Alternative polyadenylation of tumor suppressor genes in small intestinal neuroendocrine tumors

Anders Rehfeld1†, Mireya Plass3†, Kristina Døssing1, Ulrich Knigge3, Andreas Kjær4, Anders Krogh2 and Lennart Friis-Hansen1*

1 Center for Genomic Medicine, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
2 Department of Biology, The Bioinformatics Centre, University of Copenhagen, Copenhagen, Denmark
3 Department of Surgical Gastroenterology and Endocrinology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
4 Department of Clinical Physiology, Nuclear Medicine and PET, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

*Correspondence:
Lennart Friis-Hansen, Center for Genomic Medicine, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, Copenhagen DK 2100, Denmark
e-mail: lfh@dadlnet.dk
†Anders Rehfeld and Mireya Plass have contributed equally to this work.

INTRODUCTION

Small intestinal neuroendocrine tumors (SI-NETs) are a heterogeneous group of tumors originating primarily from the serotonin producing enterochromaffin neuroendocrine cells dispersed as solitary cells between the mucosal cells in the small intestine (1). Most SI-NETs tend to be slow growing with a low Ki-67 proliferation index (2). The primary tumor, regional node metastases, and related mesenteric fibrosis may give rise to local symptoms as bowel obstruction and/or vascular encasement. The presence of liver metastases may give rise to the carcinoid syndrome including diarrhea, facial flushing, and on sight cardiac and pulmonary disease, caused by release of serotonin and vasoactive peptides from the liver metastases (2).

Both the incidence and prevalence of SI-NETs have been increasing during the last three decades (3). Nowadays, SI-NETs are the second most prevalent gastrointestinal cancer after colonic cancer (4). Radical surgical intervention is the only curative treatment. However, SI-NETs are difficult to diagnose, as the symptoms often are unspecific. Therefore, many patients are first diagnosed once metastases have occurred and curative surgery is no longer an option (1).

Polyadenylation is a co-transcriptional process that consists of cleavage of the pre-messenger RNA (pre-mRNA) at a polyA site and addition of a polyA tail to the 5′ cleavage product. This process is necessary for normal mRNA 3′-end formation, transcription termination, and mRNA export from the nucleus (5, 6). Studies have shown that almost all eukaryotic pre-mRNAs and several non-coding transcripts have polyA sites and are polyadenylated. More than two-thirds of human genes have multiple polyA sites and can therefore undergo alternative polyadenylation (APA) (7). This produces two or more distinct mRNA isoforms with different 3′-terminal coding regions (10) and truncated mRNA coding regions (10). In some cases, APA also produces mRNAs containing different 3′-terminal coding regions. Therefore, APA alters both the repertoire and the expression level of proteins. Figure 1.

Alterations in polyadenylation can cause various diseases (8, 9) and several studies have linked cellular proliferation, transformation, and cancer progression to APA. Proliferation correlates with enhanced proximal polyA site usage leading to expression of mRNAs with 3′ truncated coding regions (10) and truncated 3′ UTRs (5, 10). Cell transformation is associated with 3′ UTR truncation (11–14) and changed APA pattern (10, 15), where transformed cells express mRNAs with shorter 3′ UTRs compared to non-transformed cells with similar proliferation rate (12). Furthermore, 3′ UTR shortening is linked to poor cancer prognosis (16). The APA pattern also differs according to cancer subtype.

Keywords: carcinoids, neuroendocrine tumors, gene transcription, molecular biology, polyadenylation

The tumorigenesis of small intestinal neuroendocrine tumors (SI-NETs) is poorly understood. Recent studies have associated alternative polyadenylation (APA) with proliferation, cell transformation, and cancer. Polyadenylation is the process in which the pre-messenger RNA is cleaved at a polyA site and a polyA tail is added. Genes with two or more polyA sites can undergo APA. This produces two or more distinct mRNA isoforms with different 3′ untranslated regions. Additionally, APA can also produce mRNAs containing different 3′-terminal coding regions. Therefore, APA alters both the repertoire and the expression level of proteins. Here, we used high-throughput sequencing data to map polyA sites and characterize polyadenylation genome-wide in three SI-NETs and a reference sample. In the tumors, 16 genes showed significant changes of APA pattern, which lead to either the 3′ truncation of mRNA coding regions or 3′ untranslated regions. Among these, 11 genes had been previously associated with cancer, with 4 genes being known tumor suppressors: DCC, PDZD2, MAGI1, and DACT2. We validated the APA in three out of three cases with quantitative real-time-PCR. Our findings suggest that changes of APA pattern in these 16 genes could be involved in the tumorigenesis of SI-NETs. Furthermore, they also point to APA as a new target for both diagnostic and treatment of SI-NETs. The identified genes with APA specific to the SI-NETs could further be tested as diagnostic markers and drug targets for disease prevention and treatment.
referred to either 3′ truncation of mRNA coding regions or 3′ UTRs.

