

Cell Genomics, Volume 5

Supplemental information

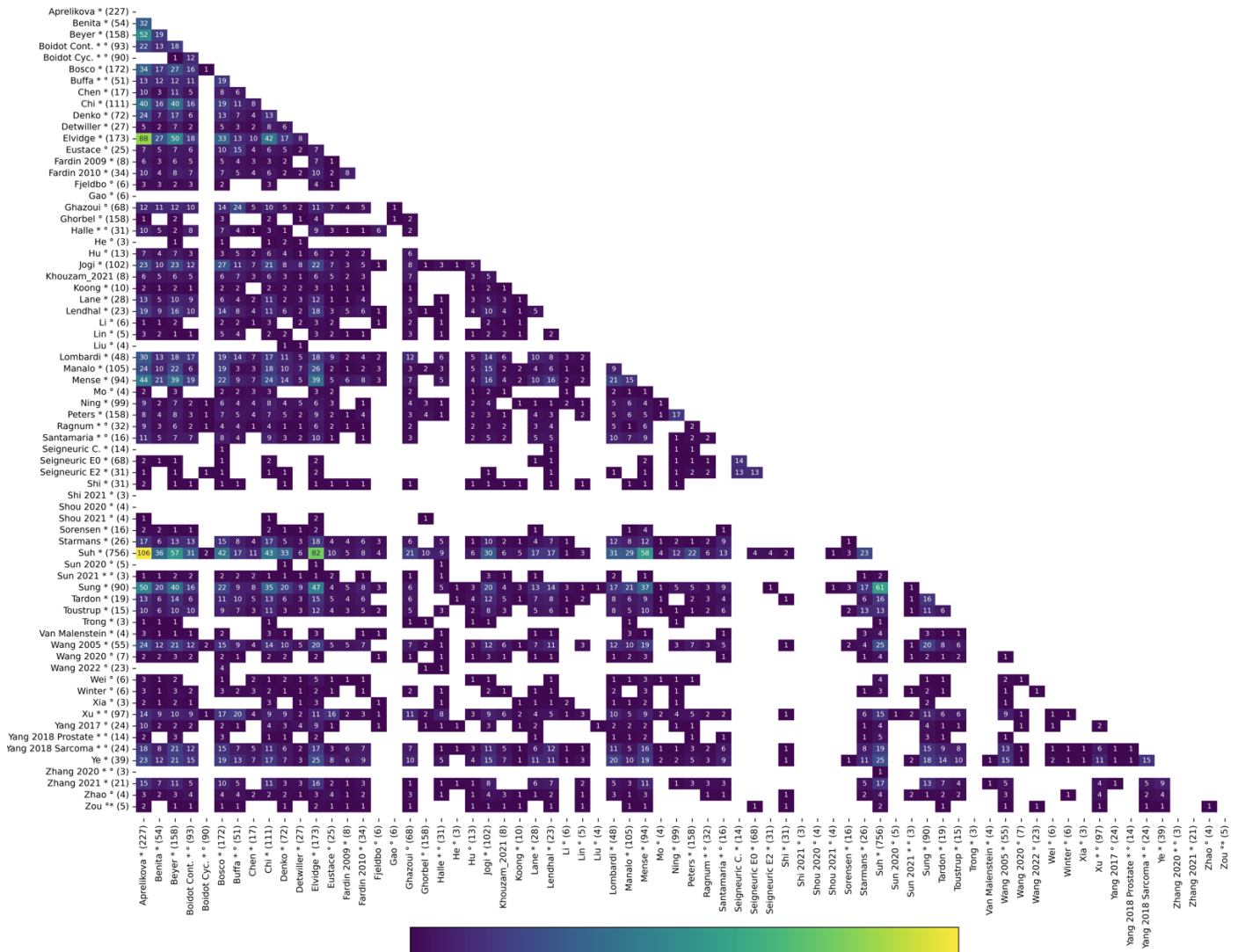
**Defining hypoxia in cancer: A landmark evaluation
of hypoxia gene expression signatures**

Matteo Di Giovannantonio, Fiona Hartley, Badran Elshenawy, Alessandro Barberis, Dan Hudson, Hana S. Shafique, Vincent E.S. Allott, David A. Harris, Simon R. Lord, Syed Haider, Adrian L. Harris, Francesca M. Buffa, and Benjamin H.L. Harris

Supplemental information

Defining hypoxia in cancer: a landmark evaluation of hypoxia gene expression signatures

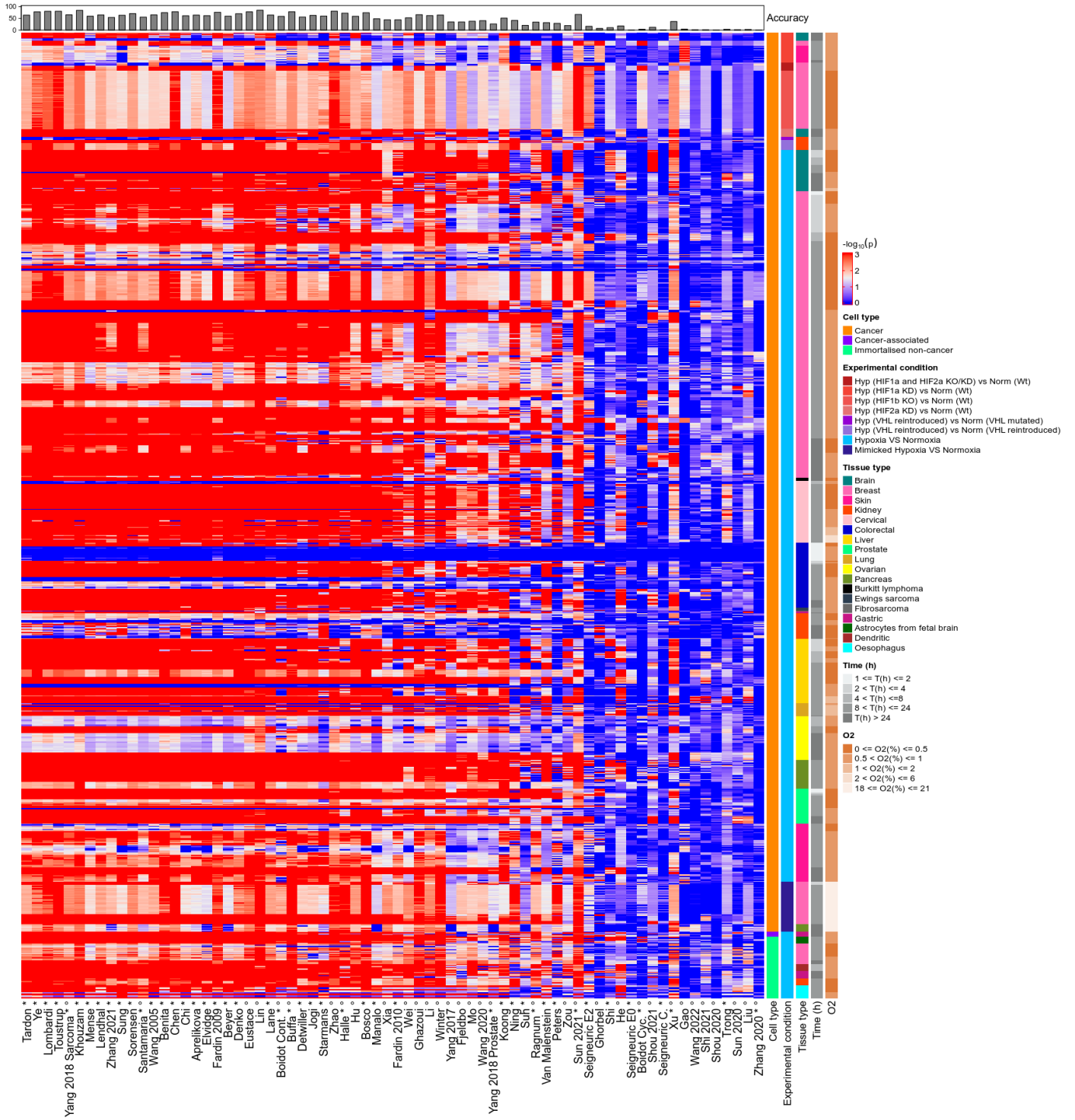
Matteo Di Giovannantonio, Fiona Hartley, Badran Elshenawy, Alessandro Barberis, Dan Hudson, Hana Shafique, Vincent E.S Allott, David A. Harris, Simon R. Lord, Syed Haider, Adrian L. Harris, Francesca M. Buffa, Benjamin H.L Harris



Yang 2018 Prostate * (14)
 Yang 2018 Sarcoma * (24)
 Ye * (39)
 Zhang 2020 * (3)
 Zhang 2021 * (21)
 Zhao * (4)
 Zou * (5)

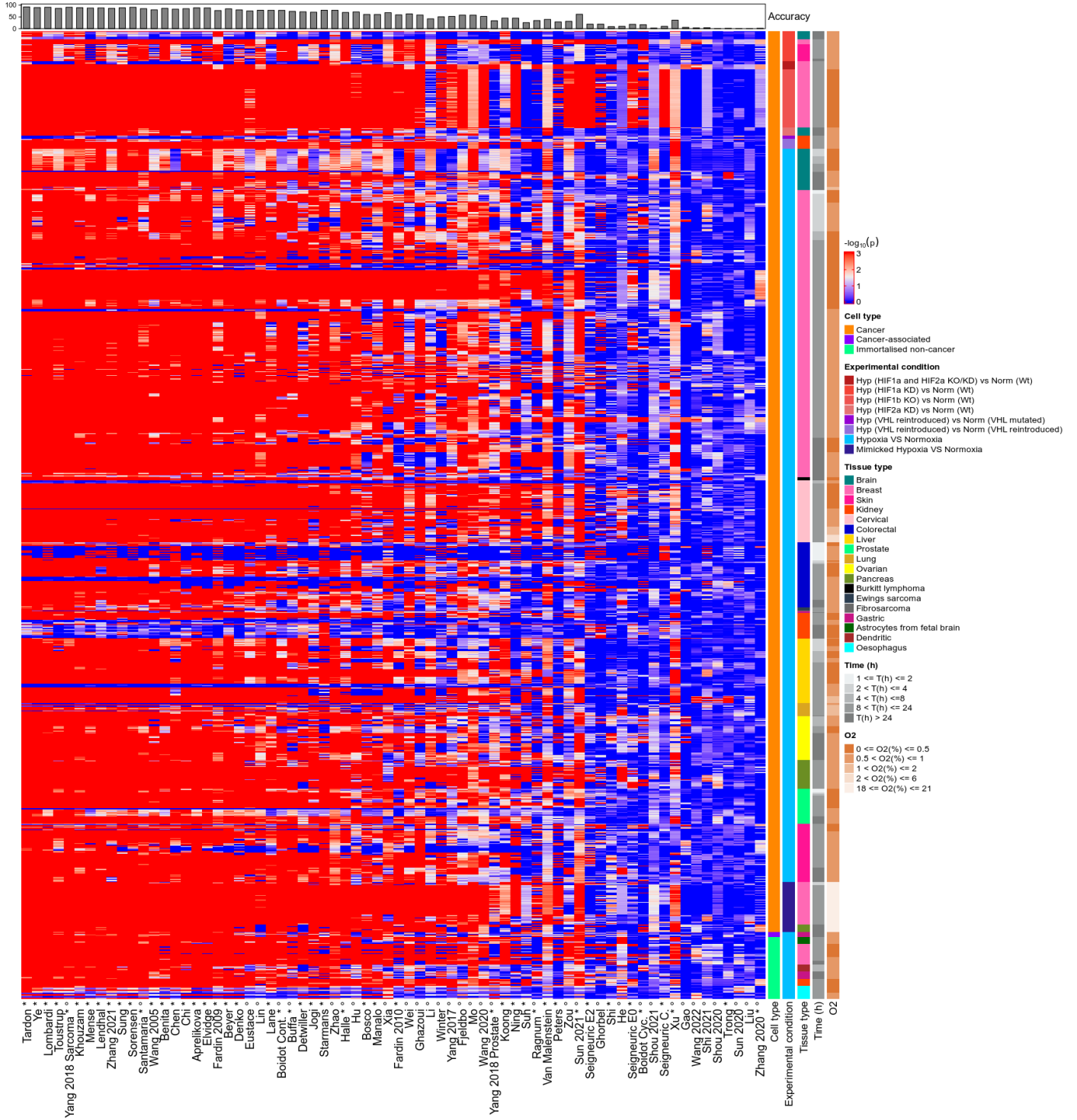
Supplementary Figure S1: The number of overlapping genes between signatures

Heatmap of gene overlap across investigated signatures. The colour intensity indicates the number of shared genes between each pair of signatures. Signatures with no overlapping genes are represented in white. The numbers in parentheses following each signature name report the final gene count for that signature after quality check, related to Figure 1.



