The Intestinal Stem Cell Signature Identifies Colorectal Cancer Stem Cells and Predicts Disease Relapse

Anna Merlos-Suárez, Francisco M. Barriga, Peter Jung, Mar Iglesias, María Virtudes Céspedes, David Rossell, Marta Sevillano, Xavier Hernando-Mombolina, Victoria da Silva-Diz, Purificación Muñoz, Hans Clevers, Elena Sancho, Ramón Mangues, and Eduard Batlle

1Oncology Program
2Biostatistics Unit
Institute for Research in Biomedicine (IRB), Baldiri Reixac 10, 08028 Barcelona, Spain
3Department of Pathology, Hospital Universitari del Mar, Universitat Autònoma de Barcelona, 08003 Barcelona, Spain
4Institute of Biomedical Research (IIB Sant Pau) and Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN, MICINN), Sant Antoni M. Claret, 167, 08025 Barcelona, Spain
5Catalan Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institut (IDIBELL), Gran Via de L'Hospitalet 199-203, 08907 Barcelona, Spain
6Hubrecht Institute and University Medical Center Utrecht, Uppsalalaan 8, 3584CT Utrecht, the Netherlands
7Institució Catalana de Recerca i Estudis Avançats (ICREA)

*Correspondence: eduard.batlle@irbbarcelona.org
DOI 10.1016/j.stem.2011.02.020

SUMMARY

A frequent complication in colorectal cancer (CRC) is regeneration of the tumor after therapy. Here, we report that a gene signature specific for adult intestinal stem cells (ISCs) predicts disease relapse in CRC patients. ISCs are marked by high expression of the EphB2 receptor, which becomes gradually silenced as cells differentiate. Using EphB2 and the ISC marker Lgr5, we have FACs-purified and profiled mouse ISCs, crypt proliferative progenitors, and late transient amplifying cells to define a gene program specific for normal ISCs. Furthermore, we discovered that ISC-specific genes identify a stem-like cell population positioned at the bottom of tumor structures reminiscent of crypts. EphB2 sorted ISC-like tumor cells display robust tumor-initiating capacity in immunodeficient mice as well as long-term self-renewal potential. Taken together, our data suggest that the ISC program defines a cancer stem cell niche within colorectal tumors and plays a central role in CRC relapse.

INTRODUCTION

Tissue regeneration in the intestine is ultimately sustained by intestinal stem cells (ISCs) that reside at the base of mucosal invaginations called crypts (Figure 1A). The progeny of ISCs, termed transient amplifying (TA) cells, is expanded through several rounds of mitosis while it migrates upwards along the crypt axis. Close to the intestinal lumen, TA cells undergo cell-cycle arrest and terminal differentiation. Stem cells of the mouse small intestine (Barker et al., 2007), colon (Barker et al., 2007), and stomach (Barker et al., 2010) can be identified by the specific expression of Lgr5, a G protein-coupled receptor of unknown function. Lgr5+ ISCs are long lived, proliferate continuously, and generate all cell types present in the gut. The onset of intestinal tumorigenesis is driven in most cases by activating mutations in the Wnt signaling pathway. Mouse Lgr5+ cells give rise to intestinal tumors with higher efficiency than other intestinal epithelial cell populations upon mutational activation of the Wnt pathway (Barker et al., 2008). This observation suggests that ISCs represent the cell of origin for colorectal cancer (CRC).

CRC is the second cause of death by cancer. The current therapeutic strategy for most CRC patients includes surgical resection of the tumor and chemotherapeutic treatment. After curative therapy, a large proportion of CRC patients remain clinically free of disease for months or years. Yet, after such period of remission, cancer recurs in 30%–50% of all cases, generally in the form of metastasis. The risk of cancer recurrence is linked to the stage of the disease at the time of diagnosis (American Joint Cancer Committee [AJCC] staging system). Most Stage I CRC patients (95%–98%) will not develop recurrent cancer, whereas 20%–25% of Stage II and 40%–50% of Stage III CRC patients will experience disease relapse upon treatment.

Although a large proportion of deaths by CRC are caused by recurrent tumors, the biological basis for disease relapse remains unknown. Cancer relapse is caused by isolated tumor cells that disseminated prior to resection of the primary tumor and that resisted chemotherapeutic treatments. In some patients, resilient cancer cells eventually resume growth and regenerate a new tumor mass.

Recent studies have evidenced that not all cell populations present in primary CRCs are capable of regenerating a tumor as isolated entities. CRCs contain subsets of cells that can effectively propagate the disease when implanted in an immunodeficient host. These cells, termed tumor-initiating cells, can be identified by the expression of surface markers such as cd133

Cell Stem Cell 8, 511–524, May 6, 2011 ©2011 Elsevier Inc. 511
Cd133+ or cd44+ cell populations represent a particular subset of aggressive CRC cells. In addition, some indirect evidences suggest a connection between tumor-initiating cells and ISCs. In particular, cd133 and cd44 are expressed in the proliferative compartment of normal crypts (progenitors and ISCs) (Snippert et al., 2009; Zeilstra et al., 2008). Also, purified cd133+ tumor-initiating cells can be expanded in vitro as an undifferentiated population that upon xenotransplantation generates tumors that contain differentiated-like cells (Vermeulen et al., 2008). Despite these suggestive observations, a relationship between ISCs and cd133+/cd44+ tumor cells remains speculative. Also, there is lack of evidence supporting a role of tumor-initiating cells in cancer recurrence.

