Research Article

Enhancing Cancer Treatment and Understanding Through Clustering of Gene Responses to Categorical Stressors

Christopher El Hadi[1], George Hilal[2], Rita Aoun[3]

1. Faculty of Medicine, Saint Joseph University of Beirut, Beirut, Lebanon; 2. Cancer and Metabolism Laboratory, Faculty of Medicine, Saint Joseph University of Beirut, Beirut, Lebanon; 3. Saint Joseph University of Beirut, Beirut, Lebanon

Cancer cells have unique metabolic activity in the glycolysis pathway compared to normal cells, which allows them to sustain their growth and proliferation. Therefore, inhibiting glycolytic pathways may provide a promising therapeutic approach to cancer treatment. In this first-of-its-kind study, we analyzed the genetic responses of cancer cells to stressors, particularly drugs that target the glycolysis pathway. Gene expression data for experiments on different types of cancer cells were retrieved from the Gene Expression Omnibus and expression fold-change was then clustered after dimensionality reduction. We identified four response clusters, the first and third are affected the most by anti-glycolytic drugs, consisting mainly of squamous and mesenchymal tissues, showing higher mitotic inhibition and apoptosis. Drugs acting on several glycolytic targets at once resulted in such responses. The second and fourth clusters were relatively unaffected by the treatments, succumbing the least to glycolysis inhibitors. These clusters are mainly gynecological and hormone-sensitive, with drugs acting on hexokinases mainly inducing this response. This study highlights the importance of analyzing the molecular states of cancer cells to identify potential targets for personalized cancer treatments and to improve our understanding of the disease.

Corresponding author: Christopher El Hadi, chriselhadi@gmail.com

Introduction

In the 1920s, Otto Warburg and his colleagues observed a notable uptake of glucose in tumors in comparison to that observed in the surrounding tissues. Subsequent investigation revealed glucose fermentation, resulting in lactate production even in the presence of oxygen, a phenomenon termed

 doi.org/10.32388/1T88E3.3

"aerobic glycolysis" [1][2]. This paradoxical prevalence of glycolysis in cancer cells is referred to as the "Warburg effect." The catabolism of glucose to lactate yields a relatively low amount of energy, thereby necessitating high rates of glucose consumption to meet the energetic and anabolic demands of cancer cells [3]. In addition to providing energy, glycolysis generates metabolic intermediates that enable *de novo* synthesis of nucleotides, amino acids, lipids, and NADPH, all of which are essential for rapid cell proliferation. Cancer cells thus possess a distinct metabolic profile that permits a high rate of proliferation and resistance to apoptosis signals [4].

Inhibition of glycolytic pathways could be a promising selective approach in cancer research for developing targeted anticancer agents. Chen [5] and Abdel-Wahab [6] and colleagues have evaluated various inhibitory drug modalities that forestall glucose utilization. These drugs have been tested on cancer cells, followed by gene expression evaluation to identify more potent agents with a lethal metabolic impact on cancer cells. Currently, many glycolytic inhibitors are undergoing preclinical and clinical studies, demonstrating promising results [5][7][8].

While many studies have described the effects of glycolysis inhibitory drugs on various cancers, there has been no effort to identify a common thread among the outcomes of these treatments. This study aims to fill that gap by examining how cancer cells respond to stressful environments, specifically those induced by glycolysis inhibitory drugs. By making use of fundamental cancer characteristics such as the drugs applied and their molecular targets, along with cancer “iClusters”extracted using an integrative clustering technique proposed by Shen *et al.* [9][10], we hope to uncover hidden patterns that summarize cancer cell behavior under such conditions. The iClusters, which were obtained in a 2018 article by Hoadly *et al.*, represent a new way of classifying cancer based on molecular signatures rather than histology [11]. In their latest research, Hoadly et al. expanded on this concept by analyzing approximately 33 tumor types or 10,000 samples and identifying 28 distinct iClusters using various types of data, including chromosome arm-level aneuploidy, DNA methylation, mRNA and miRNA expression, and protein expression[13]. With the insights gained from this study, we can hopefully make strides towards developing more effective cancer treatments [12].

Results

Two matrices were constructed, the first containing 199 experiments and the second 666. The rows amount to 33622, representing the studied genes as expression foldchanges. The *adjust_matrix()*

function of the *cola* package was used to clean the matrices.

k	*1-PAC*	*Mean Silhouette*	*Concordance*
2	0.989	0.97	0.987
3	0.87	0.907	0.958
4	1	0.963	0.98
5	0.949	0.888	0.949
6	0.851	0.802	0.887

Table 1. Clustering evaluation indices for a maximum k of 6

In the main matrix of 199 columns, 4 treatments were removed because more than 80% of their values were missing, leaving us 195 samples and 16127 genes filtered for analysis. Consensus partitioning recommended grouping the experiments into 5 classes, and alternatively into 2 or 4. Table 1 shows the 1-PAC, mean silhouette, and concordance evaluation indices for all classes with a maximum of 6. Having calculated the product of 1-PAC, concordance, and silhouette for all 3 suggested k, k = 4 classes was the highest k with a product ≥ 0.90 and is hence preferred. The cumulative distribution function (CDF) curves for a maximum k of 6 and the PCA plot for k equal to 4 are shown in Figure 1a and Figure 1b. Only treatments with silhouette scores ≥0.5 were retained in their corresponding classes, resulting in a new total of 193 analyzable experiments.