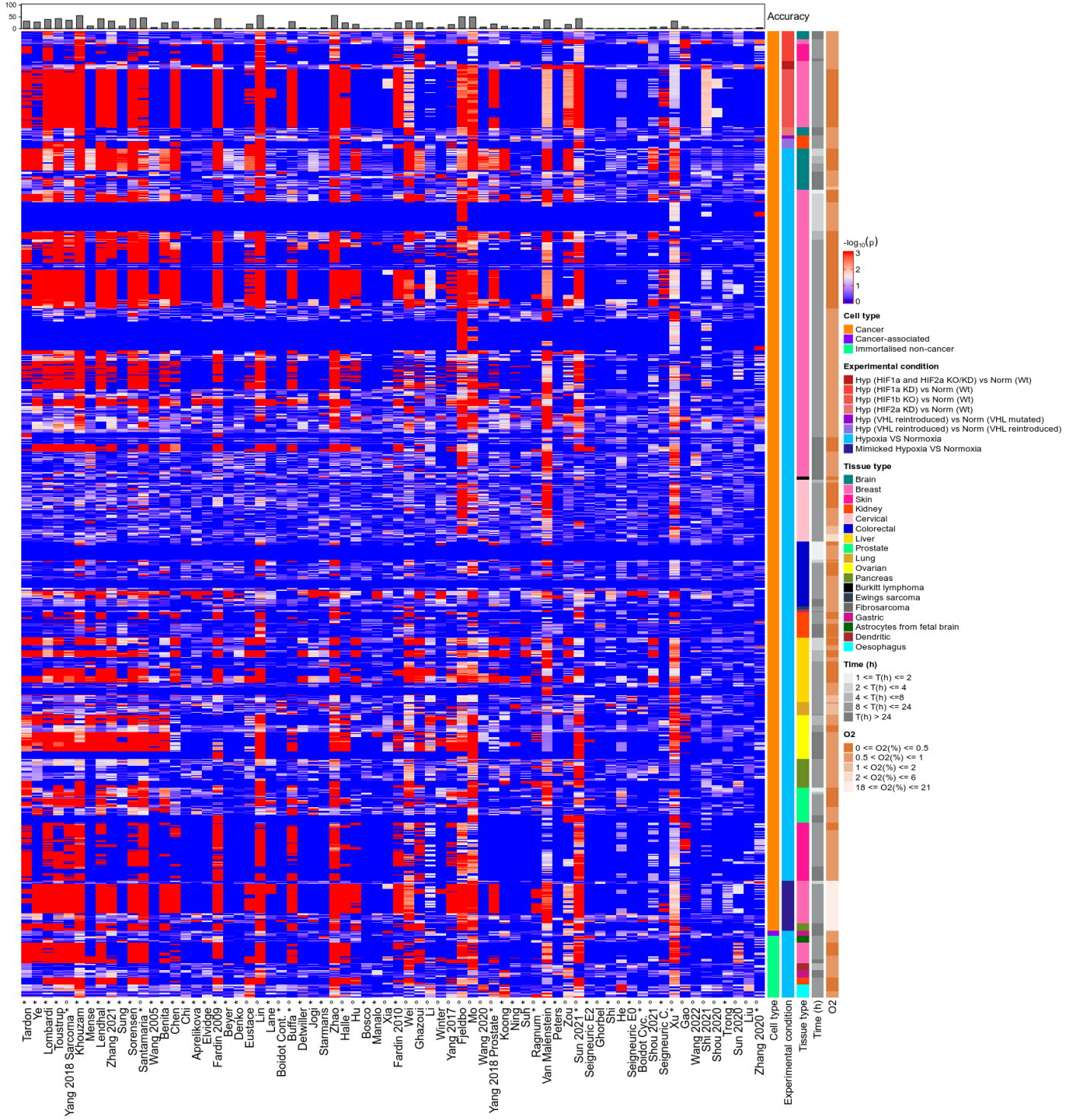
Supplementary Figure S2: Complex heatmap (mean)

Complex heatmaps of signatures p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data. Shades of red and blue denote low and high p-values, respectively, as derived from permutation tests and represent the likelihood that a random gene signature would outperform a hypoxia signature. The top bar plot summarises signature accuracy, calculated as the number of pairwise combinations with p-values < 0.05 over the total number of pairwise combinations. The right-hand legend provides additional sample features, with titles located at the beginning of the x-axis, related to Figure 3.



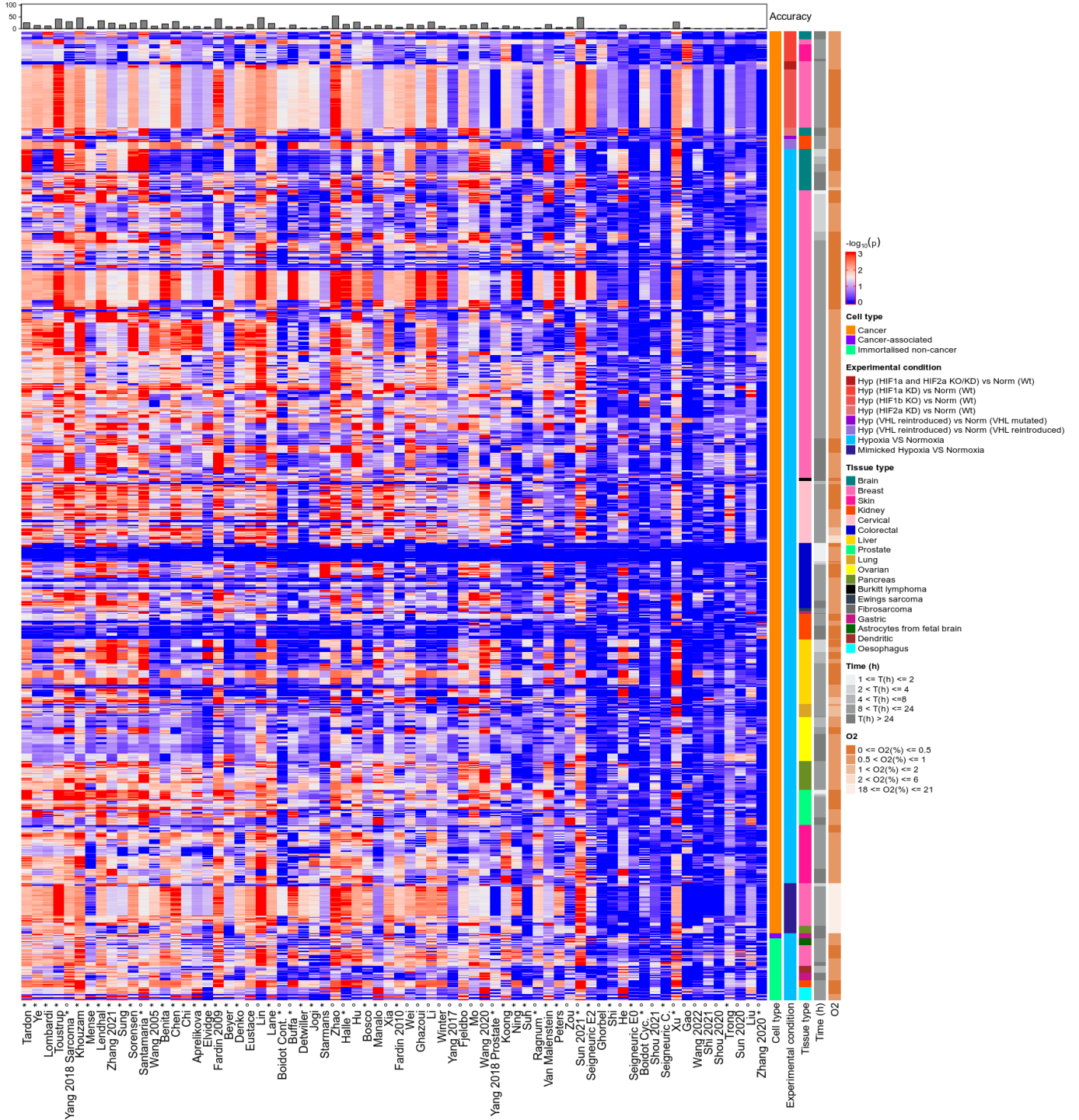
Supplementary Figure S3: Complex heatmap (median)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



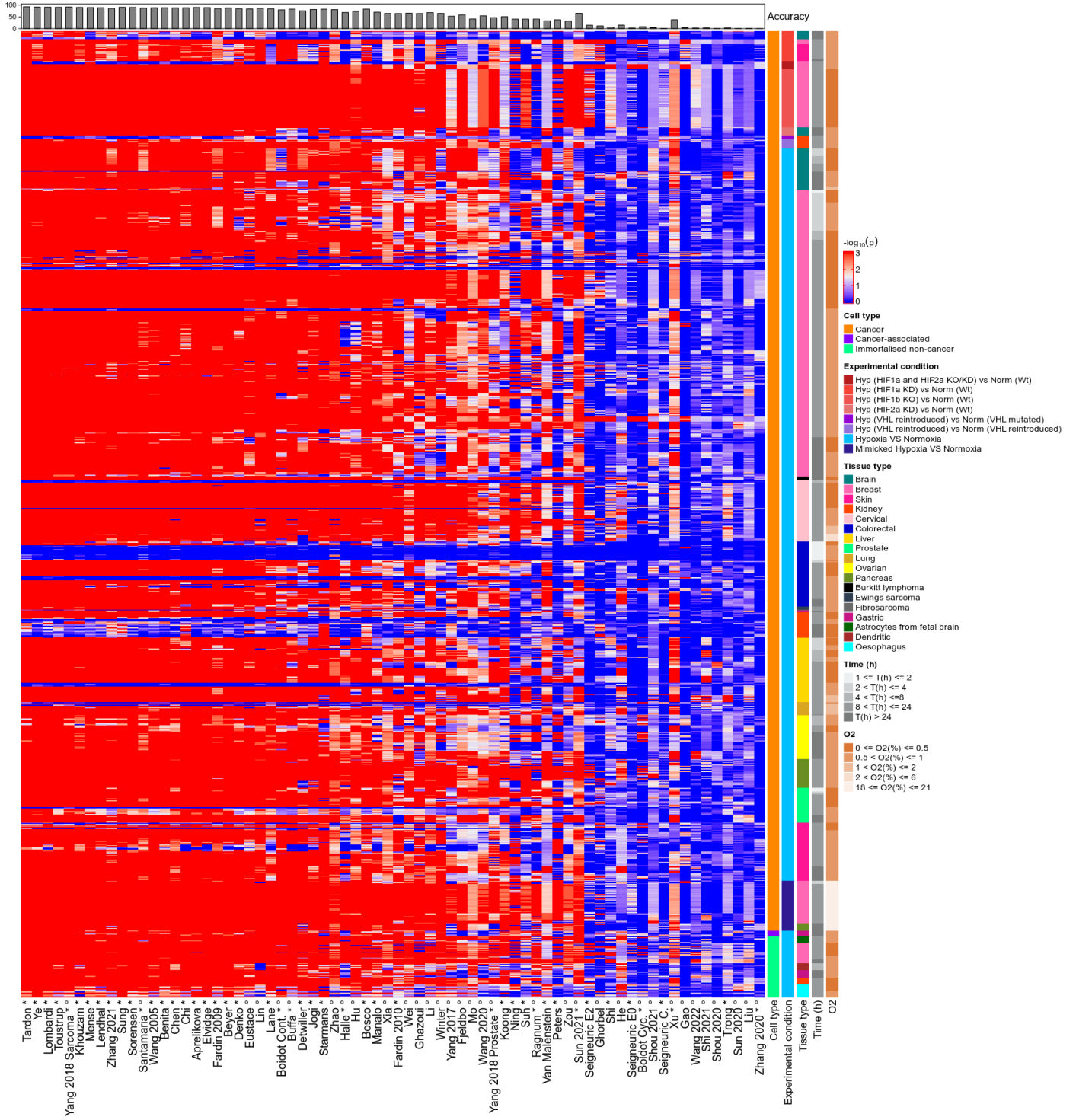
Supplementary Figure S4: Complex heatmap (mode)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



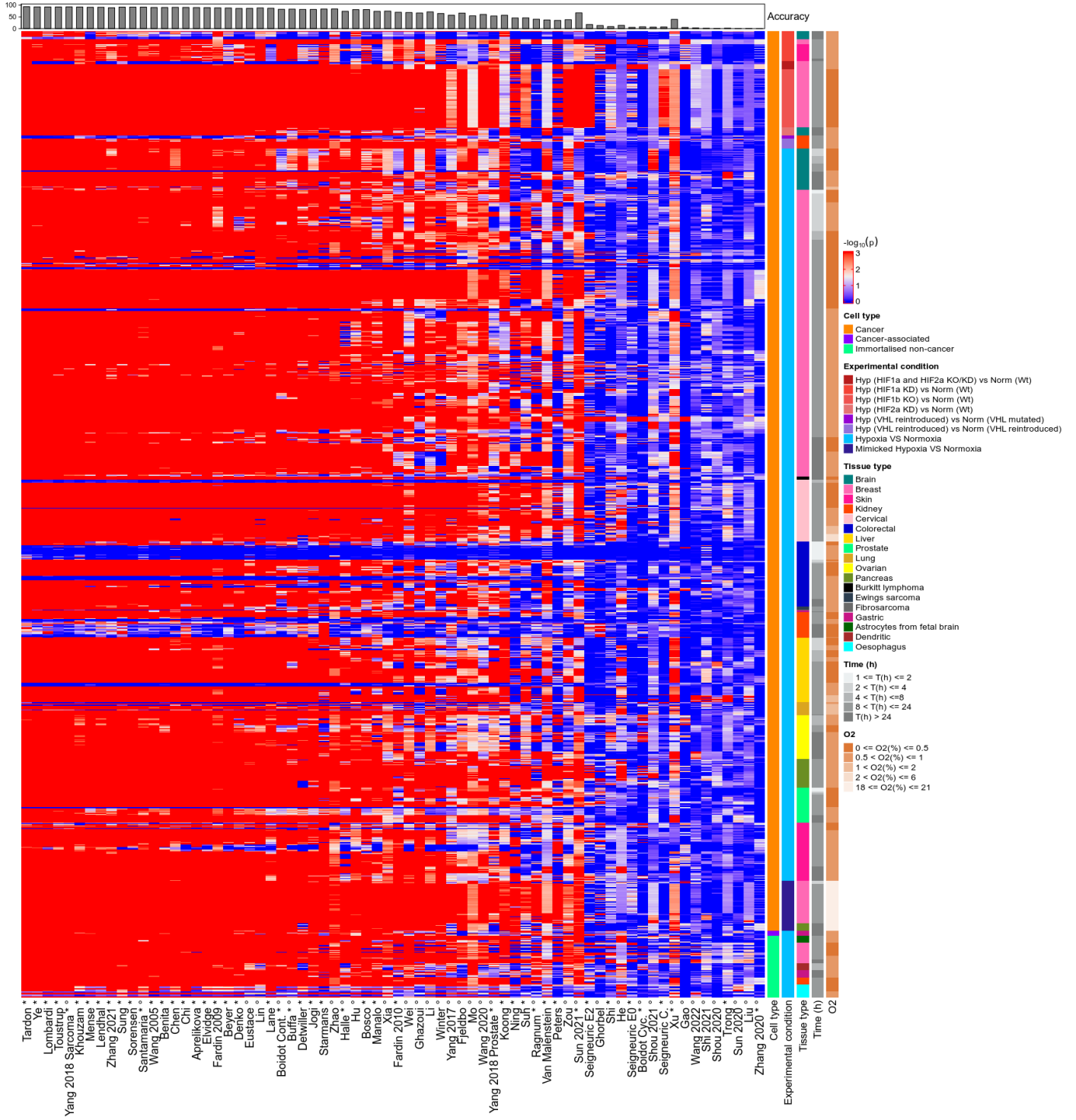
Supplementary Figure S5: Complex heatmap (midrange)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



