

Supplementary Materials

DeepSGP: Deep Learning for Gene Selection and Survival Group Prediction in Glioblastoma

Ritaban Kirtania, Subhashis Banerjee, Sayantan Laha,
B. Uma Shankar, Raghunath Chatterjee and Sushmita Mitra

S1. Subset of gene IDs selected before application of the seven filters

Table S1: Subset of gene IDs selected before application of the seven filters.

ENSG00000127993	ENSG00000110203	ENSG00000140263	ENSG00000169429	ENSG00000204241	ENSG00000105650
ENSG00000149380	ENSG00000248528	ENSG00000142748	ENSG00000169436	ENSG00000204962	ENSG00000186940
ENSG00000159167	ENSG00000111321	ENSG00000143632	ENSG00000169894	ENSG00000204963	ENSG00000213139
ENSG00000101825	ENSG00000231394	ENSG00000174567	ENSG00000170160	ENSG00000205358	ENSG00000198719
ENSG00000234707	ENSG00000124721	ENSG00000144057	ENSG00000231427	ENSG00000206129	ENSG00000234336
ENSG00000007312	ENSG00000112759	ENSG00000204381	ENSG00000170837	ENSG00000211892	ENSG00000106537
ENSG00000077274	ENSG00000112874	ENSG00000144810	ENSG00000170955	ENSG00000211895	ENSG00000106683
ENSG00000197467	ENSG00000113070	ENSG00000145014	ENSG00000215644	ENSG00000224650	ENSG00000107438
ENSG00000135407	ENSG00000113248	ENSG00000145283	ENSG00000171681	ENSG00000214110	ENSG00000107984
ENSG00000261409	ENSG00000113303	ENSG00000145536	ENSG00000171819	ENSG00000214193	ENSG00000197084
ENSG00000029534	ENSG00000113389	ENSG00000145545	ENSG00000172460	ENSG00000215474	ENSG00000136720
ENSG00000029559	ENSG00000113532	ENSG00000145623	ENSG00000261625	ENSG00000223475	ENSG00000136944
ENSG00000034152	ENSG00000113600	ENSG00000145777	ENSG00000173702	ENSG00000224057	ENSG00000137507
ENSG00000205212	ENSG00000114993	ENSG00000146592	ENSG00000173805	ENSG00000224596	ENSG00000137673
ENSG00000124882	ENSG00000115239	ENSG00000146729	ENSG00000174226	ENSG00000225698	ENSG00000137731
ENSG00000048545	ENSG00000115604	ENSG00000147206	ENSG00000174403	ENSG00000225886	ENSG00000138316
ENSG00000049656	ENSG00000116745	ENSG00000148942	ENSG00000174799	ENSG00000225940	ENSG00000272369
ENSG00000213214	ENSG00000117152	ENSG00000149656	ENSG00000229762	ENSG00000226833	ENSG00000139269
ENSG00000054356	ENSG00000117625	ENSG00000149968	ENSG00000175426	ENSG00000226939	ENSG00000198707
ENSG00000180347	ENSG00000118785	ENSG00000150275	ENSG00000175782	ENSG00000227165	ENSG00000140262
ENSG00000061337	ENSG00000118849	ENSG00000151468	ENSG00000176124	ENSG00000271287	ENSG00000165511
ENSG00000064218	ENSG00000120696	ENSG00000152592	ENSG00000176771	ENSG00000228061	ENSG00000165730
ENSG00000070495	ENSG00000123892	ENSG00000153561	ENSG00000176907	ENSG00000228522	ENSG00000166348
ENSG00000071282	ENSG00000123989	ENSG00000153976	ENSG00000227496	ENSG00000230006	ENSG00000166415
ENSG00000174206	ENSG00000173918	ENSG00000154080	ENSG00000177565	ENSG00000231419	ENSG00000167123
ENSG00000123342	ENSG00000124491	ENSG00000154127	ENSG00000178297	ENSG00000232006	ENSG00000167779
ENSG00000161610	ENSG00000124568	ENSG00000155719	ENSG00000178718	ENSG00000232788	ENSG00000168077
ENSG00000078142	ENSG00000125538	ENSG00000156463	ENSG00000178752	ENSG00000233093	ENSG00000168118
ENSG00000080166	ENSG00000125735	ENSG00000157150	ENSG00000179008	ENSG00000233125	ENSG00000168283
ENSG00000082196	ENSG00000125740	ENSG00000158477	ENSG00000179168	ENSG00000233818	ENSG00000242950
ENSG00000083782	ENSG00000233791	ENSG00000158485	ENSG00000179454	ENSG00000234513	ENSG00000241825
ENSG00000086544	ENSG00000126838	ENSG00000158578	ENSG00000180458	ENSG00000235109	ENSG00000197588
ENSG00000086967	ENSG00000166225	ENSG00000159674	ENSG00000181885	ENSG00000236449	ENSG00000198400
ENSG00000088340	ENSG00000128045	ENSG00000160180	ENSG00000237164	ENSG00000237125	ENSG00000198491
ENSG00000088386	ENSG00000128262	ENSG00000160224	ENSG00000183439	ENSG00000237945	ENSG00000198650
ENSG00000137959	ENSG00000128285	ENSG00000161013	ENSG00000183760	ENSG00000239440	ENSG00000198795
ENSG00000144136	ENSG00000129354	ENSG00000162944	ENSG00000183798	ENSG00000248187	ENSG00000223891
ENSG00000170421	ENSG00000130829	ENSG00000163209	ENSG00000184139	ENSG00000249755	ENSG00000203943
ENSG00000095752	ENSG00000131015	ENSG00000185633	ENSG00000229544	ENSG00000250295	ENSG00000203965
ENSG00000099869	ENSG00000131142	ENSG00000163497	ENSG00000185215	ENSG00000250815	ENSG00000203995

