

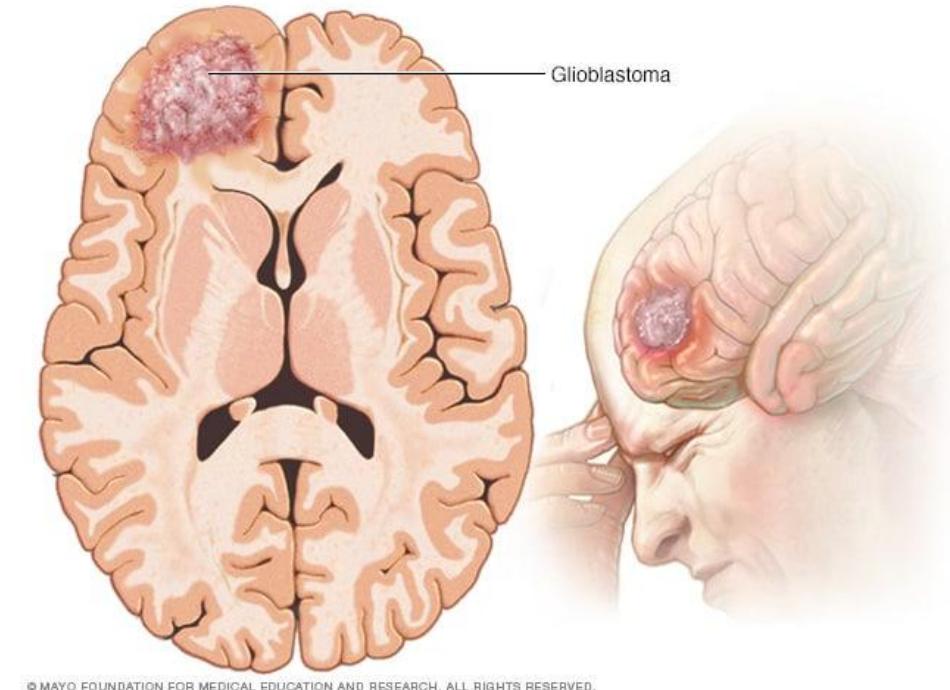
ARSI as a potential novel biomarker in Glioblastomas

Chyi Ricketts

CI Assignment

Glioblastoma

- Glioblastoma (GBM) is the most aggressive primary brain tumor with poor survival despite standard therapy
 - Type IV
 - Incidence rate around 3 per 100,000 people
 - 5-year survival rate is only 6.8%
- Molecular profiling defines key subtypes:
 - Classical
 - Mesenchymal
 - Proneural
 - Neural



Methods

Data Collection and
Preprocessing using
TCGA GBM Data

Univariate Analysis

Clinical Correlation

Multivariate
Analysis
by Subtype

Copy Number
Variation Analysis

Methylation Probe
Analysis

ECM and
Mesenchymal
Associated Genes
Panel Analysis

RPPA dataset

Univariate Analysis

Gene	HR	LCI	UCI	PVAL	FDR
NSUN5	2.188749	1.568035	3.055175	4.15E-06	0.083119
LRRC61	1.225964	1.119266	1.342834	1.16E-05	0.115874
NKX3-1	1.480686	1.224185	1.790933	5.25E-05	0.116802
C21orf96	2.505914	1.609502	3.901582	4.76E-05	0.116802
RPL39L	1.252385	1.125155	1.394002	3.83E-05	0.116802
OR2AG1	28.7064	5.881135	140.1188	3.32E-05	0.116802
STATH	151.4185	13.44155	1705.722	4.85E-05	0.116802
DNAJC30	2.378203	1.57078	3.600662	4.24E-05	0.116802
OSMR	1.327997	1.15855	1.522226	4.64E-05	0.116802
BMPR1A	0.506892	0.363727	0.706409	6.01E-05	0.120235
...
ARSI	1.16895	1.06169	1.28705	0.00148	0.176389

Clinical Analysis and Multivariate Analysis

Variable	coef	exp(coef)	se(coef)	z	Pr(> z)
age	0.025113	1.02543	0.007139	3.52E+00	4.36E-04

ARSI Correlation

Analysis	Type	HR	LCI	UCI	PVAL	FDR
Univariate	UNI	1.16895	1.06169	1.28705	0.00148	0.17639
Univariate Mesenchymal	UNI	1.39855	1.17428	1.66565	0.00017	0.41729
Multivariate	MULTI	1.13295	1.02452	1.25286	0.01502	0.54655
Multivariate Mesenchymal	MULTI	1.39674	1.16841	1.66969	0.00024	0.38676

