Molecular properties of side population-sorted cells from mouse small intestine

Ajay S. Gulati,1 Scott A. Ochsner,2 and Susan J. Henning1,3
1Department of Pediatrics, 2Huffington Center on Aging, and 3Department of Molecular and Cellular Biology, Baylor College
of Medicine, Houston, Texas

Submitted 14 September 2007; accepted in final form 9 November 2007

Gulati AS, Ochsner SA, Henning SJ. Molecular properties of side population-sorted cells from mouse small intestine. Am J Physiol Gastrointest Liver Physiol 294: G286–G294, 2008.—The high rate of turnover of the intestinal epithelium is maintained by a group of stem cells that reside at the base of the crypts of Lieberkuhn. Whereas the existence of these intestinal epithelial stem cells has been well established, their study has been limited due to the inability to isolate them. Previous work has utilized side population (SP) sorting of the murine small intestine to isolate a viable fraction of cells enriched for putative intestinal epithelial stem cells. In the present study, we have used gene expression profiling techniques to characterize the molecular features of this potential stem cell population. Further in situ hybridization studies reveal that transcripts enriched in the SP tend to localize to the intestinal crypt base/progenitor cell zone, while de-enriched transcripts localize outside of this region. From a functional standpoint, gene ontology and pathway mapping analyses demonstrate that immune, mesenchymal, and differentiated epithelial cells are depleted in the SP fraction, while putative progenitor cells are enriched in this cell population. Furthermore, the significance of the maturity onset diabetes of the young pathway in these cells suggests that enteroendocrine progenitors are enriched in this cell fraction as well. In conclusion, SP sorting of mouse small intestinal mucosa does appear to isolate cells with progenitor characteristics. These findings provide the foundation for membrane protein-based sorting procedures that can be used to further fractionate these cells for transplantation experiments in the future.

The surface area of the adult human intestine is ~200 m², nearly 100 times that of the external surface of the body. The epithelial lining that covers this vast surface turns over every 3–4 days (30), faster than any other in the mammalian organism. This high rate of turnover is maintained by a group of intestinal epithelial stem cells (IESCs) located deep in the crypts of Lieberkuhn (11, 36). These cells give rise to the four differentiated cell lineages of the small intestinal epithelium: enterocytes, Paneth cells, goblet cells, and enteroendocrine cells (7). Because many diseases of the gastrointestinal tract cause significant damage to the intestinal epithelium, there is a pressing need to design treatment strategies that can harness the restorative capacity of these multipotent cells.

To date, a comprehensive study of the IESCs has been hampered due to the lack of tools needed to identify and isolate these cells. Of particular significance is the paucity of specific markers for IESCs. Various groups have identified an array of candidate genes, but a definitive marker has not come to bear. For example, genes such as EphB2 (5), EphB3 (5), CD44 (6), Fgfr3 (45), and Sox9 (8) have a crypt-base expression pattern, but are not specific to IESCs. Analyses of various Wnt pathway components and their target genes have identified Sfrp5 (19) and Ascl2 (44) as possible IESC markers, while study of bone morphogenetic protein signaling in the gut has shown that phospho-phosphatase and tensin homolog deleted on chromosome 10 and phospho-AKT are expressed in the supra-Paneth cells of the intestinal stem cell zone (20). Musashi-1 (Msi1), an RNA-binding protein involved in Notch signaling, is also expressed in these supra-Paneth cells, as well as crypt base columnar cells of the stem cell zone (26, 37). Unfortunately, none of these genes encode for proteins that are easily amenable to cell sorting procedures. Because such an antibody-mediated approach is not possible, investigators have had to turn to more indirect, marker-independent techniques to isolate and further study the IESCs.

One technique that can be used to isolate cells without the use of specific markers is laser capture microdissection (LCM). LCM has been used to capture putative IESCs based on their known position within the crypt (42). These studies were done using a mouse strain genetically depleted of Paneth cells to increase the fractional representation of presumed stem cells within the crypt base. Further gene expression analyses of these putative progenitors found numerous transcripts involved in c-myc signaling to be significantly expressed (42). A subsequent study by the same group described the construction and sequencing of an extensive cDNA library generated from these laser captured, putative IESCs (17). Functional genomics analyses of this library found numerous networks, including the Wnt/β-catenin and phosphoinositol-3/Akt kinase pathways, to be significantly represented in these cells. Further immunohistochemical studies showed the proteins Dcamk1l and Mapk14 to be expressed within the stem cell zone (17). Again, neither Dcamk1l nor Mapk14 are cell surface markers; therefore, the isolation of viable cells using antibodies to these proteins is precluded.

