SUPPLEMENTAL FIGURES

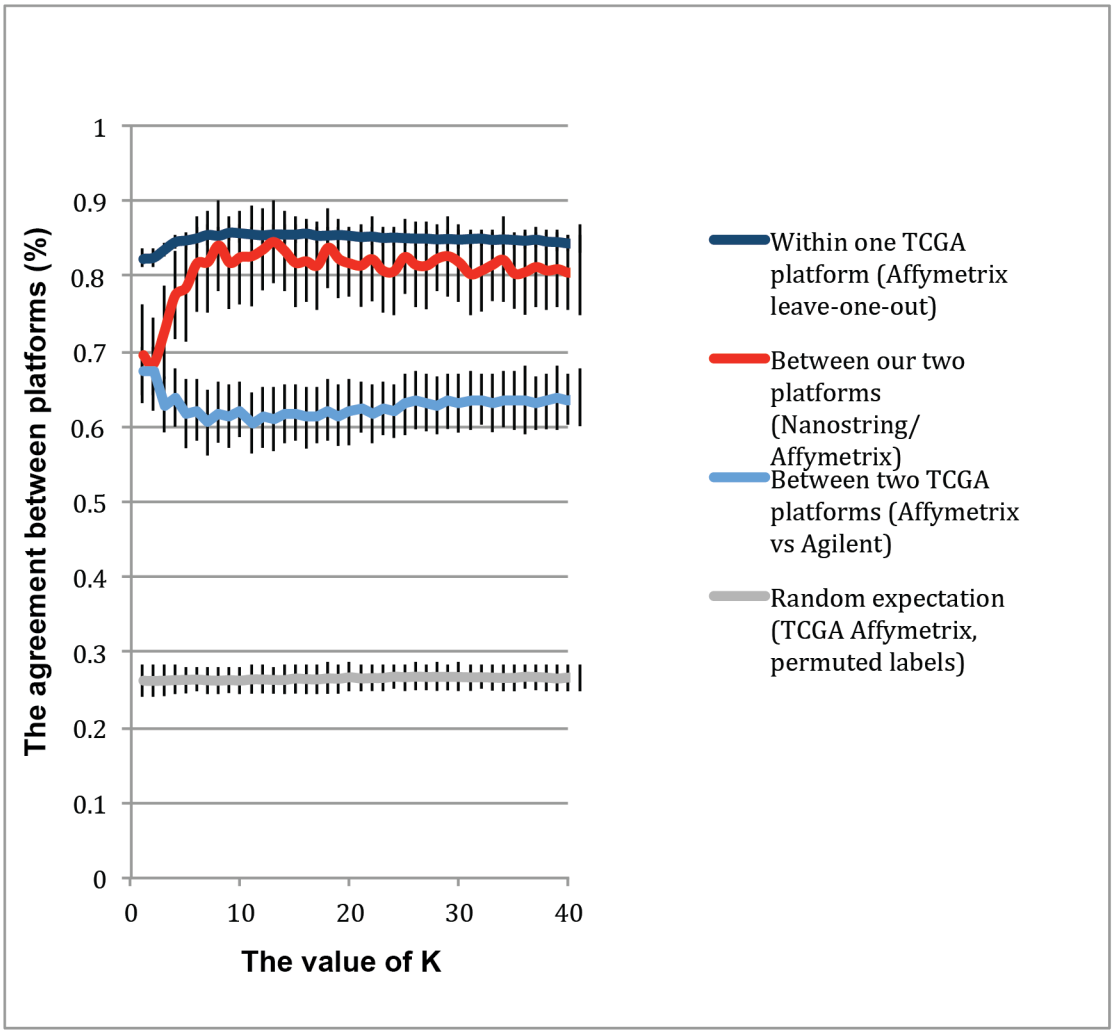

Figure S1. The K nearest neighbor classifier assigns molecular subclass to each sample based on the K most similar reference samples in the TCGA data. Related to figure 2. We noted that the classifications differed somewhat between our Nanostring and Affymetrix profiles (red curve). Depending on the value of K, the agreement is >80% for K>6 (red curve). The optimum of the curve agrees well with the value of K=11 selected from leave-one-out agreement within one platform (dark blue curve). To address whether 80% is an acceptable level of agreement, we consider two points of reference. First, we consider the agreement in classification when using the two different datasets available from the TCGA, collected on Affymetrix and Agilent platforms for the same set of samples: in this case, the agreement is somewhat lower (between 60 and 68%, light blue curve). The classification stability is thus higher between our Nanostring and Affymetrix profiles, than between TCGA's own two platforms. As a second point of reference, note that KNN classification among four approximately equally sized classes has only ~25% expected agreement.

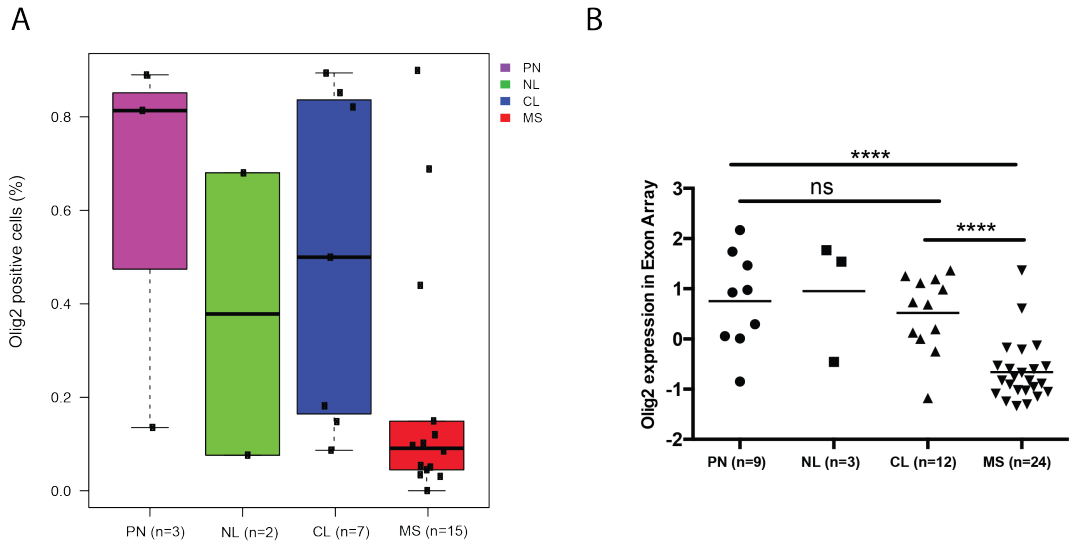

Figure S2. The protein (A) and mRNA (B) levels of Olig2 expression in Proneural (PN), Neural (NL), Classical (CL) and Mesenchymal (MS) subtypes. Related to Figure 4.