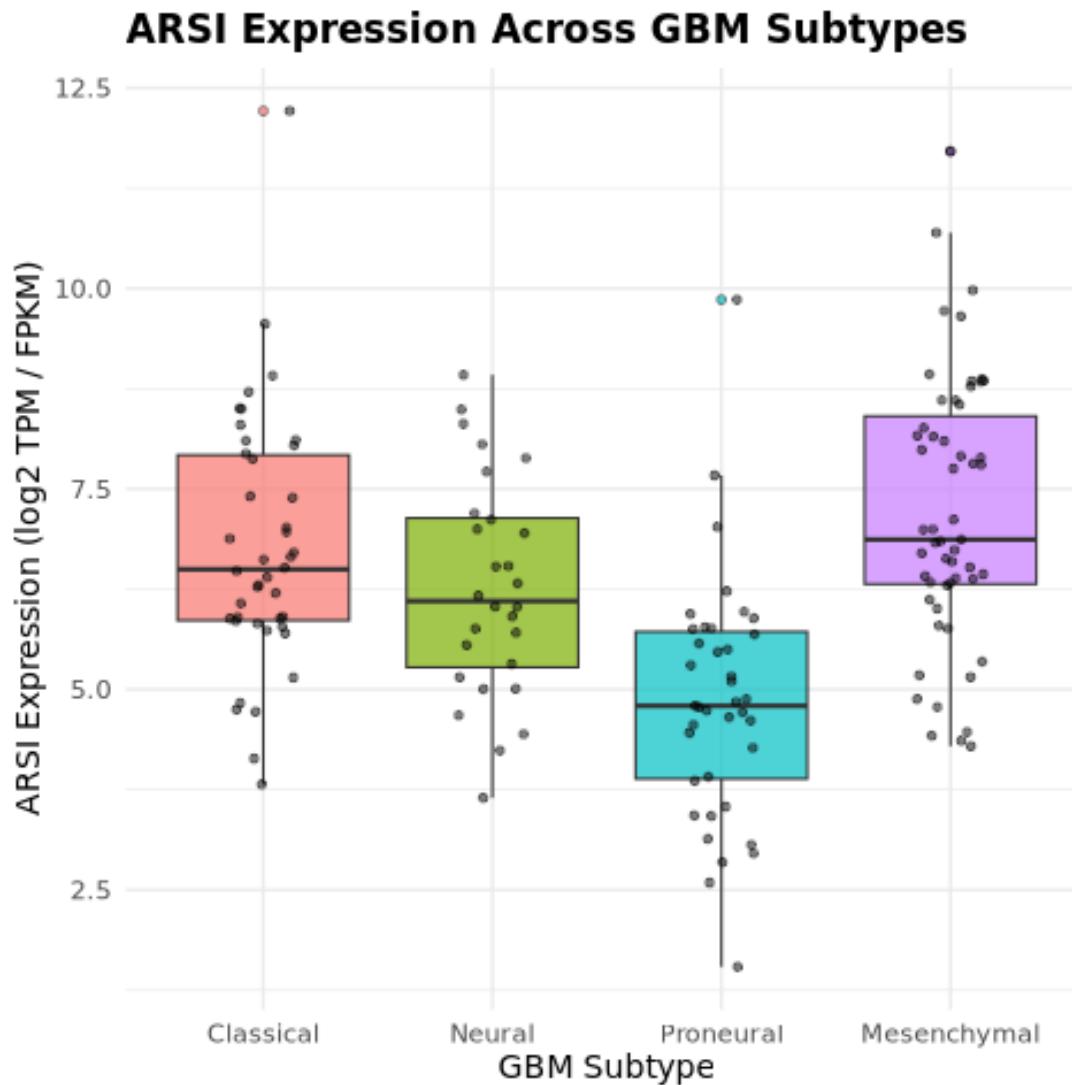
ARSI Expression

ARSI: Arylsulfatase I

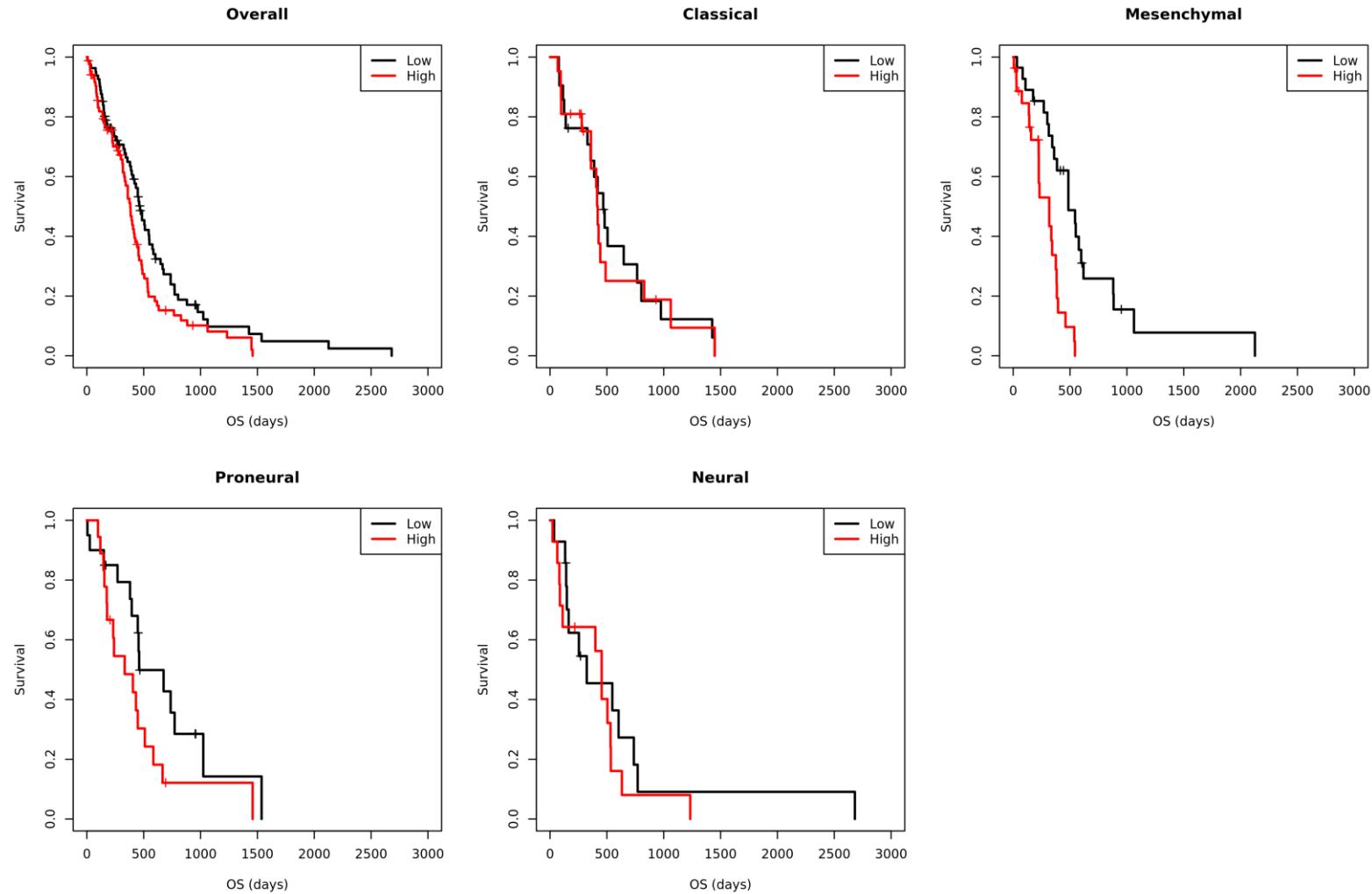
- Removes sulfate groups from ECM molecules

Low expression in most tissues

