Research Paper

Identification of dual-purpose therapeutic targets implicated in aging and glioblastoma multiforme using PandaOmics - an AI-enabled biological target discovery platform

Andrea Olsen[1,*], Zachary Harpaz[1,2,*], Christopher Ren[3,*], Anastasia Shneyderman[4], Alexander Veviorskiy[4], Maria Dralkina[4], Simon Konnov[4], Olga Shcheglova[4], Frank W. Pun[4], Geoffrey Ho Duen Leung[4], Hoi Wing Leung[4], Ivan V. Ozerov[4], Alex Aliper[4], Mikhail Korzinkin[4], Alex Zhavoronkov[4]

[1]The Youth Longevity Association, Sevenoaks, NA, United Kingdom
[2]Pine Crest School Science Research Department, Fort Lauderdale, Florida 33334, USA
[3]Shanghai High School International Division, Shanghai 200231, China
[4]Insilico Medicine Hong Kong Ltd., Hong Kong Science and Technology Park, New Territories, Hong Kong, China
*Equal contribution

Correspondence to: Mikhail Korzinkin; **email:** mike@insilicomedicine.com
Keywords: aging, target discovery, GBM, glioblastoma, PandaOmics
Received: February 7, 2023 **Accepted:** April 9, 2023 **Published:** April 26, 2023

ABSTRACT

Glioblastoma Multiforme (GBM) is the most aggressive and most common primary malignant brain tumor. The age of GBM patients is considered as one of the disease's negative prognostic factors and the mean age of diagnosis is 62 years. A promising approach to preventing both GBM and aging is to identify new potential therapeutic targets that are associated with both conditions as concurrent drivers. In this work, we present a multi-angled approach of identifying targets, which takes into account not only the disease-related genes but also the ones important in aging. For this purpose, we developed three strategies of target identification using the results of correlation analysis augmented with survival data, differences in expression levels and previously published information of aging-related genes. Several studies have recently validated the robustness and applicability of AI-driven computational methods for target identification in both cancer and aging-related diseases. Therefore, we leveraged the AI predictive power of the PandaOmics TargetID engine in order to rank the resulting target hypotheses and prioritize the most promising therapeutic gene targets. We propose cyclic nucleotide gated channel subunit alpha 3 (*CNGA3*), glutamate dehydrogenase 1 (*GLUD1*) and sirtuin 1 (*SIRT1*) as potential novel dual-purpose therapeutic targets to treat aging and GBM.

INTRODUCTION

Glioblastoma Multiforme (GBM) is the most deadly and aggressive brain tumor that exists, accounting for 16% of all primary brain tumors [1]. GBM is caused by genetic mutations leading to uncontrolled growth of glial stem or progenitor cells. This cancer is divided into primary GBM (*de novo*) and secondary GBM (progression from lower-grade tumors) [2]. Secondary GBM occurs most often in younger patients with a mean age of 45 years, whereas primary GBM most often happens in patients with a mean age of 62 years. Secondary GBM is, furthermore, the predominant GBM diagnosis in pediatric cases [1]. Primary GBM differs from secondary glioblastomas in its mutations, proteins, and pathways involved, as well as the distribution of molecular subtypes among patients (Proneural, Neural, Classical, Mesenchymal) [3, 4].

Current treatment for GBM includes maximal surgical resection, radiotherapy, and chemotherapy, such as concomitant and maintenance temozolomide, and these approaches usually lead to tumor resurgence and a median survival of 15 months [5]. Recently, tumor-treating fields (TTF), was proposed as a novel glioblastoma therapy that uses alternating electric fields to disrupt the division of cancer cells. TTF therapy has been shown to improve overall survival in patients with newly diagnosed or recurrent glioblastoma when used in combination with standard treatments such as chemotherapy and radiation therapy [6]. It is important to highlight some of the immunotherapies are either already approved for the treatment of GBM or on the way to being approved. For example, Pembrolizumab and Nivolumab are two checkpoint inhibitors that are being tested in clinical trials for glioblastoma [7]. However, despite significant advances in treating glioblastoma, the current standard of care for GBM is damaging to the brain and results in an overall 10-year survival rate (after the diagnosis) of 0.71% [8].

Most targeted therapeutics for GBM account for 1 to 2 targets in the GBM tumor, such as epidermal growth factor receptor (EGFR) and isocitrate dehydrogenase (IDH1), allowing cancer cells without expressing specific targets to survive and eventually kill the patient [9, 10]. More genetic targets for GBM must be identified and a method to personalize gene therapy for GBM patients may be useful for treating the many varieties of mutations present in GBM. Furthermore, identifying potentially effective combinations of drugs for the treatment of GBM patients using automated drug discovery approaches is an important task [11].

Due to the variety of mutations and development of GBM across different ages [3], this study utilized artificial intelligence (AI) and age-correlation analysis with the goal to identify dual-purpose targets associated with aging and GBM leveraging the power of AI-driven PandaOmics platform, which integrates analysis of multi-omics and literature data. Recently, the platform has been demonstrated to predict novel age-associated targets for the purpose of drug discovery [12].

Identifying therapeutic targets is crucial for successful drug development, as erroneous targets can lead to costly programs and failed clinical trials. While traditional computational approaches for target and biomarker discovery are important for success, they face limitations due to complex data and batch effects [13]. However, AI-driven approaches have shown efficacy in this area, using pathway analysis and algorithms on multiomics data to identify new targets and biomarkers, even with insufficient prior evidence [14–17].

Insilico Medicine scientists have pioneered the use of generative artificial intelligence in biology since 2016 [18, 19]. Many of the summary of the original generative biology approaches utilizing the generative systems for generation of synthetic biological data and the first demonstration of PandaOmics platform were first presented at the Interdisciplinary Workshop and the National Institute of Aging [20]. In addition to the target discovery using generative biology approaches, Insilico developed the capabilities in generative chemistry with the first peer reviewed publication in 2016 [21–24]. These approaches have been successfully applied to various diseases and have demonstrated their potential in identifying novel compounds and accelerating drug development [25, 26]. As such, generative chemistry and biology approaches are becoming increasingly important for the current and future drug discovery efforts. This paper demonstrates the application of generative biology approach to the complex interplay between GBM and Aging.

Aging is the exponential decline in homeostatic capabilities, which ultimately leads to an increased risk of age-related diseases and death [27]. With the rise of the longevity field, a novel approach has been undertaken to target diseases not just separately, with the goal to slow aging, but multiple diseases at once. Aging has, furthermore, been studied in the context of specific diseases, as we did in the present study. GBM has particularly strong aging-related effects on patients (including decreased resilience for aggressive treatment, comorbidity, and age-dependent immune system status), at the same time, this disease is prompted by old age [28]. By investigating GBM from both the traditional cancer-targeting approach and with the novel aging-oriented approach, we hope to find new genetic targets, which may alleviate both the onset of GBM as well as slow aging in patients.

RESULTS

The starting point of our analysis was to collect relevant GBM datasets and combine them into a single multi-omics project inside the PandaOmics system. The final project consisted of data obtained from 29 different studies: 25 RNA-seq/microarray, 3 methylation and 1 proteomics (Supplementary Table 1). These data were subsequently used to generate genesets mined from age correlation analysis, survival rates, differences in expression levels and previously published information on aging-related genes [12]. Intersection of those sets following three different strategies and ranking the resulting lists by PandaOmics TargetID AI-driven scores led us to the identification of the most promising dual-purpose therapeutic target candidates (Figure 1).

