

Effect of Obesity-Associated Genes on Survival Rates of Brain Tumor Patients

Aarav Sehdev

Basis Phoenix, 13231 North 22nd St, Phoenix, AZ 85022, USA; aaravsehdev8@gmail.com

ABSTRACT: Brain tumors are a heterogeneous group of tumors, including low-grade gliomas (LGG) and glioblastoma multiforme (GBM). Obesity increases the risk of developing cancer, including brain tumors. However, there is conflicting evidence from epidemiological studies regarding the effect of obesity on brain tumors. We conducted a retrospective cohort study using publicly available databases to study the effect of obesity-associated genomic alterations on overall survival (OS) and disease-free survival (DFS) in brain tumor patients. We used the genome-wide association studies (GWAS) catalog to compile a list of genes associated with obesity. We hypothesized that brain tumor patients with alterations in obesity-associated genes (altered group) demonstrate better OS as compared to brain tumor patients without alterations (unaltered group). We found a significant improvement in the OS and a suggestion of improvement in the DFS in the altered group as compared to the unaltered group. Our study has several limitations due to its retrospective nature and the lack of availability of body mass index, thus inhibiting our ability to directly correlate clinical obesity status with genomic findings. Our analysis highlights the need for future preclinical and clinical research to elucidate the mechanisms of how obesity affects the survival outcomes in brain tumor patients.

KEYWORDS: Medical and Health Sciences, Oncology, Molecular Biology, Genetics, Bioinformatics.

■ Introduction

Brain tumors, including low-grade gliomas (LGG) and glioblastoma multiforme (GBM), are a heterogeneous group of tumors that arise from different types of glial or precursor cells, such as astrocytes, oligodendrocytes, and neural stem cells.¹ There are 24,820 estimated new cases of brain tumors, with 18,330 estimated deaths in the year 2025.² The 5-year relative survival is 33%.² These tumors are more common in males than in females due to hormonal and immunological differences.^{3,4} The World Health Organization (WHO) classifies brain tumors into 4 grades, with grade 1 being benign (e.g., pilocytic astrocytoma) and grade 4 being the most malignant and aggressive brain tumor (e.g., glioblastoma multiforme). Grade 2 tumors are low-grade with usually a good prognosis after surgical excision, but with a 50% probability of anaplastic transformation (e.g., low-grade astrocytoma), whereas grade 3 tumors are more aggressive with a median survival of 24 months (e.g., anaplastic oligodendroglioma).⁵ The development and progression of brain tumors is associated with environmental exposures such as ionizing radiation, carcinogens (such as N-nitroso compounds, pesticides, and air pollutants), and radiofrequency electromagnetic waves.⁶ There is emerging evidence for an association between brain tumors and obesity.⁷

Obesity is now recognized as a major public health problem and is reaching epidemic proportions.⁸ Obesity increases the risk of developing cancer, including brain tumors such as GBM and LGG; however, there is conflicting evidence from epidemiological studies.^{9–11} Furthermore, emerging data suggest that taking glucagon-like peptide-1 receptor agonists lowers the risk of cancer, including meningioma, in obese or overweight

individuals.¹² There is epidemiological and experimental evidence suggesting that obesity, measured by body mass index (BMI), is associated with a lower incidence and mortality from brain tumors – described as the “obesity paradox.”^{13–15} This is paradoxical because obesity is a known risk factor for multiple diseases such as diabetes, cardiovascular disease, and cancer. There are multiple plausible explanations for this paradox, including less aggressive cancer biology, a stronger immune system, and better tolerance to toxic treatments such as chemotherapy.^{16–19} The molecular mechanisms underlying the outcomes in obese patients with brain tumors are not well understood. Preclinical and clinical evidence suggest that obesity promotes insulin resistance, resulting in increased levels of insulin and insulin-like growth factor (IGF), which act as potent mitogens by stimulating insulin receptor substrate (IRS).²⁰ In addition, adipose tissue can make a variety of pro-inflammatory cytokines and adipokines.²¹ The altered levels of these cell-extrinsic factors regulate cellular metabolism through complex interactions with multiple signal transduction pathways.^{20,21}

Genome-wide association studies (GWAS) have shown several genes to be associated with obesity.^{22–24} Body mass index (BMI) was predominantly used as the proxy for obesity (BMI ≥ 30 kg/m²) in these studies.^{25–27} Over the last 2 decades, more than 1000 single-nucleotide polymorphisms have been shown to have a significant association with obesity.²⁸ Additionally, several meta-analyses have further shown the strength of this association.^{29–32} Therefore, it is plausible that the molecular alterations in the obesity-associated genes affect the survival of brain tumor patients.

Our hypothesis for this study is that brain tumor patients with alterations in obesity-associated genes demonstrate better

overall survival as compared to brain tumor patients without alterations in obesity-associated genes. Therefore, we conducted a retrospective cohort study with the primary aims of: i) describing the prevalence of obesity-associated genomic alterations in patients with brain tumors (LGG and GBM); ii) assessing the effect of obesity-associated genes on overall survival (OS) of patients with brain tumors; and iii) assessing the effect of obesity-associated genes on disease-free survival (DFS) of patients with brain tumors.

■ Methods

Patient Samples and Data Sources:

To test our hypothesis of correlation between obesity-associated genes and brain tumors, including LGG (astrocytoma grade II-III, anaplastic astrocytoma grade III, oligoastrocytoma grade III, anaplastic oligodendroglioma grade III, and anaplastic oligoastrocytoma grade III) and glioblastoma multiforme (GBM), we used the Cancer Genome Atlas (TCGA) data.^{33,34} The TCGA data were accessed through cBioPortal, which is a widely available and easily usable online tool for clinical and genomics datasets for many different cancer types.^{35–38} The Genomic Data Commons (GDC) data of brain tumor patients were obtained, which includes grade II to IV brain tumors, including astrocytoma, anaplastic astrocytoma, oligoastrocytoma, anaplastic oligodendroglioma, anaplastic oligoastrocytoma, and GBM.³⁹

Using the TCGA data, brain tumor patients were analyzed for alterations in morbid obesity-associated genes. We used morbid obesity-associated genes instead of obesity-associated genes because of the stronger theoretical and statistical association, as well as the total number of genes showing association (53 unique genes for “morbid obesity,” whereas 1122 genes for “obesity”).

All available data, including gene mutation, copy number variations, and mRNA gene expression, were used to categorize the entire cohort into an altered group (those with genomic alterations) and an unaltered group (those without genomic alterations). Only non-synonymous mutations were used for gene mutation analyses. Synonymous mutations are not currently available in cBioPortal. The patients were classified into high and low mRNA expression using an RNA-seq z-score threshold of 2. The RNA-seq z-score is a normalization method that is used to determine how many standard deviations a gene's expression in a sample differs from the average expression across all samples. For instance, a z-score of 2 indicates the mean; a z-score higher than 2 indicates overexpression (high), and a z-score lower than 2 means underexpression (low). The final data lock date was December 6, 2025.

Primary Gene List: To identify obesity-associated genes, we used the Human GWAS Catalog. The National Human Genome Research Institute-European Bioinformatics Institute (NHGRI-EBI) Catalog of human GWAS was accessed using the search term “morbid obesity” to identify the studies that have found significant gene association with morbid obesity.⁴⁰ The GWAS catalog provides both locus-level and mapped gene-level data in a tabular format. We used the mapped gene-level data. A total of 53 unique genes were identified,

which were used for primary survival analysis and referred to as the “primary gene list” (Supplemental Table 1). Notably, the NHGRI-EBI Catalog is the largest known genome-wide association study database.

Alternate Gene Lists: To further assess the relationship between the obesity-associated genes and brain tumors, we also compiled 3 alternate gene lists of obesity-associated genes using other GWAS databases (Supplemental Table 2). These are described below:

i) The Centers for Disease Control and Prevention (CDC) Public Health Genomics and Precision Health Knowledge Base, version 10.0 (PHGKB).⁴¹ PHGKB is based on a searchable online database of published scientific literature of genotype-phenotype associations. PHGKB is maintained by the Centers for Disease Control and Prevention (CDC). We found 15 genes showing association with obesity when we queried the PHGKB database for the search term “morbid obesity”.

ii) Phenome-wide association study Web (PheWeb) is based on International Classification of Diseases (ICD) billing codes extracted from electronic health records (EHR) for genome-wide associations.⁴² The PheWeb primarily has White British participants from the United Kingdom. We found 68 genes showing association with obesity when we queried the PheWeb database for the search term “obesity”.

iii) A recent publication by Huang et al., where they analyzed 14 different tumor types using TCGA data to correlate molecular variances by BMI.³⁸ In this paper, they have reported 82 genes to be associated with higher BMI across 14 different tumor types.

Statistical Analysis:

Baseline characteristics between the altered group (brain tumor patients with genomic alterations in obesity-associated genes) and the unaltered group (brain tumor patients without genomic alterations in obesity-associated genes) were compared using a chi-square test for categorical variables and a Wilcoxon test for continuous variables. Q-value was calculated where relevant to control for the false discovery rate due to multiple hypothesis testing. Overall survival (OS) and disease-free survival (DFS) were estimated using the Kaplan-Meier method. The log-rank test was used to compare survival between the two groups (brain tumor patients with and without alterations in obesity-associated genes) using a P value < 0.05. A total of 10 patients who overlapped between the two groups were excluded from patient-level analysis. All statistical analyses were performed using cBioPortal or Microsoft® Excel for Mac Version 16.104.

■ Results

Study Participants:

A total of 530 LGG (diffuse glioma) and 539 GBM (glioblastoma multiforme) patients were in the study, of which 683 were in the altered and 386 were in the unaltered group. Within the altered group, there were significantly more patients with LGG compared to the unaltered group (406 vs. 124, respectively, $p < 10^{-10}$, $q < 10^{-10}$). Similarly, there were more

GBM patients in the altered group as compared to the unaltered group (277 vs. 262, respectively, $p < 10^{-10}$, $q < 10^{-10}$). The year of diagnosis for the altered group was between 1992 and 2013 versus 1989 and 2013 for the unaltered group ($p = 2.33e^{-4}$, $q = 4.95e^{-4}$). The patients in the altered group had less prior treatment compared to the unaltered group (3 vs 17, $p = 1.62e^{-5}$, $q = 4.25e^{-5}$). Altered and unaltered groups showed no significant difference between the mean mutation count (79.69 vs. 33.79, respectively, $p = 0.94$, $q = 0.96$) or mean fraction of genome altered (0.16 vs. 0.16, respectively, $p = 0.37$, $q = 0.53$). However, there were no significant differences between the two groups with respect to other demographic or clinical characteristics (Table 1).

Table 1: Characteristics of brain tumor patients with and without alterations in obesity genes. Importantly, there was no significant difference between the altered and unaltered groups in either mean mutation count or the mean fraction of the genome altered.

