranging, through HIF-dependent and -independent mechanisms, that it is very important to accurately and precisely evaluate the results of GSEA analyses before reaching the conclusion that an experiment resulted in positive or negative enrichment of the HALLMARK_HYPOXIA gene set.

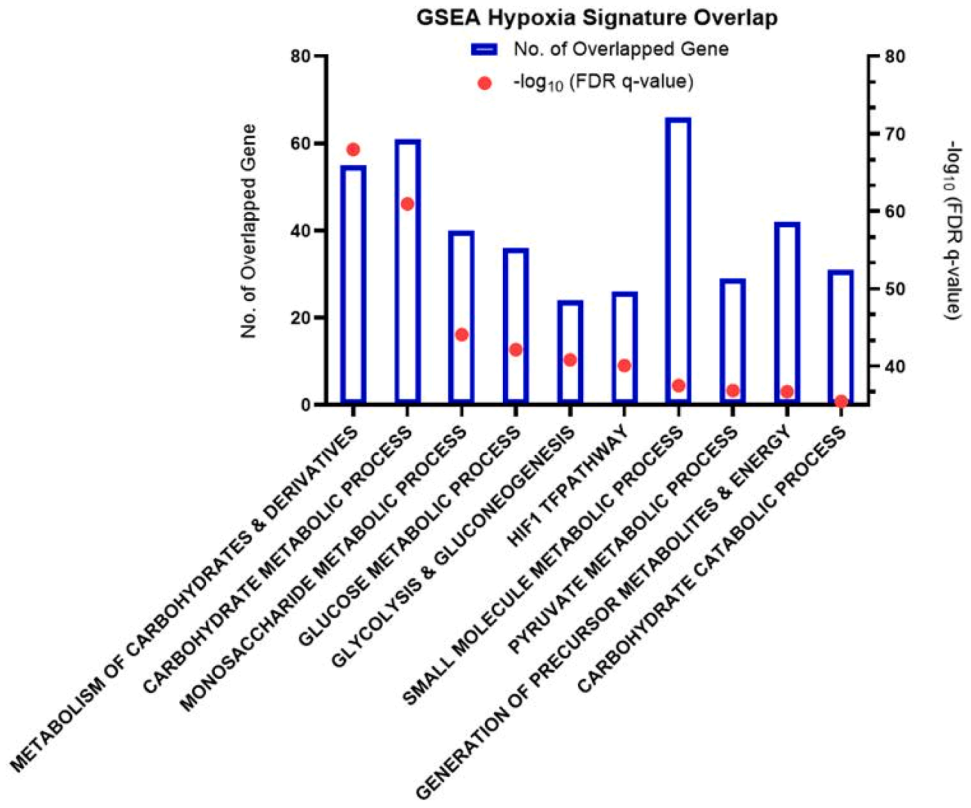

Fig. 1. . **Gene Ontology enrichment of genes in the HALLMARK_HYPOXIA gene set.** The 200 genes in the HALLMARK_HYPOXIA gene set were used for Gene Ontology (GO) analysis, and some GO categories were significantly enriched in the gene set and are displayed here. This diagram visualizes both the number of genes and the associated -$\log_{10}$(FDR q-value) for each GO category that is significantly enriched in the HALLMARK_HYPOXIA gene set. The y-axis on the left refers to the number of genes in each GO category and corresponds to the blue bars. The Y axis on the right refers to negative $\log_{10}$(FDR q-value), which are represented as red dots for each category listed on the x axis.