Genes transcriptionally regulated by the Wnt pathway in the intestinal epithelium are expressed with patterns ranging from restricted expression to ISCs (e.g., Lgr5) to broad expression in proliferative cells (e.g., Myc) (Van der Flier et al., 2007). One of such Wnt target genes, the receptor tyrosine kinase EphB2, is expressed in a decreasing gradient from the crypt base toward the differentiated cell compartment (Batlle et al., 2002). EphB activity is required to establish the position of the different cell types in the crypts as mice mutant for EphB2, EphB3 (Batlle et al., 2002), or their ligand ephrin-B1 (Cortina et al., 2007) display defects in intestinal cell compartmentalization. Here, we take advantage of EphB2 graded expression to purify and obtain the expression profiles of ISCs and other crypt cell types. We discover that expression of ISC-specific genes associates with...
the risk of developing recurrent CRC. Furthermore, paralleling the results in normal crypts, the graded expression of EphB2 in CRCs identifies tumor cell populations displaying ISC-like or differentiated-like phenotypes.

RESULTS

Purification and Expression Profiling of ISCs and Other Crypt Cell Types

The transcriptional profiles of the different cell types present in the intestinal epithelium remain essentially uncharacterized, which precludes an in-depth analysis of intestinal cell type-specific gene program expression in CRC. We developed a cell isolation method based on surface expression of EphB2, which is expressed following a decreasing gradient from the crypt base (Figure 1A). ISCs (i.e., crypt base columnar cells) showed highest expression of membrane EphB2 (Figure 1A, black arrowheads), whereas TA cells progressively decreased EphB2 protein levels as they migrated toward the surface. Differentially cells in the villus and Paneth cells were negative for EphB2 expression (Battle et al., 2002). We stained suspensions of disaggregated mouse intestinal cells with a monoclonal antibody directed against the extracellular domain of EphB2 (Mao et al., 2004), and by FACS, we isolated epithelial cells (EpCAM+, cd45−, cd31−, cd11b−) displaying different EphB2 surface expression (Figure 1B). Analysis of marker genes revealed similar levels of Ki67, Myc, FoxM1, (Figure 1C) and other genes characteristic of proliferating cells in EphB2hi and EphB2med cell subpopulations (Figure S1 available online). However, EphB2hi cells expressed around 5-fold higher levels of ISC-specific genes Lgr5 (Barker et al., 2007), Ascl2 (van der Flier et al., 2009), and Olfm4 (van der Flier et al., 2009) (Figure 1C). These results suggest that the EphB2hi cell population is enriched in ISCs, whereas EphB2med cells correspond mainly to crypt TA cells.

To test the functionality of each crypt cell population, freshly isolated cells from mouse small intestine were cultured in conditions that mimic the intestinal stem cell niche as previously described (Sato et al., 2009). Single cells expressing high or medium levels of EphB2 developed into complex multicellular 3D structures (Figure 1D). These in vitro organoids recapitulated the organization of the intestinal epithelium, including the presence of multiple crypt-like structures (Figures 1D and 1F, arrowheads) distributed around a central lumen. EphB2lo cells showed 3- to 5-fold higher efficiency than EphB2med cells in the formation of intestinal organoids (Figure 1E). This result parallels the enrichment in ISC marker gene expression between EphB2hi and EphB2med populations shown in Figure 1C. On the contrary, EphB2lo and EphB2neg cells never grew (Figure 1E). Organoids derived from single EphB2hi cells displayed proliferative Ki67+ cells within crypt-like evaginations (Figure 1G, empty arrowheads) and cell-cycle-arrested Krt20+ differentiated cells forming large lumens (Figure 1H, arrows). Importantly, staining for lineage-specific markers revealed the presence of all mature intestinal cell types (Figures 1I–1K). Organoids were expanded by mechanical disaggregation every 4 to 5 days over a 2 month period without decrease in growth rates. The efficiency of organoid generation from isolated EphB2hi cells was in the range of 0.5%–1%, similar to that obtained from Lgr5-GFPhi cells in our laboratory (1%–2%; data not shown) and slightly lower than reported before (Sato et al., 2009). Overall, these observations support the idea that the EphB2hi cell population is enriched in ISCs.

