*Article*

# Multi-omic meta-analysis of transcriptomes and the bibliome uncovers *GPR146* as the novel hypoxia-inducible gene

**Yoko Ono [1] and Hidemasa Bono [1,*]**

1 Program of Biomedical Science, Graduate School of Integrated Sciences for Life, Hiroshima University, 3-10-23 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-0046, Japan
* Correspondence: bonohu@hiroshima-u.ac.jp; Tel.: +81-82-424-4013

**Abstract:** Hypoxia is a condition in which cells, tissues, or organisms are deprived of sufficient oxygen supply. Aerobic organisms have the hypoxic response system, represented by hypoxia-inducible factor 1-$\alpha$ (HIF1A), to avoid this condition. Because of publication bias, the genes other than well-known signature hypoxia-stimulating genes may not be focused. Therefore, in this study, we performed meta-analysis to identify novel hypoxia-responsive genes. We searched publicly available transcriptome databases to obtain hypoxia-related experimental data, retrieved the metadata, and manually curated it. We selected the genes that are differentially expressed by hypoxic stimulation, and evaluated their relevance in hypoxia by performing enrichment analyses. We then calculated the number of reports and similarity coefficient of each gene for HIF1A (surrogate for hypoxia) using gene2pubmed that provides information about the relationship between gene and publication. Enrichment analysis confirmed that the selected genes were related to hypoxic stimulation, as expected. We then performed gene2pubmed analysis to investigate the genes that had not been fully studied under hypoxia. In this data-driven study, we report that the G protein-coupled receptor 146 gene is upregulated by hypoxic stimulation.

**Keywords:** hypoxia; RNA-seq; ChIP-seq; gene2pubmed; meta-analysis; signature genes; *GPR146*; enrichment analysis

**Copyright:** © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

With the development of genetic engineering techniques such as genome editing, their use by molecular biologists to test various hypotheses has increased. This has made possible to test hypotheses that are constructed on the basis of the knowledge of each researcher, including literature information, by performing biological experiments. Hypothesis construction is triggered by literature surveys and the collective knowledge of researchers, and is influenced by publication bias. For example, when the frequency of about 600,000 publications annotated with about 20,000 human coding genes was calculated, more than 9,000 publications reported the p53 gene, whereas over 600 genes were not mentioned at all [1]. In contrast, the development of microarray and high-throughput sequencing technologies has enabled the comprehensive acquisition of thousands of gene expression profiles at a time. Published transcriptome data are archived in public databases such as Gene Expression Omnibus (GEO) of the U.S. National Center for Biotechnology Information (NCBI) [2], ArrayExpress (AE) of the European Bioinformatics Institute [3], and Genomic Expression Archive (GEA) of the DNA Data Bank of Japan [4]. At present, although approximately 60,000 expression data series of human genes have been registered, we estimate that they are used in a few discoveries. We believe that meta-analysis of these transcriptome data may provide additional insights into biology.

Among the various fields of biological research, more the well-studied field, greater is the impact of publication bias. We have considered areas where meta-analysis based on

gene expression data can uncover new findings buried in publication bias. The analyses were performed by considering the following conditions: (1) hypoxia-inducible factor 1-$\alpha$ (HIF1A) is a representative transcription factor involved in hypoxic stimulus-response, and (2) it has been categorized in Gene Ontology, which is a field where gene expression data can be easily obtained. In our previous study, we studied hypoxic stimulus-responses using public databases [5]. The present study is the continuation of our previous study to further build on the knowledge of hypoxic stimulus-response.

Oxygen is essential for respiration in aerobic organisms. It is used to produce ATP by oxidative phosphorylation and electron transfer systems. Hypoxia is a condition in which cells, tissues, or organisms are deprived of sufficient oxygen supply as the amount of oxygen in the blood decreases. The tissues and cells of higher organisms including humans have a hypoxic response system, regulated by HIF1A to avoid this oxygen-deprive condition caused by hypoxic stimuli. Patients suffering from coronavirus disease 2019 (COVID-19) may also suffer through the condition called happy hypoxia [6]. Patients with happy hypoxic become severely hypoxic, and do not recognize their own oxygen depletion, and this condition becomes life-threatening accompanied with increased respiratory rate. However, the detailed mechanism underlying happy hypoxia remains unclear.

Therefore, we focused on hypoxic stimulus-response and aimed to identify novel genes by performing meta-analysis using public databases. We selected hypoxia-responsive genes based on the gene expression ratio of approximately 500 pairs of hypoxia-stimulated samples and evaluated the similarity coefficient of each gene for HIF1A, a representative factor in hypoxia research, using gene2pubmed that shows the relationship between gene and publication.

## 2. Materials and Methods

### *2.1. Curation of Public Gene Expression Data*

To check the accession of hypoxia-related gene expression data from public databases, we used a graphical web tool called All Of gene Expression (AOE) [7,8]. AOE integrates metadata from not only GEO, AE, and GEA but also RNA sequencing (RNA-seq) data that are exclusively archived in Sequence Read Archive (SRA) [9]. We searched for a list of experimental data series related to hypoxia from GEO using the following search formulas: "hypoxia"[MeSH Terms] OR "hypoxia"[All Fields] AND "Homo sapiens"[porgn] AND "gse"[Filter]. We downloaded this list on August 17, 2020. We obtained the metadata from the Series Matrix files in GEO, from articles, or downloaded the metadata from SRA using the Python package pysradb (v 0.11.1). We then curated the data into the comparable sample pairs of hypoxia and normoxia (HN-pairs) [10]. We collected the SRA format data from the NCBI using prefetch (version 2.9.6).

### *2.2. Gene Expression Quantification*

Since the downloaded data were in the SRA format, we used the fasterq-dump program in SRA Toolkit [11] to convert the data into the FASTQ formatted files for expression quantification. We then quantified single-end and paired-end RNA-seq reads using ikra (version 1.2.3) [12], an RNA-seq pipeline with default parameters. Ikra automates the RNA-seq data analysis process, which includes the quality control of reads (Trim Galore version 0.6.3) and transcript quantification (Salmon version 0.14.0 [13] with reference transcript sets in GENCODE release 30). We chose this workflow because it is suitable for quantifying gene expression from heterogeneous RNA-seq data obtained from various instruments and laboratories. In this study, the data acquisition and quality control processes required approximately 1 month to obtain the current dataset. Quantitative RNA-seq data are accessible at figshare [14].