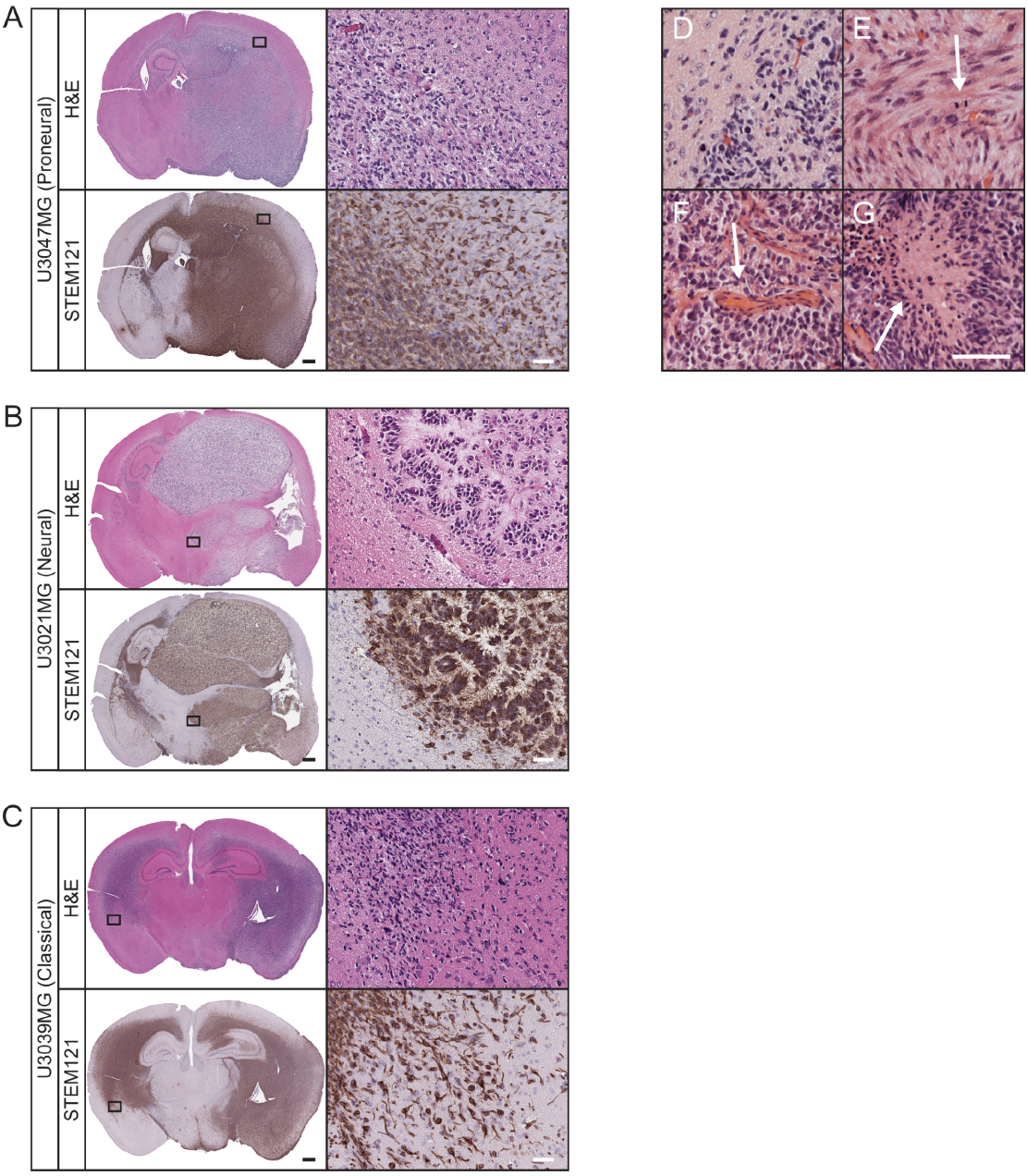

Figure S3. Secondary tumors in NOD-SCID mice. Related to Figure 5.
Representative sections from tumors induced by a proneural GC line (A, U3047MG), a neural GC line (B, U3021MG), and a classical GC line (C, U3039MG) were stained with H&E and antibodies against STEM121. Black scale bar=500 μm, white scale bar=50 μm. (D-G) H&E staining demonstrating (D) diffuse invasion, (E) mitosis (arrow), (F) microvascular proliferation (arrow) and (G) pseudopalisading necrosis (arrow), four features of high-grade glioma in the human GC xenografts. Scale bar=50 μm.

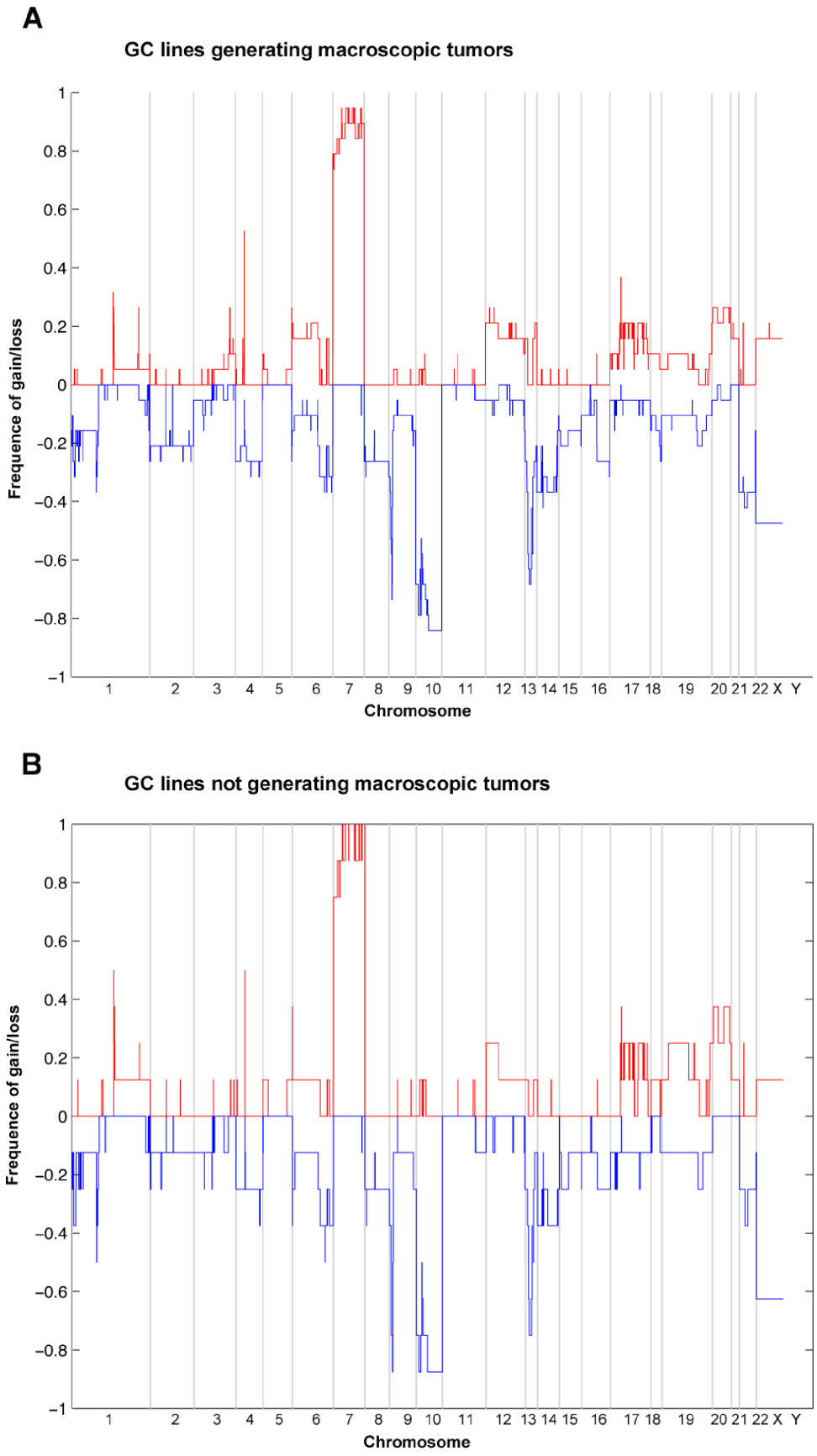

Figure S4. Analysis of the variation in genomic copy number. Related to Figure 6.
(A) GC lines generating macroscopic tumors, (n=20) and (B) GC lines not generating such tumors, (n=10).