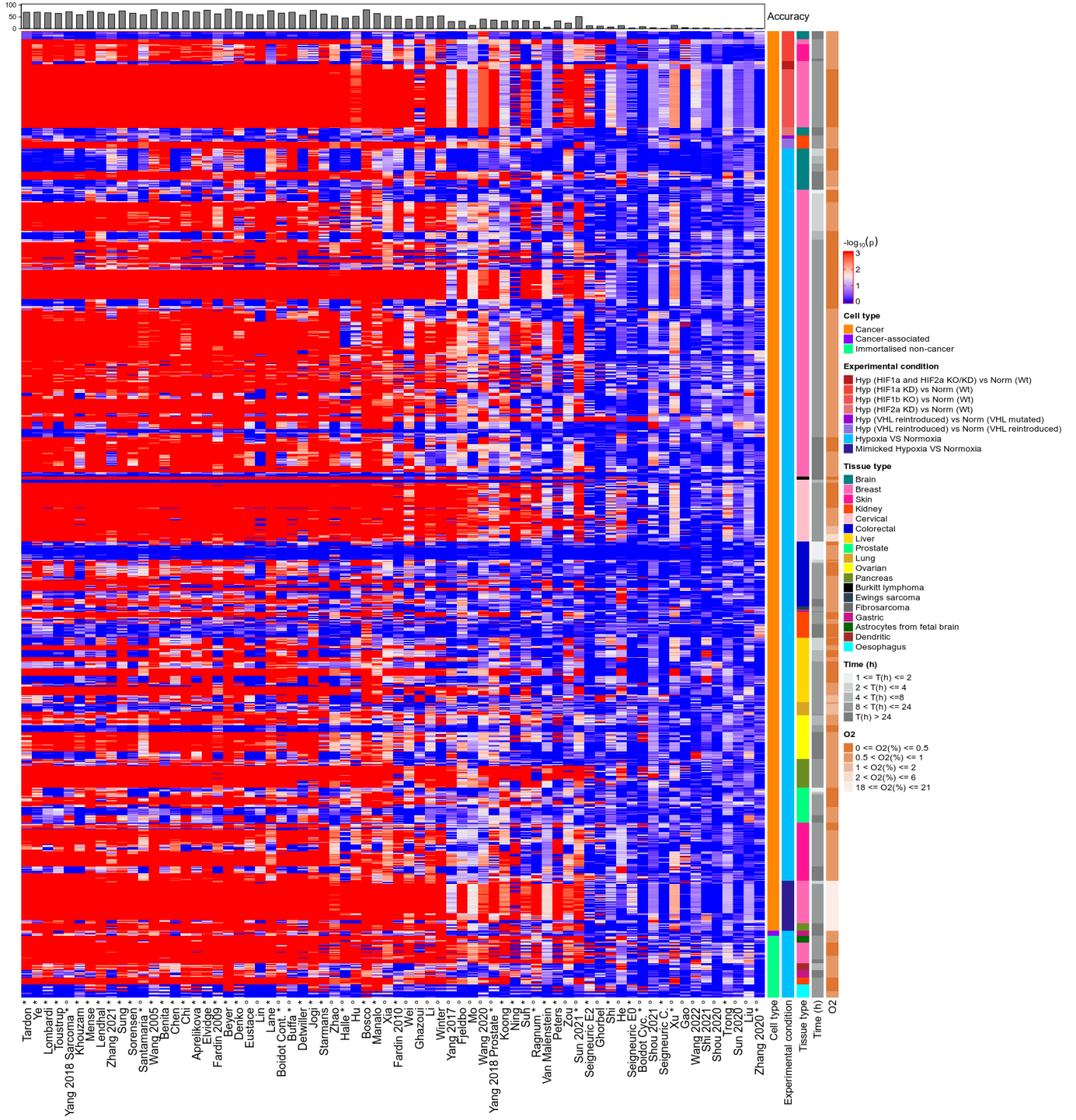
Supplementary Figure S6: Complex heatmap (midhinge)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



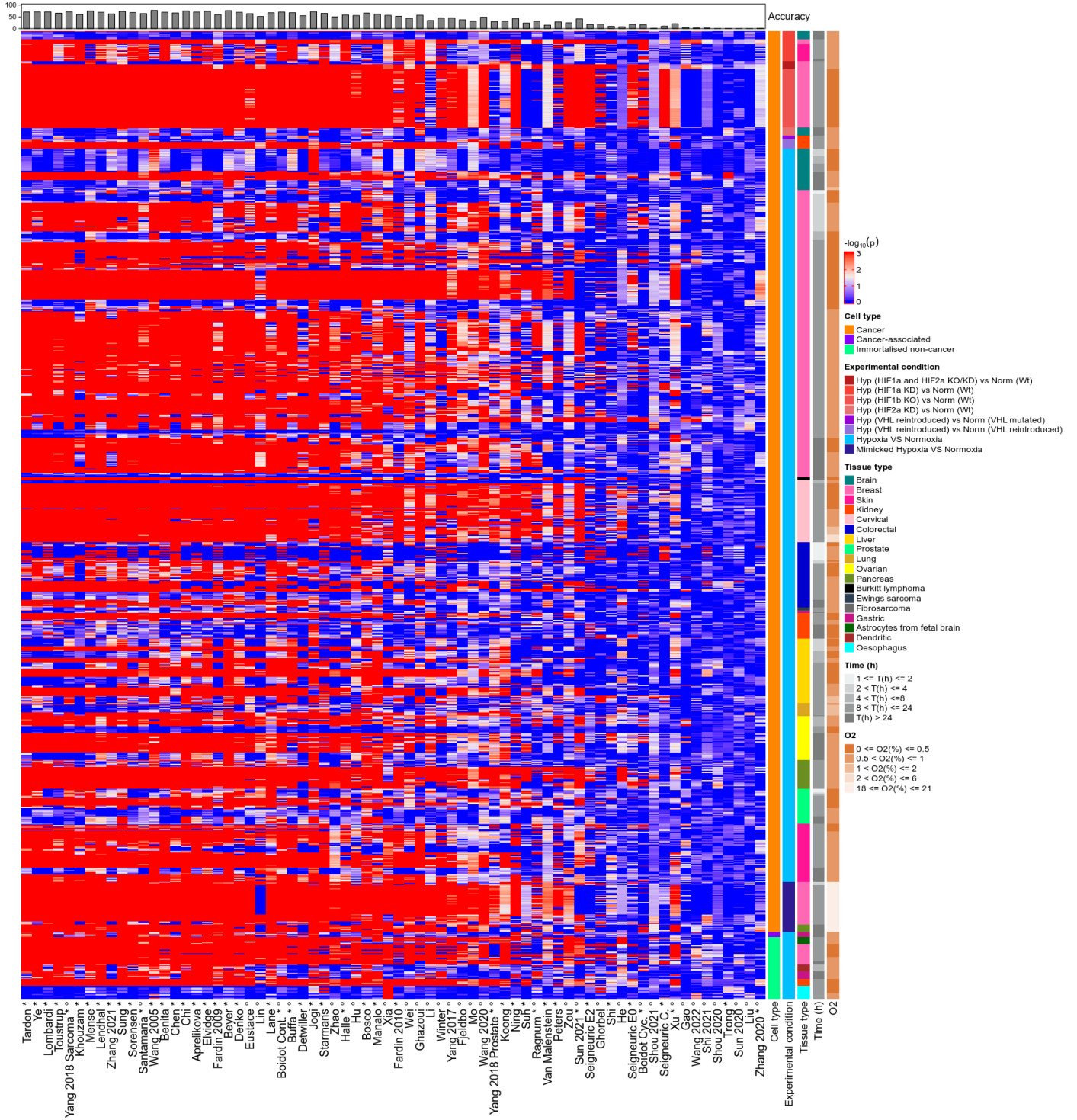
Supplementary Figure S7: Complex heatmap (trimean)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



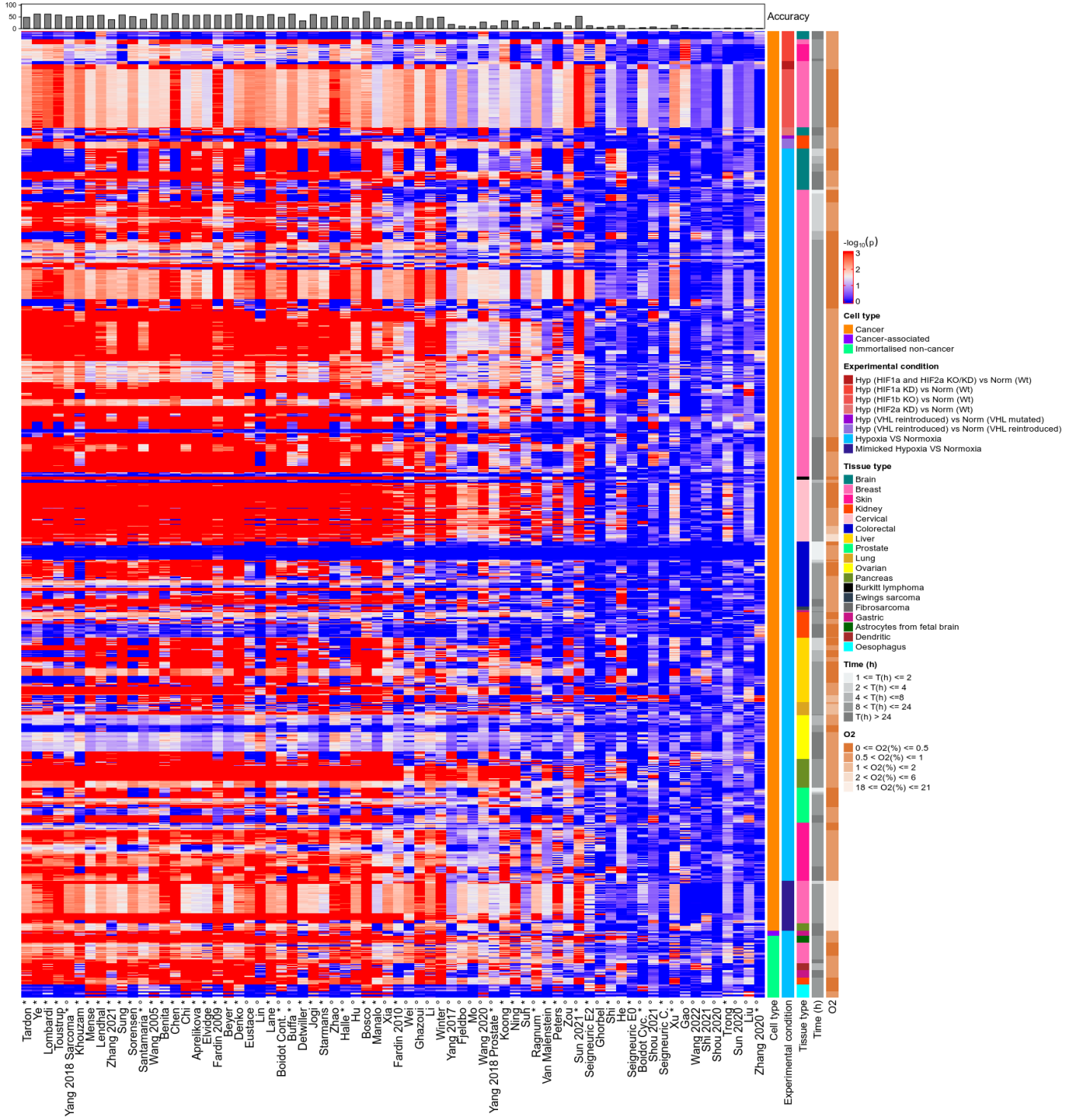
Supplementary Figure S8: Complex heatmap (IQR)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



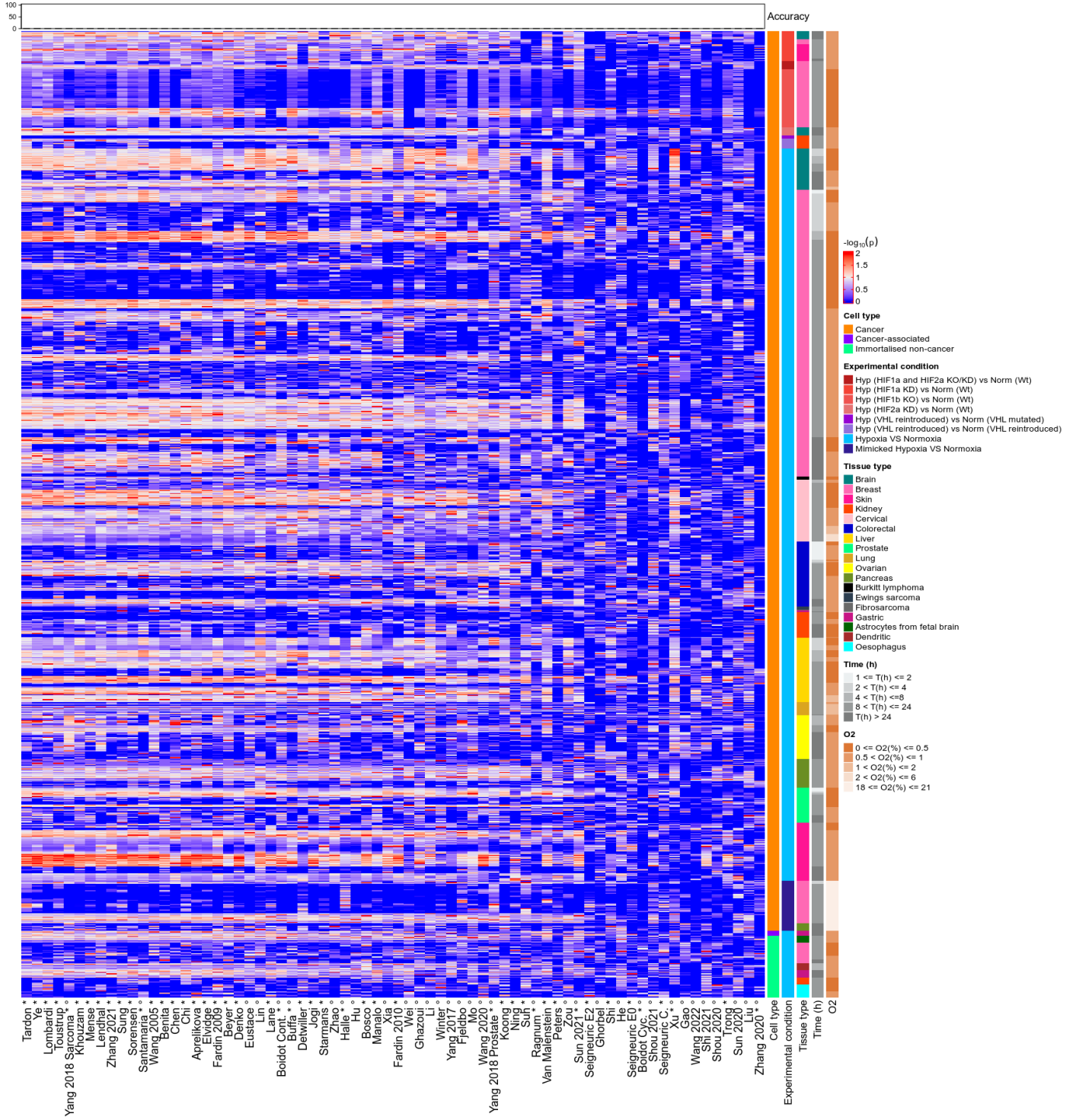
Supplementary Figure S9: Complex heatmap (MAD)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



