Stem Cell Reports
Report
OPEN ACCESS

Genetic and Functional Diversity of Propagating Cells in Glioblastoma

Sara G.M. Piccirillo,[1] Sue Colman,[2] Nicola E. Potter,[2] Frederik W. van Delft,[2,5] Suzanne Lillis,[3] Maria-Jose Carnicer,[2] Lyndal Kearney,[2] Colin Watts,[1,4,*] and Mel Greaves[2,*]
[1]John van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0PY, UK
[2]Centre for Evolution and Cancer, Division of Molecular Pathology, The Institute of Cancer Research, London SM2 5NG, UK
[3]Molecular Diagnostics, The Centre for Molecular Pathology, The Royal Marsden NHS Foundation Trust, London SM2 5PT, UK
[4]Division of Neurosurgery, Department of Clinical Neurosciences, University of Cambridge, Box 167 Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, UK
[5]Present address: Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne NE2 4AD, UK
*Correspondence: cw209@cam.ac.uk (C.W.), mel.greaves@icr.ac.uk (M.G.)
http://dx.doi.org/10.1016/j.stemcr.2014.11.003

SUMMARY

Glioblastoma (GBM) is a lethal malignancy whose clinical intransigence has been linked to extensive intraclonal genetic and phenotypic diversity and the common emergence of therapeutic resistance. This interpretation embodies the implicit assumption that cancer stem cells or tumor-propagating cells are themselves genetically and functionally diverse. To test this, we screened primary GBM tumors by SNP array to identify copy number alterations (a minimum of three) that could be visualized in single cells by multicolor fluorescence in situ hybridization. Interrogation of neurosphere-derived cells (from four patients) and cells derived from secondary transplants of these same cells in NOD-SCID mice allowed us to infer the clonal and phylogenetic architectures. Whole-exome sequencing and single-cell genetic analysis in one case revealed a more complex clonal structure. This proof-of-principle experiment revealed that subclones in each GBM had variable regenerative or stem cell activity, and highlighted genetic alterations associated with more competitive propagating activity in vivo.

INTRODUCTION

Glioblastoma (GBM), the most common primary brain cancer, is characterized by genetic instability and complex evolutionary dynamics. Histopathological diversity generates various clinical phenotypes whose common feature is the rapid emergence of treatment resistance to radiotherapy and chemotherapy. Dominant clonal populations that emerge as a result of genetic and epigenetic changes, which confer a tumor survival advantage, drive tumor growth (Nowell, 1976). Intratumor genetic and phenotypic heterogeneity is a hallmark of most cancers (Greaves and Maley, 2012; Marusyk et al., 2012) and is particularly marked in GBM (Bonavia et al., 2011; Sottoriva et al., 2013; Patel et al., 2014). Karyotypic (Shapiro et al., 1981) comparative genomic hybridization (Jung et al., 1999) and fluorescence in situ hybridization (FISH) screens (Little et al., 2012; Snuderl et al., 2011; Szerlip et al., 2012), as well as ultradeep, targeted sequencing (Nickel et al., 2012), have documented the intraclonal diversity of recurrent genetic abnormalities in GBM, including the amplified receptor tyrosine kinase genes *EGFR*, *PDGFRA*, and *MET*. Genetic diversity of subclones is likely to contribute to the clinical intransigence of GBM (Nicholas, 2007), and therapeutic resistance of critical tumor-propagating or stem cells is presumed to be pivotal to this issue (Bao et al., 2006; Chen et al., 2010, 2012).

Subclonal evolution and cancer stem cells (CSCs) have been considered as alternative mechanisms for disease progression (Shackleton et al., 2009), but it is also argued that they are part of the same process because clonal diversity is likely to be generated and sustained by genetically distinct CSCs, which provide the units for evolutionary selection (Greaves, 2013; Kreso and Dick, 2014). Evidence supporting this notion is found in acute lymphoblastic leukemia, where subclonal genetic architecture has been linked to the presence of genetically distinct stem cells assayed in vivo by serial xenotransplantation (Anderson et al., 2011; Notta et al., 2011). It seems likely a priori that GBM would similarly harbor genetically diverse tumor stem/propagating cell populations. This is supported by a previous study in which it was shown that distinct regions of the same dissected GBM tumors had chromosomally distinct (but clonally related) diversity, but were all expandable in vitro under serum-free stem cell conditions and transplantable in vivo as a readout of CSC (Piccirillo et al., 2009). These data raise an important question about the evolution of genetic diversity within the complex subclonal structure of GBMs. Here, we conducted an analysis at single-cell resolution of the genomic changes that occur in GBM, and used competitive clonal phylogenies (Anderson et al., 2011) before and after xenotransplantation to infer the genetics of stem/propagating cells in GBM.

RESULTS AND DISCUSSION

Using the neurosphere protocol (Piccirillo et al., 2009), we established stable patient-derived cell cultures from

dissociated tumor tissue from 12 cases of primary GBM (Table S1 available online). In order to exclude the possibility of in vitro aberrations and culture selection, as soon as the primary cells formed neurospheres, they were dissociated into single cells and used for intracerebral transplantation (and retransplantation) into NOD-SCID mice. We used high-resolution SNP arrays performed on DNA extracted from the primary GBM tumor cells to identify "driver" copy number alterations (CNAs), defined as recurrent regions of amplification or deletion (Table S2). FISH probes were designed for selected (preferably focal) CNAs, and three-color FISH was applied to neurosphere cells as well as cells isolated from tumors after secondary transplantation in NOD-SCID mice. We then compared the subclonal genetic architecture and clonal phylogenies of the neurospheres and the tumors generated in the mice. In all cases, the driver CNAs chosen from analysis of tumor DNA were present in the neurospheres and subclones that were present in the mouse xenografts could be backtracked to the original tumor, confirming the validity of this approach for investigating clonal progression. A schematic overview of the workflow is given in Figure S1.

Ten of the original 12 GBM neurosphere cultures resulted in tumors in the mice (Table S3). Four of these (GBM 2, GBM 5, GBM 8, and GBM 11) had at least three "driver" lesions that could be tracked by FISH in both the neurospheres and secondary xenografts in the mice. The remaining cases were not included because they had fewer than three "trackable" lesions by SNP array (GBM 3 and GBM 6), because various aneuploid conditions were observed in the derived neurosphere cell line (GBM 1) (Table S2), or because there were too few cells for FISH at secondary transplantation (GBM 4, GBM 7, and GBM 9). In all four cases studied by multicolor FISH, there was genetic heterogeneity in the neurosphere cells, and each case showed a unique, branched phylogenetic architecture. In each case, more than one subclone was capable of propagating tumors in secondary transplanted mice (Figures 1, 2, 3, and 4).

Analyses of clonal architecture by multicolor FISH for CNAs inevitably underestimate the extent of clonal diversity (Anderson et al., 2011). In GBM 5, we had sufficient material for a more detailed genetic analysis. The SNP arrays of primary tumor GBM 5 revealed high-level, focal amplification of *EGFR*; homozygous loss of *CDKN2A* (one large deletion and one small focal deletion); and loss of *TP53* due to a deletion of 17p (Figure S2). We observed seven subclones in the neurosphere cells and a branching subclonal structure. The major clone in the neurospheres had high-level focal amplification of *EGFR*, heterozygous *TP53*, and homozygous *CDKN2A* loss. Only two subclones read out in the secondary transplant tumor cells, and both had high-level *EGFR* amplification and homozygous *CDKN2A* loss. In contrast to the neurospheres, the major clone in the secondary mouse xenograft had two copies of *TP53*. We performed mutation screening of the *TP53* gene by capillary electrophoresis single-strand conformation analysis in primary patient tumor DNA, followed by Sanger sequencing to characterize any mobility shifts thus identified. This revealed a mutation in exon 5: c.454C > T: p.152S. The same *TP53* mutation was also found in xenograft cells after secondary transplantation (mouse 1 and mouse 3). Both wild-type and mutated *TP53* sequences were present in the tumor DNA, but only the *TP53* mutated sequence was present in the xenograft DNA, indicating that the mutation was present in all subclones of the mouse xenografts. In order to investigate this further, we carried out whole-exome sequencing and single-cell analysis for the simultaneous occurrence of CNAs and selected single-nucleotide variants (SNVs) in this case. For the latter, we used multiplex-targeted DNA amplification of flow-sorted single cells followed by high-throughput quantitative PCR (qPCR) using the BioMark HD microfluidic platform (Potter et al., 2013).

Whole-exome sequencing of tumor DNA from GBM 5 identified a total of 32 SNVs (Table S3). A subset of these SNVs was selected for single-cell analysis based on putative gene function and variant allele fractions encompassing high, low, and intermediate frequencies. The genomic targets selected included SNVs in *KCNH5*, *PLCB2*, *GDF5*, *TRMT5*, *TP53*, and *PALB2*, and CNAs in *CDKN2A*, *TP53*, and *EGFR*.

Single-cell analysis for the simultaneous presence of six SNVs and three CNAs was carried out on flow-sorted neurosphere and xenograft tumor cells from GBM 5 (a representative heatmap of the qPCR data from the BioMark HD is given in Figure S3). A comparison of the clonal phylogeny and subclonal architecture of neurosphere and xenograft cells is shown in Figure 4. Homozygous *CDKN2A* deletion, gain of *EGFR* (up to four copies), and *KCN5*, *PLCB2*, *GDF5*, and *TRMT5* mutations all occurred early and were present in all subclones of the neurospheres (Figure 4A). Loss of one *TP53* wild-type allele occurred after *EGFR* amplification of more than four copies. Heterozygous *TP53* and *PALB2* mutations occurred after further *EGFR* amplification. According to the chromosome 7 copy number as assessed by FISH (Figure S4) and single-cell data (not shown), *EGFR* gain was uncoupled from chromosome copy number at three or four copies of chromosome 7. Subsequently, there was an increasing gain of *EGFR*, consistent with the formation of extrachromosomal double minutes. All of the cells in the secondary mouse xenograft possessed all of the mutations, including heterozygous *TP53* and *PALB2* mutations, and most likely derived from two of the most evolved subclones in the neurospheres (being present in only 3.3% and 4%, respectively; Figure 4B). The