SUPPLEMENTAL TABLES

Table S1. The patient cohort and cell lines established

Diagnosis	Number of patients	Sex		Age (years)		Number of recurrent samples	Cell lines from reccurrent samples	Total number of samples	Total number of cell lines	Overall success rate (%)
		Male	Female	Mean	Range					
All glioma	**101**	**56**	**45**	**60.7**	**21-82**	**9**	**1**	**114**	**54**	**47.4**
Grade IV	**83**	**48**	**35**	**62.7**	**21-82**	**7**	**1**	**95**	**53**	**55.8**
Glioblastoma	82	47	35	62.5	21-82	7	1	94	53	56.4
Gliosarcoma	1	1	0	78	-	0	-	1	0	0
Grade III	**13**	**4**	**9**	**50.5**	**26-73**	**2**	**0**	**13**	**1**	**7.7**
Anaplastic astrocytoma	4	0	4	45.8	29-63	1	-	4	1	25
Anaplastic oligodendroglioma	9	4	5	52.6	26-73	1	0	9	0	0
Grade II	**5**	**4**	**1**	**55.6**	**42-68**	**0**	**0**	**6**	**0**	**0**
Oligodendroglioma	4	3	1	55.3	42-68	0	-	5	0	0
Pleomorphic xanthoastrcytoma	1	1	0	57	-	0	-	1	0	0

Table S2. Comprehensive GC line information

GC line	Patient data				Cells				Tissue Subtype	IDH1 status	Tumor[c]	Expression of cell line markers (%)				Proliferative capacity[d]
					Passages		Subtype									
	Age (years)	Sex	Survival (days)	STR	EA[a]	NS[b]	EA[a]	NS[b]	NS[b]			Sox2	Nestin	GFAP	Olig2	
U3002MG	74	F	596	N	11	ND	CL	ND	ND	WT	ND	ND	ND	ND	ND	LP
U3004MG	70	M	134	Y	6	13	CL	CL	CL	WT	+	100	98.7	91.8	85.2	ND
U3005MG	65	M	26	N	5	ND	PN	ND	ND	WT	+	ND	ND	ND	ND	IP
U3008MG	64	F	38	Y	9	12	NL	MS	CL	WT	+	100	99.5	92.2	7.8	HP
U3009MG	60	M	174	Y	11	12	CL	CL	MS	WT	+	100	99.6	2.1	89.4	IP
U3013MG	78	F	122	Y	6	ND	PN	ND	ND	WT	+	100	99.4	24.6	13.5	HP
U3016MG	70	M	136	N	7	ND	PN	ND	ND	WT	ND	ND	ND	ND	ND	HP
U3017MG	68	M	442	Y	4	8	CL	CL	NL	WT	+	99.4	99.5	87.2	14.7	IP
U3019MG	82	F	117	N	11	ND	PN	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3020MG	68	M	160	N	10	ND	MS	ND	ND	WT	ND	ND	ND	ND	ND	HP
U3021MG	50	M	387	Y	4	ND	NL	ND	ND	WT	+	100	99.6	93.6	68.2	HP
U3024MG	73	F	170	Y	5	11	MS	MS	PN	WT	+	98.8	94.8	36.7	68.8	HP
U3027MG	44	M	766	Y	7	ND	MS	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3028MG	72	F	496	N	6	ND	CL	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3029MG	70	F	58	Y	5	ND	MS	ND	ND	WT	+/-	ND	ND	ND	ND	ND
U3031MG	65	F	468	Y	7	ND	MS	ND	ND	WT	+	100	99.1	7.2	3.1	ND
U3033MG	58	M	178	Y	5	ND	PN	ND	ND	WT	ND	ND	ND	ND	ND	IP
U3034MG	73	M	539	Y	11	13	MS	MS	MS	WT	+/-	100	98.4	20.9	4.5	IP
U3035MG	55	F	578	Y	6	9	MS	MS	MS	WT	+/-	100	100	45.7	3.4	IP
U3037MG	60	M	333	Y	5	7	MS	MS	NL	WT	-	100	99.5	6.1	9.7	IP
U3039MG	38	M	994	Y	8	12	CL	CL	CL	WT	+	100	98.0	65.8	50.0	LP
U3042MG	67	M	202	Y	4	8	CL	CL	CL	WT	+	100	100	80.2	82.1	LP
U3046MG	73	M	186	Y	4	7	MS	CL	NL	WT	+	100	99.5	56.2	44.0	IP
U3047MG	66	F	496	Y	5	ND	PN	ND	ND	WT	+	100	5.9	7.9	81.4	IP
U3048MG	77	M	279	Y	7	ND	CL	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3051MG	72	M	493	N	6	9	MS	MS	PN	WT	+/-	100	100	100	89.9	ND
U3053MG	64	M	277	Y	6	8	MS	MS	PN	WT	+/-	100	100	88.1	14.9	LP
U3054MG[e]	60	F	611	Y	5	ND	MS	ND	ND	WT	+	100	84.9	0.2	0.0	HP
U3056MG	68	F	110	N	6	ND	CL	ND	ND	WT	ND	ND	ND	ND	ND	IP

Table S2. Comprehensive GC line information

U3060MG	60	M	493	Y	10	ND	MS	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3062MG	76	F	166	N	5	ND	MS	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3065MG	77	M	127	Y	10	9	MS[f]	PN	PN	WT	+	99.0	100	24.8	18.1	IP
U3067MG	67	F	305	Y	7	11	MS	MS	MS	WT	-	100	71.9	1.7	23.9	ND
U3068MG	69	M	397	Y	10	ND	MS	ND	ND	WT	+/-	100	100	16.6	8.5	LP
U3071MG	65	M	309	Y	6	8	MS	MS	MS	WT	+/-	100	98.1	12.0	11.9	ND
U3073MG	71	M	481	Y	9	7	MS	MS	PN	WT	+	100	100	1.0	5.4	IP
U3078MG	51	M	716	Y	10	5	MS	CL	NL	WT	+	100	100	0.5	10.2	LP
U3082MG	70	F	314	Y	5	8	PN	PN	MS	WT	+	100	36.0	0.7	89.0	IP
U3084MG[g]	72	M	444	Y	10	5	CL	CL	CL	WT	+	ND	ND	ND	ND	ND
U3085MG	68	M	110	Y	9	ND	NL	ND	ND	WT	ND	ND	ND	ND	ND	IP
U3086MG[g]	72	M	444	Y	7	9	CL	CL	CL	WT	+	99.5	97.5	3.2	8.7	LP
U3088MG	67	F	464	Y	11	7	MS	CL	PN	WT	+/-	99.5	90.5	78.7	5.1	ND
U3101MG	73	M	284	N	6	ND	MS	ND	ND	WT	ND	ND	ND	ND	ND	IP
U3110MG	58	F	243	N	5	ND	MS	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3117MG[h]	57	M	793	N	5	ND	PN	ND	ND	WT	ND	ND	ND	ND	ND	HP
U3118MG[h]	57	M	793	N	4	ND	PN	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3121MG	64	M	479	N	5	ND	MS	ND	ND	WT	ND	ND	ND	ND	ND	ND
U3123MG	64	M	925[i]	N	3	ND	CL	ND	ND	WT	ND	ND	ND	ND	ND	LP

M: male, F: female, Y: yes, N:no, ND: not determined, CL: classical subtype, MS: mesenchymal subtype, NL: neural subtype, PN: proneural subtype, WT: wild type

[a] Exon array
[b] NanoString
[c] Tumor present at the 20 week endpoint.
[d] LP: low proliferative capacity, IP: intermediate proliferative capacity, HP: high proliferative capacity.
[e] Surgical sample from recurrent tumor
[f] Subtype assignment based on HTA array
[g] Surgical sample from the same patient
[h] Surgical sample from the same patient
[i] Patient alive 2015-02-04