### *2.3. Calculation of HN-ratio and HN-score*

Based on the transcriptome data paired with hypoxia and normoxia, we calculated the ratio of expression value for each gene (termed as HN-ratio) [15]. HN-ratio $R$ is calculated using the following equation:

$$R = \frac{T_{hypoxia} + 1}{T_{normoxia} + 1} \quad (1)$$

where $T_{hypoxia}$ is the gene expression value called scaled transcripts per million (TPM) [16] under hypoxic conditions and $T_{normoxia}$ is TPM under normoxic conditions, paired with hypoxic conditions. To reduce the effect of small variations in extremely low gene expression values, we calculated HN-ratio by adding 1 to the denominator and numerator.

We then classified each gene into three groups based on its HN-ratio. When the HN-ratio was greater than the threshold, the gene was considered upregulated, and when the ratio was less than the inverse of the threshold, the gene was considered downregulated. If a gene is neither upregulated nor downregulated, it was classified as unchanged. For the classification of upregulated and downregulated genes, we tested 1.5-and 2-fold thresholds, of which we then considered the 1.5-fold threshold to classify upregulated and downregulated genes. For evaluating hypoxia-inducible genes, we calculated HN-score for each coding gene in humans. HN-score was calculated by subtracting the number of samples with downregulated gene from the number of samples with upregulated gene.

### 2.4. *Enrichment Analysis*

We used Metascape [17,18] and ChIP-Atlas [19,20] for Gene Set Enrichment Analysis. ChIP-Atlas is a comprehensive and integrated database for visualization and utilization of publicly available chromatin immunoprecipitation sequencing (ChIP-seq) data. In this study, we performed conventional 'express analysis' using Metascape. In the ChIP-Atlas settings, we set the "Select dataset to be compared" item to "Refseq coding genes (excluding user data)" and set other items to default.

### 2.5. *Meta-Analysis of ChIP-Seq Data*

Public ChIP-seq data were collected, curated, and pre-calculated for reuse in the ChIP-Atlas database [19,20]. We used the "Target Genes" function of ChIP-Atlas to obtain average Model-based Analysis of ChIP-seq (MACS2) scores for three antigens associated with hypoxia at a distance of ±5k from transcription start site. These antigens were HIF1A, endothelial PAS domain-containing protein 1 (EPAS1, also known as hypoxia-inducible factor-2$\alpha$), and aryl hydrocarbon receptor nuclear translocator (ARNT). We combined the ChIP-seq data with the gene expression data described above, using the names of genes to combine the two datasets.

### 2.6. *Calculation of the Number of Publications for Each Gene and Similarity Coefficient for HIF1A*

We calculated the number of publications and Simpson similarity coefficient per human gene using Python (version 3.8.6) [21]. The Simpson similarity coefficient $S$ was calculated using the following equation:

$$S(X,Y) = \frac{|X \cap Y|}{min(|X|,|Y|)} \quad (2)$$

where $|X \cap Y|$ is the number of PubMed IDs that overlaps between HIF1A and a gene in gene2pubmed, whereas $min(|X|,|Y|)$ is the number of PubMed IDs of HIF1A or the gene in gene2pubmed, whichever is less. Gene2pubmed [22] and Gene Info [22], which were required for the above calculations, were downloaded on January 4, 2021.

### 2.7. *Visualization and Integrated Functional Analysis of Genes*

For generating scatter plots and box plots, we used TIBCO Spotfire Desktop version 11.0.0 (TIBCO Spotfire, Inc., Palo Alto, CA, USA).

## 3. Results

### 3.1. Overview

This study was performed in three main steps (Fig. 1). In step 1, we obtained the RNA-seq data related to hypoxic stimulation from public gene expression database. After the manual curation of metadata, UP 100 and DOWN 100 gene lists related to hypoxia stimulation were obtained through the quantification of gene expression in the corresponding datasets. The UP 100 and DOWN 100 gene lists included top 100 and bottom 100 genes in HN-score, respectively. In step 2, we verified hypoxic stimulus-response-related genes in UP 100 and DOWN 100 gene lists. Finally, in step 3, we performed two types of analyses for our new findings. In the first, we studied the relationship between hypoxic response and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection response by performing enrichment analysis in COVID. In the other, we investigated novel hypoxia-responsive genes using gene2pubmed.

