Exploratory Meta-Analysis of Hypoxic Transcriptomes Using Precise Transcript Reference Sequence Set

Yoko Ono[1] and Hidemasa Bono[1,2*]

[1] Laboratory of Genome Informatics, Graduate School of Integrated Sciences for Life, Hiroshima University, Japan

[2] Laboratory of Bio-DX, Genome Editing Innovation Center, Hiroshima University, Japan

[*] Correspondence: bonohu@hiroshima-u.ac.jp; Tel.: +81-82-424-4013.

Running title: Hypoxic Transcriptome Meta-Analysis

Keywords: FANTOM-CAT, hypoxia, non-coding RNA, transcript annotation

Summary blurb

Exploratory meta-analysis of hypoxic RNA-seq data using FANTOM-CAT as a reference transcriptome facilitates the evaluation of relationship between sense-antisense transcripts from the same locus.

Abstract

Gene expression studies are intrinsically biased, with many studies influenced by concomitant information such as gene-disease associations. This limitation can be overcome using a data-driven analysis approach without relying on ancillary information. The FANTOM CAGE Associated Transcriptome (FANTOM-CAT) project provides a comprehensive meta-assembly of the human transcriptome using coding and non-coding genes. Hypoxia strongly influences gene expression; additionally, non-coding RNA (ncRNA) metabolism is downregulated in response to hypoxic stimuli. We evaluated the differential response of various transcripts to hypoxia by determining their hypoxia responsiveness scores. Enrichment analysis revealed that several genes associated with ncRNA metabolism, particularly those involved in ribosomal RNA processing, were downregulated in response to hypoxia. Previously published information from the FANTOM-CAT project was suitable for meta-analysis of the transcriptome sequencing data from both coding and ncRNAs, and evaluate the hypoxia responsiveness of target transcripts and relationship between sense-antisense transcripts from the same locus. Our results may facilitate functional annotation of various transcripts including ncRNAs, allowing for both sense and antisense and coding and non-coding evaluations.

Introduction

Genomic research is inherently biased as it is heavily influenced by the known associations between specific genes and their families and the disease of interest (Mihai et al, 2021). Meanwhile,

growing evidence from global molecular biology labs has facilitated the comprehensive functional annotation of much of the genome. Therefore, further development requires to move away from those familiar genes into the unknown using a data-driven approach. Our previous study revealed a clear publication bias regarding gene expression variation in response to hypoxia and we identified a specific group of genes that had not been investigated in previous reports (Ono and Bono 2021).

Oxygen is essential for the maintenance of vital functions, and cells have strict molecular mechanisms designed to help them cope with hypoxia. The study of these responses was dramatically advanced by the discovery of hypoxia-inducible factor-1 (HIF-1) in the 1990s (Semenza and Wang 1992; Wang and Semenza 1995; Wang et al, 1995). Under normoxic conditions, HIF is hydroxylated by prolyl hydroxylase (PHD) and factor inhibiting HIF-1 (FIH-1) and then degraded by the ubiquitin-proteasome system to inhibit transcriptional activation (Jaakkola et al, 2001; Mahon et al, 2001). However, hypoxia reduces PHD and FIH-1 activity, allowing HIF to escape its hydroxylation and promoting its interactions with the aryl hydrocarbon receptor nuclear translocator (ARNT) and transcriptional co-factor, CREB, which in turn induces downstream gene expression (Ebert and Bunn 1998).

Our previous study revealed that genes related to the non-coding RNA (ncRNA) metabolic process were responsive to hypoxia (Ono and Bono 2021).

Here, we evaluated the expression of genes related to ncRNA metabolism in response to hypoxia despite a lack of functional data describing their specific effects. In addition, while there is a balance in the data available for non-coding and coding RNAs we believe that some changes in the expression of ncRNAs can only be revealed via data-driven analysis.

One tool facilitating these evaluations in the FANTOM5 project was designed to provide a reliable 5' human lncRNA data set for transcriptomic evaluations. This data was then combined with the data from the FANTOM5 Cap Analysis of Gene Expression (CAGE) project to produce the FANTOM CAGE-associated transcriptome (FANTOM-CAT) database. This dataset can be used to evaluate patterns in ncRNA expression using data from these and various coding genes. FANTOM CAGE data was used to identify approximately 28,000 lncRNA genes from the human genome with approximately 20,000 of these transcripts likely to have some specific function (Hon et al, 2017).

Here we used FANTOM-CAT as a reference and comprehensively scored the hypoxic response of various transcripts for both coding and non-coding genes. The data obtained in this study will help in the functional annotation of various transcripts, including ncRNAs, especially where there are fewer clues to their function.

RESULTS

Overview

The analyses were designed to evaluate hypoxia-responsive genes and transcripts based on two reference transcript sets (Fig. 1).

In the analyses using two reference transcripts, one describes the quantitative gene expression data annotated using GENCODE release 30 and focuses only on protein-coding genes, and the other describes the expression data for both protein-coding and non-coding transcripts using FANTOM-CAT.

Initial evaluations of the coding-only dataset revealed that hypoxic stimulation suppressed the ncRNA metabolic process, especially genes involved in rRNA processing. Given this, we went on to expand the scope of our evaluations to include both protein-coding and non-coding transcripts, and comprehensively evaluated the effects of hypoxia on both types of transcripts using a FANTOM-CAT dataset as a reference.

This dataset and its construction were described in detail in our previous study (Ono and Bono 2021). The hypoxic conditions of the analyzed data were summarized for 495 samples (Fig. 2A, B). Normoxic conditions were considered as a 20% oxygen concentration, although some samples were not mentioned, whereas hypoxic conditions ranged from 0.1 to 5% oxygen concentrations and some cases of hypoxia induced by chemicals such as $CoCl_2$. The treatment time ranged from 1 h to 3 months. The most common condition of hypoxia in the dataset was a 1% oxygen concentration for 24 h of treatment. Sixty-five percent of all samples were of cancer origin, with breast cancer as the most common tissue of origin. Gene expression in each tissue under representative conditions (cancer,

92 oxygen concentration 1%, 24 h treatment) is shown as the log2-transformed fold-change compared to
93 under normoxic condition (Fig. 2C).

94 **Down-regulation of ncRNA metabolism-related genes by the hypoxic response**

95 From a sample of hypoxia-related datasets obtained from public databases, we selected sample
96 pairs of hypoxia and normoxia pairs (HN-pairs). Based on these HN-pairs, plus 1 was calculated for
97 each HN-pair if the expression variation was greater than 1.5-fold compared to normoxic sample and
98 minus 1 if the variation was less than the reciprocal of 1.5-fold. These values were summed for each
99 gene and designated as the hypoxia-normoxia score (HN-score). The HN-score qualitatively reflects
100 hypoxia-responsiveness, unlike the method used in general gene expression analysis. This HN-score
101 was named as the HNg-score when the reference sequence was GENCODE and HNf-score when the
102 sequence was based on FANTOM-CAT. We subjected 200 genes with high HNg-scores and 200
103 genes with low HNg-scores to enrichment analysis and evaluated whether the genes selected based on
104 the HNg-score could be used to examine hypoxia responsiveness.

105 Enrichment analysis focusing only on protein-coding genes revealed that the top 200 genes
106 identified by HNg-score, a measure of susceptibility to hypoxia, were enriched for hypoxia response-
107 related gene sets (Fig. 3A). However, enrichment analysis of the bottom 200 genes identified based on
108 their HNg-score revealed suppression of expression of ncRNA metabolism- related genes (Fig. 3B).
109 Detailed investigation showed that most of these genes were associated with rRNA processing.

110 In enrichment analysis, the ncRNA metabolic process was further subdivided into the following
111 Gene Ontology (GO) terms: ncRNA processing, ribosome biogenesis, ribonucleoprotein complex
112 biogenesis, rRNA processing, and rRNA metabolic process (Fig. 3C). The results of rRNA GO
113 network analysis were visualized in Upset plots, which showed that the ncRNA metabolic process is
114 heavily reliant on most of these rRNA metabolic genes (Fig. 3D).

115 **Survey of the number of papers referring FANTOM**

116 FANTOM-CAT is a comprehensive catalog of human lncRNAs that improves upon existing lncRNA
117 transcript models. The results obtained using this dataset have greatly contributed to life science
118 research (Imada et al, 2020; Ramilowski et al, 2020; Imada et al, 2021). Despite the usefulness of this
119 dataset, we predicted that it is used less frequently for RNA-sequencing (RNA-seq) quantification
120 compared to GENCODE; we quantitatively examined this hypothesis. If FANTOM-CAT is used less
121 frequently, results obtained using this database may show a lower influence of publication bias (Ono
122 and Bono 2021), as demonstrated in our previous report. The number of articles in PubMed Central
123 (PMC) that contained references to GENCODE, FANTOM, and RNA-seq was calculated for each
124 year, with the number of RNA-seq articles exceeding 30,000 in 2021, while the number of articles
125 referencing both FANTOM and RNA-seq was 0.5%, about one-tenth of the number referencing RNA-
126 seq and GENCODE (Table 1).