Correlation analysis

Among the list of 25 transcriptomics datasets, age metadata were available for 12 which were consequently used for the correlation analysis (Supplementary Table 2). Spearman's correlation coefficients between gene expression and age, as well as corresponding *p*-values, were calculated for each gene.

Our next step was separating the genes whose expression was positively and negatively correlated with age. Intersections between gene lists from 12 datasets were performed and only those genes that correlated with age and had the same correlation sign in at least 11 out of 12 datasets were selected for further analysis. This resulted in 76 genes that were positively correlated with age and 170 that were negatively correlated with age (Figure 2). Furthermore, we applied Stouffer's method for combining *p*-values to calculate the significance of the correlation coefficients. Only 38 out of 76 genes were significantly (Stouffer's combined *p*-value < 0.05) positively correlated with age, and 92 out of 170 genes appeared to be significantly negatively correlated with age.

Survival analysis

We performed a survival analysis for the genes that were significantly correlated with age, using the TCGA-GBM [29] dataset. First, we divided patients into three cohorts: young (< 45 years), middle-aged (from 45 to 60 years) and senior (> 60 years). Then for each cohort, we performed survival analysis according to the expression level of the analyzed gene. 16 out of 38 significantly positively correlated genes and 22 out of 92 significantly negatively correlated genes were able to stratify patients by survival according to their expression level (Figure 3). It is important to mention that most genes were able to significantly stratify the cohort of young patients.

Target selection strategies guided by AI-powered PandaOmics TargetID engine

Finally, we determined three main strategies for target identification and utilized the PandaOmics TargetID engine in order to prioritize the best target hypotheses based on predictive AI-scores and additional assessment in terms of druggability, safety and brain tissue-specific expression.

Strategy 1: Expression and survival analysis of positively-correlated with age genes

The first strategy involved finding targets that met the criteria of being positively correlated with age and which high expression is significantly associated with a worse survival rate. This approach assumed that inhibiting such targets could bring benefits to GBM patients. We were able to find a set of 16 genes that

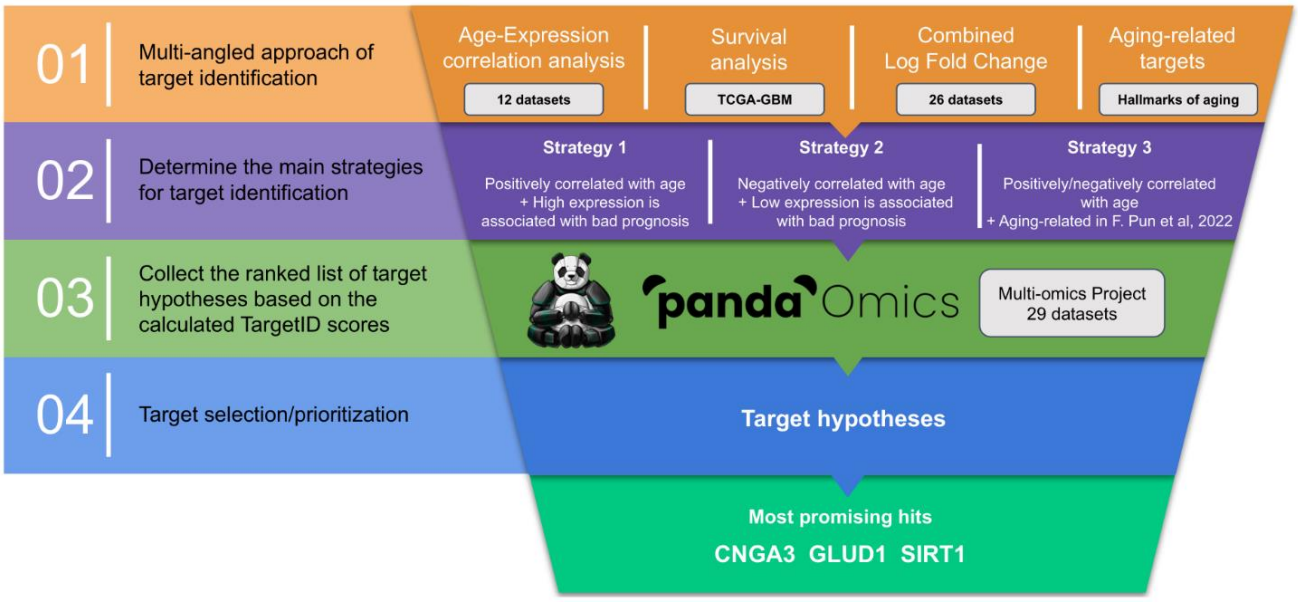


Figure 1. Overall workflow of the study. Current pipeline is designed to prioritize dual-purpose therapeutic targets by combining several data modalities following three distinct strategies of target identification. Potential target hypotheses are ranked using AI-driven scores obtained via PandaOmics TargetID engine and information regarding the combined expression, druggability, safety, novelty and accessibility by small molecules.

satisfy both requirements and passed those to the PandaOmics TargetID engine. After manual curation of the results, we nominated *CNGA3* as a putative target. Although *CNGA3* did not top the list of the most promising targets based on PandaOmics predictions (Figure 4A), it was highly ranked by the Expression score (score value 0.91). In fact, *CNGA3* gene was significantly upregulated in 16 out of 26 expression datasets and had a positive combined log fold change of 1.47 (FDR corrected *p*-value < 0.01, Supplementary Figure 1A). Additional arguments in favor of *CNGA3* choice were its high accessibility by small molecules, specific expression in brain tissue and absence of the safety red flags.

Strategy 2: Expression and survival analysis of negatively-correlated with age genes

The second strategy represented an inverted approach: we sought genes that were negatively correlated with age and whose low expression was associated with bad prognosis, in an attempt to find a target whose activation could benefit GBM patients. A set of 20