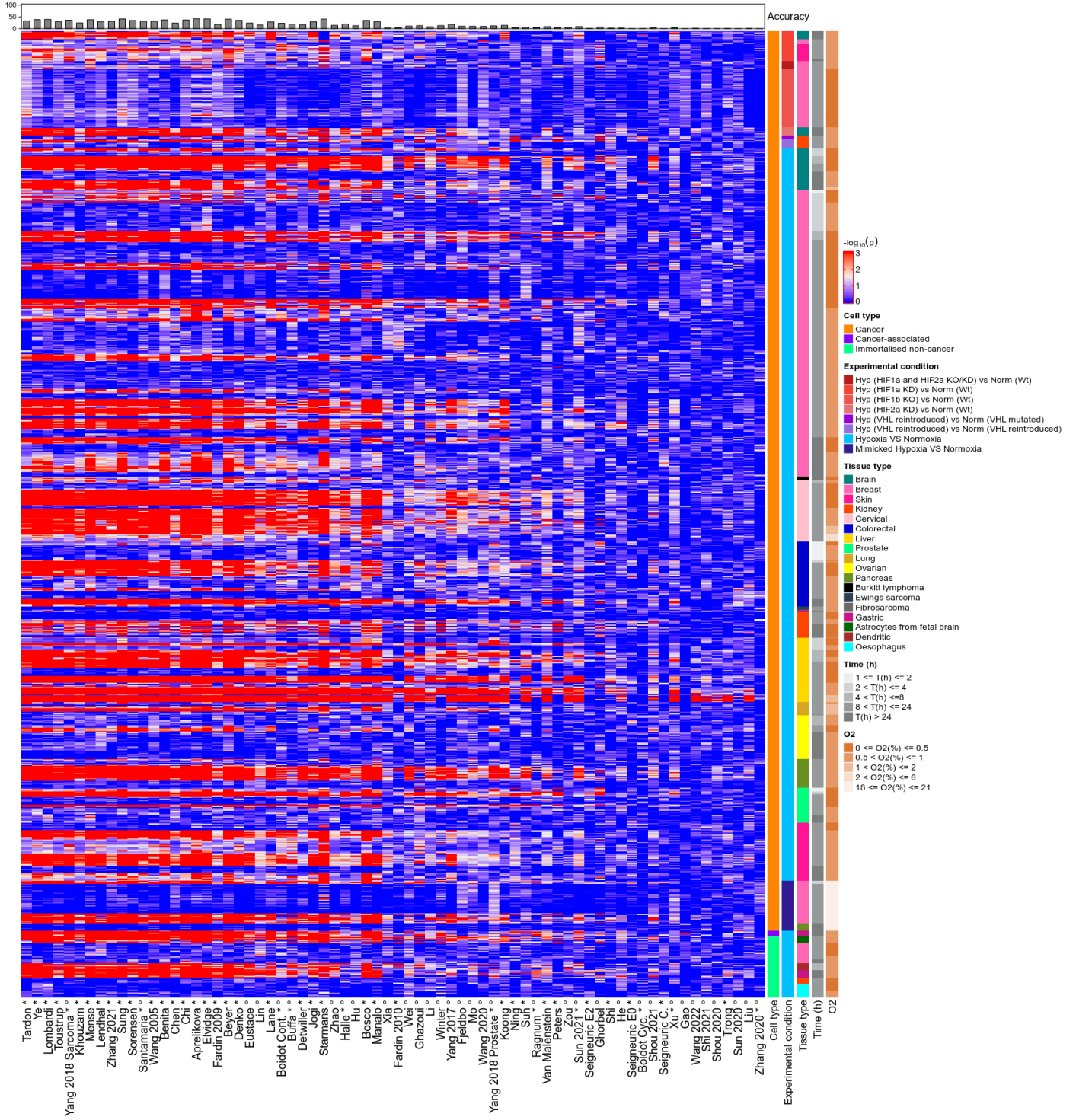
Supplementary Figure S10: Complex heatmap (AAD)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



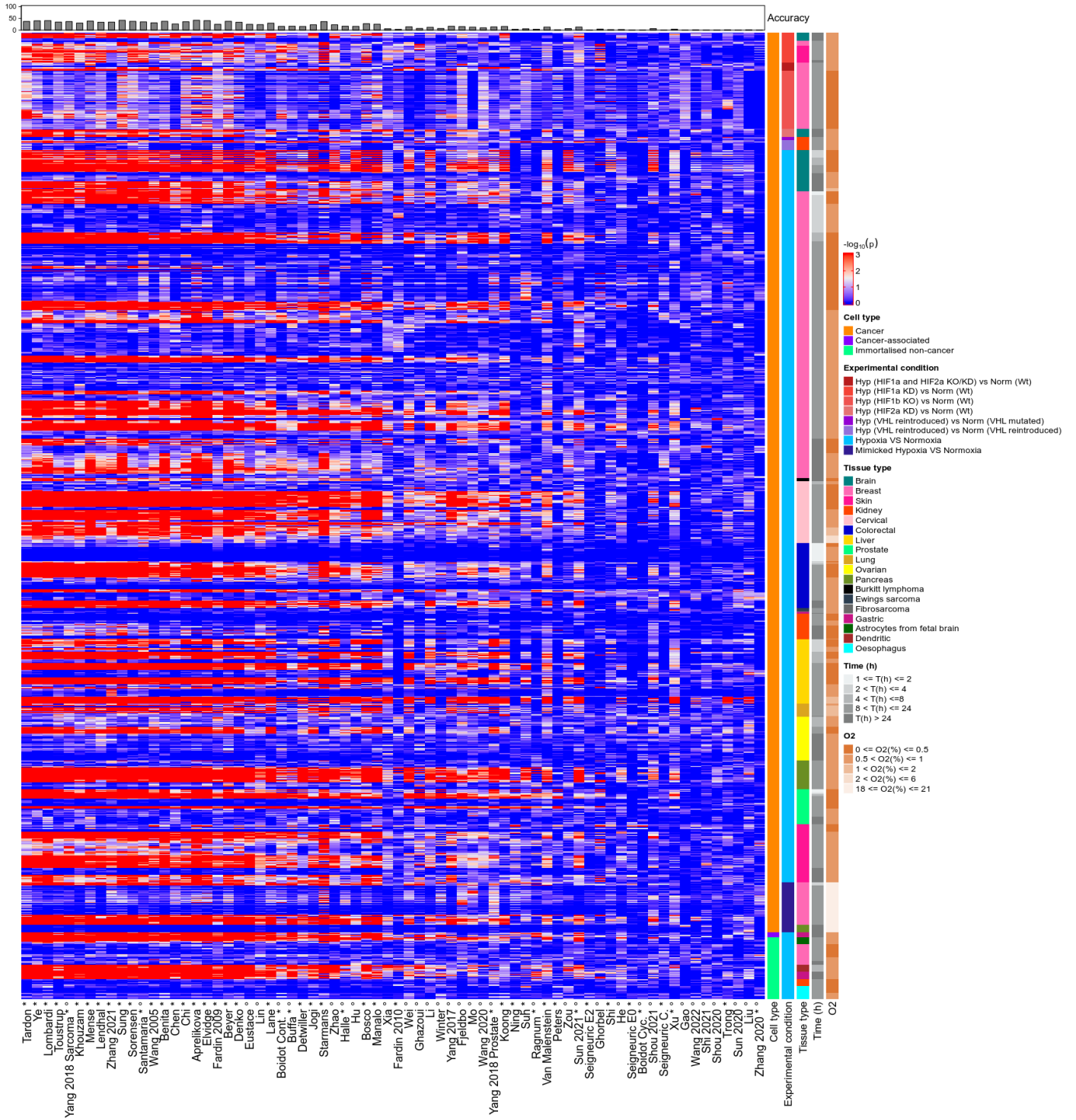
Supplementary Figure S11: Complex heatmap (ssGSEA)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



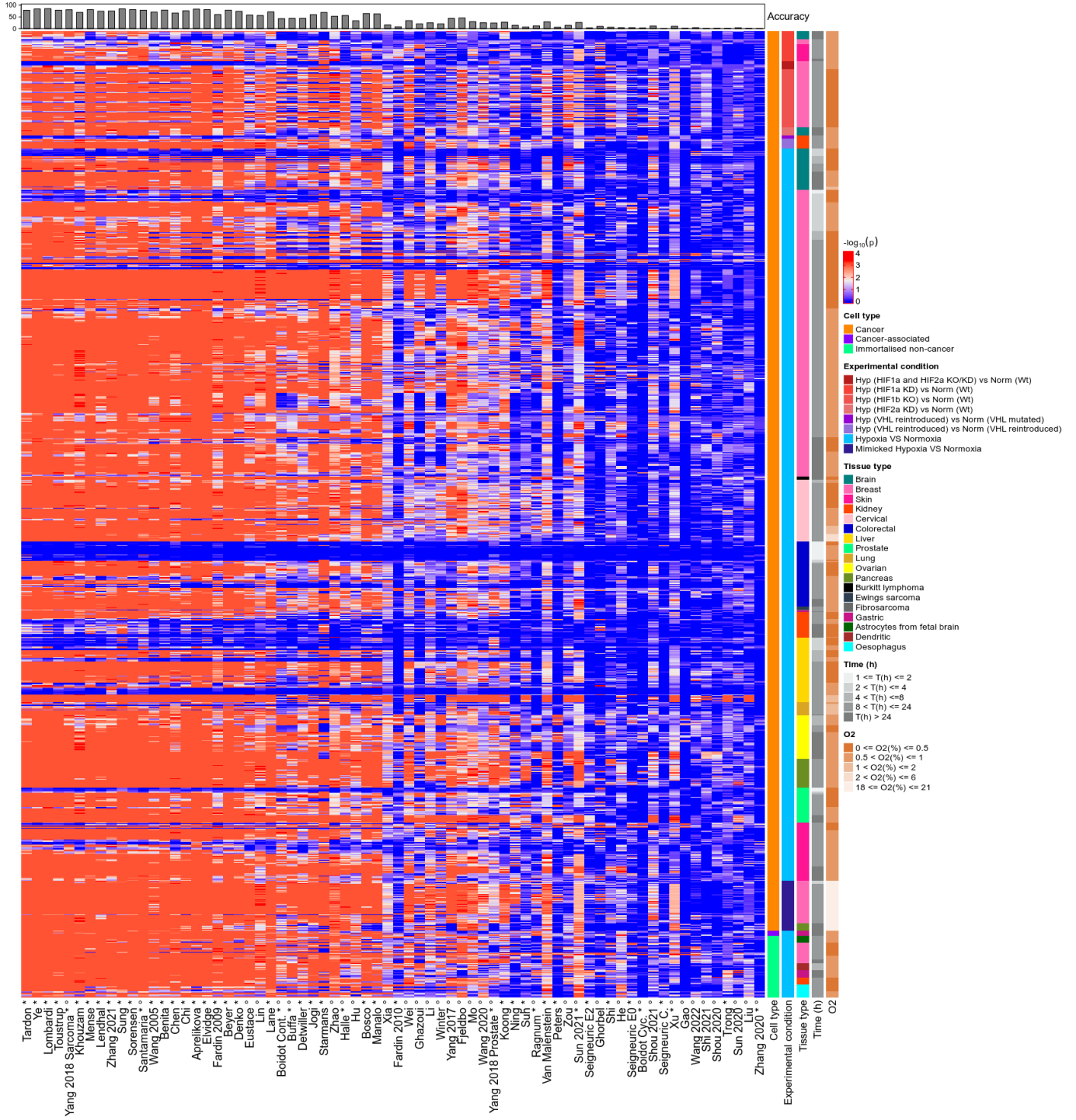
Supplementary Figure S12: Complex heatmap (GSVA)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



Supplementary Figure S13: Complex heatmaps (Z-score)

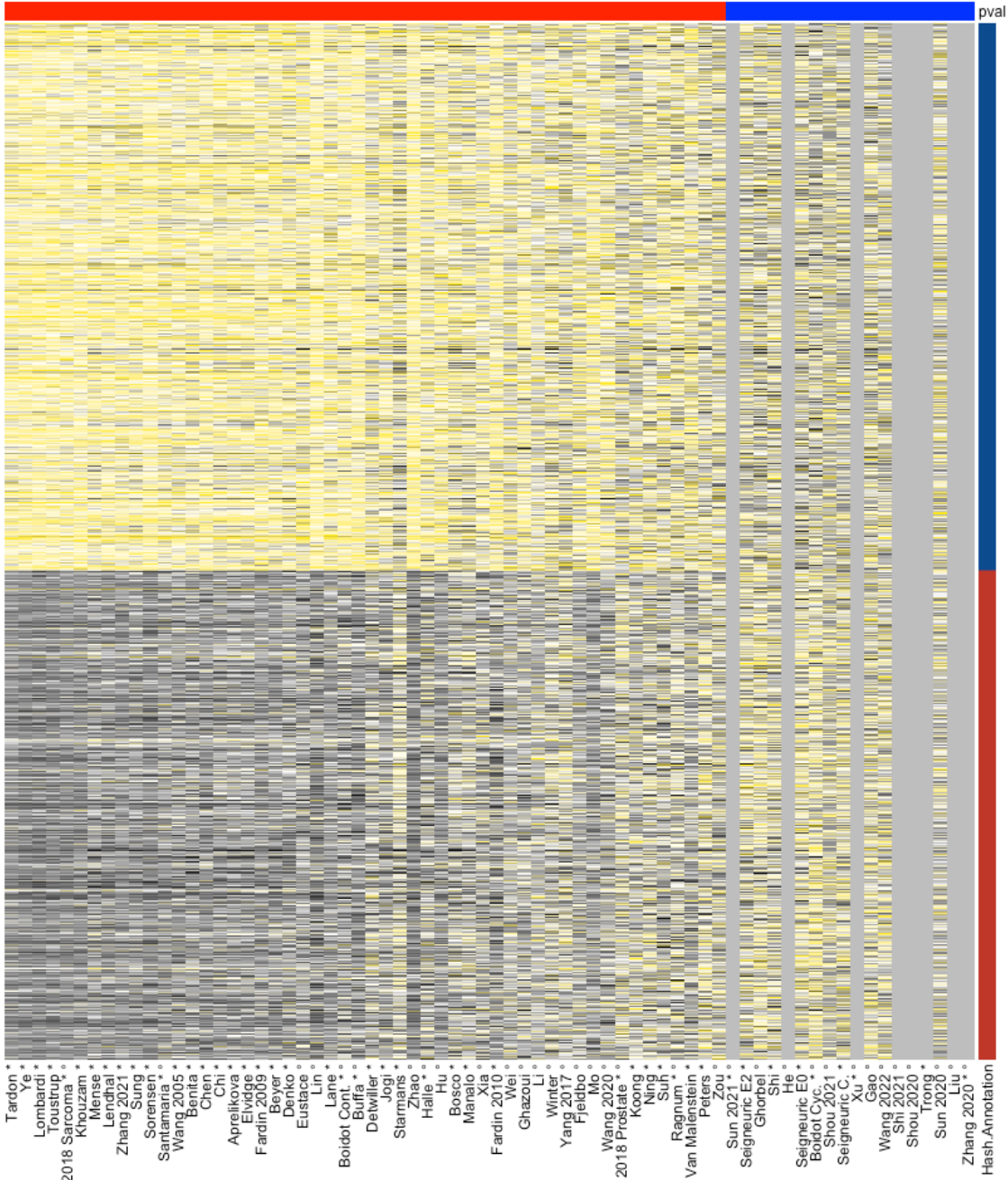
Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.



