COLD SPRING HARBOR
Molecular Case Studies

RESEARCH REPORT

Surgical and molecular characterization of primary and metastatic disease in a neuroendocrine tumor arising in a tailgut cyst

Jennifer Erdrich,[1,8,11] Kurt B. Schaberg,[2,9,11] Michael S. Khodadoust,[3] Li Zhou,[4] Andrew A. Shelton,[1] Brendan C. Visser,[1] James M. Ford,[3,5] Ash A. Alizadeh,[3] Stephen R. Quake,[6] Pamela L. Kunz,[3] and John F. Beausang[7,10]

[1]Department of Surgery, Stanford University School of Medicine, Stanford, California 94305, USA; [2]Department of Pathology, Stanford University School of Medicine, Stanford, California 94305, USA; [3]Department of Medicine/Oncology, Stanford University School of Medicine, Stanford, California 94305, USA; [4]Stanford Cancer Institute, Stanford University, Stanford, California 94305, USA; [5]Department of Genetics; Stanford University School of Medicine, Stanford, California 94305, USA; [6]Department of Bioengineering and Department of Applied Physics, Stanford University and Chan Zuckerberg Biohub, Stanford, California 94305, USA; [7]Department of Bioengineering, Stanford University, Stanford, California 94305, USA

Abstract Neuroendocrine tumors (NETs) arising from tailgut cysts are a rare but increasingly reported entity with gene expression profiles that may be indicative of the gastrointestinal cell of origin. We present a case report describing the unique pathological and genomic characteristics of a tailgut cyst NET that metastasized to liver. The histologic and immunohistochemical findings were consistent with a well-differentiated NET. Genomic testing indicates a germline frameshift in *BRCA1* and a few somatic mutations of unknown significance. Transcriptomic analysis suggests an enteroendocrine L cell in the tailgut as a putative cell of origin. Genomic profiling of a rare NET and metastasis provides insight into its origin, development, and potential therapeutic options.

Corresponding author:
john.beausang@gmail.com

Ontology terms: neoplasm of the gastrointestinal tract; neuroendocrine neoplasm

Published by Cold Spring Harbor Laboratory Press

doi: 10.1101/mcs.a003004

[Supplemental material is available for this article.]

INTRODUCTION

Tailgut cysts, also called retrorectal cystic hamartomas, are rare, congenital lesions that arise from persistent remnants of the postanal gut (Song et al. 2004; Patsouras et al. 2015; Shetty et al. 2015). Half of all cases are asymptomatic. When symptomatic, the most common presentation includes pain, constipation, palpable rectal mass, infection with fistula, recurrent perianal sepsis, and urinary retention (Lim et al. 2011; Patsouras et al. 2015). Tailgut cysts are uncommon, but they are increasingly recognized after the largest series of 53 cases was published (Hjermstad and Helwig 1988). Malignant transformation is rare but most likely to be adenocarcinoma with fewer cases of neuroendocrine tumors (NETs) reported (Abukar et al. 2014). Two cases of metastatic NETs arising in a tailgut cyst have been published (Song et al. 2004; Spada et al. 2011).

[8]Present address: Department of Surgery, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA
[9]Present address: Department of Pathology, University of Kentucky, Lexington, Kentucky 40506, USA
[10]Present address: GRAIL, Inc., Menlo Park, California 94025, USA
[11]These authors contributed equally to this work.

Genomic test results have not been published on tailgut cyst malignancies. Exome sequencing from small intestine NETs indicates a relatively low somatic mutation rate (~0.1 variants/Mb) across a range of cancer genes (Francis et al. 2013) with a frameshift mutation in *CDKN1B* the most common in 8% of cases (Francis et al. 2013). RNA-based analysis of small intestine NETs includes the search for biomarkers (Modlin et al. 2014) and expression profiling to identify subtypes (Kidd et al. 2014; Andersson et al. 2016). Recent work characterizing the RNA profile of single cells from intestinal organoids (Grün et al. 2015) represents initial steps toward molecular characterization of the diversity of neuroendocrine cells in healthy tissue (Furness et al. 2013), potentially enabling identification of NETs of unknown origin (Polish et al. 2011; Visvader 2011) and new therapeutic targets.

We present a case of retrorectal well-differentiated NET arising in a tailgut cyst with hepatic metastases. To our knowledge, this is the first up-front resection of both primary and metastatic disease and the first report using molecular assays to characterize the malignancy and identify the putative cell of origin.

RESULTS

A 77-yr-old man with a history of squamous cell carcinoma, prostate cancer, and melanoma presented with left hip pain radiating to the knee. A strong family history of cancer included a son with glioblastoma multiforme, father with prostate cancer, maternal aunt with lung cancer, maternal grandmother with ovarian cancer, and maternal grandfather with colon cancer. MRI of the spine was ordered because of suspected nerve compression causing sciatica. This detected a 6.3-cm presacral mass. CT similarly characterized a hypervascular presacral mass abutting the right bony pelvis and displacing the rectum (Fig. 1A). It also showed two hypervascular hepatic lesions. FNA of the presacral mass showed a nest of epithelioid cells with delicate cytoplasm concerning for recurrent prostatic adenocarcinoma. However, immunohistochemical stains were strong for synaptophysin and weak for chromogranin, consistent with neuroendocrine differentiation. Overall, these findings were diagnostic of a well-differentiated NET with provisional WHO grade-1 based on Ki-67 index of <1% and low mitotic index. As this is not a common site for a primary NET, it was unclear if this presacral mass represented a primary tumor or metastasis. Percutaneous biopsy of one hepatic lesion also showed well-differentiated NET with a provisional WHO grade-2 based on Ki-67 index 5.1%. Serum chromogranin, CEA, and CA 19-9 were within normal limits. Urine 5-HIAA was not performed given lack of hormonal symptoms.

