**Giraud et al, Supplemental Figure Legends**

**Supplemental Figure S1: *ALDHhigh CRC cellsdisplay CSC characteristics and elevated progastrin expression.***

**A,** Aldefluor FACS profiles of ALDHhigh cells compared to their respective DEAB-treated negative controls (top panel). Percentages of ALDHhigh cells for each cell type are provided in the bottom panel (n=3). **B,** Quantification of *ALDH1A1* mRNA following FACS-mediated separation of “ALDHhigh” (dark bars) and “ALDHneg” (light bars) cell populations. *\** P <0.05, \*\* P < 0.01, \*\*\* P < 0.001, ns = non significant, Mann Whitney test compared to matching ALDHneg population. Data in the graph bars are presented as the mean with SEM of 3 independent experiments except for HT29 (#, n = 1). **C,** Quantification of *GAST* mRNA using RT-qPCR ALDHhigh and ALDHneg cells CPP19 CRC cells, BJ human foreskin fibroblasts (a kind gift from Prof. Ricky Jonhstone, Peter Mac Callum Cancer Centre, Melbourne, Australia) and T98 and U118 human glioblastoma cells (a kind gift from Dr. Theo Mantamadiotis and Dr Paul Daniel, Department of Pathology, the University of Melbourne, Australia). A quantitative PCR signal corresponding to the expected amplicon is detected in ALDHhigh and ALDHneg cells. ALDHneg cells are detected with a Ct of 30-32. In contrast, BJ, T98 and U118 cells generate a non-specific signal (as shown by their respective melting curves) around 36-39 Ct. **D,** Representative FACS profiles of CD44/EpCAM-stained T84 and CPP25 colorectal cancer cells; Mean percentage of double positive cells is indicated in the upper right quadrant (n = 3). **E**, mRNA quantification for CD44 and GAST performed on CD44/EpCAM-double positive and negative cells purified as in (**D**). **F,** Representative FACS profiles of LGR5-stained T84 and CPP1 colorectal cancer cells; Mean percentage of positive cells is indicated in the right quadrant (n = 3). **G**, mRNA quantification for LGR5 and GAST performed on LGR5-positive and negative cells purified as in (**F**). **H-I,** representative FACS profiles (**H**) and quantification (**I**, percentages) of CD133-positive populations in multiple CRC cells. **J**-**K**, RT-qPCR-mediated quantification of mRNAs encoding *CD133* (**J**) and progastrin (*GAST)* (**K**) according to the CD133 sorting group (CD133neg, clear bars; unsorted, grey bars; CD133high, dark bars) in different cell lines. **L**, CSC gene expression level quantified by qRT-PCR in multiple CRC cells grown as colonopheres in the absence of serum (3D), expressed as a ratio of their expression in matching cells grown as monolayer with serum. Values above 1 reflect a higher expression in colonospheres than in monolayer conditions. **M**, Colonosphere number formed by CPP1 or T84 cells according to their ALDH activity (ALDHneg, white bars; ALDHhigh, black bars; unsorted, grey bars). Data are presented as the mean +/- SEM from a minimum of three independent experiments. \*\* P < 0.01, \*\*\*\* P <0.0001, one way ANOVA test.

**Supplemental Figure S2: *Progastrin inhibition decreases colonosphere incidence, CSC frequency and ALDHhigh cell proportion in vitro.***

**A**, RT-qPCR validation of *GAST* gene down-regulation 24h after transient siRNA transfection (sigal, black bars; siPG, white bars) in T84, CPP1 and CPP19 cells grown as colonospheres. \* P < 0.05 compared to sigal, two-tailed Student t-test, n=5. **B**, sphere size measured 7 days after CPP1 or T84 cells transfection with sigal (dark dots) or siPG (clear dots). \*\*\*\* P < 0.0001, Mann-Whitney test, n=3. **C**, Western blot detection (Top Panel) of phospho(Y418)-Src and total Src protein following exposure of HT-29 colorectal cancer cells to increasing concentrations of progastrin (0.1-5nM) for 15min at 37C. Bottom panel, quantification of Western blots performed as shown above (\*, p<0.05, one-way ANOVA with Bonferroni post-hoc test, n=3). **D,** HTRF-based quantification of p38MAPK phosphorylation in T84 colorectal cancer cells incubated for 30min with control or progastrin-selective polyclonal antibodies (1.5g/ml) with or without recombinant human progastrin (hPG, 5nM), as indicated, in comparison with anisomycin positive control (Pos). \*, p < 0.05 compared to control Ab alone, #, p <0.05 compared to control Ab + recombinant progastrin, ANOVA + Bonferroni post-hoc test. **E,** anti-progastrin polyclonal antibody (antiPG, white bars) reduces the ability of additional CRC cell lines to form colonospheres 7 days after treatment (3µg/mL daily), in comparison with control antibody treatment (CT, black bars). \* P < 0.05, Mann-Whitney test compared to respective CT. Data are presented as the mean +/- SEM. **F**-**G**, RT-qPCR quantification of *GAST* mRNA (**F**) and calculation of CSC frequency (**G**) in CPP1 and T84 cells treated with doxycyclin in order to induce the expression of inducible shRNAs (shCT, black bars; shPG, white bars). \* P < 0.05, Mann-Whitney test compared to respective shCT with DOX.