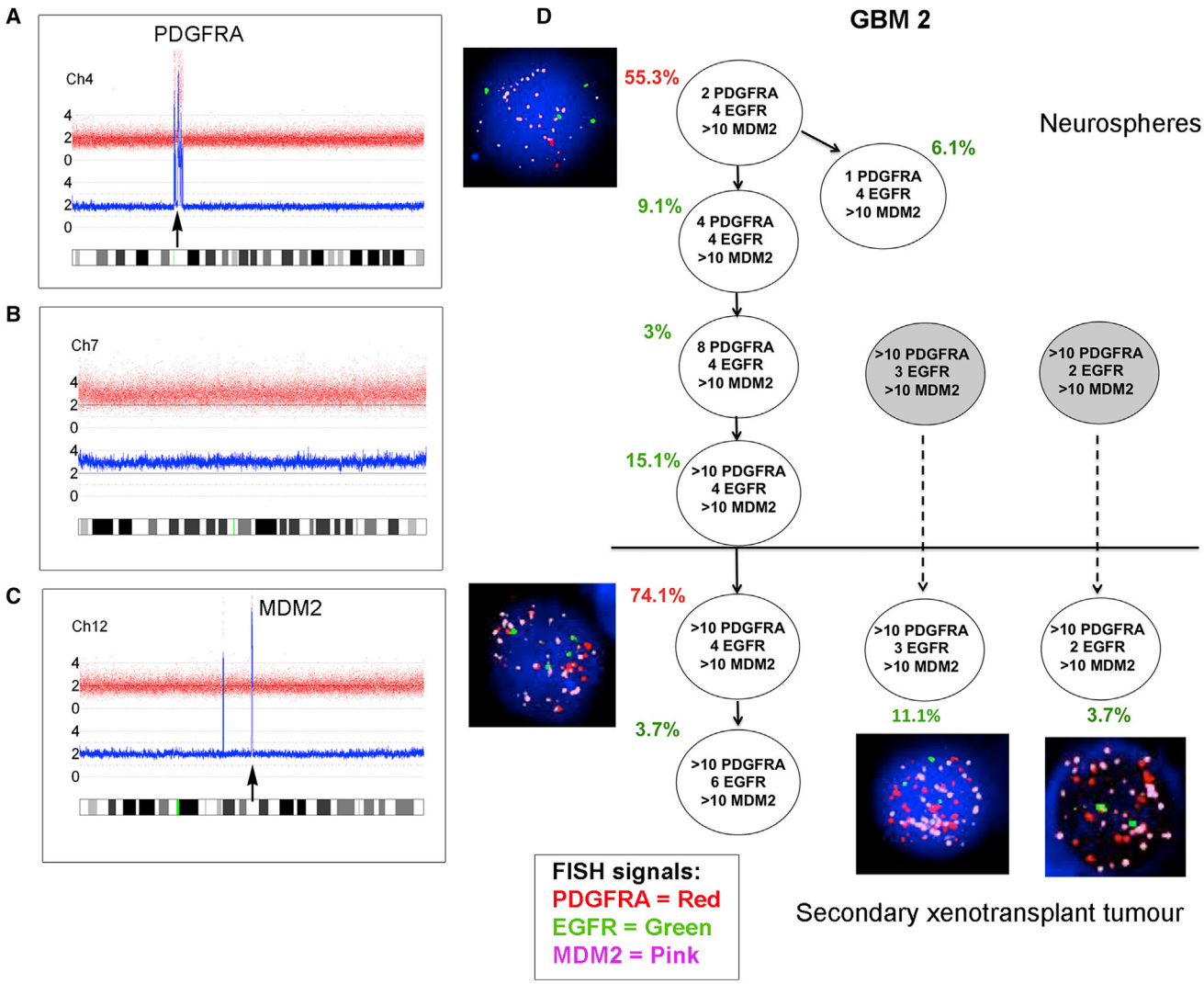


Figure 1. Subclonal Genetic Structure of Neurosphere Cells and Tumor-Propagating Cells Derived from Primary Tumor GBM 2

(A–C) SNP 6 array profiles of DNA from GBM 2 primary tumor showing high-level *PDGFRA* amplification (A), whole chromosome 7 gain (B), and high-level *MDM2* amplification (C).

(D) Subclonal genetic structure in the neurospheres (top) and after secondary transplantation in a single mouse (m3) (bottom). Subclones represented by gray circles were not present in the neurospheres above the threshold detection level. FISH images are shown next to their respective genotype. Red type indicates the major clone. Solid arrows show probable derivation of subclones. Dashed arrows indicate possible alternative derivation of subclones. FISH images were captured at 100× magnification.

See also Figure S1 and Tables S1, S2, S3, and S4.

subclone with one copy of *TP53* mutant and more than ten copies of *EGFR* evolved further in the xenograft cells by acquiring two mutated copies of *TP53*, and all subclones evolved to show high-level amplification of *EGFR* (>100 copies) (Figure 4B).

We used secondary transplantation as a more stringent measure of stem cell renewal (Dick et al., 1997). In five cases (GBM 1, GBM 5, GBM 8, GBM 9, and GBM 11), we observed a statistically shorter time to tumor formation in the secondary transplant than in the primary xenograft tumor. This pattern of evolution is consistent with the typical pattern of disease progression seen in patients and would be consistent with the genetically more evolved subclonal structure observed in the neurospheres of GBM 5, GBM 8, and GBM 11, and with the presence of *TP53* mutations in GBM 5 and GBM 8.

Clones with *EGFR* amplification consistently read out after serial transplantation, and usually further evolved with an incremental gain of more copies of *EGFR*. Other investigators have demonstrated a mosaic pattern of growth

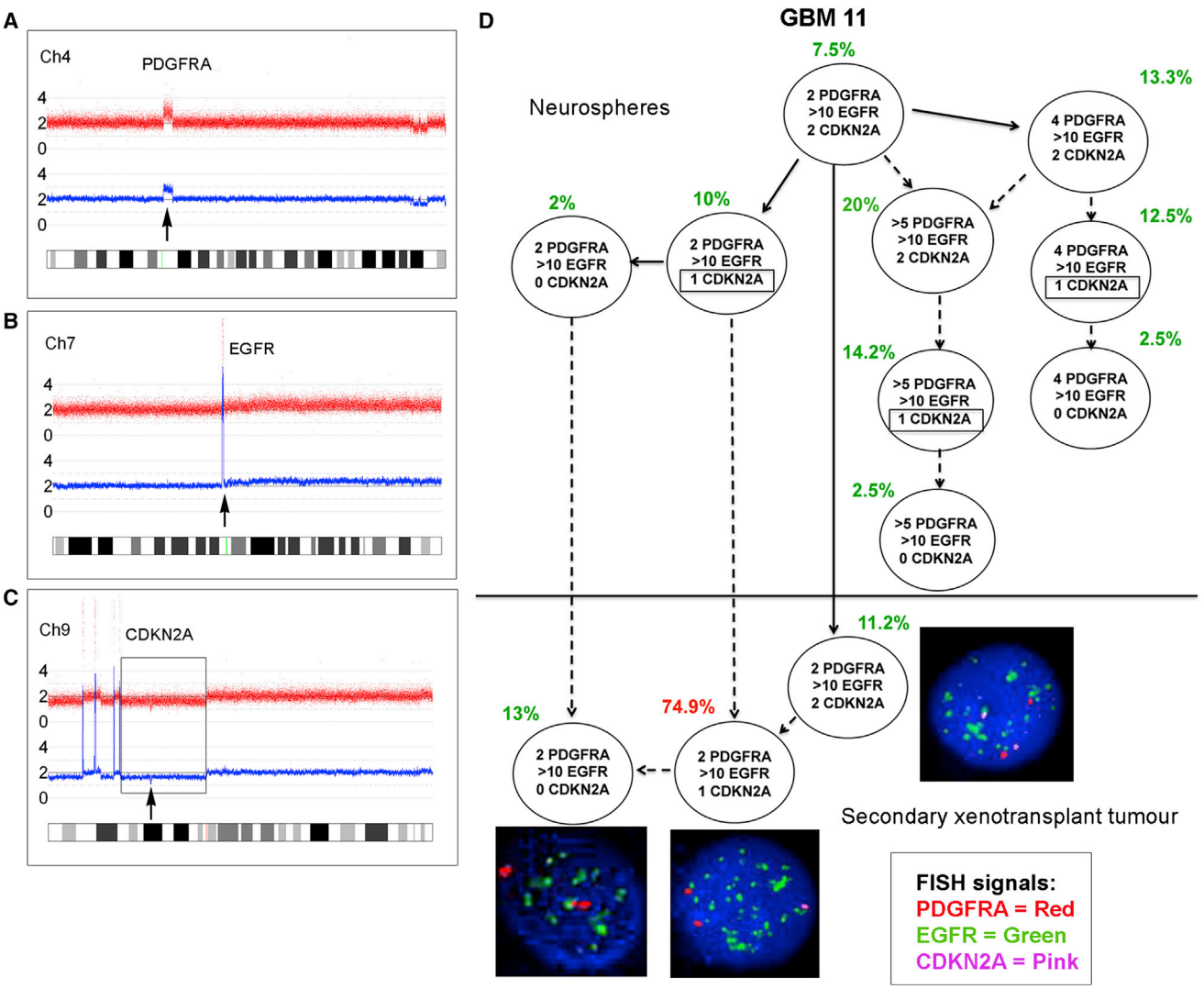


Figure 2. Subclonal Genetic Structure of Neurosphere Cells and Tumor-Propagating Cells Derived from Primary Glioblastoma GBM 11

(A–C) SNP 6 array profiles of DNA from GBM 11 primary tumor showing *PDGFRA* amplification (A), high-level *EGFR* amplification (B), and homozygous *CDKN2A* deletion (C) comprised of a large deletion of one allele (box) and focal deletion of the second allele (arrow).

(D) Subclonal structure of neurosphere cells (top) and tumor cells after secondary transplant in a single mouse (m3) (bottom). Potentially three different *CDKN2A* deletions occur in different subclones in the neurospheres (indicated by boxes). Solid arrows show probable derivation of subclones. Dashed arrows indicate possible alternative derivation of subclones. FISH images were captured at 100× magnification.

