Neuro-Oncology
25(2), 303–314, 2023 | https://doi.org/10.1093/neuonc/noac171 | Advance Access date 8 July 2022

Brain radiotoxicity-related 15CAcBRT gene expression signature predicts survival prognosis of glioblastoma patients

Jesús Reyes-González†, Francisco Barajas-Olmos†, Humberto García-Ortiz, Lorena Magraner-Pardo, Tirso Pons, Sergio Moreno, Lucinda Aguirre-Cruz, Andy Reyes-Abrahantes, Angélica Martínez-Hernández, Cecilia Contreras-Cubas, Jorge Barrios-Payan, Henry Ruiz-Garcia, Rogelio Hernandez-Pando, Alfredo Quiñones-Hinojosa, Lorena Orozco, and María del Carmen Abrahantes-Pérez

Precision Translational Oncology Laboratory, National Institute of Genomic Medicine, Mexico City, Mexico (J.R.G., A.R.A., M.C.A.P.); Immunogenomics and Metabolic Diseases Laboratory, National Institute of Genomic Medicine, Mexico City, Mexico (F.B.O., H.G.O., A.M.H., C.C.C., L.O.); Gene Function Team, The Institute of Cancer Research (ICR), London, UK (L.M-P.); Department of Immunology and Oncology, National Center for Biotechnology, Spanish National Research Council (CNB-CSIC), Madrid, Spain (T.P.); Radioneurosurgery Unit, National Institute of Neurology and Neurosurgery; Mexico City, Mexico (S.M.); Neuroendocrinology Laboratory, National Institute of Neurology and Neurosurgery; Mexico City, Mexico (L.A.C.); Department of Pathology, National Institute of Medical Sciences and Nutrition Salvador Zubiran, Mexico City, Mexico (R.H.P., J.B.P.); Department of Neurosurgery and Brain Tumor Stem Cell Research Laboratory, Mayo Clinic, Jacksonville, Florida, USA (H.J.R.G., A.Q.H.)

Corresponding Author: María del Carmen Abrahantes-Pérez, PhD, Precision Translational Oncology Laboratory, National Institute of Genomic Medicine, Periférico Sur 4809, Tlalpan, Mexico City C.P. 14610, Mexico (mcabrahantes@inmegen.gob.mx).
†These authors contributed equally to this work.

Abstract

Background. Glioblastoma is the most common and devastating primary brain cancer. Radiotherapy is standard of care; however, it is associated with brain radiation toxicity (BRT). This study used a multi-omics approach to determine whether BRT-related genes (RGs) harbor survival prognostic value and whether their encoded proteins represent novel therapeutic targets for glioblastoma.

Methods. RGs were identified through analysis of single-nucleotide variants associated with BRT (R-SNVs). Functional relationships between RGs were established using Protein-Protein Interaction networks. The influence of RGs and their functional groups on glioblastoma prognosis was evaluated using clinical samples from the Glioblastoma Bio-Discovery Portal database and validated using the Chinese Glioma Genome Atlas dataset. The identification of clusters of radiotoxic and putative pathogenic variants in proteins encoded by RGs was achieved by computational 3D structural analysis.

Results. We identified the BRT-related 15CAcBRT molecular signature with prognostic value in glioblastoma, by analysis of the COMT and APOE protein functional groups. Its external validation confirmed clinical relevance independent of age, *MGMT* promoter methylation status, and *IDH* mutation status. Interestingly, the genes *IL6*, *APOE*, and *MAOB* documented significant gene expression levels alteration, useful for drug repositioning. Biological networks associated with 15CAcBRT signature involved pathways relevant to cancer and neurodegenerative diseases. Analysis of 3D clusters of radiotoxic and putative pathogenic variants in proteins coded by RGs unveiled potential novel therapeutic targets in neuro-oncology.

Conclusions. 15CAcBRT is a BRT-related molecular signature with prognostic significance for glioblastoma patients and represents a hub for drug repositioning and development of novel therapies.

Key Points

- RGs encode a prognostic molecular signature for glioblastoma patients
- 3D protein clusters of radiotoxic and pathogenic variants are potential therapeutic targets
- Proteins encoded by RGs relate to cancer and neurodegenerative molecular pathways

Importance of the Study

Glioblastoma (GBM) is the most common and devastating primary brain cancer. Although radiotherapy is standard of care, it is associated with brain radiation toxicity (BRT). Our study unveils the interplay between BRT-related genes (RGs), GBM prognosis, and potential new therapeutic targets. We show 15CAcBRT is a novel BRT-related molecular signature that harbor prognostic value in GBM, independent of the *IDH* mutation status, *MGMT* promoter methylation status, and age. 15CAcBRT-associated biological networks involve cancer and neurodegenerative molecular pathways, which further support our findings. By analyzing networks of interactions of proteins encoded by RGs, we show that COMT-APOE cluster-specific associations related to 15CAcBRT harbor potential hubs for novel therapeutic targets and drug repositioning. Furthermore, our data revealed 3D clusters of R-SNVs and putative pathogenic variants in proteins (COMT, APOE, CYP1B1, MGMT, and POR) coded by RGs that might represent a relevant part of a neuro-oncology molecular target.

Glioblastoma (GBM) is the most common and lethal primary brain cancer in adults.[1] Despite multimodal treatment including maximal safe resection followed by radiotherapy and chemotherapy,[2] the prognosis remains poor with a median overall survival of merely 14.6 months.[3–5] Unfortunately, therapeutic radiation can sometimes induce damage to healthy brain tissue, also known as brain radiotoxicity (BRT),[6] which is associated with poorer survival in high-grade glioma patients.[7]

To date, prognostic signatures of GBM survival have been identified mostly by differential expression analysis between tumor and normal tissue,[8] based on the biological information of GBM,[9] or in association with its tumor microenvironment.[10] However, clinically relevant molecular signatures related to brain radiobiology have not been described yet in GBM. This study evaluates if BTR-related tumor transcriptomic may help provide independent prognosis of survival in patients with GBM and harbor potential novel therapeutic targets.

BRT has been linked to single-nucleotide variants (SNVs) of particular genes, designated as radiotoxic genes (RGs) in this work.[11] However, very little is known regarding the influence that these genes may have on the prognosis of GBM patients. Thus, we decided to use genes that were related to BRT as a starting, because though radiotoxicity predicts poorer overall survival in subjects with high-grade gliomas,[7] and genomic signatures linking radiotoxicity to GBM survival have not been described.

In this study, we hypothesized that RGs transcripts may generate a molecular signature with prognostic value for GBM patients, and that specific R-SNVs could serve as potential therapeutic molecular targets. To test this hypothesis, we identified R-SNVs and their RGs using a systematic review. We then used Protein-Protein Interaction (PPI) networks to establish functional relationships between these RGs. The gene expression profiles of RGs and their functional groups were evaluated as predictors of GBM prognosis using the "Glioblastoma Bio-Discovery Portal" database (GBM-BioDP, https://gbm-biodp.nci.nih.gov), from The Cancer Genome Atlas project (TCGA), identifying a robust predictive survival signature, 15CAcBRT. The prognostic value of our newly identified molecular signature was validated using the Chinese Glioma Genome Atlas (CGGA). Additionally, we also analyzed the pathways involved in the biological networks related to RGs, as well as their relationship with cancer. Furthermore, the distribution of new variants across the protein sequence and 3D structure of these candidate RGs, as well as their relevance as disease-promoting variants, were evaluated using data from Mexican population. The novel molecular signature 15CAcBRT may represent a reliable prognostic tool in future clinical practice and could be used to help design individualized therapies for GBM.

Materials and Methods

Collection of SNVs in RGs

SNVs associated with BRT identified via systematic review[12,13] were named as R-SNVs (see Supplementary Materials and Methods), and the genes carrying them were termed as RGs. The correspondence between genes and R-SNVs was verified in the Single Nucleotide Polymorphism Database (dbSNP) (https://www.ncbi.nlm.nih.gov/snp/).