MATERIALS AND METHODS

HUMAN CLINICAL SAMPLES AND RNA ISOLATION

Surgical specimens from SI-NETs were obtained from three patients undergoing surgery at the Department of Surgical Gastroenterology, Rigshospitalet (Table 1). The study was approved by the regional scientific ethical committee (journal number 01 313726). Signed, informed consent was obtained from the patients. Immediately after tumor resection, tumor tissue samples from the surgical specimens were placed in RNAlater® (Applied Biosystems, Carlsbad, CA, USA) for overnight incubation. Samples were subsequently stored at −80°C until RNA was extracted using TRIzol (Life Technologies, Carlsbad, CA, USA). The reference sample (PIT), human pituitary RNA, was purchased from BioChain (Hayward, CA, USA).

SEQUENCING

Total RNA from the three SI-NETs and the reference sample (PIT) were sequenced by single-molecule direct RNA-sequencing (DRS) at Helicos BioSciences (Boston, MA, USA) (17). The 3′ blocking reaction was performed with 3′ deoxyATP (Jena Biosciences, Germany) by incubating the reaction mixture at 37°C for 30 min.

READ PROCESSING AND MAPPING

Raw DRS reads starting with one or more Ts were trimmed as these can be due to inefficient locking of the RNA with the polyT primer at the beginning of the sequencing (17). Trimmed reads were mapped to the genome (hg19) with HeliSphere software (http://sourceforge.net/projects/openhelisphere/) using a seed length of 18 nt and allowing only one mismatch in the seed region. Only uniquely mapped reads with a minimum length of 21 nt and a minimum score of 4 were retained for further analysis. We additionally discarded all reads that may arise from internal priming events. DRS reads are sequenced by synthesis, and thus, they are reverse complementary to the original RNA fragments (17). We checked the 10-nt immediately downstream of the read end in the complementary strand for adenines. In the case that the first six or more of these nucleotides were adenines the read was discarded (14). A summary of the data after the different steps can be seen in Table 2.

IDENTIFICATION OF polyA SITES

It is known that the cleavage process is inefficient and can vary within a few nucleotides (18). To account for this variability, we cluster iteratively the 5′ end of mapped reads that were >24 nt apart (14,18). For each of these clusters, the cleavage site is defined as the median position of the cleavage sites in the cluster (Figure 2). As previous works have shown that APA is very common and that even very infrequently used polyA sites are real (7), we kept all polyA sites that have at least one read in at least two samples, or that have two or more reads. Next, each of the clusters was annotated according to the location of the cleavage site using the genome annotation from Ensembl66 (19). PolyA sites were classified as 5′ UTR, CODING, or 3′ UTR if overlapping these regions of coding genes; ncGENE if overlapping a non-coding gene; PROMOTER if they were located in the 1000 bases upstream of a gene; DOWNSTREAM if they were located in the 1000 bases downstream of a gene; and INTERGENIC otherwise (Figure 3A). PolyA sites found in both introns and exons of the coding region are classified as CODING, as usage of these changes the coding region of the mRNA. As DOWNSTREAM polyA sites (≤1 kB) have many more reads than other intergenic polyA sites (Figure 3B), we associated the polyA sites in this proximal intergenic region (≤1 kB) to the upstream gene and included them in the analysis of APA.