The studies described thus far have given us some insight into the biology of IESCs, but have not led to the isolation of viable cells that can be further studied using in vitro or in vivo experimental models. As such, it has not been possible to evaluate the capacity of isolated cells to proliferate and differentiate, ultimately showing they are truly stem cells. To date, the report by Dekaney et al. (12) from our laboratory remains the only published study describing the successful isolation of a viable cell population enriched for putative IESCs. This study...
utilized a technique known as side population (SP) sorting to isolate presumptive IESCs based on their ability to efflux the DNA-binding dye Hoechst 33342. SP sorting was originally described to isolate bone marrow–derived stem cells (18), but has since been applied to isolate stem cells from various other tissues, such as skeletal (3) and cardiac muscle (33), liver (40, 47), and breast (2, 46). Using SP sorting, Dekaney et al. (12) have isolated a population of cells enriched for the IESC marker Msi1. These intestinal SP cells are also cytokeratin positive, do not express markers typically found on bone marrow–derived stem cells, and remain viable for up to 14 days (12). Such results support the idea that SP sorting does enrich for a viable population of IESCs. In this study, we have sought to further characterize the molecular features of this intestinal SP cell fraction. Using gene expression profiling, we have identified numerous transcripts differentially expressed in the SP relative to intact jejunum. Further analyses of these transcripts using in situ hybridization (ISH) and functional genomics have allowed us to explore the biology of these cells and gain a better understanding of the intestinal SP cells at a molecular level.

METHODS

Isolation of intestinal SP cells. All mice for this study were housed in our animal facility under a 12:12-h light-dark cycle, and all had unrestricted access to rodent Lab Chow no. 5001 (PMI Nutrition International, Brentwood, MO), as well as acidified tap water. The Institutional Animal Care and Use Committee of Baylor College of Medicine approved all experimental procedures. Putative IESCs were isolated from the jejunum of adult male C57Bl/6 mice using dual argon laser fluorescence activated cell sorting, as described by Dekaney et al. (12). For each experiment, a total of ~30 mg of pooled, intact jejunal tissue was first collected from three animals and stored at 4°C in RNA later (Ambion, Austin, TX). Mucosal digests were next prepared from the remaining pooled jejunum using collagenase and disperse. These were then subject to SP sorting via Hoechst 33342 dye exclusion. Cells were also stained with fluorescein isothiocyanate (FITC) labeled anti-CD45 antibody (BD Pharmingen, San Diego, CA) to eliminate bone marrow–derived cells from the populations to be analyzed. Propidium iodide was used to exclude dead cells. Both CD45-negative SP (CD45−/SP) and CD45-negative non-SP (CD45−/nSP) cell fractions were collected from each sort for RNA extraction. Therefore, three experiment groups were obtained for each sort: CD45−/SP, CD45−/nSP, and intact jejunum. Ultimately four sorts were performed, thus providing four biological replicates for each experimental group.

RNA isolation and quality control. CD45−/SP and CD45−/nSP fractions were collected and spun for 5 min at 700 g to pellet the cells. Total cellular RNA was extracted using an RNAeasy Mini Kit (Qiagen, Valencia, CA) per the manufacturer’s instructions. Additional steps completed in the protocol included homogenization of the lysed sample using a QIAshredder spin column (Qiagen) and treatment with RNase-free DNase I (Qiagen). To isolate total RNA from intact jejunum specimens, the stored tissue was homogenized in a guanidine thiocyanate–containing lysis buffer (Qiagen) using a 2-ml Dounce homogenizer (Kontes Glass, Vineland, NJ). This was followed by spinning the sample through a QIAshredder column. The remainder of the extraction was performed using the RNAeasy Mini Kit per the manufacturer’s instructions, including treatment with RNase-free DNase I. To ensure adequate quality of the total RNA before microarray analysis, each sample was evaluated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) following the manufacturer’s instructions. Samples were deemed acceptable if they showed distinct 18S and 28S rRNA peaks on bioanalyzer electropherograms, indicating minimal RNA degradation. Representative electropherograms are shown in supplemental Fig. 1. (The online version of this article contains supplemental data.)

Microarray analyses. Once adequate quality of the total RNA was ensured, a 50-ng aliquot of each sample was submitted to the Baylor College of Medicine Microarray Core Facility for processing. As stated, the three experimental groups for this study were the CD45−/SP cells, CD45−/nSP cells, and intact jejunum, each with four biological replicates; therefore, 12 RNA samples were obtained and submitted for microarray analysis. Each sample underwent two rounds of in vitro transcription using the Affymetrix Two Cycle Target Labeling Kit (Affymetrix, Santa Clara, CA), according to standard Affymetrix protocols. Biotinylated cRNAs were prepared from 5 μg of total RNA. Following fragmentation, 15 μg of biotinylated cRNA from each sample were hybridized for 16 h at 45°C to Affymetrix Mouse Genome 430 2.0 GeneChips. GeneChips were washed and stained using the Affymetrix Fluidics station 400, according to standard Affymetrix protocols. Laser scanning of the chips was performed using the Affymetrix GeneChip Scanner 3000. The image files were analyzed for probe intensities and converted to tabular formats (CEL files) using the Microarray Suite Expression Analysis software (version 5.0) from Affymetrix.