ENSG00000100181	ENSG00000131355	ENSG00000163661	ENSG00000185630	ENSG00000251533	ENSG00000264345
ENSG00000163735	ENSG00000131389	ENSG00000163687	ENSG00000187140	ENSG00000253831	ENSG00000265203
ENSG00000164509	ENSG00000132749	ENSG00000163736	ENSG00000187398	ENSG00000253953	ENSG00000268088
ENSG00000100985	ENSG00000132906	ENSG00000163909	ENSG00000187527	ENSG00000254919	ENSG00000270069
ENSG00000101017	ENSG00000133048	ENSG00000163975	ENSG00000187676	ENSG00000256128	ENSG00000274426
ENSG00000198768	ENSG00000237212	ENSG00000164047	ENSG00000187889	ENSG00000256940	ENSG00000274735
ENSG00000101888	ENSG00000133063	ENSG00000164530	ENSG00000188107	ENSG00000257838	ENSG00000276170
ENSG00000101892	ENSG00000254545	ENSG00000239839	ENSG00000188522	ENSG00000258227	ENSG00000280890
ENSG00000101972	ENSG00000139767	ENSG00000164877	ENSG00000188620	ENSG00000258376	ENSG00000281162
ENSG00000102221	ENSG00000135925	ENSG00000165078	ENSG00000189067	ENSG00000258484	ENSG00000281183
ENSG00000102970	ENSG00000136014	ENSG00000165272	ENSG00000189127	ENSG00000259616	ENSG00000103044
ENSG00000136160	ENSG00000165322	ENSG00000196890	ENSG00000279881	ENSG00000275630	ENSG00000136352
ENSG00000165474	ENSG00000197057	ENSG00000261801			

S2. Subset of gene IDs selected after application of the seven filters

S2.1. CIFE

ENSG00000145014, ENSG00000145623, ENSG00000117625, ENSG00000140263, ENSG00000214110, ENSG00000106537, ENSG00000196890, ENSG00000112874, ENSG00000114993, ENSG00000189067, ENSG00000101888, ENSG00000198795, ENSG00000232788, ENSG00000171681, ENSG00000179454, ENSG0000048545, ENSG00000152778, ENSG00000203943, ENSG00000113248, ENSG00000120696, ENSG00000168077, ENSG00000168283, ENSG00000153561, ENSG00000214193, ENSG00000167123, ENSG00000177565, ENSG00000163909, ENSG00000136160, ENSG00000137731, ENSG00000148942, ENSG00000225886, ENSG00000115239, ENSG00000176771, ENSG00000173805, ENSG00000107438, ENSG00000226833, ENSG00000258227, ENSG0000049656, ENSG00000136720, ENSG00000137959, ENSG00000175782, ENSG00000204241, ENSG00000178718, ENSG00000235109, ENSG00000223891, ENSG00000281162, ENSG00000131015, ENSG00000183615, ENSG00000101017, ENSG00000146592.