127 **HNf-score evaluations of transcripts using FANTOM-CAT**

128 The HNf-scores of all transcripts, including ncRNA, were quantified using FANTOM-CAT as a
129 reference. We then went on to evaluate both the most increased and decreased transcripts using these
130 values as a guide (Fig. 4 A). A total of 25 transcripts for both the UP and DOWN data sets, as
131 delimited by HNf-score, are listed in Table 2 and a complete list of genes can be found within our
132 archived figshare file (https://doi.org/10.6084/m9.figshare.19679493.v1).
133 This investigation revealed the fact that there were several cases where hypoxia-responsive
134 transcripts with high HNf-scores were located within the antisense regions of genes with high HNg-
135 scores (*BHLHE40-AS1*, *SLC2A1-AS1*). In addition, these evaluations uncovered that the majority of
136 low HNf-score transcripts were associated with mitochondrial genes (*MT-ND2*, *MT-ATP8*). Hypoxic
137 stimuli were shown to affect mitochondria. Visualization of the HNf-score for each chromosome
138 showed that the median HNf-score was negative for mitochondria-derived transcripts (Fig. 4B).
139 **Investigating the HNf-score of antisense transcripts**
140 We investigated each antisense transcript individually, as we could not perform enrichment analysis
141 which requires information comprehensively evaluated the effects of sense and antisense transcripts.
142 The antisense transcripts for each of the UP200 and DOWN 200 hypoxia-responsive coding genes
143 were classified into one of four categories based on their gene expression behavior (Fig. 5A, Table 3).
144 These evaluations revealed that approximately 75% of the antisense transcripts mirrored the
145 behavior of their sense counterparts. However, there was still a small number of sense-antisense pairs

146 presenting with opposite expression behaviors. This indicated that they were under different
147 transcriptional control.
148 We also investigated transcripts with high or low HNf-scores to qualitatively evaluate hypoxia
149 responsiveness. This evaluation also revealed that the Top 10 antisense transcripts in the UP200 gene
150 list were dominated by both short and long ncRNAs (lncRNAs), while the bottom 10 transcripts from
151 this group were shown to be predominantly mRNA transcripts.
152 Interestingly, among the antisense transcripts from the DOWN200 gene set, phosphoglycerate
153 kinase 1 (*PGK1*), which was also included in the UP200 gene set, had the highest HNf-score.
154 Given this, we homed in on *PGK1* and its antisense TATA-box binding protein associated factor 9b
155 (*TAF9B*), evaluating their chromosomal localization using UCSC Genome Browser and HN-score at
156 the gene level (Fig. 5B, C). This analysis revealed that while these two genes are not described as
157 having a sense-antisense relationship when evaluated using the UCSC or GENCODE datasets,
158 FANTOM-CAT-based analysis suggests that these two transcripts are actually in a sense-antisense
159 relationship. By using FANTOM-CAT as a reference, we were able to present for the first time the
160 possibility that these two genes affect the regulation of expression.
161
162 **Discussion**
163 Human gene research is commonly biased towards known pathogenic genes and pathways that are
164 fairly well established (Stoeger and Amaral 2022). However, given the development of novel data-

165 driven tools, it is now possible to move beyond the targets identified in previous papers and into the
166 realm of big data. Our previous studies focusing on coding genes have revealed a clear publication
167 bias when evaluating gene expression changes in response to hypoxia and identified a novel set of
168 largely uncharacterized hypoxia-responsive elements (Ono and Bono 2021). Despite this, the details
169 of the "ncRNA metabolic process" related genes whose gene expression was suppressed in response
170 to hypoxic conditions remain unclear.

171 Thus, we designed this study to further clarify the hypoxic response using a two-pronged approach.
172 First, we completed a detailed investigation of the "ncRNA metabolic process" genes focused on only
173 the coding genes in the dataset, and then we expanded this evaluation to any of the hypoxia-
174 responsive genes annotated in the FANTOM-CAT database, which includes both coding and non-
175 coding elements (Fig. 1). All these data are available to everyone on figshare
176 (https://doi.org/10.6084/m9.figshare.c.5971218.v2). We believe that these data sets and this analysis
177 approach will help to further elucidate the mechanisms regulating the hypoxic response.

178 We previously reported that the expression of several genes associated with the "ncRNA metabolic
179 process" were downregulated by hypoxic stimuli. However, the specific metabolic processes remain
180 unknown. Given this, we started our evaluations by completing a gene ontology evaluation in an
181 effort to identify the likely functions of these target transcripts. Subsequent enrichment analysis of the
182 top and bottom 200 genes, identified using their HNg-score which was calculated using GENCODE,
183 revealed that the top 200 genes were commonly associated with hypoxia response-related transcripts,

184 while the bottom 200 genes were enriched for ncRNA metabolic processes (Fig. 3 A,B). Then GO
185 network analysis revealed a strong relationship between the ncRNA metabolic process and rRNA
186 processing genes (Fig. 3 C,D) (https://www.ebi.ac.uk/QuickGO/GTerm?id=GO: 0034660). This
187 suggests that amongst the most common downregulated transcripts most targeted or formed part of
188 the rRNA processing pathway.

189 rRNA is a critical part of the ribosome known to facilitate protein synthesis. rRNA has also been
190 linked to the degradation of cytoplasmic lncRNAs following their interaction with the ribosomes
191 (Carlevaro-Fita et al, 2016), and several other reports describe ncRNA mediated silencing of the
192 rRNA (Schmitz et al, 2010). All of which suggests an intimate relationship between these two sets of
193 RNA mediators. However, given our previous focus on only the coding genes, the effects of hypoxia
194 on the total transcript population, including the ncRNAs, remains unclear. Given this, we went on to
195 calculate the more comprehensive HNf-score, which evaluates all RNA transcripts including ncRNAs
196 via the application of a FANTOM-CAT reference.

197 FANTOM-CAT is expected to provide additional insights through integration with other studies,
198 such as its use in a comprehensive expression atlas across the broad human transcriptome called FC-
199 R2 (Imada et al, 2020). The application of this dataset allows for a more comprehensive evaluation of
200 changes in the gene expression data from human experiments, expanding the evaluations to the
201 ncRNA transcripts. However, to the best of our knowledge, FANTOM-CAT is only published as a
202 General Transfer Format (GTF) file and is not available in FASTA format, thus the underlying data is

203 not publicly available. As this dataset may not be commonly applied as a reference point for RNA-
204 seq, we investigated the number of articles referencing both "RNA-seq" and "FANTOM" in PMC.
205 To investigate how often RNA-seq quantification is performed using FANTOM-CAT, the number of
206 RNA-seq publications using FANTOM-CAT and GENCODE was searched and quantified in PMC.
207 The number of papers with both "RNA-seq" and "FANTOM" accounted for less than 1% of the total
208 that describes "RNA-seq" data even after 2017 when the FANTOM-CAT paper was first published
209 (Table 1). This data also indicated that there were only a few reports describing gene expression
210 quantification using FANTOM-CAT as a reference and that most studies continue to focus on more
211 well-established gene sets, potentially overlooking critical actors in the under-evaluated data (Ono and
212 Bono 2021). When we combined these observations, we decided that a comprehensive analysis,
213 including the ncRNAs, was likely worth the effort when discussing hypoxia. Given this, we went on
214 to complete a FANTOM-CAT-based evaluation of the hypoxia data described above.
215
216 To this end, we created a new FASTA file using the FANTOM-CAT genome annotation GTF file
217 and then added the HNf-score to each of the FANTOM-CAT lv4 transcripts (Fig. 4A). Subsequent
218 analysis of this file then revealed that the bottom 25 genes in the HNf-score dataset were
219 predominantly mitochondrial DNA-derived transcripts. Visualization of this patterning then revealed
220 likely regulation of gene expression by chromosome (Fig. 4B, Table 2), and confirmed that these
221 effects were most pronounced for mitochondrial transcripts. The downregulation of mitochondrial

222 genes such as *MT-ND2*, *MT-ATP8*, and *MT-ND* by hypoxic stimulation is already well known (Arnaiz
223 et al, 2021) validating our data. It has also been reported (Zhang et al, 2008) that mitochondrial
224 autophagy is commonly induced in response to hypoxia and that this process requires hypoxia-
225 dependent factor-1-dependent BNIP3 expression, which also supports our results. On the contrary, the
226 top 25 transcripts from the HNf-score dataset were dominated by antisense reads, such as *BHLHE40-*
227 *AS1* and *SLC2A1-AS1* (Table 2). This was an interesting observation, and its implications were
228 supported by the fact that the sense transcripts for both *BHLHE40* and *SLC2A1* are well-established
229 hypoxia response genes. In addition, RP1-261G23.7, which had the highest HNf-score, is a short
230 ncRNA localized to the antisense of vascular endothelial growth factor A (*VEGFA*), a well-known
231 hypoxia response gene. This short ncRNA has been reported to be functionally involved in the
232 transcriptional regulation of *VEGFA* during hypoxia (Nieminen et al, 2018). Based on these findings,
233 we thought that the transcripts from the antisense strands of known hypoxia-responsive genes were
234 likely worth further investigation.
235
236 Antisense transcripts are transcribed from the complementary strand of the DNA molecule and
237 usually overlap with other sense transcripts encoding both proteins and noncoding regulators. It is also
238 worth noting that approximately 30% of human (Ozsolak et al, 2010) and mouse (Katayama et al,
239 2005) transcripts originate from the antisense strand. Given this, we went on to select 200 hypoxia-

240 responsive genes from both the UP and DOWN regulated gene sets using their HNg-score. We then
241 evaluated the HNf-score of all antisense transcripts of these genes.
242 This revealed that 75% of the UP and DOWN regulated antisense transcripts presented with a similar
243 pattern of expression to their sense partner, while the other 25% of antisense transcripts were shown
244 to be regulated in opposition to their sense partner (Fig. 5A). Several functions are regulated in the
245 same direction. Some antisense transcripts enhance the translational function of sense transcripts or
246 repress degradation and translational repression (Carrieri et al, 2012; Faghihi et al, 2010).
247 Further evaluation then revealed that the bottom 10 antisense transcripts in the DOWN200 gene set
248 were dominated by mitochondrial genes (Table 2, Fig. 4B). Interestingly, the Top 10 transcripts from
249 the DOWN200 gene set included *PGK1* and antisense *TAF9B* whose expression was suppressed in
250 response to hypoxia (Fig. 5C). This is of specific interest as *PGK1* catalyzes the reversible conversion
251 of 1,3 -diphosphoglycerate to 3-phosphoglycerate in the glycolytic system to generate ATP (Valentin
252 et al, 1998) and this gene is well known for its increased expression in response to HIF1A signaling
253 following hypoxic induction (Li et al, 1996). Meanwhile, TAF9B is one of the proteins associated
254 with the TFIID and plays an important role in transcriptional initiation (Tora 2002; Frontini et al,
255 2005). TAF9B is also known to be upregulated in response to HIF1A knockdown and downregulated
256 under hypoxic conditions (Mathieu et al, 2011). Using FANTOM-CAT, the relationship between
257 *PGK1* and *TAF9B* was shown to be that of a sense-antisense pair. This relationship is not present in
258 either the UCSC gene or GENCODE annotations (Fig. 5B).