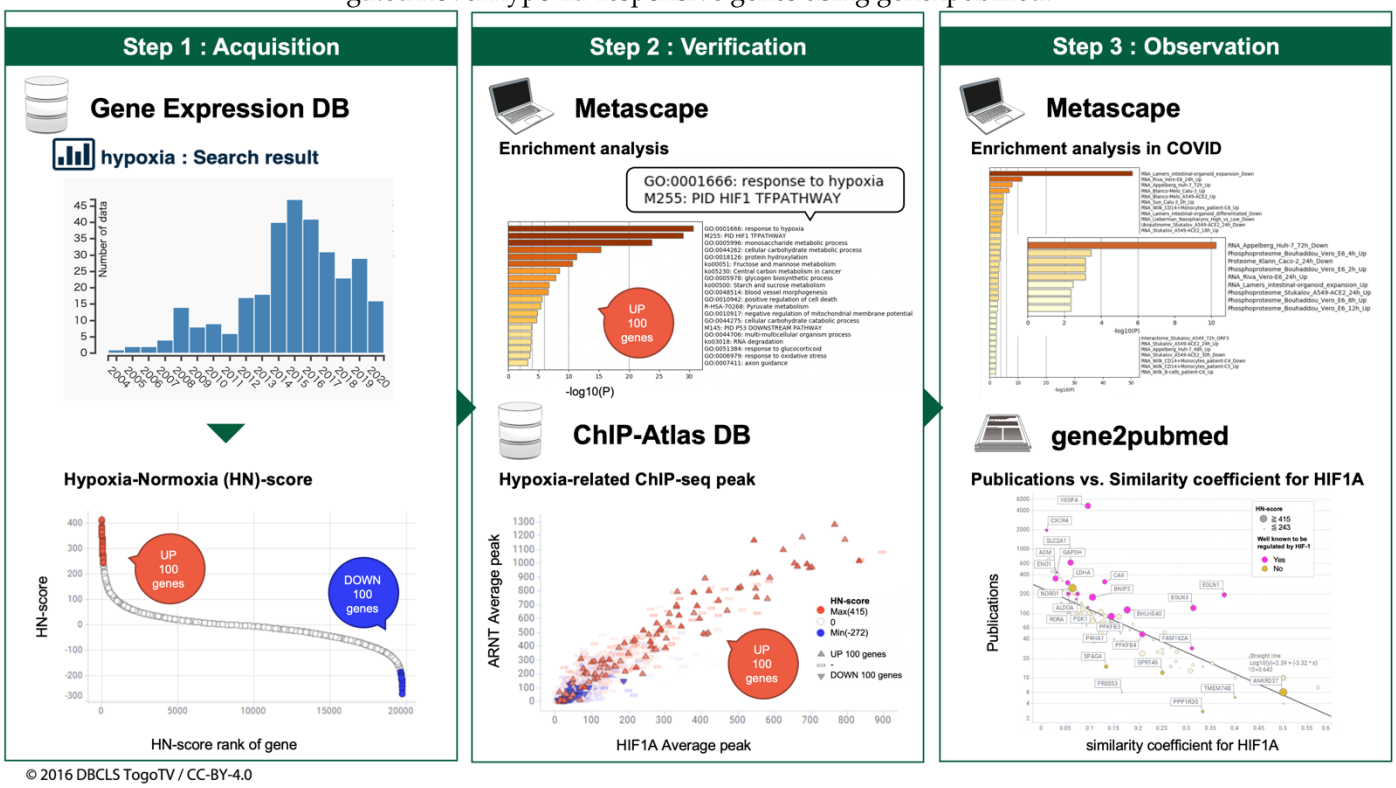



**Figure 1.** Schematic view of hypoxic transcriptome meta-analysis. Step 1. Evaluation and listing of upregulation and downregulation of hypoxia-responsive genes. Step 2. Confirmation of known hypoxic stimulation-related genes. Step 3. Discovery of novel genes related to hypoxic stimulus−response.

### 3.2. Curation of Hypoxic Transcriptome Data in Public Databases

Initially, we used AOE to check whether a large number of hypoxia-related data were registered. We then curated the metadata provided by the NCBI for further analysis. The criteria for curation were that the specimens with the experimental conditions of hypoxia and normoxia should be in the same data series so as to set HN-pair and to derive RNA-seq reads from human cell lines or tissue specimens. Next, we obtained the SRA IDs for 69 data series and 495 HN-pairs of GEO-drawn samples after integration with hypoxia-related data lists provided in previous reports and elimination of duplicates [10]. In 495 HN-pair data, the hypoxic conditions were ranging between 0.1% and 5% $O_2$ concentration, including some chemical hypoxic conditions where $CoCl_2$ was used to induce a

hypoxia-related state under normoxic conditions. The treatment time ranged from 1 hour to maximum 3 months. The most common hypoxic condition among the data was 1% $O_2$ (266 HN-pairs, 53.7%) and 24 hours of treatment (234 HN-pairs, 47.3%). The most common cell type was cancer (324 HN-pairs, 65.5%), and the most common tissue among them was breast cancer (112 HN-pairs, 22.6%).

### *3.3. Meta-Analysis of Hypoxia-Responsive Genes*

After the quantification of gene expression data obtained from RNA-seq, the number of conditions under which each gene was upregulated, downregulated, and remained unchanged was estimated on the basis of HN-ratio. The complete lists of meta-analyzed results are accessible from figshare [23]. We calculated the HN-score of each gene and quantified by RNA-seq to evaluate responsiveness of each gene toward hypoxia. For example, in the case of vascular endothelial growth factor A *(VEGFA)*, the counts showed 406 UP, 25 DOWN, and 64 unchanged, and therefore, its HN-score was 381. Full lists of counts (upregulated/downregulated/unchanged) with HN-scores for all genes are accessible from figshare [23]. We selected the top 100 [24] and bottom 100 [25] HN-scored genes and used these genes as the UP 100 and DOWN 100 gene lists, respectively, for further analysis.

### *3.4. Evaluation of Hypoxic Stimuli Response Genes*

To confirm whether the UP 100 and the DOWN 100 gene lists were affected by hypoxia-related regulators such as HIF1A, we performed enrichment analysis using ChIP-Atlas. Enrichment analysis showed that the UP 100 gene list was related to hypoxia-related antigens such as HIF1A, ARNT, and EPAS1, whereas the DOWN 100 gene list was related to epigenetic regulators such as SAP30 and HDAC1, and MYC, an oncogene involved in cell cycle progression (Table 1). Similarly, enrichment analysis by Metascape showed that hypoxia-related gene sets and genes related to the metabolic processes of non-coding RNAs (ncRNAs) were listed in the UP 100 gene list and DOWN 100 gene list, respectively (Fig. 2 A, B). The scatter plots of the mean MACS2 peaks of HIF1A, EPAS1, and ARNT, which were the top-ranked genes in the enrichment analysis of UP 100 gene list, were plotted. The genes having higher MACS2 peaks also had higher HN-score (Fig. 2 C, D). These results indicate that the genes regulated by hypoxia-related factors are listed in the UP 100 gene list, and genes regulated by SAP, MYC, and HDAC1 are listed in the DOWN 100 gene list.

**Table 1.** Enrichment analysis of ChIP-Atlas

| Input list | Antigen | ID | Log P-val | Log Q-val | Fold Enrichment |
|---|---|---|---|---|---|
| UP 100 gene list | HIF1A | SRX4802348 | -88.8246 | -84.064 | 35.6249 |
| | ARNT | SRX4802353 | -76.4303 | -72.3686 | 83.3136 |
| | EPAS1 | SRX3051209 | -73.1987 | -69.2831 | 34.9928 |
| DOWN 100 gene list | SAP30 | SRX116447 | -34.1844 | -30.0149 | 4.96916 |
| | MYC | SRX1497384 | -31.4158 | -27.5474 | 2.97453 |
| | HDAC1 | SRX186644 | -27.4205 | -24.1541 | 3.3231 |