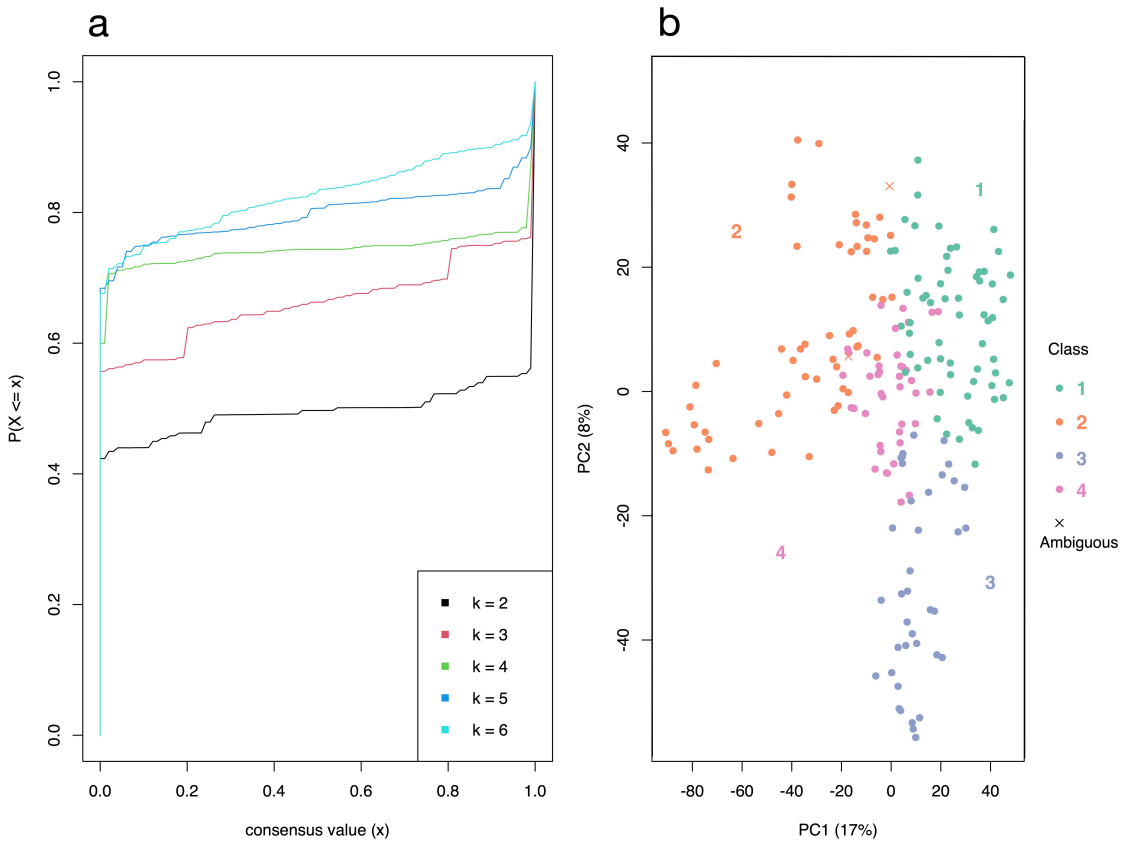

Figure 1. eCDF curve and PCA plot of the consensus clustering. (a) eCDF curve of the consensus matrix from partition by ATC:skmeans, a good k can be selected by aiming at the flatness of the eCDF curve. (b) PCA on 5000 rows with the highest ATC scores, 193 out of 195 confident samples were kept in their classes (silhouette > 0.5)

Samples were not unequally distributed (p = 0.054), with class 1 having 61 experiments, class 2 having 54, class 3 with 38, and class 4 having 40 experiments. The treatments were then classified by molecular mechanisms and the goodness-of-fit test was applied to the frequency of experiments in each class. Table 2 lists the treatments used and their respective molecular mechanism as explained by Abdel-Wahab. Treatments acting on Hexokinase 2 (HK2) showed a significant preference for class 2, and those acting on variants of mTORC (1/2) demonstrated high affinity for class 1. Combination of treatments acting on the mitochondrion (e.g., biguanides) with those acting on HK2 or with glucose restricting, constituting a type of "pan-glycolysis" inhibition, had a preference for class 3. This classification in class 3 was also seen with glucose starvation, a true "pan-glycolysis" inhibition, which additionally showed an affinity for class 1. Nevertheless, it should be noted that drugs targeting

doi.org/10.32388/1T88E3.3

mitochondrial function and the KRAS pathway, alone, did not show a predilection for any class, proving their wide range of activity. The results described are visualized in Figure 2.

Action	Drug	Action	Drug
Akt	Afuresertib	***Mitochondria***	Buformin
GLUT	Silybin		Metformin
Glucose	No Glucose		Phenformin
HK2	2DG	***mTORC***	Everolimus
	Genistein		MLN128
	Resveratrol		PP242
KRAS Mutations	Selumetinib		BEZ235
	Sorafenib		Rapamycin
	Trametinib		Temsirolimus
		PDK	DCA