Additional staging with novel somatostatin receptor scinitigraphy, Ga68 DOTATATE PET/CT revealed mild uptake in the pancreatic head in addition to the hepatic and presacral sites, compounding the location of the primary source. Pancreatic EUS with biopsy showed chronic pancreatitis with no clear tumor. Given that the disease was limited to the presacral mass and liver, a cytoreductive surgery was performed (Fig. 1B). In the operating room, a long midline incision was used to allow access for both surgical teams. The liver resection was performed first to avoid any potential contamination from the colorectal portion of the operation. A 2-cm lesion in segment 4b and a 0.5-cm lesion in segment 3 were identified by intraoperative ultrasound and resected. Cholecystectomy was performed because the segment 4b lesion was close to the gallbladder fossa. The Colorectal team then mobilized the rectum and circumferentially dissected the tumor free. It was not invading any nearby structures. It was well-encapsulated and shelled out of the presacral space intact. The patient recovered on the regular surgical ward on the Early Recovery After Surgery (ERAS) pathway and was discharged to home on the third postoperative day.

Histology of the presacral mass showed nests and trabeculae of epithelioid cells with eosinophilic cytoplasm, stippled chromatin, and inconspicuous nucleoli consistent with a

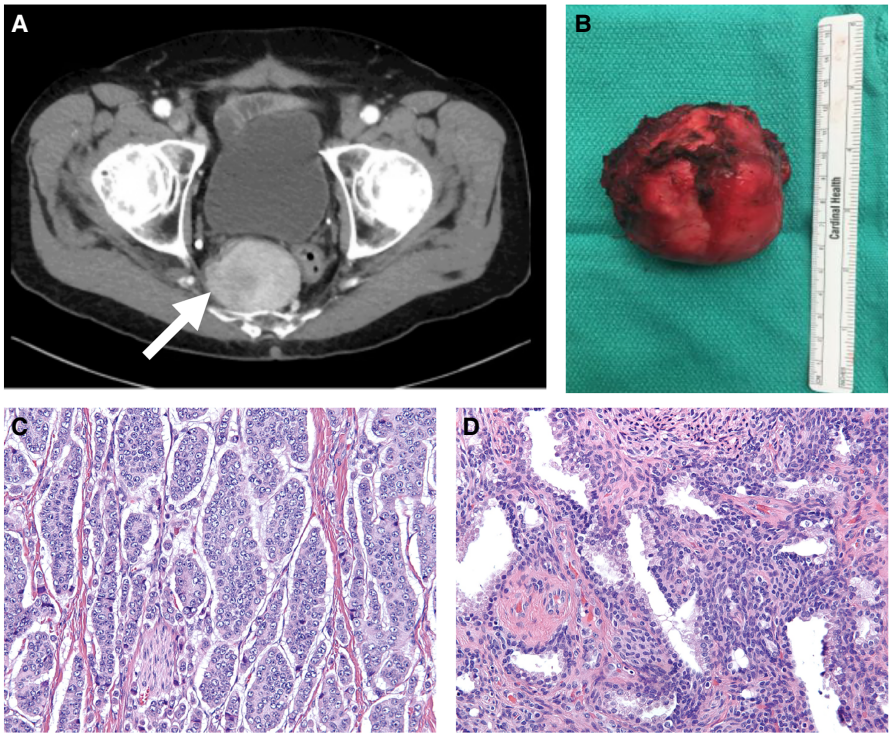

Figure 1. Imaging and histology of presacral mass. (*A*) Multiphasic CT demonstrating a hypervascular presacral mass (*white arrow*) measuring 4 × 6.3 × 4.6 cm. It displaces but does not appear to invade the adjacent rectum. (*B*) Presacral mass, 7 cm, firm and well-encapsulated, tethered to the sacrum, removed intact. (*C*) Well-differentiated neuroendocrine tumor component of presacral mass resection. H&E, 20× magnification. (*D*) Tailgut cyst component of presacral mass demonstrating lumina lined by prostatic-type and transitional epithelium with intervening fibrous tissue and disorganized smooth muscle. H&E, 20× magnification.

well-differentiated NET (Fig. 1C). The Ki-67 staining within the primary tumor was 8.6% with <1 mitosis/10 HPF consistent with WHO grade-2. The liver lesions showed identical morphologic features and were WHO grade-2 with a Ki-67 index of 6.4%. At the periphery of the presacral malignancy were multiple small cysts lined by a variety of epithelial cells including transitional, prostate-like, and mucinous epithelium (Fig. 1D) with admixed bundles of smooth muscle and fibrous tissue diagnostic of a tailgut cyst, which harbored the primary NET.