Clinical Attribute		Altered Group (N=683)	Unaltered Group (N=386)	Statistical Test	p-Value	q-Value
Diagnosis Age	Range	18-89	18-88	Wilcoxon Test	0.24	0.39
	Median	51	52			
	Mean	50.12	51.34			
Ethnicity Category	Not Hispanic or Latino	571	308	Chi-Squared Test	0.71	0.81
	NA	71	54			
	Hispanic or Latino	31	14			
Gender	Female	273	159	Chi-Squared test	0.50	0.62
	Male	398	213			
	NA	2	4			
Race	White	600	329	Chi-Squared Test	0.69	0.80
	Black or African American	42	25			
	NA	16	17			
	Asian	14	5			
	American Indian or Alaska Native	1	0			
Tumor Category	TCGA-LGG (Low-grade glioma/diffuse glioma)	406	124	Chi-Squared Test	<10 ⁻¹⁰	<10 ⁻¹⁰
	TCGA-GBM (Glioblastoma multiforme)	277	262			
Morphology	Glioblastoma, grade 4	270	252	Chi-Squared test	<10 ⁻¹⁰	<10 ⁻¹⁰
	NA	2	4			
	Diffuse astrocytoma, grade 2	41	24			
	Anaplastic astrocytoma, grade 3	106	25			
	Oligodendroglioma, grade 2	87	26			
	Oligoastrocytoma, grade 2	96	37			
	Anaplastic oligodendroglioma, grade 3	71	8			
Patient Primary Tumor site	Central nervous system (brain, spinal cord, and retina)	401	120	Chi-Squared test	<10 ⁻¹⁰	<10 ⁻¹⁰
	NA	2	4			
	Brain	270	252			
Year of Diagnosis	Minimum	1992	1989	Wilcoxon Test	2.33e ⁻⁴	4.95e ⁻⁴
	Maximum	2013	2013			
Mutation Count	Range	1-11177	1-131	Wilcoxon Test	0.94	0.96
	Median	30	31			
	Mean	79.69	33.79			
Fraction Genome Altered	Range	0-0.95	0-0.93	Wilcoxon Test	0.37	0.53
	Median	0.14	0.15			
	Mean	0.16	0.16			
Prior Treatment	False	665	355	Chi-Squared Test	1.62e ⁻⁵	4.25e ⁻⁵
	True	3	17			
	NA	5	4			
Prior Malignancy	False	400	136	Chi-Squared Test	0.96	0.96
	True	5	1			
	NA	268	239			

Genomic Alterations:

Analyzing the genomic alterations in obesity-associated genes, we found that 76.94% of LGG patients had alterations in these genes as compared to 51.33% of GBM patients. Patients with GBM showed a higher percentage of mutations (5.13% vs. 0.97%), amplifications (4.18% vs. 0.58%), and deep deletions (1.71% vs. 0.58%) as compared to LGG patients. However, a higher percentage of gene expression (62.4% vs. 30.23%) and multiple alterations (9.5% vs. 9.32%) were seen in LGG patients as compared to GBM patients (Table 2; Figure 1).

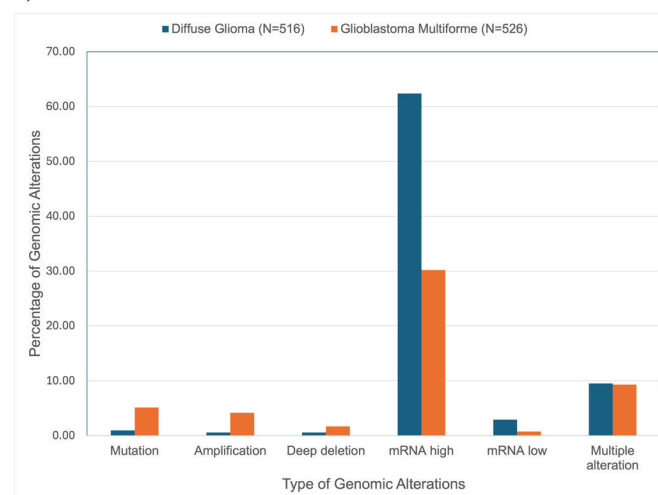


Figure 1: Cancer type and summary of genomic alterations in the altered vs. unaltered group. GBM patients had higher percentages of mutations, amplifications, and deep deletions, whereas LGG patients showed higher percentages of gene expression.

Table 2: Genomic alterations in obesity genes between diffuse glioma and glioblastoma multiforme. GBM patients showed more mutations, amplifications, and deep deletions, while LGG patients showed higher gene expression.

	Diffuse Glioma (N=516)	Glioblastoma Multiforme (N=526)
Mutation	0.97% (5 patients)	5.13% (27 patients)
Amplification	0.58% (3 patients)	4.18% (22 patients)
Deep deletion	0.58% (3 patients)	1.71% (9 patients)
mRNA high	62.4% (322 patients)	30.23% (159 patients)
mRNA low	2.91% (15 patients)	0.76% (4 patients)
Multiple alteration	9.5% (49 patients)	9.32% (49 patients)
Total	76.94% (397 patients)	51.33% (270 patients)

In the analysis of gene frequencies within the altered and unaltered group (as classified by the obesity-associated genes), we found that in the unaltered group the genes including CDKN2A (39.22% vs 27.29%, $p = 6.80e^{-5}$, $q = 2.36e^{-3}$), CDKN2B (38.18% vs 25.81%, $p = 3.02e^{-5}$, $q = 1.26e^{-3}$), UBA52P6 (37.92% vs 25.66%, $p = 3.89e^{-5}$, $q = 1.54e^{-3}$), CDKN2A-AS1 (37.40% vs 25.81%, $p = 9.40e^{-5}$, $q = 2.95e^{-3}$), CDKN2B-AS1 (35.32% vs 23.89%, $p = 8.99e^{-5}$, $q = 2.95e^{-3}$) showed significantly greater frequency as compared to the altered group (Table 3, Figure 2). Furthermore, we found that there were genes in the altered group, IDH1 (45.72% vs 31.17%, $p = 3.72e^{-6}$, $q = 4.95e^{-4}$), which showed significantly greater frequency in the altered group compared to the unaltered group (Table 3; Figure 2). Notably, the genes shown here (Table 3; Figure 2) are the genes with the highest frequency of alterations within these 2

groups (not obesity-associated genes but rather any gene). In other words, after stratifying the cohort based on obesity-associated gene status, Figure 2 highlights the most frequently mutated genes in each group.

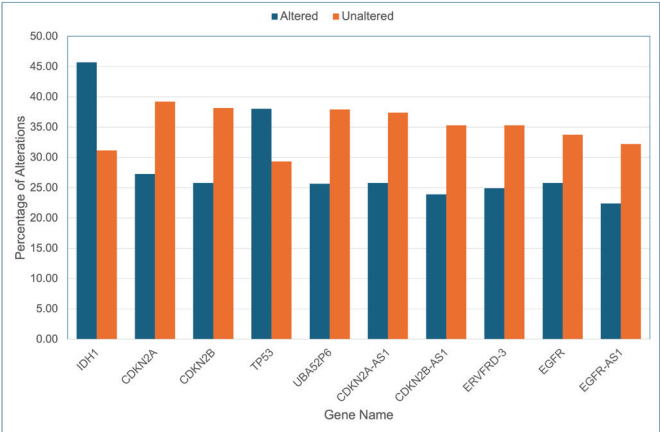


Figure 2: Genes with the highest frequency of variations in the altered and unaltered group. CDKN2A, CDKN2B, UBA52P6, CDKN2A-AS1, and CDKN2B-AS1 showed significantly greater frequency of modifications in the unaltered group, whereas IDH1 showed significantly greater frequency of modifications in the altered group.

Table 3: Genes with the highest frequency of alterations in the altered and unaltered group. CDKN2A, CDKN2B, UBA52P6, CDKN2A-AS1, and CDKN2B-AS1 were more frequently modified in the unaltered group, whereas IDH1 modifications were more common in the altered group.

Gene Name	Altered (N=678)	Unaltered (N=385)	P-value	Q-value
IDH1	45.72% (310)	31.17% (120)	3.72e-6	4.95e-4
CDKN2A	27.29% (185)	39.22% (151)	6.80e-6	2.36e-3
CDKN2B	25.81% (175)	38.18% (147)	3.02e-5	1.26e-3
TP53	38.05% (258)	29.35% (113)	4.89e-3	0.09
UBA52P6	25.66% (174)	37.92% (146)	3.89e-6	1.54e-3
CDKN2A-AS1	25.81% (175)	37.40% (144)	9.40e-6	2.95e-3
CDKN2B-AS1	23.89% (162)	35.32% (136)	8.99e-6	2.95e-3
ERVFRD-3	24.93% (169)	35.32% (136)	4.10e-4	0.01
EGFR	25.81% (175)	33.77% (130)	7.28e-3	0.12
EGFR-AS1	22.42% (152)	32.21% (124)	6.10e-4	0.02

Survival Analysis Using Primary Gene List:

The OS analysis included 1022 patients, 652 in the altered group and 370 in the unaltered group (Figure 3; Table 4). The median OS was significantly better in the altered group as compared to the unaltered group (33.21 *vs.* 20.37 months; HR = 0.70; 95% CI = 0.58-0.83; *p* = 2.76e-5; *q* = 5.53e-5).

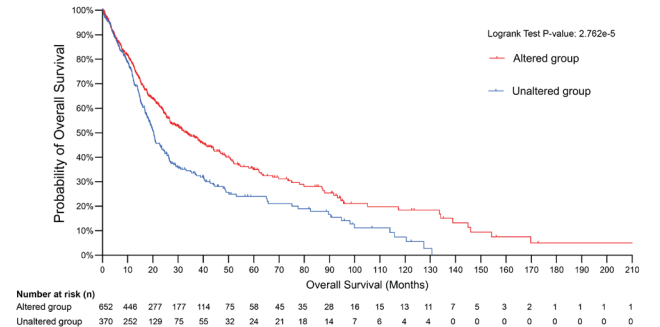


Figure 3: Overall survival (OS) in the altered vs. unaltered group. The median OS was significantly better in the altered group as compared to the unaltered group.

Table 4: Overall survival (OS) plot summary. Median OS was significantly higher in the altered group than in the unaltered group.

	Number of Cases, Total	Number of Events	Median Months Overall (95% CI)
Altered group	652	310	33.21 (26.91-41.10)
Unaltered group	370	226	20.37 (18.33-23.49)

The DFS analysis included 133 patients, 98 in the altered group and 35 in the unaltered group (Figure 4; Table 5). The median DFS was not statistically different between the two groups (median DFS not available; HR = 0.71; 95% CI = 0.27-1.87; *p* = 0.45; *q* = 0.45). Notably, Kaplan-Meier curves provide median DFS only when 50% of the cohort has experienced the event. As seen in Figure 4, this threshold has not been reached in our dataset for either group. Therefore, the median DFS cannot be estimated.

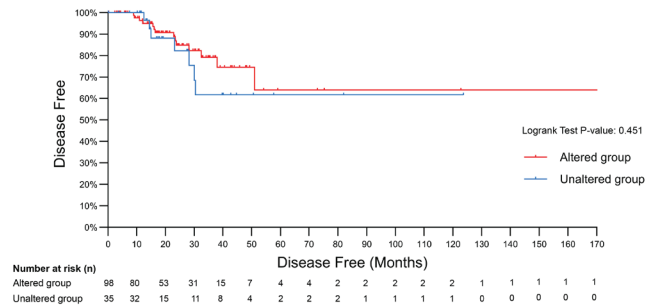


Figure 4: Disease-free survival (DFS) in the altered vs. unaltered group. The median DFS was not statistically different between the two groups.

Table 5: Disease-free survival (DFS) plot summary. Median DFS did not differ significantly between the two groups.

	Number of Cases, Total	Number of Events	Median Months Disease Free (95% CI)
Altered group	98	14	NA
Unaltered group	35	7	NA

Survival Analysis Using an Alternate Gene List:

i) *PHGKB*: The OS analysis included 1019 patients, 389 in the altered group and 630 in the unaltered group. The median OS was significantly better in the altered group as compared to the unaltered group (40.80 *vs.* 21.88 months; HR = 0.67; 95% CI = 0.57-0.80; *p* = 1.79e-5; *q* = 3.58e-5).

The DFS analysis included 132 patients, 75 in the altered group and 57 in the unaltered group. The median DFS was not statistically different between the two groups (median DFS not available; HR = 1.08; 95% CI = 0.45-2.60; *p* = 0.86; *q* = 0.86).

ii) *PheWeb*: The OS analysis included 1023 patients, 716 in the altered group and 307 in the unaltered group. The median OS was significantly better in the altered group as compared to the unaltered group (32.42 *vs.* 19.84 months; HR = 0.63; 95% CI = 0.52-0.76; *p* = 1.64e-7; *q* = 3.27e-7).