Table 2. Actions of the studied drugs on the glycolytic pathway

doi.org/10.32388/1T88E3.3

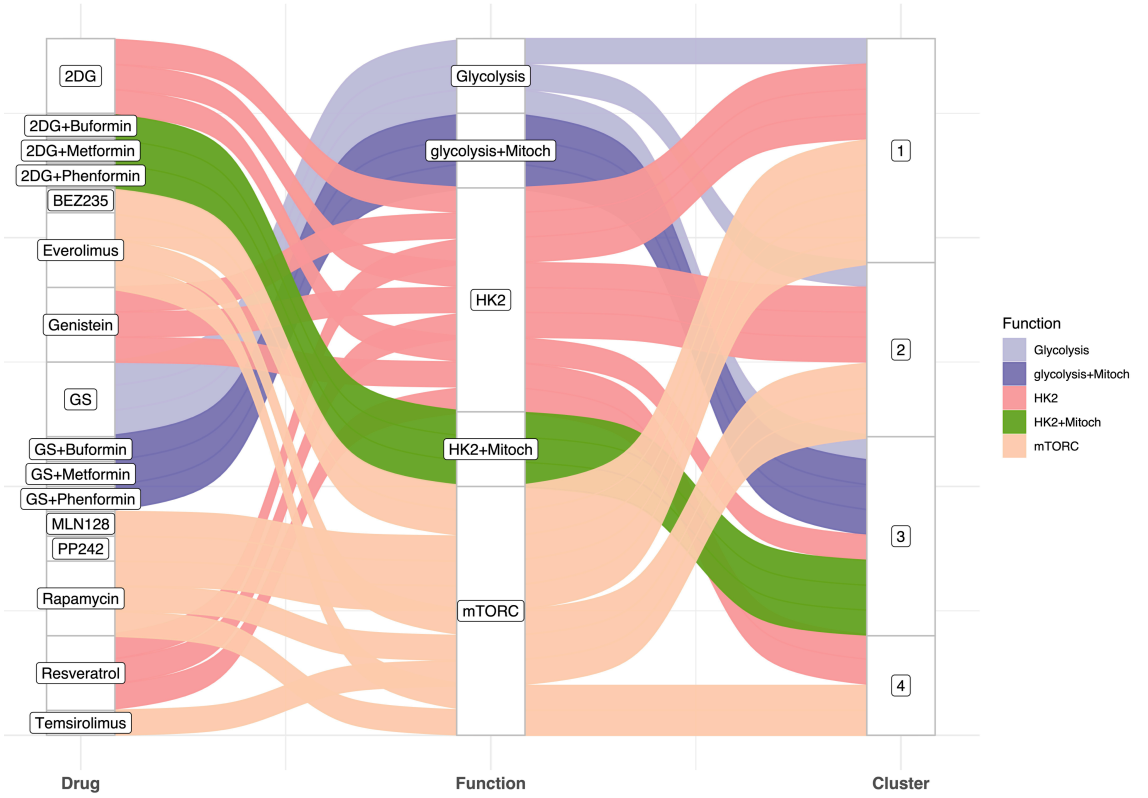

Figure 2. Relationship between the drugs studied, their molecular activity, and response clusters. The Sankey diagram shows molecular activity of each drug. Cluster designations are shown on the right.

In terms of cell type, Cluster 1 appears to harbor the majority of cells, comprising iClusters 15 (C15) and C24, almost exclusively comprising SKCM (skin/new melanoma) and LAML (myeloid leukemia), respectively. C15 is affected by UVB exposure and has the highest median mutation per megabase, and C24 is known to have the lowest median mutation per megabase among all iClusters. Similarly, C10, C25 and C27 (partitioning all squamous cell cancers) were predominantly classified in group 1 alongside group 3, as were C22, C7, C3, enriched in SARC (sarcomas, primarily bone). One cell type predominates in C10 and C27, namely lung and cervical squamous cell cancers, the latter being HPV-related. Moreover, C10, C25, and C27 were enriched with squamous-cell and proliferation-related pathways, relatively high level of hypoxia, immune-related signaling and basal signaling. Of note, C25 was specifically enriched with chr11 amplification. As for C22, C7, and C3, these iClusters were predominantly mesenchymal, most commonly having a chr9 deletion, and were enriched in gene programs representing PD1, CTLA4, and JAK2/STAT1, 3, 6 signaling. Cluster 2 was primarily enriched in C8 and C19, themselves known to cluster BRCA and UCEC (breast and uterus), with C8 also found

doi.org/10.32388/1T88E3.3

almost exclusively in cluster 4. C8 is also a homogeneous iCluster, dominated by a single tumor type, with the POLE mutation present in highly somatically mutated cells. Together with C19, they both demonstrate hormone and cell of origin dependence, with C19 furthermore expressing the estrogen signaling gene program. Figure 3 summarizes this repartitioning and Supplementary file S3 illustrates the analysis process described for the drugs and tissue group distribution.

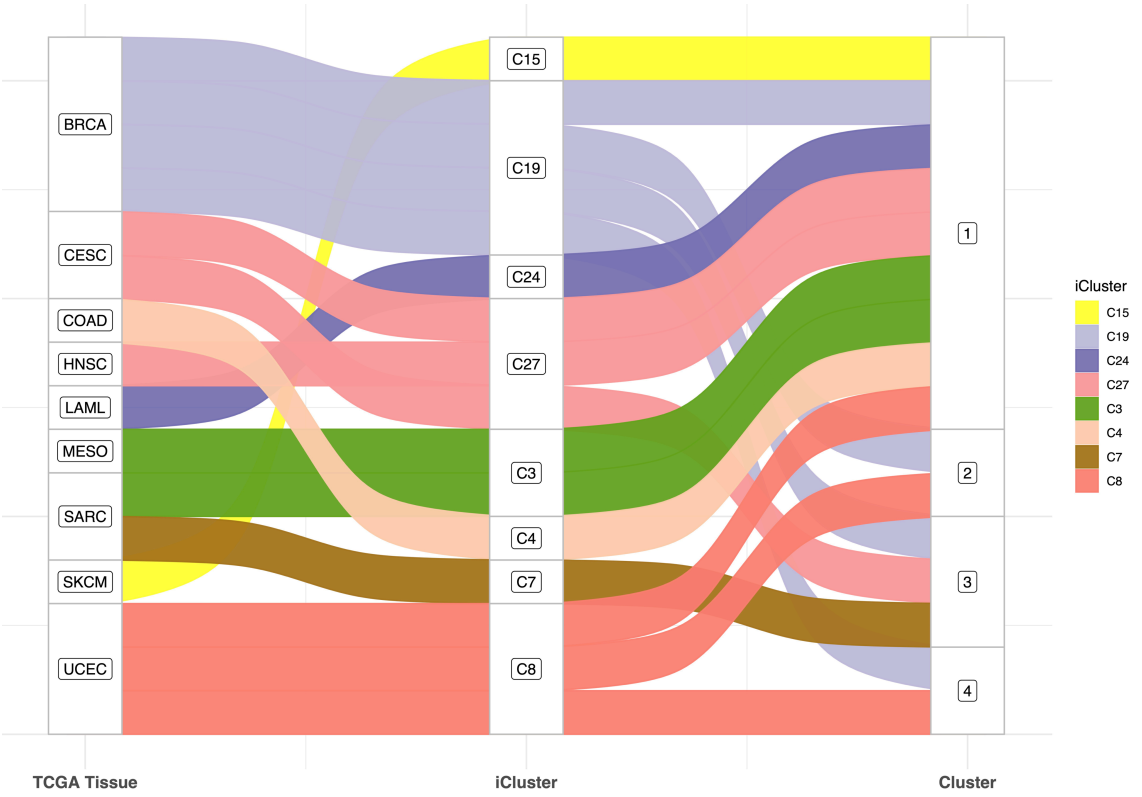

Figure 3. Relationship between TCGA tumor type, iCluster, and response clusters. The Sankey diagram shows the tumor type composition of each iCluster. Cluster designations are shown on the right.