Gene Expression Signatures that Define ISCs, Proliferative Cells, and Late TA Cells

By microarray analysis, we identified the global expression profile of EphB2hi ISCs. We selected probe sets enriched at least 2-fold in EphB2hi compared with EphB2med cells, and that were further downregulated in EphB2lo cells. Only 71 probes (54 annotated genes) passed this stringent criterion (Table S1), including the previously described ISC-specific genes Lgr5 (Barker et al., 2007) and Ascl2 (van der Flier et al., 2009). The expression of several randomly selected genes from this list (Smoc2, Pcdh8, and Igfbp4) was restricted to small, wedge-like cells intermingled with Paneth cells at the crypt-bottom most positions (Figure 1L; data not shown). This pattern is characteristic of Lgr5+ cells (Barker et al., 2007), implying that these three genes represent new ISC-specific markers. To broaden this analysis, we analyzed the expression of the EphB2hi ISC signature in Lgr5-GFP knockin mice (Barker et al., 2007). The vast majority (98%) of genes contained in the EphB2hi ISC gene expression signature were upregulated (>1.5-fold, p < 0.05) in Lgr5-GFPhi cells relative to Lgr5-GFPneg cells (Figures S2A–S2C) (Barker et al., 2007). Of note, the EphB2hi ISC gene signature did not contain core pluripotency genes (i.e., Oct4, Sox2, nanog, klf4) or the Myc-driven gene module characteristic of ESCs (Kim et al., 2010). It was not enriched in genes directly involved in proliferation owing to equal levels in ISCs and TA cells (see below).

We also established the gene program that defines the proliferation status of crypt cells by selecting probes expressed at equivalent levels in EphB2hi and EphB2med cells that were silenced in EphB2hi cells (Table S1). This gene module was enriched in known mediators of DNA replication, mitosis, DNA-damage repair, and proliferation (Figure 1C and Figure S1). Genes with higher expression levels in EphB2hi compared to EphB2med cells that were further downregulated in EphB2hi subpopulation constituted the gene expression signature of late progenitor cells (Table S1). This set contained several well-established markers of intestinal differentiation such as Krt20, An Pep, or Ada (Figure 1C, Figure S1 and Table S1).

Aggressive CRCs Are Enriched in ISC Gene Expression

Each gene expression signature described above was instrumental to investigate whether the global transcriptional profile of human CRCs resembled that of any particular crypt cell type. Crypt proliferation genes were overexpressed in all CRCs compared with normal tissue (Figure 2A). Remarkably, 38% of the genes from the EphB2hi ISC gene program were also overexpressed in CRCs, including the bona fide ISC markers Lgr5 and Ascl2 (Figure 2B) (average fold change > 1.5, p < 0.05). The signature of late TA cells was silenced overall (Figure 2A), and Krt20 and other markers of intestinal differentiation were downregulated relative to normal tissue (Figure 2B and data not shown).

We next investigated the association of each crypt cell gene-expression signature with clinical progression of CRC. We selected for analysis a representative pooled cohort of 340 primary colon cancer patients treated at three different hospitals.
Figure 2. Aggressive CRCs Are Enriched in ISC Gene Expression

(A) Unsupervised hierarchical clustering analysis of transcriptional profiles of normal mucosa (NM) and CRC samples according to different crypt expression signatures.

(B) Plots depict relative expression levels of Lgr5, Ascl2, Ki67 and Krt20 in NM and CRC samples according to microarray measurements.

(C,D) GSEA results for the EphB2hi ISC, Proliferation and Late TA signatures in Stage III-IV compared to Stage I-II colon cancers (C) or in well-differentiated versus poorly differentiated colon cancers (D). ES, enrichment score; NES, normalized enrichment score; FDR, false discovery rate.

See also Figure S2 and Tables S2, S3, and S4.
Expression of ISC-Specific Genes Predicts Disease Relapse in CRC Patients

GSEA also revealed that the EphB2-derived ISC signature was highly enriched in genes that positively predicted the relative risk of developing recurrent cancer upon intended curative therapy (FDR < 1 x 10^-8; Figure 3A). We used ISC gene expression levels to stratify CRC patients into three groups (Figure 3B). Patients bearing primary CRCs with high average expression of ISC genes had a relative risk of relapse 10-fold higher than those with low levels (p < 0.0001). The EphB2-ISC medium expression group displayed an intermediate risk (medium versus low; HR = 6.28; p = 0.0182). Remarkably, the EphB2-ISC low group displayed an intermediate risk (medium versus low; HR = 1 x 10^-8; data not shown) as well as in poorly differentiated compared with well-differentiated tumors (Figure S2E). Altogether, these results suggest that a distinctive feature of aggressive colon cancers is high expression of an ISC-like gene program.

Humanized Colon Crypt ISC Signatures Identify Recurrent CRCs

The gene expression signature obtained from late TA cells was enriched in genes that predicted no relapse over follow-up, albeit with borderline statistical significance (FDR = 0.03; Figure 3A). In addition, several well-established markers of intestinal differentiation associated inversely with tumor recurrence, including MUC2 (p = 0.024; HR = -1.37), TFF3 (p = 0.011; HR = -1.38), MATH1 (p = 0.0001; HR = -1.54), CDX2 (p = 0.0005; HR = -1.54), or Villin (p = 0.02; HR = -1.31). GSEA also demonstrated that the crypt proliferation signature was inversely associated with disease relapse (Figure 3A) as previously reported (Anjomshoa et al., 2008). This observation reinforces the notion that ISC and proliferation genes measure independent features of the tumor.