Furthermore, when a variant was located between two genes, both genes were included in the RG list, generating extended RGs (E-RGs). The proteins encoded by RGs were named radiotoxic proteins (RPs). Novel variants of R-SNVs identified utilizing prospectively collected data from Mexican population[14] were also included in our analysis (see Supplementary Materials and Methods).

Computational Structural Analysis of the Novel Variants Identified in Mexican Population

We investigated the distribution of the novel variants identified in Mexican population along protein domains and 3D structure. The experimentally determined 3D structure and theoretical models were retrieved from the PDB (https://www.rcsb.org/) and ModBase (https://modbase.compbio.ucsf.edu/) databases, respectively. For more details, see Supplementary Table S1. The spectrum of novel variants in protein domains is represented as a lollipop plot. The protein isoform associated with the longer transcript of each gene was used. The position and clustering of the variants in the 3D structure of the proteins were also studied with the mutation 3D program (http://mutation3d.org), as recommended by the developers.[15] The parameters were set as follows: diameter or CL-distance = 30 Å and number of bootstrapping iterations = 10 000. The identified clusters contained at least three individual positions within spheres with a 15-30 Å diameter.

We also integrated functional annotations, such as ligand binding sites, catalytic residues, posttranslational modifications of proteins, and residues in PPI interfaces, for the variants using the Structure-PPi system.[16] Structure-PPi considers residues in physical proximity (at a 5 Å distance) to amino acid changes found in human diseases.

Construction of the PPI Networks Associated with BRT

We used the Search Tool for the Retrieval of Interacting Genes/Proteins database (STRING v.11, https://string-db.org/)[17] to construct the PPI network associated with BRT. The E-RGs were introduced as input in STRING to generate the PPI network (score ≥0.4, for *homo sapiens*) among the proteins produced by these genes. This network was called "BRT-PPI."

Topological analyses of the search node of the BRT-PPI network were performed. With use of the hub node and its edges as input, we implemented STRING again to construct the "COMT-extended-PPI network" associated with hub. The settings used were: (1) no more than 50 interactors and no more than 10 interactors in the first shell and second shell, respectively, and (2) required an interaction score ≥0.4. The COMT-extended-PPI network was then imported with the highest confidence for an interaction score ≥0.9 into Cytoscape Consortium (v3.8.0) (https://github.com/cytoscape/cytoscape/releases/3.8.0/) and clustered by the Molecular Complex Detection method.

The biological functions of each of the clusters were determined by BiNGO v 3.0.4 (https://apps.cytoscape.org/). The BRT-PPI network was extended in a similar manner to the COMT-extended-PPI network. The giant components (enrichment of PPI with more than 100 nodes) of the extended-PPI networks were generated by STRING through enrichment.

Identification of novel GBM survival prognostic profiles related to the expression of genes involved in BRT, using GBM-BioDP

Analysis of RGs expression and mRNA coding for protein clusters derived from the COMT-extended-PPI network (interaction score ≥0.9) was performed using the GBM-BioDP tool (https://gbm-biodp.nci.nih.gov) with data from the TCGA (https://tcga-sata.nci.nih.gov/tcga/). Heatmaps for visualizing the correlation matrix and Kaplan-Meier (KM) plots were also generated with GBM-BioDP. The strategy to identify molecular signatures and their prognostic relevance for GBM patients in GBM-BioDP was carried out according to Varghese et al.[18] The datasets used herein include: (1) HT_HG-U133A from Broad Institute and (2) HuEx-1_0-st-v2 from the Berkeley Laboratory. Associations between R-SNVs, radiotherapy, and clinical output could not be evaluated, because in this point our interest were the gene expression levels.

Additionally, to evaluate the impact of functional SNV on gene expression, we used the Genotype-Tissue Expression portal (https://gtexportal.org). See Supplementary Materials and Methods.

External Validation of the 15CAcBRT Signature Using the CGGA Dataset

The gene expression and clinical data were obtained from the CGGA (Set ID: mRNA_array_301) from http://www.cgga.org.cn/download.jsp. With these data, a clustering analysis was performed based on expression levels of the genes that conformed 15CAcBRT signature (15 mRNAs of COMT + APOE protein clusters associated with BRT). Only primary GBM with overall survival data were included for analysis. The patients were stratified according to the heatmap cluster membership to compute a Cox regression analysis, using R CRAN project 4.0.0. Additionally, a uni- and multivariate analysis was performed using age, IDH mutation status, MGMT promoter methylation status, chemotherapy treatment, and heatmap cluster membership.

Ethics Statement

This project (research protocol No. 116/19) was reviewed and approved by the Clinical Research Committee (No. DIC/264/2020) and by the Research Ethics Committee (No CEI/029/21) of the National Institute of Neurology and Neurosurgery of Mexico. It was also reviewed and approved by the Research and Ethics Committees of the National Institute of Genomic Medicine of Mexico.

Statistical Analysis

Fisher's exact tests were performed using GraphPad Prism version 8.0.0 for Windows (GraphPad Software, San Diego, CA, USA; www.graphpad.com) to compare the allelic

frequencies of R-SNV in the Mexican population with those of other populations of phase 3 of the 1000 Genomes Project.

Results

Genotypes Associated with BRT

We performed a systematic literature review (Supplementary Figure S1) and identified 36 R-SNVs. Twenty-nine R-SNVs belonged to 11 genes (RGs) selected from the systematic review (Table 1, Supplementary Table S2), and seven R-SNVs belonged to six genes selected from the dbSNP (https://www.ncbi.nlm.nih.gov/snp/). *APOC1* was also included in our analysis because it is in a 19q13.32 chromosome cluster with *APOE*,[23] which was previously identified as an RG. These 18 genes were named as extended (E-RGs). The E-RGs included *APOE*, *NECTIN2*, *APOC1P1*, *TOMM40*, *POR*, *MAT1A*, *CYP1B1*, *COMT*, *ARVCF*, *TXNRD2*, *BDNF*, *BDNF-AS*, *DTNBP1*, *MGMT*, *CEP128*, *DISC1FP1*, *KCTD1*, and *APOC1* (Supplementary Table S3).

PPI Networks with Functional Clusters Associated with BRT

The BRT-PPI network (Figure 1A) consisted of 12 nodes and 17 edges (Supplementary Table S3). The global network topological and stats measurements (Figure 1B) showed that the proteins have significantly more interactions among one another than expected for a random set of proteins of similar size (P-value = 2.3×10^{-12}). Such enrichment indicates that the BRT-PPI network is considered as a nonrandom graph, and that these proteins are at least partially biologically connected. The results of the topological analyses of each node are listed in Supplementary Table S4, which showed that COMT was the "hub" (with the largest degree k = 7) and "bottleneck" (with the highest Betweenness Centrality value of 0.73) in this network, thus suggesting a central regulatory role. COMT could therefore be strongly involved in cellular processes associated with BRT or be relevant in connecting regulatory molecules. Interestingly, the node corresponding to APOE had the second most relevant topological attributes in biological terms of the BRT-PPI network.

Next, the COMT-extended-PPI network was generated with COMT (hub of the BRT-PPI network) and its direct edges, and then expanded by STRING (Figure 1C). The topological characteristics of the COMT-extended-PPI network (Figure 1B) indicated that the nodes have significantly more interactions than expected (P-value < 1×10^{-16}), as previously found in the BRT-PPI network. Based on these results, we identified a network of proteins (Supplementary Table S5) functionally related to BRT. The COMT-extended-PPI network (score ≥0.9) visualized by Cytoscape generated eight clusters (Figure 1D), seven of which were related to the following RPs: COMT, APOE, BDNF, MAT1A, DTNBP1, APOC1, and CYP1B1. The top four biological functions of those clusters (Figure 1D, Supplementary Figure S2) associated with RPs (correlation P-value between 1×10^{-5} and 1×10^{-12}) revealed that these proteins are part of biologically active complexes (Figure 1D).