IDENTIFICATION OF APA IN SI-NETs

We restricted the analysis to protein coding genes. For each gene, we selected only those polyA sites that have at least 10% of all reads
Table 1 | SI-NET biopsies.

| Identifier | Tissue of origin | Sex/age | Synaptophysin | Chromogranin | Serotonin | CD117 |
|------------|------------------|---------|---------------|--------------|-----------|-------|
| NE MTT     | Ileum            | M/73    |               |              |           |       |
| NE CT1     | Ileum            | F/66    |               |              | NA        | NA    |
| NE 2TC     | Ileum            | M/73    |               |              |           |       |

Table 2 | Read processing statistics.

|          | NE MTT | NE CT1 | NE 2TC | PIT  |
|----------|--------|--------|--------|------|
| Raw reads| 4895411| 1087879| 8859412| 9773879|
| Uniquely mapped reads (>21 nt) | 2012076| 4465454| 3675490| 4027944|
| Reads filtered after removing internal priming events | 1918906| 4266418| 3499253| 3816258|
| Clusters (including singletons) | 686450 | 1352151| 1296130| 1289606|
| Polya sites | 221147 | 398528 | 379890 | 385451 |

RESULTS

GLOBAL CHARACTERIZATION OF POLYA SITE USAGE REVEALS SIMILARITY BETWEEN SI-NETS

Our first objective was to map polya sites in SI-NETs genomewide and characterize APA. To perform this characterization, total RNA from three SI-NET samples (NE MTT, NE CT1, and NE 2TC) and from a reference sample (PIT) were sequenced using
FIGURE 3 | (A) Distribution of identified polyA sites across the genome for each of the samples analyzed. (B) Average number of reads per polyA site in different genomic regions.

DRS (17). DRS is a well established method, which have been used in several other studies to map polyA sites genome-wide (14, 17, 20, 21). The SI-NETs originate from neuroendocrine cells of the small intestine, which are solitary interspersed cells (22). It is very difficult to identify and collect enough normal neuroendocrine cells as a control sample, and scrapings of the mucosa adjacent to the tumors are thus normally used as reference samples (23). However, we chose to use pituitary as reference sample as we expected that the neuroendocrine cells of the pituitary are phenotypically closer to the neuroendocrine cells of the small intestine than the mixture of different cell types from the mucosal scrapings. From the DRS we obtained a total of 4895411 (NE MTT), 10878794 (NE CT1), and 8859412 (NE 2TC) reads. The final filtered reads were mapped to 221147 (NE MTT), 398528 (NE CT1), and 379890 (NE 2TC) polyA sites. These sites mapped to various genomic regions in the samples (Figure 3). Fisher exact tests found 11031 polyA sites, located in 97% of the expressed protein coding genes, with no statistical difference in usage between the individual SI-NETs (FDR >0.1). This shows similarity between the three SI-NET samples. The sequencing data obtained by DRS have been deposited to the National Center for Biotechnology Information (NCBI)’s Gene Expression Omnibus (GEO), with accession number: GSE56657.

ALTERNATIVE USAGE OF polyA SITES IN THE SI-NETs

The second aim of the project was to identify protein coding genes in SI-NETs with a differential APA pattern compared to the reference sample. Using fisher exact tests we identified 61 genes with statistically significant APA between the SI-NET group and the reference sample (Table S1 in Supplementary Material). From these 61 genes, 45 genes were removed manually, to select only the genuine cases of APA, which lead to either 3' truncation of mRNA coding regions or 3' UTRs (Table S2 in Supplementary Material). The 16 selected genes with APA specific to the SI-NETs are described below, including a description of the possible implications for protein isoform and expression (Table S3 and Figure S1 in Supplementary Material). Interestingly, 11 out of these 16 genes have previously been associated with cancer, with 4 of the genes being known tumor suppressors. These 16 genes with APA specific to the SI-NETs could be tested as diagnostic markers and drug targets for disease prevention and treatment.

GENES WITH UTR-APA SPECIFIC TO THE SI-NETs

Of the 16 genes with APA specific to the SI-NETs, 9 genes had significant UTR-APA leading to 3' UTR shortening in a larger proportion of mRNAs than in the reference sample. These genes are: PSMD8, DACT2, TM9SF3, CD59, ANKH, CIAO1, SRSF5, MRSP16, and NDUFA6. The shortening of 3' UTRs plays an important part in post-transcriptional regulation, as the 3' UTR contains target sites for microRNAs (miRNAs), which control mRNA turnover and translation rate (24). Because of miRNA target site loss, mRNAs with shorter 3' UTRs are more stable and produce more protein (5, 11, 12). We mapped the target sites of all human miRNAs to the 3' UTRs of genes displaying UTR-APA (Figure S2 in Supplementary Material). The miRNA target site predictions were downloaded from microRNA.org (25). Only conserved miRNAs
with a good score were mapped. We then compared the miRNA target sites in the 3′ UTRs of the nine genes to miRNA-profiles of five primary SI-NET samples from another study (26). We considered only the miRNAs that were detected in a minimum of four out of these five primary SI-NET samples. For all nine genes, except DACT2, the 3′ UTR shortening leads to loss of miRNA target sites for miRNAs detected in SI-NETs, suggesting an upregulated expression for these genes. See Table 3 for a detailed description of the genes with UTR-APA.