Initial GeneChip quality control analyses were done using the dChip (version 2005) software platform (29). All array images were manually examined to ensure that there were no significant baseline problems due to contamination or improper hybridization. Chips were subsequently normalized to the array with the median overall intensity using the invariant set normalization method, and probe set expression values were calculated using the perfect match/mismatch difference model (29). The data set was filtered using a coefficient of variance between 0.7 and 1.00 to identify genes that were changing across the sample set. Hierarchical clustering was performed using a distance measure of 1 minus the Pearson correlation coefficient between samples coupled to complete agglomeration using the R stats package. To lend further support to the integrity of the data, the Affycoretools package from Bioconductor (16) was used with the R statistical software platform (22) to assess array quality and perform principal component analyses with robust multiarray normalized gene expression values (23). Finally, dChip was utilized to identify transcripts whose levels were differentially expressed in the CD45−/SP cells by at least twofold relative to intact jejunum, with a P value ≤ 0.05, as calculated by Student’s t-test. This fold change cutoff was chosen based on previous studies that utilized microarray analyses to evaluate gene expression in the intestinal tract (4, 42). The data discussed in this publication have been deposited in NCBI’s Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE9013.

ISH. For ISH, intestinal tissue from adult male C57Bl/6 mice was collected and flushed with ice-cold phosphate buffered saline (Invitrogen, Carlsbad, CA). A blunt syringe was then used to fill the lumen with optimum cutting temperature (OCT) compound (Tissue-Tek, Santa Cruz, CA), and the tissue was subsequently cut into sections ~1.5 cm in length. The segments were next aligned in parallel in a Cryomold (Tissue-Tek) that had been filled with a thin layer of OCT compound. Additional OCT compound was added to cover the tissue and then frozen in Cytocool II (Richard-Allen Scientific, Kalamazoo, MI) on dry ice. Cryostat sectioning was performed at ~16°C to generate 16-μm sections, as described by Yaylaoglu et al. (48). Digoxigenin-labeled riboprobes for ISH were generated using in vitro transcription, as previously described (48). Sequence data for these probes can be obtained from the Allen Brain Atlas (www.brainmap.org) (28). Probe detection was accomplished by catalyzed reporter deposition using biotinylated tyramide with subsequent colorimetric detection of biotin using an avidin-alkaline phosphatase conjugate (9, 48). The final step results in a dark blue precipitate, which is proportional in abundance to the number of detected transcripts (9, 48).
Functional genomics analyses. Gene ontology (GO) analyses were performed using the Database for Annotation, Visualization and Integrated Discovery, which can be accessed at http://niaid-abcc.ncifcrf.gov (13, 21). To determine GO term enrichment, the proportion of test genes that mapped to a particular GO term was compared with the proportion of genes from the entire GeneChip that mapped to the same term. Statistical significance was calculated using a modified Fisher exact test. Terms with P values ≤ 0.01 were considered to be significantly represented. Initial GO analyses were performed by evaluating the complete list of CD45+/SP enriched transcripts using the “biological processes” arm of the GO. From the extensive list of descriptors generated, the most relevant terms were selected as follows. First, the 100 most highly enriched transcripts in the CD45+/SP cells were analyzed, again using the “biological processes” arm of the GO. Those GO descriptors that proved to be significant for both the enriched and highly enriched gene lists were selected. Second, additional GO analyses were performed using the “cellular component” and “molecular function” arms of the GO. This allowed for the selection of identifiers significantly represented in multiple arms of the GO. Similar analyses and criteria were used to identify the most relevant CD45−/SP deenriched GO terms.

Pathway mapping analyses were performed using the Kyoto Encyclopedia of Genes and Genomes, which can be found at www.genome.jp/kegg. The Kyoto Encyclopedia of Genes and Genomes Pathway database contains pathways representing current knowledge regarding metabolism, cellular processes, and human disease. The Database for Annotation, Visualization and Integrated Discovery system was again used to determine enrichment of various pathways within our gene list. P values were calculated in the same manner as described above for the GO.