See also Figure S1 and Tables S1, S2, S3, and S4.

factor amplification in GBM tumors, with *EGFR*, *MET*, and *PDGFRA* gain occurring in distinct populations of cells (Snuderl et al., 2011; Szerlip et al., 2012). In the present study, there was one case (GBM 11) with subclones in the neurosphere cells that showed concurrent *PDGFRA* and *EGFR* gain in the same cell, as well as subclones with only *EGFR* gain (Figure 2). However, only the subclones with high-level *EGFR* amplification repopulated the mouse xenograft; none of the subclones with *PDGFRA* gain were present. These observations reveal the dynamic complexity of subclonal interactions in GBM and provide deeper insight into the role of *PDGFRA*. We previously showed that amplification of *PDGFRA* occurs in the midphase of GBM evolution (Sottoriva et al., 2013) rather than as a primary driver event. Evidence suggests that tumor-propagating clones may arise from a common precursor, with key early events including genetic alterations in *EGFR*, *CDKN2A/B*, and *TP53* (Goodenberger and Jenkins,

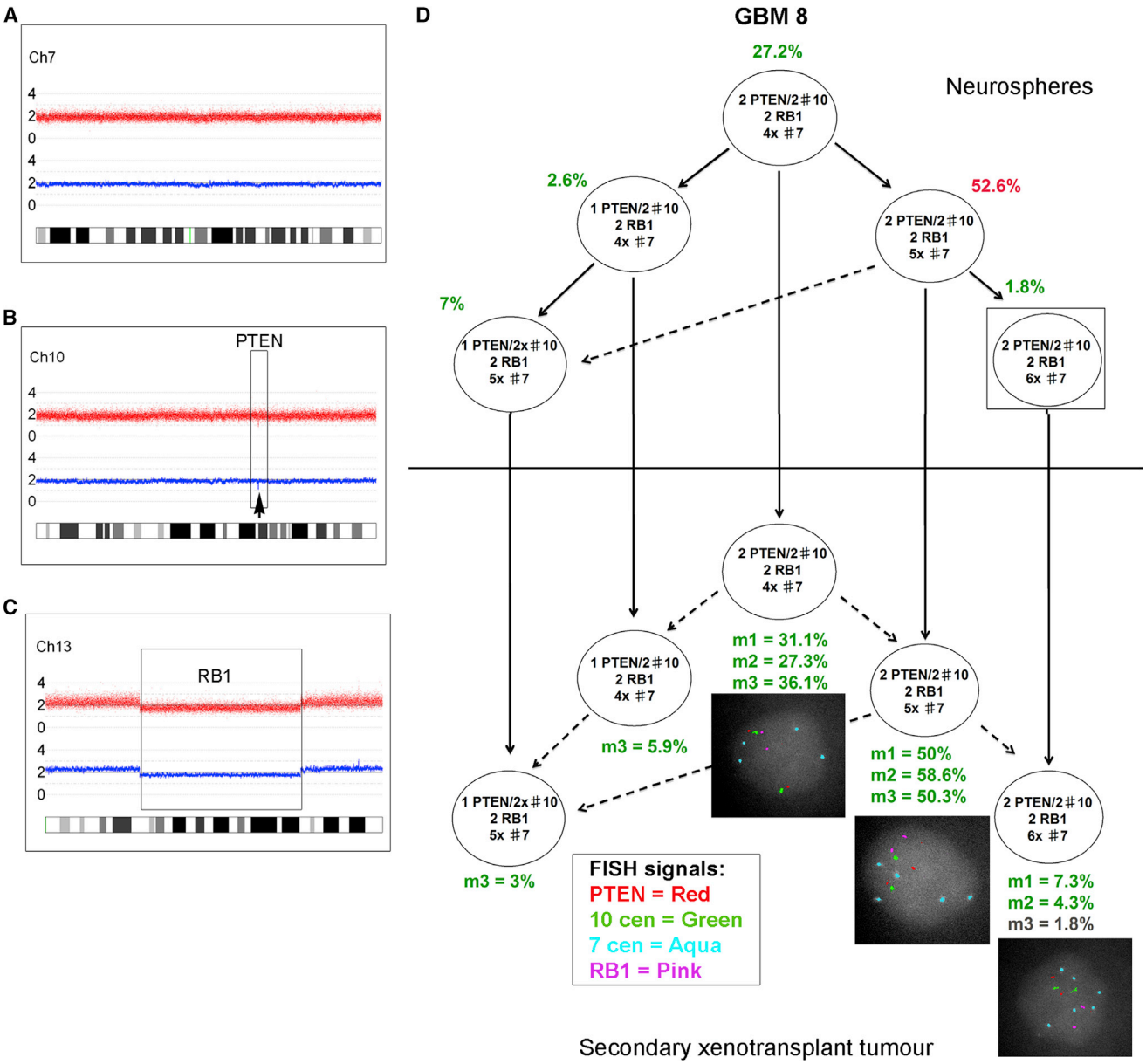


Figure 3. Identical Subclonal Genetic Structure of Neurosphere Cells and Tumor-Propagating Cells Derived from Primary Tumor GBM 8

(A–C) SNP 6 array profiles showing chromosome 7 (A), focal PTEN loss (B), and a large deletion of chromosome 13, including RB1 (C). There was also loss of heterozygosity (LOH) for the whole of 17p (including the TP53 gene) without any copy number change (uniparental disomy [UPD]).

(D) FISH analysis of GBM 8 neurospheres using a range of centromere probes revealed that these were nearly triploid, with two *PTEN* and *RB1* signals corresponding to a loss of one copy of each locus. The neurospheres showed a branched subclonal structure with four subclones above the FISH detection threshold (2%) at the time of injection into primary mice (top). All of these read out in the tumors of at least one secondary transplanted mouse (bottom). One further subclone detected in all mice was present in the neurospheres at a level below the cutoff for FISH (1.8%) (box). m1, m2, m3: three replicate mice, each injected with 1×10^6 neurosphere cells. Solid arrows show the probable derivation of subclones. Dashed arrows indicate the possible alternative derivation of subclones. FISH images were captured at 100× magnification.

See also Figure S1 and Tables S1, S2, S3, and S4.

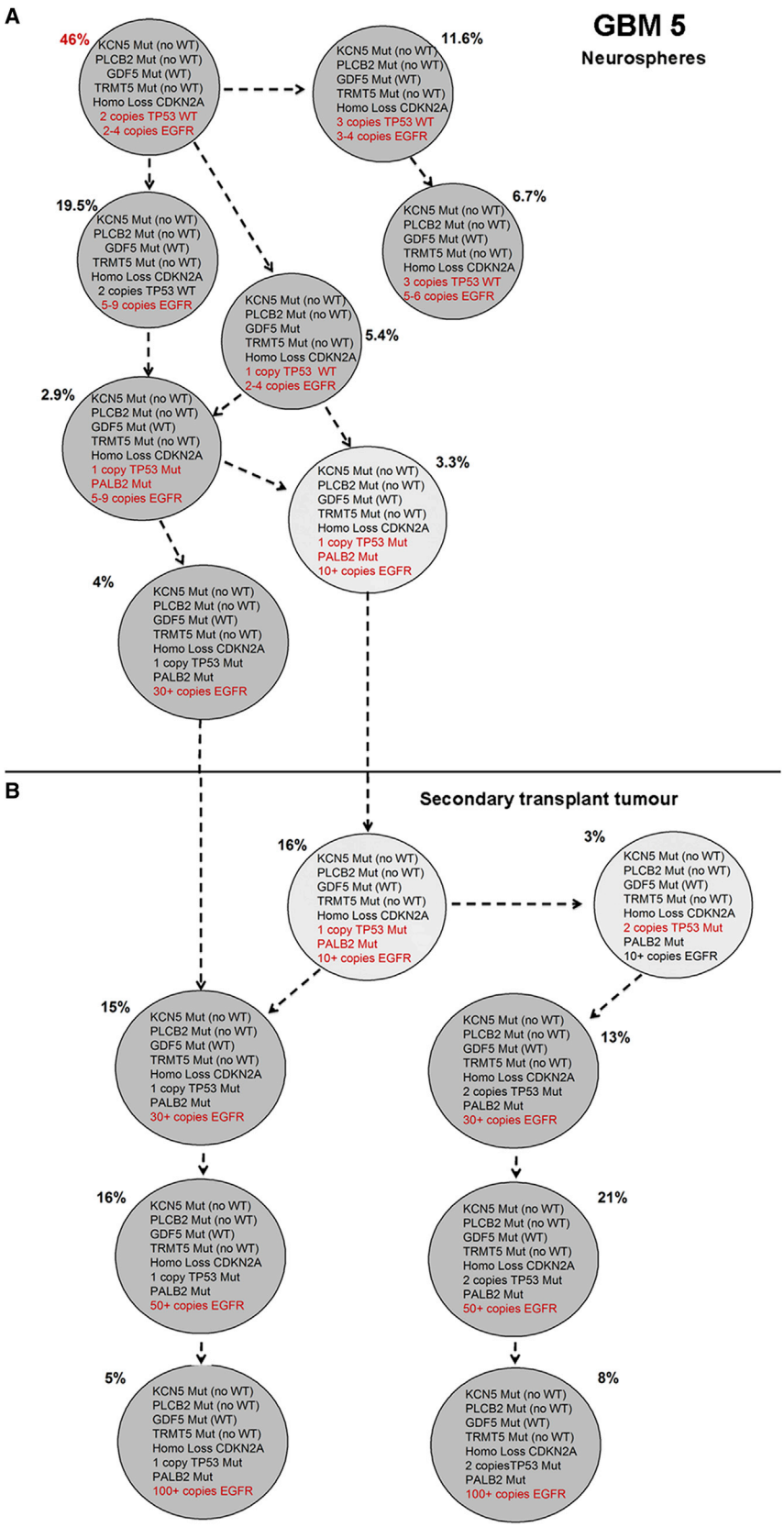


Figure 4. Single-Cell Analysis of Selected Mutations and CNAs Identified by Exome Sequencing in GBM 5

(A and B) Subclonal genetic architecture in neurospheres (A) and tumor-propagating cells derived from GBM 5 after xenotransplantation (B, mouse 2). A total of 240 neurosphere cells and 100 cells from the secondary xenotransplant tumor were evaluated. Mutations and CNAs are given within the circles; additional mutations and CNAs in individual subclones are indicated in red.

See also Figures S2–S4 and Tables S1, S2, S3, and S4.

2012; Snuderl et al., 2011; Sottoriva et al., 2013). These observations are supported by data from glioma susceptibility studies that revealed prominent roles for alterations in *EGFR*, *CDKN2A*, and *TP53* in glioma evolution (Andersson et al., 2010; Egan et al., 2012; Shete et al., 2009; Stacey et al., 2011; Wrensch et al., 2009).

Our data confirm previously unrecognized levels of temporal diversity and complexity in the subclonal landscape of GBM. In all cases analyzed, genetically distinct subclones had variable serial repopulating activity in vivo. We can exclude the possibility that the presence of in vitro aberrations played a role in the clonal architecture of the mouse xenografts, given the culture conditions used and the number of in vitro cell passages. Since the in vivo readout is likely to be a functional activity of self-renewing CSCs, this suggests that the competitive self-renewal ability of tumor-propagating stem cells in GBM varies on the basis of frequency and/or quantitative features (e.g., proliferation rates and growth factor dependence). This is in line with the principle that the extensive replicative potential of CSCs allows subclonal evolution (Greaves and Maley, 2012).

Our interrogation of subclonal genetic diversity of human GBM has revealed that tumor-propagating cells in GBM are genetically heterogeneous and have a variable competitive capacity for tumor propagation in vivo. The link forged among genetic diversity, clonal architecture, and propagating activity in vivo may facilitate the characterization of mutational variants that are responsible for disease recurrence and therapeutic resistance in patients (Johnson et al., 2014).