259 In addition, the regulation of genes by their antisense transcripts has been extensively studied
260 (Pelechano and Steinmetz 2013; Okada et al, 2008; Katayama et al, 2005), with many studies
261 focusing on antisense transcript-mediated epigenetic regulation and the possibility that when
262 transcription is initiated simultaneously from both sense and antisense start sites, the factors involved
263 in their transcription may interact to repress gene expression. These results suggest that the 3′ end of
264 *PGK1* affects the expression of *TAF9B*.
265
266 This data sets used in these evaluations were not derived from strand-specific RNA sequencing.
267 However, these findings suggest that strand-specific RNA-seq or knockdown experiments may
268 facilitate a better understanding of the hypoxic response and should be considered in the future.
269
270 Thus, this paper describes the use of FANTOM-CAT and other open data sources to facilitate the
271 exploratory analysis of existing data. We first evaluated the ncRNA metabolic process and provided a
272 dataset that will help generate hypotheses for clarifying specific ncRNA intracellular responses. We
273 used this approach to further characterize hypoxia-responsive transcripts and overcome the native bias
274 in genome research towards known genes. Our study used the increased processing power of
275 computers to enable the application of big data analysis to drive discovery. We believe that this data-
276 driven approach is increasingly necessary to help advance genomic research. Our data suggest that
277 FANTOM-CAT has not been widely utilized in research despite its obvious value. Thus, we designed

this study to showcase the value of FANTOM-CAT derived transcript annotation as a means to uncover new targets for investigation and validation. Our evaluation of the hypoxia data uncovered a novel connection between *PGK1* and *TAF9B* which was not reported in any other evaluations despite their critical value in hypoxia.

Thus, we suggest that this approach can be applied to other questions to facilitate more nuanced, and less biased evaluations of large RNA-seq data sets and propose that FANTOM-CAT should be commonly evaluated as an indispensable tool in transcriptome evaluations.

Materials and Methods

Evaluating the number of papers with descriptions of FANTOM

We used the following queries across our PMC web search (conducted on April 9, 2022)

RNA-Seq **:** “rna-seq”[MeSH Terms] OR “rna-seq”[All Fields] OR (“rna”[All Fields] AND “seq”[All Fields]) OR “rna seq”[All Fields].

RNA-Seq AND GENCODE : (“rna-seq”[MeSH Terms] OR “rna-seq”[All Fields] OR (“rna”[All Fields] AND “seq”[All Fields]) OR “rna seq”[All Fields]) AND “gencode”[All Fields]

RNA-Seq AND FANTOM : (“rna-seq”[MeSH Terms] OR “rna-seq”[All Fields] OR (“rna”[All Fields] AND “seq”[All Fields]) OR “rna seq”[All Fields]) AND “fantom”[All Fields] to identify any

296 relevant publications. Once identified we downloaded the relevant article information in MEDLINE

297 format and calculated the number of publications per year based on the Date of Publication.

298

299 **Public gene expression data**

300 We then collected the RNA-seq data (https://doi.org/10.6084/m9.figshare.14141219.v1) from 495

301 pairs of human hypoxia-Normoxia samples evaluated in our previous study (Ono and Bono 2021).

302

303 **Gene expression quantification**

304 The expression of coding genes was quantified using ikra 1.2.3 (https://github.com/yyoshiaki/ikra)

305 as previously reported. The following procedure was used to quantify gene expression when

306 FANTOM-CAT was used as the reference sequence. As there were no FASTA-formatted files for

307 these data recorded in the FANTOM-CAT repository, we created a FASTA file based on the General

308 Transfer Format file data using gffread v0.12.1 and the following command:

309 > gffread FANTOM_CAT.lv4_stringent.only_lncRNA.gtf -g hg19.fa -w lv4.fa

310 We used salmon 0.14.0 for index creation and quantification

311 (https://github.com/no85j/hypoxia_code/tree/master/salmon/salmon_v0.14.0), with some

312 modifications to Pitagora-cwl (https://github.com/pitagora-network/pitagora-

313 cwl/tree/master/tools/salmon), and tximport 1.18.0

314 (https://github.com/no85j/hypoxia_code/tree/master/tximport) to match as closely as possible with the

315 quantification method used in our previous studies (Ono and Bono 2021; Bono and Hirota 2020). We
316 then calculated scaledTPM for use this as a representative value for gene expression (Soneson et al,
317 2016). Our quantitative RNA-seq data are accessible at figshare
318 (https://doi.org/10.6084/m9.figshare.19679520.v1)
319
320 **HN-score**
321 The fold change of the [scaledTPM]+1 values for each of the 495 sample pairs was calculated and
322 differential expression was determined at a 1.5-fold cut-off. We then produced the Hypoxia-Normoxia
323 (HN)-score for this data by taking the number of samples demonstrating significant upregulation and
324 subtracting the number of samples with significant downregulation of their overall expression profile.
325 Here we added a second HN-score calculation method using different reference transcripts to allow
326 for a more nuanced evaluation. These scores were then referred to as HNg (coding genes) and HNf
327 (coding and non-coding gene), respectively. This first score was designed to evaluate the coding gene
328 and was calculated from the previously published data set
329 (https://doi.org/10.6084/m9.figshare.14141135.v1) and the previously described HN1.5 cut-off value.
330 This means that the HNf-score is effectively a descriptor of the HN-score for any FANTOM-CAT
331 transcripts. This HNf-score was calculated as described above except that FANTOM-CAT was used
332 as a reference. Transcript annotation was obtained from the FANTOM-CAT website

333 (https://fantom.gsc.riken.jp/cat/v1/#/genes) and the HNf-score data are accessible at figshare

334 (https://doi.org/10.6084/m9.figshare.19679508.v2)

335

336 **Enrichment analysis**

337 We created UP200 and Down200 gene sets for the top and bottom 200 gene lists of the HNg-score

338 respectively and performed enrichment analysis in Metascape (https://metascape.org/)(Zhou et al,

339 2019). We used the default parameters for Metascape settings.

340 **Antisense transcript**

341 We used the bedtools (v2.30.0) program to complete the following evaluations (Quinlan and Hall

342 2010), including the creation of a bed file in which all the regions of each gene were merged based on

343 the transcript id of the UP200 and DOWN200 genes identified in the coding gene dataset. We then

344 used this bed file to identify the transcripts in the reverse strand using the intersect command.

345

346 **Visualization**

347 We used TIBCO Spotfire Desktop version 11.5.0 (TIBCO Spotfire, Inc., Palo Alto, CA, USA) to

348 produce all of our bar and scatter plots and we used the UpSetR (1.4.0) package from R (4.0.3) to

349 create our upset plots. We then used the pandas (1.2.5), matplotlib (3.2.2), and seaborn (0.11.0)

350 packages in Python (3.8.10) to complete our violin plotting and visualized the genome coordinates by

351 adding "FANTOM5 summary Tracks" to the UCSC genome browser.

DATA AVAILABILITY

Identifiers of the RNA-seq data used in this study are listed in the link to the left [https://doi.org/10.6084/m9.figshare.14141219.v1].

Acknowledgments

The authors would like to thank Drs. Takeya Kasukawa and Masaki Suimye Morioka for their input and discussion.

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.

Conflict of interest

The authors report that there are no conflicts of interest.

References

371 Arnaiz E, Miar A, Dias Junior AG, Prasad N, Schulze U, Waithe D, Nathan JA, Rehwinkel J, Harris
372 AL (2021) Hypoxia regulates endogenous double-stranded RNA production via reduced
373 mitochondrial DNA transcription. *Front Oncol* **11**: 779739; doi:10.3389/fonc.2021.779739.
374 Bono H, Hirota K (2020) Meta-analysis of hypoxic transcriptomes from public databases.
375 *Biomedicines* **8**: 10; doi:10.3390/biomedicines8010010.
376 Carlevaro-Fita J, Rahim A, Guigó R, Vardy LA, Johnson R (2016) Cytoplasmic long noncoding
377 RNAs are frequently bound to and degraded at ribosomes in human cells. *RNA* **22**: 867–882;
378 doi:10.1261/rna.053561.115.
379 Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, Pesce E, Ferrer I, Collavin L,
380 Santoro C, et al (2012) Long non-coding antisense RNA controls Uchl1 translation through
381 an embedded SINEB2 repeat. *Nature* **491**: 454–457.

382 Ebert BL, Bunn HF (1998) Regulation of transcription by hypoxia requires a multiprotein complex
383 that includes hypoxia-inducible factor 1, an adjacent transcription factor, and p300/CREB
384 binding protein. *Mol Cell Biol* **18**: 4089–4096; doi: 10.1128/MCB.18.7.4089.
385 Faghihi MA, Zhang M, Huang J, Modarresi F, Van der Brug MP, Nalls MA, Cookson MR, St-
386 Laurent G, Wahlestedt C (2010) Evidence for natural antisense transcript-mediated
387 inhibition of microRNA function. *Genome Biol* **11**: R56.

388 Frontini M, Soutoglou E, Argentini M, Bole-Feysot C, Jost B, Scheer E, Tora L (2005) TAF9b
389 (formerly TAF9L) is a bona fide TAF that has unique and overlapping roles with TAF9. *Mol*
390 *Cell Biol* **25**: 4638–4649; doi:10.1128/MCB.25.11.4638-4649.2005.

391 Hon C-C, Ramilowski JA, Harshbarger J, Bertin N, Rackham OJL, Gough J, Denisenko E, Schmeier
392 S, Poulsen TM, Severin J, et al (2017) An atlas of human long non-coding RNAs with accurate
393 5′ ends. *Nature* **543**: 199–204; doi:10.1038/nature21374.