RESULTS

SP sorting of small intestinal mucosal digests reveals SP and nSP cell fractions with distinct gene expression profiles. The use of SP sorting to isolate putative IESCs has been previously described by Dekaney et al. (12). This study showed that SP sorting can be used to isolate distinct populations of cells from intestinal mucosa digested with collagenase and dispase. As shown in Fig. 1A, we have successfully reproduced this technique and have isolated both SP and nSP fractions from such digests. Further purification using FITC-labeled anti-CD45 antibodies (Fig. 1B) allowed for the exclusion of bone marrow-derived CD45-positive cells from further analyses. Total RNA samples from the CD45−/SP and nSP fractions were then subject to microarray analyses, as described above.

The integrity of the microarray expression data was evaluated using hierarchical clustering and principle component analyses. These tools represent examples of unsupervised clustering and are meant to ensure that biological replicates are similar to one another, while experimental groups are discernibly different from one another. As shown in Fig. 2A, hierarchical clustering analysis reveals three experimental groups, corresponding to the intact jejunum, CD45−/nSP, and CD45−/SP fractions. As expected, the CD45−/nSP and SP fractions are more closely related to one another than they are to the intact jejunum. Each group also contains four subbranches representing the four biological replicates for each group. Similar results are depicted in the principal components analysis (Fig. 2B), which again identifies three experimental groups, each with four biological replicates. These analyses demonstrate that SP sorting is a reliable technique that reproducibly yields distinct fractions of CD45−/SP and nSP cells.

Identification of transcripts differentially expressed in the CD45−/SP cells. Because the CD45−/SP fraction has been previously suggested to include putative IESCs (12), further analyses in the present study are focused on this particular group of cells. Using the dChip software platform, the gene expression profile of the CD45−/SP cells was compared with that of the intact jejunum. Intact jejunum was chosen for this purpose because RNA from this source can be reliably obtained from any laboratory, thus facilitating further comparisons in the future. When restricted by a fold change greater than or equal to ±2, and a P value of <0.05, 2,679 probe sets...
were found to be differentially expressed in the CD45−/SP cells relative to intact jejunum. Of these, 1,572 (supplemental Table 1) were enriched in the CD45−/SP relative to intact jejunum, while 1,107 (supplemental Table 2) were found to be deenriched. These lists provide the substrate for downstream analyses such as ISH, GO, and pathway mapping to better characterize the CD45−/SP cell fraction.

**ISH of selected transcripts shows localization of the CD45−/SP cells to the stem cell zone.** Figure 3 shows ISH results for representative transcripts differentially expressed within the CD45−/SP. Notch1, a transcription factor known to be important in intestinal cell fate decision making (41), is enriched approximately fourfold in the CD45−/SP cell fraction. Fgfr3, which has also been shown to play a role in intestinal stem cell biology (45), is similarly enriched roughly fourfold in the CD45−/SP cells. Both of these transcripts are rather distinctly localized to the intestinal crypt base/stem cell zone (Fig. 3, A and B). Converse localization is seen for transcripts deenriched in the CD45−/SP (Fig. 3, C and D). The transcription factor Mtf1 plays a critical role in the metal-inducible transcriptional activation of metallothionein and other genes involved in metal homeostasis (38). Its mRNA...
expression is fourfold less in the CD45−/SP cells relative to intact jejunum. As seen in Fig. 3C, Mtfl expression is localized primarily to the villi, outside the intestinal stem cell zone. Figure 3D shows ISH results for the calmodulin regulator Pcp4, with expression restricted to the muscular layers of the gut. As expected, it too is deenriched in the CD45−/SP cells, by approximately fourfold. We have gone on to perform ISH for a total of 48 selected transcripts differentially expressed in the CD45−/SP (Table 1). Of these, 36 are enriched in the CD45−/SP cells relative to intact jejunum, while 12 are deenriched. Of the 36 enriched transcripts, 32 (89%) localize within the intestinal crypt base/stem cell zone. In contrast, 10 (83%) of the 12 deenriched transcripts localize to areas outside the crypt base/stem cell zone, such as the villi, smooth muscle, and lamina propria. These results clearly indicate that SP sorting does enrich for cells that are found within the intestinal stem cell zone.

Table 1. Localization of CD45−/side population enriched and deenriched transcripts by in situ hybridization