Signature genes were identified using a false discovery rate (FDR) threshold of <0.001. Clustering of genes by K-means yielded 5 signature functions, 2 of which were biologically more significant than the others. Functional analysis of the first function-group using Gene Ontology (GO) terms for biological properties resulted in 17 term clusters (Figure 4). Function A addresses cell cycle inhibition, including (i) DNA replication and repair, as well as RNA splicing and telomere maintenance, and (ii) mitotic cell cycle and regulation. Function B addresses cancer adaptation and molecular quality control mechanisms describing (i) apoptosis and the unfolded protein response (UPR) and related aspects of (ii) RNA and protein transport and localization in cells, (iii) organ development (perception

doi.org/10.32388/1T88E3.3

and learning), and (iv) ion homeostasis. The activity of these functions in each group relative to the others (maximum

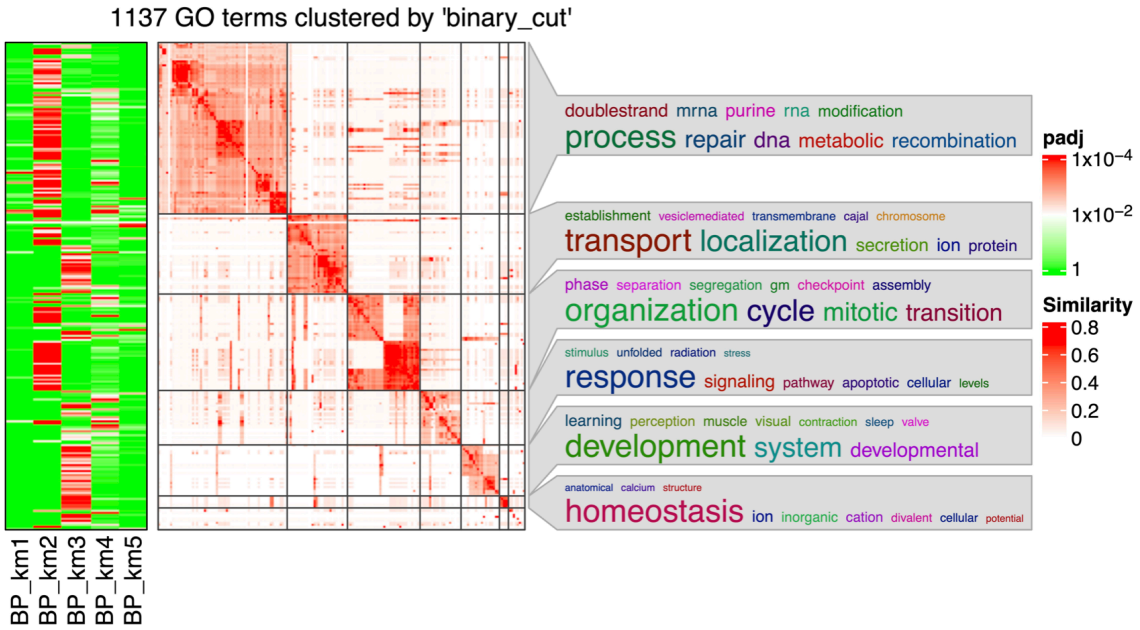

Figure 4. Heatmap of the GO term similarities from the 195-sample gene list. The terms resulted from the enrichment of the genes in either group (using FDR < 0.01). The green-red columns show for which group the respective GO terms are significant. The word cloud keywords highlight the biological functions in each GO group.