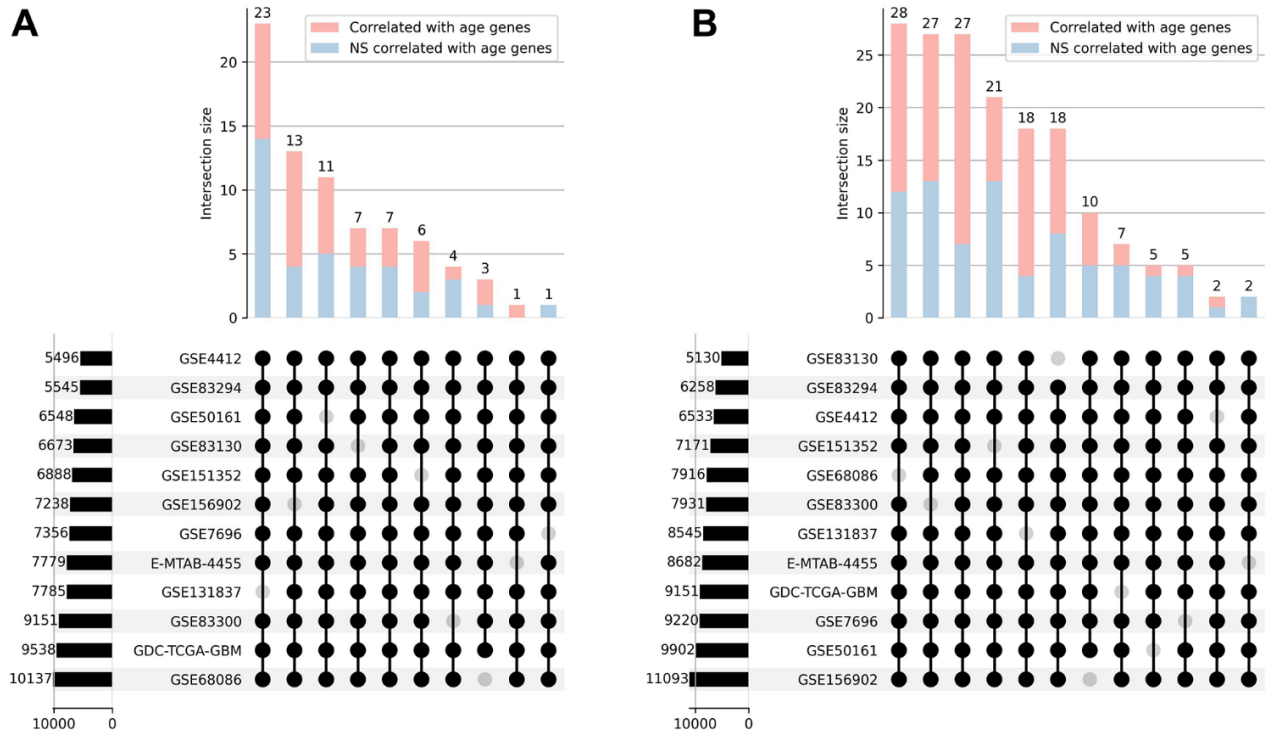


Figure 2. Correlation analysis. UpSet plots [30] representing the overlap of positively (**A**) and negatively (**B**) correlated with age genes across 12 transcriptomic datasets. The combination matrix identifies the intersections, while the bars on top represent the size of each intersection divided into significantly and not significantly (NS) correlated genes. Bars on the left depict the overall amount of correlated genes in each dataset.

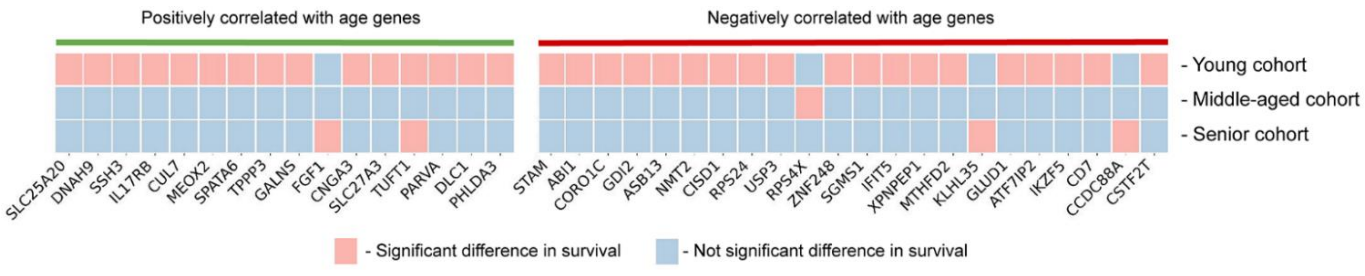


Figure 3. Survival analysis results. Each of the genes that is significantly correlated with age (n = 38) was tested for significant or not significant difference in survival rates with respect to the high and low expression levels in young, middle-aged and senior patient cohorts from TCGA-GBM dataset.

genes matching these conditions was analyzed and *GLUD1* was nominated as a putative target since it had the highest rank in PandaOmics TargetID analysis (Figure 4B). Additionally, we observed significantly decreased *GLUD1* levels in 15 GBM datasets which were reflected in a negative combined log fold change of -0.68 (FDR corrected p-value < 0.01, Supplementary Figure 1B).

Strategy 3: Intersecting age-correlated GBM genes with potential aging targets

The third strategy was to investigate whether the age-correlated genes from GBM patients identified in this study are related to aging in general. For this, we intersected our set of 130 age-correlated genes with the list of previously identified dual-purpose disease and age-associated targets associated with the Hallmarks of aging predicted by Pun et al. [12]. The authors performed a comprehensive analysis using a variety of target identification and prioritization techniques offered by the PandaOmics platform and proposed a list of promising aging-associated targets that may be used for drug discovery. Five genes were identified in the intersection and all of them were identified as negatively correlated with age in GBM patients, namely *MTHFD2, DDX21, KDM7A, GALNT1* and *SIRT1*. The latter was nominated as a putative target following PandaOmics TargetID ranking (Figure 4C).

DISCUSSION

Recent advances in the application of AI-powered algorithms in the target discovery field has proven to be a robust and viable method, which was demonstrated in several recent studies focused on cancer and aging research [12, 31]. However, the search for dual-purpose therapeutic targets implicated in both conditions simultaneously is still a challenge. Below we give the assessment of the targets revealed by a multi-angled AI-guided approach of target identification based on the previously published data.

CNGA3

Among genes that are significantly positively correlated with age in GBM patients and whose high expression is associated with worse survival, we selected *CNGA3*.

CNGA3 codes for an ion channel, which belongs to the cyclic nucleotide-gated cation channel family. *CNGA3* is involved in visual signal transduction and is essential for the generation of light-evoked electrical responses in photosensitive cones, located in the retina of the eye.

Our results are consistent with the research of Pollak et al. who showed that high expression of *CNGA3* was significantly associated with reduced median survival in GBM patients. By contrast, when the authors examined gene expression patterns in the Ivy Glioblastoma Atlas Project (Ivy GAP) database, and considered expression restricted to the central solid tumor region, high *CNGA3* was associated with increased, rather than decreased, survival. The authors suggest that this discrepancy may reflect the complexity of ion channel functions in GBM, as well as differences in sample composition between databases and bias introduced by restricting the Ivy GAP analysis to the central tumor region [32].

In another study [33], *CNGA3* was mentioned among the genes whose overexpression was associated with worse GBM patients survival and provides a predictive value of nitrosoureas treatment resistance.

In summary, *CNGA3* may play a dual role through its effects on GBM development and progression as well as aging. Further investigation is needed to confirm our hypothesis and explore specific molecular mechanisms of *CNGA3* involvement in both processes.

GLUD1

In the group of genes significantly negatively correlated with age in GBM patients and low expression is associated with bad prognosis, the *GLUD1* gene had the highest rank in our PandaOmics TargetID analysis.

GLUD1 encodes mitochondrial glutamate dehydrogenase 1 - a mitochondrial matrix enzyme which catalyzes the conversion of L-glutamate into alpha-ketoglutarate. In the nervous tissue, *GLUD1* is involved in healthy learning and memory creation by increasing the turnover of glutamate, an excitatory neurotransmitter. Distorted *GLUD1* function plays a role in several psychiatric and neurological disorders [34].

Several studies focusing on glutamate metabolism changes during aging in the brain demonstrated that there is a gradual rise in extracellular glutamate in the brain and an increase in the sensitivity of certain neurons to the cytotoxic effects of glutamate during aging [35, 36]. The decline in neuronal function during aging may result from increases in extracellular glutamate, glutamate-induced neurotoxicity, and altered mitochondrial metabolism [37].