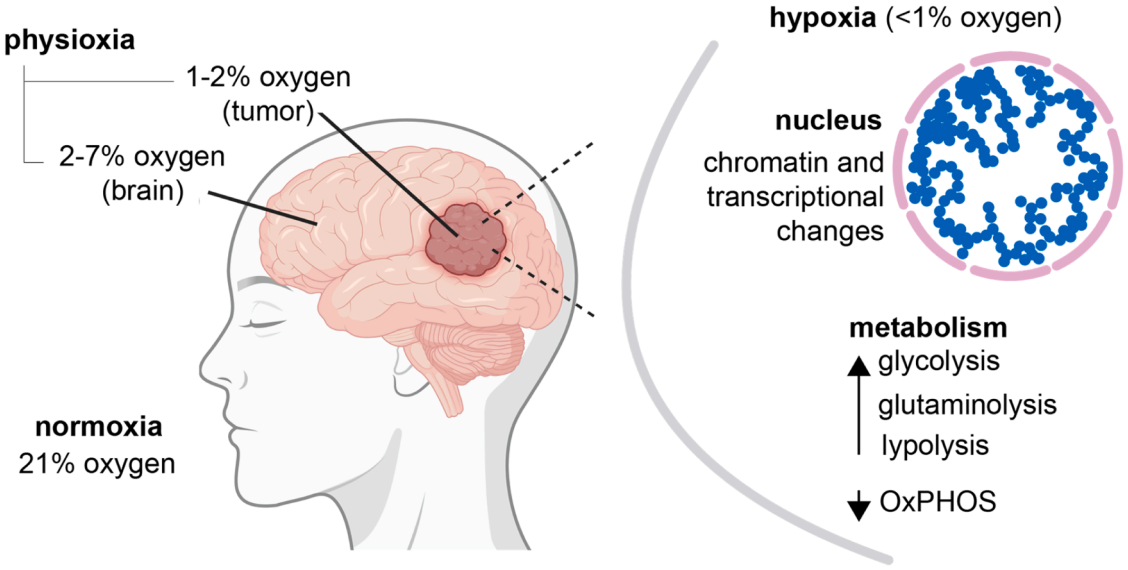

Fig. 2. . **Hypoxia has wide-ranging effects on cellular molecular profiles.** In this illustration, we showcase the need to define hypoxia in the context of the physioxic state of a specific tissue or tumor type. For instance, oxygen concentration in brain tissue is in the 2-7% range, whereas oxygen concentration in brain tumors (eg glioblastoma) is 1-2%. In comparison, atmospheric oxygen concentration at sea level is about 21% (normoxia). Hypoxia has pleiotropic effects on cell physiology, including alterations of epigenetic marks leading to chromatin re-organization and transcriptional and metabolic changes.

Critical interpretation of enrichment of hypoxia signatures

We want to emphasize that informatically-determined enrichment for a hypoxia signature cannot be taken at face value to indicate hypoxia per se. Unfortunately, we have witnessed a trend toward accepting the output of computational studies as the ground truth. This attitude is dangerous, because it leads to the uncritical acceptance of informatic results as biologically valid and to the dissemination of this

misconception until it becomes entrenched. Computational studies are prone to error, over-interpretation and bias as much as – or more than – functional wet lab assays. In the specific context of GSEA, users must be cautious and critical about the outcomes of their analyses. They need to familiarize themselves with the genes that are present in a specific gene set, research their biological functions and become aware of the caveats inherent in their analyses. Our fear is that incorrect interpretations of biological processes are spreading in the community and are becoming cemented as truth, confounding the premises and designs of future studies.

Here, we would like to highlight some practical steps that should be taken when performing GSEA, or other informatic analyses of biological data:

Ensuring that gene sets are applicable to the cell type and experimental platform being considered. Using what we have discussed above about the HALLMARK_HYPOXIA gene set, it is fundamental to determine whether a specific gene set is relevant to brain tumor cells, for instance. Researchers need to understand how the gene set was determined, what experimental data were used to derive the gene set, and how it was manually curated. Based on available experimental data (reviewed in [14]), a threshold of at most 1% oxygen concentration should be used as definition of hypoxia from brain tumor analyses. Physiological relevance to a cell type or system being investigated cannot be assumed. In most cases, the experimenter might need to perform carefully-controlled experiments with relevant tumor models to derive an accurate hypoxia signature for the malignancy being investigated.

(1) Critical assessment of gene sets. What are the known biological functions of the genes in the list? Are there any genes that are clearly either involved in pleiotropic functions or that are linked to a biological function without proper experimental support?
(2) Declaration of limitation of informatic analyses. The caveats of each analyses need to be properly stated in a research paper.

These three critical steps are basic pillars of scientific inquiry that we should expect of every experiment. This approach is expected of wet lab experiments, but for some reason the community has been more lax with informatic analyses. This might just reflect the relative newness of bioinformatics compared to molecular and biochemical approaches, and to the more restricted expertise required for computational studies. However, we feel it is important to emphasize the need for more critical approaches to bioinformatics, given the increasing role they have in fundamental research and the growing number of researchers performing these analyses.

CRediT authorship contribution statement

Rebecca K Simmons: Writing – review & editing, Formal analysis. **Pratima Raut:** Writing – review & editing, Data curation. **Subhi Talal Younes:** Writing – review & editing, Data curation. **Marco Gallo:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Formal analysis, Conceptualization.

Declaration of competing interest

The authors have no competing interests.

Acknowledgements

We would like to acknowledge funding from the Ependymoma Research Foundation (#A6M05172501) and Alex's Lemonade Stand Foundation (#1436154) to MG. Some elements of Fig. 2 were created in BioRender [Simmons, R. (2026)].