### GENES WITH CR-APA SPECIFIC TO THE SI-NETs

Seven of the genes identified had significant CR-APA leading to 3′ truncated coding regions in a larger proportion of mRNAs than in the reference sample. These genes are: DCC, PDZD2, MAGI1, ZCWPW2, LRREFP1, RIC3, and THADA. CR-APA does not lead to nonsense-mediated decay and the expressed miRNAs thus translate into C-terminally truncated protein isoforms that could be non-functional or have functions different from the full-length protein (33). The seven genes with CR-APA are described in detail in Table 4.

### Q-RT-PCR VALIDATION OF APA

Three genes were selected for Q-RT-PCR validation: DCC, PDZD2, and LRREFP1. cDNA was synthesized from total RNA from all four samples. Primers were designed as sets targeting the regions just upstream of a proximal and a distal polyA site with significant

### Table 3 | Genes with UTR-APA specific to the SI-NETs.

| Gene   | Function                                      | Role in cancer | APA                  | Functional consequence                                                                 |
|--------|-----------------------------------------------|----------------|----------------------|----------------------------------------------------------------------------------------|
| **PSMD8** | Subunit of the proteasome                     | Overexpressed in breast cancer (27) | Increased usage of a proximal polyA site in the 3′ UTR | Shortening of 3′ UTR by 325 nt with loss of miRNA target sites for three miRNAs present in SI-NETs: hsa-mir-22*, hsa-mir-192*, and hsa-mir-1263 |
| **DACT2** | Negative regulator of the TGF-β/nodal signaling pathway | Tumor suppressor (28) | Increased usage of two proximal polyA sites in the 3′ UTR | Shortening of 3′ UTR by up to 12113 nt without loss of miRNA target sites |
| **TM9SF3** | Transmembrane protein with unknown function    | Unknown         | Increased usage of two proximal polyA sites in the 3′ UTR | Shortening of 3′ UTR by up to 2848 nt with loss of miRNA target sites for 13 miRNAs present in SI-NETs: hsa-mir-1826, hsa-mir-338-5p, hsa-mir-552, hsa-mir-352-3p, hsa-mir-7-2*, hsa-mir-7-1*, hsa-mir-10b*, hsa-mir-140-3p, hsa-mir-193b*, hsa-mir-550*, hsa-mir-200c*, hsa-mir-17*, and hsa-mir-93* |
| **CD59**  | Membrane complement regulatory protein         | Overexpressed and anti-apoptotic (29) | Increased usage of two proximal polyA sites in the 3′ UTR | Shortening of 3′ UTR by up to 4707 nt with loss of miRNA target sites for one miRNA present in SI-NETs: hsa-mir-629 |
| **ANKH**  | Pyrophosphate transport regulator              | Amplified and suggested overexpressed (30) | Increased usage of a proximal polyA site in the 3′ UTR | Shortening of 3′ UTR by up to 4879 nt with loss of miRNA target sites for eight miRNAs present in SI-NETs: hsa-mir-330-3p, hsa-mir-29b-2*, hsa-mir-1275, hsa-mir-625, hsa-mir-452, hsa-mir-628-3p, hsa-mir-30a*, and hsa-mir-30e* |
| **CIAO1** | Role in iron-sulfur protein biogenesis         | Interacts with tumor suppressor WT1 (31) | Increased usage of two proximal polyA sites in the 3′ UTR | Shortening of 3′ UTR by up to 2581 nt with loss of miRNA target sites for nine miRNAs present in SI-NETs: hsa-mir-550*, hsa-mir-200c*, hsa-mir-214*, hsa-mir-423-3p, hsa-mir-680, hsa-mir-29b-1*, hsa-mir-1246, hsa-mir-7-1*, and hsa-mir-7-2* |
| **SRSF5** | Splice factor                                  | Overexpressed in breast cancer and anti-apoptotic (32) | Increased usage of a proximal polyA site in the 3′ UTR | Shortening of 3′ UTR by 312 nt with loss of miRNA target sites for seven miRNAs present in SI-NETs: hsa-mir-338-5p, hsa-mir-628-3p, hsa-mir-409-3p, hsa-mir-183*, hsa-mir-330-3p, hsa-mir-7-1*, and hsa-mir-7-2* |
| **MRSP16** | Mitochondrial ribosomal protein                | Unknown         | Increased usage of a proximal polyA site in the 3′ UTR | Shortening of 3′ UTR by 1888 nt with loss of miRNA target sites for 16 miRNAs present in SI-NETs: hsa-mir-671-5p, hsa-mir-193b*, hsa-mir-651b*, hsa-mir-338-5p, hsa-mir-30b*, hsa-mir-125b-2*, hsa-mir-330-3p, hsa-mir-7-1*, hsa-mir-7-2*, hsa-mir-188-5p, hsa-mir-769-5p, hsa-mir-324-3p, hsa-mir-552, hsa-mir-193-5p, hsa-mir-501-5p, and hsa-mir-362-5p |
| **NDUFA6** | NADH dehydrogenase                            | Unknown         | Increased usage of a proximal polyA site in the 3′ UTR | Shortening of 3′ UTR by 563 nt with loss of miRNA target sites for six miRNAs present in SI-NETs: hsa-mir-552, hsa-mir-21*, hsa-mir-664*, hsa-mir-505*, and hsa-mir-500* |
Table 4 | Genes with CR-APA specific to the SI-NETs.