In another study, Franco et al. investigated the protein expression profile of the key regulators of glutaminolysis, including *GLUD1*, in a cohort of astrocytomas of different malignancy grades and non-neoplastic brain samples. *GLUD1* expression was shown to be downregulated in GBM, and upregulated in lower grades of astrocytoma (AGII-AGIII). Significant

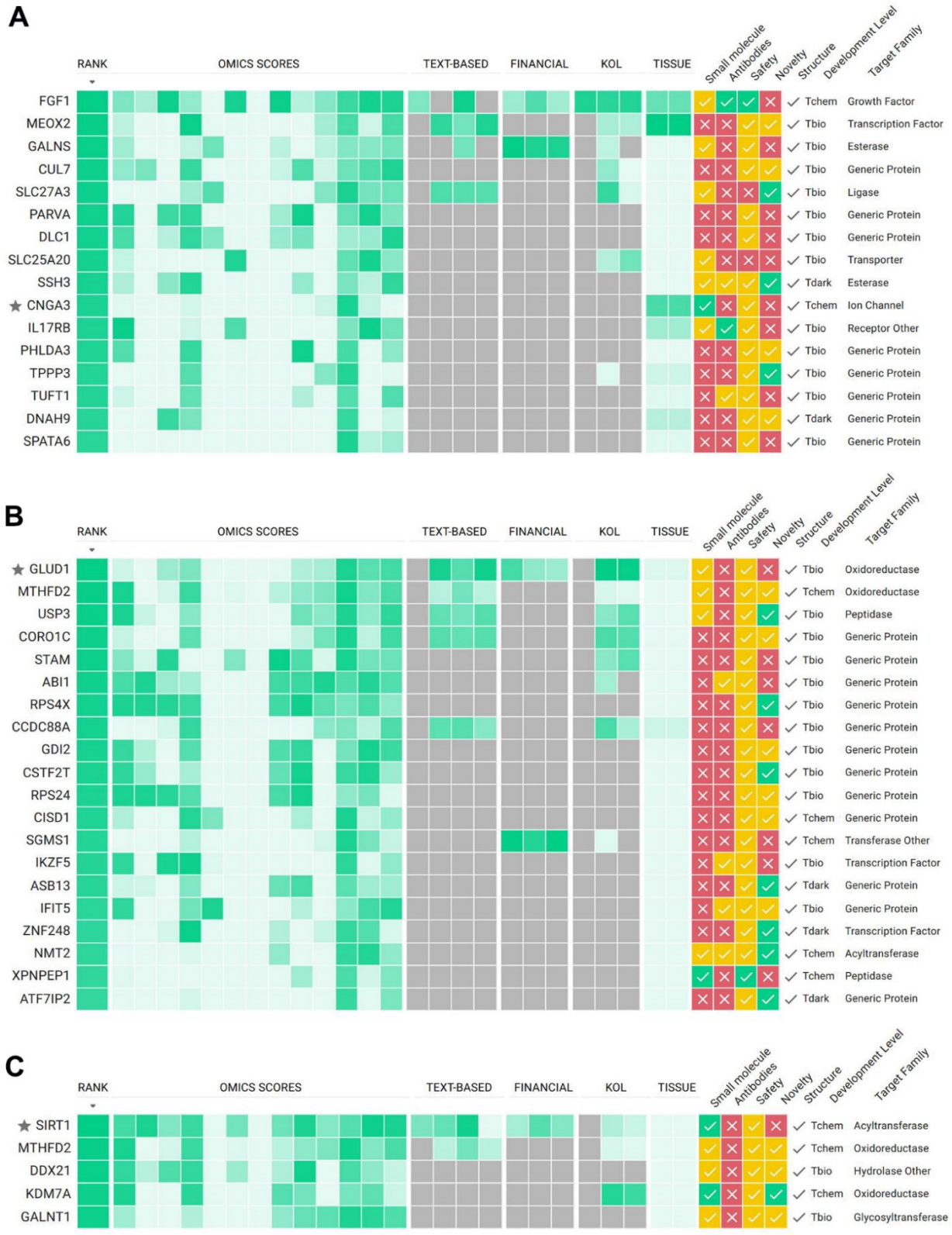


Figure 4. PandaOmics TargetID scoring approach for potential targets selected according to strategy 1 (**A**), strategy 2 (**B**) and strategy 3 (**C**). Target hypotheses are ranked according to the scores obtained from different AI-powered predictive models: omics-, text-, key opinion leaders (KOLs) and funding- based. For each target additional information on tissue specific expression, accessibility by small molecules and antibodies, safety, novelty, structure availability, development level and protein family are provided. For a detailed description of all scores and filters see Materials and Methods section, as well as the user manual at https://insilico.com/pandaomics/help.

low *GLUD1* protein levels were observed in the mesenchymal subtype of GBM. The authors suggested that the downregulation of *GLUD1* in GBM increased the source of glutamate for glutathione synthesis and enhanced tumor cell fitness due to increased anti-oxidative capacity [38].

SIRT1

SIRT1 is among the most studied genes in aging research. *SIRT1* was identified as a target connected to all hallmarks of aging, due to its wide range of interactions with aging-associated pathways [12].

SIRT1 encodes the NAD-dependent protein deacetylase sirtuin-1 that links transcriptional regulation directly to intracellular energetics and participates in the coordination of several separated cellular functions such as the cell cycle, response to DNA damage, metabolism, apoptosis and autophagy.

The significance of the sirtuin pathways in longevity has been extensively reported [39]. *SIRT1* activity was found to be attenuated in the cerebellum during aging, leading to alterations of the epigenetic landscape, thereby changing gene expression that interferes with motor function [40]. In addition, activation of *SIRT1* suppressed aging by ensuring telomere integrity [41], antagonizing oxidative stress [42], regulating nutrient signaling by inhibiting the mTOR pathway [43] and mitochondrial unfolded protein response [44].

SIRT1 is extensively studied in the context of cancer as well. Accumulating evidence has recently revealed that *SIRT1* may act as a tumor suppressor in several types of cancer, thus, activating *SIRT1* would represent a possible therapeutic strategy. It was recently identified as a key prognostic factor in brain cancer. A small-molecule activator of *SIRT1* showed a therapeutic potential on GBM *in vitro* and *in vivo* by inducing autophagy and mitophagy [45].

CONCLUSIONS

Our study provides an example of a pipeline development intended to identify dual-purpose therapeutic targets. This is achieved by using the synergy of several data modalities and *in silico*-based approaches. In this regard, the application of AI-powered algorithms, such as PandaOmics, may accelerate subsequent gene target discovery not only for GBM but for a broader range of age-associated diseases. Through the three selected strategies as well as combining the GBM correlation analysis with survival analysis and AI-proposed GBM targets, we identified three potential therapeutic targets: *CNGA3*, *GLUD1*, and *SIRT1*, which we propose to investigate in further studies. The next steps towards implementation of the identified therapeutic targets into the clinic would involve a generation of small molecules and their optimisation with further validation and preclinical testing to determine their safety, efficacy, and potential side effects.