- ARSI Expression Across GBM Subtypes
- Highest Expression in Mesenchymal Subtype

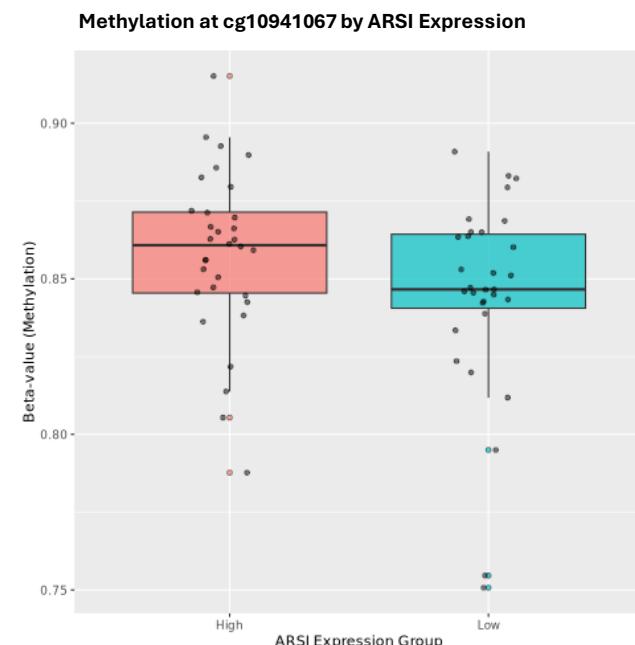
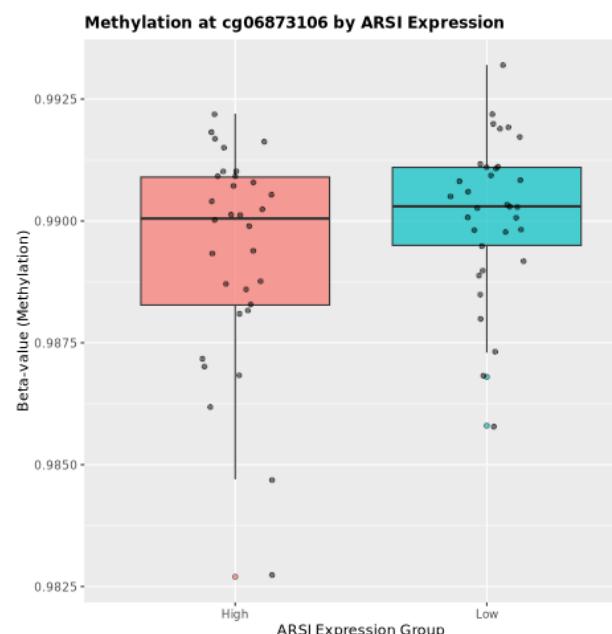
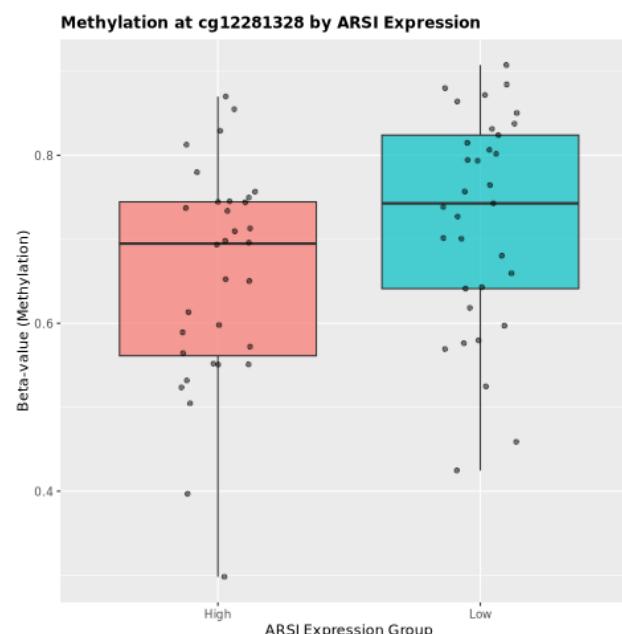