Table S3. STR profile information

GC line	Amelogenin		CS1FPO		D13S317		D16S539		D18S51		D19S433		D21S11		D2S1338		D3S1358		D5S818		D7S820		D8S1179		FGA		TH01		TPOX		vWA	
	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2
U3004MG	X	Y	9	11	11	12	11	11	16	19	15.2	16.2	29	31.2	19	24	15	15	11	12	9	10	13	14	20	23	7	7	8	8	16	17
U3008MG	X	X	11	12	9	11	11	12	14	18	15	15	30	33.2	19	25	14	15	11	12	11	12	11	15	21	25	8	9	8	11	16	16
U3009MG	X	Y	10	12	8	12	9	9	13	15	14	15	29	31	19	25	16	17	13	13	9	11	14	14	19	22.2	9	9.3	8	8	16	17
U3013MG	X	X	10	12	14	14	10	10	13	14	14	14	31.2	32.2	25	26	17	19	12	12	9	11	13	14	18	24	6	9.3	8	8	16	16
U3017MG	X	Y	11	11	11	11	9	12	14	18	14	16.2	29	31.2	20	25	15	16	11	12	9	10	10	15	20	26	6	9	8	9	14	18
U3019MG	X	X	10	12	14	14	10	10	13	14	14	14	31.2	32.2			17	19	12	12	9	11	13	14	18	24			8	8	16	16
U3020MG	X	Y	10	11	9	12	11	13	16	16	13	15	31.2	31.2	23	24	15	15	11	13	10	12	12	13	21	21	6	7	8	11		
U3021MG	X	Y	12	12	11	12	9	12	13	15	14	15	28	29	17	23	16	16	10	12	8	10	13	13	21	25	6	9	8	8	16	18
U3024MG	X	X	10	12	14	14	12	13	15	17	14	15	32.2	32.2	17	17	15	18	10	12	8	12	12	15	19	23	8	9.3	8	8	14	14
U3027MG	X	Y	12	12	11	13	12	12	16	18	13	14	30	31.2	21	24	17	18	11	12	9	12	12	14	22	23	6	7	8	10	18	19
U3029MG	X	X	10	10	11	11	12	12	12	14	14	15	29	30	17	18	17	17	12	13	8	8	13	15	21	23	6	8	8	11	18	18
U3031MG	X	X	11	12	11	12	11	12	14	18	14.2	14.2	30	33.2	18	19	15	17	11	12	9	10	11	12	22	24	6	9.3	9	11	18	20
U3033MG	X	Y	11	12	12	12	10	12	17	17	14	14	29	32.2	16	24	18	18	11	12	8	11	12	14	22	22	7	9	8	11	16	18
U3034MG	X	Y	11	12	11	12	11	14	13	16	13	13.2	27	28			18	18	11	12	8	11	15	16	19	22	9.3	9.3	11	11	15	17
U3035MG	X	X	10	12	12	12	11	14	14	18	14	15	28	30.2	17	17	16	17	10	11	10	10	11	12	21	25	7	9	8	11	16	18
U3037MG	X	Y	11	11	11	13	11	14	12	16	15	15	30	31.2	25	25	14	17	12	12	11	11	12	13	21	25	7	9.3	8	9	16	17
U3039MG	X	Y	12	14	9	13	11	11	12	14	12	15	29	33.2	17	20	14	16	11	12	11	12	11	14	21	22	7	9.3	8	9	14	18
U3042MG	X	Y	11	11	12	13	12	13	16	19	14	15	31	31	17	20	15	15	11	11	9	10	13	13	21	23	6	7	8	12	17	18
U3046MG	X	Y	10	12	9	9	11	11	14	16	13	14	28	31	20	22	14	18	12	12	10	12	13	15	23	25	9	9.3	8	11	17	17
U3047MG	X	X	10	11	8	13	9	9	14	20	14	15.2	28	31	23	23	16	17	12	12	10	10	13	13	21	23	7	9.3	8	8	14	19
U3048MG	X	Y	11	11	8	8	11	13	13	18	15	15.2	30	31	23	23	14	15	11	12	9	9	13	14	23	23	7	9.3	8	12	14	16
U3053MG	X	Y	10	10	8	12	11	12	14	18	14	15	28	31	16	24	15	18	12	12	9	10	10	12	21	22	9.3	9.3	8	8	16	16
U3054MG	X	X	12	12	9	9	9	12	12	12	14	14.2	31	32.2	23	23	15	15	11	11	10	12	13	17	21	23	6	7	8	8	16	17
U3060MG	X	Y	11	11	13	13	11	13	17	18	14	15	28	29	17	24	15	16	12	12	8	9	10	14			9	9.3	8	8	15	17
U3065MG	X	Y	10	13	8	10	11	12	12	15	13	13	28	30.2	19	21	16	17	11	11	8	11	14	15	21	25	9.3	9.3	8	12	14	16
U3067MG	X	X	10	12	11	12	11	13	14	17	13	14	27	29	17	18	15	15	12	12	10	12	14	14	23	23	9	10	8	9	19	19
U3085MG	X	Y	10	12	13	13	11	11	12	14	14	14	28	30	17	18	15	17	11	12	10	12	12	15	21	24	6	9.3	10	11	15	17
U3071MG	X	Y	9	9	11	11	11	12	15	15	14	14.2	30.2	31	20	23	16	18	11	13	10	10	13	13	19	23	9.3	10	8	11	15	18
U3073MG	X	Y	10	10	10	10	12	12	12	16	13	15	29	30.2	20	25	17	17	11	12	9	11	10	12	19	21	6	8	8	9	17	17
U3078MG	X	Y	10	11	10	12	11	12	14	15	13	15	28	29	17	24	14	17	11	11	8	10	13	13	23	25	9	9.3	8	9	15	17
U3082MG	X	X	12	14	9	12	12	13	12	12	14	15	28	31.2	16	19	15	16	9	12	9	13	12	13	22	25	7	9.3	8	8	16	16
U3084MG	X	Y	10	12	8	11	11	14	12	18	13	13	29	30	17	20	15	17	11	14	8	10	13	14	22	22	6	9.3	8	12	16	17
U3085MG	X	Y	10	12	13	13	11	11	12	14	14	14	28	30	17	18	15	17	11	12	10	12	12	15	21	24	6	9.3	10	11	15	17
U3086MG	X	Y	10	12	8	11	11	14	12	18	13	13	29	30	17	20	15	17	11	14	8	10	13	14	22	22	6	9.3	8	12	16	17
U3088MG	X	X	9	12	12	12	11	11	17	18	15	16.2	27	30	19	24	15	16	10	11	11	12	13	14	23	24	9.3	9.3	8	8	15	17

A1: allele 1, A2: allele 2

Table S4. NanoString probes

Gene name
CDH4
CDK6
CHI3L1
CLEC2B
COL1A2
DLL3
DNM3
EGFR
EMP3
GRIA2
GRIK1
IGFBP6
LGALS3
NCAM1
OLIG2
PDPN
PHF11
SERPINA1
SLC4A4
SOCS2
SOX9
SPRY2
TIMP1

Table S5. IDH1 mutation analysis

	Nucleotide state[a]			
GC line	A	C	G	T
U3002MG	0	191	1	0
U3004MG	0	218	0	0
U3004MG	0	205	1	0
U3005MG	0	255	1	0
U3008MG	0	207	0	0
U3008MG	0	228	0	0
U3009MG	0	198	0	0
U3009MG	0	184	0	0
U3013MG	0	132	0	0
U3013MG	0	145	0	0
U3016MG	0	200	0	0
U3017MG	0	224	0	0
U3017MG	0	242	0	0
U3019MG	0	157	0	0
U3020MG	0	251	0	0
U3021MG	0	227	0	1
U3024MG	0	244	0	0
U3024MG	0	233	0	0
U3027MG	0	188	0	0
U3028MG	0	225	0	0
U3029MG	0	250	0	1
U3031MG	0	206	0	1
U3031MG	0	221	0	0
U3033MG	0	211	0	0
U3034MG	0	87	0	0
U3034MG	0	174	0	0
U3035MG	0	217	0	0
U3037MG	1	249	0	0

	Nucleotide state[a]			
GC line	A	C	G	T
U3039MG	0	228	1	0
U3042MG	0	233	1	0
U3046MG	0	167	0	0
U3047MG	0	164	1	0
U3048MG	0	226	0	0
U3051MG	0	203	0	0
U3053MG	0	186	1	0
U3054MG	0	254	0	0
U3056MG	0	302	0	0
U3060MG	0	234	0	0
U3062MG	0	236	0	0
U3065MG	0	254	0	0
U3067MG	0	240	1	0
U3068MG	0	232	0	0
U3071MG	0	236	0	0
U3073MG	0	196	0	1
U3078MG	0	193	1	0
U3082MG	0	138	0	1
U3084MG	0	170	1	0
U3085MG	0	206	1	0
U3086MG	0	268	1	0
U3088MG	0	267	0	3
U3101MG	0	253	0	0
U3110MG	0	262	0	0
U3117MG	0	213	2	0
U3118MG	0	229	0	0
U3121MG	0	309	0	1
U3123MG	0	208	0	0

[a] Nucleotide state at Chr2:209,113,112 is given as number of reads on the Ion Proton sequencer.