The DFS analysis included 134 patients, 113 in the altered group and 21 in the unaltered group. The median DFS was not statistically different between the two groups (median DFS not available; HR = 1.75; 95% CI = 0.54-0.57; *p* = 0.44; *q* = 0.44).

iii) *Huang et al.*: The OS analysis included 1032 patients, 959 in the altered group and 73 in the unaltered group. The median OS was significantly better in the altered group as compared to the unaltered group (27.2 *vs.* 17.64 months; HR = 0.57; 95% CI = 0.41-0.80; *p* = 2.38e-5; *q* = 4.77e-5).

The DFS analysis could not be done due to not enough patients in the representative groups.

■ Discussion

In this large retrospective cohort study of 1069 patients with brain tumors, we report an observation of significant improvement in the OS and a suggestion of improvement in the DFS in patients with alterations in obesity-associated genes. Our results show that patients in the altered group have a significantly better prognosis as compared to the unaltered group. This has been described as the “obesity paradox” in brain tumors as well as in other cancers.^{43–49} There are multiple plausible explanations for why obese patients have better outcomes. First, higher BMI in cancer patients could be reflective of less aggressive cancer biology, leading to better outcomes.¹⁶ Second, most studies assess obesity using BMI, which can be misleading because elevated BMI may reflect greater muscle mass rather than increased adiposity.^{17,18,43} Third, studies have shown a stronger immune system in obese patients, resulting in protective effects.¹⁹ Lastly, patients with higher BMI may have greater nutrient reserves, potentially improving their tolerance to toxic treatments such as chemotherapy and radiation therapy.^{16,43}

Our results showed better OS in the altered group compared to the unaltered group. One caveat in interpreting these results is the tumor type distribution between the altered and unaltered groups. Notably, the altered group included more LGG patients as compared to the unaltered group (406 vs. 124); however, this is unlikely to fully account for our findings, as the altered group also had a higher number of grade 3 brain tumor patients as compared to the unaltered group (177 vs. 32), partially balancing differences in tumor grade distribution. Therefore, we believe the distribution of LGG vs. GBM is probably not a significant factor by itself to account for the OS differences.

Strengths and Limitations :

There are several strengths of our study. First, we have used multiple gene lists that have shown association with obesity in different GWAS databases. Notably, there is not much overlap between these four different obesity-associated gene lists. For instance, the GWAS Catalog-based gene list (primary gene list) had only 7 genes in common with PHGKB, 3 genes in common with PheWeb, and 1 gene in common with Huang *et al.*³⁸ Finding significantly better OS results using different obesity-associated gene lists from different GWAS databases strengthens the accuracy of our findings. Second, our study is unique as we used publicly available databases to study the correlation between obesity-associated genes and brain tumors. To our knowledge, this correlation has not been previously reported. Additionally, this is the largest study examining the effect of obesity-associated genes and brain tumors in the existing literature. Third, we highlight the use of public databases to conduct research by high school students. Research opportunities for high-school and pre-medical undergraduate students remain limited, and many students drop out due to declining interest.⁵⁰ Increasing the availability of research opportunities

can be a way to improve the interest of these students.⁵¹ Using publicly available databases to conduct biomedical research can lead to high-impact research, as we demonstrate here.⁵² Similar methods can be used to ask other important research questions by high-school and undergraduate students.

Our study has some limitations. First, we do not have the BMI data available on these patients, and therefore, we cannot be certain if the obesity-associated genes were more altered in the high BMI patients. However, there are no publicly available databases that have BMI and survival data available for brain tumor patients. Future research will need to be conducted with the collection of height and weight to confirm our findings. Second, this is a retrospective study and may have limitations such as selection bias and confounding. However, we have attempted to minimize selection bias by using a very large cohort and including all patients with brain tumors in the TCGA database. Third, there is potential bias from multiple hypothesis testing. However, we have calculated the false discovery rate and provided q-values where applicable. Fourth, it is possible that we have not been able to account for some of the post-translational effects of obesity on brain tumor patients. However, these studies will require laboratory experimentation and verification, which can be the next step in this area of research. Fifth, we used the GWAS catalog to identify obesity-associated genes based on mapped associations. However, the GWAS catalog primarily reports locus-level variant associations that are subsequently mapped to genes; this approach may introduce annotation bias and limit precision in gene-level analyses. Sixth, we recognize that more comprehensive analyses, such as multivariate approaches including the Cox proportional hazards model, could better account for potential confounding factors as well as pleiotropic effects of genes. However, the analyses presented in this manuscript rely on built-in tools available within cBioPortal, which currently do not support multivariate modeling. Nonetheless, our findings should be considered hypothesis-generating and may inform future investigations. Lastly, it is beyond the scope of the current work to conduct multivariate analyses, as this will require integrating multiple patient-level datasets and applying advanced statistical models, requiring expertise and resources not available for this student-led study.

Our study provides a holistic picture of the importance of obesity-associated genes in the prognosis of brain tumor patients. Given the limitations of our study (as described above) and the complex pathophysiology of obesity, our results are not clinically applicable as such. However, with further prospective research using clinical and biospecimen data collection, there will be a probable improvement in the outcomes of brain tumor patients. Therefore, we would like to propose the collection of such data in brain tumor patients. Additionally, we believe the advent of artificial intelligence (AI) and machine learning will tremendously improve the scope of healthcare research using publicly available databases by high school and undergraduate students.⁵³ Our study would not have been possible without the available tools in cBioPortal. However, cBioPortal only provides basic statistical tools and interactive exploration, but does not currently implement any built-in AI

or machine learning tools. We are hopeful that future improvements in cBioPortal will make it possible to use more advanced AI analytical tools. Using these tools will make analyses of multi-omic datasets and identification of correlative markers more feasible, thereby improving patient outcomes.

■ Conclusion

In conclusion, in a large study using publicly available databases, we demonstrate that brain tumor patients with genomic alterations in obesity-associated genes have better overall survival than those without genomic alterations. Further preclinical and clinical studies are needed to elucidate the mechanisms of how obesity affects the survival outcomes in brain tumor patients. Surrogate markers for obesity (adiposity) such as BMI, fat mass index (FMI), and waist-to-hip ratio (WHR) should be collected during clinical studies in brain tumor patients.⁵⁴ This study highlights the importance of hypothesis-generating research using publicly available databases. This study can serve as a framework for other high-school and undergraduate students to develop and explore meaningful research questions. Lastly, enhancements in cBioPortal with built-in AI and machine learning tools will be immensely helpful to promote future complex research questions.

■ Acknowledgments

I would like to express my sincere gratitude to Dr. Yi Han Ng for her assistance in hypothesis formation, data collection, data interpretation, and manuscript writing. Additionally, I am deeply grateful to my science teachers, Ms. Kelsey Cooper and Dr. Thomas Carpenter, for teaching me the fundamentals of science and scientific writing.

■ References

- Wiedmann, M. K. H.; Brunborg, C.; Di Ieva, A.; Lindemann, K.; Johannesen, T. B.; Vatten, L.; Helseth, E.; Zwart, J. A. The Impact of Body Mass Index and Height on the Risk for Glioblastoma and Other Glioma Subgroups: A Large Prospective Cohort Study. *Neuro-Oncol.* **2017**, *19* (7), 976–985. <https://doi.org/10.1093/neuonc/now272>
- Surveillance Research Program, National Cancer Institute. Cancer Stat Facts: Brain and Other Nervous System Cancer. <https://seer.cancer.gov/statfacts/html/brain.html> (accessed 2025-11-28)
- Yin, J.; Liu, G.; Zhang, Y.; Zhou, Y.; Pan, Y.; Zhang, Q.; Yu, R.; Gao, S. Gender Differences in Gliomas: From Epidemiological Trends to Changes at the Hormonal and Molecular Levels. *Cancer Lett.* **2024**, *598*, 217114. <https://doi.org/10.1016/j.canlet.2024.217114>
- Yang, W.; Warrington, N. M.; Taylor, S. J.; Whitmire, P.; Carrasco, E.; Singleton, K. W.; Wu, N.; Lathia, J. D.; Berens, M. E.; Kim, A. H.; Barnholtz-Sloan, J. S.; Swanson, K. R.; Luo, J.; Rubin, J. B. Sex Differences in GBM Revealed by Analysis of Patient Imaging, Transcriptome, and Survival Data. *Sci. Transl. Med.* **2019**, *11* (473), eao5253. <https://doi.org/10.1126/scitranslmed.aao5253>
- Arulmani, D.; Manickam, D. Brain Tumors. *J. Stud. Res.* **2024**, *13* (2). <https://doi.org/10.47611/jsrhs.v13i2.6694>
- Vienne-Jumeau, A.; Tafani, C.; Ricard, D. Environmental Risk Factors of Primary Brain Tumors: A Review. *Rev. Neurol. (Paris)* **2019**, *175* (10), 664–678. <https://doi.org/10.1016/j.neurol.2019.08.004>
- Ahn, S.; Han, K.; Lee, J.-E.; Jeun, S.-S.; Park, Y.-M.; Yang, S. H. Associations of General and Abdominal Obesity with the Risk of Glioma Development. *Cancers* **2021**, *13* (12), 2859. <https://doi.org/10.3390/cancers13122859>
- Chamarthi, V. S.; Daley, S. F. Comprehensive Assessment and Diagnosis of Metabolic and Biomechanical Complications in Obesity. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2025.
- Ahn, S.; Han, K.; Lee, J.-E.; Jeun, S.-S.; Park, Y.-M.; Yang, S. H. Associations of General and Abdominal Obesity with the Risk of Glioma Development. *Cancers* **2021**, *13* (12), 2859. <https://doi.org/10.3390/cancers13122859>
- Sergentanis, T. N.; Tsvigoulis, G.; Perlepe, C.; Ntanas-Stathopoulos, I.; Tzanninis, I.-G.; Sergentanis, I. N.; Psaltopoulou, T. Obesity and Risk for Brain/CNS Tumors, Gliomas and Meningiomas: A Meta-Analysis. *PLOS ONE* **2015**, *10* (9), e0136974. <https://doi.org/10.1371/journal.pone.0136974>
- Zhang, D.; Chen, J.; Wang, J.; Gong, S.; Jin, H.; Sheng, P.; Qi, X.; Lv, L.; Dong, Y.; Hou, L. Body Mass Index and Risk of Brain Tumors: A Systematic Review and Dose-Response Meta-Analysis. *Eur. J. Clin. Nutr.* **2016**, *70* (7), 757–765. <https://doi.org/10.1038/ejcn.2016.4>
- Dai, H.; Li, Y.; Lee, Y. A.; Lu, Y.; George, T. J.; Donahoo, W. T.; Lee, K. P.; Nakshatri, H.; Allen, J.; Guo, Y.; Sun, R. C.; Guo, J.; Bian, J. GLP-1 Receptor Agonists and Cancer Risk in Adults With Obesity. *JAMA Oncol.* **2025**, *11* (10), 1186–1193. <https://doi.org/10.1001/jamaoncol.2025.2681>
- Obesity and Brain Tumors. In *The Link Between Obesity and Cancer*, Academic Press, 2023; pp 203–217. <https://doi.org/10.1016/B978-0-323-90965-5.00011-8>
- Valente Aguiar, P.; Carvalho, B.; Vaz, R.; Linhares, P. Body Mass Index as an Independent Prognostic Factor in Glioblastoma. *Cancer Causes Control* **2021**, *32* (4), 327–336. <https://doi.org/10.1007/s10552-020-01388-9>
- Wang, J.; Hao, Z.; Li, R.; Wu, W.; Huang, N.; Zhang, K.; Hao, S.; Feng, J.; Chu, J.; Ji, N. Association of Body Mass Index with Clinical Outcome of Primary WHO Grade 4 Glioma. *Front. Oncol.* **2024**, *14*. <https://doi.org/10.3389/fonc.2024.1318785>
- Trestini, I.; Carbognin, L.; Bonaiuto, C.; Tortora, G.; Bria, E. The Obesity Paradox in Cancer: Clinical Insights and Perspectives. *Eat. Weight Disord. EWD* **2018**, *23* (2), 185–193. <https://doi.org/10.1007/s40519-018-0489-y>
- Lennon, H.; Sperrin, M.; Badrick, E.; Renehan, A. G. The Obesity Paradox in Cancer: A Review. *Curr. Oncol. Rep.* **2016**, *18* (9), 56. <https://doi.org/10.1007/s11912-016-0539-4>
- Park, Y.; Peterson, L. L.; Colditz, G. A. The Plausibility of the Obesity Paradox in Cancer-Point. *Cancer Res.* **2018**, *78* (8), 1898–1903. <https://doi.org/10.1158/0008-5472.CAN-17-3043>
- Naik, A.; Monjazebe, A. M.; Decock, J. The Obesity Paradox in Cancer, Tumor Immunology, and Immunotherapy: Potential Therapeutic Implications in Triple Negative Breast Cancer. *Front. Immunol.* **2019**, *10*, 1940. <https://doi.org/10.3389/fimmu.2019.01940>
- Kim, Y.-K.; Song, J. Metabolic Imbalance and Brain Tumors: The Interlinking Metabolic Pathways and Therapeutic Actions of Antidiabetic Drugs. *Pharmacol. Res.* **2025**, *215*, 107719. <https://doi.org/10.1016/j.phrs.2025.107719>
- O'Flanagan, C. H.; Bowers, L. W.; Allott, E. H.; Hursting, S. D. Molecular and Metabolic Mechanisms Underlying the Obesity–Cancer Link. In *Energy Balance and Obesity*; Romieu, I., Dossus, L., Willett, W. C., Eds.; IARC Working Group Reports; International Agency for Research on Cancer: Lyon (FR), 2017.
- Goodarzi, M. O. Genetics of Obesity: What Genetic Association Studies Have Taught Us about the Biology of Obesity and Its