<table>
<thead>
<tr>
<th>Enriched (n = 36)</th>
<th>Deenriched (n = 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crypt base/stem cell zone</td>
<td>Villus</td>
</tr>
<tr>
<td>Lhx1</td>
<td>Mrg1</td>
</tr>
<tr>
<td>1700086119Rik</td>
<td>Slc25a22</td>
</tr>
<tr>
<td>Neurog3</td>
<td>Slc4a4</td>
</tr>
<tr>
<td>Slc18a2</td>
<td>Treh</td>
</tr>
<tr>
<td>Isl1</td>
<td>Kirf15</td>
</tr>
<tr>
<td>Neurod1</td>
<td>Slc15a3</td>
</tr>
<tr>
<td>Pax6</td>
<td>CytoF</td>
</tr>
<tr>
<td>Sytl3</td>
<td>Lamina propria</td>
</tr>
<tr>
<td>Ssbp2</td>
<td>Reln</td>
</tr>
<tr>
<td>Vim</td>
<td>Paneth cells</td>
</tr>
<tr>
<td>Scarb1</td>
<td>1810030j14Rik</td>
</tr>
<tr>
<td>Slc20a1</td>
<td>Muscle</td>
</tr>
<tr>
<td>Slc38a2</td>
<td>Pcp4</td>
</tr>
<tr>
<td>Pde4b</td>
<td>Crypt base/stem cell zone</td>
</tr>
<tr>
<td>Chga</td>
<td>Birc5</td>
</tr>
<tr>
<td>Cldna</td>
<td>Top2a</td>
</tr>
<tr>
<td>Sox4</td>
<td>Pde1c</td>
</tr>
<tr>
<td>Mdm2</td>
<td>Mtf1</td>
</tr>
<tr>
<td>Slc12a8</td>
<td>Neurog3</td>
</tr>
<tr>
<td>Socs2</td>
<td>Slc25a22</td>
</tr>
<tr>
<td>Ets1</td>
<td>Top2a</td>
</tr>
<tr>
<td>Cck</td>
<td>Trpc4</td>
</tr>
<tr>
<td>Junb</td>
<td>Rela</td>
</tr>
<tr>
<td>Sox9</td>
<td>1110061a14Rik</td>
</tr>
<tr>
<td>Fgf3</td>
<td>Sort1</td>
</tr>
<tr>
<td>Notch1</td>
<td>Upper crypt/lower villus</td>
</tr>
<tr>
<td>Skl29a1</td>
<td>Ldlr</td>
</tr>
<tr>
<td>Trpc4</td>
<td>Zdhhc13</td>
</tr>
<tr>
<td>Rela</td>
<td>Villus</td>
</tr>
<tr>
<td>1110061a14Rik</td>
<td>Ephb3</td>
</tr>
<tr>
<td>Sort1</td>
<td>Scattered muscle/nerve cells</td>
</tr>
<tr>
<td>Upper crypt/lower villus</td>
<td>Ctnb1</td>
</tr>
</tbody>
</table>

Localization of transcripts (n) is based on highest intensity of expression signal by in situ hybridization.

GO analyses of transcripts differentially expressed in the CD45−/SP. Once the location of the CD45−/SP cells was established, we sought to further explore the biology of these cells. Using the controlled vocabulary of the GO, biological processes significant in the CD45−/SP enriched (complete list) and highly enriched (top 100) transcripts were identified. Complete results of this analysis can be found in supplemental Table 3. Figure 4A shows selected biological processes enriched within the CD45−/SP cell fraction. Selection criteria for these processes are delineated in the METHODS section. The most significant GO term depicted in Fig. 4A is “regulation of transcription from RNA polymerase II.” Transcription in general seems to be a key process in the CD45−/SP cell fraction, as GO descriptors involved in this function appear on both the enriched and highly enriched GO lists. Interestingly, the biological process of “exocytosis” is next in significance within the CD45−/SP cell fraction. This differentiated cell function contrasts with the biological processes of “cell differentiation” and “neurogenesis,” which are also significantly enriched within the CD45−/SP cells, and describe more undifferentiated cells. Finally, the biological process of “phosphorus metabolism” is also shown in Fig. 4A, indicating a high degree of secondary messenger signaling in the CD45−/SP cells. When taken as a whole, these data provide evidence that the CD45−/SP cell fraction is heterogeneous, but is enriched for transcriptionally active cells with stemlike characteristics (i.e., differentiation potential).

Having identified biological processes enriched in the CD45−/SP cells, a similar analysis was next performed for transcripts deenriched (complete list) and highly deenriched (bottom 100) in this cell population (supplemental Table 4). Figure 4B depicts selected processes deenriched in the CD45−/SP cells. The most significant of these GO descriptors is the term “immune response,” which is decidedly significant on both the deenriched and highly deenriched GO lists. Similarly, the biological process “hemopoietic or lymphoid organ development” is also deenriched in the CD45−/SP cell fraction, indicating that immune-type cells are likely depleted from this cell population. The biological process of “mitosis” is next in significance in the CD45−/SP deenriched transcripts. As the most mitotically active cells in the small intestinal epithelium are the rapidly dividing transit amplifying cells, it is likely these too are relatively depleted from the CD45−/SP cell fraction. Finally, the biological processes “inorganic anion transport” and “collagen catabolism” are also deenriched in the CD45−/SP cells. These findings respectively support the depletion of absorptive cells and collagen-producing mesenchymal cells from the CD45−/SP cell fraction. In summary, the deenriched biological processes illustrated in Fig. 4B provide evidence for the exclusion of various immunological, mesenchymal, and differentiated absorptive cells from the CD45−/SP cell fraction.