394 Imada EL, Sanchez DF, Collado-Torres L, Wilks C, Matam T, Dinalankara W, Stupnikov A, Lobo-
395 Pereira F, Yip C-W, Yasuzawa K, et al (2020) Recounting the FANTOM CAGE-Associated
396 Transcriptome. *Genome Res* **30**: 1073–1081.

397 Imada EL, Sanchez DF, Dinalankara W, Vidotto T, Ebot EM, Tyekucheva S, Franco GR, Mucci LA,
398 Loda M, Schaeffer EM, et al (2021) Transcriptional landscape of PTEN loss in primary
399 prostate cancer. *BMC Cancer* **21**: 856.

400 Jaakkola P, Mole DR, Tian YM, Wilson MI, Gielbert J, Gaskell SJ, von Kriegsheim A, Hebestreit
401 HF, Mukherji M, Schofield CJ, et al (2001) Targeting of HIF-alpha to the von Hippel-Lindau
402 ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* **292**: 468–472;
403 doi:10.1126/science.1059796.

404 Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC,
405 Suzuki M, Kawai J, et al (2005) Antisense transcription in the mammalian transcriptome.
406 *Science* **309**: 1564–1566; doi:10.1126/science.1112009.

Li H, Ko HP, Whitlock JP (1996) Induction of phosphoglycerate Kinase 1 gene expression by hypoxia. *J Biol Chem* **271**: 21262–21267; doi:10.1074/jbc.271.35.21262.

Mahon PC, Hirota K, Semenza GL (2001) FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity. *Genes Dev* **15**: 2675–2686; doi: 10.1101/gad.924501.

Mathieu J, Zhang Z, Zhou W, Wang AJ, Heddleston JM, Pinna CMA, Hubaud A, Stadler B, Choi M, Bar M, et al (2011) HIF Induces Human Embryonic Stem Cell Markers in Cancer Cells. *Cancer Res* **71**: 4640–4652; doi: 10.1158/0008-5472.CAN-10-3320.

Mihai IS, Das D, Maršalkaite G, Henriksson J (2021) Meta-analysis of gene popularity: less than half of gene citations stem from gene regulatory networks. *Genes* **12**.doi: 10.3390/genes12020319.

Nieminen T, Scott TA, Lin F-M, Chen Z, Yla-Herttuala S, Morris KV (2018) Long non-coding RNA modulation of VEGF-A during hypoxia. *Non-Coding RNA* **4**: 34; doi:10.3390/ncrna4040034.

Okada Y, Tashiro C, Numata K, Watanabe K, Nakaoka H, Yamamoto N, Okubo K, Ikeda R, Saito R, Kanai A, et al (2008) Comparative expression analysis uncovers novel features of endogenous antisense transcription. *Hum Mol Genet* **17**: 1631–1640; doi: 10.1093/hmg/ddn051.

Ono Y, Bono H (2021) Multi-Omic meta-analysis of transcriptomes and the bibliome uncovers novel hypoxia-inducible genes. *Biomedicines* **9**: 582; doi: 10.3390/biomedicines9050582.

424 Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, John B, Milos PM (2010)
425 Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative
426 polyadenylation. *Cell* **143**: 1018–1029; doi: 10.1016/j.cell.2010.11.020.
427 Pelechano V, Steinmetz LM (2013) Gene regulation by antisense transcription. *Nat Rev Genet* **14**:
428 880–893; doi: 10.1038/nrg3594.
429 Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features.
430 *Bioinformatics* **26**: 841–842; doi:10.1093/bioinformatics/btq033.
431 Ramilowski JA, Yip CW, Agrawal S, Chang J-C, Ciani Y, Kulakovskiy IV, Mendez M, Ooi JLC,
432 Ouyang JF, Parkinson N, et al (2020) Functional annotation of human long noncoding RNAs
433 via molecular phenotyping. *Genome Res* **30**: 1060–1072.
434 Schmitz K-M, Mayer C, Postepska A, Grummt I (2010) Interaction of noncoding RNA with the
435 rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev* **24**:
436 2264–2269; doi:10.1101/gad.590910.
437 Semenza GL, Wang GL (1992) A nuclear factor induced by hypoxia via de novo protein synthesis
438 binds to the human erythropoietin gene enhancer at a site required for transcriptional activation.
439 *Mol Cell Biol* **12**: 5447–5454; doi: 10.1128/mcb.12.12.5447-5454.1992.
440 Soneson C, Love MI, Robinson MD (2016) Differential analyses for RNA-seq: transcript-level
441 estimates improve gene-level inferences. *F1000Res* **4**: 1521; doi:
442 10.12688/f1000research.7563.2.

443 Stoeger T, Amaral LAN (2022) The characteristics of early-stage research into human genes are
444 substantially different from subsequent research. *PLoS Biol* **20:** e3001520; doi:
445 10.1371/journal.pbio.3001520.
446 Tora L (2002) A unified nomenclature for TATA box binding protein (TBP)-associated factors
447 (TAFs) involved in RNA polymerase II transcription. *Genes Dev* **16**: 673–675;
448 doi:10.1101/gad.976402.
449 Valentin C, Birgens H, Craescu CT, Brødum-Nielsen K, Cohen-Solal M (1998) A phosphoglycerate
450 kinase mutant (PGK Herlev; D285V) in a Danish patient with isolated chronic hemolytic anemia:
451 mechanism of mutation and structure-function relationships. *Hum Mutat* **12**: 280–287;
452 doi:10.1002/(SICI)1098-1004(1998)12:4<280::AID-HUMU10>3.0.CO;2-V.
453 Wang GL, Jiang BH, Rue EA, Semenza GL (1995) Hypoxia-inducible factor 1 is a basic-helix-loop-
454 helix-PAS heterodimer regulated by cellular O2 tension. *Proc Natl Acad Sci U S A* **92**: 5510–
455 5514; doi: 10.1073/pnas.92.12.5510.
456 Wang GL, Semenza GL (1995) Purification and characterization of hypoxia-inducible factor 1. *J Biol*
457 *Chem* **270**: 1230–1237; doi:10.1074/jbc.270.3.1230.
458 Zhang H, Bosch-Marce M, Shimoda LA, Tan YS, Baek JH, Wesley JB, Gonzalez FJ, Semenza GL
459 (2008) Mitochondrial autophagy is an HIF-1-dependent adaptive metabolic response to hypoxia.
460 *J Biol Chem* **283**: 10892–10903; doi: 10.1074/jbc.M800102200.

461 Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK

462 (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level

463 datasets. *Nat Commun* **10**: 1523; doi: 10.1038/s41467-019-09234-6.

464

465 **Figure legends**

466 **Figure 1 - Schematic overview describing our exploratory study of hypoxia-responsive**

467 **transcripts.**

468 Data source: human hypoxia RNA-seq data from the NCBI Gene Expression Omnibus database.

469 Protein coding genes only: evaluated using GENCODE as the reference. Evaluation of the genes

470 downregulated by hypoxia revealed that they were enriched for transcripts involved in rRNA

471 processing. Protein coding and non-coding transcripts: Evaluated using FANTOM-CAT as reference.

472 Number of transcripts derived from mitochondrial DNA and antisense analysis was reduced,

473 suggesting a transcriptional regulation relationship in both sense and antisense genes (*PGK1* and

474 *TAF9B*).

475 **Figure 2 - Contents of hypoxia-related data set.**

476 The dataset is **(A)** summarized for oxygen concentration and **(B)** treatment time. **(C)** Gene expression

477 information for *PGK1* derived from cancer cells treated for 24 h with a 1% oxygen concentration, the

478 most representative condition, is displayed in a boxplot.

479 **Figure 3 - Hypoxic stimulation suppresses the expression of non-coding RNA (ncRNA)**

480 **metabolic process-related genes involved in rRNA processing.**

481 Enrichment analysis for (A) the 200 UP regulated gene list and (B) the 200 DOWN regulated gene

482 list. (C) Gene Ontology (GO) network and (D) Upset plot showing the number of genes included in

483 GO in this analysis describing the relative changes in ncRNA metabolism associated genes in the

484 downregulated gene list. In the GO Network, the ncRNA metabolic process and rRNA processing

485 focused on in this study are shown in yellow.

486 **Figure 4 - Hypoxia-normoxia (HN)-score of coding and non-coding genes (HNf-score)**

487 **evaluation of transcript using FANTOM-CAT.**

488 **(A)** HNf-score distribution of FANTOM-CAT transcripts. **(B)** Violin plot of the HNf-score of each

489 chromosome.

490 **Figure 5. Hypoxia-normoxia (HN)-score of coding and non-coding genes (HNf-score)** antisense

491 transcripts in connection to their sense counterparts.

492 **(A)** Distribution of HNf-scores for each antisense transcript from each gene in the UP200 and

493 DOWN200 gene lists. **(B)** Visualization of *PGK1* and *TAF9B* loci in the UCSC

494 genome browser. In FANTOM-CAT, *PGK1* transcripts are a sense-antisense pair for *TAF9B*, but

495 GENCODE does not show this relationship. **(C)** HN-score for coding genes (HNg-score) for *PGK1*

496 and *TAF9B*.

