Supplementary Figure 1. CD44 and ALDH1A1 is prognostic when highest quartile is used to separate the groups. The highest quartile of expression was used to divide the patient set in two groups after which the relapse-free survival probability was analyzed for each group. Only CD44 and ALDH1A1 displayed the capacity to predict prognosis, whereas the other CSC markers did not. **A.** CD24, **B.** CD29, **C.** CD44, **D.** CD166, **E.** Lgr5, **F.** ALDH1A1.

Supplementary Figure 2. Correlation in expression levels between different CSC markers. Expression levels of different CSC markers were plotted against each other. Linear regression was used to study if there is a significant correlation. **A.** CD133 versus CD24, **B.** CD133 versus CD29, **C.** CD133 versus CD44, **D.** CD133 versus CD166, **E.** CD133 versus Lgr5, **F.** CD133 versus ALDH1A1, **G.** CD24 versus CD29, **H.** CD24 versus CD44, **I.** CD24 versus CD166, **J.** CD24 versus Lgr5, **K.** CD24 versus ALDH1A1, **L.** CD29 versus CD44, **M.** CD29 versus CD166, **N.** CD29 versus Lgr5, **O.** CD29 versus ALDH1A1, **P.** CD44 versus CD166, **Q.** CD44 versus Lgr5, **R.** CD44 versus ALDH1A1. **S.** CD166 versus Lgr5, **T.** CD166 versus ALDH1A1, **U.** Lgr5 versus ALDH1A1.

Supplementary Figure 3. CD133 expression of DLD1 and DKO4 without 5-aza treatment. A. CD133 mRNA expression of DLD1 and DKO4 cells. **B.** FACS analysis of CD133 expression of DLD1 and DKO4 cells. **C.** Western blot for phospho-ERK and ERK showing that MEK inhibition with U0126 is effective in the three lines tested. **D.** FACS analysis for phospho-ERK on LM5 for which limited cell numbers where available again shows effective MEK inhibition with U0126. **E.** Inhibition of downstream MEK by U0126 inhibitor for 24h reduces CD133 expression in K-Ras mutant CSC line LM5 on mRNA as well as protein shown by FACS analysis.

Supplementary Figure 4. Gene set enrichment analysis after K-means clustering with activated K-Ras gene expression signature. An activated K-Ras gene expression signature was used to segregate ours as well as other patient collections in two groups by k-means clustering (left panel). CD133 expression consequently studied in these two clusters (middle panel). To identify which cluster resembled the activated K-Ras profile, gene set enrichment analysis was used (right panel). **A.** our CRC patient cohort (n=85), **B.** a neuroblastoma collection (n=88), **C.** a glioma cohort (n=153), **D.** a breast cancer cohort (n=204, **E.** an ovarian cancer collection (n=90). ES: enrichment score; NES: normalized enrichment score,

FDR: false discovery rate. **F.** Scatter plot of CD133 expression in the different brain tumor types of the glioma cohort as described in (**C**).

Supplementary Figure 5. Gene set enrichment analysis after dividing the patient cohort in CD133 low versus CD133 high for activated K-Ras gene expression signature. The patient cohorts were divided in two equal groups based upon the median of the CD133 expression (left panel). Then, by gene set enrichment analysis, it was studied which cluster had the most overlap with the mutant K-Ras gene expression profile (right panel). **A.** our CRC patient cohort (n=85), **B.** a neuroblastoma collection (n=88), **C.** a glioma cohort (n=153), **D.** a breast cancer cohort (n=204) and **E.** an ovarian cancer collection (n=90). ES: enrichment score; NES: normalized enrichment score, FDR: false discovery rate.