Kaplan-Meier Survival Curves: by Subtype



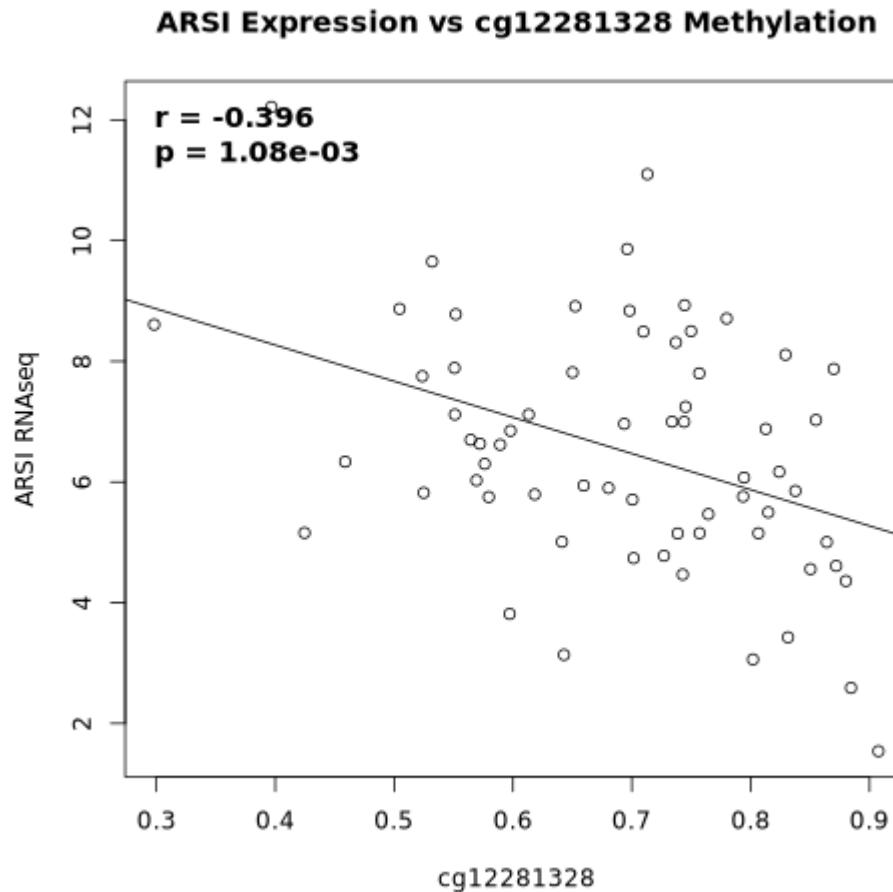
Methylation Analysis

CpG ID	Correlation	P-value	FDR	Slope	Mean ARSI High	Mean ARSI Low	Position	Island/Shore	UCSC_RefGene Group
cg12281328	-0.39625	0.001085	0.013017	-5.99321	0.656769	0.723245	1.5E+08	Island	1stExon
cg06873106	-0.29085	0.018749	0.112493	-302.786	0.989353	0.990145	1.5E+08	Island	Body
cg10941067	0.252179	0.042707	0.170829	6.634206	0.858031	0.820479	1.5E+08	Island	Body