MATERIALS AND METHODS

Data collection

Gene expression data originally from Gene Expression Omnibus, ArrayExpress, and PRIDE were collected in PandaOmics, an AI-driven target discovery platform. Twenty-five transcriptomics (GSE86202, GSE151352, GSE59612, GSE68086, GSE156902, GDC-TCGA-GBM, GSE103227, GSE90886, GSE22866, GSE50161, GSE4290, GSE108474, GSE10878, GSE90598, E-MTAB-3892, GSE83130, GSE68848, GSE153746, GSE42656, GSE7696, GSE119102, GSE65626, GSE72269, GSE13276, GSE15824), 3 methylations (GSE60274, TCGA-GBM, GSE123678) and 1 proteomics (PXD017943) datasets with a total number of 2,627 samples (case = 2,027, control = 600) were analyzed. All omics datasets were pre-processed according to the PandaOmics pipeline, which automatically defines data type and normalizes the data for further analysis. Upper-quartile normalization and log2-transformation were applied for all transcriptomics datasets used. All named datasets were further used for target prioritization and hit identification using PandaOmics TargetID approach (see "PandaOmics TargetID platform for target prioritization" section). Collected transcriptomics and proteomics datasets were used for differential expression analysis (see differential expression analysis and combined log-fold changes section). Finally, 12 out of 25 transcriptomics datasets (GSE156902, GSE68086, GSE83130, GSE50161, GSE151352, GDC-TCGA-GBM, GSE7696, GSE4412, GSE83294, E-MTAB-4455, GSE131837, GSE83300) were used to conduct a correlation analysis (see age-correlation analysis section) as age metadata were available only for these datasets.

Differential expression analysis and combined log-fold changes

Differential expression analysis was performed in PandaOmics using the *limma* R package. Each dataset has been processed according to *limma* standard protocols. Obtained gene-wise *p*-values were corrected by the Benjamini-Hochberg procedure. Combined log-fold changes (LFC) between transcriptomics and proteomics datasets were calculated in the meta-analysis section of PandaOmics. The meta-analysis section in PandaOmics enabled us to calculate combined LFC values and Q-

values across all datasets used for the analysis, using minmax normalization for LFC values and Stouffer's method combining *p*-values with further FDR correction.

Age-expression correlation analysis

Normalized gene expression matrixes along with patients' age metadata were collected from PandaOmics for 12 transcriptomics datasets. Spearman's rank correlation coefficients between gene expression and age, as well as corresponding *p*-values, were calculated for each gene separately for each dataset using only case samples. Lists with positively and negatively correlated genes were analyzed and plotted independently using the *upsetplot* python package with min_degree = 11. In order to calculate the significance of the correlation coefficients across all datasets, Stouffer's method was applied for previously obtained Spearman's correlation *p*-values. Finally, 2 lists of genes significantly correlated with age (Stouffer's combined *p*-value < 0.05) were obtained.

Survival analysis

Survival analysis was performed for the genes that were significantly correlated with age, obtained from PandaOmics GDC-TCGA-GBM dataset. Patients were divided into three cohorts: young (< 45 years), middle-aged (from 45 to 60 years) and senior (> 60 years) and for each cohort survival analysis was performed. Briefly, survival analysis was performed using the Kaplan-MeierFitter function from *lifelines* python package. The Median function was applied to normalize gene expression data and the median value for each gene of interest was used as a threshold for patients' stratification. Patients with the expression value of the gene of interest ≥ or < than the median value were considered as patients with "high" or "low" expression of a particular gene, respectively. Log-rank test was used to calculate the statistical significance. The probability of survival outcome for each cohort was plotted on a heatmap using *seaborn* python package and colored as red if there was a significant difference in survival between patients with high and low expression of a gene, and colored blue if there was no significant difference.

PandaOmics TargetID platform for target prioritization

The *in silico*-based PandaOmics TargetID approach was performed on all collected omics datasets, including transcriptomics, methylomics and proteomics, to prioritize and identify the most promising GBM therapeutic targets from the lists of genes obtained through different strategies. This approach is based on combining multiple ranking scores derived from text and omics data. Text-based scores represent how strongly a particular target is associated with a disease based on sources like scientific publications, grants, patents, clinical trials and key opinion leaders. Omics scores are based on differential expression, GWAS studies, somatic and germline mutations, interactome topology, signaling pathway perturbation analysis algorithms, knockout/overexpression experiments and more omics-data sources and, thus, represent the target-disease association based on molecular connections between the proposed target and disease of interest. Regardless of the methodology, all models output ranked lists of target hypotheses. The combination of described scores leads to a ranked list of potential targets for a disease that can be filtered out based on their novelty, safety, accessibility by molecule or antibodies and availability of PDB structure. Detailed descriptions of all scores and filters can be found in PandaOmics' User Manual (https://insilico.com/pandaomics/help). For the current study, lists with potential targets derived from strategies 1, 2 and 3 were combined and passed into TargetID. To identify the most promising hits, additional filtering was performed on PandaOmics TargetID page. After filtering, the list with the most promising hits was obtained.

Abbreviations

AI: artificial intelligence; CNGA3: cyclic nucleotide gated channel subunit alpha 3; EGFR: epidermal growth factor receptor; Ivy GAP: Ivy Glioblastoma Atlas Project; IDH1: isocitrate dehydrogenase; KOLs: key opinion leaders; Log FC: logarithmic fold-changes; GBM: glioblastoma multiforme; GLUD1: glutamate dehydrogenase 1; SIRT1: sirtuin 1; TTF: tumor-treating fields.

AUTHOR CONTRIBUTIONS

AO, ZH and CR provided conceptual design, performed data analysis, result interpretation and manuscript writing. AS prepared data curation, methodology, and result interpretation. AV performed statistical data analysis and visualization. MD, SK, and OS were responsible for the omics data aggregation and curation. FP, GL and HL reviewed the manuscript. IO performed methodology, software and manuscript writing. MK performed data analysis, participated in result interpretation and visualization, overall project administration and manuscript writing. AA and AZ provided conceptualization, reviewed the manuscript, provided resources, and supervised the project. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

AS, AV, MD, SK, OS, FP, GL, HL, IO, AA, MK and AZ are affiliated with Insilico Medicine, a commercial

company developing AI solutions for aging research, drug discovery, and longevity medicine.

FUNDING

This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Thakkar JP, Dolecek TA, Horbinski C, Ostrom QT, Lightner DD, Barnholtz-Sloan JS, Villano JL. Epidemiologic and molecular prognostic review of glioblastoma. Cancer Epidemiol Biomarkers Prev. 2014; 23:1985–96. https://doi.org/10.1158/1055-9965.EPI-14-0275 PMID:25053711

2. Kleihues P, Ohgaki H. Primary and secondary glioblastomas: from concept to clinical diagnosis. Neuro Oncol. 1999; 1:44–51. https://doi.org/10.1093/neuonc/1.1.44 PMID:11550301

3. Li Z, Wang Y, Yu J, Guo Y, Zhang Q. Age groups related glioblastoma study based on radiomics approach. Comput Assist Surg (Abingdon). 2017; 22:18–25. https://doi.org/10.1080/24699322.2017.1378722 PMID:28914549

4. Li R, Li H, Yan W, Yang P, Bao Z, Zhang C, Jiang T, You Y. Genetic and clinical characteristics of primary and secondary glioblastoma is associated with differential molecular subtype distribution. Oncotarget. 2015; 6:7318–24. https://doi.org/10.18632/oncotarget.3440 PMID:25821160

5. Fernandes C, Costa A, Osório L, Lago RC, Linhares P, Carvalho B, Caeiro C. Current Standards of Care in Glioblastoma Therapy. In: De Vleeschouwer S, editor. Glioblastoma. Brisbane (AU): Codon Publications; 2017. Chapter 11. PMID:29251860