Expression of RGs Associated with BRT Reveals a Putative Molecular Profile Associated with Survival in GBM Patients

We used mRNA expression levels of *COMT*, *APOE*, *BDNF*, *MAT1A*, *MGMT*, *POR*, and *CYP1B* (the remaining RGs were not included in the platform) to determine if there was any transcriptomic profile that might associate with survival in patients diagnosed with GBM. Data from 277 patients included in the GMB-BioDP (HT_HG-U133A platform) were analyzed as detailed in Figure 2A. An unsupervised hierarchical clustering analysis performed with the mRNA expression profiles of the seven genes above mentioned revealed two clusters of GBM patients (A and B). Cluster A showed a significantly shorter survival time (12.3 months) when compared with cluster B (17.4 months) (Log-Rank P-value = .043). The cluster analysis also identified three different RGs subgroups (Figure 2A). However, analysis of these RGs subgroups did not identify any prognostic cluster with impact on patient survival (Supplementary Figure S3). Overall, our results identified a novel molecular profile with implications on GBM patient prognosis, which we called the 7BRT (seven genes associated with BRT) molecular profile.

3D Clusters of SNVs in Proteins Encoded by 7BRT Molecular Profile Genes Include Putative Pathogenic Variants Associated with Cancer Biology

We examined the distribution of the novel variants across protein sequence and 3D structure of E-RGs in a Mexican population (genomic data previously obtained by our group),[14] (Supplementary Tables S6–S8). According to the structural analysis, 3D clusters containing at least three variants in spheres with 15-30 Å diameters were discovered in APOE, COMT, CYP1B1, MGMT, and POR (Figure 3). The functional annotation of the identified variants according to the Structure-PPI system (https://rbbt.bsc.es/StructurePPI/) is provided in Supplementary Table S9.

Our structural analysis in APOE revealed one cluster (diameter = 17.78 Å and P-value = .080) that contains the variants p.V140M, p.G145D, and p.R154S (Figure 3A). These variants are positioned in a hotspot region with deleterious effects on the protein function and contain pathogenic variants from other diseases (Supplementary Table S6).[24,25]

For the COMT, we identified one cluster (diameter = 24.16 Å and P-value = .025) that contains the variants p.A72S, p.E84K, p.A102T, p.R125C, p.R128H, and p.V158M (Figure 3B), located in the dimer interface affecting the PPI. The p.A72S variant correlates with reduced enzyme activity and is considered a risk factor for schizophrenia,[26,27] and lung and breast cancer (COSM6354285). p.E84K was identified in lung cancer (COSM2892544), while p.A102T is classified as benign in ClinVar. The p.V158M is associated with low enzyme activity and thermolability[28] and identified in lung,

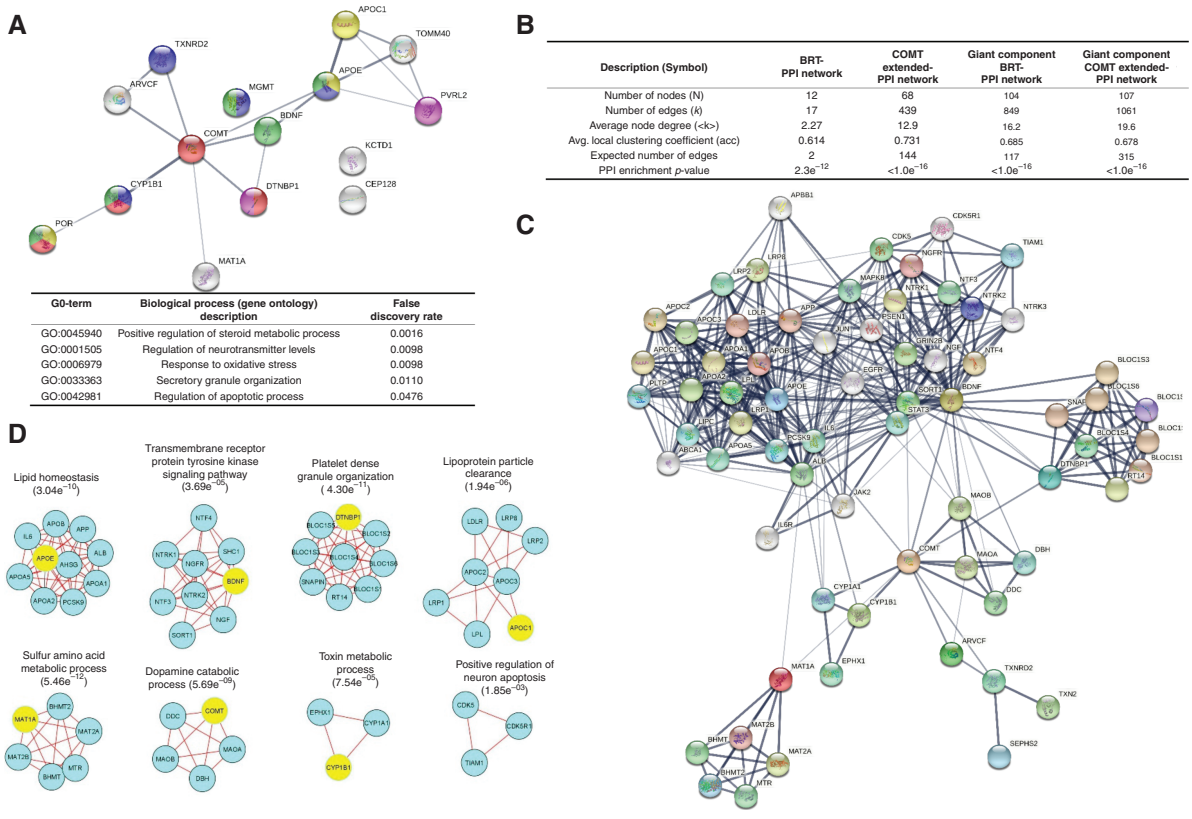
Table 1 Molecular and Phenotypic Characteristics of Germline Variants Selected From the Systematic Review

First Author, Year[Ref.]	Gene(s) Selected (Articles)	SNVs Selected (Articles)	Gene(s) Selected (dbSNP)	Role	Alleles (dbSNP)	Minor Allele	Country/ Region	Ethnicity	GWAS or CGS	Adverse Effect
Correa, 2014[19]	*APOE*	rs405509		Upstream	T>G	T	USA	Caucasian	CGS	Lower scores in recognition memory, in attention and executive functions, in recognition memory and in learning
		rs429358		Coding exon	T>C	C				
		rs7412		Coding exon	C>T	T				
		rs72654473		Downstream	C>A	A				
		rs439401		Downstream	T>C	T				
		rs6857	*NECTIN2*	3′-UTR	C>T	T				
		rs769446		Upstream	T>C	C				
		rs5112	***APOC1P1***	Intron	C>G	C				
		rs405697	*TOMM40*	Intron	A>G	A				
		rs449647		Upstream	A>T	T				
Lombardi, 2015[20]	*POR*	rs17685		3′-UTR	G>A	A	Italy	Caucasian	CGS	Higher risk to develop severe myelotoxicity
	MAT1A	rs17102596		Intron	T>C	C				
		rs7087728		3′-UTR	G>A	A				
	CYP1B1	rs1056837		Coding exon	A>G	A				
		rs1056836		Coding exon	G>C	C				
Correa, 2016[21]	*COMT*	rs4680		Coding exon	G>A	A	USA	Overall: Caucasian, Asian and Black	CGS	Lower scores in delayed recall, tests of attention, working memory, and executive functions
		rs4646316		Intron	C>T	T				
		rs9332377		Intron	C>T	T				
		rs165815	*ARVCF*	Coding exon	C>A	C				
		rs4818		Coding exon	C>G	G				
		rs5746847	*TXNRD2*	5′-UTR	T>A	T				
		rs4646312		Intron	T>C	C				
		rs5993883		Intron	T>G	G				
		rs6269		5′-UTR	A>G	G				
	BDNF	rs10767664		Upstream	T>A	T				Lower scores on tests of attention, working memory, and executive functions
		rs11030104*	*BDNF-AS*	Intron	A>G	G				
		rs10835210*	*BDNF-AS*	5′-UTR	C>A	A				
		rs2030324		Intron	A>G	G				
	DTNBP1	rs742106		3′-UTR	G>A	A				
Altinoz, 2017[22]	*MGMT*	rs2308321		Coding exon	A>G	G	Turkey	Caucasian	CGS	Higher risk to develop myelotoxicity: neutropenia and thrombocytopenia
		rs2308327		Coding exon	A>G	G				
		rs12917		Coding exon	C>T	T				
Wang, 2019[8]	*CEP128*	rs17111237		Promoter	A>G	G	China	Asian	GWAS	Higher risks to develop radiation-induced TLI
		rs162171		Intron	A>C	A				
	DISC1FP1	rs10501719		Intron	A>G	G				
	KCTD1	rs9304497		Intron	G>A	A				