*2 outliers removed for better visualization

Methylation Scatterplot of cg12281328



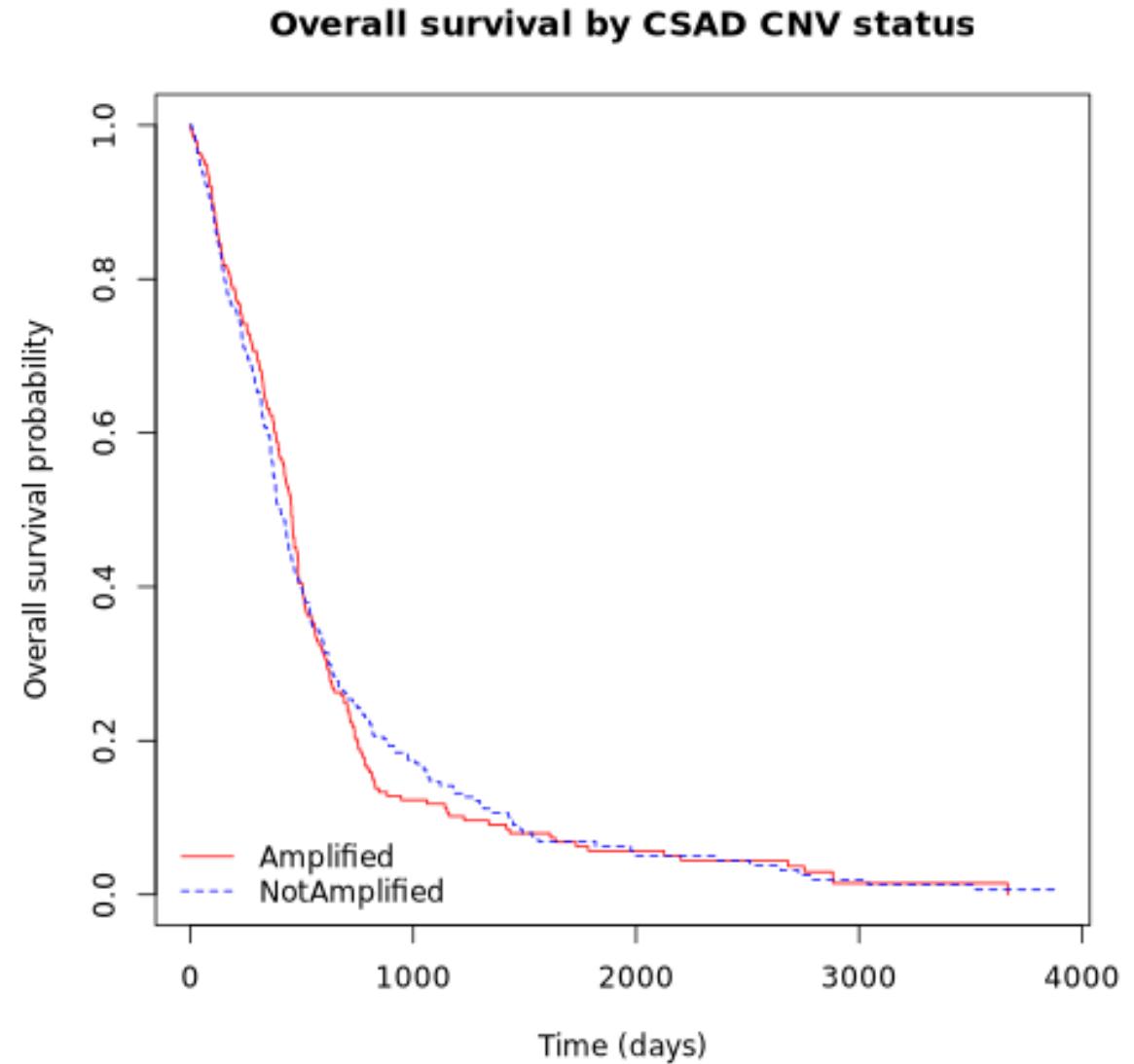
Several ARSI-linked CpG sites show hypomethylation