S2.2. DISR

ENSG00000145014, ENSG00000145623, ENSG00000125740, ENSG00000181885, ENSG00000061337, ENSG00000163661, ENSG00000167779, ENSG00000204962, ENSG00000261801, ENSG00000130829, ENSG00000163735, ENSG00000144136, ENSG00000239440, ENSG00000254919, ENSG00000242950, ENSG00000237125, ENSG00000169436, ENSG00000234354, ENSG00000183798, ENSG00000254545, ENSG00000118113, ENSG00000086544, ENSG00000100181, ENSG00000137959, ENSG00000250295, ENSG00000113303, ENSG00000078142, ENSG00000166348, ENSG00000136014, ENSG00000230006, ENSG00000157150, ENSG00000137731, ENSG00000107984, ENSG00000113532, ENSG00000117625, ENSG00000124882, ENSG00000179008, ENSG00000187527, ENSG00000197467, ENSG00000132906, ENSG00000180458, ENSG00000168077, ENSG00000088386, ENSG00000103855, ENSG00000150275, ENSG00000224650, ENSG00000160180, ENSG00000162458, ENSG00000179168, ENSG00000145423.

S2.3. ICAP

ENSG00000145014, ENSG00000145623, ENSG00000117625, ENSG00000078142, ENSG00000101972, ENSG00000137731, ENSG00000185630, ENSG00000214110, ENSG00000196890, ENSG00000189067, ENSG00000106537, ENSG00000173805, ENSG00000113248, ENSG00000146592, ENSG00000198795, ENSG00000226833, ENSG00000140263, ENSG00000176124, ENSG00000171681, ENSG00000136720, ENSG00000144810, ENSG00000165730, ENSG00000281183, ENSG00000163909, ENSG00000214193, ENSG00000205358, ENSG00000148942, ENSG00000178718, ENSG00000155719, ENSG00000183615, ENSG00000177565, ENSG00000172460, ENSG00000164877, ENSG00000153976, ENSG00000235109, ENSG00000143845, ENSG00000168118, ENSG00000131015, ENSG00000197057, ENSG00000114993, ENSG00000152778, ENSG00000170837, ENSG00000168283, ENSG00000179454, ENSG00000196850, ENSG00000146729, ENSG00000187676, ENSG00000007312, ENSG00000105650, ENSG00000166415.

S2.4. JMI

ENSG00000145014, ENSG00000145623, ENSG00000117625, ENSG00000078142, ENSG00000137731, ENSG00000136160, ENSG00000181885, ENSG00000179168, ENSG00000168077, ENSG00000185630, ENSG00000153561, ENSG00000198795, ENSG00000261801, ENSG00000233818, ENSG00000176124, ENSG00000101972, ENSG00000136720, ENSG00000070495, ENSG00000249755, ENSG00000189067, ENSG00000214110, ENSG00000162944, ENSG00000103855, ENSG00000101017, ENSG00000086544, ENSG00000254919, ENSG00000106537, ENSG00000120696, ENSG00000112874, ENSG00000146592, ENSG00000166348, ENSG00000111321, ENSG00000102221, ENSG00000137959, ENSG00000167779, ENSG00000049656, ENSG00000173805, ENSG00000196890, ENSG00000178718, ENSG00000165322, ENSG00000204241, ENSG00000034152, ENSG00000144136, ENSG00000183798, ENSG00000164877, ENSG00000163661, ENSG00000136014, ENSG00000061337, ENSG00000151468, ENSG00000168283.

S2.5. MIFS

ENSG00000145014, ENSG00000181885, ENSG00000130829, ENSG00000204962, ENSG00000178752, ENSG00000113600, ENSG00000124568, ENSG00000166225, ENSG00000205212, ENSG00000171819, ENSG00000129354, ENSG00000135407, ENSG00000083782, ENSG00000198719,

ENSG00000234336, ENSG00000124721, ENSG00000198707, ENSG00000198491, ENSG00000253953, ENSG00000280890, ENSG00000102970, ENSG00000118849, ENSG00000139767, ENSG00000231427, ENSG00000198400, ENSG00000127993, ENSG00000128045, ENSG00000257910, ENSG00000161610, ENSG00000231394, ENSG00000237212, ENSG00000099869, ENSG00000101892, ENSG00000224057, ENSG00000082196, ENSG00000112759, ENSG00000234707, ENSG00000077274, ENSG00000274426, ENSG00000173702, ENSG00000215644, ENSG00000229762, ENSG00000187889, ENSG00000197084, ENSG00000229544, ENSG00000236449, ENSG00000188107, ENSG00000272369, ENSG00000118113, ENSG00000164509.