497

498

499 **Tables**

500 **Table 1. Numbers of reports describing GENCODE, FANTOM, and RNA-seq data**

Year	“RNA-seq”	“RNA-seq” &”GENCODE” (%)	“RNA-seq” &”FANTOM” (%)
2017	17,373	769 (4.4%)	123 (0.7%)
2018	20,400	996 (4.9%)	127 (0.6%)
2019	25,668	1,254 (4.9%)	130 (0.5%)
2020	33,607	1,593 (4.7%)	176 (0.5%)
2021	40,757	1,976 (4.8%)	210 (0.5%)
2022	9,420	406 (4.3%)	39 (0.4%)

501

502

503 **Table 2. Top 25 UP and DOWN regulated transcripts identified using their HNf-scores**

		Transcript ID	HNf-score	Chr	Gene Name	Gene Class	Antisense
Top 25 upregulated transcripts	1	ENST00000607600.1	321	chr6	RP1-261G23.7	Short ncRNAs	AS_VEGFA
	2	ENST00000481651.1	309	chr7	RP11-61L23.2	Pseudogenes	
	3	ENST00000307365.3	301	chr10	DDIT4	Protein-coding mRNAs	
	4	FTMT21000000269.1	300	chr3	BHLHE40-AS1	lncRNA, intergenic	AS_BHLHE40
	5	ENST00000368636.4	289	chr10	BNIP3	Protein-coding mRNAs	
	6	ENST00000335174.4	275	chr4	ANKRD37	Protein-coding mRNAs	
	7	ENST00000290573.2	272	chr2	HK2	Protein-coding mRNAs	
	8	FTMT24300004891.1	267	chr11	LDHA	Protein-coding mRNAs	
	9	HBMT00000242044.1	265	chr11	SBF2	Protein-coding mRNAs	AS_ADM
	10	FTMT24200000803.1	265	chr11	CATG00000004979.1	lncRNA, antisense	AS_LDHA
	11	ENST00000380629.2	265	chr8	BNIP3L	Protein-coding mRNAs	
	12	ENCT00000004975.1	263	chr1	SLC2A1-AS1	lncRNA, divergent	AS_SLC2A1
	13	ENST00000543445.1	257	chr11	LDHA	Protein-coding mRNAs	
	14	ENST00000378357.4	257	chr9	CA9	Protein-coding mRNAs	
	15	ENST00000453116.1	254	chr10	MXI1	Protein-coding mRNAs	
	16	ENST00000471240.1	253	chr10	DDIT4	Protein-coding mRNAs	
	17	ENST00000460806.1	251	chr3	BHLHE40	Protein-coding mRNAs	
	18	FTMT20200001505.1	250	chr1	SLC2A1	Protein-coding mRNAs	
	19	HBMT00000734702.1	250	chr19	CATG00000040757.1	lncRNA, divergent	AS_GPI
	20	FTMT21100041760.1	249	chr3	BHLHE40	Protein-coding mRNAs	
	21	ENST00000250457.3	246	chr14	EGLN3	Protein-coding mRNAs	
	22	HBMT00000734730.1	245	chr19	CATG00000040757.1	lncRNA, divergent	AS_GPI
	23	ENST00000426263.3	243	chr1	SLC2A1	Protein-coding mRNAs	
	24	HBMT00000727478.1	242	chr19	RAB11B-AS1	lncRNA, divergent	AS_ANGPTL4
	25	ENST00000534464.1	241	chr11	ADM	Protein-coding mRNAs	
Top 25 Downregulated transcripts	1	ENST00000320270.2	-248	chr8	RRS1	Protein-coding mRNAs	
	2	ENST00000368232.4	-246	chr1	GPATCH4	Protein-coding mRNAs	
	3	ENST00000361453.3	-245	chrM	MT-ND2	Protein-coding mRNAs	
	4	ENST00000361851.1	-230	chrM	MT-ATP8	Protein-coding mRNAs	
	5	ENST00000361227.2	-221	chrM	MT-ND3	Protein-coding mRNAs	
	6	ENST00000361390.2	-215	chrM	MT-ND1	Protein-coding mRNAs	
	7	ENCT00000264196.1	-200	chr20	IDH3B	Protein-coding mRNAs	AS_NOP56
	8	ENST00000362079.2	-194	chrM	MT-CO3	Protein-coding mRNAs	

	9	FTMT30000000001.1	-193	chrM	MT-ND5	Protein-coding mRNAs	AS_MT-ND6
	10	ENCT00000020267.1	-188	chr1	CATG00000042732.1	lncRNA, divergent	
	11	FTMT20300078215.1	-188	chr1	APOA1BP	Protein-coding mRNAs	AS_GPATCH4
	12	FTMT26100009680.1	-188	chr16	POLR3K	Protein-coding mRNAs	
	13	FTMT24100001922.1	-187	chr11	H2AFX	Protein-coding mRNAs	AS_HMBS
	14	MICT00000370386.1	-185	chrM	MT-TA	Structural RNAs	AS_MT-ATP6 etc.
	15	ENST00000416718.2	-177	chr1	RP5-857K21.11	Pseudogenes	
	16	HBMT00000533809.1	-176	chr16	ALG1	Protein-coding mRNAs	AS_EEF2KMT
	17	FTMT22300045607.1	-176	chr6	SRSF3	Protein-coding mRNAs	
	18	ENST00000585075.1	-175	chr17	RP11-649A18.12	lncRNA, divergent	AS_SLC25A19
	19	ENST00000295304.4	-175	chr2	CHAC2	Protein-coding mRNAs	
	20	MICT00000370388.1	-175	chrM	MT-TA	Structural RNAs	AS_MT-ATP6 etc.
	21	ENST00000361899.2	-175	chrM	MT-ATP6	Protein-coding mRNAs	
	22	ENST00000458605.1	-174	chr22	RRP7B	Pseudogenes	
	23	HBMT00000611042.1	-172	chr17	METTL23	Protein-coding mRNAs	
	24	ENST00000371538.3	-170	chr1	SELRC1	Protein-coding mRNAs	
	25	ENST00000293860.5	-169	chr16	POLR3K	Protein-coding mRNAs	

506 **Table 3. Top and bottom 10 antisense transcripts for UP200 and DOWN200 HNg-score gene**

507 **sets, as identified based on their HNf-scores**

			Transcript ID	HNf-score	Gene Name	Gene Class	Antisense
Antisense transcripts from the UP200 gene set	Top10	1	ENST00000607600.1	321	RP1-261G23.7	Short ncRNAs	AS_VEGFA
		2	FTMT21000000269.1	300	BHLHE40-AS1	lncRNA, intergenic	AS_BHLHE40
		3	HBMT00000242044.1	265	SBF2	Protein-coding mRNAs	AS_ADM
		4	FTMT24200000803.1	265	CATG00000004979.1	lncRNA, antisense	AS_LDHA
		5	ENCT00000004975.1	263	SLC2A1-AS1	lncRNA, divergent	AS_SLC2A1
		6	HBMT00000734702.1	250	CATG00000040757.1	lncRNA, divergent	AS_GPI
		7	HBMT00000734730.1	245	CATG00000040757.1	lncRNA, divergent	AS_GPI
		8	HBMT00000727478.1	242	RAB11B-AS1	lncRNA, divergent	AS_ANGPTL4
		9	ENCT00000211059.1	241	RAB11B-AS1	lncRNA, divergent	AS_ANGPTL4
		10	HBMT00000734676.1	234	CATG00000040757.1	lncRNA, divergent	AS_GPI
	Bottom10	1	ENCT00000244163.1	-152	RP11-259N19.1	lncRNA, divergent	AS_HK2
		2	HBMT00000507967.1	-124	CATG00000025826.1	Protein-coding mRNAs	AS_ISG20
		3	ENST00000309424.3	-115	CD3EAP	Protein-coding mRNAs	AS_PPP1R13L
		4	FTMT21900007493.1	-111	SRFBP1	Protein-coding mRNAs	AS_LOX
		5	ENST00000524270.1	-94	SPSB2	Protein-coding mRNAs	AS_TPI1
		6	ENST00000357429.6	-87	C7orf50	Protein-coding mRNAs	AS_GPR146
		7	ENST00000462901.1	-81	CGGBP1	Protein-coding mRNAs	AS_ZNF654
		8	ENST00000439489.1	-81	DIABLO	Protein-coding mRNAs	AS_B3GNT4
		9	MICT00000230070.1	-80	PPIL2	Protein-coding mRNAs	AS_YPEL1
		10	ENST00000422285.2	-79	PDHA1	Protein-coding mRNAs	AS_MAP3K15
Antisense transcripts from the DOWN200 gene set	Top10	1	MICT00000376640.1	175	PGK1	Protein-coding mRNAs	AS_TAF9B
		2	ENCT00000468735.1	145	PGK1	Protein-coding mRNAs	AS_TAF9B
		3	ENST00000229270.4	128	TPI1	Protein-coding mRNAs	AS_SPSB2
		4	FTMT29100015042.1	120	PGK1	Protein-coding mRNAs	AS_TAF9B
		5	ENST00000566326.1	114	MAP2K1	Protein-coding mRNAs	AS_SNAPC5
		6	ENST00000329078.3	99	SPNS2	Protein-coding mRNAs	AS_MYBBP1A
		7	ENST00000222256.4	87	RAB3A	Protein-coding mRNAs	AS_MPV17L2
		8	ENCT00000059259.1	76	PDZD7	Protein-coding mRNAs	AS_TWNK
		9	MICT00000139943.1	71	SPNS2	Protein-coding mRNAs	AS_MYBBP1A
		10	FTMT25300024476.1	70	ZBTB25	Protein-coding mRNAs	AS_MTHFD1
	Bottom10	1	ENCT00000264196.1	-200	IDH3B	protein-coding mRNAs	AS_NOP56

		2	FTMT30000000001.1	-193	MT-ND5	Protein-coding mRNAs	AS_MT-ND6
		3	FTMT20300078215.1	-188	APOA1BP	Protein-coding mRNAs	AS_GPATCH4
		4	FTMT24100001922.1	-187	H2AFX	Protein-coding mRNAs	AS_HMBS
		5	MICT00000370386.1	-185	MT-TA	Structural RNAs	AS_MT-ATP6
		6	MICT00000370386.1	-185	MT-TA	Structural RNAs	AS_MT-ATP8
		7	MICT00000370386.1	-185	MT-TA	Structural RNAs	AS_MT-CO3
		8	MICT00000370386.1	-185	MT-TA	Structural RNAs	AS_MT-ND1
		9	MICT00000370386.1	-185	MT-TA	Structural RNAs	AS_MT-ND2
		10	MICT00000370386.1	-185	MT-TA	Structural RNAs	AS_MT-ND3