Abbreviations: CGS, comparative genome sequence; GWAS, genome-wide association studies.
The genes that do not code for proteins are in bold.
*These SNVs are assigned to both BDNF and BDNF-AS genes in dbSNP.

pancreatic, and colorectal cancers (COSM4997949). p.V158M are linked to drug response, Parkinson's disease, and schizophrenia spectrum disorders.[27,29]

In CYP1B1 (Figure 3C), we identified only one cluster (diameter = 18.08 Å and *P*-value = .014), which contains the variants p.V432L, p.A443G, and p.N453S. Somatic variant



G0-term	Biological process (gene ontology) description	False discovery rate
GO:0045940	Positive regulation of steroid metabolic process	0.0016
GO:0001505	Regulation of neurotransmitter levels	0.0098
GO:0006979	Response to oxidative stress	0.0098
GO:0033363	Secretory granule organization	0.0110
GO:0042981	Regulation of apoptotic process	0.0476

Description (Symbol)	BRT-PPI network	COMT extended-PPI network	Giant component BRT-PPI network	Giant component COMT extended-PPI network
Number of nodes (N)	12	68	104	107
Number of edges (*k*)	17	439	849	1061
Average node degree (<k>)	2.27	12.9	16.2	19.6
Avg. local clustering coefficient (acc)	0.614	0.731	0.685	0.678
Expected number of edges	2	144	117	315
PPI enrichment *p*-value	$2.3e^{-12}$	$<1.0e^{-16}$	$<1.0e^{-16}$	$<1.0e^{-16}$

Fig. 1 Protein-Protein Interaction (PPI) networks generated by STRING. (A) The BRT-PPI network. (B) Global topological measurements and statistics of PPI networks. (C) The COMT-extended-PPI network. (D) Cluster of the COMT-extended-PPI network made with Cytoscape; proteins encoded by E-RGs are highlighted in yellow. In the upper part, the biological functions of each group are shown with the *P*-value of correlation.

p.N453S (COSM1408011) was found in colorectal cancer (Supplementary Table S10).

In MGMT (Figure 3D), another cluster with five variants was identified (diameter = 21.03 Å and *P*-value = .021). The variants were: p.V81M, p.L84F, p.S93L, p.L142I, and p.I143V. Although all of them were not annotated in the ClinVar database, MGMT mutations at position 81 with different variations, have been identified in the lung (COSM6461010), colorectal (COSM6712349), and stomach (COSM8554441) cancers. Moreover, somatic variants in contiguous positions were associated with GBM (A82V, COSM2151781), and other types of cancer (Supplementary Table S10).

In POR, we identified three different clusters: Cluster 1 (diameter = 20.99 Å and *P*-value = .020) contains the variants p.T372M, p.I444Hfs*6, and p.V492M without any annotations from ClinVar database (Figure 3E); Cluster 2 (diameter = 28.26 Å and *P*-value = .002) contains the variants p.P469A, p.V472M, p.A503V, p.I544L, and p.D575N and Cluster 3 (diameter = 29.97 Å and *P*-value = .003) contains the variants p.V85M, p.R107H, p.D165A, p.G213E, p.V230M, p.E232G, and p.R406C. The variants in these last two clusters have been associated with different types of cancer (Supplementary Table S10).

COMT and APOE RPs Functional Groups Reveal a Novel and Robust Gene Expression Profile Associated with Prognosis in GBM Patients

We analyzed the protein clusters obtained from the COMT-extended-PPI network and the combinations between them. Only COMT and APOE protein clusters showed significant prognostic value (Figure 4A). For this reason, we combined COMT and APOE protein clusters to evaluate its prognostic potential in GBM patients using the GMB-BioDP portal ($n = 277$ patients, HuEx-1_0-st-v2 platform). Thus, we identified two groups of patients with significant differences in survival (Log-Rank *P*-value <.001) (Figure 4B). The median overall survival times for clusters A and B were 13.5 and 33.3 months, respectively. Moreover, a stratified survival analysis according to GBM molecular subtypes (C: classic, M: mesenchymal, P: proneural, N: neural) was also able to identify differences in survival among clusters A and B (Figure 4C–E). Overall, patients with lower multigene expression levels (cutoff < median) survive longer regardless of the GBM molecular subtypes. Taken together, our data strongly suggest that the interaction of the COMT and APOE protein clusters associated with BRT represents a novel signature with strongly significant prognostic value in GBM. The newly-developed molecular signature, named

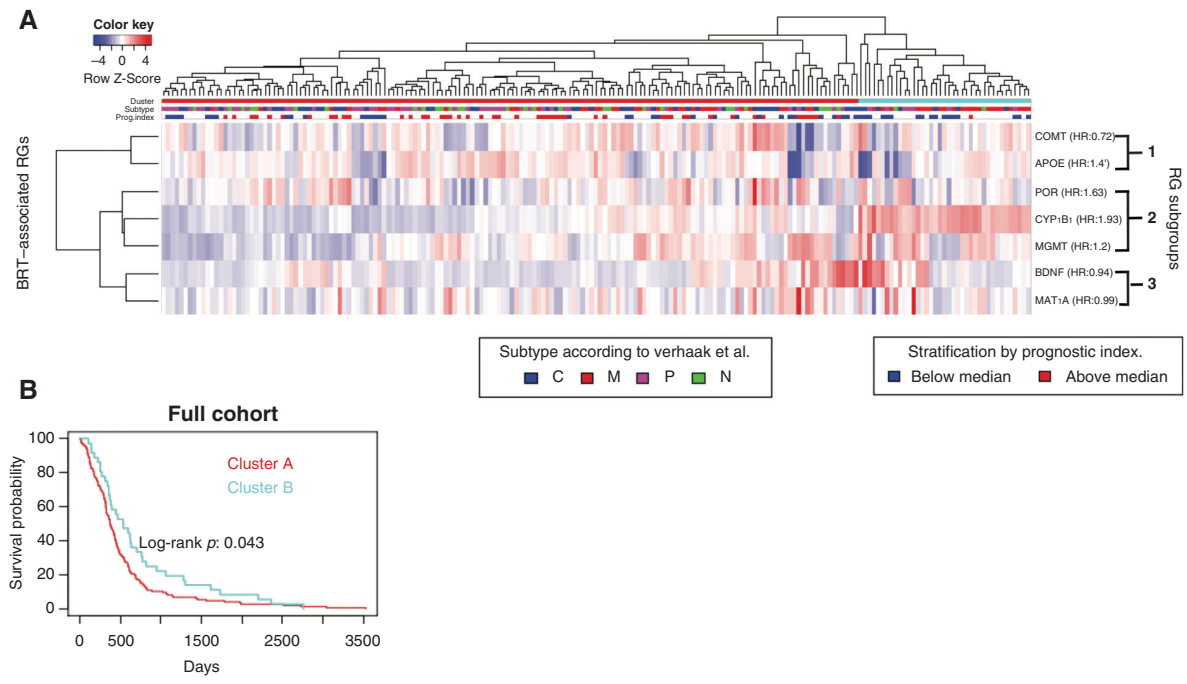


Fig. 2 RGs as a group associated with BRT molecular profile for GBM prognosis. (A) GBM patient clustering with RGs. (B) Kaplan-Meier (KM) survival analysis, using the Platform HT_HG-U133A in Bio-Discovery Portal (GBM-BioDP). Molecular names are annotated with Hazard Ratios (HR) from Cox analysis; * indicates HRs with *P*-values ≤ .1. Molecular subtypes: C (classical), M (mesenchymal), P (proneural), and N (neural).