Supplementary Figure S14: Complex heatmap (NCS)

Complex heatmap of signature p-values in hypoxic vs. non-hypoxic samples across GEO cell-line conditions in bulk RNAseq and microarray data, related to Figure 3. Conventions as in S2.

Cells



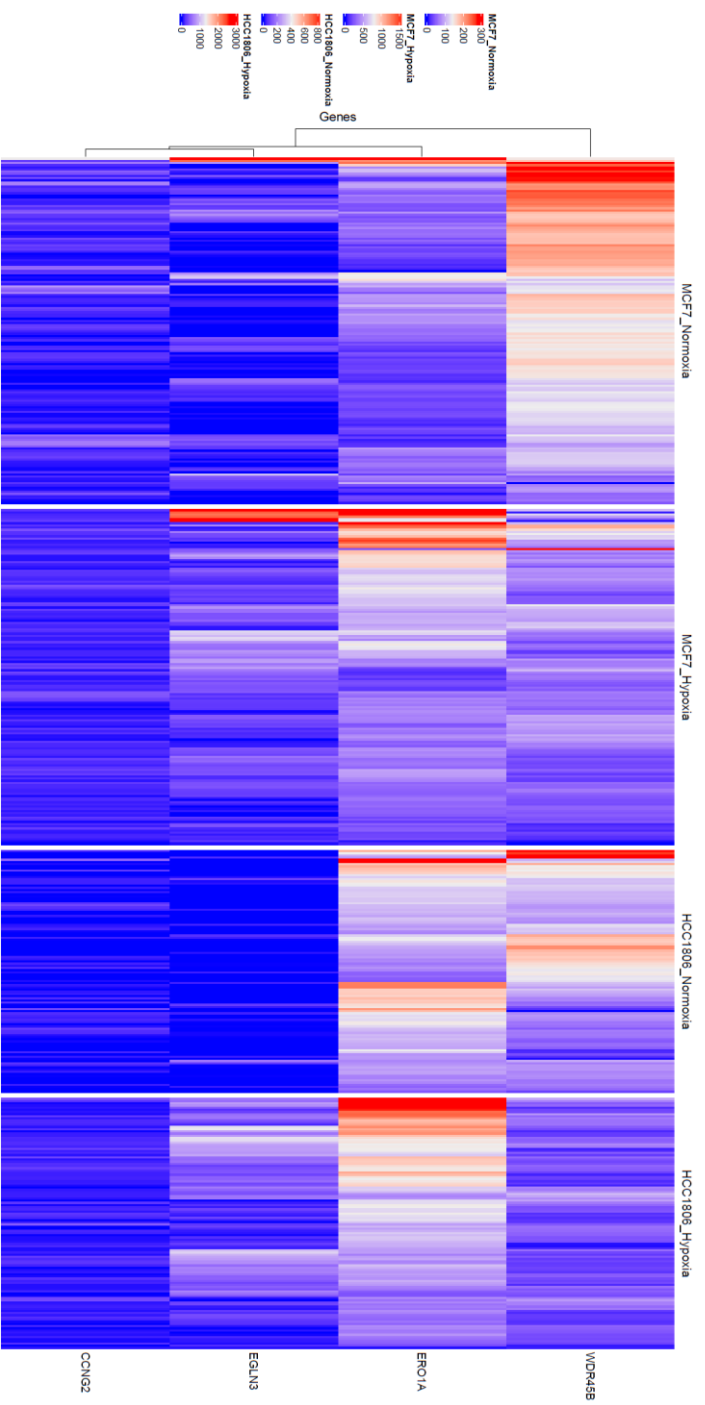
Hash.Annotation
■ NAT_kidney_hypoxia
■ NAT_kidney_normoxia

Tardon*
Ye*
Lombardi*
Toustrup*
Yang 2018 Sarcoma*
Khouzam*
Mense*
Lendhal*
Zhang 2021*
Sung*
Sorensen*
Santamaria*
Wang 2005*
Benita*
Chen*
Chi*
Aprelikova*
Elvige*
Fardin 2009*
Beyer*
Duroc*
Eustace*
Lin*
Lane*
Boidot Cort.*
Buffa*
DeWilder*
Jogi*
Starmans*
Zhao*
Halle*
Hu*
Bosco*
Mannato*
Xia*
Fardin 2010*
Wei*
Chazouli*
Li*
Winter*
Yang 2017*
Fjeldbo*
Mo*
Wang 2020*
Yang 2018 Prostate*
Koong*
Ning*
Sun*
Ragnum*
Van Malenstein*
Peters*
Sun 2021*
Seigneuric E2*
Ghorbel*
Shi*
He*
Seigneuric E0*
Boidot Cys.*
Shou 2021*
Seigneuric C.*
Xu*
Gao*
Wang 2022*
Shi 2021*
Shou 2020*
Trong*
Sun 2020*
Liu*
Zhang 2020*
Hash.Annotation

Signatures

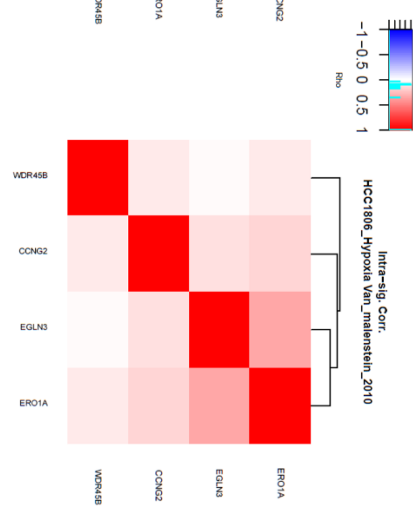
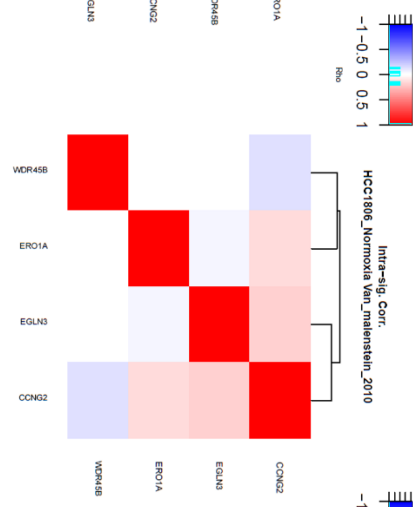
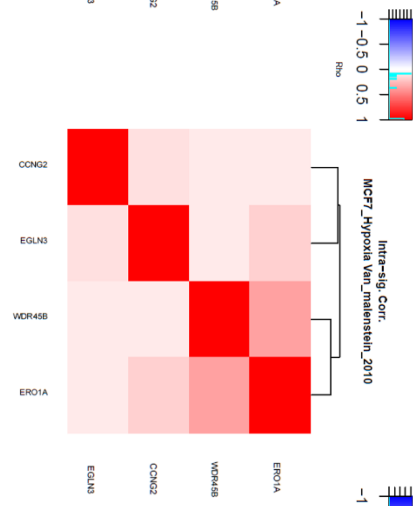
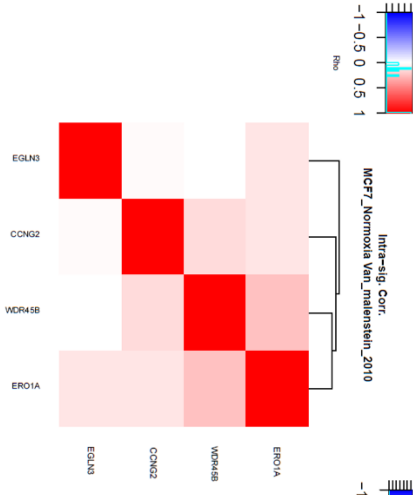
Supplementary Figure S15: Performance of the Tardon signature using the IQM in scRNA-Seq data for NAT human kidney (GSE200207) under normoxic and hypoxic conditions

Complex heatmap of IQM hypoxia score for hypoxic vs. normoxic cells in Normal Adjacent Tissue (NAT) from human kidney (GSE200207). Signatures on the x-axis are ordered based on their accuracy in the IQM score from Figure 3 for ease of comparison, related to Figure 4. The heatmap displays ranked-normalised hypoxia score values for each signature.



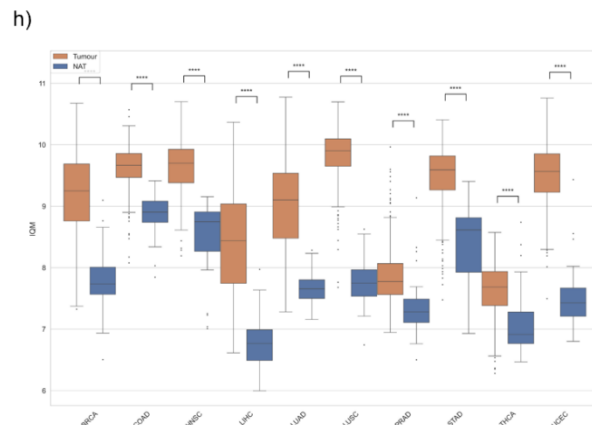
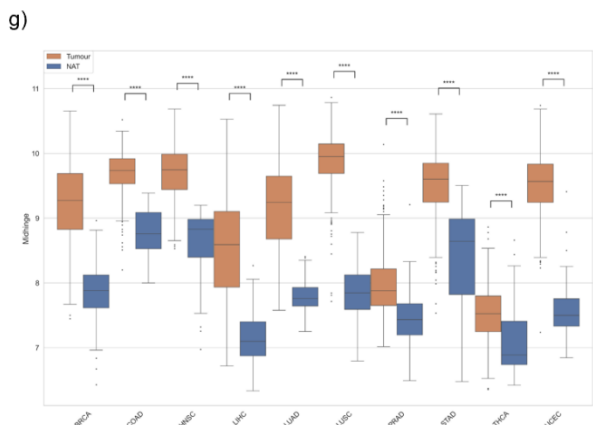
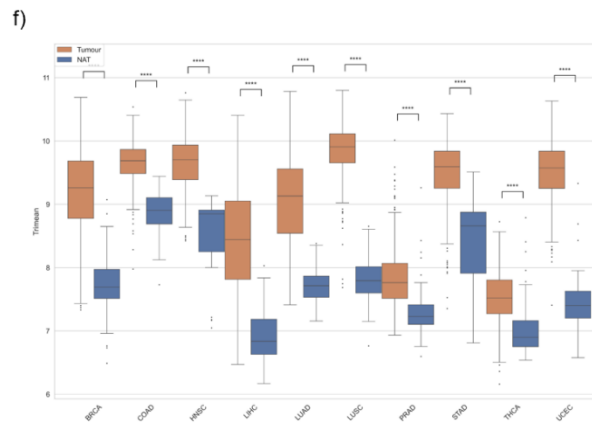
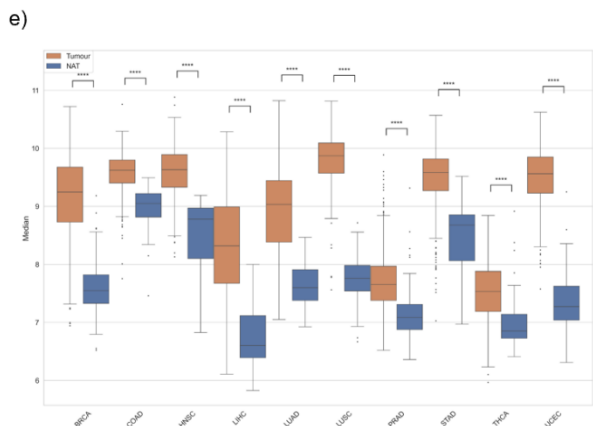
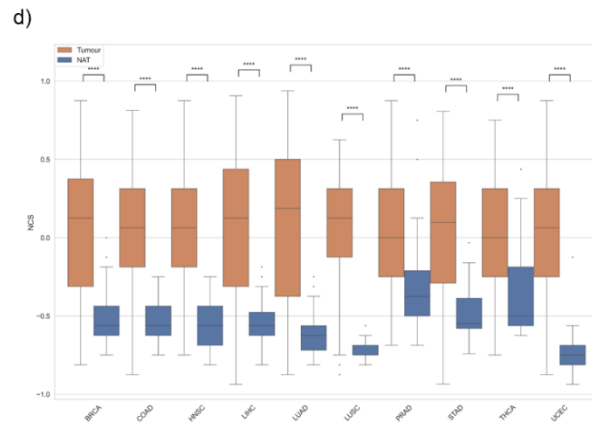
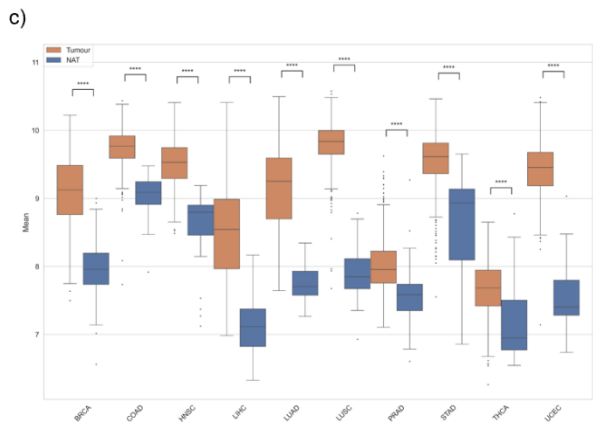
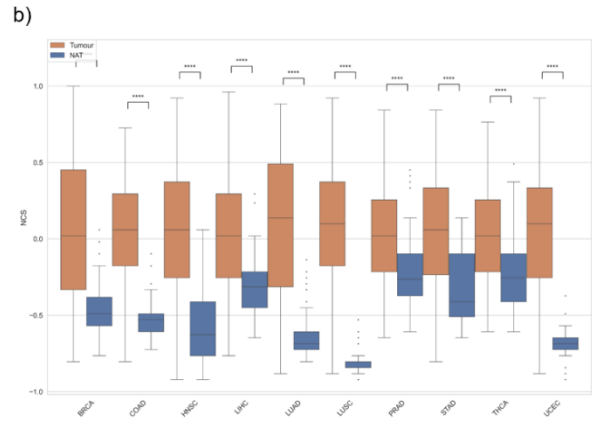
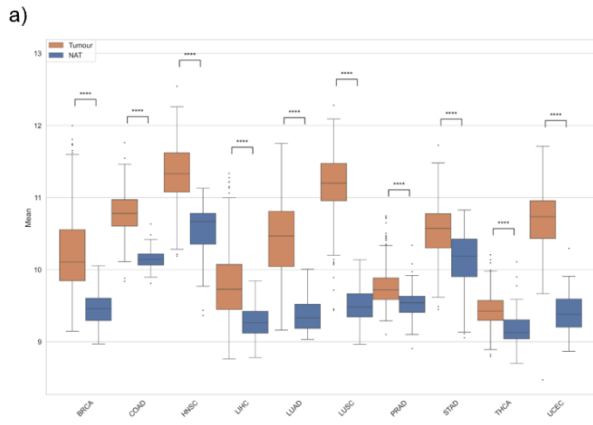
Supplementary Figure S16: van Malenstein expression heatmaps clearly show dominance of ERO1A and WDR45B