doi.org/10.32388/1T88E3.3

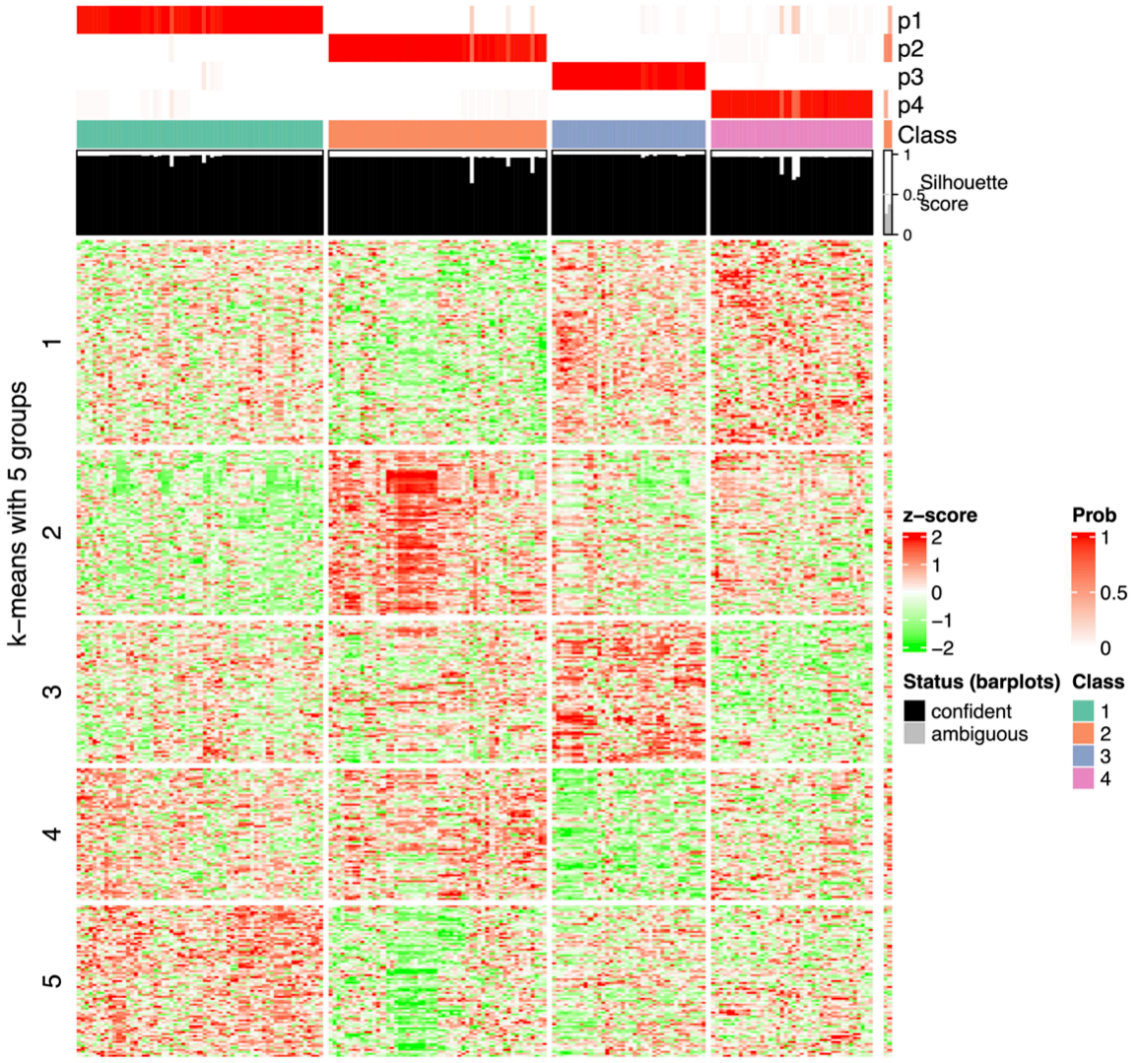

Figure 5. Heatmap of the signature genes and their regulation in each class. 9368signature genes (58.1% of total genes) chosen for an FDR < 0.01. This heat map represents the differences in expression relative to the average FC for each gene. Table 3. Summarizes these differences.

doi.org/10.32388/1T88E3.3

Cluster\Function	**A: mitotic activity & DNA repair**	**B: Quality control, UPR & apoptosis**
1	↓	–
2	↑	–
3	↓	↑
4	–	↓

Table 3. Function clusters to antiglycolytic agents

The previously listed functions describe the expression of biological pathways that are different between classes, regardless of whether they are up- or down-regulated in each experiment (significantly high FC). That said, up- or down-regulated responses remain to be explored. Because our data were quantile-normalized, a FC $\geq$ 0.5 or $\leq$ -0.5 was considered to define an up- or downregulated gene, respectively. Only classes 1 and 3 could be enriched, showing downregulation in mitotic activity and DNA replication, and class 3 showed additional upregulation in apoptotic activity and cellular stress response, cellular differentiation, and immune activation of cells.

Additional consensus clustering was performed on 341 sorafenib-treated and 325 rapamycin-treated samples (from accession of 1014 samples), with each treatment having primarily 3 available doses given for 3 exposure times. Consensus partitioning of rapamycin samples resulted in 3 classes and 4 classes of sorafenib. This classification takes into account the drug used and leaves room for classification and interpretation based on tissue, dose, and time. The classes obtained for both treatments simply split the samples almost evenly between them, which is true for tissue, treatment duration, and dose if tissue-dose and tissue-duration interactions were considered. Therefore, no hidden trends or categorization of these parameters could be inferred. It should be noted, however, that when tissue, treatment duration, and dose were not considered, there was a preference for certain classes ($P < 0.01$). Treatment signature genes were also classified and resulted in the same biological processes as described above, with organ development/differentiation being more vascular.

It was deemed unnecessary to combine the 666 experiments with the 195 and perform a clustering analysis because of the potential bias that could result from adding more than 300 samples treated with only 2 agents, to 195 samples treated with a total of 26 agents. In addition, partitioning that holds the tissue variable constant and allows for a treatment-based study was also deemed inefficient due to the limited number of samples per tissue.

Discussion

This study used gene expression data from the Gene Expression Omnibus (GEO) to cluster cancer gene responses to stressors, specifically drugs that inhibit the glycolysis pathway. Unsupervised clustering after dimensionality reduction revealed four response clusters, which showed distinct molecular characteristics and sensitivity to different classes of drugs.

Our findings first highlight the importance of "pan-glycolysis" targets, especially glucose restriction, which showed the best anti-cancer effect among all other treatments. This pan-glycolytic effect can also be achieved by combining compounds, the most potent one being those affecting simultaneously the cytoplasmic and mitochondrial phases of glucose oxidation. The only treatments seemingly effective as a monotherapy are those targeting mTOR variants, which are known to play an important role in cancer growth and metabolism [13], which stresses the importance of multi-target therapies for targeting cancer. Secondly, it came clear that patients with mesenchymal and squamous tumors may benefit from mTOR and pan-glycolytic targets. However, patients with hormone-sensitive tumors probably do not benefit from any anti-glycolytic drug but can consider those acting on mitochondrial proteins and KRAS mutations. One can also consider the latter cells to adapt well to hypoglycemic and glycolysis-challenged environments.

Clusters 1 and 3 showed significant gene regulation that is worth elaborating upon. On the first hand, tumors classified in these clusters showed downregulation in mitosis, DNA replication, and all associated pathways and quality control checkpoints. This should mean that the drugs induced Warburgian stress that disrupted normal cell cycle processes, potentially slowing down or halting cell growth and division. On the other hand, cluster 3 showed the detection of immune-related gene expression. Knowing that cell culture lacks immune cells from its microenvironment, this expression can be due to various factors, including autocrine immune-related signaling [14], stress-induced signaling and death associated with exposure and production of damage-associated molecular patterns (DAMPs) [15], and persistence of immune cell from the original tissue. The additional

detection of genes related to development and differentiation can be due to the initiation of adaptative responses to the stressful environment by evolutionary favoring of more resistant mutants through, for instance, incomplete epithelial-mesenchymal transition (EMT) or mesenchymal-epithelial transition (MET) [16][17], and the ability of cancer cells to dedifferentiation or transdifferentiation into different cell types [18].