Table S6. Number of mice injected with each GC line

GC line	Number of mice
U3004MG	6
U3005MG	11
U3008MG	8
U3009MG	9
U3013MG	6
U3017MG	6
U3021MG	8
U3024MG	8
U3029MG	7
U3031MG	7
U3034MG	9
U3035MG	6
U3037MG	6
U3039MG	8
U3042MG	7
U3046MG	14
U3047MG	9
U3051MG	11
U3053MG	6
U3054MG	10
U3065MG	8
U3067MG	11
U3068MG	6
U3071MG	6
U3073MG	7
U3078MG	11
U3082MG	7
U3084MG	6
U3086MG	5

Table S7. Copy number alteration frequencies compared between HGCC and TCGA

Regions of interest (Brennan et al., 2013)	HGCC-Up	TCGA-Up	HGCC-Down	TCGA-Down	Adjusted p-value[a]
10q23.31	0	0,001751313	0,916666667	0,887915937	1
13q14.2	0,0625	0,010526316	0,520833333	0,435087719	0,179031096
13q21.2	0,0625	0,014035088	0,416666667	0,414035088	1
22q13.31	0,020833333	0,056042032	0,375	0,369527145	1
9p21.3	0,041666667	0,036777583	0,770833333	0,749562172	1
14q24.3	0,0625	0,054290718	0,395833333	0,309982487	1
14q13.1	0,125	0,054290718	0,375	0,323992995	1
6q26	0,0625	0,042031524	0,3125	0,320490368	1
15q14	0,104166667	0,038528897	0,3125	0,267950963	1
6q22.2	0,104166667	0,042105263	0,25	0,256140351	1
1p36.23	0,083333333	0,126315789	0,3125	0,254385965	1
11p15.5	0,083333333	0,03502627	0,125	0,217162872	1
17p13.1	0,166666667	0,078809107	0,104166667	0,162872154	1
19q13.33	0,125	0,288966725	0,1875	0,166374781	1
16q23.1	0,104166667	0,070175439	0,166666667	0,173684211	1
4q35.1	0,020833333	0,050788091	0,270833333	0,157618214	1
16p12.1	0,083333333	0,064912281	0,1875	0,157894737	1
2q37.1	0,083333333	0,052539405	0,166666667	0,103327496	1
17q11.2	0,333333333	0,09314587	0,041666667	0,133567663	0,0000582
18q22.3	0,1875	0,113835377	0,229166667	0,15061296	1
12p13.1	0,375	0,11033275	0,104166667	0,119089317	0,000049
3p21.31	0,291666667	0,10877193	0,145833333	0,092982456	0,010605268
3q29	0,354166667	0,143607706	0,083333333	0,117338004	0,026041574
3q13.31	0,354166667	0,11558669	0,083333333	0,096322242	0,00077578
2q22.1	0,083333333	0,052631579	0,166666667	0,066666667	0,679533013
1p22.1	0,1875	0,143607706	0,1875	0,082311734	0,758991374
1p33	0,1875	0,147110333	0,166666667	0,061295972	0,40013598
7p11.2	0,9375	0,865148862	0,041666667	0,010507881	0,669319705
7q34	0,9375	0,828371278	0	0,012259194	1
7q31.2	0,895833333	0,817863398	0	0,00525394	1
7q21.2	0,9375	0,833625219	0	0,010507881	1
12q14.1	0,4375	0,229422067	0,083333333	0,084063047	0,170670978
4q12	0,1875	0,180385289	0,166666667	0,071803853	1
1q32.1	0,291666667	0,266666667	0,020833333	0,028070175	1
12q15	0,4375	0,143607706	0,0625	0,112084063	0,000047
12p13.32	0,375	0,119089317	0,0625	0,087565674	0,000234404
3q26.33	0,458333333	0,182136602	0,041666667	0,047285464	0,00115865
1q44	0,25	0,176882662	0	0,061295972	1
19q12	0,229166667	0,375438596	0,0625	0,073684211	1
2p24.3	0,020833333	0,07530648	0,1875	0,064798599	0,146734462
8q24.21	0,208333333	0,12084063	0,1875	0,078809107	0,160782873
1p36.21	0,083333333	0,168421053	0,166666667	0,112280702	1
13q34	0,166666667	0,047368421	0,3125	0,301754386	0,080638435
17q25.1	0,229166667	0,150877193	0,020833333	0,075438596	1
11p13	0,145833333	0,050788091	0,083333333	0,171628722	0,371203133

[a] p-values for differences in alteration frequencies between HGCC and TCGA using a Chi-square test and Holms correction for multiple testing.

Table S8. Copy number alteration frequencies per subtype compared between HGCC and TCGA