15CAcBRT (15 mRNAs of COMT + APOE protein clusters related to BRT), includes: *APOE*, *COMT*, *APP*, *MAOB*, *PCSK9*, *MAOA*, *IL6*, *APOA1*, *APOA5*, *DDC*, *DBH*, *ALB*, *APOA2*, *AHSG*, and *APOB* (Supplementary Table S4). Additionally, we identify that the COMT and APOE protein clusters, connected by COMT-APOE and MAOB-APP protein interactions (Figure 4F), were increased for all GBM molecular subtypes (Supplementary Tables S7 and S11). The analysis of survival with a Cox model multivariate documented that 15CAcBRT molecular signature was independent of *MGMT* promoter methylation status (*P*-value = .167) performed in GBM-BioDP.

Additionally, we found that brain expression levels of E-RGs and genes involved in 15CAcBRT signature are affected by R-SNVs, among them, COMT_rs4680 and MATIA_rs7087728 described as an eQTL, (Supplementary Information, Supplementary Figures S4 and S5).

External Validation of the 15CAcBRT Signature

We obtained the microarray dataset mRNA_array_301 from the CGGA portal (http://www.cgga.org.cn/download.jsp), to validate the 15CAcBRT signature in a cohort with different ancestry using the same approach of multi-gene survival analysis of GBM-BioDP. Similar to our previous results (see section "COMT and APOE RPs Functional Groups Reveal a Novel and Robust Gene Expression Profile Associated with Prognosis in GBM Patients"), this analysis revealed three different GBM patient clusters (group 1, $n = 23$; group 2, $n = 83$; and group 3, $n = 2$), with significant differences in survival between group 1 vs group 2 (*P*-value= .007) and group 1 vs group 3 (*P*-value = .108) (Supplementary Figure S6A and B). Furthermore, in the most aggressive molecular subtypes (classical and mesenchymal),[30] better prognosis was found in patients with lower gene expression profiles, in line with the results from our GBM-BioDP cohort (Supplementary Figure S7A and B). A description of the patients by group is presented in Supplementary Table S12. Multivariate Cox regression analysis confirmed that membership to group 2 of 15CAcBRT signature was an independent predictor of survival (*P*-value = .006) after accounting for confounding covariates, such as age (*P*-value = .836), *MGMT* promoter methylation status (*P*-value = .362), *IDH* mutation status (*P*-value = .955) (Log-Rank *P*-value = 8.0×10^{-4}), Supplementary Table S13. Interestingly *IL6*, *APOA5*, *APOA1*, *APOA2*, *APOB*, *APOE*, *DCC*, *MAOB*, and *PCSK9* genes, belonging to 15CAcBRT signature, showed significant differences in gene expression levels among groups, when the patients were stratified by heat map cluster membership (Supplementary Figure S6C). Additionally, this signature was able to predict an increase in the probability of survival in patients under chemotherapy treatment (Log-Rank *P*-value = 3×10^{-5}). (Supplementary Figure S7C and D).

Genes Associated with BRT Are Functionally Associated With Cancer and Neurodegenerative Disease Pathways

To better understand the molecular mechanisms involved in BRT, we performed a functional enrichment analysis of giant components of PPI networks associated with BRT.

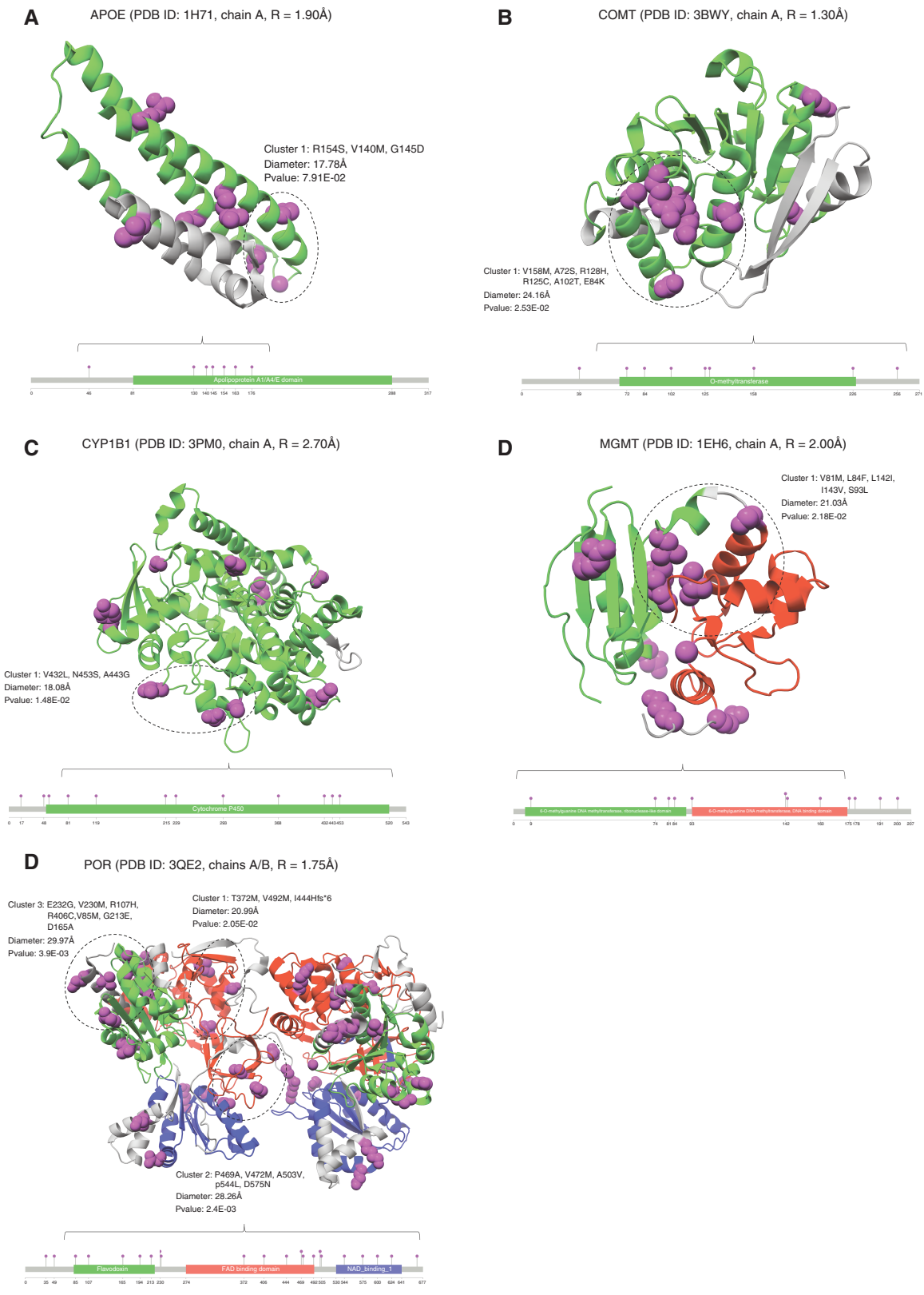


Fig. 3 Mapping of novel variants identified in the Mexican population onto protein sequence and 3D structure. The location of the variants and the spatial 3D clusters identified are depicted for APOE (A), COMPT (B), CYP1B1 (C), MGMT (D), and POR (E).