Molecular pathways differentially expressed in the CD45−/SP cell population. While the GO analyses described above provide some biological insight into the CD45−/SP cells, further understanding of their function is gained by identifying key molecular pathways active in this cell population. When the complete list of transcripts enriched in the CD45−/SP cells is subject to pathways analysis, the MAPK signaling pathway emerges as the only significant pathway enriched in the CD45−/SP fraction (P value = 3 × 10−5). The
function of this pathway is consistent with the enriched descriptors of “transcription” and “phosphorus metabolism” identified in our GO studies. In contrast, analysis of the highly enriched probe sets expressed in the CD45−/SP cells identifies the maturity onset diabetes of the young (MODY) pathway as the only significant pathway ($P$ value $= 4 \times 10^{-5}$) in this group of transcripts. Additional evidence supporting the importance of these pathways in the CD45−/SP cell fraction is revealed by analysis of the transcription factors enriched in these cells. As shown in Fig. 4A, the process of transcription is significantly enriched in the CD45−/SP cell fraction. When pathway mapping is applied to the transcription factors of this cell population, both the MODY and MAPK pathways are again found to have high levels of significance, with $P$ values of $8 \times 10^{-5}$ and $3 \times 10^{-4}$, respectively.

To complete our pathway mapping studies, similar analyses were performed for transcripts deenriched in the CD45−/SP cell fraction. These analyses reveal ECM receptor interaction ($P$ value $= 0.001$), Type 1 diabetes mellitus ($P$ value $= 0.002$), and focal adhesion ($P$ value $= 0.006$) pathways to be deenriched in the CD45−/SP cells. The majority of genes included in the Type 1 diabetes mellitus pathway are immune system related and involved in the generation of auto-antibodies to pancreatic islet cells. The other two pathways, ECM receptor interaction and focal adhesion, both involve cellular interactions with the extracellular matrix. Such findings lend further support to the depletion of immunological and mesenchymal cells from the CD45−/SP cell fraction.

**DISCUSSION**

The small intestinal epithelium is a complex tissue capable of very diverse biological processes. It must function as a mechanical and immunological barrier to harmful antigens, while concurrently allowing for the digestion and absorption of essential nutrients. When damaged, these functions must be restored promptly to maintain the well-being of the organism. The process of epithelial repair involves three stages: restitution, epithelial cell proliferation, and epithelial cell differentiation (14). While restitution is mediated by fully differentiated epithelial cells and is not dependent on the production of new cells (32), the latter two processes rely on adequate stem cell function. This concept underlies the success of epithelial stem cell transplantation in tissues such as the cornea (35) and the epidermis (34). Because of its rapid rate of turnover, the small intestinal epithelium is a strong candidate for regenerative stem cell therapies; however, if we are to use this therapeutic tool for gastrointestinal disease, understanding the intestinal stem cell at a molecular level is imperative.

Numerous investigators have utilized classical molecular biology techniques to study the role of individual genes and proteins as they pertain to intestinal stem cells. While such methods have given us some insight into the function of these cells, they have not allowed for a comprehensive analysis of IESCs. For this reason, we have approached the study of the SP-derived putative IESCs using gene profiling techniques that allow for the global analysis of multiple transcripts simultaneously. This approach demonstrates the utility of intestinal SP sorting as a reproducible technique that yields very distinct populations of cells. Furthermore, our extensive ISH studies and functional genomics analyses provide a comprehensive characterization of the molecular features of this putative stem cell population.

First described as a means to isolate hematopoietic stem cells (18), SP sorting has now been used to extract putative stem cells from numerous solid tissues (10), including the small intestine (12). As with any isolation technique, the ability to yield reproducible cell fractions is essential. The difficulty of extracting quality total RNA from cells that have undergone prolonged manipulation lends further complexity to the reproducibility of downstream transcript analyses. Despite these
challenges, our results show that intestinal SP sorting can be used as a reliable cell isolation technique. The similarity of the flow cytometry plots shown in Fig. 1 to those previously described (12) supports the reproducibility of the SP sorting procedure. Furthermore, the unsupervised clustering analyses illustrated in Fig. 2 reveal that the intact jejunum, CD45−/SP, and CD45−/inSP cell fractions have distinct gene expression profiles that distinguish these sample groups from one another. These analyses also show that sample biological replicates tend to group together, again reinforcing the reproducibility of this technique.