**Supplemental Figure S3: *Progastrin regulates the self-renewal of ALDHhigh colorectal CSCs.* A,** HTRF-based quantification of p38MAPK phosphorylation in ALDHneg or ALDHhigh T84 and CPP19 colorectal cancer cells incubated for 30min with control or with progastrin-selective polyclonal antibodies (1.5g/ml) with or without recombinant human progastrin (hPG, 5nM), as indicated, in comparison with Anisomycin positive control (Pos). \*, p < 0.05 and \*\*, p < 0.01 compared to control Ab alone, #, p <0.05 and ##, p < 0.01 compared to control Ab + recombinant progastrin, ANOVA + Bonferroni post-hoc test. **B,** Colonosphere formation 7 days after ALDHhigh CPP1 cell purification and growth in suspension in the presence of antibodies (3µg/mL daily, anti-PG antibody, white bar; control antibody, dark grey bar) or of recombinant progastrin (hPG 10nM, clear grey bar), in comparison with untreated controls (-, black bar). \* P < 0.05, \*\*\*\* p < 0.0001, ns = non significant, one way ANOVA test.

**Supplemental Figure S4: *shRNA-mediated down regulation of progastrin in vitro and in vivo.*** **A**-**B**, representative FACS profiles (**A**) and quantification (percentages, **B**) of ALDHhigh T84 cells 48 hours after treatment with doxycycline (+ DOX), in comparison with untreated cells (shCT, black bars; shPG, white bars). \*\*\* P < 0.001, Mann-Whitney test. Data are presented as the mean of 4 independent experiments +/- SEM. **C,** RT-qPCR quantification of the *GAST* mRNA in the first generation of tumors initiated by T84 ALDHhigh cells (shCT, black bar; shPG, white bar). \*\*\* P < 0.001 compared to shCT, Student t-test. Data are presented as the mean +/- SEM, n=6 mice. **D,** RT-qPCR quantification of *GAST* mRNA expression in CPP19, SW480 shCT and SW480 shPG colorectal cancer cells stably transfected with a control (pBABE-Zeo, white bars) or a full-length progastrin-expressing (pBABE-Zeo + PG, black bars) construct. \*, p < 0.05, \*\*, p < 0.01 vs matching pBABE-Zeo control, t-test, n=3). **E,** Tumor intake in balb/C nude CD1 mice 30 days after injection with 2500 ALDHhigh SW480 or CPP19 cells as indicated. No tumours were detected for up to 60 days in mice injected with 50 or 250 cells.

**Supplemental Figure S5: *Promotion of a glycolytic profile by progastrin is essential for the self-renewal ability of ALDHhigh cells.* A,** Hierarchical clustering of 18 samples and 140 genes differentially expressed between GAST siRNA (siPG) and control siRNA (siβgal)-transfected ALDHhigh cell-derived T84 colonospheres.Hierarchical clustering was applied to the 18 mRNA samples (triplicate of each time point (6, 24 and 72h) for cells transfected with siPG and siβgal) and to the 140 genes identified as differentially expressed at the 72h time point. We used centroid-linkage clustering and Pearson correlation as similarity metrics. Each row represents a gene and each column represents a sample. The expression level of each gene in a single sample is relative to its median abundance across all samples and is depicted according to a color scale shown at the bottom. Red and green indicate expression levels respectively above and below the median. The magnitude of deviation from the median is represented by the color saturation. The dendrogram of samples (above matrix) represents overall similarities in gene expression profiles. Characteristics of samples are indicated between the dendrogram and the data matrix and are color-coded as indicated. **B,** Time-course of glycolysis (ECAR, left panel) and of oxygen consumption (OCR, right panel) measurements in ALDHhigh cells transfected with sigal (black curves) or siPG (light curves) for 96h, in comparison with sigal-transfected ALDHneg cells (grey curves). **C-D**, Quantification of glycolysis (ECAR, expressed as mpH per minute) in HT-29 **(C)** and HCT-116 **(D)** ALDHhigh cells incubated with control (black bars) or progastrin-selective polyclonal antibodies (grey bars) for 96h, in comparison with ALDHneg cells (white bars). Data represent the mean +/- SEM from 3-6 replicates of representative experiments. \*, p < 0.05 compared to ALDHhigh + control antibody and #, p < 0.05 compared to ALDHhigh + PG antibody, 2-way ANOVA with Newman-Keuls post-hoc test. **E,** CSC frequency in ALDHhigh cells 7 days after transfection of a siRNA selectively directed against the *GAST* gene (siPG) or against a control sequence (sigal), in comparison with ALDHneg cells. Cells were supplemented or not with vehicle or with 0.05microM Antimycin A (AA). Statistically significant differences were tested using a Chi-Square test.

**Supplemental Figure S6: *Annexin II expression is not enriched in cancer stem cell populations.*** **A**, RT-qPCR quantification of *ANXAII* mRNA in multiple CRC cells grown as colonospheres in the absence of serum (3D), compared to their respective counterparts grown as monolayers with serum. Values above or below 1 respectively signify that *ANXAII* is more or less expressed in colonospheres than in monolayer condition. **B**, RT-qPCR quantification of *ANXAII* mRNA according to ALDH activity levels in CRC cells (ALDHneg, light bars; ALDHhigh, dark bars). *\** P <0.05, \*\* P < 0.01, ns = non significant, Mann Whitney test compared to matched ALDHneg cells. Data in the graph bars are presented as the mean +/- SEM from 3 independent experiments.