- Complications. *Lancet Diabetes Endocrinol.* **2018**, 6 (3), 223–236. [https://doi.org/10.1016/S2213-8587\(17\)30200-0](https://doi.org/10.1016/S2213-8587(17)30200-0)
23. Wang, K.; Li, W.-D.; Zhang, C. K.; Wang, Z.; Glessner, J. T.; Grant, S. F. A.; Zhao, H.; Hakonarson, H.; Price, R. A. A Genome-Wide Association Study on Obesity and Obesity-Related Traits. *PLoS ONE* **2011**, 6 (4), e18939. <https://doi.org/10.1371/journal.pone.0018939>
 24. Krishnan, M.; Anwar, M. Y.; Justice, A. E.; Chittoor, G.; Chen, H.-H.; Roshani, R.; Smit, R. A. J.; Preuss, M. H.; Chami, N.; Hadad, B. S.; Parra, E. J.; Cruz, M.; Hui, Q.; Wilson, P. W. F.; Sun, Y. V.; Zhang, X.; Linchangco, G. V.; Kardia, S. L. R.; Faul, J. D.; Weir, D. R.; Bielak, L. F.; Highland, H. M.; Young, K. L.; Qi, B.; Wang, Y.; Fornage, M.; Haiman, C.; Cheng, I.; Peters, U.; Kooperberg, C.; Buyske, S.; McCormick, J. B.; Fisher-Hoch, S. P.; Lona-Durazo, F.; Peralta, J.; Gomez-Zamudio, J.; Rich, S. S.; Ferrier, K. R.; Lange, E. M.; Gignoux, C. G.; Kenny, E. E.; Wojcik, G. L.; Cho, K.; Gaziano, M. J.; Djousse, L.; Liu, S.; Vaidya, D.; Mutsert, R. de; Josyula, N. S.; Bauer, C. R.; Zhao, W.; Walker, R. W.; Smith, J. A.; Lange, L. A.; Meyer, M. C.; Liu, C.-T.; Yanek, L. R.; Lee, M.; Raffield, L. M.; Loos, R. J. F.; Gordon-Larsen, P.; Below, J. E.; North, K. E.; Graff, M. Genome-Wide Association Study Provides Novel Insight into the Genetic Architecture of Severe Obesity. *medRxiv* February 26, 2025, p 2025.02.25.25322870. <https://doi.org/10.1101/2025.02.25.25322870>
 25. Krishnan, M.; Anwar, M. Y.; Justice, A. E.; Chittoor, G.; Chen, H.-H.; Roshani, R.; Scartozzi, A.; Dickerson, R. R.; Smit, R. A. J.; Preuss, M. H.; Chami, N.; Hadad, B. S.; Parra, E. J.; Cruz, M.; Hui, Q.; Wilson, P. W. F.; Sun, Y. V.; Zhang, X.; Linchangco, G. V.; Kardia, S. L. R.; Faul, J. D.; Weir, D. R.; Bielak, L. F.; Highland, H. M.; Young, K. L.; Qi, B.; Wang, Y.; Fornage, M.; Haiman, C.; Cheng, I.; Peters, U.; Kooperberg, C.; Buyske, S.; McCormick, J. B.; Fisher-Hoch, S. P.; Lona-Durazo, F.; Peralta, J.; Gomez-Zamudio, J.; Rich, S. S.; Ferrier, K. R.; Lange, E. M.; Gignoux, C. R.; Kenny, E. E.; Wojcik, G. L.; Cho, K.; Gaziano, M. J.; Djousse, L.; Liu, S.; Vaidya, D.; de Mutsert, R.; Josyula, N. S.; Bauer, C. R.; Zhao, W.; Walker, R. W.; Smith, J. A.; Lange, L. A.; Meyer, M. C.; Liu, C.-T.; Yanek, L. R.; Lee, M.; Raffield, L. M.; Loos, R. J. F.; Gordon-Larsen, P.; Below, J. E.; North, K. E.; Graff, M. Genome-Wide Association Study Provides Novel Insight into the Genetic Architecture of Severe Obesity. *PLOS Genet.* **2025**, 21 (9), e1011842. <https://doi.org/10.1371/journal.pgen.1011842>
 26. Pulit, S. L.; Stoneman, C.; Morris, A. P.; Wood, A. R.; Glastonbury, C. A.; Tyrrell, J.; Yengo, L.; Ferreira, T.; Marouli, E.; Ji, Y.; Yang, J.; Jones, S.; Beaumont, R.; Croteau-Chonka, D. C.; Winkler, T. W.; Hattersley, A. T.; Loos, R. J. F.; Hirschhorn, J. N.; Visscher, P. M.; Frayling, T. M.; Yaghootkar, H.; Lindgren, C. M. Meta-Analysis of Genome-Wide Association Studies for Body Fat Distribution in 694 649 Individuals of European Ancestry. *Hum. Mol. Genet.* **2019**, 28 (1), 166–174. <https://doi.org/10.1093/hmg/ddy327>
 27. Pei, Y.-F.; Zhang, L.; Liu, Y.; Li, J.; Shen, H.; Liu, Y.-Z.; Tian, Q.; He, H.; Wu, S.; Ran, S.; Han, Y.; Hai, R.; Lin, Y.; Zhu, J.; Zhu, X.-Z.; Papasian, C. J.; Deng, H.-W. Meta-Analysis of Genome-Wide Association Data Identifies Novel Susceptibility Loci for Obesity. *Hum. Mol. Genet.* **2014**, 23 (3), 820–830. <https://doi.org/10.1093/hmg/ddt464>
 28. Ang, M. Y.; Takeuchi, F.; Kato, N. Deciphering the Genetic Landscape of Obesity: A Data-Driven Approach to Identifying Plausible Causal Genes and Therapeutic Targets. *J. Hum. Genet.* **2023**, 68 (12), 823–833. <https://doi.org/10.1038/s10038-023-01189-3>
 29. Yengo, L.; Sidorenko, J.; Kemper, K. E.; Zheng, Z.; Wood, A. R.; Weedon, M. N.; Frayling, T. M.; Hirschhorn, J.; Yang, J.; Visscher, P. M.; GIANT Consortium. Meta-Analysis of Genome-Wide Association Studies for Height and Body Mass Index in ~700000 Individuals of European Ancestry. *Hum. Mol. Genet.* **2018**, 27 (20), 3641–3649. <https://doi.org/10.1093/hmg/ddy271>
 30. Xu, M.; Li, M.; Zhang, Y.; Li, L.; Shen, Y.; Hu, G. Contributions of Clinical Obesity and Preclinical Obesity to the All-Cause Mortality Risk: Findings From the UK Biobank Cohort. *Diabetes Metab. Res. Rev.* **2025**, 41 (7), e70095. <https://doi.org/10.1002/dmrr.70095>
 31. Smit, R. A. J.; Wade, K. H.; Hui, Q.; Arias, J. D.; Yin, X.; Christiansen, M. R.; Yengo, L.; Preuss, M. H.; Nakabuye, M.; Rocheleau, G.; Graham, S. E.; Buchanan, V. L.; Chittoor, G.; Graff, M.; Guindo-Martínez, M.; Lu, Y.; Marouli, E.; Sakaue, S.; Spracklen, C. N.; Vedantam, S.; Wilson, E. P.; Chen, S.-H.; Ferreira, T.; Ji, Y.; Karaderi, T.; Lüll, K.; Machado, M.; Malden, D. E.; Medina-Gomez, C.; Moore, A.; Rüeger, S.; Akiyama, M.; Allison, M. A.; Alvarez, M.; Andersen, M. K.; Appadurai, V.; Arbeeve, L.; Bartell, E.; Bhaskar, S.; Bielak, L. F.; Bis, J. C.; Bollepalli, S.; Bork-Jensen, J.; Bradfield, J. P.; Bradford, Y.; Brandl, C.; Braund, P. S.; Brody, J. A.; Broeckel, U.; Burgdorf, K. S.; Cade, B. E.; Cai, Q.; Camarda, S.; Campbell, A.; Cañadas-Garre, M.; Chai, J.-F.; Chesi, A.; Choi, S. H.; Christofidou, P.; Couture, C.; Cuellar-Partida, G.; Danning, R.; Degenhardt, F.; Delgado, G. E.; Delitala, A.; Demirkan, A.; Deng, X.; Dietl, A.; Dimitriou, M.; Dimitrov, L.; Dorajoo, R.; Eichelmann, F.; Eliassen, A. U.; Engmann, J. E.; Erdos, M. R.; Fairhurst-Hunter, Z.; Farmaki, A.-E.; Faul, J. D.; Fernandez-Lopez, J.-C.; Forer, L.; Frank, M.; Freitag-Wolf, S.; Fritsche, L. G.; Fuchsberger, C.; Galesloot, T. E.; Gao, Y.; Geller, F.; Giannakopoulou, O.; Giulianini, F.; Gjesing, A. P.; Goel, A.; Gordon, S. D.; Gorski, M.; Grove, J.; Guo, X.; Gustafsson, S.; Haessler, J.; Hansen, T. F.; Havulinna, A. S.; Haworth, S. J.; Heard-Costa, N.; Hemerich, D.; Highland, H. M.; Hindy, G.; Ho, Y.-L.; Hofer, E.; Holliday, E.; Horn, K.; Hornsby, W. E.; Hottenga, J.-J.; Huang, H.; Huang, J.; Huerta-Chagoya, A.; Huo, S.; Hwang, M. Y.; Hwu, C.-M.; Iha, H.; Ikeda, D. D.; Isono, M.; Jackson, A. U.; Jansen, I. E.; Jiang, Y.; Johansson, I.; Jonsson, A.; Jørgensen, T.; Kalafati, I. P.; Kanai, M.; Kanoni, S.; Kärhus, L. L.; Kasturiratne, A.; Katsuya, T.; Kawaguchi, T.; Kember, R. L.; Kentistou, K. A.; Kim, D.; Kim, H.-N.; Kim, Y. J.; Kleber, M. E.; Knol, M. J.; Kurbasic, A.; Lauzon, M.; Le, P.; Lea, R.; Lee, J.-Y.; Lee, W.-J.; Leonard, H. L.; Li, H.; Li, S. A.; Li, X.; Li, X.; Liang, J.; Lin, H.; Lin, K.; Liu, J.; Liu, X.; Lo, K. S.; Long, J.; Lores-Motta, L.; Luan, J.; Lyssenko, V.; Lyytikäinen, L.-P.; Mahajan, A.; Malik, M. Z.; Mamakou, V.; Mangino, M.; Manichaikul, A.; Marten, J.; Mattheisen, M.; McDaid, A. F.; Mei, Q.; Meiselbach, H.; Melendez, T. L.; Milaneschi, Y.; Miller, J. E.; Millwood, I. Y.; Mishra, P. P.; Mitchell, R. E.; Møllehave, L. T.; Mononen, N.; Mucha, S.; Munz, M.; Mykkänen, J.; Nakatochi, M.; Nardone, G. G.; Nelson, C. P.; Nethander, M.; Nho, C. W.; Nielsen, A. A.; Nolte, I. M.; Nongmaithem, S. S.; Noordam, R.; Ntalla, I.; Nutile, T.; Pandit, A.; Pauper, M.; Petersen, E. R. B.; Petersen, L. V.; Piluso, F.; Polašek, O.; Poveda, A.; Pyarajan, S.; Raffield, L. M.; Rakugi, H.; Ramirez, J.; Rasheed, A.; Raven, D.; Rayner, N. W.; Riveros, C.; Rohde, R.; Ruggiero, D.; Ruotsalainen, S. E.; Ryan, K. A.; Sabater-Lleal, M.; Santin, A.; Saxena, R.; Scholz, M.; Shen, B.; Shi, J.; Shin, J. H.; Sidore, C.; Sidorenko, J.; Sim, X.; Sliker, R. C.; Smith, A. V.; Smith, J. A.; Smyth, L. J.; Southam, L.; Steinthorsdottir, V.; Sun, L.; Takeuchi, F.; Taylor, K. D.; Tayo, B. O.; Tcheandjie, C.; Terzikhan, N.; Tesolin, P.; Teumer, A.; Theusch, E.; Thompson, D. J.; Thorleifsson, G.; Timmers, P. R. H. J.; Trompet, S.; Turman, C.; Vaccargiu, S.; van der Laan, S. W.; van der Most, P. J.; van Klinken, J. B.; van Setten, J.; Verma, S. S.; Verweij, N.; Veturi, Y.; Wang, C. A.; Wang, C.; Wang, J.-S.; Wang, L.; Wang, Y. X.; Wang, Z.; Warren, H. R.; Bin Wei, W.; Wen, W.; Wheeler, W. A.; Wickremasinghe, A. R.; Wielscher, M.; Winsvold, B. S.; Wong, A.; Wuttke, M.; Xia, R.; Yamamoto, K.; Yang, J.; Yao, J.; Young, H.;