cg12281328 shows a strong hypomethylation signal

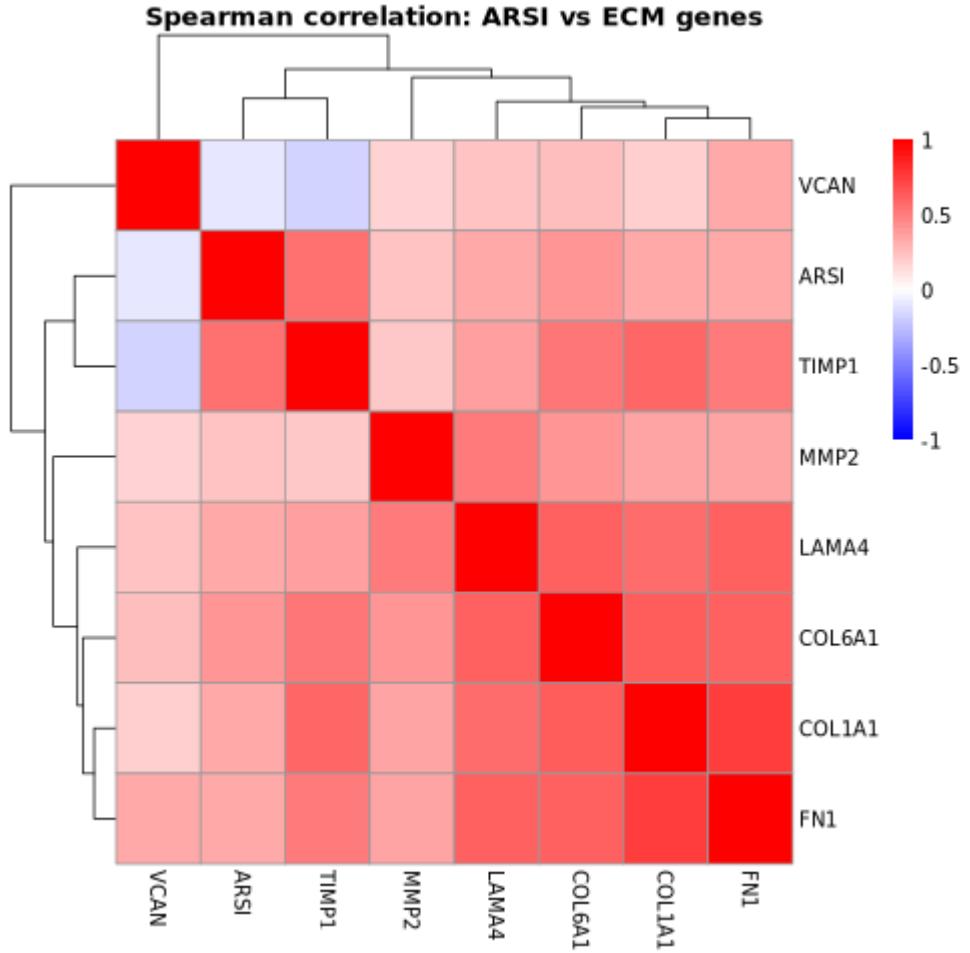
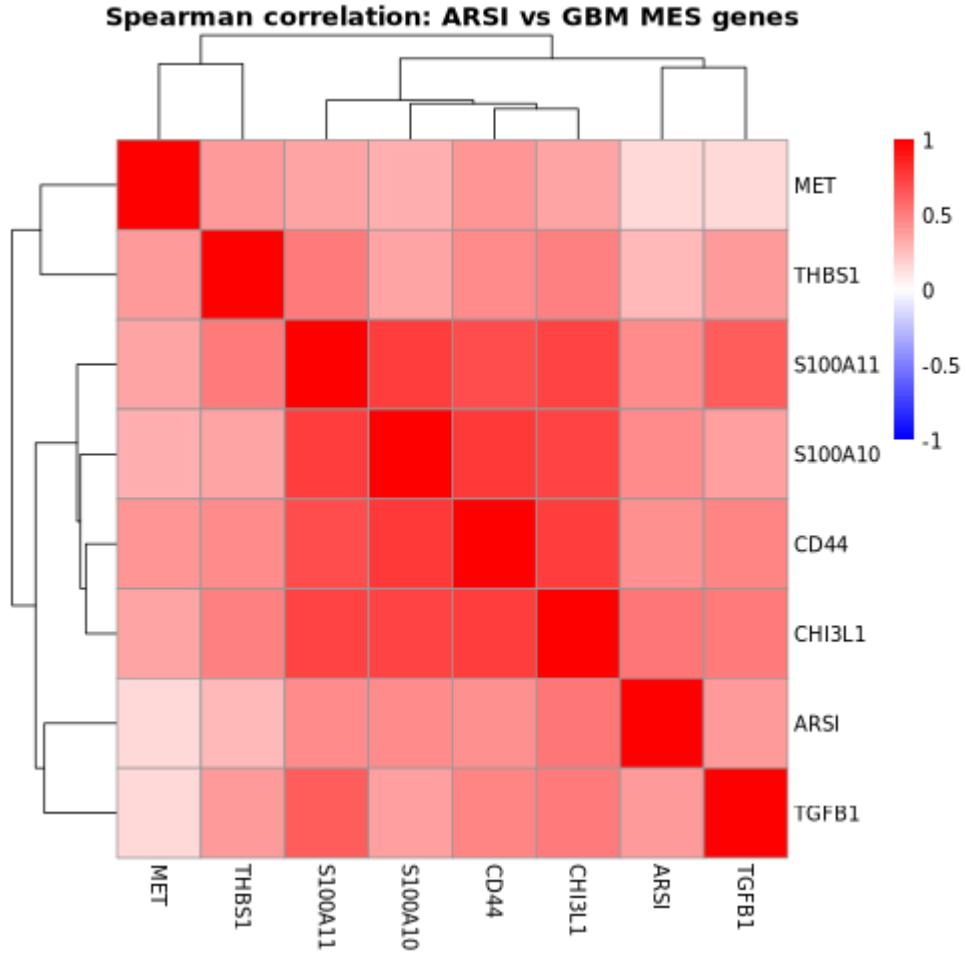
The pattern suggests that specific CpG sites drive ARSI epigenetic regulation in this tumor type