Heatmaps depicting expression levels of van Malestein features in normoxic and hypoxic subsets from MCF7 and HCC1806 single-cell datasets. Rows represent signature features, columns represent individual cells, and colour intensities indicate library size-adjusted expression values. Notably, ERO1A exhibits elevated expression in hypoxia and reduced in normoxia, while WDR45B displays the inverse. This contrasting expression pattern leads to a diminished intra-signature correlation, affecting the signature's performance in single-cell data, related to Figure 4.



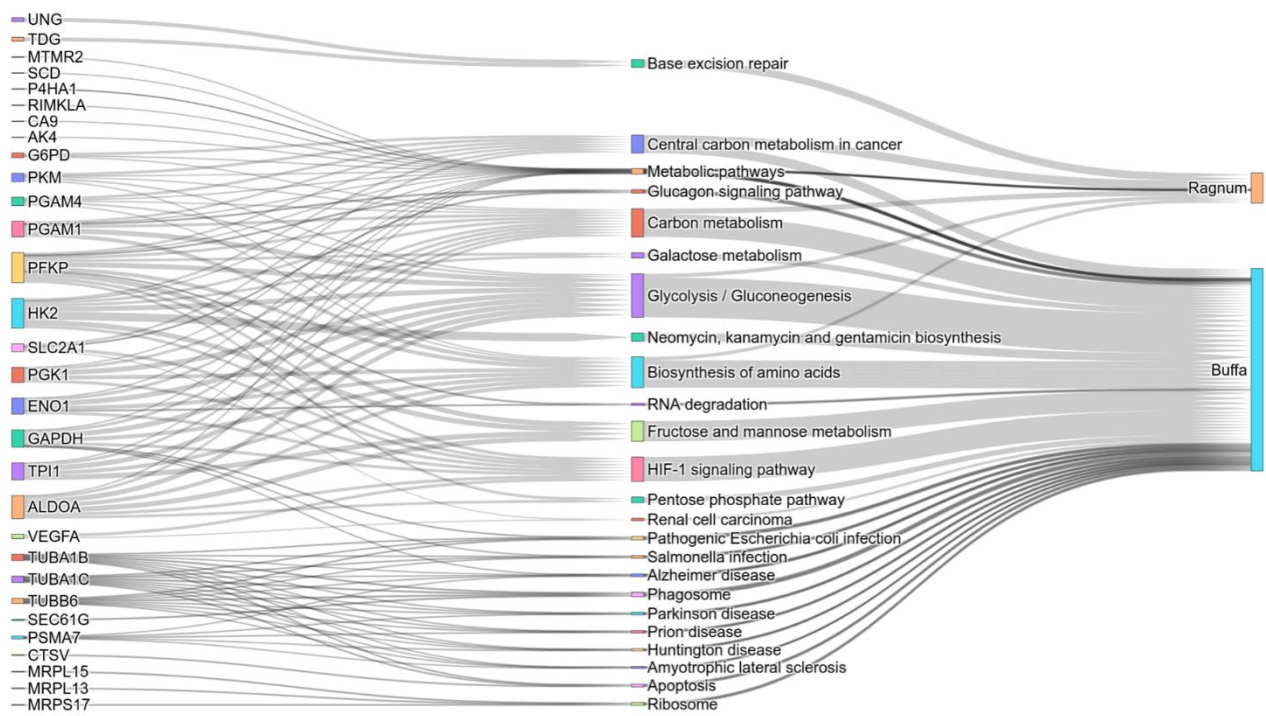
Supplementary Figure S17: Poor intra-signature correlation is the primary driver of aberrant signature performance in single-cell data.

Heatmaps displaying intra-signature auto-correlation for normoxic and hypoxic subsets from MCF7 and HCC1806 single-cell datasets. Both rows and columns represent van Malestein signature features, while colour intensities indicate correlation levels. The evident weak intra-signature correlation greatly influences the disparity in signature performance between bulk and single-cell datasets, related to Figure 4.



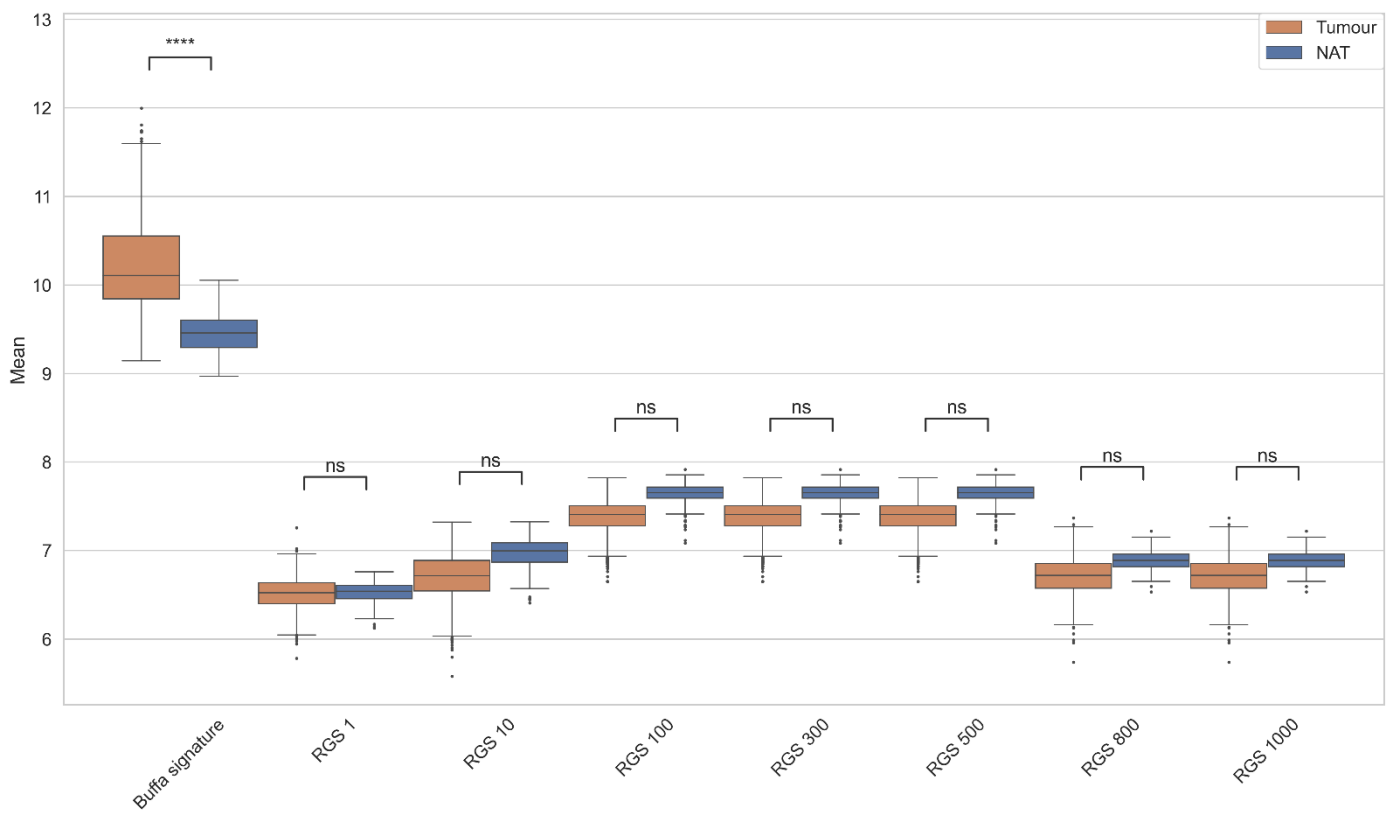
Supplementary Fig. S18: Two signatures, Buffa and Ragnum, showed higher scores in tumours compared to NAT across all 10 cancer types

Boxplots depicting hypoxia score differences in Buffa and Ragnum hypoxia signatures in 10 tumour types and normal adjacent tissue (NAT) in the The Cancer Genome Atlas (TCGA), related to Figure 5. Two signatures, Buffa and Ragnum, showed higher scores in tumours compared to NAT across all 10 cancer types using six scoring metrics. These results are shown in the eight panels, with orange denoting tumours and blue NAT: a) Buffa mean, b) Buffa NCS, c) Ragnum mean, d) Ragnum NCS, e) Ragnum median, f) Ragnum trimean, g) Ragnum midhinge, h) Ragnum NCS. The p-values were calculated using the Mann-Whitney-Wilcoxon test (****, $p \leq 1.00E-04$).



Supplementary Fig. S19: Gene-pathway associations in the Buffa and Ragnum hypoxia signatures

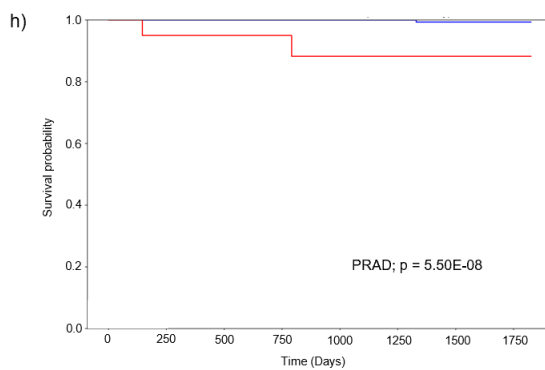
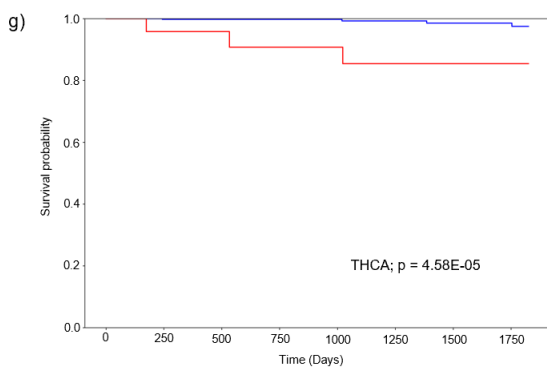
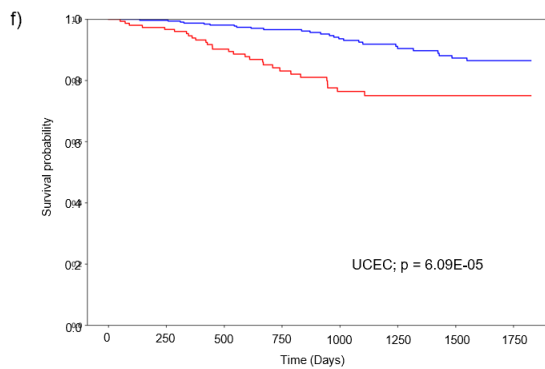
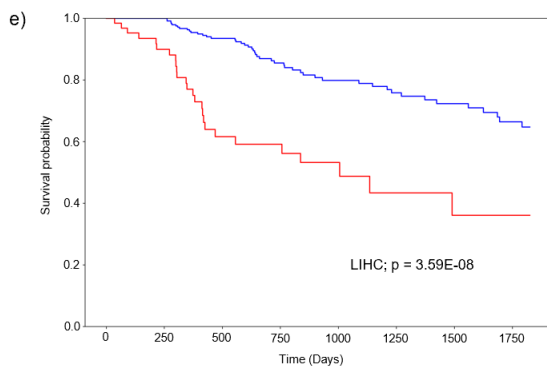
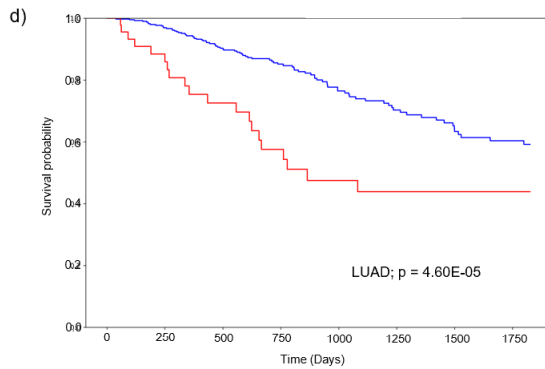
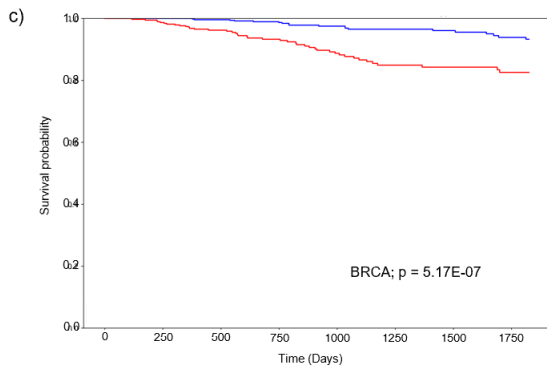
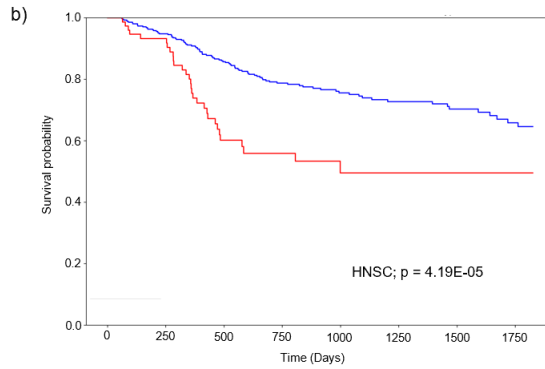
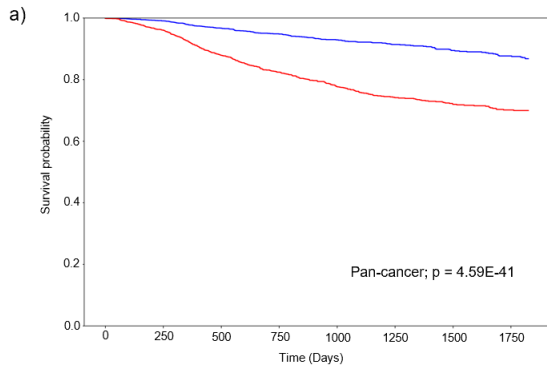
The Sankey diagram delineates the relationships between individual genes from Buffa and Ragnum hypoxia signatures and the biological pathways they are involved in. Each gene is linked to one or more pathways, denoting its functional role within the cellular processes. The pathways, in turn, converge into the hypoxia signatures of Buffa or Ragnum, highlighting the contribution of each pathway to the respective signature. The width of the bands reflects the number of genes shared between the entities, emphasising the significance of each gene and pathway to the hypoxia signatures' composition. The diagram reveals a significant enrichment of both signatures in pathways such as glycolysis/gluconeogenesis, central carbon metabolism in cancer, and broader carbon metabolism processes, related to Figure 5.