S2.6. MIM

ENSG00000145014, ENSG00000145623, ENSG00000181885, ENSG00000117625, ENSG00000179168, ENSG00000078142, ENSG00000261801, ENSG00000185630, ENSG00000086544, ENSG00000144136, ENSG00000167779, ENSG00000103855, ENSG00000137731, ENSG00000168077, ENSG00000249755, ENSG00000233818, ENSG00000101972, ENSG00000061337, ENSG00000163661, ENSG00000254919, ENSG00000070495, ENSG00000176124, ENSG00000111321, ENSG00000136160, ENSG00000153561, ENSG00000198795, ENSG00000254545, ENSG00000163735, ENSG00000169436, ENSG00000162944, ENSG00000166348, ENSG00000101017, ENSG00000136720, ENSG00000120696, ENSG00000102221, ENSG00000183798, ENSG00000137959, ENSG00000174226, ENSG00000112874, ENSG00000189067, ENSG00000049656, ENSG00000034152, ENSG00000151468, ENSG00000130829, ENSG00000107984, ENSG00000113303, ENSG00000161013, ENSG00000183439, ENSG00000157150, ENSG00000164877.

S2.7. MRMR

ENSG00000145014, ENSG00000181885, ENSG00000078142, ENSG00000163661, ENSG00000130829, ENSG00000145623, ENSG00000183798, ENSG00000117625, ENSG00000137731, ENSG00000239440, ENSG00000261801, ENSG00000254919, ENSG00000237125, ENSG00000204962, ENSG00000101972, ENSG00000179168, ENSG00000168077, ENSG00000157150, ENSG00000167779, ENSG00000242950, ENSG00000250295, ENSG00000111321, ENSG00000061337, ENSG00000137959, ENSG00000163735, ENSG00000136014, ENSG00000249755, ENSG00000234354, ENSG00000113303, ENSG00000166348, ENSG00000174226, ENSG00000169436, ENSG00000254545, ENSG00000070495, ENSG00000086544, ENSG00000113248, ENSG00000224650, ENSG00000154080, ENSG00000176124, ENSG00000179008, ENSG00000230006, ENSG00000144136, ENSG00000153561, ENSG00000125740, ENSG00000183439, ENSG00000187527, ENSG00000118113, ENSG00000268088, ENSG00000233818, ENSG00000136160.

S2.8. Overlapping genes IDs between different subsets selected by filters

ENSG00000145014, ENSG00000117625, ENSG00000137731, ENSG00000145623, ENSG00000078142, ENSG00000137959, ENSG00000168077, ENSG00000181885, ENSG00000061337, ENSG00000086544, ENSG00000101972, ENSG00000130829, ENSG00000136160, ENSG00000136720, ENSG00000144136, ENSG00000153561, ENSG00000163661, ENSG00000166348, ENSG00000167779, ENSG00000176124, ENSG00000179168, ENSG00000183798, ENSG00000189067, ENSG00000198795, ENSG00000254919, ENSG00000261801, ENSG00000049656, ENSG00000070495, ENSG00000101017, ENSG00000103855, ENSG00000106537, ENSG00000111321, ENSG00000112874, ENSG00000113248, ENSG00000113303, ENSG00000118113, ENSG00000120696, ENSG00000136014, ENSG00000146592, ENSG00000157150, ENSG00000163735, ENSG00000164877, ENSG00000168283, ENSG00000169436, ENSG00000173805, ENSG00000178718, ENSG00000185630, ENSG00000196890, ENSG00000204962, ENSG00000214110.

S3. The IDs of the subset of shortlisted 100 genes for pathway analysis

Table S2: The IDs of the subset of shortlisted 100 genes for pathway analysis.