A

Protein clusters	*p* values (log-rank)							
		APOE	BDNF	MAT1A	DTNBP1	APOC1	CYP1B1	CK5
COMT	**<0.001**	**<0.001**	**0.003**	**<0.001**	**0.012**	**<0.001**	**0.001**	**<0.001**
APOE	**0.030**		**0.017**	0.107	0.062	**0.001**	0.838	**0.017**
BDNF	0.810			0.669	0.275	0.187	0.548	0.756
MAT1A	0.854				0.749	0.526	0.941	0.379
DTNBP1	0.649					0.7621	0.705	0.881
APOC1	0.597						0.077	0.098
CYP1B1	0.627							0.798
CK5	0.815							

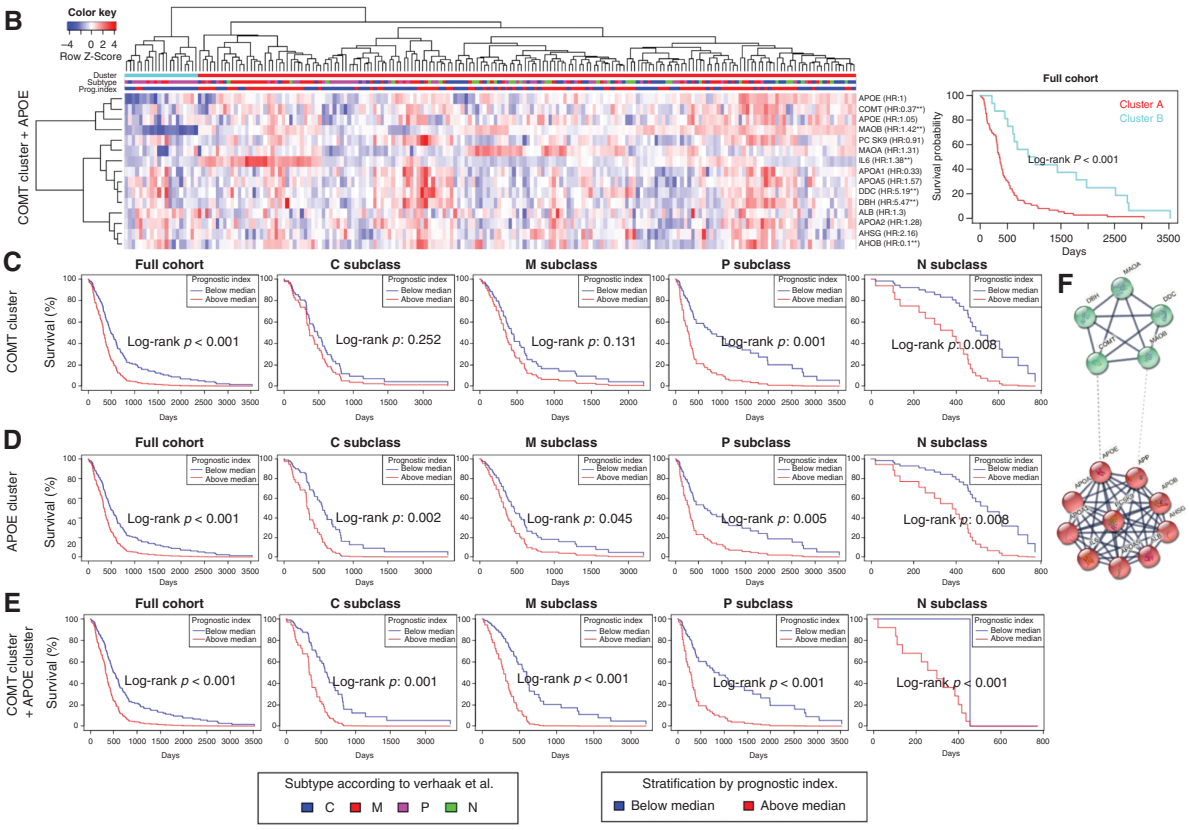


Fig. 4 COMT and APOE protein clusters associated with BRT present novel prognostic signatures for GBM. (A) Protein clusters associated with RGs and their combinations. Statistically significant results are in bold. (B) GBM patient clustering with COMT + APOE protein clusters and Kaplan-Meier (KM) survival analysis using the Platform HuEx-1_0-st-v2 in Bio-Discovery Portal (GBM-BioDP). Molecular names are annotated with Hazard Ratios (HR) from Cox analysis; * indicates HRs with *P*-value ≤ .1; ** indicates HRs with *P*-value ≤ .05. (C–E) Survival analysis based on the impact of the multi-gene prognostic index by molecular subtype: classical (C), mesenchymal (M), proneural (P), and neural (N); for the COMT cluster, APOE cluster, and COMT + APOE clusters, respectively. (F) PPI subnetwork compound by COMT + APOE clusters.

For this purpose, giant components (Figure S8) of the BRT-PPI and COMT-extended-PPI networks were generated (Supplementary Tables S14 and S15). Their global topological and stats measurements (Figure 1B) showed that these giant components can be useful for functional enrichment analysis. In the giant component of COMT-extended-PPI network, cancer-associated pathways were the majority (38%) (Figure 5A, red bars). However, neurodegenerative diseases were predominant (50%) in the giant component BRT-PPI network, where Alzheimer's, Huntington's, and Parkinson's diseases were the most relevant (Figure 5B, red bars).

The human diseases associated with the giant component BRT-PPI network were covered by the giant component COMT-extended-PPI network. Interestingly, the COMT-extended-PPI network linked neurodegenerative diseases with cancer pathways (Figure 5A and C). All pathways are listed in Supplementary Tables S16 and S17.

Discussion

In the current study, we identified the novel 15CAcBRT brain radiotoxicity-related gene expression signature that can classify GBM patients with survival differences. This signature is unique as it showed that: (1) it is composed of two protein clusters which interactions could be potentially regulated with targeted therapy or drug repositioning in GBM patients (Figure 4F) and (2) it is a novel predictor of survival, independent of *MGMT* methylation and *IDH* mutation status.

The extended COMT-PPI network with its direct interactors generated the protein clusters of the 15CAcBRT molecular signature. A relevant protein in the 15CAcBRT was COMT, which has a frequent worldwide eQTL (rs4680, Val158Met; Supplementary Table S6) that decreases its activity, increasing the dopamine (monoamine catecholamine neurotransmitter) levels in the prefrontal cortex.[31] It has been proposed that high levels of dopamine generate antitumor effects on GBM.[32] This suggests that the COMT genotype could be decisive for the prognosis of the GBM, as well as its functional cluster, which is consistent with our findings showing the COMTrs4680 as eQTLs. Furthermore, Structure-PPI analysis showed that variants located at the COMT dimer interface affect its interactions with other proteins and can therefore be used to design new drugs that regulate the interaction between COMT and APOE clusters. This could lead to a significant alteration in the expression of genes involved in the 15CAcBTR signature, whose decreased expression resulted in increased survival rate for all GBM molecular subtypes. Otherwise, *MGMT* promoter methylation status has relevance as a clinical biomarker in GBM,[33] and it is associated with better response to chemoradiation and longer overall survival.[34,35] In this work, we identified a hotspot region in the MGMT protein with significant impacts