6. Rominiyi O, Vanderlinden A, Clenton SJ, Bridgewater C, Al-Tamimi Y, Collis SJ. Tumour treating fields therapy for glioblastoma: current advances and future directions. Br J Cancer. 2021; 124:697–709. https://doi.org/10.1038/s41416-020-01136-5 PMID:33144698

7. Huang J, Liu F, Liu Z, Tang H, Wu H, Gong Q, Chen J. Immune Checkpoint in Glioblastoma: Promising and Challenging. Front Pharmacol. 2017; 8:242. https://doi.org/10.3389/fphar.2017.00242 PMID:28536525

8. Tykocki T, Eltayeb M. Ten-year survival in glioblastoma. A systematic review. J Clin Neurosci. 2018; 54:7–13. https://doi.org/10.1016/j.jocn.2018.05.002 PMID:29801989

9. Binder ZA, Thorne AH, Bakas S, Wileyto EP, Bilello M, Akbari H, Rathore S, Ha SM, Zhang L, Ferguson CJ, Dahiya S, Bi WL, Reardon DA, et al. Epidermal Growth Factor Receptor Extracellular Domain Mutations in Glioblastoma Present Opportunities for Clinical Imaging and Therapeutic Development. Cancer Cell. 2018; 34:163–77.e7. https://doi.org/10.1016/j.ccell.2018.06.006 PMID:29990498

10. Karpel-Massler G, Nguyen TTT, Shang E, Siegelin MD. Novel IDH1-Targeted Glioma Therapies. CNS Drugs. 2019; 33:1155–66. https://doi.org/10.1007/s40263-019-00684-6 PMID:31768950

11. Saraf R, Agah S, Datta A, Jiang X. Drug target ranking for glioblastoma multiforme. BMC Biomed Eng. 2021; 3:7. https://doi.org/10.1186/s42490-021-00052-w PMID:33902757

12. Pun FW, Leung GHD, Leung HW, Liu BHM, Long X, Ozerov IV, Wang J, Ren F, Aliper A, Izumchenko E, Moskalev A, de Magalhães JP, Zhavoronkov A. Hallmarks of aging-based dual-purpose disease and age-associated targets predicted using PandaOmics AI-powered discovery engine. Aging (Albany NY). 2022; 14:2475–506. https://doi.org/10.18632/aging.203960 PMID:35347083

13. Ozerov IV, Lezhnina KV, Izumchenko E, Artemov AV, Medintsev S, Vanhaelen Q, Aliper A, Vijg J, Osipov AN, Labat I, West MD, Buzdin A, Cantor CR, et al. *In silico* Pathway Activation Network Decomposition Analysis (iPANDA) as a method for biomarker development. Nat Commun. 2016; 7:13427. https://doi.org/10.1038/ncomms13427 PMID:27848968

14. Broner EC, Trujillo JA, Korzinkin M, Subbannayya T, Agrawal N, Ozerov IV, Zhavoronkov A, Rooper L, Kotlov N, Shen L, Pearson AT, Rosenberg AJ, Savage PA, et al. Doublecortin-Like Kinase 1 (DCLK1) Is a Novel NOTCH Pathway Signaling Regulator in Head and Neck Squamous Cell Carcinoma. Front Oncol. 2021; 11:677051. https://doi.org/10.3389/fonc.2021.677051 PMID:34336664

15. Stamatas GN, Wu J, Pappas A, Mirmirani P, McCormick TS, Cooper KD, Consolo M, Schastnaya J, Ozerov IV, Aliper A, Zhavoronkov A. An analysis of gene expression data involving examination of signaling pathways

activation reveals new insights into the mechanism of action of minoxidil topical foam in men with androgenetic alopecia. Cell Cycle. 2017; 16:1578–84. https://doi.org/10.1080/15384101.2017.1327492 PMID:28594262

16. Pasteuning-Vuhman S, Boertje-van der Meulen JW, van Putten M, Overzier M, Ten Dijke P, Kiełbasa SM, Arindrarto W, Wolterbeek R, Lezhnina KV, Ozerov IV, Aliper AM, Hoogaars WM, Aartsma-Rus A, Loomans CJ. New function of the myostatin/activin type I receptor (ALK4) as a mediator of muscle atrophy and muscle regeneration. FASEB J. 2017; 31:238–55. https://doi.org/10.1096/fj.201600675R PMID:27733450

17. Solanki HS, Raja R, Zhavoronkov A, Ozerov IV, Artemov AV, Advani J, Radhakrishnan A, Babu N, Puttamallesh VN, Syed N, Nanjappa V, Subbannayya T, Sahasrabuddhe NA, et al. Correction: Targeting focal adhesion kinase overcomes erlotinib resistance in smoke induced lung cancer by altering phosphorylation of epidermal growth factor receptor. Oncoscience. 2021; 8:108–9. https://doi.org/10.18632/oncoscience.546 PMID:34589558

18. Aliper A, Plis S, Artemov A, Ulloa A, Mamoshina P, Zhavoronkov A. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. Mol Pharm. 2016; 13:2524–30. https://doi.org/10.1021/acs.molpharmaceut.6b00248 PMID:27200455

19. Shayakhmetov R, Kuznetsov M, Zhebrak A, Kadurin A, Nikolenko S, Aliper A, Polykovskiy D. Molecular Generation for Desired Transcriptome Changes With Adversarial Autoencoders. Front Pharmacol. 2020; 11:269. https://doi.org/10.3389/fphar.2020.00269 PMID:32362822

20. Moore JH, Raghavachari N, and Workshop Speakers. Artificial Intelligence Based Approaches to Identify Molecular Determinants of Exceptional Health and Life Span-An Interdisciplinary Workshop at the National Institute on Aging. Front Artif Intell. 2019; 2:12. https://doi.org/10.3389/frai.2019.00012 PMID:33733101

21. Kadurin A, Nikolenko S, Khrabrov K, Aliper A, Zhavoronkov A. druGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties *in Silico*. Mol Pharm. 2017; 14:3098–104. https://doi.org/10.1021/acs.molpharmaceut.7b00346 PMID:28703000

22. Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of Deep Learning in Biomedicine. Mol Pharm. 2016; 13:1445–54. https://doi.org/10.1021/acs.molpharmaceut.5b00982 PMID:27007977

23. Putin E, Asadulaev A, Ivanenkov Y, Aladinskiy V, Sanchez-Lengeling B, Aspuru-Guzik A, Zhavoronkov A. Reinforced Adversarial Neural Computer for de Novo Molecular Design. J Chem Inf Model. 2018; 58:1194–204. https://doi.org/10.1021/acs.jcim.7b00690 PMID:29762023

24. Polykovskiy D, Zhebrak A, Vetrov D, Ivanenkov Y, Aladinskiy V, Mamoshina P, Bozdaganyan M, Aliper A, Zhavoronkov A, Kadurin A. Entangled Conditional Adversarial Autoencoder for de Novo Drug Discovery. Mol Pharm. 2018; 15:4398–405. https://doi.org/10.1021/acs.molpharmaceut.8b00839 PMID:30180591

25. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, Terentiev VA, Polykovskiy DA, Kuznetsov MD, Asadulaev A, Volkov Y, Zholus A, Shayakhmetov RR, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nat Biotechnol. 2019; 37:1038–40. https://doi.org/10.1038/s41587-019-0224-x PMID:31477924