Expression of ISC-Specific Genes Predicts Disease Relapse in CRC Patients

GSEA also revealed that the EphB2-derived ISC signature was highly enriched in genes that positively predicted the relative risk of developing recurrent cancer upon intended curative therapy (FDR < 1 x 10^-8; Figure 3A). We used ISC gene expression levels to stratify CRC patients into three groups (Figure 3B). Patients bearing primary CRCs with high average expression of ISC genes had a relative risk of relapse 10-fold higher than those with low levels (p < 0.0001). The EphB2-ISC medium expression group displayed an intermediate risk (medium versus low; HR = 6.28; p = 0.0182). Remarkably, the EphB2-ISC low group displayed an intermediate risk (medium versus low; HR = 1 x 10^-8; data not shown) as well as in poorly differentiated compared with well-differentiated tumors (Figure S2E). Altogether, these results suggest that a distinctive feature of aggressive colon cancers is high expression of an ISC-like gene program.

Humanized Colon Crypt ISC Signatures Identify Recurrent CRCs

The gene expression signatures used in this study were obtained from mouse intestinal cells. We reason that genes involved in essential functions for crypt cells must be largely conserved between human and rodents. Yet, to rule out the possibility that the use of mouse-derived gene lists may impose artifacts in the clinical association studies, we analyzed their overall expression patterns in human colon mucosa. We interrogated a publicly available data set of transcriptomic profiles of normal human colon mucosa in which the top crypt and bottom compartments were microdissected and analyzed separately (Kosinski et al., 2007). GSEA showed that EphB2-ISC, Lgr5-ISC, and Proliferation signatures were largely enriched in crypt human bottom compared to top region compartments, whereas the Late TA signature followed the reverse pattern (Figure 4B). Thus, to a large extent, mouse-derived signatures used in this study are expressed by human colon crypts with equivalent patterns. Using the expression profiles of human colon crypts as filters, we excluded from each mouse crypt-derived cell signature those genes that were not differentially expressed between human colon crypt-bottom and crypt-top compartments (see Experimental Procedures). The resulting “humanized” gene lists (Table S5) associated with disease progression with equivalent strength and specificity than the unrefined mouse crypt-derived gene lists (Figures 4C-4F and Figure S3). In particular, Hu-EphB2-ISC and Hu-Lgr5-ISC signatures were enriched in metastatic CRC (Stage IV versus the rest, Hu-EphB2-ISC; FDR = 0.04 and Hu-Lgr5-ISC; FDR < 10^-16, data not shown) as well as in recurrent tumors (Figure 4C and Figure S3A). Analogous to the signature of mouse crypt proliferative cells, the Hu-Proliferation signature was enriched in nonrecurrent CRCs (FDR < 10^-16, data not shown). In addition, GSEA confirmed a further enrichment of Hu-ISC signatures in tumors that will relapse relatively late (>2 years) (Figure 4D and Figure S3D). Both Hu-EphB2-ISC and Hu-Lgr5-ISC signatures had approximately linear effect on the risk of recurrence (Figure 4E and Figure S3D). Multivariate analysis using the Cox Proportional Hazards Model demonstrates that the power of the humanized ISC signatures to predict relapse is, to a large extent, independent of AJCC staging (Figure 4F and Figure S3C and S3E). Overall, these results strengthen the specificity of the association of the ISC expression program with the clinical progression of CRC.
Figure 3. Expression of ISC Genes Predict CRC Relapse

(A) GSEA results for crypt signatures in recurrent (Rec) versus nonrecurrent (No Rec) CRC samples.
(B) Kaplan-Meier representation of recurrence-free survival probability over time for patients bearing colon cancers with average high, medium, or low expression of EphB2-derived ISC signature.

---

**Table 1.**

<table>
<thead>
<tr>
<th>All Recurrences</th>
<th>HR</th>
<th>95% CI</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EphB2 ISC (+1 sd)</td>
<td>3.78</td>
<td>(1.06 – 13.42)</td>
<td>0.0452</td>
</tr>
<tr>
<td>AJCC Stage</td>
<td>1.29</td>
<td>(1.01 – 1.65)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Stage 2 vs. 1</td>
<td>3.13</td>
<td>(-1.41 – 13.78)</td>
<td>0.1322</td>
</tr>
<tr>
<td>Stage 3 vs. 1</td>
<td>7.78</td>
<td>(1.86 – 32.49)</td>
<td>0.0049</td>
</tr>
</tbody>
</table>

**Table 2.**

<table>
<thead>
<tr>
<th>Late Rec (&gt;2y)</th>
<th>HR</th>
<th>95% CI</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EphB2 ISC (+1 sd)</td>
<td>21.41</td>
<td>(2.30 – 99.98)</td>
<td>0.0068</td>
</tr>
<tr>
<td>AJCC Stage</td>
<td>1.71</td>
<td>(1.18 – 2.48)</td>
<td>0.0541</td>
</tr>
<tr>
<td>Stage 2 vs. 1</td>
<td>3.13</td>
<td>(-1.41 – 13.78)</td>
<td>0.4624</td>
</tr>
<tr>
<td>Stage 3 vs. 1</td>
<td>7.78</td>
<td>(1.86 – 32.49)</td>
<td>0.0880</td>
</tr>
</tbody>
</table>
Intestinal Stem-like Cells in Colorectal Cancer Relapse

Figure 4. Humanized ISC Signatures Identify Recurrent CRCs

(A) Representation of a normal colonic crypt.
(B) GSEA of each mouse expression signature in human colon crypt-top versus crypt-bottom regions.
(C) GSEA for “humanized” Hu-EphB2-ISC signature comparing recurrent (Rec) with nonrecurrent (No Rec) colon cancers.
(D) GSEA for Hu-EphB2-ISC signature comparing tumors that recurred 2 years or later after therapy (late) with nonrecurrent tumors (No Rec).
(E) log-Hazard Ratio (HR) of recurrence as a smooth function (penalized quartic splines) of the average Hu-EphB2-ISC expression. Dashed lines, 95% confidence intervals.
(F) Upper table: multivariate analysis using Cox Proportional Hazards Model to assess dependency of different tumor variables in the prediction of CRC relapse. Hu-EphB2-ISC expression modeled as a linear effect, as supported by (E). Lower table: multivariate analysis of Hu-EphB2-ISC signature for late CRC relapse (>2 years).