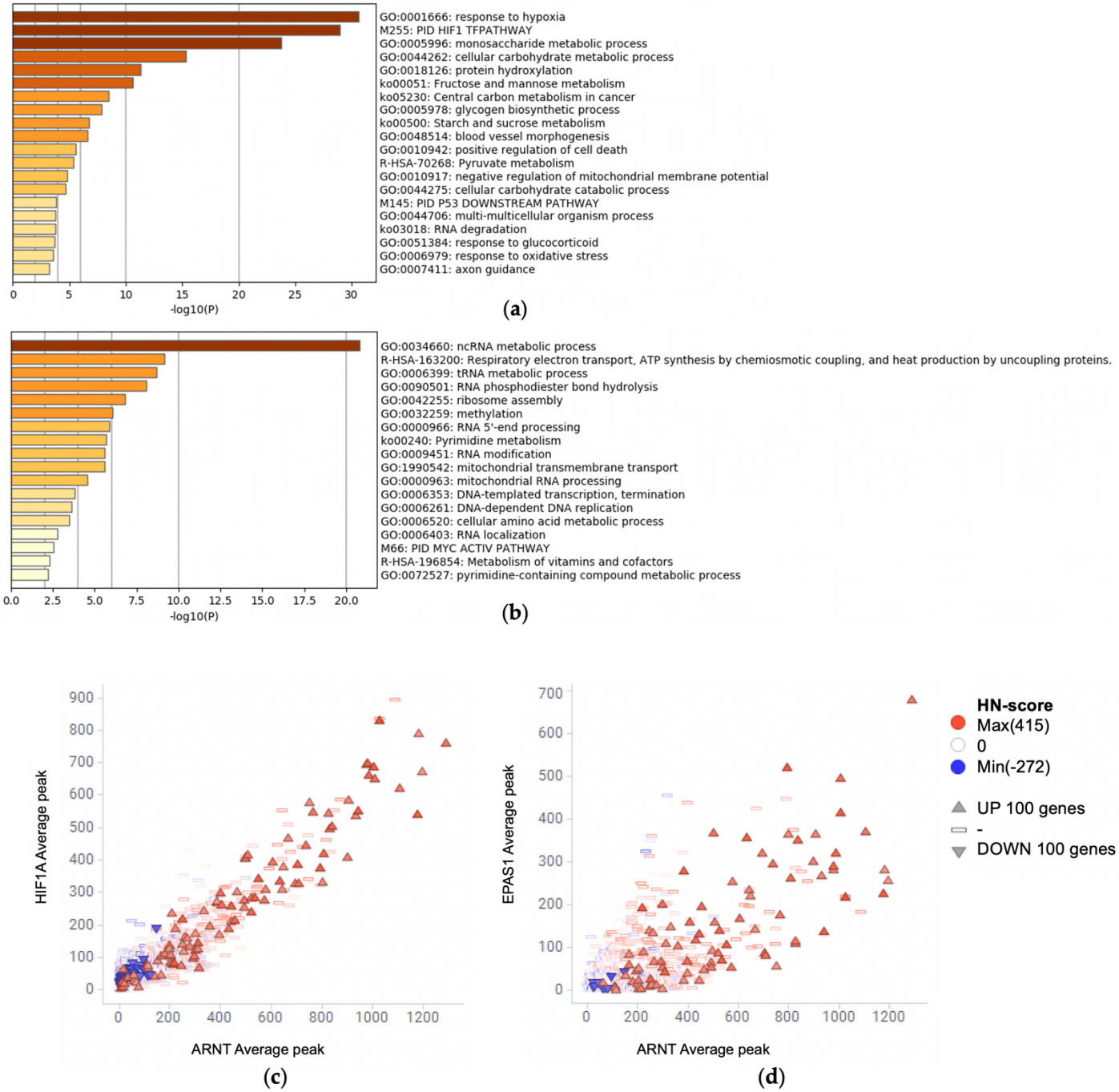



**Figure 2.** (a, b) Confirmation of known hypoxic stimulation-related genes. Enrichment analysis for (a) the UP 100 gene list and (b) the DOWN 100 gene list, (c, d) Scatter plot of ChIP-seq average peaks of hypoxic related antigens, (c) HIF1A vs. ARNT and (d) EPAS1 vs. ARNT colored by HN-score.

### *3.5. Enrichment Analysis in COVID using Metascape*

To investigate the relationship between COVID-19 and hypoxic response, we performed enrichment analysis in COVID for UP and DOWN 100 gene lists. In the UP 100 gene list, COVID027: RNA_Lamers_intestinal-organoid_expansion_Down was highly ranked (Fig. 3A). On the other hand, the DOWN 100 gene list suggested the association of COVID005: RNA_Appelberg_Huh-7_72h_Down (Fig. 3B). The results of COVID analysis in UP 100 gene list showed that these gene expression changes in the intestinal organoid infected with SARS-CoV-2 was less than that in the uninfected group.