All of the studies we reviewed for data interrogation, which investigated the response of specific cell lines to a single dose and duration of exposure of an antiglycolytic drug, demonstrated similar biological functions to ours (GEO accession numbers of the studies that demonstrated deregulation of class 1 and 3: GSE31058, GSE97346, GSE59882, GSE79316, GSE73923, GSE59228, GSE36847, GSE25412, GSE114060, GSE96794, GSE79246, GSE9008, GSE62663, GSE116387, GSE137553, GSE112079; and class 2 and 4 dysregulation: GSE59704, GSE5200, GSE85257, GSE112079. See supplementary Table S1).

The distribution of cancer cells and drugs is not binary and must be considered when interpreting results due to the complex and diverse nature of the experiments. The differences are attributed primarily to the heterogeneity of the experiments, acting on the unique metabolism of each of the 79 cell lines tested by exposing them to different doses and time frames of 26 reagents. The varying levels of activity of biological pathways in each cell are due to developmental and differentiation programs, and epigenetic states of the cells of origin in conjunction with exogenous factors, such as mutagenic exposures, pathogens, and inflammation. This explanation should be extended to cover cancer's microenvironment when translating the findings on cancer responses to *in vivo* studies. We should therefore mention that cancer's microenvironment, which includes neighboring cells, blood vessels, and extracellular matrix, can affect the expression of glycolysis-related genes [19][20], limit drug delivery to the tumor cells, and alter the immune response to the tumor [21][22].

It is reasonable to mention at this point that further *in silico* studies are needed to validate our observations. Future studies should concentrate on quantifying the reactions in a cell through genomic, proteomic, and other omic module detection, rather than weighing the intensity of gene expression, and then comparing those biological modules between cells [23]. Asking for more *in vitro* experiments to obtain more statistically significant results is theoretically advantageous, but practically useless knowing that we are asking for more combinations of experiments that distances us from the study's purpose. We believed that there must be a hidden pattern, imperceptible to the human brain, that links human cancers together allowing their positioning on a categorizable

continuum. This explicates why we used iClusters in place of tissues to categorize cells and explore what is beyond simplistic tissue-based classifications.

Conclusions

While several studies have examined the effects of a single antiglycolytic drug on specific cell lines, none, to our knowledge, have used response clustering to analyze cancer. Our study aimed to fill this gap by identifying categories of cancer cell responses to a wide range of antiglycolytic agents. This classification can provide insight into the expected effects of any drug targeting glycolysis, helping to predict the response of cancer cells to such drugs.

Our results revealed unique molecular characteristics and varying levels of sensitivity to different drug classes, highlighting the importance of multi-targeted therapies in cancer treatment. Patients with mesenchymal and squamous tumors may benefit from mTOR and pan-glycolytic targets, whereas those with hormone-sensitive tumors are unlikely to benefit from anti-glycolytic drugs; however, drugs such as biguanides and those targeting KRAS mutations may be considered for the latter. Finally, it is important to consider the cancer microenvironment when applying these findings to *in vivo* studies.

Methods

Data Processing. Gene expression profiles were downloaded from the GEO (Gene Expression Omnibus) database. The terms entered in the search engine were "Cancer" followed by the name of the agent, or any of the alternative names for the same agent. All the dataset accession numbers and their respective drugs and cell lines can be found in Supplementary Table S1. The raw data were also provided in Supplementary Tables S3 and S4. The filters used with every search were "Homo sapiens" for organisms, "Expression profiling by array" for the study type, and "Series" under the entry type. The inclusion criteria we based our judgment on for choosing the datasets were:

1. Expression data should only be from single-channel microarray chips. Agilent, Affymetrix, and Illumina are the only accepted manufacturers
2. Cells should be treated in an artificial *in vitro* milieu (no mice xenografts etc.)
3. Cells should not have been transfected or mutated to affect the targeted pathway or been made resistant to treatment (only wild cell types under the influence of chemicals)

4. Cells should not have undergone secondary treatments (other than the vehicle and the drug of interest) before RNA extraction
5. Coding RNA, particularly cytoplasmic or total extractions, is preferred
6. The glucose deprivation medium should contain less than 1g/L, i.e., 5.6 mM, of glucose
7. Cells should be cancerous, not benign, with no other superimposed disease
8. The data series should contain both controls and treatments.

The raw files downloaded were either in CEL or TXT format. Few TXT files contained processed data, and these were simply used as-is. For the Illumina Beadchips, R's *limma* [24][25] package was used and the function *neqc()* was employed to perform background correction followed by quantile normalization. It should be noted that negative control probes were not always available, therefore the detection p-values of each probe were exploited instead. For probe annotation, the *annotate* package was chosen [26], with *illuminaHumanv3.db* as the annotation data [27], and "gene symbol" as the annotation. For Affymetrix chips, *limma*, *affy* [28], and *oligo* [29] were used to process the CEL files. The *rma()* function normalized the microarray signals, and the annotation data packages were adapted to the version of Affymetrix chips available [30][31][32][33][34][35][36][37]. The Agilent chips were read using likewise the *limma* package, annotated according to the chip versions [38][39][40], background-corrected following the normal-exponential method, and quantile normalized.

After reading and normalizing the signals, the controls and their respective treatments were exported into excel files. This resulted in a total of 798 microarrays (517 treated assays and 281 controls) comprising 73 cell lines from 17 different tissues treated with 26 different anti-glycolytic agents. An additional 1014 microarrays were also downloaded on only one study (666 treated assays and 348 controls), which included 59 cell lines from 11 tissues treated with either sorafenib or rapamycin at different drug doses and exposure durations. The log base 2 foldchange (FC) was then calculated for each treated chip, and all the FCs were lastly fitted into two matrices: the first consisted of 199 foldchange retrievals, the second 666. The matrices thus contain expression data in form of FCs.

Consensus clustering. Consensus clustering was performed using the *cola* package (version 1.8.1) [41]. This data classification technic was, foremost, done on the 199 experiments found on the GEO, to find a common denominator function for all glucose-challenged cells. Clustering was also performed on specific treatments, namely rapamycin and sorafenib from the 1014-sample accession, to study the effect of sample classification, while holding the treatment constant. Before performing the

consensus partitioning, an important step was to clean up and perform data imputation to the input matrix. The *adjust_matrix()* function provided by *cola was hence used*

After cleaning, quantile normalization was performed on each sample. This scaling method was used as it demonstrated better results compared to other scaling methods. For dimensionality reduction, the top 5000 features were first selected by the ATC top-value method (or "Ability To Correlate to other rows"). Then, the matrix was scaled by the selected rows and randomly sampled; these samples were partitioned by the skmeans (or "Spherical k-means") clustering method, a variant of the standard k-means clustering. Finally, the sampling and partitioning process was repeated 50 times to obtain a list of partitions.