| Gene      | Function                                      | Role in cancer                      | APA compared                                                                 | Functional consequence                                                                 |
|-----------|-----------------------------------------------|-------------------------------------|------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| DCC       | Receptor for the axon guidance molecule nextrin | Tumor suppressor (34)               | Increased usage of a proximal polyA site located in intron 5                  | Usage of this polyA site leads to expression of an mRNA lacking 24 exons in the 3' end, which translates into a highly C-terminally truncated DCC protein containing only 5 Ig-like domains and lacking both the nextrin binding domain, the transmembrane domain, and the intracellular domain. Thus, it lacks the parts responsible for the tumor suppressive functions (34) |
| PDZD2     | PDZ domain-containing protein                 | Can be cleaved near its C-terminus to generate a secreted form sPDZD2 with tumor suppressive function (35, 36) | Increased usage of two proximal polyA sites located in intron 2                | Usage of these polyA sites leads to expression of an mRNA lacking 23 exons in the 3' end, which translates a highly C-terminally truncated PDZD2 protein, lacking the ability to form sPDZD2 and thus the tumor suppressive function (35) |
| MAGI1     | Scaffolding protein localized at cell-cell contacts, containing PDZ domains | Tumor suppressive and anti-metastatic, through β-catenin recruitment to the cell membrane and thus inhibition of the Wnt/β-catenin signaling pathway (37) | Increased usage of two proximal intronic polyA sites located in the intron 2 | Usage of these polyA sites leads to expression of an mRNA lacking 21 exons in the 3' end, which translates a highly C-terminally truncated MAGI1 protein, lacking all PDZ domains and thus the tumor suppressive function (37) |
| ZCWPW2    | Zinc finger protein                           | Unknown                             | Increased usage of a proximal polyA site located in intron 5                  | Usage of this polyA site leads to expression of an mRNA lacking five exons in the 3' end, which translates into a highly C-terminally truncated ZCWPW2 protein |
| LRRFIP1   | Transcriptional repressor                     | Overexpressed in cancer cell lines (39) and pro-metastatic (39) | Increased usage of a proximal polyA site located in the fifth last intron | Usage of this polyA site leads to the expression of an mRNA lacking five exons in the 3' end. The 3'-end of this mRNA isoform corresponds to the 3'-end of the 4.2-kB mRNA overexpressed in cancer cell lines (39) |
| RIC3      | Chaperone protein                             | Unknown                             | Increased usage of a proximal intronic polyA site located in intron 1         | Usage of this polyA site leads to the expression of an mRNA lacking five exons in the 3' end, giving rise to C-terminally truncated RIC3 protein |
| THADA     | Unknown                                       | Frequently truncated in thyroid adenomas and the truncated allele is hypothesized to play a role in the thyroid tumorigenesis (40) | Increased usage of a proximal polyA site located in intron 22                  | Usage of this polyA site leads to the expression of an mRNA lacking 16 exons in the 3' end, i.e., more than in the thyroid adenomas (40), which translates into highly C-terminally truncated THADA protein |