508

509

Human RNA-seq

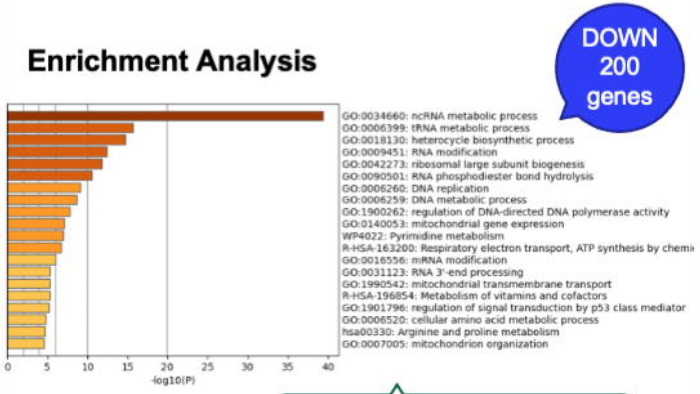
Hypoxia-related datasets

© 2016 DBCLS TogoTV / CC-BY-4.0

Protein coding genes only

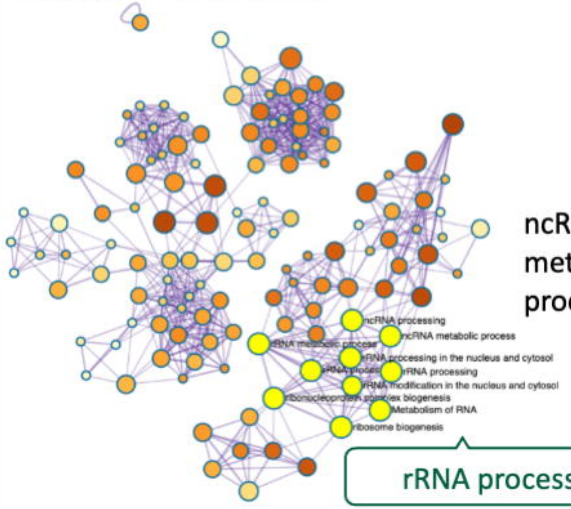
Reference transcript : GENCODE release 30

Enrichment Analysis



ncRNA metabolic process

Down200 GO Network



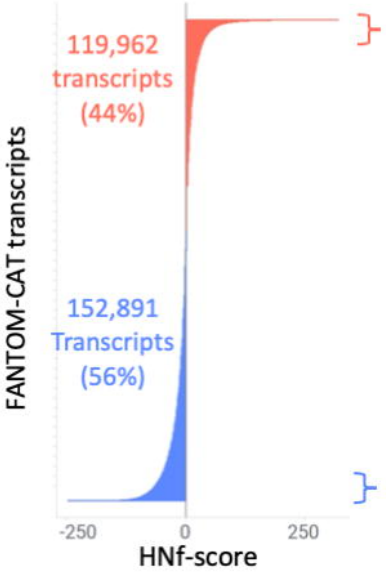
ncRNA metabolic process

rRNA processing

Protein coding and non-coding transcripts

Reference transcript : FANTOM-CAT

Hypoxia-Normoxia (HNf)-score of all transcripts



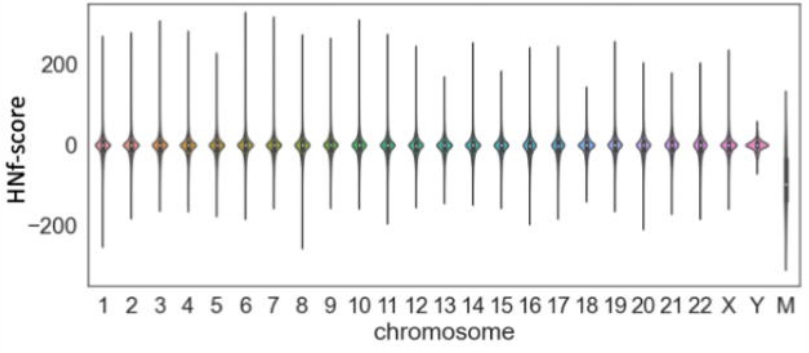
Gene name	HNf-score
BHLHE40-**AS1**	300
SLC2A1-**AS1**	263
RAB11B-**AS1**	242

Antisense

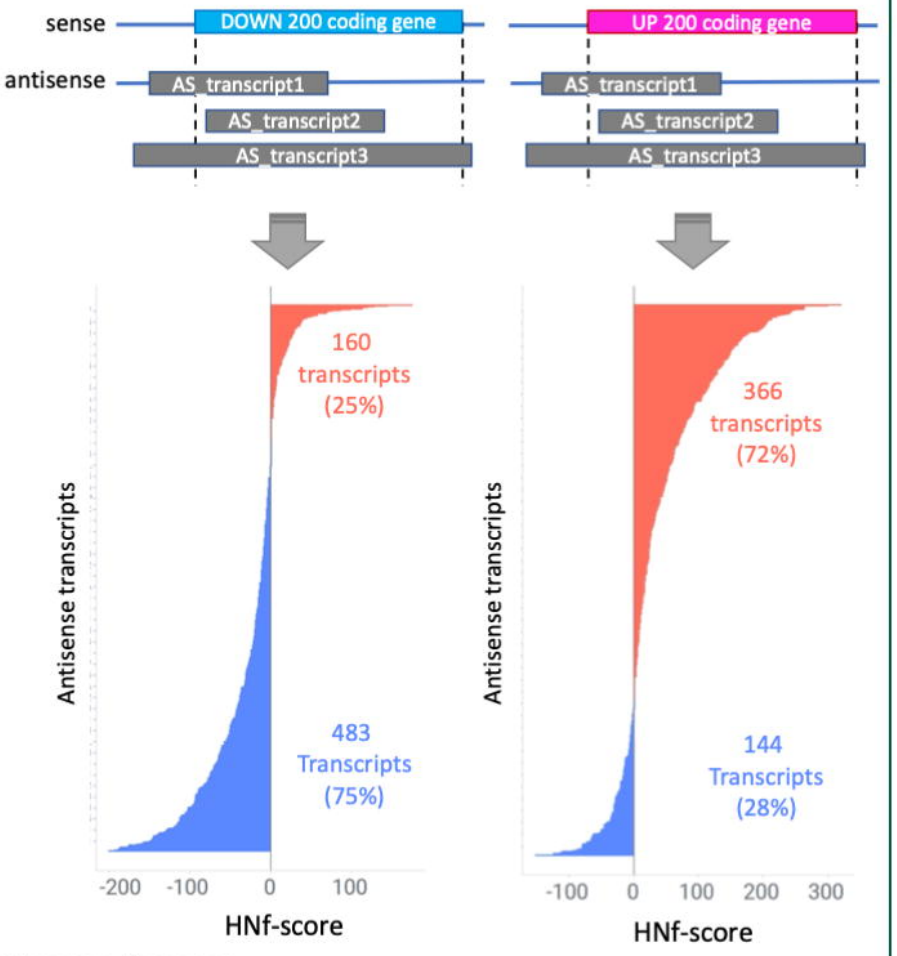
Gene name	HNf-score
MT-ND2	-245
MT-ATP8	-230
MT-ND3	-221

