Electronic Supplementary Material (ESI) for Molecular BioSystems

Supplementary Figure Legends

Supplementary Figure 1. Gating strategy used to identify single Glioblastoma tumor cells. Gate 1: forward scatter area (FSC-A) and side scatter area (SSC-A) were used to enrich for intact cells. Gate 2: DAPI staining recorded in the logarithmic scale (FL9 Log) was used to identify events that contained intact nuclei (red gate), effectively excluding cell debris that did not contain intact nuclei. Gate 3: Side scatter pulse width and forward scatter pulse width were used to identify single cell events. A density plot with a very tight gate (shown in blue) provided the most stringent single cell gate. Gate 4: The far red channel (FL8) was used to identify both dead cells (Invitrogen LIVE/DEAD® Fixable Cell Stain) and leukocyte cells using the leukocyte common antigen CD45 (BD Bioscience Alexa Fluor® 700 Mouse Anti-Human CD45, Catalogue number 560566). Gate 5: The NFF diploid control cells were identified based on positive CFSE staining, with the GBM tumor cells being CFSE negative. Gate 6: Neural and glial cells within the CFSE negative gate were identified using neural cell adhesion molecule 1 (NCAM-1) staining (BD Bioscience Alexa Fluor® 647 Mouse Anti-Human CD56 Catalogue number 563443) on the red channel (FL6). DNA content was assessed by DAPI staining recorded on the FL9 linear scale, and the presence of polyploidy was identified using the 4n+ gate as shown in the final panel. Isotype and single stained controls were used for compensation

Supplementary Figure 2. Graphing 2n DNA peaks using FlowJo and Prism. **(A)** The tumor cell sample is mixed with human diploid control cells pre-stained with CFSE, the mix of live cells stained with fixable LiveDead (if required), and then the cell mix is fixed, permeabilized and stained using antibodies of interest (e.g. CD45, CD56) and DAPI to identify DNA. The first gate is a crude cell gate consisting of forward scatter area (FSC A) and side scatter area (SSC A), which identifies cell populations from other debris. **(B)** Next, single cells are identified using forward scatter and side scatter time of flight (FSC TOF and SSC TOF); a contour plot is used in this example to highlight the single cell peak. **(C)** Next intact nuclei are identified using the DAPI channel recorded on the Log scale, which clearly identifies the intact nuclei (red gate) from degraded nuclei. **(D)** CFSE staining is then used to identify the unstained tumor cell population (GBM) and the CFSE-stained human diploid control (Control). Next, the DNA content of the tumor cell population **(E)** and the human diploid control population **(F)** are displayed as DNA histograms with the DNA content revealed using the DAPI Linear channel. The 2n DNA peaks are readily identifiable and are gated within the FlowJo software as shown. The raw data sets for each 2n peak, which contains the linear DAPI fluorescent intensity for each cell, is then exported from FlowJo as a Txt file. The Txt files are imported into Prism and the Mean for the diploid control calculated using the Prism software statistical analysis tool. This value is then used to normalize both the normal human diploid control and the GBM tumor sample using the Prism Normalization tool and expressing the normalized values as a fraction. In this way the DNA content of the normal diploid control has a mean value of one, and the aneuploidy tumor sample is some value relative to one, which gives a semi-quantitative value as to whether the tumour is subdiploid or hyperdiploid relative to the control sample. Panel **(G)** shows a simple overlap of the DNA histograms, with the diploid control shown in red and the tumor sample in blue, revealing the tumor sample to be slightly sub-diploid relative to the control. Panel **(H)** shows the 2n peaks of the same samples expressed graphically as box plots after normalization in Prism. Again, this reveals that the tumor 2n subpopulation is subdiploid relative to the control population, with the advantage that many samples can be clearly displayed on a single graph.

Supplementary Figure 3. Histopathology of Tumours. Typical haematoxcylin and eosin staining demonstrating sections of xenograft tumours derived from three primary parental control tumorsphere lines (left panels) and their derived hyperdiploid clonal lines (right panels) that all showed highly cellular tumors with pleomorphic astrocytic tumor cells with marked nuclear atypia and brisk mitotic activity. Numerous atypical mitotic figures are identified (yellow arrows).

Supplementary Figure 4. Hyperdiploid tumor cells are enriched within the slow-cycling, label retaining subpopulation in three primary patient tumorsphere lines. The Parental population of the three primary patient lines used in this study were stained with CFSE and then cultured for 7 days under tumorsphere culture conditions. CFSE is diluted with each cell division; after 7 days infrequently dividing cells can be identified as the label-retaining, CFSE-high subpopulation of cells. The upper panels show DNA content of the bulk population, revealing a typical cell-cycle distribution between the 2n and 4n peaks. The bottom panels show the DNA content of the label-retaining (Top 5% CFSE). The label retaining population of all three clones clearly shows an increase in the 4n and 8n populations. Analyzing the DNA content of the bulk and label retaining subpopulations revealed that both the 4n and 8n peaks increased with label-retention, showing that the proportion of hyperdiploid cells increased within the label-retaining, infrequently cycling cells for all three lines.

Supplementary Figure 5. Visualization of CGH array data for individual chromosomes. CGH data for each chromosome identifies obvious chromosomal gains and/or losses present within the polyploid clone compared to the parental control. All three clones displayed both gains and losses relative to the parental bulk population. The green line displays where a normal diploid karyotype would appear.

Supplementary Table 1. Total Pathway Analysis of Copy Number Variation.

Using the segmented CGH Array data, we identified 284 genomic regions across 16 chromosomes with an average log ratio greater than 0.3 or smaller than -0.3, corresponding to a gain or loss respectively between any clone and the parent. The pathways overrepresented by the list of common genes were analyzed using Ingenuity Pathway Analysis and the categories that are overrepresented in all three clones are itemized in this table.

Supplementary Table 2. Total list of GeneGo Pathways analysis of genes affected in all three clones.

Segmented data from CGH array was used to identify copy number variations is our Glioblastoma parental and hyperdiploid clones. The pathways over-represented in all three clones in the list of the commonly altered genes was determined using GeneGo, these data are itemized in this table.