- Yousri, N. A.; Yu, L.; Zeng, L.; Zhang, W.; Zhang, X.; Zhao, J.-H.; Zhao, W.; Zhou, W.; Zimmermann, M. E.; Zoledziewska, M.; t Hart, L. M.; Adair, L. S.; Adams, H. H. H.; Aguilar-Salinas, C. A.; Al-Mulla, F.; Arnett, D. K.; Asselbergs, F. W.; Åsvold, B. O.; Attia, J.; Banas, B.; Bandinelli, S.; Beilin, L. J.; Bennett, D. A.; Bergler, T.; Bharadwaj, D.; Biino, G.; Boerwinkle, E.; Böger, C. A.; Borja, J. B.; Bouchard, C.; Bowden, D. W.; Brandslund, I.; Brumpton, B.; Buring, J. E.; Caulfield, M. J.; Chambers, J. C.; Chandak, G. R.; Chanock, S. J.; Chaturvedi, N.; Ida Chen, Y.-D.; Chen, Z.; Cheng, C.-Y.; Cho, Y. S.; Christensen, K.; Christophersen, I. E.; Ciullo, M.; Cole, J. W.; Collins, F. S.; Concas, M. P.; Cooper, R. S.; Cruz, M.; Cucca, F.; Cutler, M. J.; Damrauer, S. M.; Dantoft, T. M.; de Borst, G. J.; de Geus, E. J. C.; de Groot, L. C. P. G. M.; De Jager, P. L.; de Kleijn, D. P. V.; de Silva, H. J.; Dedoussis, G. V.; den Hollander, A. I.; Du, S.; Easton, D. F.; Eckardt, K.-U.; Elders, P. J. M.; Eliassen, A. H.; Ellinor, P. T.; Elmstahl, S.; Erdmann, J.; Evans, M. K.; Fatkin, D.; Feenstra, B.; Feitosa, M. F.; Ferrucci, L.; Florez, J. C.; Ford, I.; Fornage, M.; Franke, A.; Franks, P. W.; Freedman, B. I.; Gieger, C.; Girotto, G.; Golightly, Y. M.; Gonzalez-Villalpando, C.; Gordon-Larsen, P.; Grallert, H.; Grant, S. F. A.; Grarup, N.; Griffiths, L.; Gudnason, V.; Haiman, C.; Hakonarson, H.; Hansen, T.; Hartman, C. A.; Hattersley, A. T.; Hayward, C.; Heid, I. M.; Heng, C.-K.; Hengstenberg, C.; Herzig, K.-H.; Hewitt, A. W.; Hishigaki, H.; Hougaard, D. M.; Hoyng, C. B.; Huang, P. L.; Huang, W.; Huang, W.-Y.; Huffman, J. E.; Hunt, S. C.; Hutri, N.; Hveem, K.; Hyppönen, E.; Iacono, W. G.; Ichihara, S.; Ikram, M. A.; Isasi, C. R.; Jarvelin, M.-R.; Jin, Z.-B.; Jöckel, K.-H.; Jonas, J. B.; Joshi, P. K.; Jousilahti, P.; Jukema, J. W.; Kähönen, M.; Kamatani, Y.; Kang, K. D.; Kaprio, J.; Kardia, S. L. R.; Karpe, F.; Kato, N.; Kavousi, M.; Kee, F.; Kessler, T.; Khera, A. V.; Khor, C. C.; Kiemeny, L. A. L. M.; Kim, B.-J.; Kim, E. K.; Kim, H.-L.; Kirchhof, P.; Kivimäki, M.; Koh, W.-P.; Koistinen, H. A.; Kokkinos, A.; Kooner, J. S.; Kooperberg, C.; Kovacs, P.; Kraaijeveld, A.; Kraft, P.; Krauss, R. M.; Kumari, M.; Kutalik, Z.; Laakso, M.; Lange, L. A.; Langenberg, C.; Launer, L. J.; Lee, H.; Lee, N. R.; Lehtimäki, T.; Lemaitre, R. N.; Li, H.; Li, L.; Lieb, W.; Lin, X.; Lind, L.; Linneberg, A.; Liu, C.-T.; Liu, J.; Loeffler, M.; London, B.; Lu, F.; Lubitz, S. A.; Mackey, D. A.; Magnusson, P. K. E.; Manson, J. E.; Marcus, G. M.; Marques Vidal, P.; Martin, N. G.; März, W.; Matsuda, F.; McCarthy, M. I.; McGarrah, R. W.; McGue, M.; McKnight, A. J.; Medland, S. E.; Mellström, D.; Metspalu, A.; Mitchell, B. D.; Mitchell, P.; Mook-Kanamori, D. O.; Mori, T. A.; Morris, A. D.; Mucci, L. A.; Munroe, P. B.; Nalls, M. A.; Nazarian, S.; Nelson, A. E.; Neville, M. J.; Newton-Cheh, C.; Nielsen, C. S.; Niinikoski, H.; Nikus, K.; Nöthen, M. M.; Ogunniyi, A.; Ohlsson, C.; Oldehinkel, A. J.; Orozco, L.; Pahkala, K.; Pajukanta, P.; Palmer, C. N. A.; Parra, E. J.; Pattaro, C.; Pedersen, O.; Pennell, C. E.; Penninx, B. W. J. H.; Perusse, L.; Peters, A.; Peyser, P. A.; Porteous, D. J.; Posthuma, D.; Power, C.; Pramstaller, P. P.; Province, M. A.; Psaty, B. M.; Qi, Q.; Qu, J.; Rader, D. J.; Raitakari, O. T.; Rallidis, L. S.; Rao, D. C.; Redline, S.; Reilly, D. F.; Reiner, A. P.; Rhee, S. Y.; Ridker, P. M.; Rienstra, M.; Ripatti, S.; Ritchie, M. D.; Rivadeneira, F.; Roden, D. M.; Rosendaal, F. R.; Rotter, J. I.; Rudan, I.; Rutter, F.; Ryu, S.; Sabanayagam, C.; Salako, B.; Saleheen, D.; Salomaa, V.; Samani, N. J.; Sanghera, D. K.; Sattar, N.; Schmidt, B.; Schmidt, H.; Schmidt, R.; Schulze, M. B.; Schunkert, H.; Scott, L. J.; Scott, R. J.; Sever, P.; Sheu, W. H. H.; Shoemaker, M. B.; Shu, X.-O.; Simonsick, E. M.; Sims, M.; Singleton, A. B.; Sinner, M. F.; Smith, J. G.; Snieder, H.; Spector, T. D.; Spedicati, B.; Stampfer, M. J.; Stark, K. J.; Strachan, D. P.; Tabara, Y.; Tai, E. S.; Tang, H.; Tardif, J.-C.; Thanaraj, T. A.; Tönjes, A.; Tuomi, T.; Tuomilehto, J.; Tusié-Luna, M.-T.; van Dam, R. M.; van der Harst, P.; Van der Velde, N.; van Duijn, C. M.; van Schoor, N. M.; Vitart, V.; Vohl, M.-C.; Völker, U.; Vollenweider, P.; Völzke, H.; Vrieze, S.; Wacher-Rodarte, N. H.; Walker, M.; Wandler, G. S.; Wareham, N. J.; Watanabe, R. M.; Watkins, H.; Weir, D. R.; Werge, T. M.; Widen, E.; Willemsen, G.; Willett, W. C.; Wilson, J. F.; Wilson, P. W. F.; Wong, T. Y.; Woo, J.-T.; Wright, A. F.; Xu, H.; Yajnik, C. S.; Yang, J.; Yokota, M.; Yuan, J.-M.; Zeggini, E.; Zemel, B. S.; Zheng, W.; Zhu, X.; Zillikens, M. C.; Zonderman, A. B.; Zwart, J.-A.; Abecasis, G. R.; Assimes, T. L.; Auton, A.; Boehnke, M.; Chasman, D. I.; Esko, T.; Stefansson, K.; Lettre, G.; Lindgren, C. M.; Ng, M. C. Y.; O'Donnell, C. J.; Thorsteinsdottir, U.; Visscher, P. M.; Walters, R. G.; Winkler, T. W.; Wood, A. R.; Deloukas, P.; Frayling, T. M.; Justice, A. E.; Kilpeläinen, T. O.; Locke, A. E.; Mohlke, K. L.; North, K. E.; Okada, Y.; Willer, C. J.; Young, K. L.; Fatumo, S.; McCaffery, J. M.; Timpson, N. J.; Hirschhorn, J. N.; Sun, Y. V.; Berndt, S. I.; Loos, R. J. F. Polygenic Prediction of Body Mass Index and Obesity through the Life Course and across Ancestries. *Nat. Med.* **2025**, *31* (9), 3151–3168. <https://doi.org/10.1038/s41591-025-03827-z>
32. Larsson, S. C.; Burgess, S. Causal Role of High Body Mass Index in Multiple Chronic Diseases: A Systematic Review and Meta-Analysis of Mendelian Randomization Studies. *BMC Med.* **2021**, *19* (1), 320. <https://doi.org/10.1186/s12916-021-02188-x>
33. The Cancer Genome Atlas Research Network. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N. Engl. J. Med.* **2015**, *372* (26), 2481–2498. <https://doi.org/10.1056/NEJMoa1402121>
34. Brennan, C. W.; Verhaak, R. G. W.; McKenna, A.; Campos, B.; Nounshmehr, H.; Salama, S. R.; Zheng, S.; Chakravarty, D.; Sanborn, J. Z.; Berman, S. H.; Beroukhi, R.; Bernard, B.; Wu, C.-J.; Genovese, G.; Shmulevich, I.; Barnholtz-Sloan, J.; Zou, L.; Vegesna, R.; Shukla, S. A.; Ciriello, G.; Yung, W.; Zhang, W.; Sougnez, C.; Mikkelson, T.; Aldape, K.; Bigner, D. D.; Van Meir, E. G.; Prados, M.; Sloan, A.; Black, K. L.; Eschbacher, J.; Finocchiaro, G.; Friedman, W.; Andrews, D. W.; Guha, A.; Iacocca, M.; O'Neill, B. P.; Foltz, G.; Myers, J.; Weisenberger, D. J.; Penny, R.; Kucherlapati, R.; Perou, C. M.; Hayes, D. N.; Gibbs, R.; Marra, M.; Mills, G. B.; Lander, E.; Spellman, P.; Wilson, R.; Sander, C.; Weinstein, J.; Meyerson, M.; Gabriel, S.; Laird, P. W.; Haussler, D.; Getz, G.; Chin, L. The Somatic Genomic Landscape of Glioblastoma. *Cell* **2013**, *155* (2), 462–477. <https://doi.org/10.1016/j.cell.2013.09.034>
35. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B. E.; Sumer, S. O.; Aksoy, B. A.; Jacobsen, A.; Byrne, C. J.; Heuer, M. L.; Larsson, E.; Antipin, Y.; Reva, B.; Goldberg, A. P.; Sander, C.; Schultz, N. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov.* **2012**, *2* (5), 401–404. <https://doi.org/10.1158/2159-8290.CD-12-0095>
36. Gao, J.; Aksoy, B. A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S. O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; Cerami, E.; Sander, C.; Schultz, N. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* **2013**, *6* (269), pl1. <https://doi.org/10.1126/scisignal.2004088>
37. de Bruijn, I.; Kundra, R.; Mastrogiacommo, B.; Tran, T. N.; Sikin, L.; Mazor, T.; Li, X.; Ochoa, A.; Zhao, G.; Lai, B.; Abeshouse, A.; Baiceanu, D.; Ciftci, E.; Dogrusoz, U.; Dufilie, A.; Erkoç, Z.; Garcia Lara, E.; Fu, Z.; Gross, B.; Haynes, C.; Heath, A.; Higgins, D.; Jagannathan, P.; Kalletta, K.; Kumari, P.; Lindsay, J.; Lisman, A.; Leenknegt, B.; Lukasse, P.; Madela, D.; Madupuri, R.; van Nierop, P.; Plantalech, O.; Quach, J.; Resnick, A. C.; Rodenburg, S. Y. A.; Satravada, B. A.; Schaeffer, F.; Sheridan, R.; Singh, J.; Sirohi, R.; Sumer, S. O.; van Hagen, S.; Wang, A.; Wilson, M.; Zhang, H.; Zhu, K.; Rusk, N.; Brown, S.; Lavery, J. A.; Panageas, K. S.; Rudolph, J. E.; LeNoue-Newton, M. L.; Warner, J. L.; Guo, X.; Hunter-Zinck, H.; Yu, T. V.; Pilai, S.; Nichols, C.; Gardos, S. M.; Philip, J.; AACR Project GENIE BPC Core Team, AACR Project