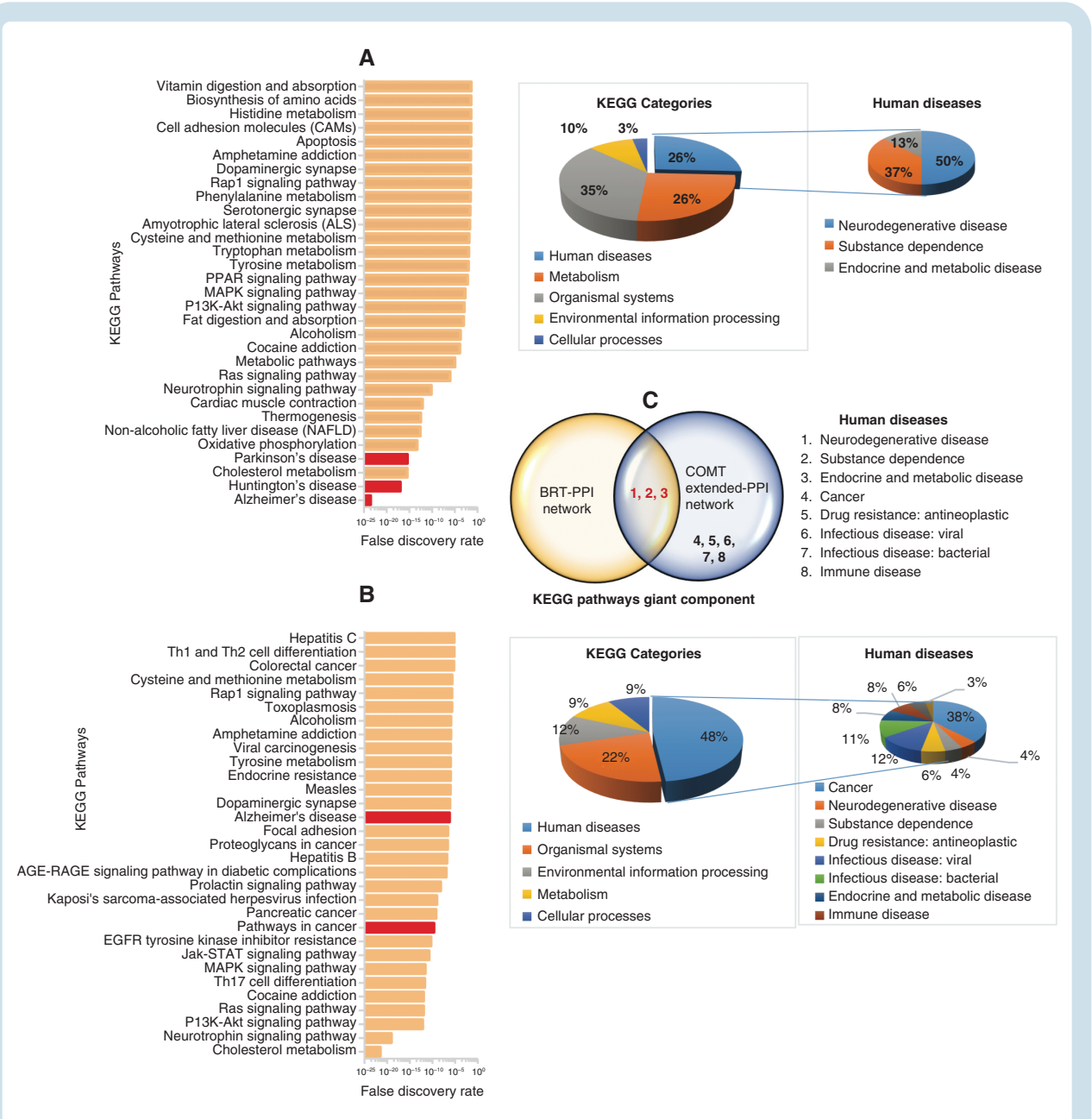


Fig. 5 Pathway enrichment of giant components of PPI networks associated with BRT: pathway top, KEGG categories, and human diseases. (A) Prediction of KEGG pathway analysis for giant components of the BRT-PPI network. (B) Prediction of KEGG pathway analysis for giant components of the COMT-extended-PPI network. (C) Venn diagram of diseases from KEGG pathway analysis of the giant components of the BRT-PPI and COMT-extended-PPI networks.

in different tumors, including GBM (COSM2151781). This reveals that regions other than the *MGMT* promoter may also be relevant to GBM patients, and their influence on response to chemoradiation and overall survival should be evaluated. Overall, the 3D structural modifications for proteins codified by BRT-associated RGs described here may not only be relevant for GBM but also important for other types of cancer (Supplementary Table S10). These are promising avenues as molecular targets for cancer therapy, although further studies should be conducted to evaluate the interactions between BRT-associated germline variants (which induce low levels of RGs expression) and somatic variants in GBM patients. This could provide different mechanisms of analysis and clinical interpretation, as current genomic data are limited to filtered somatic alterations.[36] Furthermore, although the articles included in our systematic review comprised studies containing mainly Caucasian and Asian populations, R-SNVs were found in all populations included in the 1K genomes and gnomAD databases, although they displayed different frequencies (Supplementary Table S6).

Our novel molecular signature 15CAcBRT classifies patients with significant survival differences based on the expression level of RGs (Supplementary Figure S6C), which suggest silencing or overexpression of their products might represent an unexplored therapeutic approach. 15CAcBRT signature may be useful for drug repositioning for GBM. For example, a specific antibody targeting the IL-6 receptor or anti-IL-6 treatment could be combined with MAOB inhibitors (RNA expression profile in Figure 4B, Supplementary Figure S6A and C) used for the treatment of Parkinson's disease and Temozolomide, acting on cancer pathways and neurodegenerative diseases.[37]

On the other hand, the inverse comorbidity between neurodegenerative diseases and cancer has been clinically observed, but its molecular basis is still under study. Recently, it has been documented that some microRNAs are able to link GBM to neurodegenerative disorders,[38] and expression levels of genes involved in neurodegenerative conditions, such as Alzheimer's and Parkinson's diseases, correlated with better prognosis of gliomas.[39] APOE and COMT have a proven association with these neurodegenerative diseases.[40] Their protein clusters are the components of the 15CAcBRT signature identified and validated in our work, and they form biological networks involved in the pathways of cancer and neurodegenerative diseases. Furthermore, among the most significant KEGG pathways (Figure 5) are those associated with cellular radiation response: PI3K/AKT, Ras, MAPK, and EGFR tyrosine kinase inhibitor resistance,[41,42] which evidences the relationship between the 15CAcBRT signature, radiotherapy, cancer, and neurodegenerative diseases. Consequently, our findings provide evidence at the level of molecular biological networks linking neurodegenerative diseases and cancer (Figure 5), useful for evaluating combinations of antitumor drugs with those used for neurodegenerative diseases, focused on drug repositioning for precision medicine.

Study Limitations

Associations between the clinical prognosis and the sex of the GBM patients were not evaluated.

Conclusions

Our results demonstrate that genes related to BRT have clinical relevance for GBM patients. The 15CAcBRT signature related to BRT harbors prognostic survival value in GBM patients and potential for drug repositioning and precision medicine. A relationship between the RGs, neurodegenerative diseases, and cancer was demonstrated using KEGG pathway enrichment analysis. The 3D clusters in the proteins (APOE, COMT, CYP1B1, MGMT, and POR) encoded by RGs containing R-SNVs and the putative pathogenic variants associated with BRT could be a relevant part of a neuro-oncology targeting approach. Overall, our results add evidences to the precision medicine, diagnostic interventions, and development of therapeutics for GBM.

Supplementary Material

Supplementary material is available at *Neuro-Oncology* online.

Keywords

brain radiotoxicity | COMT | glioblastoma | molecular signatures | prognosis

Funding

This work was supported by the National Institute of Genomic Medicine, Mexico and EMBO Short-Term Fellowship 8374 (L.M.-P.).

Conflicts of interest statement. The authors declare no conflict of interest.

Authorship statement. M.C.A.P., J.R.G., and L.O. conceptualized the study. M.C.A.P. and J.R.G. designed and obtained the molecular signatures. L.O., F.B.O., H.G.O., and J.R.G. obtained and analyzed the genotypes of the populations. L.M-P. and T.P. performed computational structural analysis of the germline and somatic variants. M.C.A.P., J.R.G., F.B.O., L.M-P., T.P., and L.O. designed the methodology. M.C.A.P., J.R.G., F.B.O, L.M.-P., T.P., A.R.A., A.M.H., C.C.C., L.O., R.H.P., J.B.P. and H.G.O. conducted investigations. M.C.A.P., J.R.G., F.B.O., L.M-P., and T.P. visualized the data. A.Q.H., H.J.R.G., S.M., M.L.A., M.C.A.P., and J.R.G. analyzed and interpreted the data. M.C.A.P., L.O., and L.M-P. acquired funding. M.C.A.P. was project administrator. M.C.A.P., L.O., A.Q.H., and T.P. supervised the study. M.C.A.P., J.R.G., F.B.O., L.M-P., and T.P. wrote the manuscript. All authors reviewed and edited the work. All authors have read and agreed to the published version of the manuscript.