Molecular testing was performed on the presacral mass, metastasis, germline, and plasma-derived DNA using both laboratory-grade and clinically available, CLIA-certified tests including Guardant360, FoundationOne, Stanford Tumor Actionable Mutation Panel (STAMP), and MyriadMyRisk. Mutations in BRCA1 E23fs*17 (Germline Ashkenazi founder mutation; FoundationOne) and POM121L12 (variant of unknown significance; STAMP) were confirmed by exome sequencing. RNA sequencing was performed on biopsied ("Bx") and resected ("Rs") tissue from the sacral mass and liver. The expression levels in the samples are compared for the 25 most highly expressed genes in the presacral mass (Fig. 2) and for the top 25 differentially expressed genes in the metastasis relative to normal liver (Fig. 3). Tissue-specific gene expression was determined using the Genotype-Tissue Expression (GTEx) database of healthy tissue (The GTEx Consortium et al. 2015) and indicated that many top genes were enriched in neuronal and/or endocrine tissues. Both gene lists showed characteristics of a NET with two of the top three genes, insulin-like 5 (*INSL5*) and chromogranin-B (*chrB*) common in both primary and metastasis. *INSL5*, a colon-

	Resected Liver		Sacral Mass*	
	Normal (FPKM)	Met (FPKM)	(FPKM)	Dominant Tissue
INSL5	38	1913	2063	Colon
RPPH1	636	1076	1315	Brain
CHGB	13	1010	1269	Pituitary
TTR	1860	814	1068	Liver
MALAT1	308	675	959	-
RMRP	516	493	741	Brain
GCG	1	113	687	Pancreas
CPE	16	657	628	Brain
TPT1	1235	395	627	-
B2M	680	596	617	-
HSP90B1	611	610	608	-
APOD	0	17	576	Nerve
ACTB	505	472	567	-
HSP90AB1	205	489	547	-
SCG5	12	506	535	Pituitary
CALR	612	538	518	-
EEF2	493	397	507	-
PTMA	341	449	481	-
SNORD17	15	278	470	EBV-Lymp
NPM1	287	396	468	-
ENPP2	12	334	466	-
PTMS	740	413	450	-
VIM	55	430	448	-
SNORD97	84	304	437	-
ACTG1	278	462	435	-

Figure 2. Gene expression for the top 25 genes in the presacral mass. Gene expression in fragments per kilobase per million (FPKM) in the presacral mass is sorted from high (blue) to low (white) and the top 25 genes compared in the three resected samples. Dominant Tissue indicates tissue-specific expression estimated from the GTEx database. EBV-Lymp represents expression associated with EBV-transformed lymphocytes. See the Methods section for details and top genes expressed in the metastasis (Supplemental Table S1) and normal liver (Supplemental Table S2). An asterisk indicates sort column.

	Liver						
	Normal		Metastasis		Sacral Mass	Avg. Fold Change*	
	Bx (FPKM)	Rs (FPKM)	Bx (FPKM)	Rs (FPKM)	Rs (FPKM)	$\log_2$(Met/Nrml)	Dominant Tissue
CHGB	0.1	13.1	898	1010	1269	7.9	Pituitary
INSL5	1.1	37.5	1634	1913	2063	7.6	Colon
SCG2	0.2	4.0	442	440	319	7.5	-
HEPACAM2	0.1	2.7	288	358	289	7.3	-
VGF	0.1	1.2	205	257	120	7.2	-
PCSK1N	0.2	4.2	277	442	372	7.2	Pituitary
CDH17	0.0	1.6	181	279	208	7.1	-
SCG3	0.1	2.0	211	296	287	7.1	Brain
PPP1R1B	0.1	2.5	206	291	270	7.0	Brain
SCGB2A1	0.0	0.8	132	207	95	7.0	Cervix
TFF1	0.0	2.8	214	238	145	6.9	Stomach
PYY	0.0	2.2	198	207	196	6.8	-
QPCT	0.2	1.7	158	262	196	6.8	-
KIAA1324	0.0	1.3	151	200	191	6.8	-
MS4A8B	0.0	0.7	115	164	199	6.7	NA
KIF5C	0.1	1.3	144	189	228	6.7	Brain
ADCYAP1	0.0	1.5	132	187	92	6.6	Brain
PPY	0.3	3.5	180	313	250	6.6	Pancreas
PRPH	0.1	1.5	126	187	110	6.5	-
MUC13	0.0	1.3	122	163	107	6.5	-
SEMA3D	0.7	1.5	175	194	75	6.5	-
PARM1	0.5	1.5	146	195	206	6.4	-
GUCY2C	0.3	1.3	121	177	188	6.4	-
GCG	0.0	1.1	119	113	687	6.3	Pancreas
PTPRN	0.0	0.5	76	122	80	6.3	-

Figure 3. Gene expression for the top differentially expressed genes in the liver metastasis. Comparison of the gene expression (FPKM) in the top 25 genes with increased expression in the metastasis relative to normal liver, sorted from high (blue) to low (white). Genes are sorted by the average $\log_2$ of the fold change (FC) expression in the biopsied (Bx) and resected (Rs) samples. Dominant Tissue indicates tissue-specific expression estimated from the GTEx database. See Methods for details. An asterisk indicates sort column.

specific marker of healthy gastrointestinal tract (Gremel et al. 2014), and *CDH17* have both been linked to colorectal NETs (Mashima et al. 2013; Thanasupawat et al. 2013; Gremel et al. 2014; Snow et al. 2015). Classical neuroendocrine markers chromogranin-A and synaptophysin (Klimstra et al. 2010; Kunz et al. 2013) and small intestine–specific markers secretogranins (SCG2 and SCG5) and somatostatin receptors SSTR1, 2, and 5 were all elevated (Modlin et al. 2014; Kidd et al. 2015; Andersson et al. 2016). Gene ontology analysis indicated numerous pathways associated with signaling, secretory peptides, and extracellular vesicle trafficking, as expected for NETs (Supplemental Tables S3, S4). Pearson correlation coefficients of the gene expression between the core biopsies and resected tissues was 0.97 and 0.74 for normal liver and the metastasis, respectively (Supplemental Fig. S1), indicating the diagnostic potential of RNA sequencing from preoperative samples. See the Methods section for further details on sequencing SNV results and GTEx analysis.