differential usage between the SI-NET group and PIT. The proximal primer sets targeted intronic sequences, corresponding to 3' UTRs of alternative last exons, and the distal primer sets targeted either intronic sequences, corresponding to 3' UTRs of alternative last exons, or the 3' UTRs of the last exon. As the proximal primer sets targeted intronic sequences, the detection of a product from these sets could only come from usage of the adjacent polyA site, as the intron is otherwise spliced out. Proximal and distal isoforms were thus individually measured. The genes were used as their own controls, as there were two sets of primers for each gene; one upstream of a proximal polyA site and one upstream of a distal polyA site. The Ct values of the proximal and distal amplicon were thus compared. Q-RT-PCR was run on triplicates of each sample with all six primer sets. Median Ct values from the triplicates were used for further analysis with individual Ct values deviating more than two from the median considered as outliers and removed (2 of the 72 Ct values removed). In contrast to our results obtained from DRS (Table 5), we found a higher expression of the proximal isoform of DCC and LRRFIP1, as well as a lower expression of the proximal isoform of PDZD2 in all three SI-NETs as well as in the reference sample (Table 6). However, when comparing the results for the SI-NETs with our reference sample, we found a higher expression of the proximal LRRFIP1 isoform in all three SI-NETs compared to our reference sample. For DCC and PDZD2, we found a higher expression of the proximal isoform in two out of the three SI-NETs compared to our reference sample (Table 7). These Q-RT-PCR results validate the presence of the polyA sites and the APA of these genes. They also indicate that these three
Table 5 | Expression ratio of (short isoform/long isoform) derived from DRS data.

|          | DCC proximal/ distal | PDZD2 proximal/ distal | LRRFIP1 proximal/ distal |
|----------|-----------------------|------------------------|--------------------------|
| NE 2TC   | 25                    | 14                     | 3.725490196              |
| NE CT1   | 156                   | 15.66666667            | 2.173469388              |
| NE MTT   | Inf                   | 36                     | 3.25                     |
| PIT      | 0.128205128           | 0.652631579            | 0.735294118              |

Table 6 | Expression ratio of (short isoform/long isoform) derived from Q-RT-PCR data.

|          | DCC proximal/ distal | PDZD2 proximal/ distal | LRRFIP1 proximal/ distal |
|----------|-----------------------|------------------------|--------------------------|
| NE 2TC   | 3.853506079           | 0.026305316            | 25.5536049               |
| NE CT1   | 2.887782327           | 0.012957265            | 14.2667269               |
| NE MTT   | 1.019581667           | 0.032685206            | 30.0892867               |
| PIT      | 2.030313016           | 0.025800759            | 3.645055529              |

Table 7 | Expression ratio of (short isoform/long isoform) normalized to reference sample, derived from Q-RT-PCR data.

|          | DCC             | PDZD2           | LRRFIP1           |
|----------|-----------------|-----------------|------------------|
| NE 2TC   | 1.897986        | 1.019556        | 7010524          |
| NE CT1   | 1.422334        | 0.502205        | 3.913947         |
| NE MTT   | 0.502313        | 1.265788        | 8.254755         |

genes indeed undergo differential APA leading to the expression of a larger amount of mRNAs with 3′ truncation of the coding region in the SI-NETs than in the reference sample.

**DISCUSSION**

This is the first study to map polyA sites and characterize APA genome-wide in SI-NETs. To map the polyA site usage genome-wide we sequenced total RNA from three SI-NET samples (NE MTT, NE CT1, and NE 2TC) and a reference sample (PIT) using DRS (17). DRS is a well established method, which have been used in several other studies to map polyA sites genome-wide (14, 17, 20, 21). Selected polyA sites have been confirmed in all cases with PCR in three of those studies (14, 17, 21). We selected three genes for Q-RT-PCR validation: DCC, PDZD2, and LRRFIP1. For these three genes we validated the presence of all tested polyA sites with Q-RT-PCR and thus the APA of the selected genes. However, we did not find the same ratio of usage between the proximal and distal polyA sites with Q-RT-PCR as we found with DRS (Tables 5 and 6). When comparing our Q-RT-PCR results in the SI-NETs with the reference sample, they however indicate that the three genes in the SI-NETs indeed undergo differential APA leading to expression of a larger amount of mRNAs with 3′ truncation of the coding region than in the reference sample (Table 7). The quantitative differences found in expression ratios of the short and long isoforms derived from DRS and Q-RT-PCR could be due to biases and artifacts introduced during the synthesis of cDNA, contamination of cDNA samples with genomic DNA and problems with one or more of the primer sets.