After establishing the reliability of the SP sorting process, the biological properties of the CD45−/SP cells could be explored. As described, this particular fraction has previously been shown to be enriched for Msi1 expression and is thus believed to include the putative IESCs (12). For our analysis, we chose to focus on gene expression differences in the CD45−/SP cells relative to intact jejunum. We feel this particular comparison is important for two reasons. First, the validation of our microarray studies relies on ISH, which, by its nature, compares the specific location of transcripts to an entire cross section of whole tissue. Thus using intact jejunum for our comparisons allows us to utilize the same denominator for both microarray and validation studies. Second, maintaining intact jejunum as our reference tissue allows for the comparison of our data with other groups who might utilize alternative methods for procuring preparations of putative IESCs. This is important for further validation of the CD45−/SP cell dataset.

Once differentially expressed transcripts in the CD45−/SP cells were identified, we utilized these data to determine the location of the CD45−/SP cells along the crypt-villus axis. Because the architecture of the crypt-villus unit is highly conserved, the location of a cell along this axis is very informative. The very base of the crypt is inhabited by a group of Paneth cells, while the IESCs are believed to reside just above, and possibly in between, this Paneth cell population (7, 11, 36). These IESCs give rise to a group of transit amplifying cells, which, in turn, give rise to daughter cells that migrate up the crypt and onto the villus, differentiating as they go. Our ISH results clearly show that the majority of transcripts enriched in the CD45−/SP cells localize to the crypt base/progenitor cell zone, while CD45−/SP deenriched transcripts tend to fall outside of this zone. These findings support the notion that the CD45−/SP fraction contains a population of putative progenitor cells.

After identifying the location of the CD45−/SP cells, we next utilized a functional genomics approach to further characterize the molecular features of these cells. Focusing on transcripts deenriched in the CD45−/SP fraction, as evidenced by the enrichment of transcripts such as Defcr22 (Δ 4-fold), which encodes a cryptdin expressed in Paneth cells. This is of particular importance, as the crypt base location of the CD45−/SP cells could support the inclusion of Paneth cells in this cell fraction. However, the deenrichment of Defcr22 in our microarray analyses implies that, as previously concluded by Dekaney et al. (12), Paneth cells are relatively excluded from the CD45−/SP. Finally, rapidly cycling transit amplifying cells also appear to be depleted from the CD45−/SP cell fraction, as evidenced by the significance of the term “mitosis” on the deenriched GO list.

Having described the types of cells excluded from the CD45−/SP fraction, attention can be turned to those cells enriched in this population. As shown by ISH, the location of these enriched cells supports the notion that they indeed contain one or more populations of progenitor cells. Further evidence of progenitor cells within the CD45−/SP cell fraction comes from our functional genomics analyses. An important characteristic of progenitor cells is their ability to both proliferate and differentiate in response to external signals from their environment. This process requires the translation of signals from the cell surface into appropriate changes in gene expression. The MAPK signaling cascade has been shown to function in this capacity in human intestinal cells, exerting effects on both cell cycle progression and differentiation (1). As described in our results, the MAPK pathway is significantly enriched in the CD45−/SP cells. The importance of this pathway in the CD45−/SP cell fraction, along with enriched GO terms such as “transcription” and “cell differentiation,” supports the presence of progenitor cells within this cell population.

In addition to the analyses described, further validation of the CD45−/SP dataset comes from a functional comparison of these cells to previously published intestinal stem cell studies. Specifically, LCM has been used to isolate IESCs based on their known location within the crypt (42). A cDNA library was ultimately generated from these LCM-captured cells and subject to extensive GO analyses, allowing for the identification of functional categories enriched in these putative progenitors (17). Using the 32 SP-enriched transcripts with confirmed localization to the crypt base/stem cell zone (Table 1), we have found significant overlap between the GO terms enriched in these transcripts and those identified using the LCM-derived cDNA library. In particular, the terms “regulation of metabolism,” “regulation of transcription,” “transcription,” “morphogenesis,” and “development” were found to describe both sets of genes. Again, such terms depict a transcriptionally active cell with progenitor-type features.

While our analyses provide evidence that the CD45−/SP cell fraction is enriched for progenitor cells and deenriched for the other cell types described, our results also confirm that this population is a heterogeneous one. For example, the fact that markers of immune, mesenchymal, absorptive, and Paneth cells were even detectable (although deenriched) in the CD45−/SP cells implies that trace numbers of these cells remain present in this fraction. In addition, significant numbers of mature endocrine cells appear to be represented in the CD45−/SP, as evidenced by the enrichment of transcripts such as secretin, cholecystokinin, glucagon, and chromogranin A. These findings are supported by the enrichment of the GO term “exocytosis” in the CD45−/SP cells. Similar GO results are
revealed when the gene expression profile of the CD45−/SP cells is compared with that of the CD45+/nSP cell fraction, which accounts for changes in gene expression that might arise from the stress of cell isolation and sorting. These analyses also show processes such as “exocytosis” (P value = 0.0003) as well as “hormone secretion” (P value = 0.002) to be significant in the CD45−/SP cells.