26. Ren F, Ding X, Zheng M, Korzinkin M, Cai X, Zhu W, Mantsyzov A, Aliper A, Aladinskiy V, Cao Z, Kong S, Long X, Man Liu BH, et al. AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor. Chem Sci. 2023; 14:1443–52. https://doi.org/10.1039/d2sc05709c PMID:36794205

27. Pomatto LCD, Davies KJA. The role of declining adaptive homeostasis in ageing. J Physiol. 2017; 595:7275–309. https://doi.org/10.1113/JP275072 PMID:29028112

28. Kim M, Ladomersky E, Mozny A, Kocherginsky M, O'Shea K, Reinstein ZZ, Zhai L, Bell A, Lauing KL, Bollu L, Rabin E, Dixit K, Kumthekar P, et al. Glioblastoma as an age-related neurological disorder in adults. Neurooncol Adv. 2021; 3:vdab125. https://doi.org/10.1093/noajnl/vdab125 PMID:34647022

29. Brennan CW, Verhaak RGW, McKenna A, Campos B, Noushmehr H, Salama SR, Zheng S, Chakravarty D, Sanborn JZ, Berman SH, Beroukhim R, Bernard B, Wu CJ, et al, and TCGA Research Network. The somatic genomic landscape of glioblastoma. Cell. 2013; 155:462–77. https://doi.org/10.1016/j.cell.2013.09.034 PMID:24120142

30. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. IEEE Trans Vis Comput Graph. 2014; 20:1983–92. https://doi.org/10.1109/TVCG.2014.2346248 PMID:26356912

31. Mkrtchyan GV, Veviorskiy A, Izumchenko E, Shneyderman A, Pun FW, Ozerov IV, Aliper A, Zhavoronkov A, Scheibye-Knudsen M. High-confidence cancer patient stratification through multiomics investigation of DNA repair disorders. Cell Death Dis. 2022; 13:999. https://doi.org/10.1038/s41419-022-05437-w PMID:36435816

32. Pollak J, Rai KG, Funk CC, Arora S, Lee E, Zhu J, Price ND, Paddison PJ, Ramirez JM, Rostomily RC. Ion channel expression patterns in glioblastoma stem cells with functional and therapeutic implications for malignancy. PLoS One. 2017; 12:e0172884. https://doi.org/10.1371/journal.pone.0172884 PMID:28264064

33. Menyhárt O, Fekete JT, Győrffy B. Gene expression-based biomarkers designating glioblastomas resistant to multiple treatment strategies. Carcinogenesis. 2021; 42:804–13. https://doi.org/10.1093/carcin/bgab024 PMID:33754151

34. Lander SS, Chornyy S, Safory H, Gross A, Wolosker H, Gaisler-Salomon I. Glutamate dehydrogenase deficiency disrupts glutamate homeostasis in hippocampus and prefrontal cortex and impairs recognition memory. Genes Brain Behav. 2020; 19:e12636. https://doi.org/10.1111/gbb.12636 PMID:31898404

35. Nickell J, Pomerleau F, Allen J, Gerhardt GA. Age-related changes in the dynamics of potassium-evoked L-glutamate release in the striatum of Fischer 344 rats. J Neural Transm (Vienna). 2005; 112:87–96. https://doi.org/10.1007/s00702-004-0151-x PMID:15599607

36. Brewer GJ. Neuronal plasticity and stressor toxicity during aging. Exp Gerontol. 2000; 35:1165–83. https://doi.org/10.1016/s0531-5565(00)00121-2 PMID:11113600

37. Choi IY, Lee P, Wang WT, Hui D, Wang X, Brooks WM, Michaelis EK. Metabolism changes during aging in the hippocampus and striatum of glud1 (glutamate dehydrogenase 1) transgenic mice. Neurochem Res. 2014; 39:446–55. https://doi.org/10.1007/s11064-014-1239-9 PMID:24442550

38. Moreira Franco YE, Alves MJ, Uno M, Moretti IF, Trombetta-Lima M, de Siqueira Santos S, Dos Santos AF, Arini GS, Baptista MS, Lerario AM, Oba-Shinjo SM, Marie SK. Glutaminolysis dynamics during astrocytoma progression correlates with tumor aggressiveness. Cancer Metab. 2021; 9:18. https://doi.org/10.1186/s40170-021-00255-8 PMID:33910646

39. Moskalev AA, Aliper AM, Smit-McBride Z, Buzdin A, Zhavoronkov A. Genetics and epigenetics of aging and longevity. Cell Cycle. 2014; 13:1063–77. https://doi.org/10.4161/cc.28433 PMID:24603410

40. Marton O, Koltai E, Nyakas C, Bakonyi T, Zenteno-Savin T, Kumagai S, Goto S, Radak Z. Aging and exercise affect the level of protein acetylation and SIRT1 activity in cerebellum of male rats. Biogerontology. 2010; 11:679–86. https://doi.org/10.1007/s10522-010-9279-2 PMID:20467811

41. Osum M, Serakinci N. Impact of circadian disruption on health; SIRT1 and Telomeres. DNA Repair (Amst). 2020; 96:102993. https://doi.org/10.1016/j.dnarep.2020.102993 PMID:33038659

42. Meng T, Qin W, Liu B. SIRT1 Antagonizes Oxidative Stress in Diabetic Vascular Complication. Front Endocrinol (Lausanne). 2020; 11:568861. https://doi.org/10.3389/fendo.2020.568861 PMID:33304318