Overall, the data points to epigenetic upregulation of ARSI, which may contribute to tumor subtype differences and more aggressive tumor behavior

CNV Amplification on Survival Data



Panels for Mesenchymal and ECM Genes



RPPA

Gene	logFC	AveExpr	t	P.Value	adj.P.Val	B
ARSI	1	6.322222	584.8917	7.99E-139	1.64E-134	215.4937
ANXA2P2	0.447674	11.04537	8.753244	4.43E-13	2.81E-09	19.06341
PQLC3	0.31393	8.192181	8.726516	4.98E-13	2.81E-09	18.94683
TIMP1	0.537266	12.4349	8.705092	5.47E-13	2.81E-09	18.85337
ANXA2	0.446881	11.98458	8.616827	8.06E-13	3.31E-09	18.46832
EPHB1	-0.63172	8.727037	-8.09874	7.81E-12	2.67E-08	16.20953
FAM176B	0.308958	7.841022	8.021926	1.09E-11	3.20E-08	15.87518
ZNF711	-0.3563	9.385268	-7.93537	1.60E-11	4.10E-08	15.49868
ZC4H2	-0.30467	8.259059	-7.89805	1.88E-11	4.29E-08	15.33644
SPATS2L	0.203068	11.20657	7.83181	2.51E-11	4.69E-08	15.04867

Conclusion and Clinical Implications

- Methylation
 - Proposes mechanisms of upregulation through hypomethylation
- CNV
 - Little change in high and low CNV
- ECM
 - High correlation with ECM associated genes
- Mesenchymal
 - High correlation with other Mesenchymal associated genes
- RPPA

Conclusion:

- Candidate biomarkers can guide patient stratification
- Identification of a candidate biomarker for GBM: ARSI
 - Needs to be further explored as it does not hold statistical significance
- Support targeting therapeutic research:
 - Methylation
 - ECM Pathways

Limitations and Further Research

- Size of GBM Dataset
 - Limited Statistical power
 - Wide Confidence Intervals
 - High FDR → results not statistically significant after correction
- Observational nature: Functional mechanisms need to be further explored
 - Methylation
 - ECM Matrix

Future Directions:

- Validation on an external dataset
 - CGGA: Chinese Glioma Genome Atlas
- Function/Wet-lab studies: CRISPR, knockdown/overexpression experiments
- Multi-omics integration: Combine methylation, CNV, and expression data to identify actionable pathways and refine patient stratification

References

- Barbosa LC, Machado GC, Heringer M, Ferrer VP. Identification of established and novel extracellular matrix components in glioblastoma as targets for angiogenesis and prognosis. *Neurogenetics*. 2024 Jul;25(3):249-262. doi: 10.1007/s10048-024-00763-x. Epub 2024 May 22. PMID: 38775886.
- Brown D. V., Daniel P. M., D'Abaco G. M., Gogos A., Ng W., Morokoff A. P., Mantamadiotis T. Coexpression analysis of CD133 and CD44 identifies Proneural and Mesenchymal subtypes of glioblastoma multiforme. *Oncotarget*. 2015; 6: 6267-6280. Retrieved from <https://www.oncotarget.com/article/3365/text/>
- Behnan J, Finocchiaro G, Hanna G. The landscape of the mesenchymal signature in brain tumours. *Brain*. 2019 Apr 1;142(4):847-866. doi: 10.1093/brain/awz044. PMID: 30946477; PMCID: PMC6485274.
- Cescon M, Rampazzo E, Bresolin S, Da Ros F, Manfreda L, Cani A, Della Puppa A, Braghetta P, Bonaldo P, Persano L. Collagen VI sustains cell stemness and chemotherapy resistance in glioblastoma. *Cell Mol Life Sci*. 2023 Jul 28;80(8):233. doi: 10.1007/s00018-023-04887-5. PMID: 37505240; PMCID: PMC10382393.
- Chen X, Chen X, Xie W, Ge H, He H, Zhang A, Zheng J. BRAF-activated ARSI suppressed EREG-mediated ferroptosis to promote BRAF^{V600E} (mutant) papillary thyroid carcinoma progression and sorafenib resistance. *Int J Biol Sci*. 2025 Jan 1;21(1):128-142. doi: 10.7150/ijbs.99423. PMID: 39744439; PMCID: PMC11667812.
- Shen Y, Wei Z, Zhou C, Song J, Wang J, Wang J, Wu L, Fang S, Shen Z. Arylsulfatase I is a prognostic biomarker for head and neck squamous cell carcinoma and Pan-cancer. *J Clin Lab Anal*. 2022 Sep;36(9):e24600. doi: 10.1002/jcl.a.24600. Epub 2022 Jul 23. PMID: 35870182; PMCID: PMC9459304
- Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O'Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, Hodgson JG, James CD, Sarkaria JN, Brennan C, Kahn A, Spellman PT, Wilson RK, Speed TP, Gray JW, Meyerson M, Getz G, Perou CM, Hayes DN; Cancer Genome Atlas Research Network. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*. 2010 Jan 19;17(1):98-110. doi: 10.1016/j.ccr.2009.12.020. PMID: 20129251; PMCID: PMC2818769.