Regions of interest (Brennan et al., 2013)	HGCC								TCGA								Adjusted p-values[a]			
	PN-Down	NL-Down	CL-Down	MS-Down	PN-Up	NL-Up	CL-Up	MS-Up	PN-Up	NL-Up	CL-Up	MS-Up	PN-Down	NL-Down	CL-Down	MS-Down	PN	NL	CL	MS
10q23.31	0,55555556	1	1	1	0	0	0	0	0,00757576	0	0	0	0,74242424	0,94047619	0,95138889	0,90789474	1	NA	NA	NA
13q14.2	0,66666667	0,33333333	0,33333333	0,56521739	0	0	0	0,13043478	0,01515152	0,01190476	0	0,01324503	0,56060606	0,41666667	0,22916667	0,56953642	1	1	NA	1
13q21.2	0,33333333	0,33333333	0,25	0,52173913	0	0	0	0,13043478	0,01515152	0,01190476	0,00694444	0,01324503	0,53030303	0,39285714	0,21527778	0,54304636	1	1	1	1
22q13.31	0,44444444	1	0,25	0,34782609	0	0	0	0,04347826	0,10606061	0,02380952	0,00694444	0,07894737	0,34090909	0,41666667	0,375	0,38157895	1	1	1	1
9p21.3	0,77777778	0,33333333	0,75	0,82608696	0	0	0	0,08695652	0,06060606	0,04761905	0	0,05921053	0,61363636	0,75	0,91666667	0,69736842	1	1	NA	1
14q24.3	0,55555556	0,33333333	0,16666667	0,47826087	0	0	0	0,13043478	0,04545455	0,08333333	0,05555556	0,05263158	0,46212121	0,33333333	0,15277778	0,36842105	1	1	1	1
14q13.1	0,66666667	0,33333333	0,08333333	0,43478261	0	0	0,08333333	0,2173913	0,03787879	0,04761905	0,05555556	0,05921053	0,43181818	0,39285714	0,16666667	0,36842105	1	1	1	1
6q26	0,33333333	0,33333333	0,25	0,34782609	0	0	0	0,13043478	0,07575758	0,04761905	0,00694444	0,04605263	0,28787879	0,35714286	0,375	0,28947368	1	1	1	1
15q14	0,22222222	0,66666667	0,16666667	0,39130435	0	0	0	0,17391304	0,0530303	0,02380952	0,00694444	0,07236842	0,49242424	0,28571429	0,11805556	0,23026316	1	1	1	1
6q22.2	0,33333333	0,33333333	0,16666667	0,26086957	0	0	0,08333333	0,17391304	0,08333333	0,04761905	0,00694444	0,03947368	0,26515152	0,22619048	0,27083333	0,27631579	1	1	1	1
1p36.23	0,44444444	0,66666667	0,33333333	0,2173913	0	0,33333333	0	0,13043478	0,11363636	0,17857143	0,06944444	0,17218543	0,18181818	0,26190476	0,35416667	0,23178808	1	1	1	1
11p15.5	0,11111111	0,33333333	0	0,13043478	0	0	0,08333333	0,13043478	0,01515152	0,08333333	0,01388889	0,05263158	0,49242424	0,0952381	0,07638889	0,21052632	1	1	1	1
17p13.1	0,11111111	0	0,08333333	0,13043478	0,55555556	0	0	0,13043478	0,11363636	0,0952381	0,02777778	0,07894737	0,24242424	0,10714286	0,0625	0,23026316	0,19903288	1	1	1
19q13.33	0,33333333	0,33333333	0	0,2173913	0,11111111	0	0,16666667	0,08695652	0,13636364	0,26190476	0,4375	0,26315789	0,29545455	0,14285714	0,08333333	0,15131579	1	1	1	1
16q23.1	0,33333333	0	0,16666667	0,08695652	0,11111111	0,33333333	0	0,13043478	0,10606061	0,09638554	0,01388889	0,09210526	0,18181818	0,21686747	0,11111111	0,20394737	1	1	1	1
4q35.1	0,55555556	0,66666667	0,16666667	0,17391304	0	0	0	0,04347826	0,07575758	0,04761905	0,04861111	0,04605263	0,28787879	0,07142857	0,10416667	0,13157895	1	0,35218191	1	1
16p12.1	0,11111111	0	0,16666667	0,2173913	0,11111111	0,33333333	0	0,08695652	0,10606061	0,10843373	0,00694444	0,07894737	0,14393939	0,19277108	0,13194444	0,19078947	1	1	1	1
2q37.1	0,33333333	0	0,08333333	0,13043478	0	0	0,08333333	0,13043478	0,06818182	0,04761905	0,02777778	0,06578947	0,15909091	0,10714286	0,07638889	0,09210526	1	1	1	1
17q11.2	0	0	0,08333333	0,04347826	0,55555556	0	0,16666667	0,39130435	0,15909091	0,08536585	0,04861111	0,07894737	0,09090909	0,09756098	0,0625	0,25	1	1	1	0,00659526
18q22.3	0,44444444	0,66666667	0,08333333	0,17391304	0,22222222	0	0	0,26086957	0,23484848	0,07142857	0,0625	0,10526316	0,15909091	0,20238095	0,05555556	0,17105263	1	1	1	1
12p13.1	0	0	0,08333333	0,17391304	0,66666667	0,33333333	0,16666667	0,34782609	0,15151515	0,07142857	0,10416667	0,11184211	0,15151515	0,11904762	0,08333333	0,13157895	0,08907666	1	1	1
3p21.31	0	0	0,16666667	0,2173913	0,33333333	0,33333333	0,08333333	0,34782609	0,18939394	0,13095238	0,0625	0,09868421	0,12878788	0,0952381	0,09027778	0,08552632	1	1	1	0,04375153
3q29	0	0	0,16666667	0,08695652	0,33333333	0,66666667	0,08333333	0,43478261	0,23484848	0,1547619	0,09027778	0,13815789	0,13636364	0,11904762	0,13194444	0,07894737	1	1	1	0,33316078
3q13.31	0	0	0	0,17391304	0,33333333	0,66666667	0,16666667	0,39130435	0,1969697	0,14285714	0,07638889	0,09868421	0,09090909	0,14285714	0,08333333	0,09210526	1	1	1	0,02864393
2q22.1	0,33333333	0	0,08333333	0,13043478	0	0	0	0,17391304	0,04545455	0,05952381	0,03496503	0,07894737	0,11363636	0,05952381	0,04195804	0,05921053	1	1	1	1
1p22.1	0,33333333	0	0,16666667	0,17391304	0	0	0,08333333	0,30434783	0,15151515	0,17857143	0,07638889	0,19078947	0,07575758	0,03571429	0,125	0,07236842	1	1	1	1
1p33	0,44444444	0	0,08333333	0,08695652	0,22222222	0	0,16666667	0,2173913	0,15151515	0,21428571	0,08333333	0,17763158	0,08333333	0,05952381	0,05555556	0,05921053	0,28940614	1	1	1
7p11.2	0,11111111	0	0	0,04347826	0,88888889	1	1	0,91304348	0,62121212	0,94047619	0,95138889	0,94078947	0,02272727	0,01190476	0,00694444	0	1	1	1	1
7q34	0	0	0	0	0,88888889	1	1	0,91304348	0,65151515	0,91666667	0,83333333	0,90131579	0,03030303	0	0,00694444	0,00657895	1	NA	1	1
7q31.2	0	0	0	0	0,77777778	1	1	0,91304348	0,64393939	0,89285714	0,81944444	0,90131579	0,00757576	0	0,00694444	0	1	NA	1	NA
7q21.2	0	0	0	0	0,88888889	1	0,91666667	0,95652174	0,62121212	0,91666667	0,875	0,90789474	0,03030303	0	0,00694444	0,00657895	1	NA	1	1
12q14.1	0	0	0,08333333	0,13043478	0,66666667	0,33333333	0,16666667	0,47826087	0,31818182	0,25	0,19444444	0,20394737	0,11363636	0,05952381	0,04861111	0,11184211	1	1	1	1
4q12	0,33333333	0,33333333	0,08333333	0,13043478	0,22222222	0,33333333	0,16666667	0,17391304	0,31060606	0,19047619	0,13194444	0,13157895	0,09090909	0,05952381	0,04861111	0,09210526	1	1	1	1
1q32.1	0	0	0	0,04347826	0	0,33333333	0,33333333	0,34782609	0,25757576	0,32142857	0,23776224	0,28289474	0,03787879	0	0,01398601	0,03947368	1	NA	1	1
12q15	0	0	0	0,13043478	0,66666667	0,33333333	0,16666667	0,47826087	0,14393939	0,16666667	0,11805556	0,15789474	0,20454545	0,0952381	0,04861111	0,11842105	0,05373822	1	1	0,19903288
12p13.32	0	0	0	0,13043478	0,55555556	0,33333333	0,25	0,34782609	0,15151515	0,0952381	0,11805556	0,11842105	0,09848485	0,08333333	0,04166667	0,125	1	1	1	1
3q26.33	0	0	0	0,08695652	0,66666667	0,66666667	0,16666667	0,47826087	0,28030303	0,17857143	0,125	0,17105263	0,03787879	0,08333333	0,04166667	0,03947368	1	1	1	0,22024599
1q44	0	0	0	0	0	0,33333333	0,25	0,30434783	0,20454545	0,21428571	0,125	0,19736842	0,07575758	0,02380952	0,04861111	0,05263158	1	1	1	1
19q12	0	0,33333333	0	0,08695652	0,33333333	0	0,33333333	0,13043478	0,18939394	0,3452381	0,58333333	0,34210526	0,14393939	0,04761905	0,01388889	0,07894737	1	1	1	1
2p24.3	0,33333333	0	0,08333333	0,17391304	0	0	0	0,04347826	0,11363636	0,0952381	0,03472222	0,07894737	0,11363636	0,03571429	0,04166667	0,07236842	1	1	1	1
8q24.21	0,33333333	0	0,16666667	0,17391304	0,22222222	0,33333333	0,16666667	0,17391304	0,25	0,05952381	0,07638889	0,09868421	0,0530303	0,13095238	0,0625	0,09210526	1	1	1	1
1p36.21	0,44444444	0,33333333	0	0,13043478	0	0	0	0,17391304	0,13636364	0,26190476	0,125	0,1986755	0,12878788	0,0952381	0,125	0,09271523	1	1	1	1
13q34	0,33333333	0,33333333	0,08333333	0,39130435	0,33333333	0	0	0,2173913	0,06818182	0,05952381	0,02777778	0,0397351	0,38636364	0,27380952	0,17361111	0,39072848	1	1	1	0,63307058
17q25.1	0	0	0	0,04347826	0,55555556	0,33333333	0,25	0,08695652	0,21212121	0,15662651	0,09027778	0,13815789	0,04545455	0,03614458	0,0625	0,15131579	1	1	1	1
11p13	0	0,33333333	0	0,08695652	0	0	0,16666667	0,2173913	0,02272727	0,08333333	0,04861111	0,06578947	0,35606061	0,10714286	0,04861111	0,18421053	1	1	1	1

[a] p-values for differences in alteration frequencies between HGCC and TCGA using a Chi-square test and Holms correction for multiple testing.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

GBM surgical samples and glioma cell cultures

Tumor tissue was minced with a scalpel, passed through syringes with 18- and 22-gauge needles, and then incubated in a 1:1 mixture of Accutase (eBioscience, San Diego, CA, USA) and TrypLE (Invitrogen) for 10 minutes at 37° C. The dissociated cells were washed twice with DMEM/F12 medium followed by centrifugation at 600 rpm for 8 minutes before being plated onto uncoated dishes in Neurobasal and DMEM/F12 media (1:1 mix) containing N2 and B27 supplements (Invitrogen) and human recombinant FGF2 and EGF (10 ng/ml, PEPROTECH). Five to 7 days later, the spheres were plated onto Primaria dishes (BD Biosciences) coated with mouse laminin (Sigma-Aldrich) to allow adherent growth as described previously (Pollard et al., 2009). Cells were maintained and passaged as adherent cultures and as early as possible (usually after passage 2) aliquots were taken from each passage and stored in a freezer at -150° C or in liquid nitrogen. All GC lines have reached at least 10 passages and most more than 20.