- GENIE Consortium; Kehl, K. L.; Riely, G. J.; Schrag, D.; Lee, J.; Fiandalo, M. V.; Sweeney, S. M.; Pugh, T. J.; Sander, C.; Cerami, E.; Gao, J.; Schultz, N. Analysis and Visualization of Longitudinal Genomic and Clinical Data from the AACR Project GENIE Biopharma Collaborative in cBioPortal. *Cancer Res.* **2023**, *83* (23), 3861–3867. <https://doi.org/10.1158/0008-5472.CAN-23-0816>
38. Huang, F.; Xu, P.; Yue, Z.; Song, Y.; Hu, K.; Zhao, X.; Gao, M.; Chong, Z. Body Weight Correlates with Molecular Variances in Patients with Cancer. *Cancer Res.* **2024**, *84* (5), 757–770. <https://doi.org/10.1158/0008-5472.CAN-23-1463>
 39. Heath, A. P.; Ferretti, V.; Agrawal, S.; An, M.; Angelakos, J. C.; Arya, R.; Bajari, R.; Baqar, B.; Barnowski, J. H. B.; Burt, J.; Catton, A.; Chan, B. F.; Chu, F.; Cullion, K.; Davidsen, T.; Do, P.-M.; Dompierre, C.; Ferguson, M. L.; Fitzsimons, M. S.; Ford, M.; Fukuma, M.; Gaheen, S.; Ganji, G. L.; Garcia, T. I.; George, S. S.; Gerhard, D. S.; Gerthoffert, F.; Gomez, F.; Han, K.; Hernandez, K. M.; Issac, B.; Jackson, R.; Jensen, M. A.; Joshi, S.; Kadam, A.; Khurana, A.; Kim, K. M. J.; Kraft, V. E.; Li, S.; Lichtenberg, T. M.; Lodato, J.; Lolla, L.; Martinov, P.; Mazzone, J. A.; Miller, D. P.; Miller, I.; Miller, J. S.; Miyauchi, K.; Murphy, M. W.; Nullet, T.; Ogwara, R. O.; Ortuño, F. M.; Pedrosa, J.; Pham, P. L.; Popov, M. Y.; Porter, J. J.; Powell, R.; Rademacher, K.; Reid, C. P.; Rich, S.; Rogel, B.; Sahni, H.; Savage, J. H.; Schmitt, K. A.; Simmons, T. J.; Sislow, J.; Spring, J.; Stein, L.; Sullivan, S.; Tang, Y.; Thiagarajan, M.; Troyer, H. D.; Wang, C.; Wang, Z.; West, B. L.; Wilmer, A.; Wilson, S.; Wu, K.; Wysocki, W. P.; Xiang, L.; Yamada, J. T.; Yang, L.; Yu, C.; Yung, C. K.; Zenklusen, J. C.; Zhang, J.; Zhang, Z.; Zhao, Y.; Zubair, A.; Staudt, L. M.; Grossman, R. L. The NCI Genomic Data Commons. *Nat. Genet.* **2021**, *53* (3), 257–262. <https://doi.org/10.1038/s41588-021-00791-5>
 40. MacArthur, J.; Bowler, E.; Cerezo, M.; Gil, L.; Hall, P.; Hastings, E.; Junkins, H.; McMahon, A.; Milano, A.; Morales, J.; Pendlington, Z. M.; Welter, D.; Burdett, T.; Hindorf, L.; Flicek, P.; Cunningham, F.; Parkinson, H. The New NHGRI-EBI Catalog of Published Genome-Wide Association Studies (GWAS Catalog). *Nucleic Acids Res.* **2017**, *45* (D1), D896–D901. <https://doi.org/10.1093/nar/gkw1133>
 41. Yu, W.; Gwinn, M.; Dotson, W. D.; Green, R. F.; Clyne, M.; Wulf, A.; Bowen, S.; Kolor, K.; Khoury, M. J. A Knowledge Base for Tracking the Impact of Genomics on Population Health. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **2016**, *18* (12), 1312–1314. <https://doi.org/10.1038/gim.2016.63>
 42. Exploring and visualizing large-scale genetic associations by using PheWeb | *Nature Genetics*. <https://www.nature.com/articles/s41588-020-0622-5> (accessed 2025-11-30)
 43. Cha, J.-Y.; Park, J.-S.; Hong, Y.-K.; Jeun, S.-S.; Ahn, S. Impact of Body Mass Index on Survival Outcome in Patients with Newly Diagnosed Glioblastoma: A Retrospective Single-Center Study. *Integr. Cancer Ther.* **2021**, *20*, 1534735421991233. <https://doi.org/10.1177/1534735421991233>
 44. Schiffmann, J.; Karakiewicz, P. I.; Rink, M.; Manka, L.; Salomon, G.; Tilki, D.; Budäus, L.; Pompe, R.; Leyh-Bannurah, S.-R.; Haese, A.; Hammerer, P.; Huland, H.; Graefen, M.; Tennstedt, P. Obesity Paradox in Prostate Cancer: Increased Body Mass Index Was Associated with Decreased Risk of Metastases after Surgery in 13,667 Patients. *World J. Urol.* **2018**, *36* (7), 1067–1072. <https://doi.org/10.1007/s00345-018-2240-8>
 45. Zhang, X.; Liu, Y.; Shao, H.; Zheng, X. Obesity Paradox in Lung Cancer Prognosis: Evolving Biological Insights and Clinical Implications. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* **2017**, *12* (10), 1478–1488. <https://doi.org/10.1016/j.jtho.2017.07.022>
 46. Ligibel, J. A.; Alfano, C. M.; Courneya, K. S.; Demark-Wahnefried, W.; Burger, R. A.; Chlebowski, R. T.; Fabian, C. J.; Gucalp, A.; Hershman, D. L.; Hudson, M. M.; Jones, L. W.; Kakarala, M.; Ness, K. K.; Merrill, J. K.; Wollins, D. S.; Hudis, C. A. American Society of Clinical Oncology Position Statement on Obesity and Cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2014**, *32* (31), 3568–3574. <https://doi.org/10.1200/JCO.2014.58.4680>
 47. Lee, J.; Meyerhardt, J. A.; Giovannucci, E.; Jeon, J. Y. Association between Body Mass Index and Prognosis of Colorectal Cancer: A Meta-Analysis of Prospective Cohort Studies. *PLOS ONE* **2015**, *10* (3), e0120706. <https://doi.org/10.1371/journal.pone.0120706>
 48. Chan, D. S. M.; Vieira, A. R.; Aune, D.; Bandera, E. V.; Greenwood, D. C.; McTiernan, A.; Navarro Rosenblatt, D.; Thune, I.; Vieira, R.; Norat, T. Body Mass Index and Survival in Women with Breast Cancer-Systematic Literature Review and Meta-Analysis of 82 Follow-up Studies. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **2014**, *25* (10), 1901–1914. <https://doi.org/10.1093/annonc/mdl042>
 49. Yuan, C.; Bao, Y.; Wu, C.; Kraft, P.; Ogino, S.; Ng, K.; Qian, Z. R.; Robinson, D. A.; Stampfer, M. J.; Giovannucci, E. L.; Wolpin, B. M. Prediagnostic Body Mass Index and Pancreatic Cancer Survival. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2013**, *31* (33), 4229–4234. <https://doi.org/10.1200/JCO.2013.51.7532>
 50. Lin, K. Y.; Parnami, S.; Fuhrel-Forbis, A.; Anspach, R. R.; Crawford, B.; De Vries, R. G. The Undergraduate Premedical Experience in the United States: A Critical Review. *Int. J. Med. Educ.* **2013**, *4*, 26–37. <https://doi.org/10.5116/ijme.5103.a8d3>
 51. Estrada, M.; Burnett, M.; Campbell, A. G.; Campbell, P. B.; Denetclaw, W. F.; Gutiérrez, C. G.; Hurtado, S.; John, G. H.; Matsui, J.; McGee, R.; Okpodu, C. M.; Robinson, T. J.; Summers, M. F.; Werner-Washburne, M.; Zavala, M. Improving Underrepresented Minority Student Persistence in STEM. *CBE Life Sci. Educ.* **2016**, *15* (3), e5. <https://doi.org/10.1187/cbe.16-01-0038>
 52. Hankey, W.; Zanghi, N.; Crow, M. M.; Dow, W. H.; Kratz, A.; Robinson, A. M.; Robinson, M. R.; Segarra, V. A. Using The Cancer Genome Atlas as an Inquiry Tool in the Undergraduate Classroom. *Front. Genet.* **2020**, *11*. <https://doi.org/10.3389/fgene.2020.573992>
 53. Al-Nafjan, A.; Aljuhani, A.; Alshebel, A.; Alharbi, A.; Alshehri, A. Artificial Intelligence in Predictive Healthcare: A Systematic Review. *J. Clin. Med.* **2025**, *14* (19), 6752. <https://doi.org/10.3390/jcm14196752>
 54. Khan, I.; Chong, M.; Le, A.; Mohammadi-Shemirani, P.; Morton, R.; Brinza, C.; Kiflen, M.; Narula, S.; Akhbar, L.; Mao, S.; Morrison, K.; Pigeyre, M.; Paré, G. Surrogate Adiposity Markers and Mortality. *JAMA Netw. Open* **2023**, *6* (9), e2334836. <https://doi.org/10.1001/jamanetworkopen.2023.34836>
 55. Stelzer, G.; Rosen, N.; Plaschkes, I.; Zimmerman, S.; Twik, M.; Fishilevich, S.; Stein, T. I.; Nudel, R.; Lieder, I.; Mazor, Y.; Kaplan, S.; Dahary, D.; Warshawsky, D.; Guan-Golan, Y.; Kohn, A.; Rappaport, N.; Safran, M.; Lancet, D. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinforma.* **2016**, *54*, 1.30.1–1.30.33. <https://doi.org/10.1002/cpbi.5>
 56. Centers for Disease Control and Prevention (CDC). Public Health Genomics and Precision Health Knowledge Base (PHGKB), 2024. <https://phgkb.cdc.gov/PHGKB/phgHome.action>
 57. Gagliano Taliun, S. A.; VandeHaar, P.; Boughton, A. P.; Welch, R. P.; Taliun, D.; Schmidt, E. M.; Zhou, W.; Nielsen, J. B.; Willer, C. J.; Lee, S.; Fritsche, L. G.; Boehnke, M.; Abecasis, G. R. Exploring and Visualizing Large-Scale Genetic Associations by Using PheWeb. *Nat. Genet.* **2020**, *52* (6), 550–552. <https://doi.org/10.1038/s41588-020-0622-5>