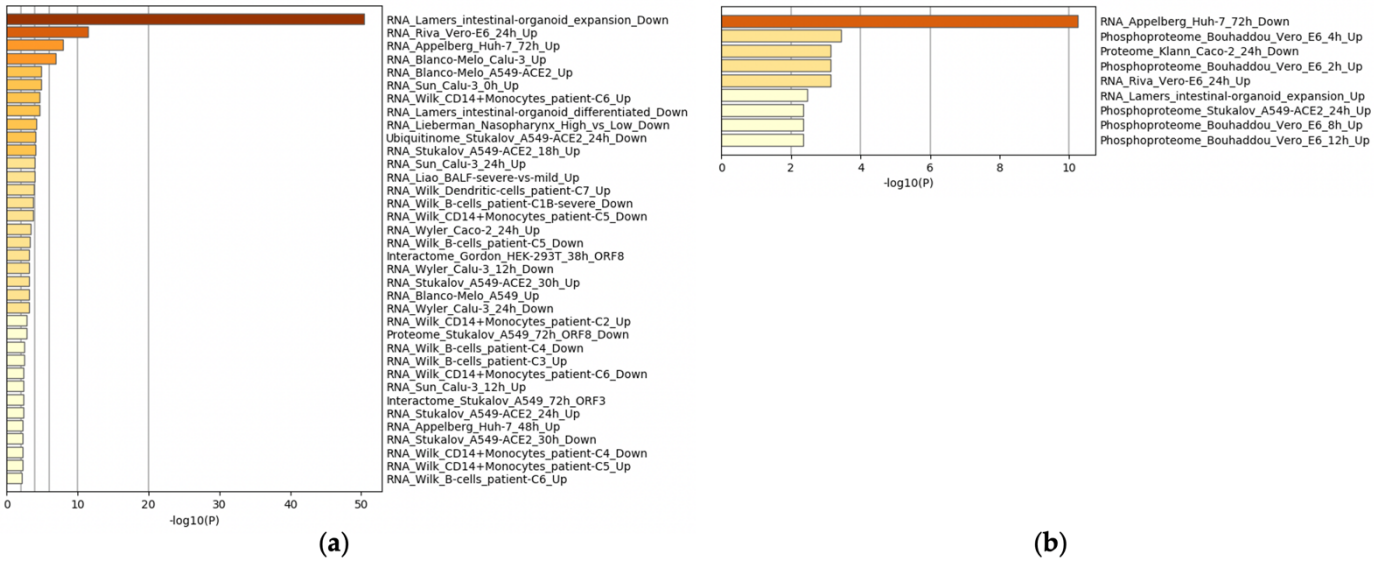


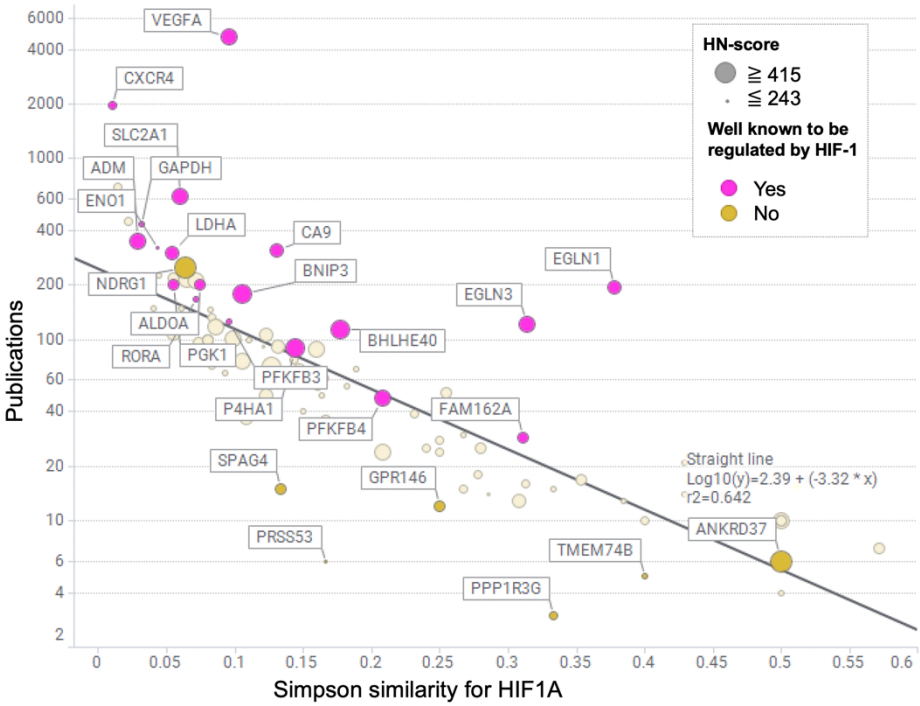

**Figure 3.** Enrichment analysis in COVID for (a) UP 100 gene list and (b) DOWN 100 gene list.

### *3.6. Evaluation of Simpson Similarity between HIF1A and Genes*

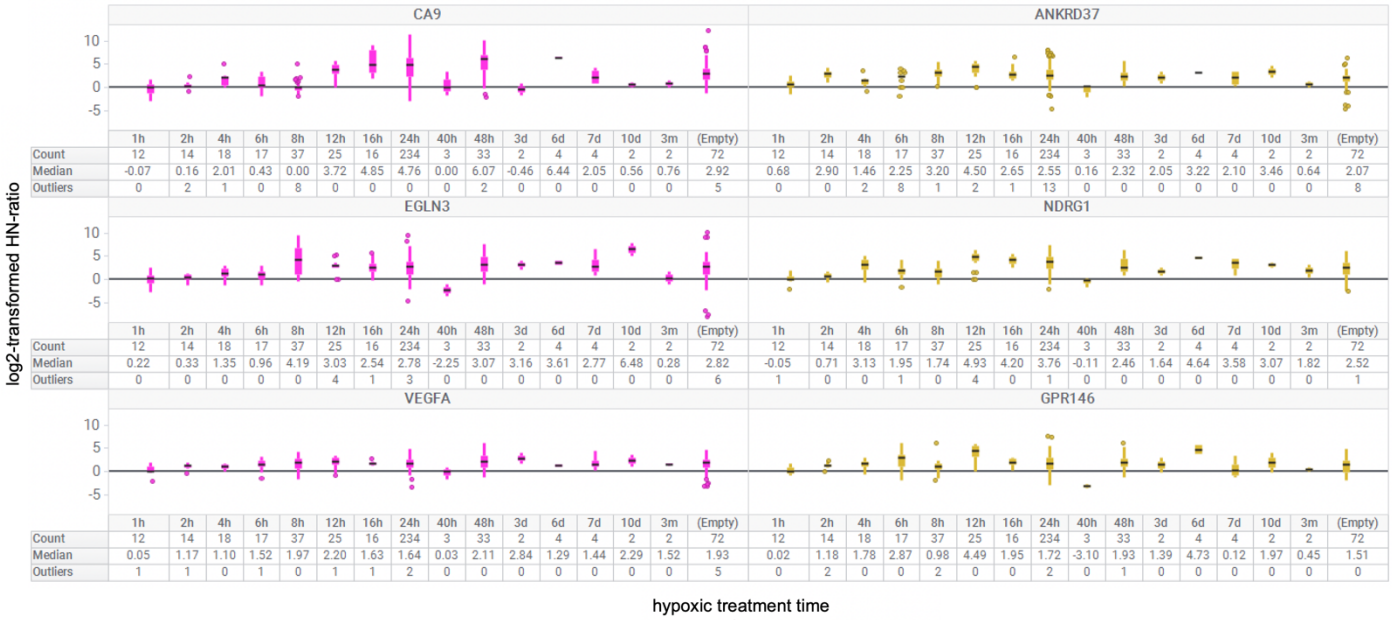
We performed bibliometric analysis to determine whether the genes of UP 100 gene list had been previously studied in relation to hypoxia. We calculated the number of publications for each gene in the UP 100 gene list [26]. We also calculated the Simpson similarity coefficient between these genes and HIF1A [27]. The Simpson similarity coefficient was determined to evaluate the extent of hypoxia-related research. The number of publications and Simpson similarity coefficients were visualized using the scatter plot (Fig 4A). In addition, we marked the genes that are directly regulated by HIF-1 as of 2006 [28] in this scatter plot. Interestingly, hypoxia-related genes were plotted on the upper side of the regression line.

### *3.7. Box Plot of HN-ratio by Treatment Time*

To check the individual HN-ratio of genes with high HN-score based on the treatment time, we visualized base-2 logarithm (log2) transformed HN-ratio using the box plot (Fig 4B). Genes in the box plot were selected from the HIF-1-regulated genes and the genes with high HN-scores in this metadata analysis. The former included carbonic anhydrase 9 *(CA9)*, egl-9 family hypoxia inducible factor 3 *(EGLN3)*, and *VEGFA*, whereas the latter included ankyrin repeat domain 37 *(ANKRD37)*, N-myc downstream regulated 1 *(NDRG1)*, and G protein-coupled receptor 146 *(GPR146)*. These six genes visualized in the box plot showed positive HN-ratio values for most of the treatment time.




(a)




(b)

**Figure 4.** Discovery of novel genes associated with hypoxic stimulus-response. (a) Scatter plot of the number of publications vs. Simpson similarity coefficients for HIF1A in the UP 100 gene list; (b) Box plot of log2-transformed HN-ratio per hypoxic treatment time for some of the UP 100 genes.

## 4. Discussion

In this study, we identified hypoxia-responsive gene by performing metadata analysis. We believe that each gene was evaluated without affecting by publication bias. Our study was feasible because of the data availability from public databases. This new finding was obtained because of the access and proper maintenance of public databases.