STR genotyping and data analysis

STR genotyping was performed using the AmpFℓSTR® Identifiler™ PCR amplification kit (Applied Biosystems), which amplifies 15 STR loci and the Amelogenin locus. These 15 loci were CSF1PO, D13S317, D16S539, D18S51, D19S433, D21S11, D2S1338, D31S1358, D5S818, D7S820, D8S1179, FGA, TH01, TPOX and vWA. To each reaction tube 5 µl DNA (0.1 ng/µl) was added to obtain a total volume of 12.5µl. The PCR reaction was performed in accordance with the manufacturer's instructions (AmpFℓSTR® Identifiler™ PCR Amplification Kit User Guide. Life Technologies Corporation. 2012). The data were processed by a 3730XL

Analyzer (Applied Biosystems) with GeneScan™ LIZ 500™ Size Standard (Applied Biosystem) and thereafter interpreted using the ABI PRISM® GeneMapper™ Software Version 3.0 (Applied Biosystem).

Cell line classification utilizing the *k*-nearest neighbors approach and bootstrap aggregation

From Verhaak (Verhaak et al., 2010) we obtained a list of 840 transcripts expressed differentially by the different TCGA subtypes of glioblastoma, of which 765 (91%) mapped to our expression dataset. Log2 scale TCGA mRNA expression data (n=529, Affymetrix platform) and HGCC mRNA expression data (n=48) were both normalized by Z-score transformation. Subtype assignments for the TCGA samples were obtained from the TCGA server. We applied *k*-Nearest Neighbor classification (KNN, using the Matlab function knnclassify) to assign a subtype to each of these 48 samples, employing the Pearson correlation between vectors of Z-scores for the 746 subtype marker transcripts as the similarity metric. The choice of the KNN tuning parameter *k* (set to a value of 11) and the choice of distance metric were both informed by cross-validation, among the set of TCGA reference cases (data not shown). To determine the confidence of subclass classification, we applied bootstrap simulation (Caprile et al., 2004, James et al., 2013) where each HGCC cell line was classified 1000 times, each time using a resampled TCGA dataset (529 TCGA cases randomly selected with replacement). Following this simulation, the relative weight of each subtype in each cell line (Figures 2 and 6B) was defined as the proportion of the 1000 simulations in which that cell line was assigned to that class.

Isomap analysis of the GC lines

In addition to classification (supervised analysis), Isomapping (Tenenbaum et al., 2000), a non-linear procedure for dimensionality reduction that has previously been successfully been applied to classification of cancer (Nilsson et al., 2004), was used to visualize the space of the GC lines and TCGA samples in two dimensions. Normalized expression values across the 48 cell lines and 529 TCGA tissue samples for 765 of the 840 classifier genes (Verhaak et al., 2010) represented on both Affymetrix Exon 1.0ST and U133A microarray platforms were transformed into *z*-scores (for each gene in all samples in each dataset). The *z*-scores were combined and the data matrix used for isomapping was generated as described (http://isomap.stanford.edu/). The resulting isomap provides a good assessment of the similarities and differences between samples and allows visual inspection and comparison of the assignment of subtypes to the GC lines and TCGA tumors. All data manipulations were performed in the R (R Core Team, 2014) and MATLAB (The MathWorks, Inc., Natick, MA, United States) software.

Whole-exome sequencing and data analysis of IDH1 codon R132

For each sample, 100 ng of DNA was distributed to 12 separate primer pools and amplified according to the Ion AmpliSeq™ Exome Library Preparation protocol (Revision A.0, Life Technologies). The separate PCR products were pooled and primer sequences were partially digested. Adaptors were ligated and resulting amplicons were purified using Agencourt AMPure XP Reagent (Beckman Coulter) and eluted in amplification mix (Platinum PCR SuperMix High Fidelity and Library Amplification Primer Mix, Life Technologies). Size-selection and purification was conducted using Agencourt AMPure XP Reagent. The amplicons were quantified using the Agilent Bioanalyzer instrument with Agilent High Sensitivity DNA kit.

Emulsion PCR was performed on the Ion OneTouch 2 system using Ion PI Template OT2 200 Kit v3 chemistry, followed by enrichment using Ion OneTouch ES. Samples were loaded on an Ion PI chip Kit v2 and sequenced on the Ion Proton System using Ion PI Sequencing 200 Kit v3 (200 bp read length, Life Technologies) chemistry. DNA samples for the mother and the father were barcoded and multiplexed on one PI chip, while the fetal DNA was run on a separate PI chip. The sequencing generated over 30 million reads/sample of average length >150 bp.

Alignment of reads to the human hg19 assembly and variant detection was performed using v4.0 of the Torrent Suite Software. SNPs and indel data were stored in a local installation of the CanvasDB in-house database system (Ameur et al., 2014) together with variant annotation information obtained from ANNOVAR (Wang et al., 2010) and dbSNP137. Nucleotide at chr2:209,113,112 (middle base in codon 132 of the IDH1 gene) was analyzed for substitution from C to T.

Analysis of gene expression by NanoString Technology of surgical GBM samples and GC lines

RNA extracted from 22 fresh, frozen specimens of human glioma using TRIzol and RNA from the corresponding GCs (prepared in the same manner as the Affymetrix Exon Array) were subjected to a custom made assay of gene expression by NanoString Technology. The NanoString probe set (Table S4) was designed by Drs. Cameron Brennan and Jason Huse at the Brain Tumor Center, Memorial Sloan-Kettering Cancer Center, New York. Gene expression in our GBM tissue/GC line samples was related to that of the TCGA tumor samples with known subtype to predict subtype, using the same methodology (KNN classifier) as above.

Proliferation assay

Proliferation of all the cell lines was assayed twice. For the AlamarBlue assay, 200 µl of cell suspension (2.5×10^5 cells/ml) was placed into each well of Primaria 96-well plates (BD Biosciences), performing triplicates for each time-point and cell line and three wells with media alone in each experiment for determination of background. On days 1, 4 and 7, 20 µl AlamarBlue was added to each well, the plate was incubated at 37° C for 2 more hours, and fluorescence then measured with a Wallac 1420 VICTOR plate reader (PerkinElmer).

In the case of the trypan blue test of cell viability 2 ml of cell suspension (2.5×10^4 cells/ml) was seeded into each well of Primaria 6-well plates (BD Biosciences), with duplicate wells for each time-point and cell line. On days 1, 4 and 7, the cells were detached and the number of viable cells excluding trypan blue determined with an automated cell counter (Invitrogen). The cell density on day 4 and 7 were normalized to that on day 1.

Immunofluorescence stainings

Adherent cells seeded onto polyornithine/laminin-coated coverslips were washed twice with phosphate-buffered saline (PBS) and then fixed in 4% formaldehyde/PBS for 20 min at room temperature. The cells were next blocked and permeabilized with 5% normal goat serum in PBS and 0.3% Triton-X-100 for one hour at room temperature and then incubated with primary antibodies (diluted 1:200 in PBS containing 1% bovine serum albumin and 0.3% Triton-X-100) overnight at 4° C. After being washed three times in PBS, the cells were incubated with fluorescent-

labeled secondary antibodies (diluted 1:400 in the same buffer) for 2 h 30 min at room temperature. Thereafter, the cells were washed in PBS, stained with Hoechst (2 μg/mL in PBS) for 20 min at room temperature and then washed in PBS again. Following this final wash, the coverslips were mounted with Fluoromount (DAKO).