To identify genes that distinguish the metastasis from the presacral mass, differential expression relative to the presacral mass is compared with differential expression relative to normal liver (Fig. 4). Liver-specific genes are highly expressed in both the metastasis and normal liver with low expression in the presacral mass (e.g., *ALB*, *HP*, and *FGG* and other genes from Supplemental Table S2 in the upper left quadrant). Genes highly expressed in the sacral mass (Fig. 2) are also highly expressed in the metastasis with either low (e.g., *CHGB* and *INSL5* along the positive *x*-axis) or high (e.g., *B2M*, *ACTB*, and *HSP90B1* in the center of Fig. 4) expression in the normal liver. Thirteen genes with high expression in the metastasis and low expression in both the presacral mass and normal liver tissue suggest that they are specific to the metastasis (set A in Fig. 4; Supplemental Table S10). These genes have diverse roles in protein secretion (*LCN1* and *TCN1*), cell migration (*MYOC*), intracellular transport (*MARCH11* and *LY6H*), and differentiation (*NELL1* and *ASCL1*). Three of the genes have been associated with metastasis in the human cancer metastasis database (HCMDB): *HLA-G*, *MAGEA3* (melanoma), and the proto-oncogene *RET* (thyroid cancer; Zheng et al. 2018). In addition, RET has been reported as a therapeutic target of tyrosine kinase inhibitors in small intestine NETs (Andersson et al. 2016).

Genes specific to the presacral mass (set B genes along the negative *y*-axis of Fig. 4) have moderate expression (20–190 FPKM) but very low expression in the metastasis and normal liver (Supplemental Table S10). Many of these genes (including *FBLN1*, *DES*, *ACTG2*, *COMP*, *CNN1*, *MFAP5*, and *MYH11*) have roles in muscle/motility and the cytoskeleton. Two of the three most differentially expressed genes, *SFRP2* and *PRRX1*, were identified in the HCMDB database for metastases in melanoma and colorectal cancer. *SFRP2* and fibutlin-1 (*FBLN1*) are tumor-suppressor genes in colon (Kongkham et al. 2010) and gastric cancers (Cheng et al. 2008), respectively. *PRRX1* induces the epithelial-to-mesenchymal transition, is required for cancer cells to metastasize in vivo, and is a biomarker for lack of metastasis (Ocaña et al. 2012). The reduced expression of these genes in the metastasis compared to the primary presacral mass may have contributed to progression to metastasis.

Expression levels of key genes in 12 different gastrointestinal enteroendocrine cells (Furness et al. 2013) were compared to determine a possible cell of origin for the presacral mass (Table 1). Glucagon-like peptide 1 and 2 (*GLP1* and *GLP2*), oxyntomodulin, and glucagon (*GCG*) are all posttranslational modifications from the *Gcg* gene. The prohormone convertases (PCs) that generate these products are tissue-specific, with *PCSK2* found in pancreas, producing mostly glucagon, and *PCSK1* found in brain and intestine, producing mostly *GLP-1*, oxyntomodulin, and *GLP-2* (Sandoval and D'Alessio 2015). Greater than 100-fold higher levels of PCSK1 compared to *PCSK2* suggest that the *GCG* transcripts are likely *GLP1* and/or *GLP2* and not glucagon, consistent with intestinal L cells. High levels of *GCG* and pancreatic polypeptide (*PPY*) may be associated with an islet cell tumor in the pancreas; however, low expression of insulin (INS) and high expression of colorectal markers suggest