EXPERIMENTAL PROCEDURES

GBM Sample Collection

Twelve patients diagnosed with primary GBM were administered 5-aminolevulinic acid (Medac UK) 5 hr before surgery as an oral dose of 20 mg/kg as previously described (Piccirillo et al., 2012; Stummer et al., 2006). The tissue collection protocols complied with the UK Human Tissue Act 2004 (HTA license ref. 12315) and were approved by the local regional ethics committee (LREC ref. 04/Q0108/60). Informed consent was obtained from each patient before surgery.

Cell Line Derivation and In Vivo Tumorigenicity

Primary culture and neurosphere cell line derivation were performed as previously described (Fael Al-Mayhani et al., 2009; Piccirillo et al., 2009, 2012). Briefly, primary GBM cells were plated in culture dishes directly after tumor resection from the patient and used for in vivo experiments upon formation of the first neurospheres. To evaluate in vivo tumorigenicity, serial transplantations (two in vivo passages) were performed using immunosuppressed animals. As soon as neurosphere cultures were established, mechanical dissociation to single cells was carried out and 1×10^6 cells/animal were used for intracerebral transplantation into the right striatum of 4-week-old NOD-SCID mice (Charles River), using previously described stereotactic coordinates (Piccirillo et al., 2006, 2009). From the same single-cell suspension, 1×10^6 cells were used for FISH analysis.

In total, 36 animals were injected with cells derived from 12 GBMs (n = 3 animals/GBM). Mice were sacrificed when they became symptomatic according to the Home Office guidelines. Whole mouse brains were removed and tumors were resected as previously described (Galli et al., 2004). The tissue was mechanically disaggregated and reinjected into other 4-week-old NOD-SCID mice using the same stereotactic coordinates (n = 3 animals/GBM, 34 animals in total). When symptoms appeared, the animals were sacrificed and the whole tumors were resected and disaggregated. All of the tumors harvested from mice after the secondary in vivo passage were used in FISH analyses for comparison with the single-cell suspension derived from neurosphere cultures. Disaggregated cells from the primary xenograft tumors were fixed for FISH and stored for comparison if needed. All of the in vivo experiments were performed according to UK Project license approval.

FISH

Single-cell suspensions from dissociated neurosphere cultures or mouse xenograft tumor mice were harvested and fixed in methanol-acetic acid according to standard cytogenetic methods (Horsley et al., 2008). Then, 100–200 nuclei from each cell preparation were analyzed for the presence of the relevant FISH probes signals. Interphase FISH for selected CNAs was carried out as previously described (Anderson et al., 2011) using BAC and fosmid probes for selected genes (the BACPAC Resource Center, Children's Hospital, Oakland Research Institute; http://bacpac.chori.org). Probes were labeled by nick translation with biotin-16-dUTP (Roche Diagnostics), SpectrumGreen (Vysis, Abbott Laboratories), or Cy3-dUTP (Roche Diagnostics), and hybridized in combination as previously described (Anderson et al., 2011). Hybridization and washes were performed according to the Vysis protocol, with a single layer of Cy5-conjugated streptavidin (GE Healthcare) for detection of biotinylated probes. Fluorescent signals were viewed with a Zeiss Axioskop fluorescence microscope equipped with filters for DAPI, fluorescein isothiocyanate/SpectrumGreen, SpectrumOrange, and Cy5. Images were captured and analyzed using a Hamamatsu ORCA-ER CCD camera and SmartCapture X software (Digital Scientific).

Whole-Exome Sequencing

Genomic DNA was subjected to whole-exome sequencing (Oxford Gene Technology). Exome capture was performed using the SureSelectXT Human All Exon v4 kit (Agilent) according to manufacturer's instructions and sequenced with Illumina paired-end sequencing (protocol v1.2).

Single-Cell Analysis

Single-cell sorting, qPCR, and analysis were all performed essentially as previously described (Potter et al., 2013). Briefly, single cells were sorted on a BDFACSAria1-SORP instrument (BD) directly into lysis buffer. Specific (DNA) targeted amplification was then

performed prior to qPCR. Single-cell, target-amplified DNA was interrogated by qPCR for each DNA target of interest using the 96.96 dynamic microfluidic array and the BioMark HD system (Fluidigm) as recommended by the manufacturer.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2014.11.003.

AUTHOR CONTRIBUTIONS

M.G., C.W., L.K., and S.G.M.P. conceived and designed the study. S.G.M.P., S.C., L.K., N.E.P., F.W.v.D., S.L., and M.-J.C. carried out the work, including development of the methodology. M.G., C.W., L.K., S.G.M.P., N.E.P., and F.W.v.D. analyzed and interpreted the results. L.K., S.G.M.P., M.G., C.W., N.E.P., and F.W.v.D. were involved in writing, reviewing, and/or revising the manuscript. M.G., C.W., and L.K. supervised the study.

ACKNOWLEDGMENTS

We thank Oxford Gene Technology for providing the method for whole-exome sequencing. This study was supported by the National Institute for Health Research, the Cambridge Biomedical Research Centre, the Higher Education Funding Council for England, the Royal College of Surgeons of Edinburgh, The Brain Tumour Charity, the European Commission-Seventh Framework Programme (Marie Curie Intra-European Fellowship to S.G.M.P.), and the Wellcome Trust (Strategic Award to M.G.).

Received: May 20, 2014
Revised: November 17, 2014
Accepted: November 18, 2014
Published: December 18, 2014

REFERENCES

Anderson, K., Lutz, C., van Delft, F.W., Bateman, C.M., Guo, Y., Colman, S.M., Kempski, H., Moorman, A.V., Titley, I., Swansbury, J., et al. (2011). Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature *469*, 356–361.

Andersson, U., Schwartzbaum, J., Wiklund, F., Sjöström, S., Liu, Y., Tsavachidis, S., Ahlbom, A., Auvinen, A., Collatz-Laier, H., Feychting, M., et al. (2010). A comprehensive study of the association between the EGFR and ERBB2 genes and glioma risk. Acta Oncol. *49*, 767–775.

Bao, S., Wu, Q., Sathornsumetee, S., Hao, Y., Li, Z., Hjelmeland, A.B., Shi, Q., McLendon, R.E., Bigner, D.D., and Rich, J.N. (2006). Stem cell-like glioma cells promote tumor angiogenesis through vascular endothelial growth factor. Cancer Res. *66*, 7843–7848.

Bonavia, R., Inda, M.M., Cavenee, W.K., and Furnari, F.B. (2011). Heterogeneity maintenance in glioblastoma: a social network. Cancer Res. *71*, 4055–4060.

Chen, R., Nishimura, M.C., Bumbaca, S.M., Kharbanda, S., Forrest, W.F., Kasman, I.M., Greve, J.M., Soriano, R.H., Gilmour, L.L., Rivers, C.S., et al. (2010). A hierarchy of self-renewing tumor-initiating cell types in glioblastoma. Cancer Cell *17*, 362–375.

Chen, J., Li, Y., Yu, T.S., McKay, R.M., Burns, D.K., Kernie, S.G., and Parada, L.F. (2012). A restricted cell population propagates glioblastoma growth after chemotherapy. Nature *488*, 522–526.

Dick, J.E., Bhatia, M., Gan, O., Kapp, U., and Wang, J.C. (1997). Assay of human stem cells by repopulation of NOD/SCID mice. Stem Cells *15* (*Suppl 1*), 199–203, discussion 204–207.

Egan, K.M., Nabors, L.B., Olson, J.J., Monteiro, A.N., Browning, J.E., Madden, M.H., and Thompson, R.C. (2012). Rare TP53 genetic variant associated with glioma risk and outcome. J. Med. Genet. *49*, 420–421.

Fael Al-Mayhani, T.M., Ball, S.L., Zhao, J.W., Fawcett, J., Ichimura, K., Collins, P.V., and Watts, C. (2009). An efficient method for derivation and propagation of glioblastoma cell lines that conserves the molecular profile of their original tumours. J. Neurosci. Methods *176*, 192–199.

Galli, R., Binda, E., Orfanelli, U., Cipelletti, B., Gritti, A., De Vitis, S., Fiocco, R., Foroni, C., Dimeco, F., and Vescovi, A. (2004). Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. Cancer Res. *64*, 7011–7021.

Goodenberger, M.L., and Jenkins, R.B. (2012). Genetics of adult glioma. Cancer Genet. *205*, 613–621.

Greaves, M. (2013). Cancer stem cells as 'units of selection'. Evol. Appl. *6*, 102–108.

Greaves, M., and Maley, C.C. (2012). Clonal evolution in cancer. Nature *481*, 306–313.

Horsley, S.W., Colman, S., McKinley, M., Bateman, C.M., Jenney, M., Chaplin, T., Young, B.D., Greaves, M., and Kearney, L. (2008). Genetic lesions in a preleukemic aplasia phase in a child with acute lymphoblastic leukemia. Genes Chromosomes Cancer *47*, 333–340.

Johnson, B.E., Mazor, T., Hong, C., Barnes, M., Aihara, K., McLean, C.Y., Fouse, S.D., Yamamoto, S., Ueda, H., Tatsuno, K., et al. (2014). Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science *343*, 189–193.

Jung, V., Romeike, B.F., Henn, W., Feiden, W., Moringlane, J.R., Zang, K.D., and Urbschat, S. (1999). Evidence of focal genetic microheterogeneity in glioblastoma multiforme by area-specific CGH on microdissected tumor cells. J. Neuropathol. Exp. Neurol. *58*, 993–999.

Kreso, A., and Dick, J.E. (2014). Evolution of the cancer stem cell model. Cell Stem Cell *14*, 275–291.

Little, S.E., Popov, S., Jury, A., Bax, D.A., Doey, L., Al-Sarraj, S., Jurgensmeier, J.M., and Jones, C. (2012). Receptor tyrosine kinase genes amplified in glioblastoma exhibit a mutual exclusivity in variable proportions reflective of individual tumor heterogeneity. Cancer Res. *72*, 1614–1620.

Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? Nat. Rev. Cancer *12*, 323–334.

Nicholas, M.K. (2007). Glioblastoma multiforme: evidence-based approach to therapy. Expert Rev. Anticancer Ther. *7* (12, Suppl), S23–S27.

Nickel, G.C., Barnholtz-Sloan, J., Gould, M.P., McMahon, S., Cohen, A., Adams, M.D., Guda, K., Cohen, M., Sloan, A.E., and La-Framboise, T. (2012). Characterizing mutational heterogeneity in a glioblastoma patient with double recurrence. PLoS ONE *7*, e35262.

Notta, F., Mullighan, C.G., Wang, J.C., Poeppl, A., Doulatov, S., Phillips, L.A., Ma, J., Minden, M.D., Downing, J.R., and Dick, J.E. (2011). Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. Nature *469*, 362–367.

Nowell, P.C. (1976). The clonal evolution of tumor cell populations. Science *194*, 23–28.

Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396–1401.

Piccirillo, S.G., Reynolds, B.A., Zanetti, N., Lamorte, G., Binda, E., Broggi, G., Brem, H., Olivi, A., Dimeco, F., and Vescovi, A.L. (2006). Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. Nature *444*, 761–765.

Piccirillo, S.G., Combi, R., Cajola, L., Patrizi, A., Redaelli, S., Bentivegna, A., Baronchelli, S., Maira, G., Pollo, B., Mangiola, A., et al. (2009). Distinct pools of cancer stem-like cells coexist within human glioblastomas and display different tumorigenicity and independent genomic evolution. Oncogene *28*, 1807–1811.

Piccirillo, S.G., Dietz, S., Madhu, B., Griffiths, J., Price, S.J., Collins, V.P., and Watts, C. (2012). Fluorescence-guided surgical sampling of glioblastoma identifies phenotypically distinct tumour-initiating cell populations in the tumour mass and margin. Br. J. Cancer *107*, 462–468.

Potter, N.E., Ermini, L., Papaemmanuil, E., Cazzaniga, G., Vijayaraghavan, G., Titley, I., Ford, A., Campbell, P., Kearney, L., and Greaves, M. (2013). Single-cell mutational profiling and clonal phylogeny in cancer. Genome Res. *23*, 2115–2125.

Shackleton, M., Quintana, E., Fearon, E.R., and Morrison, S.J. (2009). Heterogeneity in cancer: cancer stem cells versus clonal evolution. Cell *138*, 822–829.

Shapiro, J.R., Yung, W.K., and Shapiro, W.R. (1981). Isolation, karyotype, and clonal growth of heterogeneous subpopulations of human malignant gliomas. Cancer Res. *41*, 2349–2359.

Shete, S., Hosking, F.J., Robertson, L.B., Dobbins, S.E., Sanson, M., Malmer, B., Simon, M., Marie, Y., Boisselier, B., Delattre, J.Y., et al. (2009). Genome-wide association study identifies five susceptibility loci for glioma. Nat. Genet. *41*, 899–904.

Snuderl, M., Fazlollahi, L., Le, L.P., Nitta, M., Zhelyazkova, B.H., Davidson, C.J., Akhavanfard, S., Cahill, D.P., Aldape, K.D., Betensky, R.A., et al. (2011). Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. Cancer Cell *20*, 810–817.

Sottoriva, A., Spiteri, I., Piccirillo, S.G., Touloumis, A., Collins, V.P., Marioni, J.C., Curtis, C., Watts, C., and Tavaré, S. (2013). Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc. Natl. Acad. Sci. USA *110*, 4009–4014.

Stacey, S.N., Sulem, P., Jonasdottir, A., Masson, G., Gudmundsson, J., Gudbjartsson, D.F., Magnusson, O.T., Gudjonsson, S.A., Sigurgeirsson, B., Thorisdottir, K., et al.; Swedish Low-risk Colorectal Cancer Study Group. (2011). A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. Nat. Genet. *43*, 1098–1103.

Stummer, W., Pichlmeier, U., Meinel, T., Wiestler, O.D., Zanella, F., and Reulen, H.J.; ALA-Glioma Study Group. (2006). Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. Lancet Oncol. *7*, 392–401.

Szerlip, N.J., Pedraza, A., Chakravarty, D., Azim, M., McGuire, J., Fang, Y., Ozawa, T., Holland, E.C., Huse, J.T., Jhanwar, S., et al. (2012). Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. Proc. Natl. Acad. Sci. USA *109*, 3041–3046.

Wrensch, M., Jenkins, R.B., Chang, J.S., Yeh, R.F., Xiao, Y., Decker, P.A., Ballman, K.V., Berger, M., Buckner, J.C., Chang, S., et al. (2009). Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. Nat. Genet. *41*, 905–908.

Stem Cell Reports, Volume *4*

Supplemental Information

Genetic and Functional Diversity of Propagating Cells in Glioblastoma

Sara G.M. Piccirillo, Sue Colman, Nicola E. Potter, Frederik W. van Delft, Suzanne Lillis, Maria-Jose Carnicer, Lyndal Kearney, Colin Watts, and Mel Greaves

SUPPLEMENTARY FIGURES

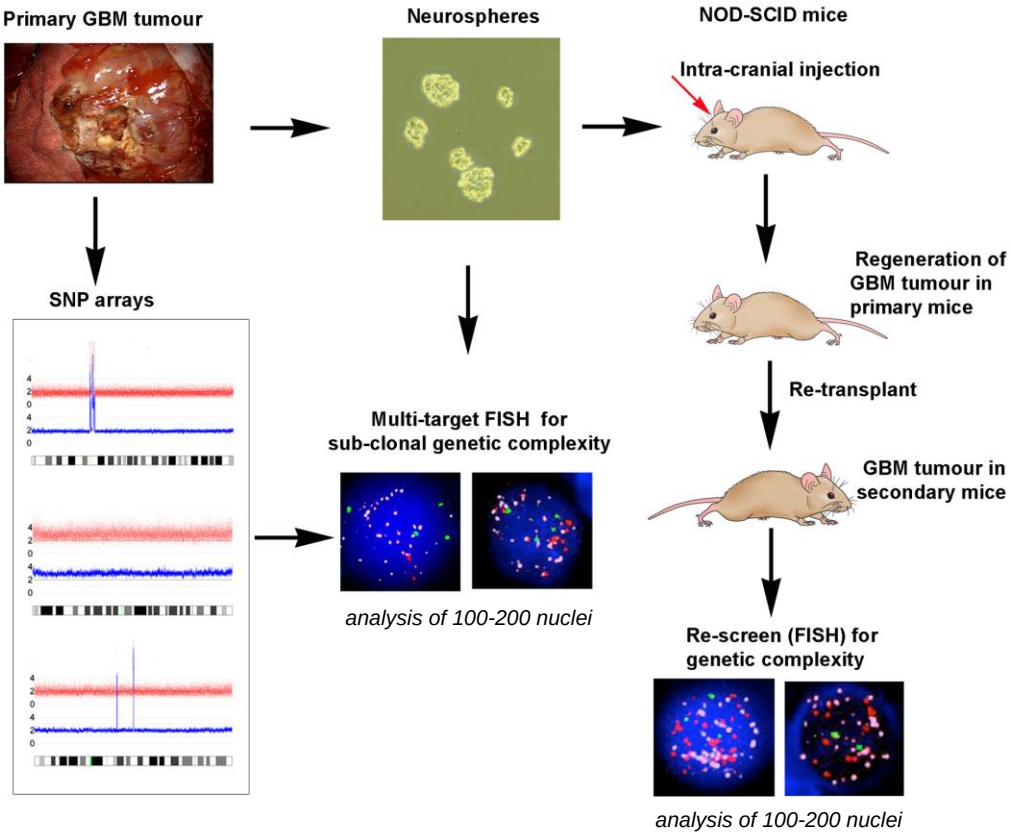


Figure S1, related to Figure 1-3. Schematic overview of the study. Primary GBM tumour was used to extract DNA for high resolution SNP (SNP6) array analysis and also to establish stable neurosphere cell cultures. High-resolution (SNP 6) arrays were performed on DNA extracted from the primary GBM tumour to identify 'driver' CNA, defined as recurrent regions of amplification or deletion. FISH probes were designed for these regions and 3 colour FISH carried out to neurosphere cells (dissociated to single cells and prepared by standard cytogenetic protocols). Neurospheres were dissociated into single cells and used for intra-cerebral transplantation (and re-transplantation) into NOD-SCID mice. FISH was carried out to the reconstituted tumours in secondary transplanted mice using the same probes used for the original FISH screen of the neurospheres. We then compared the sub-clonal genetic structure in the neurospheres and tumours generated in the mice.