ENSG00000034152	ENSG00000187676	ENSG00000113070	ENSG00000234513	ENSG00000138316	ENSG00000204241
ENSG00000137673	ENSG00000124568	ENSG00000169436	ENSG00000139269	ENSG00000224057	ENSG00000159674
ENSG00000133048	ENSG00000117625	ENSG00000145283	ENSG00000235109	ENSG00000171681	ENSG00000261625
ENSG00000261801	ENSG00000196890	ENSG00000225886	ENSG00000128285	ENSG00000166415	ENSG00000256128
ENSG00000214110	ENSG00000179168	ENSG00000101888	ENSG00000161610	ENSG00000083782	ENSG00000198795
ENSG00000188522	ENSG00000173702	ENSG00000113248	ENSG00000149968	ENSG00000100985	ENSG00000165322
ENSG00000187140	ENSG00000204963	ENSG00000137959	ENSG00000174403	ENSG00000164047	ENSG00000169429
ENSG00000233818	ENSG00000145014	ENSG00000281183	ENSG00000265203	ENSG00000254545	ENSG00000123989
ENSG00000155719	ENSG00000106537	ENSG00000164877	ENSG00000168283	ENSG00000163661	ENSG00000144057
ENSG00000125735	ENSG00000170160	ENSG00000102970	ENSG00000173918	ENSG00000223475	ENSG00000158477
ENSG00000101972	ENSG00000148942	ENSG00000205358	ENSG00000180347	ENSG00000178752	ENSG00000077274
ENSG00000233791	ENSG00000163687	ENSG00000168118	ENSG00000258227	ENSG00000198491	ENSG00000181885
ENSG00000228522	ENSG00000129354	ENSG00000236449	ENSG00000197588	ENSG00000237945	ENSG00000226833
ENSG00000162944	ENSG00000161013	ENSG00000174226	ENSG00000213214	ENSG00000049656	ENSG00000125740
ENSG00000071282	ENSG00000231419	ENSG00000061337	ENSG00000183760	ENSG00000160180	ENSG00000198768
ENSG00000188620	ENSG00000237125	ENSG00000186940	ENSG00000170837	ENSG00000178718	ENSG00000158485
ENSG00000213139	ENSG00000165078	ENSG00000088340	ENSG00000261409		

S4. List of pathways identified

Table S3: List of pathways identified.

GOID	KEGG:04657	KEGG:05219	R-HSA:3906995	R-HSA:5083635	R-HSA:5173214
GO Term	IL-17 signaling pathway	Bladder cancer	Diseases associated with O-glycosylation of proteins	Defective B3GALTL causes Peters-plus syndrome (PpS)	O-glycosylation of TSR domain-containing proteins
Ontology Source	KEGG	KEGG	REACTOME	REACTOME	REACTOME
Term PValue	0.000039	0.000619	0.000168	0.000456	0.000494
Term PValue Corrected with Benjamini-Hochberg	0.000196	0.000773	0.000558	0.000760	0.000705
Group PValue	0.000039	0.000619	0.000177	0.000177	0.000177
Group PValue Corrected with Benjamini-Hochberg	0.000157	0.000619	0.000236	0.000236	0.000236
GO Levels	[-1]	[-1]	[-1]	[-1]	[-1]
GO Groups	Group0	Group1	Group2	Group2	Group2
% Associated Genes	5.32	7.32	5.88	8.11	7.89
Nr. Genes	5	3	4	3	3
Associated Genes Found	[CCL17, CXCL8, FOSB, MMP3, MMP9]	[CXCL8, HBEGF, MMP9]	[ADAMTS14, B3GLCT, MUC13, SPON2]	[ADAMTS14, B3GLCT, SPON2]	[ADAMTS14, B3GLCT, SPON2]

Table S4: List of pathways identified.

GOID	R-HSA:1442490	R-HSA:1474290	R-HSA:1592389	R-HSA:2022090	R-HSA:9009391
GO Term	Collagen degradation	Collagen formation	Activation of Matrix Metalloproteinases	Assembly of collagen fibrils and other multimeric structures	Extra-nuclear estrogen signaling
Ontology Source	REACTOME	REACTOME	REACTOME	REACTOME	REACTOME
Term PValue	0.002263	0.000032	0.000324	0.001972	0.000271
Term PValue Corrected with Benjamini-Hochberg	0.002263	0.000318	0.000648	0.002191	0.000677
Group PValue	0.000148	0.000148	0.000148	0.000148	0.000148
Group PValue Corrected with Benjamini-Hochberg	0.000297	0.000297	0.000297	0.000297	0.000297
GO Levels	[-1]	[-1]	[-1]	[-1]	[-1]
GO Groups	Group3	Group3	Group3	Group3	Group3
% Associated Genes	4.69	5.56	9.09	4.92	5.19
Nr. Genes	3	5	3	3	4
Associated Genes Found	[MMP3, MMP7, MMP9]	[ADAMTS14, COL22A1, MMP3, MMP7, MMP9]	[MMP3, MMP7, MMP9]	[MMP3, MMP7, MMP9]	[HBEGF, MMP3, MMP7, MMP9]