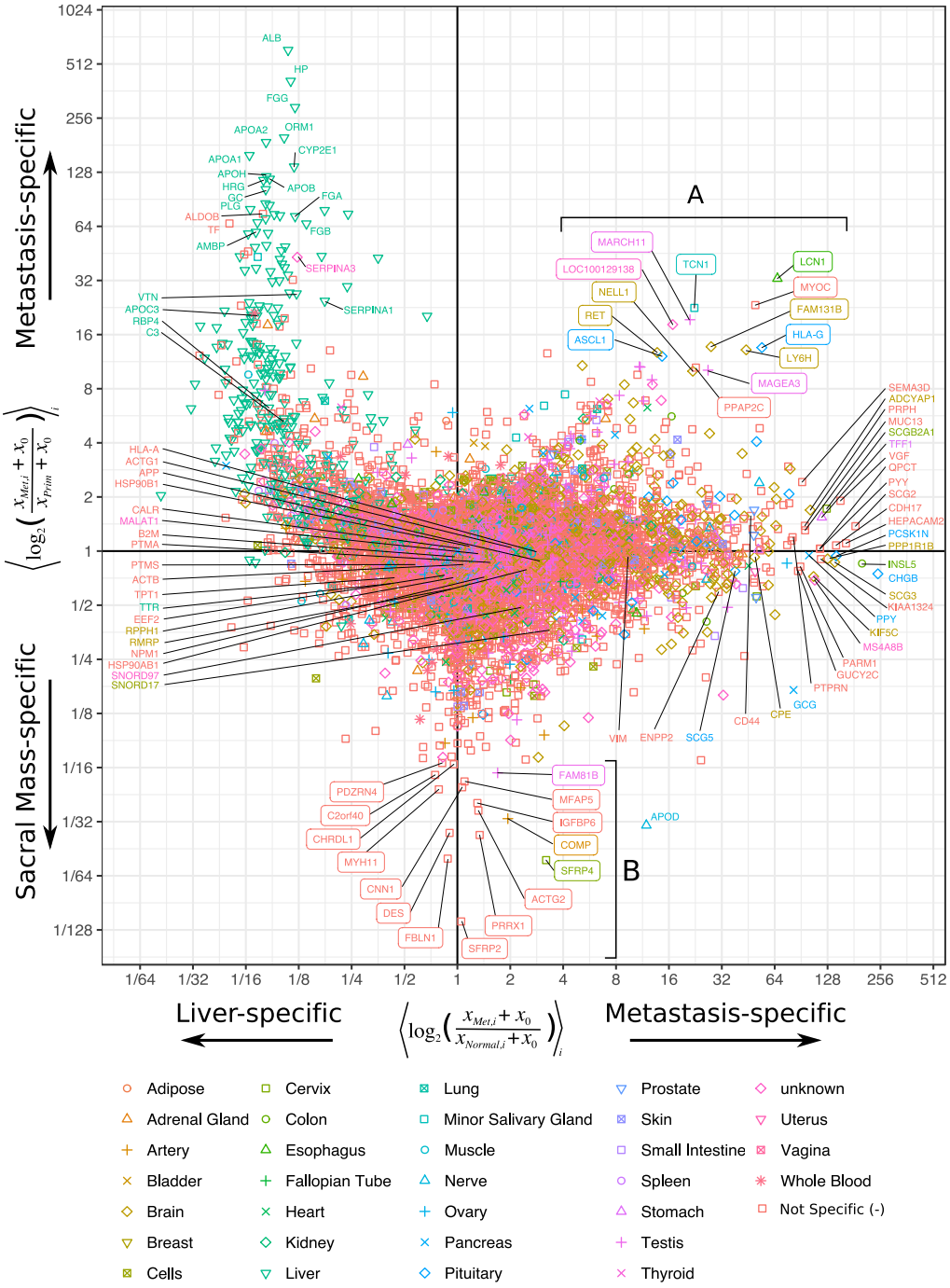

Figure 4. Differential expression of the metastasis compared to the sacral mass and normal liver. Gene expression in the liver metastasis relative to the primary sacral mass (*y*-axis) is compared to the average expression in the metastasis relative to normal liver tissue (*x*-axis). Top genes that are highly expressed in the sacral mass (Fig. 1), metastasis (Fig. 2; Supplemental Table S1), and normal liver (Supplemental Table S2) are shown as labels with color and shapes according to the dominant tissue; many genes do not meet the specificity requirements (*red squares*). Boxed labels indicate genes that are highly expressed in the metastasis compared with both the presacral mass and normal liver (set A) and decreased expression in the metastasis relative to the presacral mass (set B).

Table 1. Gene expression (FPKM) in resected samples for enteroendocrine-specific genes in the gastrointestinal tract

Cell type	Hormone/precursor name	Gene	Liver		Sacral mass
			Normal	Met	
K cells	Gastric inhibitory peptide	*GIP*	0.0	0.0	0.0
L cells	Glucagon-like peptide 1; glucagon-like peptide 2; oxyntomodulin	*GCG*	1.1	113	687
	Pancreatic polypeptide	*PPY*	3.5	313	250
	Proprotein convertase subtilisin/kexin type 1	*PCSK1*	2.1	286	103
I cells	Cholecystokinin	*CCK*	0.0	0.0	0.1
G cells	Gastrin	*GAST*	0.0	0.0	0.0
Entero-chromaffin	Tryptophan hyroxylase1 (serotonin)	*TPH1*	0.0	0.3	0.2
	Tryptophan hyroxylase2 (serotonin)	*TPH2*	0.0	6.1	0.8
Entero-chromaffin-like	Histidine decarboxylase (histidine production)	*HDC*	0.6	0.0	0.2
	Vesicular monoamine transporter subtype 2 (histidine production)	*SLC18A2*	11.1	6.4	9.8
N cell	Neurotensin	*NTS*	0.4	3.1	1.5
S cells	Secretin	*SCT*	0.0	6.8	2.0
P cells	Leptin	*LEP*	0.0	0.0	0.1
M cells	Motilin	*MLN*	0.0	0.0	6.3
D cells	Somatostatin	*SST*	0.0	0.3	3.8
A cells (X-like)	Ghrelin	*GHRL*	0.0	0.4	0.8
	Nesfatin-1	*NUCB2*	62	317	274

that an L cell in the large intestine is the more likely cell of origin. Recent studies have shown that INSL5 is an orexigenic hormone released from colonic L cells in mice (Grosse et al. 2014), supporting this hypothesis. High levels of nesftatin-1 suggest the A cell may be a candidate, but elevated levels in normal liver and low levels of ghrelin argue otherwise.

DISCUSSION

Most NETs arise in the gastrointestinal tract and lungs. Because there is no apparent source of neuroendocrine cells in the presacral space, many of these tumors likely derive from the glandular epithelium of tailgut cysts (Horenstein et al. 1998). This entity is not widely known but should be on the differential of any presacral mass. Gene expression provided a detailed description consistent with traditional pathological results and pointed toward a more detailed understanding of the tumor including a putative cell of origin that may provide a basis for future targeted therapies. Tumor heterogeneity was not assessed in this work, but transcriptomic analysis from the biopsy and resected tissue of both the primary and metastasis were highly similar and may indicate a promising source of diagnostic information in NETs.

METHODS

Sample Processing

Tissue from needle core biopsies of the liver metastasis and adjacent normal was transferred to 1.5-ml Eppendorf tubes, placed on dry ice, and stored at −80°C. After 51 d the presacral mass and liver metastasis were resected and bulk samples from each were obtained. Bulk normal liver was obtained from a dark red piece of liver ~0.5 cm from the white metastasis border. Normal tissue from the periphery of the metastasis was removed before incubating

all three samples in prechilled RNA overnight at 4°C and then stored at −20°C until use. DNA and RNA were purified from all five samples using an AllPrep DNA/RNA Mini Kit (QIAGEN). DNA and RNA concentration was determined with via fluorimetry (Qubit) and the quality of RNA assessed via RIN number using electrophoresis (Agilent Bioanalyzer).