The best-fitting number "k" of subgroups was evaluated using the average silhouette score, PAC score, concordance, and Jaccard index. The PAC score measures the proportion of ambiguous samples; 1-PAC thus represents the unambiguous sub-grouped samples. The best number of subgroups was the subgroup with the largest k having a product of the 3 scores greater than 0.9. The chi-square test was used to investigate the null hypothesis of whether the sample frequencies are equally distributed among the partitioned clusters, and the Wilson interval with continuity correction was applied to estimate the population proportions in each cluster.

Functional analysis. The “signature genes” were then identified using the F-test for differential analysis. Signatures are simply the rows that show statistically significant specificity in one or more subgroups. These were further grouped by patterns among subgroups using k-means clustering and the appropriate number of signature groups was automatically selected. Functional enrichment was applied to each group of signatures separately. Gene ontology (GO) enrichment analysis was performed by applying the R package *ClusterProfiler* [42] with *cola*’s function *functional_enrichment()*. GO terms were then clustered and visualized as heatmaps by the R package *simplifyEnrichment* [43].

Supplementary Information

- **Supplementary Table S1.** Samples accessions.
- **Supplementary Table S2.** 195-sample classification.
- **Supplementary Table S3.** 195-sample raw data.
- **Supplementary Table S4.** 666-sample raw data.

References

1. ^ *O. Warburg. (1956). On the origin of cancer cells. Science. 123(3191):309–314. doi:10.1126/science.123.3191.309PubMed PMID: 13298683*
2. ^ *O. Warburg. (1925). The Metabolism of Carcinoma Cells. J Cancer Res. 9(1):148–163. doi:10.1158/jcr.1925.148*
3. ^ *Alexei Vazquez, Jurre J. Kamphorst, Elke K. Markert, Zachary T. Schug, Saverio Tardito, et al. (2016). Cancer metabolism at a glance. J Cell Sci. 129(18):3367–3373. doi:10.1242/jcs.181016*
4. ^ *H. Liu, Y. P. Hu, N. Savaraj, W. Priebe, T. J. Lampidis. (2001). Hypersensitization of Tumor Cells to Glycolytic Inhibitors. Biochemistry. 40(18):5542–5547. doi:10.1021/bi002426w*
5. a, b *Xi-sha Chen, Lan-ya Li, Yi-di Guan, Jin-ming Yang, Yan Cheng. (2016). Anticancer strategies based on the metabolic profile of tumor cells: therapeutic targeting of the Warburg effect. Acta Pharmacol Sin. 37(8):1013–1019. doi:10.1038/aps.2016.47*
6. ^ *Ali F. Abdel-Wahab, Waheed Mahmoud, Randa M. Al-Harizy. (2019). Targeting glucose metabolism to suppress cancer progression: prospective of anti-glycolytic cancer therapy. Pharmacol Res. 150:104511. doi:10.1016/j.phrs.2019.104511*
7. ^ *Lynn Jeanette Savic, Julius Chapiro, Gregor Duwe, Jean-François Geschwind. (2016). Targeting glucose metabolism in cancer: a new class of agents for loco-regional and systemic therapy of liver cancer and beyond? Hepatic Oncol. 3(1):19–28. doi:10.2217/hep.15.36*
8. ^ *Nicholas S. Akins, Tanner C. Nielson, Hoang V. Le. (2018). Inhibition of Glycolysis and Glutaminolysis: An Emerging Drug Discovery Approach to Combat Cancer. Curr Top Med Chem. 18(6):494–504. doi:10.2174/1568026618666180523111351*
9. ^ *Ronglai Shen, Adam B. Olshen, Marc Ladanyi. (2009). Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. Bioinformatics. 25(22):2906–2912. doi:10.1093/bioinformatics/btp543*
10. ^ *Ronglai Shen, Qianxing Mo, Nikolaus Schultz, Venkatraman E. Seshan, Adam B. Olshen, et al. (2012). Integrative Subtype Discovery in Glioblastoma Using iCluster. Vladimir Brusiceditor. PLoS ONE. 7(4):e35236. doi:10.1371/journal.pone.0035236*
11. ^ *Katherine A. Hoadley, Christina Yau, Denise M. Wolf, Andrew D. Cherniack, David Tamborero, et al. (2014). Multiplatform Analysis of 12 Cancer Types Reveals Molecular Classification within and across Tissues of Origin. Cell. 158(4):929–944. doi:10.1016/j.cell.2014.06.049*

12. ^Katherine A. Hoadley, Christina Yau, Toshinori Hinoue, Denise M. Wolf, Alexander J. Lazar, et al. (2018). Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell. 173(2):291–304.e6. doi:10.1016/j.cell.2018.03.022

13. ^Zhilin Zou, Tao Tao, Hongmei Li, Xiao Zhu. (2020). mTOR signaling pathway and mTOR inhibitors in cancer: progress and challenges. Cell Biosci. 10(1):31. doi:10.1186/s13578-020-00396-1

14. ^Hendrik Ungefroren. (2021). Autocrine TGF-β in Cancer: Review of the Literature and Caveats in Experimental Analysis. Int J Mol Sci. 22(2):977. doi:10.3390/ijms22020977PubMed PMID: 33478130; PubMed Central PMCID: PMC7835898

15. ^Jitka Fucikova, Oliver Kepp, Lenka Kasikova, Giulia Petroni, Takahiro Yamazaki, et al. (2020). Detection of immunogenic cell death and its relevance for cancer therapy. Cell Death Dis. 11(11):1–13. doi:10.1038/s41419-020-03221-2

16. ^Vivek Mittal. (2018). Epithelial Mesenchymal Transition in Tumor Metastasis. Annu Rev Pathol. 13:395–412. doi:10.1146/annurev-pathol-020117-043854PubMed PMID: 29414248