Although the data described support the presence of mature endocrine cells in the CD45−/SP fraction, there are several reasons to suggest that there are substantial numbers of enteroendocrine progenitors in this cell population as well. First, the location of the CD45−/SP cells along the crypt-villus axis is more consistent with a progenitor cell population. Second, previous work by Dekaney et al. examined the morphology of cytospun CD45−/SP cells using light microscopy (12). The cells were found to have abundant amorphous cytoplasm, much more consistent with a progenitor phenotype as opposed to a differentiated enteroendocrine cell. Finally, as discussed below, the MODY pathway appears to be highly enriched in the CD45−/SP cell fraction. This pathway is significant in the CD45−/SP cells, not only compared with intact jejunum (P value = 4 × 10−5), but compared with the CD45−/nSP cells as well (P value = 2 × 10−4).

The phrase maturity onset diabetes of the young (MODY) was originally coined to describe a subset of patients with familial, juvenile onset, noninsulin-dependent diabetes (43). At a molecular level, the MODY pathway follows the maturation of a pancreatic β-cell (15), beginning with the pluripotent pancreatic stem cell, progressing through the endocrine progenitor, and culminating in a functional insulin-producing cell (39). The CD45−/SP transcripts found in the MODY pathway include Neurog3, NeuroD1, Nkx2-2, Pax4, Pax6, and glucokinnase. Many of these have been studied in the gut using knockout animals and lineage tracing experiments. Such studies reveal a key role for these transcription factors in the development of the enteroendocrine system. For example, mice lacking Neurog3 fail to produce intestinal enteroendocrine cells (24), while homozygous NeuroD1 null mice have a profound reduction in secretin and cholecystokinin-producing cells (31). Pax4 mutants have essentially no duodenal or jejunal hormone-secreting cells, whereas Pax6 appears to be more involved in the development of duodenal gastric inhibitory polypeptide cells (27). Interestingly, these transcripts are upregulated in animals such as the Hes1 knockout mouse, which has very high numbers of intestinal enteroendocrine cells (25). The enrichment of the MODY pathway in the CD45−/SP cells, therefore, lends further support to the idea that this cell fraction contains significant numbers of cells that contribute to gut endocrine system development.

In summary, in this study we have sought to characterize the molecular features of a population of cells suggested to include putative IESCs (12). Overall, our functional genomics analyses indicate that, while this fraction is heterogeneous, it is enriched for progenitor cells and deenriched for immune, mesenchymal, and differentiated epithelial cells. This conclusion is further supported by our extensive ISH data, which show that the majority of CD45−/SP enriched transcripts localize to the progenitor cell zone. Consistent with this location, the significant presence of transcripts associated with the MODY pathway suggests that the CD45−/SP cells may also be enriched for enteroendocrine progenitors. These endocrine progenitors will need to be excluded in the future, if we are to isolate a pure fraction of IESCs. We believe that further study of transcripts common to the CD45−/SP and LCM-derived datasets will lead to the development of membrane protein-based sorting strategies, allowing for fractionation of the CD45−/SP cells into viable preparations of both IESCs and enteroendocrine progenitors. We are optimistic that such studies will provide the substrate for the development of both in vitro culture and in vivo transplantation models that can be used to characterize the true potential of these important cells. Because SP sorting can be performed in a straightforward manner, without the need for complex genetic manipulation to mark and isolate putative stem cell populations, it is a valuable tool for the study of the intestinal progenitor cell compartment. As such, we believe the present study provides a comprehensive characterization of the CD45−/SP cell fraction, which will allow investigators to utilize this technique to advance the field of intestinal stem cell biology.

ACKNOWLEDGMENTS

The authors are very grateful to Sue Venable for performing the microarrays for this study, and to Christina Thaller for performing the ISH work. The authors would also like to thank Barbara Agbmame and Jose Rodriguez for additional technical assistance, and Christopher Dekaney, Noah Shroyer, and Thaddeus Stappenbeck for critical feedback and scientific discussion.

GRANTS

The microarray and ISH work were supported in part by the Gulf Coast Digestive Disease Center (P30 DK 56338). This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases grants R01 DK 069585 and T32 DK 07664.

Present addresses: A. S. Gulati, Department of Pediatrics at the University of North Carolina at Chapel Hill Chapel Hill, NC; S. A. Ochsen, Department of Molecular and Cellular Biology at Baylor College of Medicine, Houston, TX; S. J. Henning, Departments of Medicine and Cellular and Molecular Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC.

REFERENCES