See also Figures S2 and S3 and Tables S2, S3, S4, and S5.

Colorectal Cancers Contain Cells with Either Intestinal Stem Cell-like or Differentiated-like Phenotypes

The above observations indicate that most aggressive CRCs express high levels of ISC-specific genes. This phenomenon is associated to the ability of tumor cells to recur after therapy. We next sought to understand whether the ISC program was associated to the ability of tumor cells to recur after therapy.

Specific subpopulations. We used Lgr5 as a surrogate marker to identify the localization of tumor cells with ISC-like phenotype. By in situ hybridization, we detected Lgr5 mRNA at the base of human colonic crypts (Figure 5B) as well as in a large proportion of tumor samples from CRC patients (Figure 5F). Lgr5 expression was heterogeneous among different tumors. The frequency of positive cells ranged from <1% to 95% (Figure 5F), and staining...
was restricted to cell clusters dispersed throughout the tumor mass (red arrowheads in Figure 5D and Figure S4). Commonly, Lgr5+ cells localized within the most basal tumor cell layers. In tumor regions that displayed glandular organization, Lgr5+ cells resided at sites where the tumor epithelium invaginates into the surrounding stroma (dotted lines in Figure 5D and Figure S4). Remarkably, the morphology of these invaginations was reminiscent of that of normal colon crypts. As exemplified in Figure 5D, some CRCs were constituted by abundant crypt-like structures arrayed around large lumens. Other CRCs contained individual Lgr5+ crypts dispersed throughout the tumor mass (examples in Figure S4). Of note, we could not detect the
expression of Lgr5 in one-third of the samples analyzed (Figure 5F) owing to low sensitivity of ISH technique on archival material. This subset included most poorly differentiated CRCs that consistently showed low levels of Lgr5 (data not shown) even though they were, overall, enriched in expression of ISC-specific genes (Figure 2D and Figure S2E).

We also investigated the expression of intestinal differentiation markers in CRC. Krt20 codifies for a cytokertatin whose levels increase sharply in late progenitor and differentiated cells of normal crypts (Figure 5C, black arrowheads). Interestingly, we observed that the expression patterns of Krt20 and Lgr5 in CRC were mutually exclusive. Staining of serial tumor sections demonstrated that Lgr5+ cells were devoid of Krt20 staining, whereas neighboring Krt20+ cells lacked Lgr5 expression (black arrowheads in Figure 5E and Figure S4). This complementary expression pattern was found in all Lgr5+ tumors analyzed. In tumors with glandular growth patterns, Lgr5 mRNA was detected at the base of crypt-like units, whereas Krt20+ cells resided near luminal spaces (insets in Figures 5D and 5E). Some samples also contained double-negative cells positioned between Lgr5- and Krt20+ cells (white arrowheads in Figure 5H and Figure S4B), suggesting the existence of tumor cells with intermediate phenotypes. Double Lgr5/Krt20-positive tumor cells were rare (<1%) and only present in 8% of samples (data not shown). Twelve out of 14 CRC liver metastases analyzed also contained Lgr5+ tumor cells adjacent to Krt20+ cells that recreated the organization of the primary tumor. Metastatic Lgr5+ CRC cells were also located at the base of crypt-like structures (Figure 5H, red arrowheads) with Krt20+ tumor cells occupying luminal positions (Figure 5H, black arrowheads). From the limited number of marker genes analyzed, we tentatively concluded that a large proportion of CRCs contain cells with two mutually exclusive phenotypes: an ISC-like or a differentiated cell-like phenotype (Figure 5G).

**EphB2 Is a Surface Marker for ISC-like Tumor Cells**

We then investigated whether EphB2 expression could distinguish between ISC-like and differentiated-like cells in CRCs as it does in the normal intestinal mucosa. Figures 6A–6C and Figure S5 show the expression patterns of EphB2, Lgr5, and Krt20 in serial sections of representative CRCs. Most CRCs were composed of cell clusters expressing heterogeneous EphB2 levels ranging from highly positive, to intermediate, to completely negative (Figure 6A, red, blue, and green arrowheads, respectively). In the vast majority of tumors analyzed (n = 30), prominent EphB2 staining coincided with Lgr5 expression (Figures 6A and 6B, red arrowheads, and Figure S5). In glandular tumors, EphB2 and Lgr5 labeled the base of crypt-like structures that grew embedded into the tumor stroma (Figure 6A and 6B, dashed lines). On the contrary, the majority of Krt20+ tumor cells were EphB2lo or EphB2neg in all tumors (Figures 6A and 6C, green arrowheads, and Figure S5). Of note, EphB2 was silenced in areas of high grade as we previously reported (Batlle et al., 2010). Commonly, these areas also showed low levels of Lgr5 expression (data not shown).