We analyzed 69 data series related to hypoxia, selected from the entire gene expression data registered in GEO, and obtained 495 HN-pairs, which are about four times of previously reported data [5]. In this study, 65.5% of the gene expression data were derived

from cancer experiments, reflects the demand for these data in hypoxia-related cancer research and ease of handling.

First, the Metascape and ChIP-Atlas enrichment analyses of UP 100 gene list showed their association with hypoxia-related gene sets as expected, and hence, we concluded that HN-score helps select the hypoxia-responsive genes (Table 1, Fig. 2). Thus, we decided to proceed with the analysis of UP 100 gene list in subsequent analyses. On the other hand, the analysis of DOWN 100 gene list showed the association with MYC, which is repressed by HIF. The regulation of MYC by HIF is mediated by the antagonistic inhibition of MYC protein complex formation [29], direct inhibition of MYC transcriptional activity by competition for DNA binding sites [30], and promotion of MYC proteasomal degradation [31]. We believe that these indirect effects of HIF have been reflected in this result. Metascape enrichment analysis of DOWN 100 gene list suggested an association with the metabolic processes of ncRNA (Fig. 2B). This suggests that hypoxic stimulation may also affect the expression of ncRNAs, and may be related to the effect of MYC. In this study, only the protein-coding genes of GENCODE were quantitatively analyzed, and hence, ncRNAs were excluded from this analysis. Further studies are needed to identify ncRNAs involved in hypoxic stimulation.

Second, we performed enrichment analysis in COVID for UP 100 gene list, which showed a relationship to the gene set COVID027: RNA_Lamers_intestinal-organoid_expansion_Down (Fig. 3). We obtained this metadata by searching "COVID027" in Coronascape [32,33]. This dataset was the comparison of intestinal-organoid between SARS-CoV-2-infected and -uninfected samples [34]. This result suggests that the intestinal organoid without SARS-CoV-2 infection showed a hypoxic stimulus response, whereas infected organoids showed little response. *In vitro* experiments showed that Akt/mTOR/HIF-1 signaling is dysregulated in SARS-CoV-2-infected cells [35], and the present report may corroborate this result.

Gene2pubmed has been used to determine the order of genes retrieved by RefEx [36]. In this study, we used gene2pubmed to evaluate the genes that are not well-known as hypoxia-related genes. We used gene2pubmed to calculate two variables: the number of publications per gene and Simpson similarity coefficient for HIF1A. The Simpson similarity coefficient measures the strength of the co-occurrence of PubMed IDs associated with HIF1A or with a gene. If the Simpson similarity coefficient is close to 1, the relationship between PubMed IDs and HIF1A is considered to be strong; however, the less number of PubMed IDs on one side being compared indicates poor relationship. For this reason, we visualized the number of publications and Simpson similarity coefficient for the genes in the UP 100 gene list in the scatter plot. In addition, we also plotted HIF-1-regulated genes [28] in this plot (Fig. 4A). The Simpson similarity coefficient was lower in case of more publications, and might be because of the fact that more relationships were revealed, and thus the Simpson similarity coefficient is likely to be underestimated. The genes previously known to be regulated by HIF-1 scored high in the Simpson similarity coefficient and plotted on the upper side of regression line, despite a large number of publications. The genes with a small number of publications, such as *ANKRD37,* which was reported to be regulated by HIF1A [37], were also plotted above the regression line. Therefore, we hypothesized that the genes plotted below the regression line, such as sperm associated antigen 4 *(SPAG4),* GPR146, protein phosphatase 1 regulatory subunit 3G, transmembrane protein 74B, and serine protease 53 might be considered as candidates for novel hypoxia-responsive genes. Although some reports state that *SPAG4* is related to hypoxia [38,39], no reports are available in PubMed search related to hypoxia-related genes other than *SPAG4*. In the similarity coefficient analysis using gene2pubmed, three reports on the co-occurrence of PubMed IDs between HIF1A and GPR146 are available [40–42]. These were overall genetic analyses and did not focus on hypoxia. We attempted to visualize HN-ratio by focusing on *GPR146* having high HN-score. The log2-transformed HN ratio was visualized as a box plot (Fig. 4B). We focused on *CA9, EGLN3*, and *VEGFA*, which are directly regulated by HIF-1, and *ANKRD37, NDRG1*, and *GPR146,* which were among the

genes with high HN-scores in our meta-analysis. *ANKRD37* and *NDRG1* have also been reported to have high HN-scores [5]. These genes had positive HN-ratio values, as well as genes with known hypoxic responses.

We confirmed that *GPR146* was upregulated, which was plotted below the regression line in the scatter plot (Figure 4A), similar to other known hypoxia-responsive genes. *GPR146* was not reported to be associated with hypoxia in PubMed search; thus, we believe that *GPR146* is an important hypoxia-responsive gene that was not focused so far. In hypoxia studies using comprehensive gene expression data [43,44], *GPR146* is included in the list of genes upregulated by hypoxic stimulation. However, according to previous reports, *GPR146* is just a gene among many hypoxia-related genes. The inhibition of GPR146 is involved in the lowering of cholesterol level [45]. C-peptide, a putative ligand of GPR146, inhibited low $O_2$-induced ATP release in erythrocytes [46]. The biological function of GPR146 is expected to be further elucidated in future studies.

We focused on the hypoxic response and visualized the number of research reports for each gene using gene2pubmed data. We identified the novel gene by using public data containing thousands of gene expression information, which can be obtained regardless of the researcher's interest. We plan to continue using public databases to discover similar new findings. We would investigate the gene expression of ncRNAs during hypoxic condition, based on our results of the effect of hypoxia on metabolic process-related genes of ncRNA (Fig. 2 B).

## 5. Conclusions

We found that *GPR146* is upregulated in hypoxic stimulation by performing data-driven analyses such as metadata analysis based on public databases.

**Author Contributions:** Conceptualization, Y.O., H.B.; methodology, Y.O.; software, Y.O.; validation, Y.O., H.B.; formal analysis, Y.O., H.B.; investigation, Y.O.; resources, Y.O., H.B.; data curation, Y.O.; writing—original draft preparation, Y.O.; writing—review and editing, H.B.; visualization, Y.O.; supervision, H.B.; project administration, H.B.; funding acquisition, H.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the center of innovation for Bio-Digital Transformation (BioDX), program on open innovation platform for industry-academia co-creation (COI-NEXT), Japan Science and Technology Agency (JST, COI-NEXT, JPMJPF2010). This research was also supported by ROIS-DS-JOINT (010RP2020) and JSPS KAKENHI Grant Number 20H03776.

**Data Availability Statement:** The data presented in this study are openly available in figshare [47]. Source codes to replicate the study are also freely available at GitHub [21].