■ Author

Aarav Sehdev is a sophomore at Basis Phoenix High School. He is very keen on advancing his understanding of biology and

chemistry through scientific readings and experimentation. He has been elected as the Ambassador for Students Supporting Brain Tumor Research 2026-2027 and the Treasurer for the Science National Honor Society. He is vice-president and editorial board member for the Simplifications Project, which is focused on summarizing complex research papers to a high-school level. He enjoys playing tennis and guitar.

■ Supplemental Tables

Supplemental Table 1: Genome-wide association studies (GWAS) genes associated with morbid obesity.⁵⁵

Gene Symbol	Gene Name	Function
FTO	Alpha-ketoglutarate-dependent dioxygenase	This gene is a nuclear protein of the AlkB-related non-heme iron and 2-oxoglutarate-dependent oxygenase superfamily, but the exact physiological function of this gene is not known. Other non-heme iron enzymes function to reverse alkylated DNA and RNA damage by oxidative demethylation. Studies in mice and humans indicate a role in the nervous and cardiovascular systems and a strong association with body mass index, obesity risk, and type 2 diabetes.
TCF7L2	Transcription Factor 7 Like 2	This gene encodes a high mobility group (HMG) box-containing transcription factor that plays a key role in the Wnt signaling pathway. The protein has been implicated in blood glucose homeostasis. Genetic variants of this gene are associated with increased risk of type 2 diabetes. Several transcript variants encoding multiple different isoforms have been found for this gene.
GIPR	Gastric Inhibitory Polypeptide Receptor	This gene encodes a G-protein-coupled receptor for gastric inhibitory polypeptide (GIP), which was originally identified as an activity in gut extracts that inhibited gastric acid secretion and gastrin release, but subsequently was demonstrated to stimulate insulin release in the presence of elevated glucose. Mice lacking this gene exhibit higher blood glucose levels with impaired initial insulin response after oral glucose load. Defects in this gene may thus contribute to the pathogenesis of diabetes.
LINC01875	Long intergenic Non-Protein coding RNA 1875	LINC01875 (Long Intergenic Non-Protein Coding RNA 1875) is an RNA Gene, and is affiliated with the lncRNA class.
TMEM18	Transmembrane Protein 18	Predicted to enable DNA binding activity. Involved in cell migration. Located in the nuclear membrane.
ETV5	ETS Variant Transcription Factor 5	Enables DNA-binding transcription activator activity, RNA polymerase II-specific, and RNA polymerase II transcription regulatory region sequence-specific DNA binding activity. Involved in cellular response to oxidative stress; negative regulation of transcription by RNA polymerase II; and positive regulation of transcription by RNA polymerase II. Located in the nucleoplasm.
PRDX4P1	Peroxisredoxin 4 Pseudogene 1	PRDX4P1 (Peroxisredoxin 4 Pseudogene 1) is a Pseudogene.
THAP12P9	THAP Domain Containing 12 Pseudogene 9	THAP12P9 (THAP Domain Containing 12 Pseudogene 9) is a Pseudogene.
LYPLAL1-AS1	LYPLAL1 Antisense RNA 1	LYPLAL1-AS1 (LYPLAL1 Antisense RNA 1) is an RNA Gene, and is affiliated with the lncRNA class. Diseases associated with LYPLAL1-AS1 include Developmental and Epileptic Encephalopathy 28.
SLC39A8	Solute Carrier Family 39 Member 8	This gene encodes a member of the SLC39 family of solute-carrier genes, which show structural characteristics of zinc transporters. The encoded protein is glycosylated and found in the plasma membrane and mitochondria, and functions in the cellular import of zinc at the onset of inflammation. It is also thought to be the primary transporter of the toxic cation cadmium, which is found in cigarette smoke. Multiple transcript variants encoding different isoforms have been found for this gene. Additional alternatively spliced transcript variants of this gene have been described, but their full-length nature is not known.
FAIM2	Fas Apoptotic Inhibitory Molecule 2	Predicted to enable calcium channel activity. Involved in the regulation of the neuron apoptotic process. Acts upstream of or within the negative regulation of the extrinsic apoptotic signaling pathway via death domain receptors. Located in the Golgi membrane and the membrane raft.
LINC02796	Long Intergenic Non-Protein Coding RNA 2796	LINC02796 (Long Intergenic Non-Protein Coding RNA 2796) is an RNA Gene, and is affiliated with the lncRNA class.
LINC03111	Long Intergenic Non-Protein Coding RNA 3111	LINC03111 (Long Intergenic Non-Protein Coding RNA 3111) is an RNA Gene, and is affiliated with the lncRNA class.
RNU4-17P	RNA, U4 Small Nuclear 17, Pseudogene	RNU4-17P (RNA, U4 Small Nuclear 17, Pseudogene) is a Pseudogene.
AKAP6	A-Kinase Anchoring Protein 6	The A-kinase anchor proteins (AKAPs) are a group of structurally diverse proteins that have the common function of binding to the regulatory subunit of protein kinase A (PKA) and confining the holoenzyme to discrete locations within the cell. This gene encodes a member of the AKAP family. The encoded protein is highly expressed in various brain regions and cardiac and skeletal muscle. It is specifically localized to the sarcoplasmic reticulum and nuclear membrane, and is involved in anchoring PKA to the nuclear membrane or sarcoplasmic reticulum.
NYAP2	Neuronal Tyrosine-Phosphorylated Phosphoinositide-3-Kinase Adapter 2	Predicted to be involved in neuron projection morphogenesis and phosphatidylinositol 3-kinase/protein kinase B signal transduction.

MIR-5702/5702	MicroRNA 5702	microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into an RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop.
ADCY9	Adenylate Cyclase 9	Adenylate cyclase is a membrane-bound enzyme that catalyzes the formation of cyclic AMP from ATP. It is regulated by a family of G protein-coupled receptors, protein kinases, and calcium. The type 9 adenylyl cyclase is a widely distributed adenylyl cyclase, and it is stimulated by beta-adrenergic receptor activation but is insensitive to forskolin, calcium, and somatostatin.
MAP2K5	Mitogen-Activated Protein Kinase 5	The protein encoded by this gene is a dual specificity protein kinase that belongs to the MAP kinase kinase family. This kinase specifically interacts with and activates MAPK7/ERK5. This kinase itself can be phosphorylated and activated by MAP3K3/MEKK3, as well as by atypical protein kinase C isoforms (aPKCs). The signal cascade mediated by this kinase is involved in growth factor-stimulated cell proliferation and muscle cell differentiation. Three alternatively spliced transcript variants of this gene encoding distinct isoforms have been described.
LINC01741	Long Intergenic Non-Protein Coding RNA 1741	LINC01741 (Long Intergenic Non-Protein Coding RNA 1741) is an RNA Gene, and is affiliated with the lncRNA class.
SEC16B	SEC16 Homolog B, Endoplasmic Reticulum Export Factor	SEC16B is a mammalian homolog of <i>S. cerevisiae</i> Sec16 that is required for the organization of transitional endoplasmic reticulum (ER) sites and protein export
POM121C	POM121 Transmembrane Nucleoporin C	Predicted to enable nuclear localization sequence binding activity. Predicted to be a structural constituent of the nuclear pore. Predicted to be involved in RNA export from the nucleus and protein import into the nucleus. Located in the nuclear membrane and nucleoplasm.
DNAJB4	DnaJ Heat Shock Protein Family (Hsp40) Member B4	The protein encoded by this gene is a molecular chaperone, tumor suppressor, and member of the heat shock protein-40 family. The encoded protein binds the cell adhesion protein E-cadherin and targets it to the plasma membrane. This protein also binds incorrectly folded E-cadherin and targets it for endoplasmic reticulum-associated degradation. This gene is a strong tumor suppressor for colorectal carcinoma, and downregulation of it may serve as a good biomarker for predicting patient outcomes. Several transcript variants encoding different isoforms have been found for this gene.
GIPC2	GIPC PDZ Domain Containing Family Member 2	Enables identical protein binding activity. Located in the extracellular exosome.
TFAP2B	Transcription Factor AP-2 Beta	This gene encodes a member of the AP-2 family of transcription factors. AP-2 proteins form homo- or heterodimers with other AP-2 family members and bind specific DNA sequences. They are thought to stimulate cell proliferation and suppress terminal differentiation of specific cell types during embryonic development. Specific AP-2 family members differ in their expression patterns and binding affinity for different promoters. This protein functions as both a transcriptional activator and repressor. Mutations in this gene result in autosomal dominant Char syndrome, suggesting that this gene functions in the differentiation of neural crest cell derivatives.
MCM6	Minichromosome Maintenance Complex Component 6	The protein encoded by this gene is one of the highly conserved mini-chromosome maintenance proteins (MCM) that are essential for the initiation of eukaryotic genome replication. The hexameric protein complex formed by the MCM proteins is a key component of the pre-replication complex (pre-RC) and may be involved in the formation of replication forks and in the recruitment of other DNA replication-related proteins. The MCM complex, consisting of this protein and MCM2, 4, and 7 proteins, possesses DNA helicase activity and may act as a DNA unwinding enzyme. The phosphorylation of the complex by CDC2 kinase reduces the helicase activity, suggesting a role in the regulation of DNA replication. Single-nucleotide polymorphisms in the intron regions of this gene are associated with differential transcriptional activation of the promoter of the neighboring lactase gene and, thereby, influence lactose intolerance in early adulthood.
CAST	Calpastatin	The protein encoded by this gene is an endogenous calpain (calcium-dependent cysteine protease) inhibitor. It consists of an N-terminal domain L and four repetitive calpain-inhibition domains (domains 1-4), and it is involved in the proteolysis of the amyloid precursor protein. The calpain/calpastatin system is involved in numerous membrane fusion events, such as neural vesicle exocytosis and platelet and red-cell aggregation. The encoded protein is also thought to affect the expression levels of genes encoding structural or regulatory proteins. Alternatively spliced transcript variants encoding different isoforms have been described.
TDRG1	Testis Development Related 1	Predicted to be located in the cytoplasm. TDRG1 (Testis Development Related 1) is an RNA Gene, and is affiliated with the lncRNA class. Diseases associated with TDRG1 include Microvascular Complications of Diabetes 5 and Cervical Cancer.
LINC00951	Long Intergenic Non-Protein Coding RNA 951	LINC00951 (Long Intergenic Non-Protein Coding RNA 951) is an RNA Gene, and is affiliated with the lncRNA class. Diseases associated with LINC00951 include Immunodeficiency, Common Variable, 8, With Autoimmunity.
LINGO2	Leucine Rich Repeat And Ig Domain Containing 2	Predicted to act upstream of or within the positive regulation of synapse assembly. Predicted to be located in the membrane. Predicted to be active in several cellular components, including the extracellular space, glutamatergic synapse, and synaptic membrane.