We also analyzed the possibility that markers of tumor initiating cells (TICs) (i.e., cd44 and cd133) recognize ISC-like cells in CRCs. Cd44 and Lgr5 expression domains overlapped largely in most tumors (7 out 9 samples analyzed) (Figures S6A and S6B), yet Krt20+ cells were also partially included in the cd44+ domain in two-thirds of the samples (Figure S6). Cd133 antibody stained the apical membrane of luminal tumor cells whereas, in most cases, ISC-like cells were located in basal positions (Figures S6G and S6H). This localization is consistent with that described previously (Shmelkov et al., 2008). However, it is important to mention that cd133 antibodies used for FACS recognize a conformational epitope in TICs (Kemper et al., 2010).

To strengthen the link between the EphB2hi tumor cell population and the ISC-like phenotype, we analyzed EphB2 surface expression by flow cytometry on primary CRC biopsies (Figure 6D). In most tumors, epithelial cells (EpCAM+) were heterogeneous regarding EphB2 surface abundance with relative levels ranging from 1 to 104 (Figure 6D). This distribution resembled that of the normal intestinal epithelium, though different EphB2hi versus EphB2lo cell ratios characterized each tumor. We purified tumor epithelial cells expressing either the highest EphB2 relative levels (top 10%–15% brightest cells) or displaying low/negative EphB2 surface expression. Consistent with the observation on tissue sections, in 5 out of 6 samples analyzed, EphB2hi tumor cells expressed higher levels of ISC-specific genes Lgr5 and Ascl2 than EphB2lo/EphB2neg cells (Figure 6D). On the contrary, EphB2lo/EphB2neg cells were characterized by increased expression of Krt20 as well as of Krt17, another marker of mature colonic cells (Figure 6D). Combined immunophenotyping with EphB2 and cd44 antibodies showed that cd44+ cell populations were enriched in EphB2hi cells in all samples (Figure S7A). Still, EphB2 surface expression distinguished between ISC-like and differentiated-like cells within the cd44+ population (Figure S7B). These results are in agreement with the staining patterns observed in tissue sections and suggest that cd44 marks ISC-like cells albeit with lower specificity than EphB2.

**ISC-like Cells Are Tumor-Initiating Cells and Display Self-Renewal and Differentiation Capacity**

We then sought to determine the tumorigenic potential of ISC-like or differentiated-like tumor cells. As a proof of principle, we purified epithelial tumor cells (EpCAM+) expressing high, medium, or low surface EphB2 levels (Figure 7A) from a stage IV CRC. The relative expression of ISC marker genes Lgr5, Ascl2, Olfm4, and Smoc2 declined from EphB2hi to EphB2lo populations (Figure 7B). Conversely, expression of pan-intestinal differentiation markers Krt20, Krt17 and Ceacam5 was highest in EphB2hi cells and decreased progressively in EphB2med and EphB2lo cells. Importantly, the three EphB2 populations expressed equal levels of proliferation markers such as Ki67 (Figure 7B) and Myc (data not shown).

To test the tumor forming capacity of each EphB2-purified cell population we performed subcutaneous injections into NOD/Scid mice. The three tumor cell populations (EphB2hi, EphB2med, and EphB2lo) showed equivalent viability after FACS measured by propidium iodide and trypan blue exclusion (data not shown). Figure 7C shows tumor-free survival plots for mouse cohorts inoculated with different cell dosages. Mice that did not develop tumors were followed over a period of at least 5 months. EphB2hi, EphB2med, and EphB2lo cells inoculated in high numbers (5000 cells) formed tumors with similar frequency and...
latency. At lower dosages (1000 or 100 cells), EphB2hi cell population retained the capacity to generate tumors with high efficiency whereas EphB2lo/EphB2neg cells displayed reduced or null tumorigenic capacity. EphB2med cells showed an intermediate behavior. Applying the maximum-likelihood estimation method of limiting dilutions assay, we calculated a frequency of tumor initiating cells of 1 in 72 for EphB2hi, 1 in 610 for EphB2med cells and 1 in 2751 for EphB2lo cells (Figure 7F). Therefore, the enrichment in tumor initiating cells paralleled the expression of ISC genes.

EphB2hi cells generated xenografts with histological patterns that largely resembled the primary tumor including the presence of EphB2hi, EphB2med, and EphB2lo cells (representative example is shown in Figures 7D and 7E). Assessment of EphB2 surface abundance by flow cytometry confirmed that engrafted EphB2hi tumor cells gave rise to EphB2med and EphB2lo populations (Figure 7D). Remarkably, the relative expression levels of ISC, proliferation, and differentiation marker genes in the three secondary EphB2 populations were analogous to those of the primary xenograft (Figure 7E). Thus, few inoculated EphB2hi cells effectively propagated the disease in immunodeficient hosts while restoring the patterns of ISC-like and differentiated-like cells of the tumor of origin. To strengthen this conclusion, we tested the tumorigenic potential of different populations of secondary and tertiary EphB2hi-derived xenografts (Figure 7F). The frequency of tumor-initiating cells in each secondary and tertiary cell population did not differ significantly from that calculated in the tumor of origin (p > 0.05 for all comparisons) (Figure 7F). We conclude that ISC-like tumor cells hold high tumor-initiating potential as well as display long-term self-renewal and differentiation capacity.