The primary antibodies used were rabbit anti-SOX2 (Millipore AB5603), mouse anti-GFAP (Sigma G3893), rabbit anti-OLIG2 (Millipore AB9610), mouse anti-NES (Millipore MAB5326), rabbit anti-S100B (DAKO Z0311), and mouse anti-TUBB3 (Covance MMS-435P). The fluorescent-labeled secondary antibodies used were goat anti-mouse Alexa 488 (Invitrogen A11029) and goat anti-rabbit Alexa 555 (Invitrogen A21430). Images were acquired with a Zeiss AxioImager fluorescence microscope at 200X magnification. Automatic counting of nuclear stains (DAPI, SOX2, OLIG2) were performed with ImageJ software (Schneider et al., 2012) while Nestin and GFAP stainings were manually counted.

Analysis of the tumorigenicity of the GC lines

All animal experiments were performed in accordance with the rules and regulations of Uppsala University and approved by the local animal ethics committee. Neonatal non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice (P1-3) were injected intracerebrally with 1.0×10^5 human GCs of 30 different lines dissolved in 2 μl DMEM/F12 (Table S2). The animals were monitored three times a week and euthanized upon any signs of illness or, otherwise, after 20 weeks. Their brains were fixed in 4% formalin at 4° C for at least 48 h, embedded in paraffin, sectioned, stained with hematoxylin and eosin (H&E) and examined for the presence of xenograft

tumors. Images were acquired at the SciLife Lab Tissue Profiling facility using an automated slide scanner (Aperio Scanscope AT, Leica Microsystems, Germany).

Immunohistochemical analysis

Tissue sections of formalin-fixed, paraffin-embedded NOD-SCID brains injected with human GC lines were deparaffinized and dehydrated prior to antigen retrieval in acidic buffer (H-3300, Vector Labs, CA, United States) in a pressure cooker. Next, the sections were incubated with the antibody towards STEM121, a human-specific cytoplasmic marker (Stem Cell Inc., diluted 1:1000 in PBS containing 5% NGS and 0.3% Triton-X-100) overnight at 4° C. The sections were then washed 4 times with PBS containing 0.05% Tween-20 and specific binding finally visualized with a UltraVision HRP polymer detection system (Thermo Scientific), in accordance with the manufacturer's instructions. Images were acquired at the SciLife Lab Tissue Profiling facility using an automated slide scanner (Aperio Scanscope AT, Leica Microsystems, Germany).

Analysis of aberrations in DNA Copy Number

For figure 6A-B, we defined the frequency of amplification and deletion, respectively, as the fraction of samples for which copy number signal at a given genomic locus is higher than a threshold T (amplified) or lower than a threshold -T (deleted), where T was set to 1 standard deviation. (Different thresholds in the range 0.5 to 2.0 standard deviations all produced qualitatively similar results as Figure 6, correlation between TCGA and HGCC 0.7 or higher, not shown). Regions of Interest (ROI) were acquired from (Brennan et al., 2013) Table S3 (DNA Copy-Number GISTIC significant peaks). Per-gene values in these regions where averaged for HGCC and TCGA CNA data to

produce a per-region score. Each region was considered to be amplified if the score was >0.1 and to have a deletion if the score was <-0.1. In Figure 6C a subset of regions are shown, matching Figure 5A in (Brennan et al., 2013). For each ROI a Chi-squared test was used to call a significant difference in occurrence between the number of amplifications and deletions in HGCC and TCGA. These values were corrected using the Holm procedure in R (p.adjust (x,method="holm")) and significance is shown as * (<0.05), ** (<0.01) or *** (<0.001). This was done to control the Family Wise Error Rate (FWER) when calling a significant difference. The same test was conducted for each subtype separately to assess the per-subtype differences between HGCC and TCGA samples. A full list of alteration frequencies and p-values are shown in Table S7 and Table S8.

Analysis of the stability of the molecular subtypes of GCs in culture and *in vivo*

Neonatal NOD/SCID mice (P1-3) were injected intracerebrally with three different human GC lines (U3020MG, U3047MG and U3065MG), two mice with each cell line and sacrificed upon any sign of illness. The tumor tissue was dissected out and one portion fixed with formalin and embedded in paraffin for H&E staining (as described above), another portion stored at -70° C for later RNA extraction and a third portion dissociated enzymatically into single cells and for establishment of a GC culture subjected to two passages. Total RNA was extracted from the GC lines prior to transplantation, from the xenograft tumor tissue and from cultured explant cells using the miRNeasy Mini kit (Qiagen). This RNA was subsequently labeled and hybridized onto Affymetrix human transcriptome 2.0 arrays following the instructions of the manufacturer, and the expression values were RMA-normalized using the Affymetrix Expression Console software. The molecular subtypes of the samples

were determined by comparison to the reference TCGA tissue samples and depicted in two-dimensional space by isomapping, as described above. For the latter analysis, 746 subtype classifications genes (Verhaak et al., 2010) were used, that were represented on both the Affymetrix HTA 2.0 and U133 microarray platforms.

SUPPLEMENTAL REFERENCES

AMEUR, A., BUNIKIS, I., ENROTH, S. & GYLLENSTEN, U. 2014. CanvasDB: a local database infrastructure for analysis of targeted- and whole genome re-sequencing projects. *Database-the Journal of Biological Databases and Curation*.

BRENNAN, C. W., VERHAAK, R. G. W., MCKENNA, A., CAMPOS, B., NOUSHMEHR, H., SALAMA, S. R., ZHENG, S. Y., CHAKRAVARTY, D., SANBORN, J. Z., BERMAN, S. H., BEROUKHIM, R., BERNARD, B., WU, C. J., GENOVESE, G., SHMULEVICH, I., BARNHOLTZ-SLOAN, J., ZOU, L. H., VEGESNA, R., SHUKLA, S. A., CIRIELLO, G., YUNG, W. K., ZHANG, W., SOUGNEZ, C., MIKKELSEN, 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. & NETWORK, T. R. 2013. The Somatic Genomic Landscape of Glioblastoma. *Cell,* 155**,** 462-477.

CAPRILE, B., MERLER, S., FURLANELLO, C. & JURMAN, G. 2004. Exact bagging with k-nearest neighbour classifiers. *Multiple Classifier Systems, Proceedings,* 3077**,** 72-81.

JAMES, G., WITTEN, D., HASTIE, T. & TIBSHIRANI, R. 2013. *An introduction to statistical learning : with applications in R,* New York, Springer.

NILSSON, J., FIORETOS, T., HOGLUND, M. & FONTES, M. 2004. Approximate geodesic distances reveal biologically relevant structures in microarray data. *Bioinformatics,* 20**,** 874-80.

POLLARD, S. M., YOSHIKAWA, K., CLARKE, I. D., DANOVI, D., STRICKER, S., RUSSELL, R., BAYANI, J., HEAD, R., LEE, M., BERNSTEIN, M., SQUIRE, J. A., SMITH, A. & DIRKS, P. 2009. Glioma Stem Cell Lines Expanded in Adherent Culture Have Tumor-Specific Phenotypes and Are Suitable for Chemical and Genetic Screens. *Cell Stem Cell,* 4**,** 568-580.

R CORE TEAM 2014. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*.

SCHNEIDER, C. A., RASBAND, W. S. & ELICEIRI, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods,* 9**,** 671-5.

TENENBAUM, J. B., DE SILVA, V. & LANGFORD, J. C. 2000. A global geometric framework for nonlinear dimensionality reduction. *Science,* 290**,** 2319-2323.

VERHAAK, R. G. W., HOADLEY, K. A., PURDOM, E., WANG, V., QI, Y., WILKERSON, M. D., MILLER, C. R., DING, L., GOLUB, T., MESIROV, J. P., ALEXE, G., LAWRENCE, M., O'KELLY, M., TAMAYO, P., WEIR, B. A., GABRIEL, S., WINCKLER, W., GUPTA, S., JAKKULA, L., FEILER, H. S., HODGSON, J. G., JAMES, C. D., SARKARIA, J. N., BRENNAN, C., KAHN, A., SPELLMAN, P. T., WILSON, R. K., SPEED, T. P., GRAY, J. W., MEYERSON, M., GETZ, G., PEROU, C. M. & HAYES, D. N. 2010. Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell,* 17**,** 98-110.

WANG, K., LI, M. Y. & HAKONARSON, H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research,* 38.