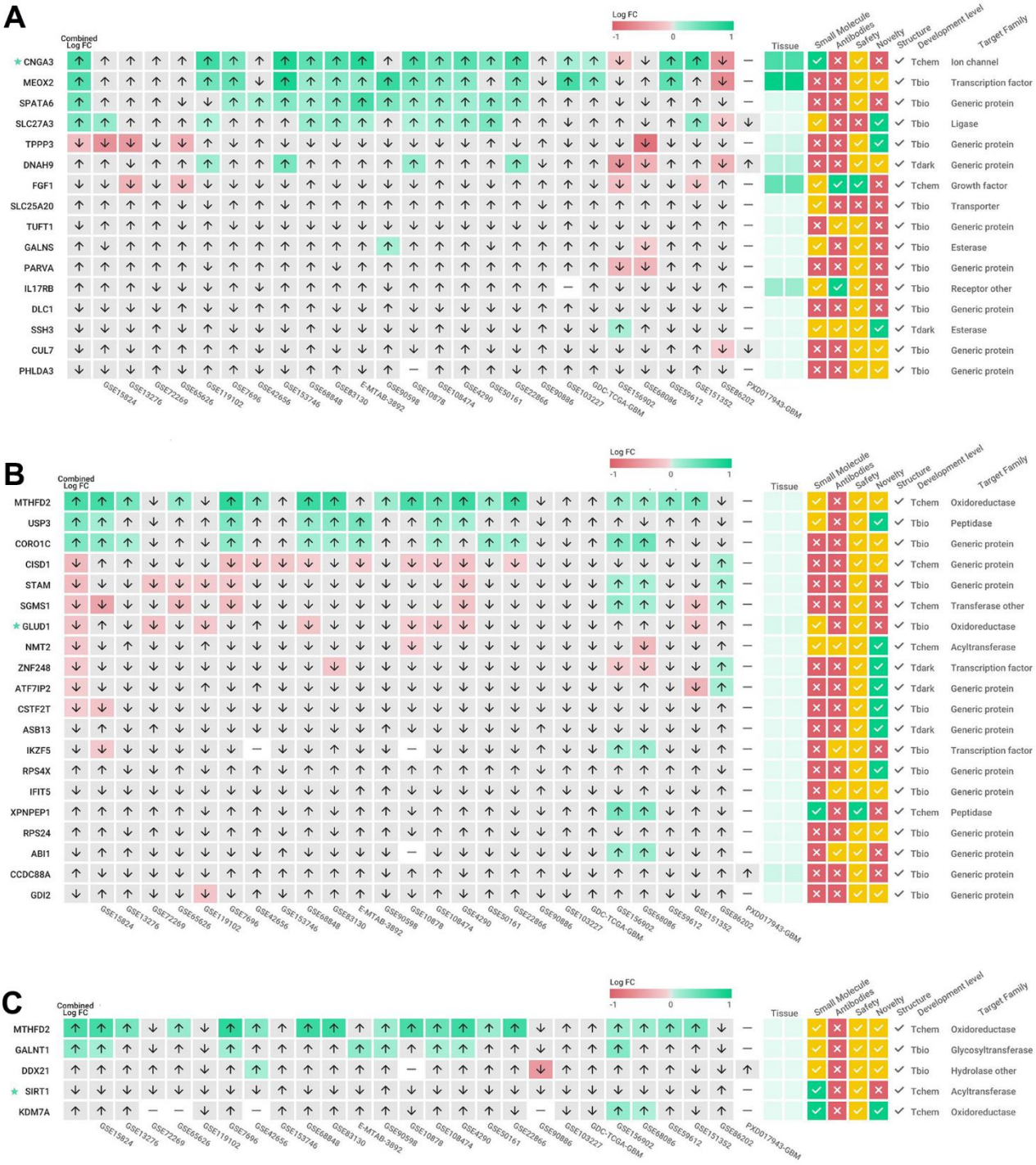
43. Ghosh HS, McBurney M, Robbins PD. SIRT1 negatively regulates the mammalian target of rapamycin. PLoS One. 2010; 5:e9199. https://doi.org/10.1371/journal.pone.0009199 PMID:20169165

44. Lin YF, Sam J, Evans T. Sirt1 promotes tissue regeneration in zebrafish through regulating the mitochondrial unfolded protein response. iScience. 2021; 24:103118. https://doi.org/10.1016/j.isci.2021.103118 PMID:34622167

45. Yao ZQ, Zhang X, Zhen Y, He XY, Zhao S, Li XF, Yang B, Gao F, Guo FY, Fu L, Liu XZ, Duan CZ. A novel small-molecule activator of Sirtuin-1 induces autophagic cell death/mitophagy as a potential therapeutic strategy in glioblastoma. Cell Death Dis. 2018; 9:767. https://doi.org/10.1038/s41419-018-0799-z PMID:29991742

SUPPLEMENTARY MATERIALS

Supplementary Figure



Supplementary Figure 1. Combined expression analysis for the genes obtained according to strategy 1 (**A**), strategy 2 (**B**) and strategy 3 (**C**). Combined logarithmic fold-changes (Log FC) and Q-values across all gene expression datasets are obtained using minmax normalization for Log FC values and Stouffer's method combining p-values with further FDR correction. Significant combined Log FC values represent both magnitude and direction of gene expression changes between disease and control across multiple datasets.

Supplementary Tables

Supplementary Table 1. Glioblastoma multiforme datasets used for PandaOmics analysis.

Dataset ID	Type	Year	Organism	Tissue	Total samples	Cases	Controls
TCGA-GBM	methylation	2015	Homo sapiens	brain	155	153	2
GSE60274	methylation	2014	Homo sapiens	brain	77	68	5
GSE123678	methylation	2018	Homo sapiens	brain	78	59	8
GSE83130	microarray	2016	Homo sapiens	brain	701	530	20
GSE7696	microarray	2007	Homo sapiens	brain	84	80	4
GSE42656	microarray	2013	Homo sapiens	brain	73	5	16
E-MTAB-3892	microarray	2016	Homo sapiens	brain	179	11	9
GSE15824	microarray	2011	Homo sapiens	brain	45	15	2
GSE90598	microarray	2017	Homo sapiens	brain	25	16	3
GSE10878	microarray	2008	Homo sapiens	brain	23	19	4
GSE108474	microarray	2018	Homo sapiens	brain	550	228	28
GSE68848	microarray	2015	Homo sapiens	brain	580	228	28
GSE4290	microarray	2006	Homo sapiens	brain	180	77	23
GSE50161	microarray	2013	Homo sapiens	brain	130	34	13
GSE22866	microarray	2011	Homo sapiens	brain	46	40	6
GSE90886	microarray	2016	Homo sapiens	brain	18	9	9
GSE13276	microarray	2009	Homo sapiens	brain	15	5	7
GSE65626	microarray	2015	Homo sapiens	brain	12	3	3
GSE103227	microarray	2017	Homo sapiens	brain	10	5	5
GSE72269	microarray	2015	Homo sapiens	brain	9	4	2
TCGA-GBM	RNA-seq	2015	Homo sapiens	brain	173	166	5
GSE59612	RNA-seq	2014	Homo sapiens	brain	92	39	17
GSE151352	RNA-seq	2020	Homo sapiens	brain	24	12	12
GSE153746	RNA-seq	2020	Homo sapiens	brain	21	6	3
GSE86202	RNA-seq	2016	Homo sapiens	brain	6	3	3
GSE119102	RNA-seq	2018	Homo sapiens	brain	6	4	2
GSE156902	RNA-seq	2020	Homo sapiens	blood	600	156	252
GSE68086	RNA-seq	2015	Homo sapiens	blood	285	40	55
PXD017943	Proteomics	2020	Homo sapiens	brain	66	12	54

Supplementary Table 2. Glioblastoma multiforme datasets used for correlation analysis.

Dataset ID	Type	Year	Organism	Tissue	Total samples	Cases	Controls	Cases with age	Controls with age	Age range cases
GSE83130	microarray	2016	Homo sapiens	brain	701	530	20	196	0	7-89
GSE7696	microarray	2007	Homo sapiens	brain	84	80	4	80	0	26-70
GSE4412	microarray	2006	Homo sapiens	brain	170	59	0	59	0	18-82
GSE83294	microarray	2016	Homo sapiens	brain	170	59	0	59	0	18-82
E-MTAB-4455	microarray	2017	Homo sapiens	brain	52	52	0	52	0	18-72
GSE131837	microarray	2019	Homo sapiens	brain	52	52	0	52	0	11-75
GSE83300	microarray	2016	Homo sapiens	brain	50	50	0	50	0	18-68
GSE50161	microarray	2013	Homo sapiens	brain	130	34	13	23	13	3-73
TCGA-GBM	RNA-seq	2015	Homo sapiens	brain	173	166	5	165	0	21-89
GSE151352	RNA-seq	2020	Homo sapiens	brain	24	12	12	12	12	17-73
GSE156902	RNA-seq	2020	Homo sapiens	blood	600	156	252	151	248	21-79
GSE68086	RNA-seq	2015	Homo sapiens	blood	285	40	55	39	46	11-84