**DISCUSSION**

The association of stem cell genes with disease outcome in several cancer types is widely accepted. This notion has spread from studies that used expression signatures derived from tumor-initiating cells (Glinsky et al., 2005). These signatures
may reflect the gene program of a particularly aggressive population rather than that expressed by adult stem cells. On the other hand, pluripotency genes expressed by embryonic stem cells (ESCs) have also been generally correlated with poor prognosis in cancer (Ben-Porath et al., 2008). Yet, the correlation of ESC genes with cancer progression has been recently undermined by the finding that a Myc-driven gene program rather than core pluripotency genes accounts for similarities between ESCs and cancer stem cells (Kim et al., 2010). Our work provides unequivocal evidence that the risk of developing recurrent CRC is proportional to the expression in the primary tumor of a set of genes specific of normal ISCs. This finding raises the possibility that tumor cells require ISC-specific functions to effectively regenerate cancer after therapy. The robust association of the ISC signature with cases of late relapse favors the idea that ISC genes endow long-term survival and regeneration potential rather than immediate tissue colonization capacity. A recent work by Di Fiore and colleagues correlates higher grading in breast cancers with higher frequencies of cancer cells expressing a breast stem cell-like phenotype (Pece et al., 2010). Our assessment of ISC gene expression levels may estimate both the relative number of ISC-like cells within a given tumor as well as the overexpression of particular ISC genes co-opted by CRC cells.

The signature of crypt proliferation shows a strong but reverse correlation with CRC recurrence. A similar association was reported independently (Anjomshoaa et al., 2008), and it is in sharp contrast to the results obtained for breast cancer where a high proliferative index indicates poor prognosis (Anjomshoaa et al., 2008). It remains to be determined whether high levels of proliferation genes in CRC modify the efficacy of chemotherapeutic drugs that target the cell division machinery. In any case, a straight interpretation of this result is that the predictive power of ISC genes for relapse does not stem from their role in...
proliferation. On the other hand, we found that expression of several individual markers of intestinal differentiation in the primary tumor associates with good prognosis. The late TA signature is also enriched in genes that inversely predicted disease relapse yet with borderline significance. We speculate that this may be due to incomplete or aberrant expression of the differentiation program by CRC cells.

In tumor tissue, ISC and differentiation programs are expressed with mutually exclusive patterns that resemble those of the normal intestinal mucosa. We identify a subpopulation of tumor cells with an ISC-like phenotype that localizes in close proximity to cells that express intestinal differentiation genes. In well- and moderately differentiated CRCs, these two cell types reside within defined structures reminiscent of intestinal crypts. In healthy intestinal mucosa, high EphB2 levels mark cells at the crypt base (i.e., ISCs), whereas the ISC progeny gradually silences EphB2 expression as it migrates away from the stem cell niche. The EphB2 expression gradient is required to compartmentalize cell populations in different territories of the cell niche. The EphB2 expression gradient is required to compartmentalize cell populations in different territories of the cell niche. (Batlle et al., 2002). FACS analysis demonstrates that CRCs display a graded expression of EphB2 reminiscent of that of intestinal crypts. ISC-like tumor cells are EphB2hi and reside closer to the stroma whereas differentiated-like cells are low or negative for EphB2 expression and face the luminal sides of tumor glands. Thus, like in normal epithelium, the identity and position of ISC-like and differentiated-like cells within the tumor correlates with different EphB2 surface levels. This pattern is common to a large proportion of CRCs except for poorly differentiated tumors that express low EphB2 levels as we previously reported (Batlle et al., 2005). Mouse colon tumors engineered to lack EphB signaling acquire a nonglandular growth pattern characterized by loss of cell polarity and an aggressive phenotype (Batlle et al., 2005; Cortina et al., 2007). Thus, EphB2 silencing in ISC-like cells may contribute to the loose architecture of poorly differentiated CRCs. Besides the downregulation of EphB2, GSEA demonstrates that poorly differentiated CRCs retain an overall ISC-like phenotype compared with well-differentiated tumors. It would be beneficial to identify additional ISC marker genes that label the EphB2neg tumor stem cell population in poorly differentiated CRCs.