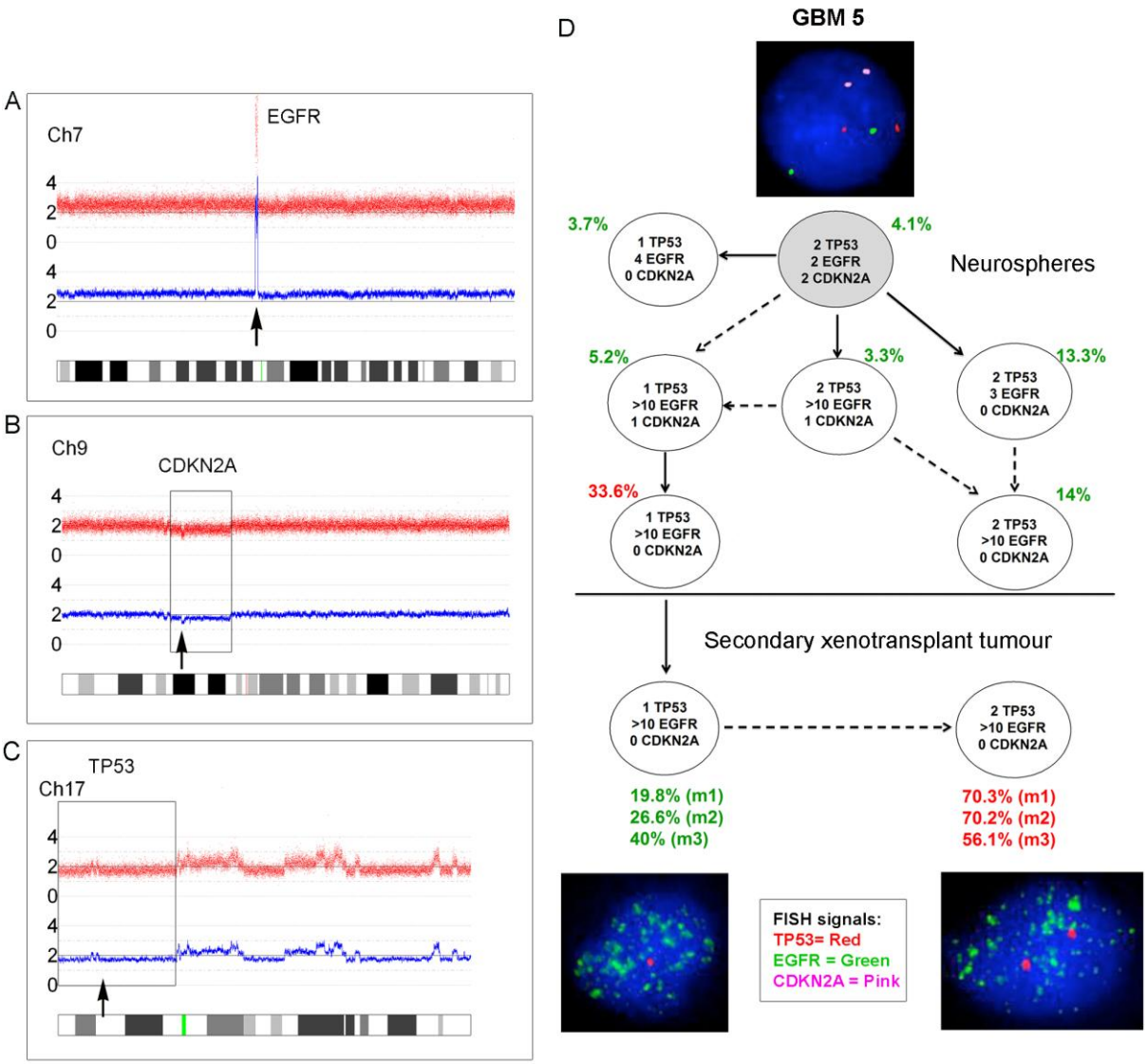


Figure S2, related to Figure 4. Sub-clonal genetic structure of neurosphere cells and tumour propagating cells derived from primary glioblastoma GBM 5 by multicolour FISH. (A-C) SNP 6 array profiles showing high level *EGFR* amplification (A), homozygous *CDKN2A* loss comprising a large deletion of one allele and a focal deletion of the second allele (B) and *TP53* loss due to a large 17p deletion (C): in each case the gene location is indicated by an arrow and the larger deletion by a box. (D) Clonal structure of neurosphere cells (top) and tumour cells after secondary transplant into NOD-SCID mice (bottom). FISH images are shown next to their respective genotype. Red type indicates the major clone; m1, m2, m3 = three replicate mice each injected with 1 x 10^6 neurosphere cells. The percentages of sub-clones with 1 or 2 or 3 copies of *TP53* determined by single cell analysis do not correspond to those determined by FISH (Figure 4). This is most likely due to the differences in probes used for the two methods. Solid arrows show probable derivation of sub-clones. Dashed arrows indicate possible alternative derivation of sub-clones. FISH images were captured at 100x magnification.

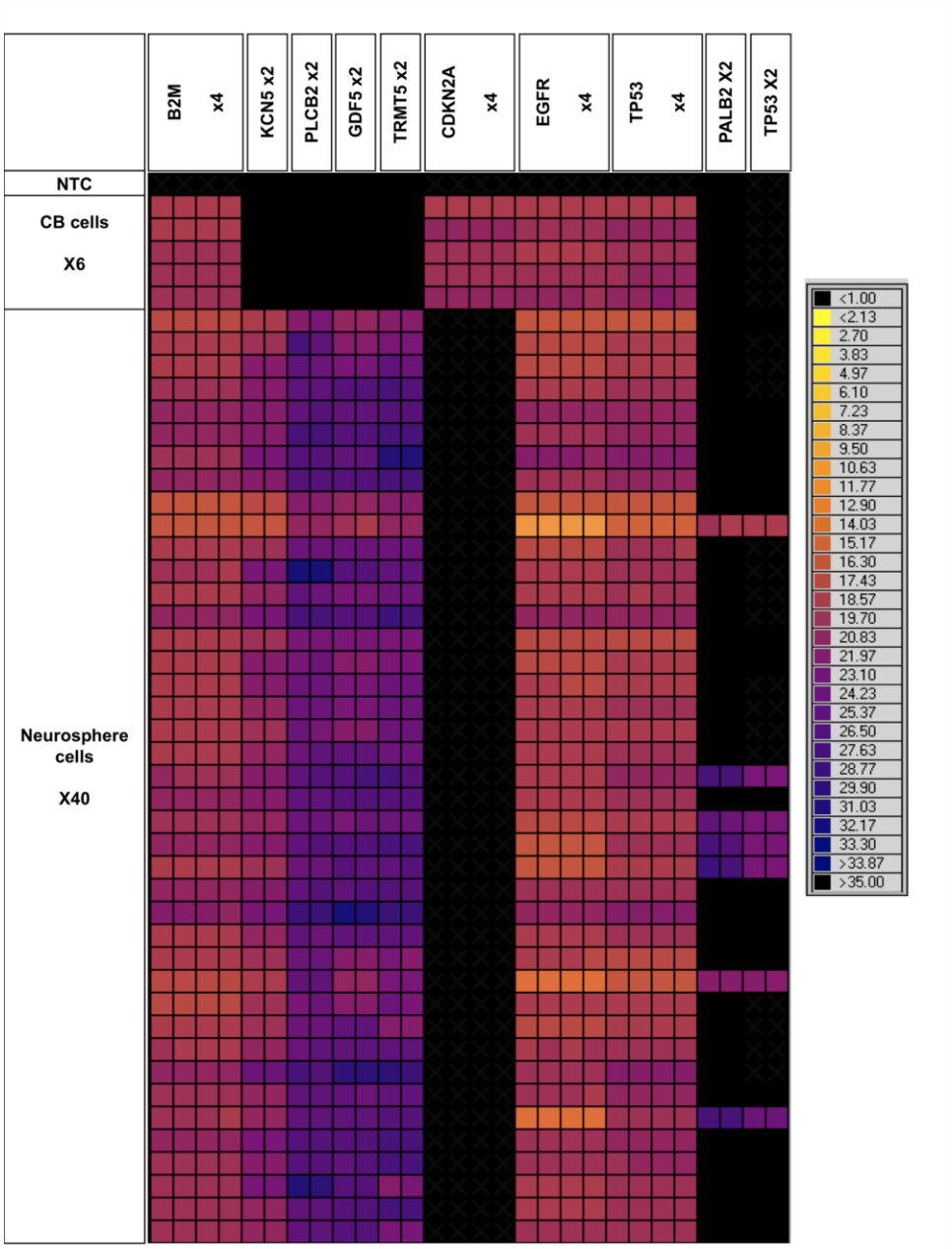


Figure S3, related to Figure 4. Heatmap depicting an example of raw Q-PCR data from the BioMark® HD. The rows represent single cells including six cord blood mononuclear cells and 40 neurosphere cells from GBM 5. The columns represent assays, each completed in quadruplicate including *β2M* (one of three assays). Mutations assessed were *KCN5*, *PLCB2 GDF5*, *TRMT5*, *PALB2* and *TP53*. Copy number alterations were *CDKN2A*, *EGFR* and *TP53*. The coloured boxes at the junction of a row and column indicate the raw C_T value (according to the key on the right) obtained for a Q-PCR reaction involving the indicated cell and assay. Assays targeting a mutation provide a definitive positive or negative result indicating the presence or absence respectively of an alteration (black = no mutation). The DNA copy number assays provide a raw C_T value which requires further analysis (standard $\Delta\Delta C_T$ method Applied Biosystems®) to attribute a DNA copy number to the target gene of interest for a single cell.

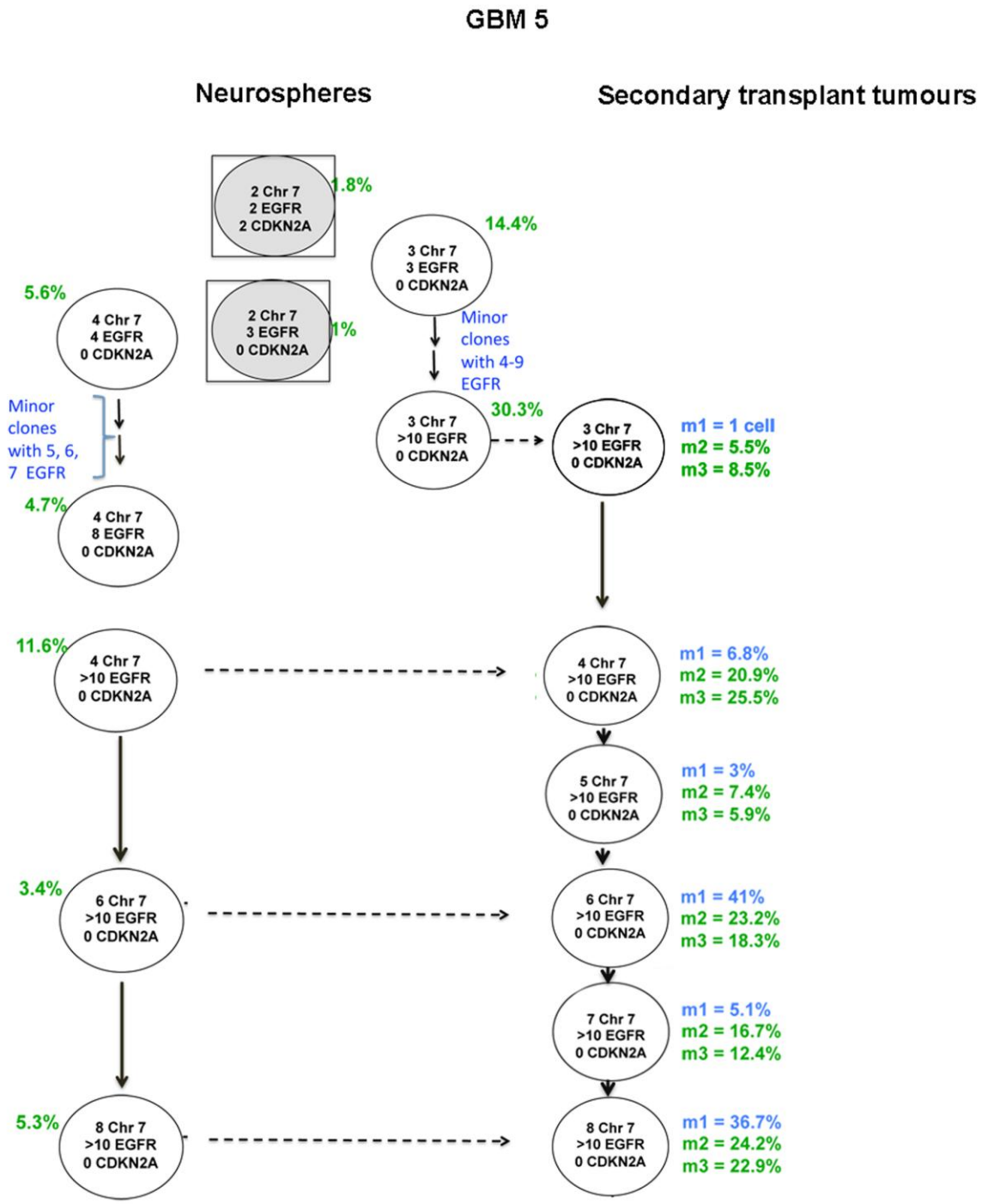


Figure S4, related to Figure 4. Sub-clonal genetic structure of GBM 5 neurospheres and secondary xenotransplant tumours using FISH probes for chromosome 7 copy number, *EGFR* amplification and *CDKN2A* loss. Stepwise clonal evolution with progressive gain of copies of chromosome 7 and EGFR amplification was observed in the neurospheres and in the tumours of secondary transplanted mice. EGFR amplification uncoupled from chromosome copy number at 3-4 copies of chromosome 7: thereafter EGFR amplification increased to very high levels, consistent with the formation of extrachromosomal double minutes (dmin). m1, m2, m3 = three replicate mice each injected with 1 x 10^6 neurosphere cells. Solid arrows show probable derivation of sub-clones. Dashed arrows indicate possible alternative derivation of sub-clones.