Supplementary Fig. S20: Hypoxia scores differences between hypoxia signatures and random gene sets in clinical samples

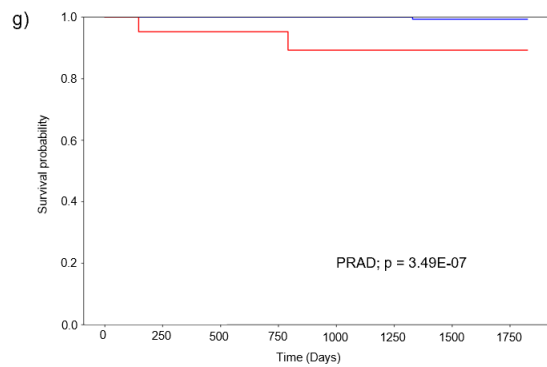
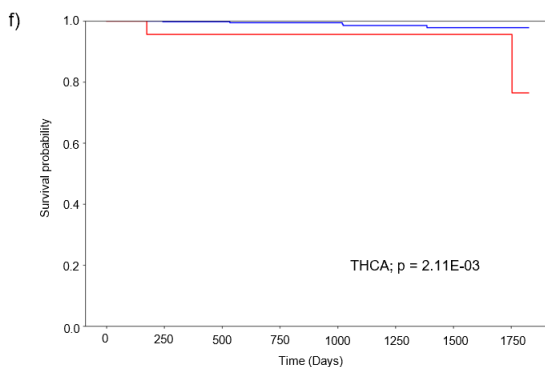
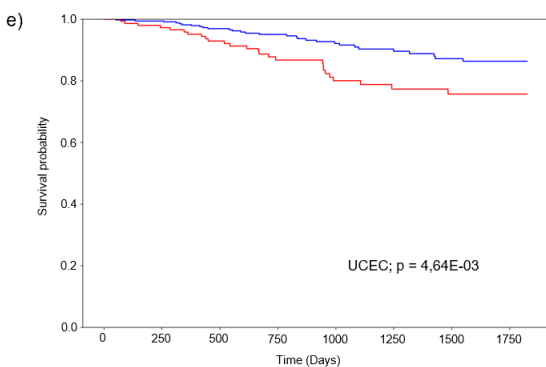
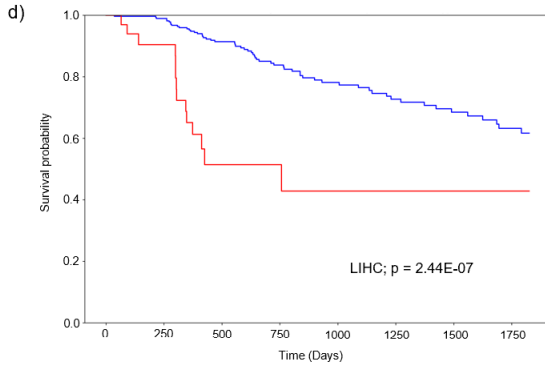
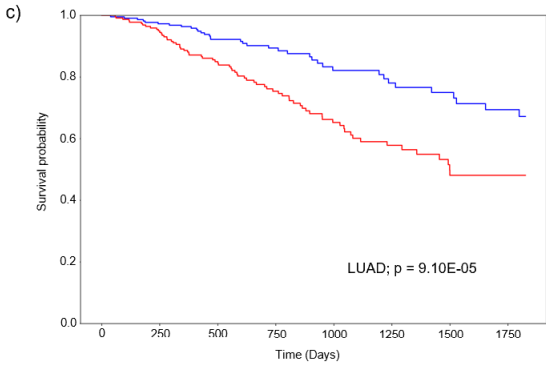
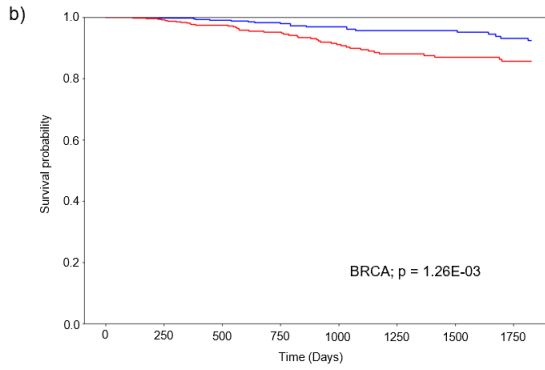
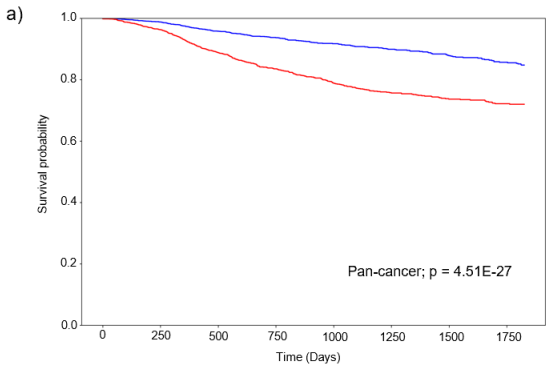
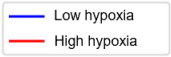
Boxplots depict the differences in hypoxia score distributions between Buffa signature and RGS using mean score in the BRCA cohort in TGCA. Hypoxia scores from RGS in tumour samples vs. normal adjacent tissue (NAT) show a lower score and a non-significant p-value as opposed to the original signature. The p-values were calculated using the Mann-Whitney-Wilcoxon test (****, $p \leq 1.00E-04$), related to Figure 5.

— Low hypoxia
— High hypoxia



Supplementary Fig. S21: Buffa/mean shows prognostic efficacy across the TCGA

Kaplan–Meier survival estimates using the Buffa hypoxia signature and the mean score were investigated across pan-cancer and across ten individual cancer types. A point of contention in the field is what type of thresholds should be used across tumour types to denote low and high hypoxic samples. Previous works have used above and below the median however this is not necessarily the most comprehensive approach. Thus, to be more thorough, the prognostic ability of Buffa/mean was examined every fifth percentile using disease specific survival (DSS). The cohorts are dichotomised into high (red) and low (blue) hypoxia groups according to the percentile that yields the most significant Log-rank p from the Kaplan-Meier survival analysis. Buffa/mean was significantly prognostic in 7/10 tumour types, including (b) HNSC [85th percentile], (c) BRCA [60th percentile], (d) LUAD [90th percentile], (e) LIHC [75th percentile], (f) UCEC [70th percentile], (g) THCA [95th percentile] and (h) PRAD [90th percentile]. Furthermore, Buffa/mean signature maintained its prognostic value in a combined pan-cancer dataset (a) [50th percentile], related to Figure 5.



Supplementary Fig. S22: Ragnum/IQM shows prognostic efficacy across the TCGA

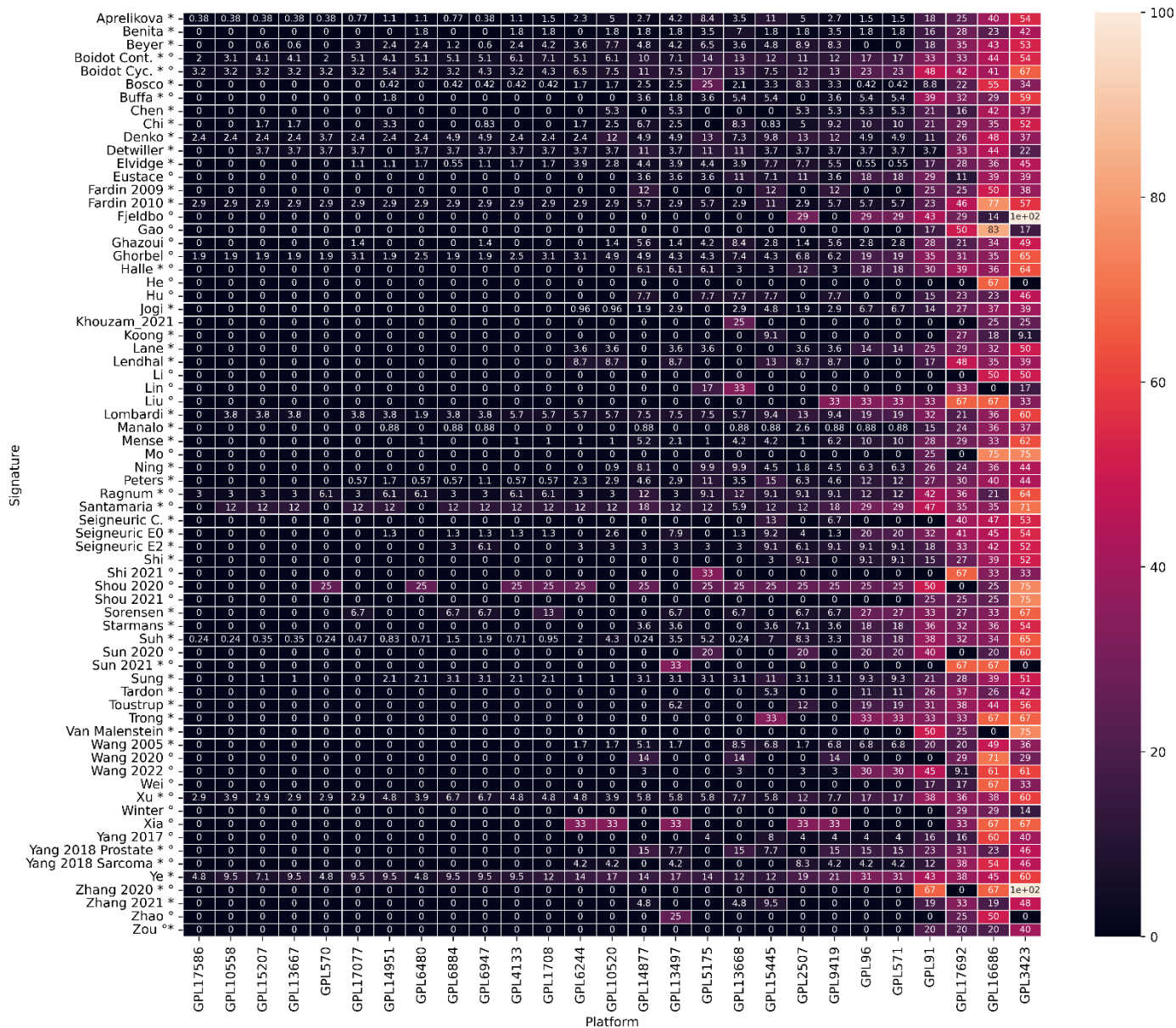
Kaplan–Meier survival estimates using the Ragnum hypoxia signature and the IQM score were investigated across pan-cancer and across ten individual cancer types across ten cancer types (conventions as in S47). Ragnum/IQM was significantly prognostic in 6/10 tumour types, including (b) BRCA [50th percentile], (c) LUAD [50th percentile], (d) LIHC [90th percentile], (e) UCEC [70th percentile], (f) THCA [50th percentile] and (g) PRAD [95th percentile]. Ragnum/IQM also maintained its prognostic value in a combined pan-cancer dataset (a) [50th percentile], related to Figure 5.