DNA Sequencing

CLIA-approved testing was performed using Guarant360, FoundationOne, STAMP, and MyriadMyRisk and verified via exome sequencing in our laboratory (Alizadeh). Pre- and postoperative testing of circulating cell-free DNA by Guardant360 revealed mutations in *ARID1A* and *SRC*. FoundationOne testing of both the primary presacral mass and liver metastasis showed mutations in *BRCA1* E23fs*17 and *ASXL1* E635fs*15. STAMP testing of the presacral mass and liver revealed variant of unknown significance (VUS) LRRTM4, which is the same as FoundationOne ASXL1. STAMP of the liver metastasis revealed another VUS, in POM121L12. Germline testing by MyriadMyRisk confirmed the *BRCA1* E23fs*17 mutation (Ashkenazi founder mutation). For exome sequencing, Illumina sequencing libraries were constructed from sheared tumor DNA and matched normal from peripheral blood mononuclear cells using KAPA HyperPrep kit. Libraries were pooled and exomes enriched using SeqCapEZ MedExome (Roche NimbleGen) kit. The final product was sequenced with 150-bp paired-end reads on an Illumina HiSeq4000 to a median depth of ~160× with an on-target rate of ~88% for each library (See Supplemental Table S5). Variant calling and annotation were performed as previously described (Khodadoust et al. 2017). Exome sequencing verifies the *BRCA1* germline mutation (from FoundationOne) and *POM121L12* (from STAMP); other mutations, including *LRRTM4* (from STAMP), *ASXL1* (from FoundationOne), and *ALK*, *ARID1A*, and *SRC* (from Guardant), were not supported. High allele fraction mutations detected by exome sequencing but none of the other assays include a homozygous *PCDHB16* missense mutation and other coding mutations in *SHANK2*, *SAMHD1*, *TNFSF8* (*CD30L*), and *NETO2*. See Supplemental Table S6 for a summary of the assay results and Table S11 for all PASS filter variants from exome sequencing.

RNA Sequencing and Analysis

RNA-seq Illumina libraries were constructed from each sample using the SMARTer Stranded Total RNA Sample Prep Kit—HI Mammalian (Clonetech #634873) kit and 43–69 million 75-bp paired-end sequencing reads obtained on the Illumina NextSeq500 (see Table 2 for details). Reads were aligned using STAR (Dobin et al. 2013) to the hg19 genome. Ht-Seq (Anders and Huber 2010) using stringent assignments was used to count uniquely mapped reads to each gene (Supplemental Table S9). Fragments per kilobase of transcript length per million reads (FPKM) values were determined for each gene by normalizing the counts by the total number of mapped reads in the sample, the length of the gene in base pairs and scaling by 1×10^9 (Supplemental Table S7). Results for the top 25 most highly expressed genes for

Table 2. RNA sequencing metrics

	Sample	Total RNA (ng)	PF clusters
1	Normal liver, biopsy	344	47,443,335
2	Normal liver, resected	800	43,283,705
3	Liver metastasis, biopsy	620	69,793,317
4	Liver metastasis, resected	800	57,589,019
5	Sacral mass, resected	800	46,676,100

the resected presacral mass, liver metastasis, and normal liver can be found in Figure 2, Supplemental Table S1, and Supplemental Table S2, respectively.

Similarity between the gene expression profiles for the five samples was determined by calculating the Pearson correlation coefficient between the gene expression values of the top 1000 genes for each pair of samples (Supplemental Fig. S1). The resected liver metastasis and sacral mass are highly correlated (0.93), and both are less correlated to the biopsied liver metastasis (0.74 and 0.66, respectively). The two normal samples are highly correlated with each other (0.97), moderately correlated to the biopsied liver metastasis (0.74 and 0.7, respectively), and weakly correlated to the resected metastasis and sacral mass (0.2–0.3). Scatter plots of the 1000 most highly expressed genes (Supplemental Fig. S1) indicate that both normal liver samples contain liver-specific genes (as determine by GTEx, see below), and the biopsied metastasis sample contains a population of liver genes, which are not seen in the resected metastasis or sacral mass, and which explain the reduced correlation between these samples. The metastasis biopsy has a component of normal liver that is lower in the resected sample and negligible in the sacral mass.

Tissue Specificity Analysis

Tissue specificity was determined from the GTEx database (GTEx Consortium 2015) by converting the raw count data for each sample to FPKM. The average expression across all samples for each gene and tissue combination was then calculated and tabulated such that each gene corresponds to a vector of tissue expression levels. Original GTEx tissue labels are truncated at the dash ("-") in the tissue name such that compartments within an organ are ignored, for example "Brain - Cerebellum" and "Brain - Cortex" are both "Brain." A gene is classified as specific if the most abundant tissue contains at least 10% of the total expression and the second most abundant tissue has at least twofold lower abundance than the top tissue. In the resected liver sample, the 25 genes with the highest expression (Supplemental Table S2) include known liver genes such as albumin, haptoglobin, and several apolipoproteins. Four genes did not pass the threshold for tissue specificity, but the remaining 21 were all determined to be specific for liver. Two liver genes are also highly expressed in the sacral mass: transthyretin (*TTR*), which is reported to transport thyroid hormone, and tumor protein translationally controlled 1 (*TPT1*), which is involved in calcium binding and microtubule stabilization (Safran et al. 2010).