**Acknowledgments:** Authors would like to thank Late Mr. Yusuke Yamada for helpful discussions in the initial stage of this study.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Carter, A.J.; Kraemer, O.; Zwick, M.; Mueller-Fahrnow, A.; Arrowsmith, C.H.; Edwards, A.M. Target 2035: Probing the Human Proteome. *Drug Discovery Today* **2019**, *24*, 2111–2115, doi:10.1016/j.drudis.2019.06.020.
2. Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M.; et al. NCBI GEO: Archive for Functional Genomics Data Sets—Update. *Nucleic Acids Research* **2013**, *41*, D991–D995, doi:10.1093/nar/gks1193.
3. Athar, A.; Füllgrabe, A.; George, N.; Iqbal, H.; Huerta, L.; Ali, A.; Snow, C.; Fonseca, N.A.; Petryszak, R.; Papatheodorou, I.; et al. ArrayExpress Update – from Bulk to Single-Cell Expression Data. *Nucleic Acids Research* **2019**, *47*, D711–D715, doi:10.1093/nar/gky964.

4. Kodama, Y.; Mashima, J.; Kosuge, T.; Ogasawara, O. DDBJ Update: The Genomic Expression Archive (GEA) for Functional Genomics Data. *Nucleic Acids Research* **2019**, *47*, D69–D73, doi:10.1093/nar/gky1002.
5. Bono, H.; Hirota, K. Meta-Analysis of Hypoxic Transcriptomes from Public Databases. *Biomedicines* **2020**, *8*, 10, doi:10.3390/biomedicines8010010.
6. Couzin-Frankel, J. The Mystery of the Pandemic′s ′Happy Hypoxia.′ *Science* **2020**, *368*, 455–456, doi:10.1126/science.368.6490.455.
7. AOE Available online: https://aoe.dbcls.jp/ (accessed on 27 February 2021).
8. Bono, H. All of Gene Expression (AOE): An Integrated Index for Public Gene Expression Databases. *PLOS ONE* **2020**, *15*, e0227076, doi:10.1371/journal.pone.0227076.
9. Kodama, Y.; Shumway, M.; Leinonen, R.; on behalf of the International Nucleotide Sequence Database Collaboration The Sequence Read Archive: Explosive Growth of Sequencing Data. *Nucleic Acids Research* **2012**, *40*, D54–D56, doi:10.1093/nar/gkr854.
10. Ono, Y.; Information on Hypoxic Conditions and Samples in the Dataset Used for the Meta-Analysis **2021**, doi:10.6084/M9.FIGSHARE.14141219.V1.
11. Download SRA Sequences from Entrez Search Results Available online: https://www.ncbi.nlm.nih.gov/sra/docs/sradownload/ (accessed on 27 February 2021).
12. Yyoshiaki/Ikra: RNAseq Pipeline Centered on Salmon Available online: https://github.com/yyoshiaki/ikra (accessed on 27 February 2021).
13. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression. *Nature Methods* **2017**, *14*, 417–419, doi:10.1038/nmeth.4197.
14. Ono, Y.; Quantitative Gene Expression Data of Hypoxia Experimental Data Sets **2021**, doi:10.6084/M9.FIGSHARE.14141252.V1.
15. Ono, Y.; Ratio of Gene Expression under Hypoxic and Normoxic Conditions(HN-Ratio) **2021**, doi:10.6084/M9.FIGSHARE.14141126.V1.
16. Soneson, C.; Love, M.I.; Robinson, M.D. Differential Analyses for RNA-Seq: Transcript-Level Estimates Improve Gene-Level Inferences. *F1000Res* **2016**, *4*, 1521, doi:10.12688/f1000research.7563.2.
17. Metascape Available online: https://metascape.org/gp/index.html#/main/step1 (accessed on 27 February 2021).
18. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner, C.; Chanda, S.K. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nature Communications* **2019**, *10*, 1523, doi:10.1038/s41467-019-09234-6.
19. ChIP-Atlas Available online: https://chip-atlas.org/ (accessed on 27 February 2021).
20. Oki, S.; Ohta, T.; Shioi, G.; Hatanaka, H.; Ogasawara, O.; Okuda, Y.; Kawaji, H.; Nakaki, R.; Sese, J.; Meno, C. ChIP-Atlas: A Data-Mining Suite Powered by Full Integration of Public ChIP-Seq Data. *EMBO reports* **2018**, *19*, e46255, doi:10.15252/embr.201846255.
21. No85j/Hypoxia_code Available online: https://github.com/no85j/hypoxia_code (accessed on 2 March 2021).
22. Index of /Gene/DATA Available online: https://ftp.ncbi.nlm.nih.gov/gene/DATA/ (accessed on 27 February 2021).
23. Ono, Y.; Score Based on the Ratio of Gene Expression between Hypoxic and Normoxic Conditions（HN-Score） **2021**, doi:10.6084/M9.FIGSHARE.14141135.V1.
24. Ono, Y.; Genelist_Top 100 Human Genes Up-Regulated under Hypoxic Conditions **2021**, doi:10.6084/M9.FIGSHARE.14141015.V1.
25. Ono, Y.; Genelist_Top 100 Human Genes Down-Regulated under Hypoxic Conditions **2021**, doi:10.6084/M9.FIGSHARE.14140997.V1.
26. Ono, Y.; Number of Publications per Human Gene **2021**, doi:10.6084/M9.FIGSHARE.14140850.V1.
27. Ono, Y.; Simpson Similarity of Each Human Gene to HIF1A Based on Gene2pubmed **2021**, doi:10.6084/M9.FIGSHARE.14140928.V1.
28. Hirota, K.; Semenza, G.L. Regulation of Angiogenesis by Hypoxia-Inducible Factor 1. *Critical Reviews in Oncology/Hematology* **2006**, *59*, 15–26, doi:10.1016/j.critrevonc.2005.12.003.
29. Gordan, J.D.; Bertout, J.A.; Hu, C.-J.; Diehl, J.A.; Simon, M.C. HIF-2$\alpha$ Promotes Hypoxic Cell Proliferation by Enhancing c-Myc Transcriptional Activity. *Cancer Cell* **2007**, *11*, 335–347, doi:10.1016/j.ccr.2007.02.006.