Supplementary: Datasets used

- TCGA Glioblastoma (GBM) (25 datasets)
- IlluminaHiSeq (n=172) TCGA Hub
- Curated survival data (n=602) TCGA Hub
- Phenotypes (n=629) TCGA Hub
- Methylation450k (n=155) TCGA Hub
- RPPA (n=244) TCGA Hub

Supplementary: Univariate Analysis

Gene	HR	LCI	UCI	PVAL	FDR
NSUN5	2.188749	1.568035	3.055175	4.15E-06	0.083119
LRRC61	1.225964	1.119266	1.342834	1.16E-05	0.115874
NKX3-1	1.480686	1.224185	1.790933	5.25E-05	0.116802
C21orf96	2.505914	1.609502	3.901582	4.76E-05	0.116802
RPL39L	1.252385	1.125155	1.394002	3.83E-05	0.116802
OR2AG1	28.7064	5.881135	140.1188	3.32E-05	0.116802
STATH	151.4185	13.44155	1705.722	4.85E-05	0.116802
DNAJC30	2.378203	1.57078	3.600662	4.24E-05	0.116802
OSMR	1.327997	1.15855	1.522226	4.64E-05	0.116802
BMPR1A	0.506892	0.363727	0.706409	6.01E-05	0.120235

Supplementary: Multivariate Analysis

Gene	HR	LCI	UCI	PVAL	FDR
SOAT2	1.618953	1.293461	2.026353	2.59E-05	0.261049
STATH	176.8973	15.85178	1974.079	2.61E-05	0.261049
CDC73	0.36683	0.219892	0.611959	1.23E-04	0.275813
GUCA1C	3.176344	1.760364	5.731295	1.24E-04	0.275813
PTPRN	1.211702	1.100345	1.33433	9.46E-05	0.275813
C21orf96	2.442283	1.556734	3.831576	1.02E-04	0.275813
CHRM5	1.364823	1.169868	1.592265	7.65E-05	0.275813
PTPRN2	1.615105	1.274721	2.04638	7.19E-05	0.275813
OR2AG1	24.45167	5.081436	117.6605	6.67E-05	0.275813
NRXN3	1.23055	1.102854	1.373031	2.06E-04	0.289626

Supplementary: ARSI Expression

Analysis	Type	HR	LCI	UCI	PVAL	FDR
Overall	Univariate	1.16895	1.06169	1.28705	0.00148	0.17639
Mesenchymal	Univariate	1.39855	1.17428	1.66565	0.00017	0.41729
Classical	Univariate	1.24562	0.92447	1.67835	0.14882	0.98821
Neural	Univariate	1.17179	0.88696	1.54811	0.26454	0.99996
Proneural	Univariate	1.03408	0.8546	1.25126	0.73044	0.93231
Overall	Multivariate	1.13295	1.02452	1.25286	0.01502	0.54655
Mesenchymal	Multivariate	1.39674	1.16841	1.66969	0.00024	0.38676
Classical	Multivariate	1.24898	0.92429	1.68772	0.14778	0.95531
Neural	Multivariate	1.05729	0.77468	1.44301	0.72553	0.99999
Proneural	Multivariate	0.92918	0.7483	1.15379	0.50609	0.9258

Supplementary: Methylation Probes Data

CpG ID	Correlation	P-value	FDR	Slope	Mean ARSI High	Mean ARSI Low	Position	Island/Shore	UCSC_RefGene_Group
cg12281328	-0.39625	0.001085	0.013017	-5.99321	0.656769	0.723245	1.5E+08	Island	1stExon
cg06873106	-0.29085	0.018749	0.112493	-302.786	0.989353	0.990145	1.5E+08	Island	Body
cg10941067	0.252179	0.042707	0.170829	6.634206	0.858031	0.820479	1.5E+08	Island	Body
cg03405473	0.221516	0.076168	0.228503	2.855841	0.6232	0.527309	1.5E+08	North Shore	Body
cg25652824	-0.16292	0.194735	0.356951	-91.1673	0.978459	0.979027	1.5E+08	North Shore	Body
cg14105467	-0.15818	0.208222	0.356951	-5.65911	0.790097	0.789018	1.5E+08	North Shore	3'UTR
cg16633129	0.166043	0.186199	0.356951	6.257584	0.852338	0.836482	1.5E+08	Island	Body
cg20873205	0.12922	0.304939	0.406585	4.98114	0.915775	0.8966	1.5E+08	Island	Body
cg26530201	-0.13441	0.28576	0.406585	-7.42263	0.078653	0.084233	1.5E+08	South Shore	TSS200
cg26292378	0.09517	0.450781	0.540937	3.957983	0.083622	0.081027	1.5E+08	South Shore	TSS200
cg02613713	0.074428	0.555706	0.555706	6.042511	0.972075	0.963288	1.5E+08	Island	Body
cg10607557	-0.07921	0.530509	0.555706	-1.24661	0.869347	0.841918	1.5E+08	North Shore	Body