When studying SI-NETs the optimal control sample would be normal small intestinal neuroendocrine cells. These solitary cells, which are interspersed among the other cells of the mucosa, are however difficult to isolate (23). In one study, human neuroendocrine cells have been isolated by pronase/collagenase digestion, gradient centrifugation, and FACs, but this method is complicated and requires at least 5–8 cm of normal ileum (41). Additionally, it is not certain if the RNA content is damaged during those processes. Because of these difficulties, scrapings from the mucosa adjacent to the tumors are normally used as reference samples. This mucosa is however composed of a mix of various cell types, primarily enteroctyes and goblet cells. We believe that this mixture of cells from the mucosa, despite a common origin with the neuroendocrine cells (42), is phenotypically farther from intestinal neuroendocrine cells than neuroendocrine cells elsewhere in the body. We therefore chose to use pituitary as a reference sample as it mainly contains neuroendocrine cells.

Using Fisher exact tests, we found 11031 polyA sites, located in 97% of the expressed protein coding genes, with no statistical difference in usage between the individual SI-NETs (FDR >0.1). This shows similarity between the three SI-NET samples. Using the same Fisher exact tests to compare APA in the SI-NETs and in the reference sample, we identified 16 genes with significant APA specific to the SI-NETs, which lead to either 3′ truncation of mRNA coding regions or 3′ UTRs (Figure S1 in Supplementary Material). These 16 genes with APA specific to the SI-NETs could be tested as diagnostic markers and drug targets for disease prevention and treatment.

Nine of the 16 genes had significant UTR-APA in the SI-NETs leading to 3′ UTR shortening in a larger proportion of mRNAs than in the reference sample. Interestingly, four genes out of the nine with UTR-APA have been found to be either overexpressed in cancer: PSMD8 (27), CD59 (29), and SRSF5 (32) or is suggested to be overexpressed: ANKH (30). The UTR-APA in these genes produces mRNA isoforms that lack one or more target sites for miRNAs present in SI-NETs. These miRNAs consequently escape a part of the miRNA regulation and, therefore, should be more stable and produce more protein (5, 11, 12). Thus, the changes predicted in these nine genes suggest that UTR-APA could contribute to the SI-NET tumorigenesis. One of these genes, DACT2, was previously found as a tumor suppressor protein (28). The UTR-APA in this gene leads to 3′ UTR shortening in a larger proportion of mRNAs in the SI-NETs than in the reference sample.

However, as no mRNA target sites are found in the shortened part of the 3′ UTR, it is unclear which role this 3′ UTR region has. Apart from miRNAs, RNA binding proteins (RBPs) also target the 3′ UTR of mRNAs and can affect the stability of mRNAs (43). Thus, it is possible that the shortening of the 3′ UTR affects the stability of DACT2 by removing putative binding sites of RBPs.

Seven of the 16 genes had significant CR-APA in the SI-NETs leading to 3′ truncated coding regions in a larger proportion of mRNAs than in the reference sample. CR-APA alters the coding region of mRNAs and thus the isoform of the produced
protein. Remarkably, three out of the seven genes with CR-APA have tumor suppressive functions: DCC (34), PDZD2 (35), and MAGI1 (37). The CR-APA in these genes produces mRNA isoforms lacking several 3′-terminal exons, translating into highly C-terminally truncated proteins. These protein isoforms lack the parts responsible for the tumor suppressive functions. The CR-APA in DCC, PDZD2, and MAGI1 could thus likely be contributing to the tumorigenesis in SI-NETs. DCC is located in chromosome 18, where there have been reported frequent copy-number losses in SI-NETs (44, 45). The copy-number losses might truncate the DCC gene, forcing the usage of the proximal polyA site. Another gene with CR-APA, THADA, has been hypothesized to contribute to the tumorigenesis of thyroid adenomas in its truncated form (40). The CR-APA found in the SI-NETs produces an mRNA even more 3′-terminally truncated and could thus also be a candidate contributing to the tumorigenesis of SI-NETs. One gene stands out in the CR-APA group, LRRFIP1, as it has been shown to be overexpressed in cancer cell lines (38) and to be pro-metastatic (39). The CR-APA of this gene in the SI-NETs gives rise to an mRNA isoform, with a 3′-end corresponding to the 3′-end of the 4.2 kbp mRNA overexpressed in cancer cell lines (38). CR-APA of LRRFIP1 could thus also be a contributor to the tumorigenesis in SI-NETs, as the major mRNA isoform expressed in the SI-NETs is likely equal to that overexpressed in the cancer cell lines.