A central question that arises from our observations is how the ISC and differentiated programs are maintained in CRC. A recent report demonstrates that tumor cells grown in vitro as spheroids display different amounts of β-catenin/Tcf signaling that, in turn, regulate the expression of stem cell or differentiation markers (Vermeulen et al., 2010). Also, in some CRCs, strongest nuclear β-catenin accumulation corresponds to tumor cells positioned closer to the tumor stroma (Brabletz et al., 2001). This localization is consistent with the Lgr5 expression domain reported here. Thus, differential Wnt signaling may account for the establishment of ISC-like or differentiated-like populations in a subset of CRCs. Other pathways involved in the balancing of the differentiation and stem cell programs in normal crypts such as Notch or BMP signaling may potentially regulate the relative numbers of ISC- and differentiated-like cells in tumors. We have shown that primary tumors and metastasis share an equivalent distribution of ISC-like and differentiated-like cell populations. Apparently, the disseminated cell that propagates the disease to the metastatic site reconstitutes the populations and organization of the tumor of origin. ISC-like tumor cells are candidates to represent tumor propagating cells in CRC as they generate tumors in immunodefficient hosts with high efficiency compared with differentiated-like cells. In addition, tumors generated from ISC-like cells recapitulate the organization of the tumor of origin. These two features parallel the self-renewal and differentiation properties of normal ISCs. Our results also indicate that the previously identified cd44+ tumor-initiating cell population is partially enriched in ISC-like cells. Overall, these discoveries strengthen the view of a hierarchical organization of CRC cells similar to that present in normal crypts and suggest that ISC-like cells are the tumor cell population that institutes recurrent cancers.

Experimental Procedures

Histological Procedures
Detailed protocols for Immunohistochemistry, in situ hybridization, and immunofluorescence are provided in Supplemental Information.

Intestinal Crypt Cell Purification and In Vitro Growth
A detailed protocol for the purification of crypt cell types according to EphB2 expression is available in the Supplementary Information section. In vitro growth of intestinal organoids and FACS purification of Lgr5-GFPhi or Lgr5-GFPlo cells were performed as previously described (Sato et al., 2009).

Generation of Crypt Cell Gene-Expression Signatures
We measured gene-expression levels of three EphB2 selected populations in two independent experiments using Affymetrix mouse302 arrays. For a detailed description of the statistical method to define each crypt expression signature, please refer to Supplemental Information.

CRC Patient Gene Expression Data Sets and Association of Crypt Expression Signatures with Clinical Outcome
At the moment of performing this study, there were six publicly available CRC gene expression datasets annotated with clinical information in the Gene Expression Omnibus at NCBI (GSE17538, GSE10402, GSE12945, GSE5206, GSE12032 and GSE14333). We compared the robustness and accuracy of these data sets (see Supplementary Information) and selected datasets GSE17538 (Smith et al., 2010) and GSE14333 (Jorissen et al., 2009) for further analysis. We combined both cohorts for a total of 340 unique CRC expression profiles. For descriptive statistics as well as association of major clinical progression parameters and clinical outcome for this cohort, please refer to Supplemental Experimental Procedures and Tables S3 and S4. Detailed descriptions of all statistical analysis employed in this manuscript are included in Supplemental Information.

Analysis of Mouse Signatures in Human Colon Crypt Top and Bottom Compartments
We used the GSE6894 dataset (Kosinski et al., 2007), which contains expression profiles of microdissected human colon crypt top and bottom compartments (n = 9 pairs from different mucosa samples). We refined the ISC, Lgr5, and proliferation signatures by selecting genes differentially expressed in human colon crypts (crypt-bottom > crypt-top for EphB2-ISC, Lgr5-ISC, and proliferation signatures and crypt-top > crypt-bottom for Late TA genes, >1.5-fold in all cases).

Tumor Disaggregation and Sorting of EphB2 Cell Populations and Assays of Tumor Initiation
Primary tumor samples were obtained from Hospital del Mar (Barcelona, Spain) or Hospital de la Santa Creu i Sant Pau (Barcelona, Spain). Detailed protocols for tumor disaggregation and tumor cell purification are described under Supplemental Experimental Procedures. Different EphB2-purified cell fractions were injected s.c. into NOD/Scid mice. Tumor growth was monitored.

Cell Stem Cell

Intestinal Stem-like Cells in Colorectal Cancer Relapse
over 5 months. Frequency of tumor-initiating cells was estimated according to published methods (Hu and Smyth, 2009).

ACCESSION NUMBERS

Microarray data are accessible at the Gene Expression Omnibus (GEO) database under accession number GSE27605.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, five tables, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.stem.2011.02.020.

ACKNOWLEDGMENTS

We thank Genentech, Inc. for providing anti-EphB2 antibody Mab 2H9, Evarist Planet (IRB, Barcelona) for assistance with statistical methods, H.G. Palmer (Vall d’Hebron Hospital) for support, M. Trias (Hospital de la Santa Creu i Sant Pau) for tumor samples, A. Calon for assistance with bioinformatics, IRB Functional Genomics Facility (IRB) for technical assistance in microarray hybridization experiments, A. Rodolosse for help with ISHs, J. Comas and R. Alvarez (UB, Barcelona) for expert technical assistance with the FACS, I. Joval for figure artwork, C. Nadal and X. Garcia (Clinic Hospital, Barcelona) for help with samples and discussions, and to all members of the Battle Lab for support. A.M.-S. holds an AECC postdoctoral fellowship, and F.M.B. holds an IRB/La Caixa PhD fellowship. This work has been supported by grants to E.B. from ERC (FP7-EU) and MICINN (Plan Nacional I+D+i and Consolider programs) and by grants PS09/00965 (MICINN) and NanoCoMets (CIBER-BBN) to R.M.

Received: April 26, 2010
Revised: January 10, 2011
Accepted: February 28, 2011
Published online: March 17, 2011

REFERENCES