Mitochondria

Visualization of HN-score per chromosome

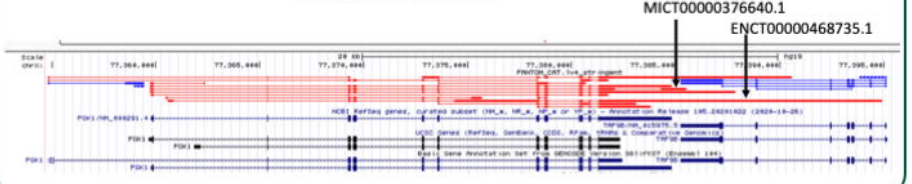


Survey of HNf-scores for sense and antisense pairs

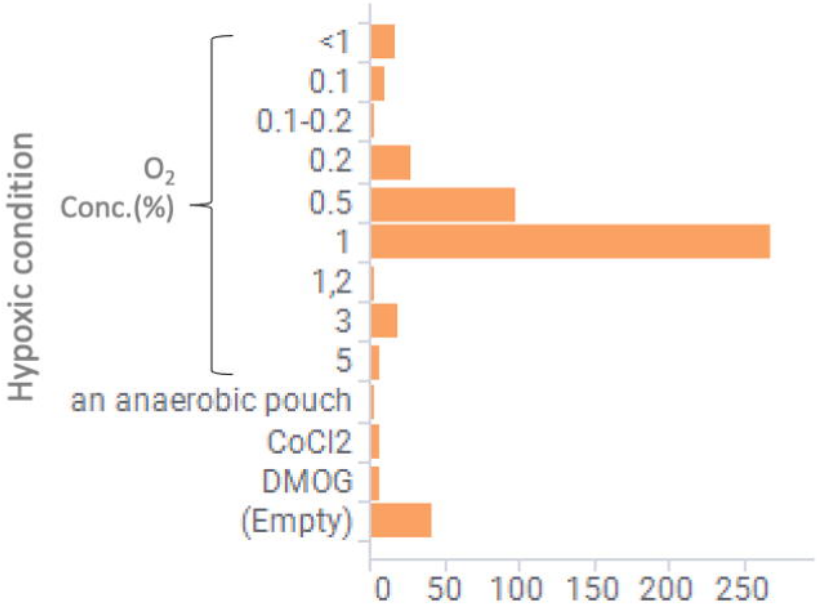


Genome Browser

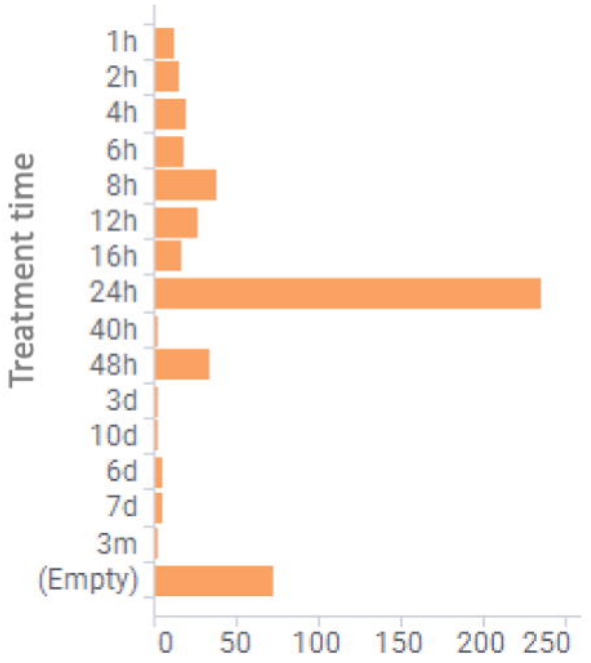
PGK1 and TAF9B



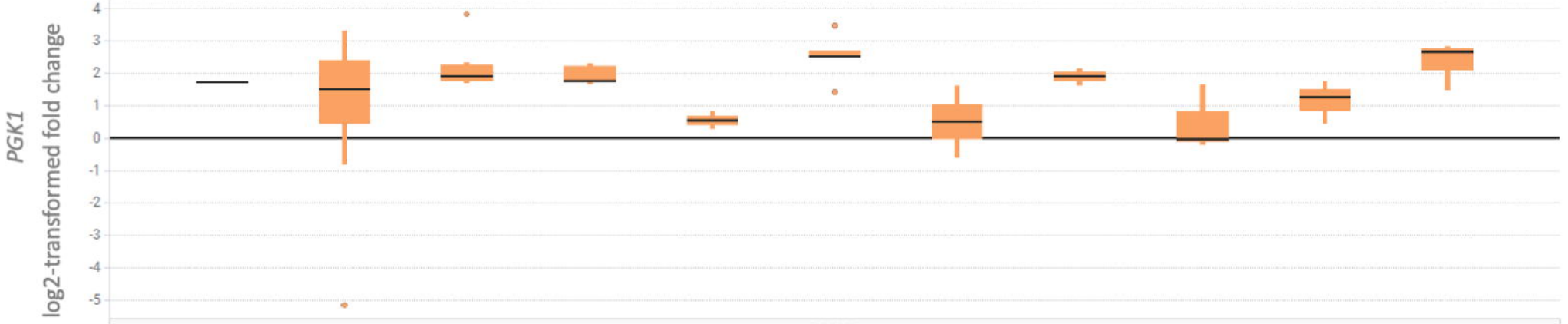
A



B

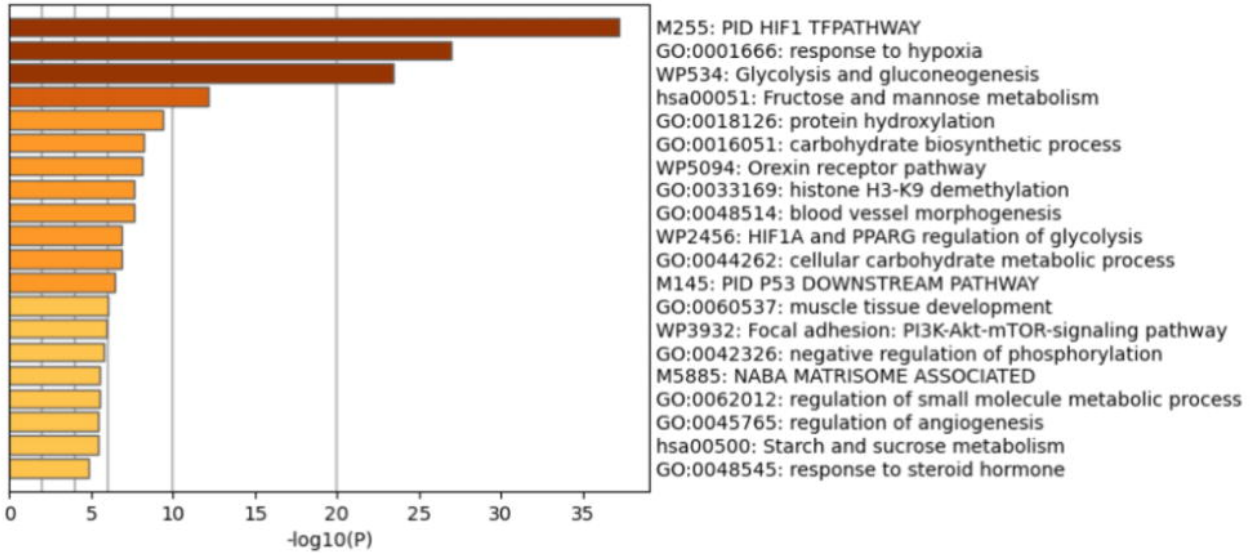


C

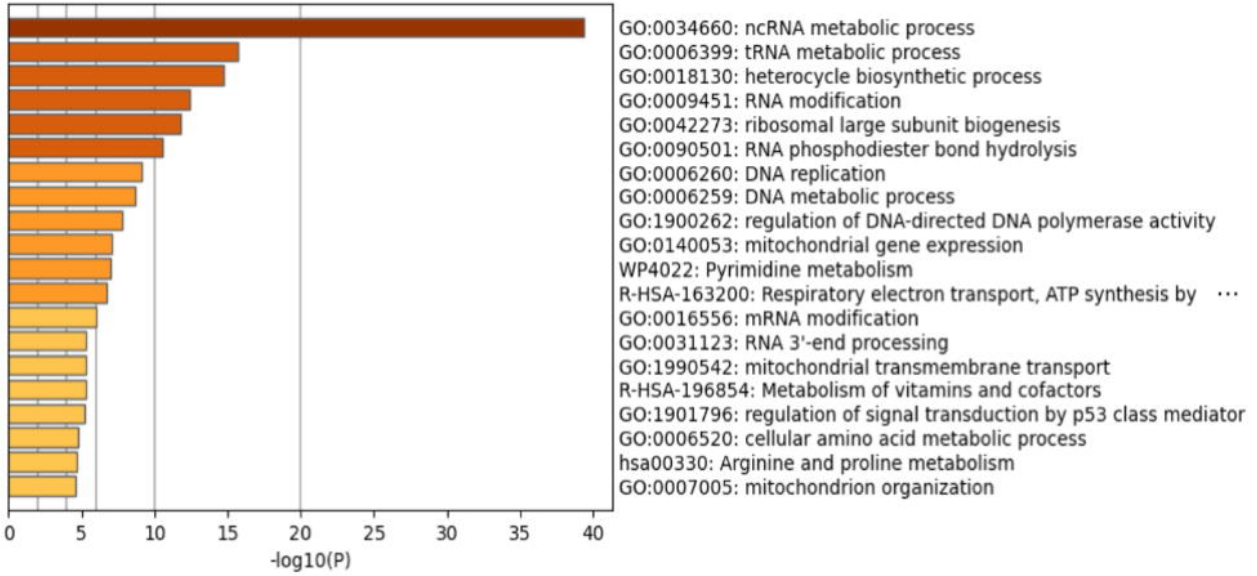


	cancer										
	brain	breast	colon	fibrosarcoma	kidney	liposarcoma	pancreas	placenta	synovial sarcoma	uterus	neurofibromatosis type 1
Count	3	38	14	5	2	5	2	2	3	3	3
Median	1.69913	1.49027	1.89857	1.74935	0.538606	2.51452	0.49413	1.88443	-0.0462301	1.23014	2.65288
Outliers	0	1	1	0	0	2	0	0	0	0	0