References

1. Cantrell JN, Waddle MR, Rotman M, et al. Progress toward long-term survivors of glioblastoma. *Mayo Clin Proc.* 2019;94(7):1278–1286.
2. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 2005;352(10):987–996.
3. Lara-Velazquez M, Al-Kharboosh R, Jeanneret S, et al. Advances in brain tumor surgery for glioblastoma in adults. *Brain Sci.* 2017;7(12):166.
4. Almeida JP, Chaichana KL, Rincon-Torroella J, Quinones-Hinojosa A. The value of extent of resection of glioblastomas: clinical evidence and current approach. *Curr Neurol Neurosci Rep.* 2015;15(2):1–13.
5. De Biase G, Garcia DP, Bohnen A, Quiñones-Hinojosa A. Perioperative management of patients with glioblastoma. *Neurosurg Clin N Am.* 2021;32(1):1–8.
6. De Ruysscher D, Niedermann G, Burnet NG, et al. Radiotherapy toxicity. *Nat Rev Dis Primers.* 2019;5(1):1–20.
7. Lawrence YR, Wang M, Dicker AP, et al. Early toxicity predicts long-term survival in high-grade glioma. *Br J Cancer.* 2011;104(9):1365–1371.
8. Zhao J, Wang L, Hu G, Wei BA. 6-gene risk signature predicts survival of glioblastoma multiforme. *Biomed Res Int.* 2019;2019:1649423.
9. Hou Z, Yang J, Wang H, Liu D, Zhang HA. potential prognostic gene signature for predicting survival for glioblastoma patients. *Biomed Res Int.* 2019;2019:9506461.
10. Xiong Y, Xiong Z, Cao H, et al. Multi-dimensional omics characterization in glioblastoma identifies the purity-associated pattern and prognostic gene signatures. *Cancer Cell Int.* 2020;20(1):1–12.
11. Wang TM, Shen GP, Chen MY, et al. Genome-wide association study of susceptibility loci for radiation-induced brain injury. *J Natl Cancer Inst.* 2019;111(6):620–628.
12. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ.* 2021;372:n160.
13. Akobeng AK. Principles of evidence based medicine. *Hum Mol Genet.* 2005;90(8):837–840.
14. Salas-Martínez MG, Saldaña-Alvarez Y, Cordova EJ, et al. Genetic variability of five ADRB2 polymorphisms among Mexican Amerindian ethnicities and the Mestizo population. *PLoS One.* 2019;14(12):e0225030e0225030.
15. Meyer MJ, Lapcevic R, Romero AE, et al. Mutation3D: cancer gene prediction through atomic clustering of coding variants in the structural proteome. *Hum Mutat.* 2016;37(5):447–456.
16. Vazquez M, Valencia A, Pons T. Structure-PPi: a module for the annotation of cancer-related single-nucleotide variants at protein-protein interfaces. *Bioinformatics.* 2015;31(14):2397–2399.
17. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47(D1):D607–D613.
18. Varghese RT, Liang Y, Guan T, et al. Survival kinase genes present prognostic significance in glioblastoma. *Oncotarget.* 2016;7(15):2014020140–20140220151.
19. Correa DD, Satagopan J, Baser RE, et al. APOE polymorphisms and cognitive functions in patients with brain tumors. *Neurology.* 2014;83(4):320–327.
20. Lombardi G, Rumiato E, Bertorelle R, et al. Clinical and genetic factors associated with severe hematological toxicity in glioblastoma patients during radiation plus temozolomide treatment. *Am J Clin Oncol.* 2015;38(5):514–519.
21. Correa DD, Satagopan J, Cheung K, et al. COMT, BDNF, and DTNBP1 polymorphisms and cognitive functions in patients with brain tumors. *Neuro Oncol.* 2016;18(10):1425–1433.
22. Altinoz MA, Elmaci I, Bolukbasi FH, et al. MGMT gene variants, temozolomide myelotoxicity and glioma risk. A concise literature survey including an illustrative case. *J Chemother.* 2017;29(4):238–244.
23. Fuior EV, Gafencu AV. Apolipoprotein C1: its pleiotropic effects in lipid metabolism and beyond. *Int J Mol Sci.* 2019;20(23):5939.
24. Wardell MR, Brennan SO, Janus ED, Fraser R, Carrell RW. Apolipoprotein E2-Christchurch (136 Arg----Ser). New variant of human apolipoprotein E in a patient with type III hyperlipoproteinemia. *J Clin Investig.* 1987;80(2):483–490.
25. Emi M, Wu LL, Robertson MA, et al. Genotyping and sequence analysis of apolipoprotein E isoforms. *Genomics.* 1988;3(4):373–379.
26. Lee SG, Joo Y, Kim B, et al. Association of Ala72Ser polymorphism with COMT enzyme activity and the risk of schizophrenia in Koreans. *Hum Genet.* 2005;116(4):319–328.
27. Okochi T, Ikeda M, Kishi T, et al. Meta-analysis of association between genetic variants in COMT and schizophrenia: an update. *Schizophr Res.* 2009;110(1–3):140–148.
28. Graf WD, Unis AS, Yates CM, et al. Catecholamines in patients with 22q11. 2 deletion syndrome and the low-activity COMT polymorphism. *Neurology.* 2001;57(3):410–416.
29. Corvol JC, Bonnet C, Charbonnier-Beaupel F, et al. The COMT Val158Met polymorphism affects the response to entacapone in Parkinson's disease: a randomized crossover clinical trial. *Ann Neurol.* 2011;69(1):111–118.
30. Fedele M, Cerchia L, Pegoraro S, Sgarra R, Manfioletti G. Proneural-mesenchymal transition: phenotypic plasticity to acquire multitherapy resistance in glioblastoma. *Int J Mol Sci.* 2019;20(11):2746.
31. Diamond A. Consequences of variations in genes that affect dopamine in prefrontal cortex. *Cereb Cortex.* 2007;17(suppl_1):i161–i170.
32. Yang K, Wei M, Yang Z, et al. Activation of dopamine receptor D1 inhibits glioblastoma tumorigenicity by regulating autophagic activity. *Cell Oncol.* 2020;43(6):1175–1190.
33. Butler M, Pongor L, Su YT, et al. MGMT status as a clinical biomarker in Glioblastoma. *Trends Cancer.* 2020;6(5):380–391.
34. Binabaj MM, Bahrami A, ShahidSales S, et al. The prognostic value of MGMT promoter methylation in glioblastoma: a meta-analysis of clinical trials. *J Cell Physiol.* 2018;233(1):378–386.
35. Dunn J, Baborie A, Alam F, et al. Extent of MGMT promoter methylation correlates with outcome in glioblastomas given temozolomide and radiotherapy. *Br J Cancer.* 2009;101(1):124–131.
36. Fernandez-de-Cossio J, Perera Y. Impact of germline susceptibility variants in cancer genetic studies. *arXiv.* 2020; arXiv:1610.09427, preprint: not peer reviewed. doi:10.13140/RG.2.2.23652.19846
37. Fowler JS, Logan J, Volkow ND, et al. Evidence that formulations of the selective MAO-B inhibitor, selegiline, which bypass first-pass metabolism, also inhibit MAO-A in the human brain. *Neuropsychopharmacology.* 2015;40(3):650–657.
38. Thomas L, Florio T, Perez-Castro C. Extracellular vesicles loaded miRNAs as potential modulators shared between glioblastoma, and Parkinson's and Alzheimer's diseases. *Front Cell Neurosci.* 2020;14:590034.
39. Gargini R, Segura-Collar B, Herránz B, et al. The IDH-TAU-EGFR triad defines the neovascular landscape of diffuse gliomas. *Sci Transl Med.* 2020;12(527):eaax1501.
40. Pau KC, Rausch R, Creek MM, et al. APOE, MAPT, and COMT and Parkinson's disease susceptibility and cognitive symptom progression. *J Parkinsons Dis.* 2016;6(2):349–359.
41. Maier P, Hartmann L, Wenz F, Herskind C. Cellular pathways in response to ionizing radiation and their targetability for tumor radiosensitization. *Int J Mol Sci.* 2016;17(1):102.
42. Domina EA, Philchenkov A, Dubrovska A. Individual response to ionizing radiation and personalized radiotherapy. *Crit Rev Oncog.* 2018;23(1–2):69–92.