References

[1] A. Subramanian, P. Tamayo, V.K. Mootha, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. 102 (2005) 15545–15550, https://doi.org/10.1073/pnas.0506580102.
[2] A. Liberzon, C. Birger, H. Thorvaldsdóttir, et al., The molecular signatures database (MSigDB) hallmark gene set collection, Cell Syst. 1 (2015) 417–425, https://doi.org/10.1016/j.cels.2015.12.004.
[3] L. Puente-Santamaria, L. Sanchez-Gonzalez, N. Pescador, et al., Formal meta-analysis of hypoxic gene expression profiles reveals a universal gene signature, Biomedicines 10 (2022) 2229, https://doi.org/10.3390/biomedicines10092229.
[4] T.A. Libermann, H.R. Nusbaum, N. Razon, et al., Amplification, enhanced expression and possible rearrangement of EGF receptor gene in primary human brain tumours of glial origin, Nature 313 (1985) 144–147, https://doi.org/10.1038/313144a0.
[5] C.W. Brennan, R.G.W. Verhaak, A. McKenna, et al., The somatic genomic landscape of glioblastoma, Cell 155 (2013) 462–477, https://doi.org/10.1016/j.cell.2013.09.034.
[6] C. Marques, T. Unterkircher, P. Kroon, et al., NF1 regulates mesenchymal glioblastoma plasticity and aggressiveness through the AP-1 transcription factor FOSL1, Elife 10 (2021) e64846, https://doi.org/10.7554/eLife.64846.
[7] M.J. Schmitt, C. Company, Y. Dramaretska, et al., Phenotypic mapping of pathologic cross-talk between glioblastoma and innate immune cells by synthetic genetic tracing, Cancer Discov. 11 (2021) 754–777, https://doi.org/10.1158/2159-8290.CD-20-0219.
[8] K. Funa, M. Sasahara, The roles of PDGF in development and during neurogenesis in the normal and diseased nervous system, J. Neuroimmune Pharmacol. 9 (2014) 168–181, https://doi.org/10.1007/s11481-013-9479-z.
[9] A. Karki, N.E. Berlow, J-A Kim, et al., Receptor-driven invasion profiles in diffuse intrinsic pontine glioma, Neurooncol. Adv. 3 (2021), https://doi.org/10.1093/noajnl/vdab039.
[10] N. Goffart, J. Kroonen, E. Di Valentin, et al., Adult mouse subventricular zones stimulate glioblastoma stem cells specific invasion through CXCL12/CXCR4 signaling, Neuro Oncol. 17 (2015) 81–94, https://doi.org/10.1093/neuonc/nou144.
[11] A. do Carmo, I. Patricio, M.T. Cruz, et al., CXCL12/CXCR4 promotes motility and proliferation of glioma cells, Cancer Biol. Ther. 9 (2010) 56–65, https://doi.org/10.4161/cbt.9.1.10342.
[12] A.B. Khan, S. Lee, A.S. Harmanci, et al., CXCR4 expression is associated with proneural-to-mesenchymal transition in glioblastoma, Int. J. Cancer 152 (2023) 713–724, https://doi.org/10.1002/ijc.34329.
[13] R.J. Youle, A. Strasser, The BCL-2 protein family: opposing activities that mediate cell death, Nat. Rev. Mol. Cell Biol. 9 (2008) 47–59, https://doi.org/10.1038/nrm2308.
[14] K. Zhang, L. Zhu, M. Fan, Oxygen, a key factor regulating cell behavior during neurogenesis and cerebral diseases, Front. Mol. Neurosci. 4 (2011) 5, https://doi.org/10.3389/fnmol.2011.00005.
[15] S.R. McKeown, Defining normoxia, physoxia and hypoxia in tumours-implications for treatment response, Br. J. Radiol. 87 (2014) 20130676, https://doi.org/10.1259/bjr.20130676.
[16] Beppu T, Kamada K, Yoshida Y, et al. Change of oxygen pressure in glioblastoma tissue under various conditions n.d.
[17] A. Chakhoyan, J-S Guillamo, S. Collet, et al., FMISO-PET-derived brain oxygen tension maps: application to glioblastoma and less aggressive gliomas, Sci. Rep. 7 (2017) 10210, https://doi.org/10.1038/s41598-017-08646-y.
[18] M. Calvo Tardón, E. Marinari, D. Migliorini, et al., An experimentally defined hypoxia gene signature in glioblastoma and its modulation by metformin, Biology 9 (2020) 264, https://doi.org/10.3390/biology9090264.
[19] M. Di Giovannantonio, F. Hartley, B. Elshenawy, et al., Defining hypoxia in cancer: a landmark evaluation of hypoxia gene expression signatures, Cell Genom. 5 (2025) 100764, https://doi.org/10.1016/j.xgen.2025.100764.
[20] R.L. Hancock, K. Dunne, L.J. Walport, et al., Epigenetic regulation by histone demethylases in hypoxia, Epigenomics 7 (2015) 791–811, https://doi.org/10.2217/epi.15.24.
[21] B. Thienpont, J. Steinbacher, H. Zhao, et al., Tumour hypoxia causes DNA hypermethylation by reducing TET activity, Nature 537 (2016) 63–68, https://doi.org/10.1038/nature19081.
[22] M. Batie, J. Frost, M. Frost, et al., Hypoxia induces rapid changes to histone methylation and reprograms chromatin, Science 363 (2019) 1222–1226, https://doi.org/10.1126/science.aau5870.
[23] R. Matuleviciute, P.P. Cunha, R.S. Johnson, et al., Oxygen regulation of TET enzymes, FEBS J. 288 (2021) 7143–7161, https://doi.org/10.1111/febs.15695.