30. Koshiji, M.; Kageyama, Y.; Pete, E.A.; Horikawa, I.; Barrett, J.C.; Huang, L.E. HIF-1$\alpha$ Induces Cell Cycle Arrest by Functionally Counteracting Myc. *The EMBO Journal* **2004**, *23*, 1949–1956, doi:10.1038/sj.emboj.7600196.
31. Wong, W.J.; Qiu, B.; Nakazawa, M.S.; Qing, G.; Simon, M.C. MYC Degradation under Low O2 Tension Promotes Survival by Evading Hypoxia-Induced Cell Death. *Molecular and Cellular Biology* **2013**, *33*, 3494–3504, doi:10.1128/MCB.00853-12.
32. COVID-19 Reference Gene Lists Available online: https://metascape.org/COVID (accessed on 27 February 2021).
33. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner, C.; Chanda, S.K. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nature Communications* **2019**, *10*, 1523, doi:10.1038/s41467-019-09234-6.
34. Lamers, M.M.; Beumer, J.; Vaart, J. van der; Knoops, K.; Puschhof, J.; Breugem, T.I.; Ravelli, R.B.G.; Schayck, J.P. van; Mykytyn, A.Z.; Duimel, H.Q.; et al. SARS-CoV-2 Productively Infects Human Gut Enterocytes. *Science* **2020**, *369*, 50–54, doi:10.1126/science.abc1669.
35. Appelberg, S.; Gupta, S.; Akusjärvi, S.S.; Ambikan, A.T.; Mikaeloff, F.; Saccon, E.; Végvári, Á.; Benfeitas, R.; Sperk, M.; Ståhlberg, M.; et al. Dysregulation in Akt/MTOR/HIF-1 Signaling Identified by Proteo-Transcriptomics of SARS-CoV-2 Infected Cells. *Emerging Microbes & Infections* **2020**, *9*, 1748–1760, doi:10.1080/22221751.2020.1799723.
36. Ono, H.; Ogasawara, O.; Okubo, K.; Bono, H. RefEx, a Reference Gene Expression Dataset as a Web Tool for the Functional Analysis of Genes. *Scientific Data* **2017**, *4*, 170105, doi:10.1038/sdata.2017.105.
37. Benita, Y.; Kikuchi, H.; Smith, A.D.; Zhang, M.Q.; Chung, D.C.; Xavier, R.J. An Integrative Genomics Approach Identifies Hypoxia Inducible Factor-1 (HIF-1)-Target Genes That Form the Core Response to Hypoxia. *Nucleic Acids Research* **2009**, *37*, 4587–4602, doi:10.1093/nar/gkp425.
38. Shoji, K.; Murayama, T.; Mimura, I.; Wada, T.; Kume, H.; Goto, A.; Ohse, T.; Tanaka, T.; Inagi, R.; van der Hoorn, F.A.; et al. Sperm-Associated Antigen 4, a Novel Hypoxia-Inducible Factor 1 Target, Regulates Cytokinesis, and Its Expression Correlates with the Prognosis of Renal Cell Carcinoma. *Am J Pathol* **2013**, *182*, 2191–2203, doi:10.1016/j.ajpath.2013.02.024.
39. Knaup, K.X.; Monti, J.; Hackenbeck, T.; Jobst-Schwan, T.; Klanke, B.; Schietke, R.E.; Wacker, I.; Behrens, J.; Amann, K.; Eckardt, K.-U.; et al. Hypoxia Regulates the Sperm Associated Antigen 4 (SPAG4) via HIF, Which Is Expressed in Renal Clear Cell Carcinoma and Promotes Migration and Invasion in Vitro. *Molecular Carcinogenesis* **2014**, *53*, 970–978, doi:https://doi.org/10.1002/mc.22065.
40. Strausberg, R.L.; Feingold, E.A.; Grouse, L.H.; Derge, J.G.; Klausner, R.D.; Collins, F.S.; Wagner, L.; Shenmen, C.M.; Schuler, G.D.; Altschul, S.F.; et al. Generation and Initial Analysis of More than 15,000 Full-Length Human and Mouse CDNA Sequences. *Proc Natl Acad Sci U S A* **2002**, *99*, 16899–16903, doi:10.1073/pnas.242603899.
41. Gerhard, D.S.; Wagner, L.; Feingold, E.A.; Shenmen, C.M.; Grouse, L.H.; Schuler, G.; Klein, S.L.; Old, S.; Rasooly, R.; Good, P.; et al. The Status, Quality, and Expansion of the NIH Full-Length CDNA Project: The Mammalian Gene Collection (MGC). *Genome Res* **2004**, *14*, 2121–2127, doi:10.1101/gr.2596504.
42. Kimura, K.; Wakamatsu, A.; Suzuki, Y.; Ota, T.; Nishikawa, T.; Yamashita, R.; Yamamoto, J.; Sekine, M.; Tsuritani, K.; Wakaguri, H.; et al. Diversification of Transcriptional Modulation: Large-Scale Identification and Characterization of Putative Alternative Promoters of Human Genes. *Genome Res* **2006**, *16*, 55–65, doi:10.1101/gr.4039406.
43. Labrecque, M.P.; Takhar, M.K.; Nason, R.; Santacruz, S.; Tam, K.J.; Massah, S.; Haegert, A.; Bell, R.H.; Altamirano-Dimas, M.; Collins, C.C.; et al. The Retinoblastoma Protein Regulates Hypoxia-Inducible Genetic Programs, Tumor Cell Invasiveness and Neuroendocrine Differentiation in Prostate Cancer Cells. *Oncotarget* **2016**, *7*, 24284–24302, doi:10.18632/oncotarget.8301.
44. Qi, J.; Nakayama, K.; Cardiff, R.D.; Borowsky, A.D.; Kaul, K.; Williams, R.; Krajewski, S.; Mercola, D.; Carpenter, P.M.; Bowtell, D.; et al. Siah2-Dependent Concerted Activity of HIF and FoxA2 Regulates Formation of Neuroendocrine Phenotype and Neuroendocrine Prostate Tumors. *Cancer Cell* **2010**, *18*, 23–38, doi:10.1016/j.ccr.2010.05.024.
45. Yu, H.; Rimbert, A.; Palmer, A.E.; Toyohara, T.; Xia, Y.; Xia, F.; Ferreira, L.M.R.; Chen, Z.; Chen, T.; Loaiza, N.; et al. GPR146 Deficiency Protects Against Hypercholesterolemia and Atherosclerosis. *Cell* **2019**, *179*, 1276-1288.e14, doi:10.1016/j.cell.2019.10.034.

46. Richards, J.P.; Yosten, G.L.C.; Kolar, G.R.; Jones, C.W.; Stephenson, A.H.; Ellsworth, M.L.; Sprague, R.S. Low O2-Induced ATP Release from Erythrocytes of Humans with Type 2 Diabetes Is Restored by Physiological Ratios of C-Peptide and Insulin. *Am J Physiol Regul Integr Comp Physiol* **2014**, *307*, R862-868, doi:10.1152/ajpregu.00206.2014.
47. Ono, Y.; Multiomic Meta-Analysis of Hypoxic Stress from Public Databases. **2021**, doi:10.6084/m9.figshare.c.5323769.v1.