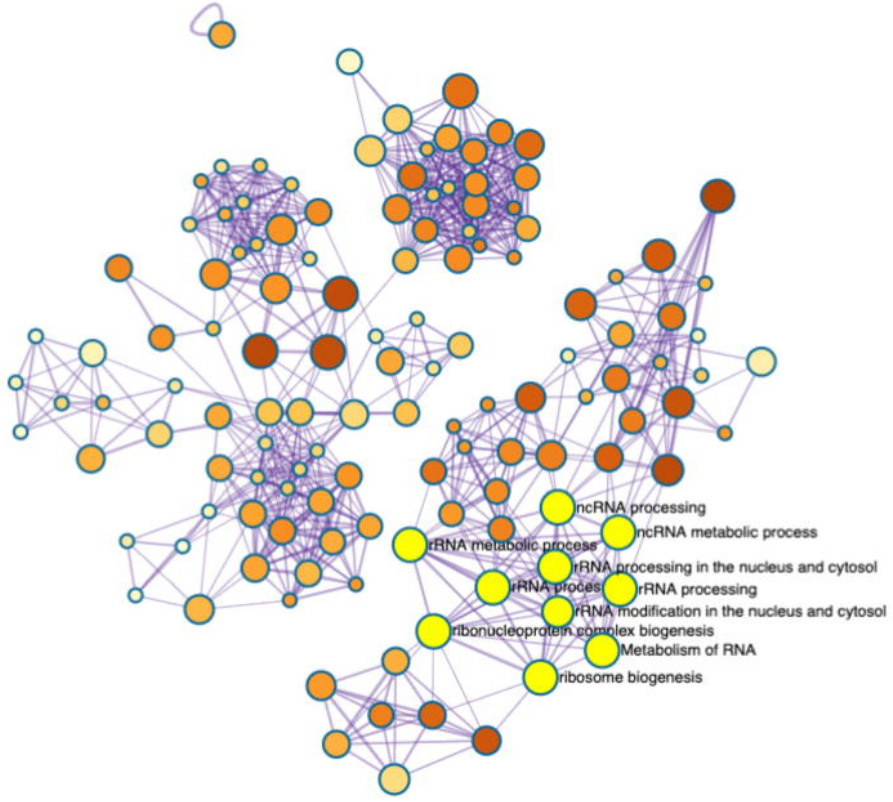
A



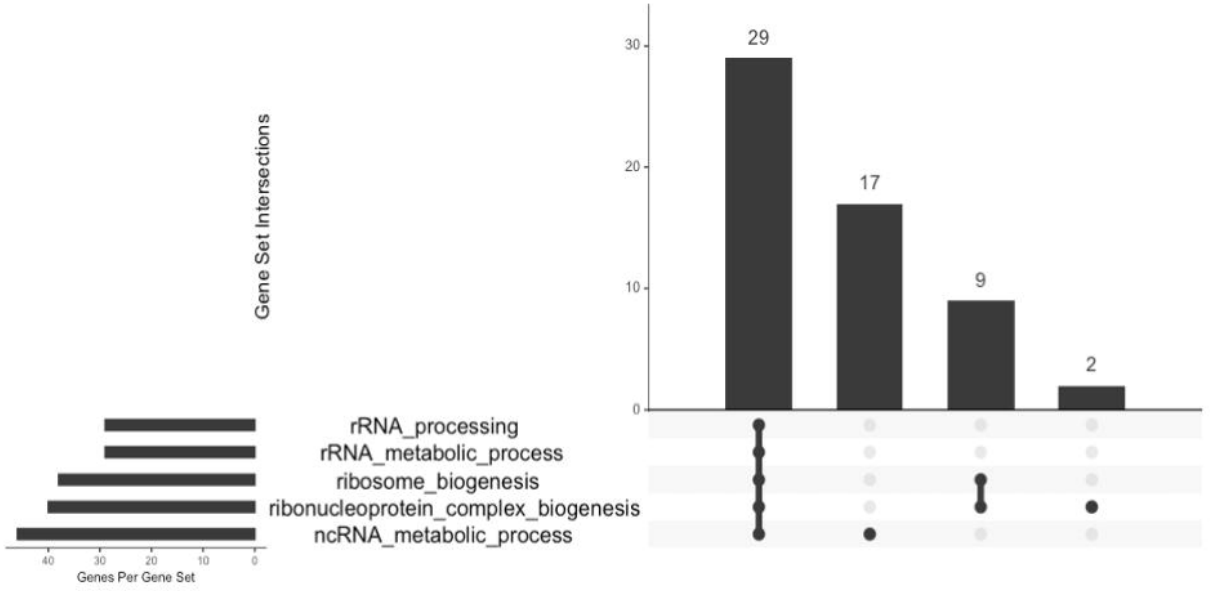
B



C

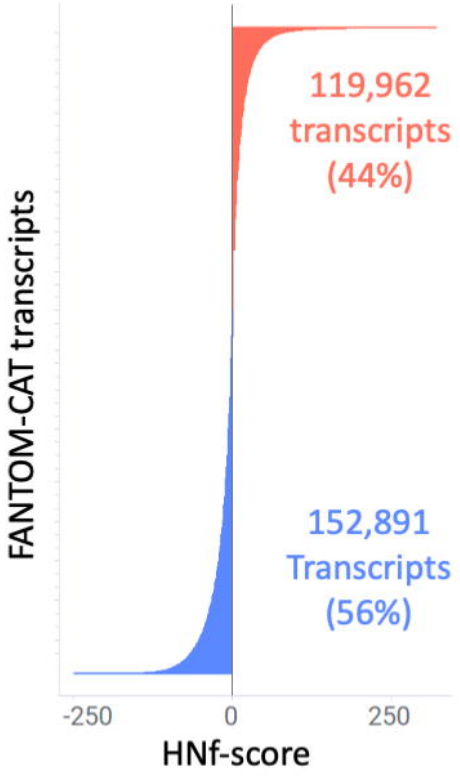


D

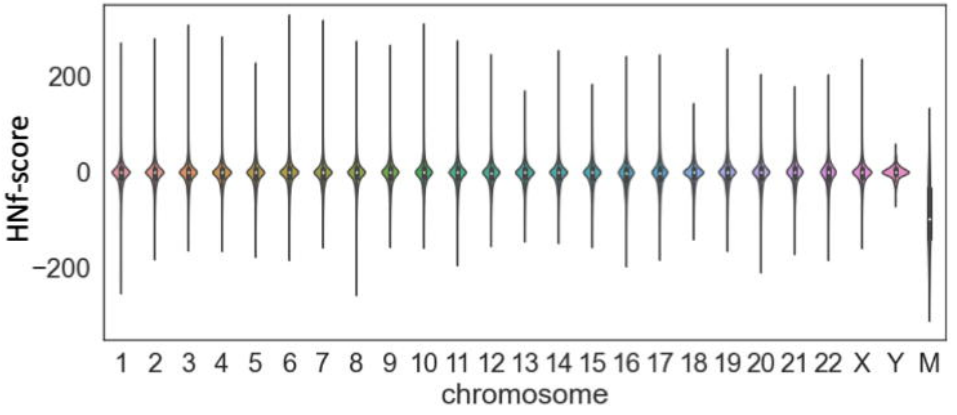


GO term	Genes
ncRNA metabolic process	DKC1|METTL1|NOP2|RRP9|DDX21|NOLC1|WDR46|LCMT2|FARSB|RCL1|NPM3|NOP56|WDR4|RPP40|UTP14A|POP1|RRS1|ADAT1|DIMT1|UTP20|RRP7A|RRP15|MRTO4|POP5|POLR3K|EXOSC4|BRIX1|LYAR|ELAC2|DUS1L|NOL6|NOC4L|METTL8|MRM1|PUS1|UTP15|MAK16|RRP36|FDXACB1|IMP4|MARS2|DTD2|KTI12|TOE1|TRMT61A|CTU2
rRNA processing	DKC1|NOP2|RRP9|DDX21|WDR46|RCL1|NOP56|RPP40|UTP14A|DIMT1|UTP20|RRP7A|NIP7|EXOSC4|PNO1|ELAC2|NOL6|NOC4L|NOL12|MRM1|UTP15|LTV1|RRP36|IMP4

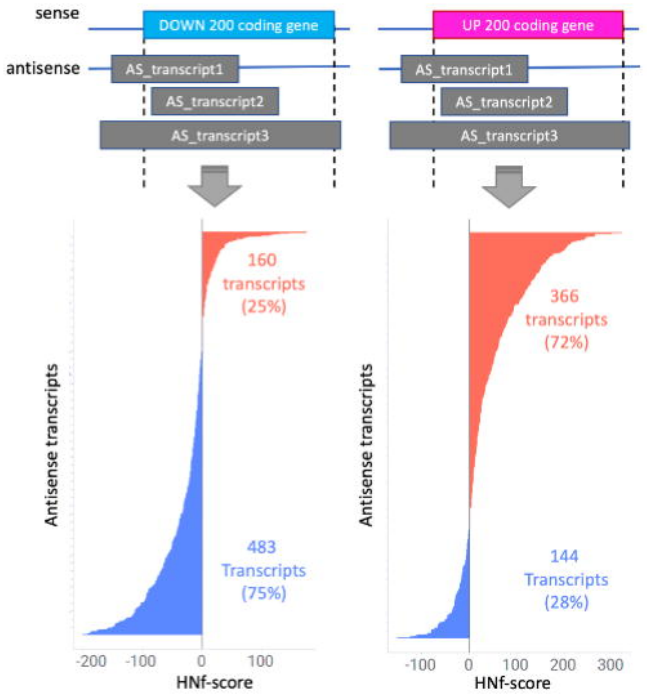
A


119,962
transcripts
(44%)
152,891
Transcripts
(56%)
FANTOM-CAT transcripts
-250
0
250
HNf-score

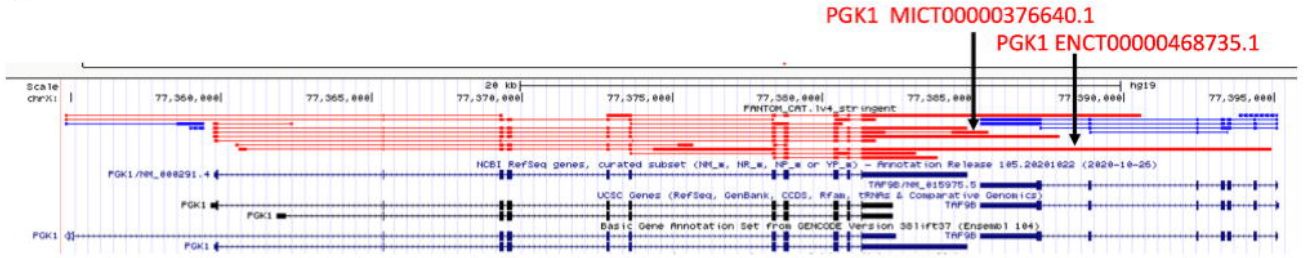
B


200
0
−200
HNf-score
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y M
chromosome

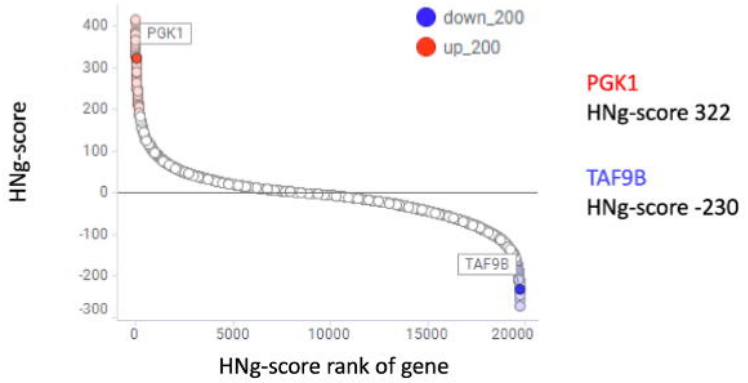
A



B



C



PGK1
HNg-score 322

TAF9B
HNg-score -230