SUPPLEMENTAL TABLES

Table S1, related to Figure 1-4. Clinical information of patients.

Sample	Age at diagnosis (y)	Sex	Performance status	Treatment post-surgery
GBM 1	72	F	Not documented	SCRT
GBM 2	61	M	0	SCRT
GBM 3	61	M	0	SCRT
GBM 4	61	M	3	SCRT
GBM 5	52	M	0	ChemoRT
GBM 6	62	M	0	SCRT
GBM 7	63	F	1	ChemoRT
GBM 8	61	M	0	ChemoRT
GBM 9	63	F	1	SCRT
GBM 10	46	F	1	SCRT
GBM 11	64	F	0	ChemoRT
GBM 12	36	M	0	ChemoRT

We provide here clinical information of 12 GBM patients. The samples obtained from these patients were used in the study. These patients were prospectively recruited through the Neuro-oncology multidisciplinary team (MDT). All patients had suspected high-grade glioma on imaging review by a consultant neuroradiologist at the MDT. This study included 7 men and 5 women. Mean age was 58.5 years at diagnosis (range 36-72). WHO performance status before surgery was 0-3. Treatment post-surgery was either short course of radio-therapy (SCRT) or chemo- and radio-therapy followed by chemotherapy (ChemoRT).

Table S2, related to Figure 1-4. SNP 6 analysis of GBM tumours.
The table summarizing the SNP 6 data of all the GBM tumours is provided as Excel file.

Table S3, related to Figure 1-4. Details of tumours in mice and whole-exome sequencing results of GBM 5 tumour.

	DAYS TO TUMOUR FORMATION					
	Primary transplant			Secondary transplant		
	mouse1	mouse 2	mouse 3	mouse1	mouse 2	mouse 3
GBM 1	159	164	164	107	133	119
GBM 2	149	164	164	43	133	133
GBM 3	163	163	163	121	121	121
GBM 4	124	124	133	77	96	166
GBM 5	147	149	149	92	69	69
GBM 6	65	65	65	28	119	107
GBM 7	119	119	119	36	149	134
GBM 8	105	113	105	71	50	52
GBM 9	120	120	120	35	35	29
GBM 10	159	189	210	n.t.f.	n.t.f.	n.t.f.
GBM 11	138	145	145	93	54	93
GBM 12	218	n.t.f.	n.t.f.	n.t.f.	n.i.	n.i.

This table summarizes the data from the *in vivo* transplantation in immunosuppressed animals. Two out of twelve GBM failed to form tumour either after the first transplantation (GBM 12, two mice out of three) or after the second transplantation (GBM 10 and GBM 12). Although we observed a trend towards shorter time to tumour formation in the secondary transplant compared to the primary xenograft tumour, a statistically significant difference was observed only in GBM 1, 5, 8, 9, 11 ($p < 0.05$). “n.t.f.”= no tumour formation, “n.i.”= not injected.

Chr.	SNV position	CNV Ref->Seq	Gene	Estimated allele burden (%)
1	29379636	G -> T	*EPB41*	28
1	57258326	G -> C	*C1orf168*	20
2	27549589	G -> A	*GTF3C2*	19
2	180348097	G -> A	*ZNF385B*	18
2	233321642	G -> A	*ALPI*	19
3	110611201	C -> A	*RP11-553A10.1*	24
6	49494443	G -> A	*GLYATL3*	27
7	16566662	C -> T	*LRRC72*	30
7	135387597	T -> C	*SLC13A4*	25
10	5966417	C -> A	*FBXO18*	22
10	79782048	G -> A	*POLR3A*	33
10	90427142	G -> C	*LIPF*	30
10	124339169	G -> T	*DMBT1*	25
11	55587748	G -> A	*OR5D18*	19
11	56128024	T -> A	*OR8J1*	22
14	**61446242**	**C -> T**	***TRMT5***	**28**
14	**63269168**	**T -> A**	***KCNH5***	**32**
15	**40591062**	**G -> A**	***PLCB2***	**24**
15	79339105	G -> T	*RASGRF1*	20
15	79748827	G -> A	*KIAA1024*	17
16	**23635396**	**A -> G**	***PALB2***	**22**
16	81934351	C -> T	*PLCG2*	24
16	82033438	G -> A	*SDR42E1*	19
17	**7578476**	**G -> A**	***TP53***	**20**
17	61623193	C -> T	*KCNH6*	36
19	8577520	C -> T	*ZNF414*	17
19	35832277	C -> G	*CD22*	17
19	36211360	G -> A	*MLL4*	20
19	45899676	C -> T	*PPP1R13L*	22
20	**34022533**	**C -> T**	***GDF5***	**15**
8	105263344	C -> T	*RIMS2*	15
9	100074426	C -> T	*C9orf174*	17

Mutation targets selected for single cell interrogation are in bold.
An allele burden of 50% indicates either a heterozygous mutation in every cell or a homozygous mutation in 25% of cells.

Table S4, related to Figure 1-4. SNP Array 6.0 call rates and primer sequences for CE-SSCA and Sequencing Analysis of *TP53*.

File ID	Gender	Call Rate (%)	Het Rate	Hom Rate
GL1.CEL	female	97.14651	27.00407	70.14243
GBM1.CEL	female	96.41983	22.32763	74.09221
GL2.CEL	male	97.09704	26.2055	70.89154
GBM2.CEL	male	97.57921	24.14025	73.43897
GL3.CEL	male	98.08635	26.00916	72.07719
GBM3.CEL	male	93.94089	28.9871	64.95379
GL4.CEL	male	97.65738	25.63307	72.02431
GBM4.CEL	male	95.93447	22.80398	73.13049
GL5.CEL	male	97.94585	26.28663	71.65922
GBM5.CEL	male	97.1453	25.80709	71.3382
GL6.CEL	male	95.51979	27.43678	68.08301
GBM6.CEL	male	94.57115	26.1391	68.43205
GL7.CEL	female	98.22553	27.00704	71.21848
GBM7.CEL	female	96.77086	27.22076	69.5501
GL8.CEL	male	98.17408	26.10953	72.06455
GBM8.CEL	male	96.87013	23.56089	73.30924
GL9.CEL	female	96.93477	27.78198	69.15279
GBM9.CEL	female	98.233	26.70879	71.52422
GL10.CEL	female	98.03644	26.57181	71.46463
GBM10.CEL	female	97.43432	25.67594	71.75838
GL11.CEL	female	98.33898	26.81795	71.52103
GBM11.CEL	female	96.73733	25.23356	71.50377
GL12.CEL	male	98.17144	25.81842	72.35302
GBM12.CEL	male	97.95739	24.96729	72.9901

GL, Germ Line; GBM, Glioblastoma tumour
Hom, homozygosity; Het, heterozygosity

Primer Name	Primer sequence (5' to 3') Labelled and unlabelled primers	Label
***TP53* ex4**	TP53 ex4 Fw GACCTGGTCCTCTGACTGCT TP53 ex4 Rv GCATTGAAGTCTCATGGAAG	**FAM**
***TP53* ex5**	TP53ex5 Fw TGTCGGCTGACTTTCAACTCT TP53ex5Rv GGCAACCAGCCCTGTCGT	**FAM**
***TP53* ex6**	TP53 ex6 Fw GCTGGGGCTGGAGAGACGA TP53 ex6 Rv CTGGAGGGCCACTGACAAC	**VIC**
***TP53* ex7**	TP53 ex7 Fw GGTCTCCCCAAGGCCCACTG TP53 ex7 Rv GGGGATGTGATGAGAGGTGGAT	**NED**
***TP53* ex8**	TP53 ex8 Fw GCCTCTTGCTTCTCTTTTCCTATC TP53 ex8 Rv GGGAGAGGAGCTGGTGTTGTT	**PET**
***TP53* ex9**	TP53 ex9 Fw AGCAGGACAAGAAGCGGTGG TP53 ex9 Rv AACGGCATTTTGAGTGTTAGACTG	**FAM**

CE-SSCA uses fluorescently labelled primers. The same primer sequences but with a universal tag instead of a fluorescent tag were used for PCR set up of individual exons prior to sequencing.

Universal tag:

Universal20-Fw	GTT GTA AAA CGA CGG CCA GT
Universal20-Rv	CAC AGG AAA CAG CTA TGA CC

***TP53* CE-SSCA PCR cycling conditions**

95 ^{o}C 7mins
94 ^{o}C 30 secs
60 ^{o}C 30 secs } x40
72 ^{o}C 30secs
72 ^{o}C 7 mins
25 ^{o}C ∞

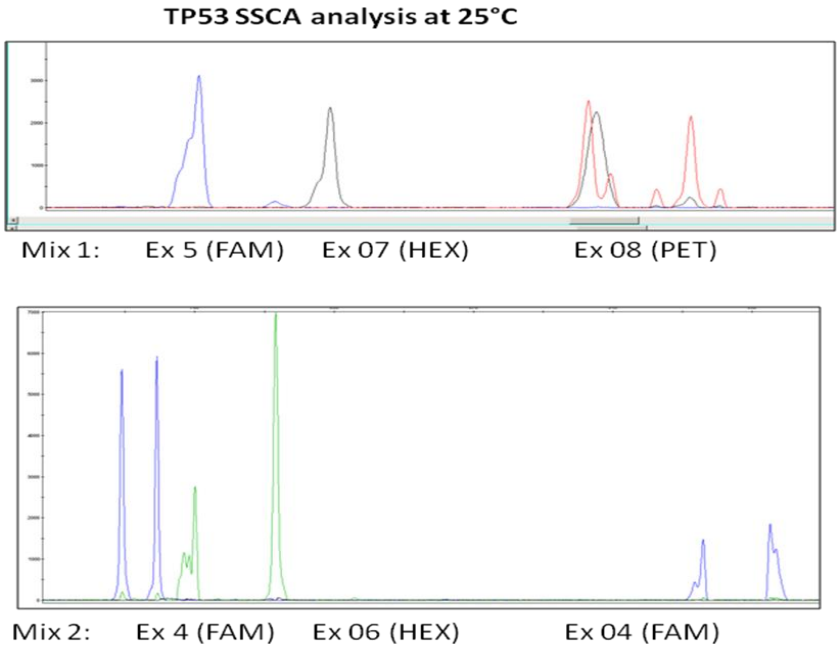
BIGDYE FAST sequencing conditions

96 ^{O}C 1 min x1
96 ^{O}C 10 secs
50 ^{O}C 5 secs x 25
60 ^{O}C 1 min 15 secs
4 ^{O}C ∞

Temperatures for CE-SSCA

Temperature 1: 20 ^{O}C
Temperature 2: 25 ^{O}C
Temperature 3: 30 ^{O}C
Temperature 4: 33 ^{O}C
Temperature 5: 35 ^{O}C

TP53 schematic of SSCA results



SUPPLEMENTAL EXPERIMENTAL PROCEDURES

DNA extraction

DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen) from the same GBMs used for cell line derivation after chopping and mixing of the tissue. Germline DNA was extracted with the same kit using blood or buffy coat. DNA quantification was performed using the picogreen assay for single-nucleotide polymorphism (SNP) analysis.