Gene Ontology

Gene ontology annotations were determined by importing the top 100 gene names for the sacral mass and top 100 differentially expressed genes in the liver metastasis into the DAVID functional analysis tool (https://david.ncifcrf.gov/). Pathways with Bonferroni adjusted P-values of <0.01 for the top genes are reported for the sacral mass (Supplemental Table S3) and liver metastasis (Supplemental Table S4).

Differentially Expressed Genes between Normal Liver and Metastasis

The $\log_2$ of the FC in differential expression for each gene is calculated as

$$\log_2(\mathrm{FC}) = \left\langle \log_2\left(\frac{x_{\mathrm{Met},i} + x_0}{x_{\mathrm{Normal},i} + x_0}\right)\right\rangle_i,$$

where x is the gene expression in FPKM, $x_0 = 1$ FPKM to minimize the impact of low abundance transcripts, and the average is over both samples in the metastasis and normal liver (i.e., $i = \{\mathrm{Biopsy, Resected}\}$). Differentially expressed genes were ranked by averaging the $\log_2$(FC) values from resected and biopsied samples and the top 25 genes with increased expression between normal liver and the metastasis reported (Fig. 3). Several of these top

genes have not been reported as neuroendocrine markers but are consistent with neuroendocrine functions and/or tumor metastasis. *VEGF* is a growth factor active in angiogenesis, vasculogenesis, and endothelial cell growth and is up-regulated in many tumors (Safran et al. 2010). *SCGB2A1* is a secretoblobin that is preferentially expressed in the cervix. It binds androgens and other steroids and has been associated with ovarian cancer (Safran et al. 2010). Trefoil factor 1 (*TFF1*) is a secretory protein expressed in healthy gastrointestinal mucosa of the stomach as well as in tumors (Safran et al. 2010). Peptide YY (*PYY*) inhibits pancreatic secretion and mobility in the gut. In the GTEx database glutaminyl-peptide cyclotransferase (*QPCT*) is expressed in several neuroendocrine tissues (adrenal gland, pituitary, and hypothalamus) and is responsible for the presence of pyroglutamyl residues in many neuroendocrine peptides (Safran et al. 2010). Eight of the genes in Figure 3 are determined to be specific for neural and/or endocrine-related tissues including brain (*SCG3*, *PPP1R1B*, *KIF5C*, and *ADCYAP1*), pituitary (*CHGB* and *PCKS1N*), and pancreas (*PPY* and *GCG*). Mucin 13 (MUC13) is a secreted cell surface glycoprotein expressed by ductal and glandular epithelial tissues and may play a role in cell signaling (Safran et al. 2010). MUC13 is not specific to a single tissue, but it is expressed primarily in colon and small intestine (GTEx Consortium 2015). Gene ontology annotations indicate processes associated with secreted proteins, signaling peptides, and extracellular regions, similar to the sacral mass (Supplemental Table S4). Full names for genes in Figures 2 and 3 can be found in Supplemental Table S8.

ADDITIONAL INFORMATION

Data Deposition and Access

The table of RNA-seq gene counts is available as Supplemental Material but raw sequencing data is not available under the IRB used to perform this work.

Ethics Statement

The patient provided written consent and was enrolled in an IRB-approved (#13535) research study at Stanford University.

Acknowledgments

The authors would like to thank Norma Neff and Jennifer Okamoto for sequencing assistance and the anonymous reviewers for their helpful suggestions and comments.

Author Contributions

A.A.S., B.C.V., J.M.F., A.A.A., P.L.K., K.B.S., S.R.Q., and J.F.B. contributed to study concept and design. All authors contributed to acquisition, analysis, or interpretation of data. A.A.S., B.C.V., J.M.F., A.A.A., S.R.Q., and P.L.K. contributed to study supervision. J.E., K.B.S., P.L.K., and J.F.B. drafted the manuscript. All authors have read and approved the final manuscript.

Competing Interest Statement

The authors have declared no competing interest.

Received March 21, 2018; accepted in revised form July 31, 2018.

Funding

S.R.Q. was supported by the Howard Hughes Medical Institute.

REFERENCES

Abukar AA, Parcell BJ, Lim CB, Patil PV, Ramsanahie A, Carey F, Steele RJC, Thaha MA. 2014. Malignancy within a tail gut cyst: a case of retrorectal carcinoid tumour. *Case Rep Surg* **2014:** 1–4.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* **11:** R106.

Andersson E, Arvidsson Y, Swärd C, Hofving T, Wängberg B, Kristiansson E, Nilsson O. 2016. Expression profiling of small intestinal neuroendocrine tumors identifies subgroups with clinical relevance, prognostic markers and therapeutic targets. *Mod Pathol* **29:** 616–629.

Cheng YY, Jin H, Liu X, Siu JM, Wong YP, Ng EK, Yu J, Leung WK, Sung JJ, Chan FK. 2008. Fibulin 1 is downregulated through promoter hypermethylation in gastric cancer. *Br J Cancer* **99:** 2083–2087.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29:** 15–21.

Francis JM, Kiezun A, Ramos AH, Serra S, Pedamallu CS, Qian ZR, Banck MS, Kanwar R, Kulkarni AA, Karpathakis A, et al. 2013. Somatic mutation of *CDKN1B* in small intestine neuroendocrine tumors. *Nat Genet* **45:** 1483–1486.

Furness JB, Rivera LR, Cho HJ, Bravo DM, Callaghan B. 2013. The gut as a sensory organ. *Nat Rev Gastroenterol Hepatol* **10:** 729–740.

Gremel G, Wanders A, Cedernaes J, Fagerberg L, Hallström B, Edlund K, Sjöstedt E, Uhlén M, Pontén F. 2014. The human gastrointestinal tract–specific transcriptome and proteome as defined by RNA sequencing and antibody-based profiling. *J Gastroenterol* **50:** 46–57.