Signature

Aprelikova *	0	1.8	3.5	3.5	3.5	4	4.4	4.8	4.8	4.2	6.2	4.6	14	14	14	14	14	4.2	4.2	5.7	14	15	8.8	21	17	20
Benita *	0	0	0	0	1.9	1.9	1.9	1.9	1.9	5.3	1.9	5.3	5.3	5.3	5.3	5.3	5.3	1.9	5.3	5.3	1.9	14	5.3	29	24	
Beyer *	0	1.9	1.9	2.6	3.2	3.9	4.5	5.8	5.8	4.2	5.2	5.4	7.7	7.7	7.7	7.7	7.7	5.4	5.4	5.8	8.3	10	14	17	21	24
Boidot Cont. *	0	8.6	9.7	9.7	9.7	11	11	15	15	10	13	14	6.1	6.1	6.1	6.1	7.1	12	12	15	11	11	19	28	19	20
Boidot Cyc. *	0	7.9	7.9	10	12	11	15	15	15	6.5	16	9.7	4.3	4.3	4.3	4.3	4.3	6.5	6.5	24	6.5	17	36	19	39	17
Bosco *	0	0	0	0	0.58	1.2	1.2	0.58	0.58	25	2.3	25	28	28	28	28	28	26	26	4.1	28	33	21	38	48	36
Buffa *	0	0	0	2	2	3.9	3.9	2	2	8.9	7.8	8.9	8.9	8.9	8.9	8.9	8.9	8.9	3.9	8.9	8.9	5.9	21	11	27	
Chen *	0	0	0	0	0	0	0	0	11	0	11	11	11	11	11	11	11	0	11	11	0	16	21	25		
Chi *	0	0.9	0.9	1.8	3.6	4.5	4.5	4.5	4.5	2.5	4.5	3.3	7.5	7.5	7.5	7.5	3.3	3.3	4.5	7.5	8.3	17	13	15	27	
Denko *	0	1.4	1.4	1.4	1.4	2.8	4.2	4.2	4.9	2.8	7.3	12	12	12	12	12	7.3	7.3	1.4	13	15	8.3	21	18	28	
Detwiller *	0	0	0	0	0	0	3.7	3.7	0	0	3.7	0	0	0	0	0	3.7	0	0	7.4	0	7.4	19	19	11	16
Elvidge *	0	1.2	1.7	1.7	2.3	3.5	3.5	4	4	0.55	4.6	0.55	5	5	5	5	5	1.7	1.7	4	6.1	6.6	9.2	15	10	18
Eustace *	0	0	0	0	0	0	0	0	0	11	8	11	11	11	11	11	11	11	0	11	11	8	29	14	32	
Fardin 2009 *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	21	
Fardin 2010 *	0	12	12	12	15	15	15	15	15	5.7	15	5.7	2.9	2.9	2.9	2.9	2.9	5.7	5.7	15	5.7	2.9	15	20	5.7	15
Fjeldbo *	0	0	0	0	17	17	17	17	17	14	17	14	14	14	14	14	14	14	14	17	14	14	17	29	8.2	
Gao *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.3
Ghazoui *	0	4.4	4.4	5.9	5.9	8.8	8.8	7.3	7.3	4.2	13	4.2	4.2	4.2	4.2	4.2	4.2	4.2	8.8	4.2	4.2	10	15	7	26	
Ghorbel *	0	2.5	1.9	5.1	5.7	7	8.2	7	7	2.5	11	4.9	2.5	2.5	2.5	2.5	2.5	2.5	10	2.5	5.6	21	15	15	32	
Halle *	0	3.2	3.2	3.2	6.5	6.5	6.5	6.5	6.5	6.1	9.7	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.5	6.1	6.1	9.7	21	12	14	
He *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	0	33	0	
Hu *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7.7	0	7.7	7.7	0	20		
Jogi *	0	0	0	0	0	0	0	0	0.96	2.9	1.9	0.96	0.96	0.96	0.96	0.96	0.96	1.9	1.9	3.9	0.96	3.9	17	6.7	21	22
Khouzam_2021 *	0	0	0	0	0	0	0	0	0	12	0	12	12	12	12	12	12	12	0	12	12	0	12	12	43	
Koong *	0	0	0	0	0	0	0	0	0	9.1	0	9.1	9.1	9.1	9.1	9.1	9.1	9.1	0	9.1	18	20	18	27	13	
Lane *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	7.1	14	17	
Lendhal *	0	0	0	0	4.3	4.3	4.3	4.3	4.3	0	8.7	0	0	0	0	0	0	0	4.3	0	4.3	0	4.3	22	0	24
Li *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	17	0	17	53		
Lin *	0	0	0	0	0	0	0	0	0	17	0	17	17	17	17	17	17	17	0	17	17	0	17	17	32	
Liu *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	67	67	67	0	0	0	
Lombardi *	0	0	0	0	0	0	2.1	2.1	9.4	4.2	11	9.4	9.4	9.4	9.4	9.4	9.4	9.4	0	9.4	9.4	2.1	15	9.4	30	
Manalo *	0	0	0	0	0.95	1.9	1.9	0.95	0.95	2.6	2.9	2.6	7.9	7.9	7.9	7.9	7.9	4.4	4.4	2.9	9.7	9.7	18	13	20	26
Mense *	0	0	1.1	1.1	2.1	2.1	2.1	2.1	3.1	3.2	3.1	3.1	3.1	3.1	3.1	3.1	3.1	4.2	4.2	2.1	3.1	3.2	8.3	4.2	26	
Mo *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	
Ning *	0	2	2	4	4	4	4	4	4	2.7	5	2.7	11	11	11	11	11	2.7	2.7	5	12	12	15	21	14	23
Peters *	0	3.2	3.2	3.2	4.4	4.4	4.4	5.1	5.1	2.3	6.3	2.9	9.2	9.2	9.2	9.2	2.3	2.3	7	9.8	12	10	17	14	25	
Ragnum *	0	0	0	0	3.1	6.2	6.2	6.2	3	9.4	6.1	3	3	3	3	3	6.1	6.1	6.2	3	6.1	9.4	15	9.1	32	
Santamaria *	0	0	0	0	6.2	6.2	12	12	5.9	12	12	5.9	5.9	5.9	5.9	5.9	5.9	5.9	6.2	5.9	6.2	18	5.9	27		
Seigneuric C. *	0	0	0	0	0	0	0	0	13	0	13	6.7	6.7	6.7	6.7	6.7	13	13	0	13	6.7	7.1	20	20	31	
Seigneuric E0 *	0	4.4	4.4	4.4	5.9	5.9	5.9	5.9	5.9	5.3	7.3	5.3	11	11	11	11	7.9	7.9	10	13	13	19	22	25	23	
Seigneuric E2 *	0	0	0	0	3.2	3.2	3.2	6.5	6.5	9.1	6.5	12	6.1	6.1	6.1	6.1	6.1	9.1	9.1	3.2	9.1	9.1	27	12	23	
Shi *	0	0	0	0	3.2	3.2	3.2	3.2	3.2	6.1	3.2	6.1	6.1	6.1	6.1	6.1	6.1	6.1	6.5	6.1	6.1	15	16	9.1	24	27
Shi 2021 *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	33	33	0	0	0	33	11	
Shou 2020 *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	25	0	25	50	29	
Shou 2021 *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	25	18	
Sorensen *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6.7	33	6.7	20	44		
Starmans *	0	0	0	0	3.9	3.9	3.9	3.9	3.9	7.1	7.7	7.1	7.1	7.1	7.1	7.1	7.1	7.1	3.9	7.1	7.1	3.9	11	7.1	20	
Sun *	0	2	2.2	3.4	4.5	5.8	5.7	6.5	6.5	3.8	8.9	4.8	11	11	11	11	11	4.7	4.7	5.6	11	11	8.5	19	12	27
Sun 2020 *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	20	0	40	3		
Sun 2021 *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	
Sung *	0	1.1	1.1	1.1	3.3	3.3	3.3	3.3	3.3	3.1	5.6	4.1	7.2	7.2	7.2	7.2	7.2	3.1	3.1	3.3	8.2	8.2	7.8	16	9.3	23
Tardon *	0	0	0	0	0	0	0	0	0	0	5.3	0	0	0	0	0	0	0	0	0	0	11	11	0	23	
Toustrup *	0	0	0	0	0	0	0	0	0	6.2	0	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	0	12	6.2	30	
Trong *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0	0	0	67	0	
Van Malenstein *	0	0	0	0	25	25	25	25	25	0	50	0	0	0	0	0	0	0	25	0	0	50	0	0	2.6	
Wang 2005 *	0	11	11	11	11	11	11	13	13	10	13	10	6.8	6.8	6.8	6.8	6.8	10	10	11	10	8.5	16	20	19	42
Wang 2020 *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	14	0	14	0	14	29
Wang 2022 *	0	0	0	0	0	0	0	0	0	30	0	30	30	30	30	30	30	30	0	30	33	26	42	61	48	
Wei *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	17	0	8.2	
Xu *	0	1	1	1	1	2.1	2.1	4.1	4.1	6.7	6.2	8.7	6.7	6.7	6.7	6.7	6.7	7.7	7.7	2.1	7.7	8.7	6.2	26	12	20
Winter *	0	0	0	0	0	0	0	0	0	14	0	14	14	14	14	14	14	14	0	14	14	33	29	39	39	
Xia *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9.5	
Yang 2017 *	0	0	0	0	0	0	0	0	0	4	0	4	4	4	4	4	4	4	4	4	4	8.3	12	8	8.5	
Yang 2018 Prostate *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7.7	7.7	23	21	
Yang 2018 Sarcoma *	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	34		
Ye *	0	2.6	2.6	5.1	7.7	5.1	7.7	10	10	7.1	10	14	7.1	7.1	7.1	7.1	7.1	7.1	10	7.1	7.1	13	14	9.5	36	
Zhang 2020 *																										

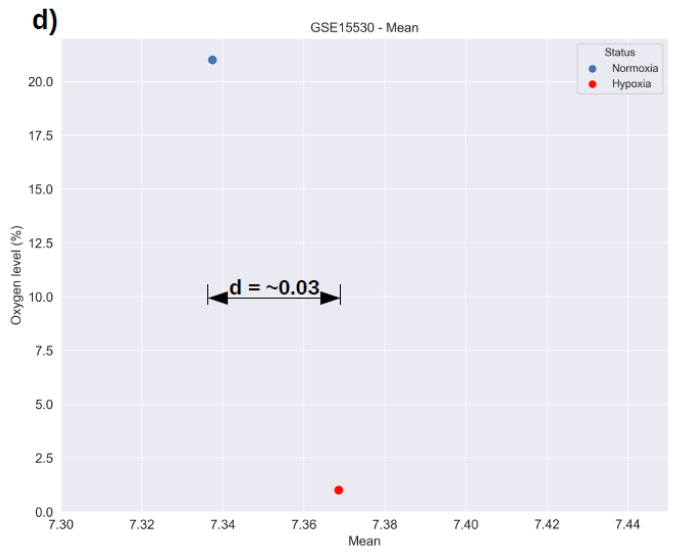
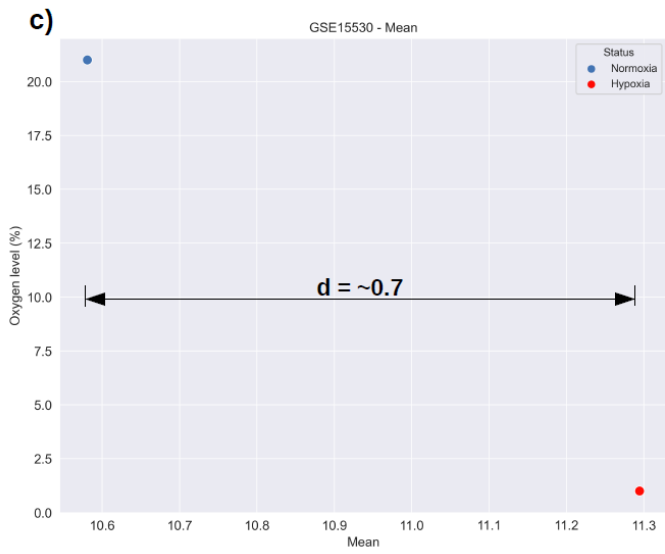
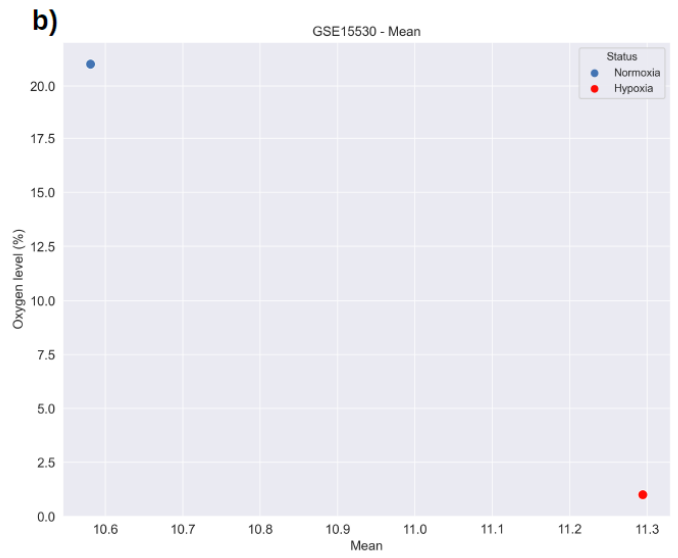
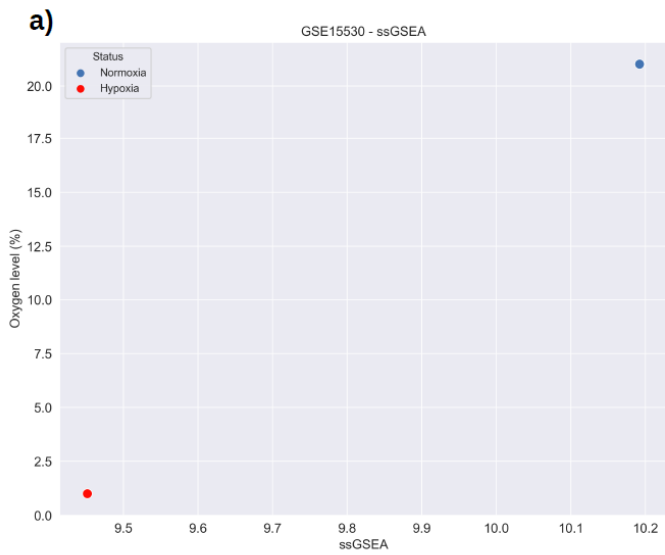
Supplementary Fig. S23: Percentage of genes missing in each signature in hypoxia-related bulk RNASeq experiments

Summary of the percentage of missing genes in each signature for all the bulk RNASeq platforms in GEO hypoxia experiments. Signature names are displayed on the y-axis, while GEO Platform (GPL) and Series (GSE) IDs are located on the x-axis. Lighter square colours indicate a higher percentage of missing genes for the respective signature, platform, and series. Individual missing-gene percentages for each signature and series are provided in the corresponding box, related to Table 1.



Supplementary Fig. S24: Percentage of genes missing in each signature in hypoxia-related microarray experiments

Summary of percentage missing genes in each signature for all the microarray platforms in GEO hypoxia experiments. Signature names are displayed on the y-axis, while GEO Platform (GPL) IDs are located on the x-axis. Since microarray data for each platform share the same manifest, Series (GSE) IDs are not reported. Lighter square colours indicate a higher percentage of missing genes for the respective signature and platform. Individual missing-gene percentages for each signature and series are provided in the corresponding box, related to Table 1.



Supplementary Fig. S25: Hypoxia scores and distance calculation against RGS

The scatterplots were generated using Buffa signature on one normoxic (GSM390200, highlighted in blue) and one hypoxic (GSM390196, highlighted in red) sample from the Series GSE15530. In *a*) the mean and *b*) the ssGSEA scores are plotted on the x-axis against the percentage of oxygen. The mean score of the hypoxic sample is greater than the score of the normoxic sample. Viceversa, the ssGSEA score of the hypoxic sample is lower than the score of the normoxic sample, resulting in a wrong classification. Panel *c*) shows that the distance $d_{SIG}(hyp, norm)$ calculated using the mean score between the two samples is $\sim 24x$ higher (left image, $d_{SIG} = \sim 0.7$) than the distance observed between the same samples using a RGS of the same length (Panel *d*, $d_{RGS} = \sim 0.03$), related to Figure 2.