Copy number and LOH analysis

Samples were analyzed with Partek Genomics Suite 6.6 and CNAG 3.3.0.1 (beta) (http://plaza.umin.ac.jp/genome/) with the use of paired tumor (test) samples with the self-reference control (reference) samples to determine copy number and LOH caused by imbalance (Nannya et al., 2005). The position of regions of LOH and gain were identified with the University of California Santa Cruz Genome Browser (Hg18, March 2006 Assembly; http://genome.ucsc.edu/cgi-bin/hgGateway).

Genome mapping analysis

Genome-Wide Human SNP Array 6.0 (Affymetrix) mapping analysis was performed with 500 ng of tumour and germline DNA from each patient essentially as previously described. Briefly, genomic DNA was digested in parallel with restriction endonucleases NspI and StyI, ligated to an adaptor, and subjected to PCR amplification with adaptor-specific primers. The PCR products were digested with DNaseI and labeled with a biotinylated nucleotide analog. The labeled DNA fragments were hybridized to the microarray, stained by streptavidin-phycoerythrin conjugates, washed with the Affymetrix Fluidics Station 450, and then scanned with a GeneChip scanner 3000 7G. The CEL files are deposited with Gene Expression Omnibus, GEO Accession no. GSE45185. SNP genotypes were obtained with the use of the BRLMM algorithm in Affymetrix Genotyping Console 4.0 software (see Table S4 for the SNP Array 6.0 call rates).

Fluorescence *in situ* hybridization

In each case, 100 - 200 nuclei were scored for the presence of the relevant probe signals. Depending on the probe, amplified gene signals were scored manually up to 5-9 signals: high-level amplifications were recorded as > 10 signals. Cut-off levels for false loss of gene signals were established using normal peripheral blood control slides as previously described (Anderson et al., 2011). For *CDKN2A*, a threshold cut-off was set at 2% for loss of a single signal and 1% for loss of two signals. Clonal architectures were established by direct scrutiny of the degree of similarity of genotypes, with the assumption that acquisition of CNA is irreversible and genotypes change unidirectionally.

Detection of *TP53* Mutations by Capillary Electrophoresis Single-Strand Conformation Analysis

We used a capillary electrophoresis single-strand conformation analysis (CE-SSCA) method to detect mutations in exons 4-9 of the *TP53* gene. Any mutations identified were further characterised by Sanger sequencing.

We used a capillary electrophoresis single-strand conformation analysis (CE-SSCA) approach for identification of mutations in exons 4-9 of the *TP53* gene. CE-SSCA involves PCR amplification using fluorescently labelled primers followed by electrophoresis through a non-denaturing capillary across a range of temperatures to increase sensitivity. *TP53* exons 5, 7 and 8 are amplified in a PCR multiplex and each primer pair is labelled with a different fluorochome (Table S4). The PCRs for exons 4, 6 and 9 are set up individually (using of two different fluorochomes) and then multiplexed prior to CE-SSCA electrophoresis on an ABI 3130xl Genetic Analyser.

The temperatures used for electrophoresis on the ABI 3130xl were 20, 25, 30, 33 and 35°C. The CE-SSCA data was analysed using Genemapper Software. The PCR product for each exon produces a specific pattern of peaks at each different temperature. The pattern for the wild type sample at each temperature was used as a reference for

analysis of patient samples. Any mobility shift identified by CE-SSCA was subsequently further tested by Sanger sequencing to elucidate the precise mutation.

For Sanger sequencing, a second PCR was performed for the specific exon (exon 5 for GBM 5 and exon 7 for GBM 8). The primer sequences used for this PCR are the same as the CE-SSCA but instead of a fluorescent label the primers are universally tagged. The PCR products were then cleaned-up using Ampure (supplied by Agencourt) and sequenced using BigDye Terminator v3.1. Ethanol, EDTA and Sodium Acetate precipitation was then performed prior to running the samples on the ABI 3730 capillary sequencer. Sequencing data was analysed using Mutation Surveyor software.

Whole-exome sequencing

Briefly, DNA was sheared by fragmentation (Covaris, Woburn, MA, USA) and purified using Agencourt AMPure XP beads (Beckman Coulter, Fullerton, CA, USA). The resulting fragments were analysed using an Agilent 2100 Bioanalyzer. Fragment ends were repaired and adaptors were ligated to the fragments. The library was purified using Agencourt AMPure XP beads and amplified by PCR before hybridisation with biotinylated RNA baits. Bound genomic DNA was purified with streptavidin coated magnetic Dynabeads (Invitrogen, Carlsbad, CA, USA) and re-amplified to include barcoding tags before pooling for sequencing on an paired-end, 100 cycle run on an Illumina HiSeq 2000 according to manufacturer's protocols. Somatic SNPs were identified using SomaticSniper (v1.0) (Larson et al., 2012) and were restricted to being called in regions known to be targeted by the exome capture kit. CNV analysis was performed with VarScan2 (Koboldt et al., 2012) using default parameters.

Single cell sorting

Single cell sorting was performed on a BDFACSAria1–SORP instrument (BD®, Franklin Lakes, NJ, USA) equipped with an automated cell deposition unit using the following settings: 100micron nozzle, 1.4bar sheath pressure, 32.6KHz head drive and a flow rate that

gave 1-200 events per second. Viable cell thawing, single cell carboxyfluorescein diacetate N-succinimidyl ester (CFSE) staining of the neurosphere sample (according to manufacturer's instructions), cell sorting parameter explanations and the assessment of single cell sorting efficiencies was completed as previously described (Potter et al., 2013). Only fixed material surplus to FISH requirements was available from the xenograft transplant. This material was resuspended in PBS and labeled with propidium iodide prior to FACS. Five plates of single cells were collected from the GBM 5 neurosphere sample and two from the fixed xenograft material each composed of a no template control (NTC), 11 control cord blood cells and 84 target cells (GBM 5).

CNA assays were completed in quadruplicate and SNV assays were completed in duplicate. A heterozygous mutation was considered to be present if the Q-PCR C_T value from the mutant and wild-type sequence probes (FAM and VIC respectively) was <28 for a single cell. A homozygous mutation was considered to be present if there was no wild-type sequence signal. To ensure robust DNA CNA data from a system that can be influenced by assay efficiency and experimental variation we employed the $\Delta\Delta C_T$ method (Applied Biosystems®) to determine a copy number for each locus with modifications to incorporate data from three distinct assays targeting the control region (*B2M)* and the region of interest. Details of this approach can be found in (Potter et al., 2013).

Single cell multiplex targeted pre-amplification and Q-PCR

Labelled single cells were sorted into 2.5 µl lysis buffer composed of 1 mg/ml proteinase K (Qiagen Ltd, Manchester, UK) and 0.5% Tween20 in HEPES buffered saline (Sigma-Aldrich®, Gillingham, UK). Lysis was carried out for 50 min at 60^{o}C followed by 10 min at 98^{o}C. Specific (DNA) targeted amplification (STA) was then performed prior to Q-PCR. This multiplex STA reaction was composed of 5µl pre-amplification master mix (Life Technologies Ltd) and 2.5ul 1:40 primer mix (containing all primers for the gene targets of interest designed in house according to ABI guidelines). Denaturation was completed at 95^{o}C for 15 min, followed by 24 cycles of amplification at 95^{o}C for 15 sec and 60^{o}C for 4 min. The STA

product was then diluted 1:6 using DNA suspension buffer (TekNova®, Surrey, UK). Finally, 2.7 µl of the single cell target amplified DNA was interrogated by Q-PCR for each DNA target of interest using the 96.96 dynamic microfluidic array and the BioMark™ HD as recommended by the manufacturer; thermal phase 70^{o}C for 1800 sec, 25^{o}C for 60^{o}C sec followed by a hot start phase of 95^{o}C for 60 sec. This was followed by 35 cycles of 96^{o}C for 5 seconds and 60^{o}C for 20 sec. CNA assays were completed in quadruplicate and SNV assays were completed in duplicate.

Single cell Q-PCR analysis

The BioMark™ HD generates a C_T value for each reaction. A heterozygous mutation was considered to be present if the signals from the mutant and wild-type sequence probes (FAM and VIC respectively) had a C_T value <28 in a single cell. A homozygous mutation was considered to be present if there was no wild-type sequence signal. To ensure robust DNA CNA data from a system that can be influenced by assay efficiency and experimental variation we employed the $\Delta\Delta C_T$ method (Applied Biosystems®) to determine a copy number for each locus with modifications to incorporate data from three distinct assays targeting the control region (*B2M)* and the region of interest. The $\Delta\Delta C_T$ value was calculated for every target gene assay using each of the three reference gene C_T values generating nine estimated DNA copy number results for a region of interest. A confidence metric was assigned to the estimated copy number inferring the confidence with which an estimated copy number could be deemed true (according to Applied Biosystems CopyCaller® Software v2). Details of this approach can be found in Potter et al., 2013. The weighted mean of the nine estimated DNA copy numbers (for a region of interest) was used as the final DNA copy number taking into consideration the confidence metric attributed to each. This reduced the contribution of less reliable estimated DNA copy numbers to the final DNA copy number. Estimated copy number results were not considered if the confidence value was less than 50% or the estimated copy number was greater than four (with only quadruplicates per

assay the results are not robust enough to accurately detect DNA copy numbers greater than four. At least two of the nine estimated copy numbers must have a confidence value above 50% to calculate the final copy number for a region of interest.

SUPPLEMENTAL REFERENCES

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome research *22*, 568-576.

Larson, D.E., Harris, C.C., Chen, K., Koboldt, D.C., Abbott, T.E., Dooling, D.J., Ley, T.J., Mardis, E.R., Wilson, R.K., and Ding, L. (2012). SomaticSniper: identification of somatic point mutations in whole genome sequencing data. Bioinformatics *28*, 311-317.

Nannya, Y., Sanada, M., Nakazaki, K., Hosoya, N., Wang, L., Hangaishi, A., Kurokawa, M., Chiba, S., Bailey, D.K., Kennedy, G.C.*, et al.* (2005). A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer research *65*, 6071-6079.