Grosse J, Heffron H, Burling K, Akhter Hossain M, Habib AM, Rogers GJ, Richards P, Larder R, Rimmington D, Adriaenssens AA, et al. 2014. Insulin-like peptide 5 is an orexigenic gastrointestinal hormone. *Proc Natl Acad Sci* **111:** 11133–11138.

Grün D, Lyubimova A, Kester L, Wiebrands K, Basak O, Sasaki N, Clevers H, van Oudenaarden A. 2015. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525:** 251–255.

Hjermstad BM, Helwig EB. 1988. Tailgut cysts. Report of 53 cases. *Am J Clin Pathol* **89:** 139–147.

Horenstein MG, Erlandson RA, Gonzalez-Cueto DM, Rosai J. 1998. Presacral carcinoid tumors: report of three cases and review of the literature. *Am J Surg Pathol* **22:** 251–255.

Khodadoust MS, Olsson N, Wagar LE, Haabeth OA, Chen B, Swaminathan K, Rawson K, Liu CL, Steiner D, Lund P, et al. 2017. Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature* **543:** 723–727.

Kidd M, Modlin IM, Drozdov I. 2014. Gene network–based analysis identifies two potential subtypes of small intestinal neuroendocrine tumors. *BMC Genomics* **15:** 595.

Kidd M, Drozdov I, Modlin I. 2015. Blood and tissue neuroendocrine tumor gene cluster analysis correlate, define hallmarks and predict disease status. *Endocr Relat Cancer* **22:** 561–575.

Klimstra DS, Modlin IR, Coppola D, Lloyd RV, Suster S. 2010. The pathologic classification of neuroendocrine tumors. *Pancreas* **39:** 707–712.

Kongkham PN, Northcott PA, Croul SE, Smith CA, Taylor MD, Rutka JT. 2010. The SFRP family of WNT inhibitors function as novel tumor suppressor genes epigenetically silenced in medulloblastoma. *Oncogene* **29:** 3017–3024.

Kunz PL, Reidy-Lagunes D, Anthony LB, Bertino EM, Brendtro K, Chan JA, Chen H, Jensen RT, Kim MK, Klimstra DS, et al. 2013. Consensus guidelines for the management and treatment of neuroendocrine tumors. *Pancreas* **42:** 557–577.

Lim SW, Huh JW, Kim YJ, Kim HR. 2011. Laparoscopy-assisted resection of tailgut cysts: report of a case. *Case Rep Gastroenterol* **5:** 22–27.

Mashima H, Ohno H, Yamada Y, Sakai T, Ohnishi H. 2013. INSL5 may be a unique marker of colorectal endocrine cells and neuroendocrine tumors. *Biochem Biophys Res Commun* **432:** 586–592.

Modlin IM, Oberg K, Taylor A, Drozdov I, Bodei L, Kidd M. 2014. Neuroendocrine tumor biomarkers: current status and perspectives. *Neuroendocrinology* **100:** 265–277.

Ocaña OH, Córcoles R, Fabra A, Moreno-Bueno G, Acloque H, Vega S, Barrallo-Gimeno A, Cano A, Nieto MA. 2012. Metastatic colonization requires the repression of the epithelial–mesenchymal transition inducer Prrx1. *Cancer Cell* **22:** 709–724.

Patsouras D, Pawa N, Osmani H, Phillips RK. 2015. Management of tailgut cysts in a tertiary referral centre: a 10-year experience. *Colorectal Dis* **17:** 724–729.

Polish A, Vergo MT, Agulnik M. 2011. Management of neuroendocrine tumors of unknown origin. *J Natl Compr Canc Netw* **9:** 1397–1402.

Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H, et al. 2010. GeneCards Version 3: the human gene integrator. *Database (Oxford)* **2010:** baq020.

Sandoval DA, D'Alessio DA. 2015. Physiology of proglucagon peptides: role of glucagon and GLP-1 in health and disease. *Physiol Rev* **95:** 513–548.

Shetty AS, Loch R, Yoo N, Mellnick V, Fowler K, Narra V. 2015. Imaging of tailgut cysts. *Abdom Imaging* **40:** 2783–2795.

Snow AN, Mangray S, Lu S, Clubwala R, Li J, Resnick MB, Yakirevich E. 2015. Expression of cadherin 17 in well-differentiated neuroendocrine tumours. *Histopathology* **66:** 1010–1021.

Song DE, Park JK, Hur B, Ro JY. 2004. Carcinoid tumor arising in a tailgut cyst of the anorectal junction with distant metastasis: a case report and review of the literature. *Arch Pathol Lab Med* **128:** 578–580.

Spada F, Pelosi G, Squadroni M, Lorizzo K, Farris A, de Braud F, Fazio N. 2011. Neuroendocrine tumour arising inside a retro-rectal tailgut cyst: report of two cases and a review of the literature. *Ecancermedicalscience* **5:** 201.

Thanasupawat T, Hammje K, Adham I, Ghia JE, Del Bigio MR, Krcek J, Hoang-Vu C, Klonisch T, Hombach-Klonisch S. 2013. INSL5 is a novel marker for human enteroendocrine cells of the large intestine and neuroendocrine tumours. *Oncol Rep* **29:** 149–154.

The GTEx Consortium. 2015. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348:** 648–660.

Visvader JE. 2011. Cells of origin in cancer. *Nature* **469:** 314–322.

Zheng G, Ma Y, Zou Y, Yin A, Li W, Dong D. 2018. HCMDB: the human cancer metastasis database. *Nucleic Acids Res* **46:** D950–D955.