17. ^Dianbo Yao, Chaoliu Dai, Songlin Peng. (2011). Mechanism of the Mesenchymal–Epithelial Transition and Its Relationship with Metastatic Tumor Formation. Mol Cancer Res. 9(12):1608–1620. doi:10.1158/1541-7786.MCR-10-0568

18. ^João Carvalho. (2020). Cell Reversal From a Differentiated to a Stem-Like State at Cancer Initiation. Front Oncol. 10. Available from: https://www.frontiersin.org/articles/10.3389/fonc.2020.00541

19. ^Sophie Taylor, Enrico Pierluigi Spugnini, Yehuda G. Assaraf, Tommaso Azzarito, Cyril Rauch, et al. (2015). Microenvironment acidity as a major determinant of tumor chemoresistance: Proton pump inhibitors (PPIs) as a novel therapeutic approach. Drug Resist Updat. 23:69–78. doi:10.1016/j.drup.2015.08.004

20. ^Kate M. Bailey, Jonathan W. Wojtkowiak, Arig Ibrahim Hashim, Robert J. Gillies. (2012). Targeting the metabolic microenvironment of tumors. Adv Pharmacol San Diego Calif. 65:63–107. doi:10.1016/B978-0-12-397927-8.00004-XPubMed PMID: 22959024; PubMed Central PMCID: PMC3796340

21. ^Lei Zhu, Fugui Yang, Xinrui Li, Qinchuan Li, Chunlong Zhong. (2021). Glycolysis Changes the Microenvironment and Therapeutic Response Under the Driver of Gene Mutation in Esophageal Adenocarcinoma. Front Genet. 12. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2021.743133

22. ^Chu Xiao, He Tian, Yujia Zheng, Zhenlin Yang, Shuofeng Li, et al. (2022). Glycolysis in tumor microenvironment as a target to improve cancer immunotherapy. Front Cell Dev Biol. 10. Available from: https://www.frontiersin.org/articles/10.3389/fcell.2022.1013885

23. ^Christopher El Hadi, Georges Ayoub, Yara Bachir, Michèle Haykal, Nadine Jalkh, et al. (2022). Polygenic and Network-based studies in risk identification and demystification of cancer. Expert Rev Mol Diagn. 22(4):427–438. doi:10.1080/14737159.2022.2065195

24. ^R: The R Project for Statistical Computing. [cited 17 Sep 2021]. Available from: https://www.r-project.org/

25. ^Matthew E. Ritchie, Belinda Phipson, Di Wu, Yifang Hu, Charity W. Law, et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43(7):e47–e47. doi:10.1093/nar/gkv007

26. ^R. Gentleman. annotate. Bioconductor 2017. doi:10.18129/B9.BIOC.ANNOTATE

27. ^Andy Lynch Mark Dunning. illuminaHumanv3.db. Bioconductor 2017. doi:10.18129/B9.BIOC.ILLUMINAHUMANV3.DB

28. ^L. Gautier, L. Cope, B. M. Bolstad, R. A. Irizarry. (2004). affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 20(3):307–315. doi:10.1093/bioinformatics/btg405

29. ^Benilton S. Carvalho, Rafael A. Irizarry. (2010). A framework for oligonucleotide microarray preprocessing. Bioinformatics. 26(19):2363–2367. doi:10.1093/bioinformatics/btq431

30. ^Marc Carlson. hgu133plus2.db. Bioconductor 2017. doi:10.18129/B9.BIOC.HGU133PLUS2.DB

31. ^James W. MacDonald. hugene10sttranscriptcluster.db. Bioconductor 2017. doi:10.18129/B9.BIOC.HUGENE10STTRANSCRIPTCLUSTER.DB

32. ^James W. MacDonald. hugene11sttranscriptcluster.db. Bioconductor 2017. doi:10.18129/B9.BIOC.HUGENE11STTRANSCRIPTCLUSTER.DB

33. ^James W. MacDonald. hugene20sttranscriptcluster.db. Bioconductor 2017. doi:10.18129/B9.BIOC.HUGENE20STTRANSCRIPTCLUSTER.DB

34. ^James W. MacDonald. hugene21sttranscriptcluster.db. Bioconductor 2017. doi:10.18129/B9.BIOC.HUGENE21STTRANSCRIPTCLUSTER.DB

35. ^Bioconductor Core Team. human.dbo. Bioconductor 2017. doi:10.18129/B9.BIOC.HUMAN.DBo

36. ^Marc Carlson. hthgu133a.db. Bioconductor 2017. doi:10.18129/B9.BIOC.HTHGU133A.DB

37. ^The Bioconductor Project. hgu219cdf. Bioconductor 2017. doi:10.18129/B9.BIOC.HGU219CDF

38. ^Marc Carlson. RnAgilentDesigno28282.db. Bioconductor 2017. doi:10.18129/B9.BIOC.RNAGILENTDESIGNo28282.DB

39. ^Marc Carlson. HsAgilentDesign026652.db. Bioconductor 2017. doi:10.18129/B9.BIOC.HSAGILENTDESIGN026652.DB

40. ^Marc Carlson. hgug4112a.db. Bioconductor 2017. doi:10.18129/B9.BIOC.HGUG4112A.DB

41. ^Zuguang Gu, Matthias Schlesner, Daniel Hübschmann. (2021). cola: an R/Bioconductor package for consensus partitioning through a general framework. Nucleic Acids Res. 49(3):e15–e15. doi:10.1093/nar/gkaa1146

42. ^Guangchuang Yu, Li-Gen Wang, Yanyan Han, Qing-Yu He. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS J Integr Biol. 16(5):284–287. doi:10.1089/omi.2011.0118PubMed PMID: 22455463; PubMed Central PMCID: PMC3339379

43. ^Zuguang Gu, Daniel Hübschmann. (2021). simplifyEnrichment: an R/Bioconductor package for Clustering and Visualizing Functional Enrichment Results. bioRxiv. :2020.10.27.312116. doi:10.1101/2020.10.27.312116

Supplementary data: available at https://doi.org/10.32388/1T88E3.3

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.