SF3A2	Splicing Factor 3a Subunit 2	This gene encodes subunit 2 of the splicing factor 3a protein complex. The splicing factor 3a heterotrimer includes subunits 1, 2, and 3 and is necessary for the <i>in vitro</i> conversion of 15S U2 snRNP into an active 17S particle that performs pre-mRNA splicing. Subunit 2 interacts with subunit 1 through its amino-terminus, while the single zinc finger domain of subunit 2 plays a role in its binding to the 15S U2 snRNP. Subunit 2 may also function independently of its RNA splicing function as a microtubule-binding protein.
DTX2P1	DTX2 Pseudogene 1	DTX2P1 (DTX2 Pseudogene 1) is a Pseudogene.
UPK3BP1	Uroplakin 3B Pseudogene 1	UPK3BP1 (Uroplakin 3B Pseudogene 1) is a Pseudogene.
PMS2P11	PMS1 Homolog 2, Mismatch Repair System Component Pseudogene 11	Predicted to enable kinase activity. PMS2P11 (PMS1 Homolog 2, Mismatch Repair System Component Pseudogene 11) is a Pseudogene.
LINC01122	Long Intergenic Non-Protein Coding RNA 1122	LINC01122 (Long Intergenic Non-Protein Coding RNA 1122) is an RNA Gene, and is affiliated with the lncRNA class.
PPARG	Peroxisome Proliferator-Activated Receptor Gamma	This gene encodes a member of the peroxisome proliferator-activated receptor (PPAR) subfamily of nuclear receptors. PPARs form heterodimers with retinoid X receptors (RXRs), and these heterodimers regulate transcription of various genes. Three subtypes of PPARs are known: PPAR-alpha, PPAR-delta, and PPAR-gamma. The protein encoded by this gene is PPAR-gamma and is a regulator of adipocyte differentiation. Additionally, PPAR-gamma has been implicated in the pathology of numerous diseases, including obesity, diabetes, atherosclerosis, and cancer. Alternatively spliced transcript variants that encode different isoforms have been described.
RARB	Retinoic Acid Receptor Beta	This gene encodes retinoic acid receptor beta, a member of the thyroid-steroid hormone receptor superfamily of nuclear transcriptional regulators. This receptor localizes to the cytoplasm and to subnuclear compartments. It binds retinoic acid, the biologically active form of vitamin A, which mediates cellular signaling in embryonic morphogenesis, cell growth, and differentiation. It is thought that this protein limits the growth of many cell types by regulating gene expression. The gene was first identified in a hepatocellular carcinoma, where it flanks a hepatitis B virus integration site. Alternate promoter usage and differential splicing result in multiple transcript variants.
BDNF	Brain-Derived Neurotrophic Factor	This gene encodes a member of the nerve growth factor family of proteins. Alternative splicing results in multiple transcript variants, at least one of which encodes a preproprotein that is proteolytically processed to generate the mature protein. The binding of this protein to its cognate receptor promotes neuronal survival in the adult brain. Expression of this gene is reduced in Alzheimer's, Parkinson's, and Huntington's disease patients. This gene may play a role in the regulation of the stress response and in the biology of mood disorders.
BDNF-AS	BDNF Antisense RNA	BDNF-AS (BDNF Antisense RNA) is an RNA Gene, and is affiliated with the lncRNA class. Diseases associated with BDNF-AS include Glioblastoma and Multiple Myeloma.
RAPSN	Receptor Associated Protein Of The Synapse	This gene encodes a member of a family of proteins that are receptor-associated proteins of the synapse. The encoded protein contains a conserved cAMP-dependent protein kinase phosphorylation site and plays a critical role in clustering and anchoring nicotinic acetylcholine receptors at synaptic sites by linking the receptors to the underlying postsynaptic cytoskeleton, possibly by direct association with actin or spectrin. Mutations in this gene may play a role in postsynaptic congenital myasthenic syndromes. Alternatively spliced transcript variants encoding multiple isoforms have been observed for this gene.
RPL35AP22	Ribosomal Protein L35a Pseudogene 22	RPL35AP22 (Ribosomal Protein L35a Pseudogene 22) is a Pseudogene.
ASTN2	Astrotactin 2	This gene encodes a protein that is expressed in the brain and may function in neuronal migration, based on functional studies of the related astrotactin 1 gene in human and mouse. A deletion at this locus has been associated with schizophrenia. Multiple transcript variants encoding different proteins have been found for this locus.
TMEM161B-DT	TMEM161B Divergent Transcript	Predicted to be a structural constituent of the ribosome. Predicted to be located in the ribosome. TMEM161B-DT (TMEM161B Divergent Transcript) is an RNA Gene, and is affiliated with the lncRNA class. Diseases associated with TMEM161B-DT include Glioma.
MIR9-2HG	MIR9-2 Host Gene	This is an evolutionarily conserved gene that produces alternatively spliced long non-coding RNAs that may be expressed predominantly in the brain and visual cortex. These transcripts may be involved in tumorigenesis, as depletion by siRNA suppressed glioma cell division. Transcripts may also bind to and regulate the activity of miR-411-5p and argonaute 2, thereby altering the expression of genes involved in tumor growth.
GLP1R	Glucagon-Like Peptide 1 Receptor	This gene encodes a 7-transmembrane protein that functions as a receptor for glucagon-like peptide 1 (GLP-1) hormone, which stimulates glucose-induced insulin secretion. This receptor, which functions at the cell surface, becomes internalized in response to GLP-1 and GLP-1 analogs, and it plays an important role in the signaling cascades leading to insulin secretion. It also displays neuroprotective effects in animal models. Polymorphisms in this gene are associated with diabetes. The protein is an important drug target for the treatment of type 2 diabetes and stroke. Alternative splicing of this gene results in multiple transcript variants.
DNAH8	Dynein Axonemal Heavy Chain 8	The protein encoded by this gene is a heavy chain of an axonemal dynein involved in sperm and respiratory cilia motility. Axonemal dyneins generate force through hydrolysis of ATP and binding to microtubules.
MIR1915HG	MIR1915 Host Gene	MIR1915HG (MIR1915 Host Gene) is an RNA gene and is affiliated with the lncRNA class.
SKIDA1	SKI/DACH Domain Containing 1	Predicted to enable DNA-binding transcription factor activity. Predicted to be involved in the regulation of transcription by RNA polymerase II.
AS3MT	Arsenite Methyltransferase	AS3MT catalyzes the transfer of a methyl group from S-adenosyl-L-methionine (AdoMet) to trivalent arsenical and may play a role in arsenic metabolism

BORCS7-ASMT	BORCS7-ASMT Readthrough (NMD Candidate)	This locus represents naturally occurring read-through transcription between the neighboring C10orf32 (chromosome 10 open reading frame 32) and AS3MT (arsenic, +3 oxidation state, methyltransferase) genes. The read-through transcript is a candidate for nonsense-mediated mRNA decay (NMD) and is therefore unlikely to produce a protein product.
SH2B1	SH2B Adaptor Protein 1	This gene encodes a member of the SH2-domain containing mediators family. The encoded protein mediates activation of various kinases and may function in cytokine and growth factor receptor signaling and cellular transformation. Alternatively spliced transcript variants have been described.
LINC02790	Long Intergenic Non-Protein Coding RNA 2790	LINC02790 (Long Intergenic Non-Protein Coding RNA 2790) is an RNA Gene, and is affiliated with the lncRNA class.
ZRANB3	Zinc Finger RANBP2-Type Containing 3	Enables ATP-dependent DNA/DNA annealing activity, DNA endonuclease activity, and K63-linked polyubiquitin modification-dependent protein binding activity. Involved in DNA repair, replication fork reversal, and response to UV. Located in the nuclear replication fork and nucleoplasm.

Supplemental Table 2: Gene lists from The CDC Public Health Genomics and Precision Health Knowledge Base, version 10.0 (PHGKB), and Phenome-wide association study Web (PheWeb), and Huang *et al.*^{38,56,57}

PHGKB	PheWeb	Huang <i>et al.</i>
TMEM18	FTO	PTEN
GNPDA2	MC4R	TXNIP
ETV5	DEFB136	TGFB2
FAIM2	TMEM18	RUNX1
MTCH2	SBK1	KIF24
NEGR1	DOCK11	PCED1A
SH2B1	SEMA3F	LIPC
MAF	SPANXB1	YTHDC1
NPC1	KCNJ15	ACTB
ENPP1	ZNF668	NUP93
BDNF	FAR1	KMT2A
KCTD15	ANXA1	MYH9
PRRC2A	PINX1	CUL1
SEC16B	MSRA	TMC04
PTER	PRKG1	FCRL6
	RGS18	MEGF8
	MCHR2	IKZF2
	SNF1	HLA-A
	EDIL3	KLF5
	KCNJ6	SLC27A5
	OLFAM4	DIAPH1
	C15orf48	TP53
	PPP3CA	ARAF
	TRAPPC9	AMER1
	FOXR2	PCBP1
	SEL1L	BRAF
	WNT7A	DYNC111
	CTSC	USP9X
	FGF1	CREB3
	ZNF722	ELF3
	XKR6	CTNNB1
	EPS8L2	ATP1B4
	BTBD3	DST
	PCDH11X	PTPN11
	LIPG	FAS
	BCL11B	ARID1A
	MRPS28	CD70
	LMO3	TMEM30A
	RSBN1	CIITA
	SMAD2	H1-2
	NDUFS4	DNAH10
	RGS7	HLA-B
	PCDH9	BIRC6
	GADL1	CDC42BPA
	PARK7	P2P
	GPC5	ERBB2
	MRLN	CUL3
	PSMG1	BAP1
	PLEKHM1	FAM20C
	ZNF716	ASPRV1
	ZNF236	TXNIP
	NALF1	KDM6A
	ZNF722	GTF2E1
	PTER	KRAS

	<i>FAM133A</i>	TRIM71
	<i>GGT1</i>	IDE
	<i>CFAP53</i>	PPT2
	<i>CD163</i>	CDKN1B
	<i>PABPC1L2B</i>	ARHGEF17
	<i>KLHL1</i>	ADAM12
	<i>LPP</i>	KCTD20
	<i>COX7C</i>	USP7
	<i>FAM220A</i>	MFSD3
	<i>MS4A10</i>	SIRPD
	<i>OLFM3</i>	KDM6A
	<i>ZFP37</i>	DNAH14
	<i>AIG1</i>	RHCG
	<i>ADGRG6</i>	ZFP36L2
		TCF7L2
		BCL9
		CLCN5
		SMAD4
		PCDHGB4
		PTEN
		SGMS1
		CDKN2A
		RIPOR3
		CALHM4
		IDH1
		TNFSF14
		CFHR4
		CYSLTR2

Abbreviations: PHGKB, Public Health Genomics and Precision Health Knowledge Base; PheWeb, Phenome-wide association study Web.