The suggested participation of APA in the tumorigenesis of SI-NETs could lead to new medical treatments of SI-NETs. This is particularly interesting as polyA sites of interest can be very specifically and effectively blocked and patterns of APA can thus be changed. Different techniques utilizing antisense elements have already been used to effectively inhibit polyA site usage. They include antisense oligonucleotides (46), siRNAs (47), and U1 modifications (48–52). Such techniques could, e.g., be used to force the expression of the full-length mRNAs of the tumor suppressor genes DCC, PDZD2, and MAGI1.

As this study was performed on only three SI-NET samples the findings should be confirmed in a larger cohort. Preferably, future studies on APA in SI-NETs should also include alternative reference tissues and use different methods for characterizing polyA(antisense) site usage. The findings could also be tested in other types of NETs. If the findings are confirmed, the effects of blocking the implicated proximal polyA sites with antisense elements should be tested experimentally.

In conclusion, we have mapped for the first time polyA sites genome-wide in SI-NETs. We found 11031 polyA sites, located in 97% of the expressed protein coding genes, with no statistical difference in usage between the individual SI-NETs. This shows similarity between the three SI-NET samples. In the SI-NETs, we identified 16 genes with significant APA leading to either 3′ truncation of coding regions or 3′ UTRs in a larger proportion of mRNAs than in the reference sample (Tables 3 and 4). The APA was confirmed in three genes using Q-RT-PCR: DCC, PDZD2, and LRRFIP1. However, we did not find the same expression ratios between the proximal and distal isoforms for these genes with Q-RT-PCR as with DRS (Tables 5 and 6).

As 11 of the 16 identified genes have previously been associated with cancer, it is likely that APA in these genes could take part in the SI-NET tumorigenesis. Future studies will have to confirm these findings in a larger cohort. The identified 16 genes with APA specific to the SI-NETs could be tested as diagnostic markers and as drug targets for disease prevention and treatment.

AUTHOR CONTRIBUTIONS
Study concept and design (Anders Rehfeld, Lennart Friis-Hansen). Acquisition of data (Anders Rehfeld, Lennart Friis-Hansen). Analysis and interpretation of data (Anders Rehfeld, Mireya Plass, Anders Krogh, Lennart Friis-Hansen). Drafting of the manuscript (Anders Rehfeld, Mireya Plass, Lennart Friis-Hansen). Critical revision of the manuscript for important intellectual content (Anders Rehfeld, Mireya Plass, Kristina Dissing, Ulrich Knigge, Andreas Kjer, Anders Krogh, Lennart Friis-Hansen). Statistical analysis (Mireya Plass, Anders Krogh). Obtained funding (Anders Rehfeld, Anders Krogh, Lennart Friis-Hansen). Administrative, technical, or material support (Lennart Friis-Hansen). Study supervision (Lennart Friis-Hansen).

ACKNOWLEDGMENTS
The work in laboratories of the authors was supported by the Danish MRC, Rigshospitalet. Funding: this work was supported by, The Danish Council for Independent Research – Medical Sciences, The Danish Council for Strategic Research, The Novo Nordisk Foundation, and Rigshospitalet.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at http://www.frontiersin.org/Journal/10.3389/fendo.2014.00046/abstract

REFERENCES
1. Oberk B, Kniege U, Kvekkeboon D, Perren A. Neuroendocrine gastro-enteropancreatic tumors: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol (2012) 23(Suppl 7):vii124–38. doi:10.1093/annonc/mds295
2. Janson ET, Sørbye H, Welin S, Federispial B, Gronbaek H, Hellman P, et al. Nordic guidelines 2010 for diagnosis and treatment of gastroenteropancreatic neuroendocrine tumours. Acta Oncol (2010) 49:740–56. doi:10.3109/0282484X.2010.492791
3. Lawrence B, Gustafsson BI, Chan A, Svejda B, Kidd M, Modlin IM. The epidemiology of gastroenteropancreatic neuroendocrine tumors. Endocr Relat Cancer (2011) 18:1–18. doi:10.1530/erc.110.012805
4. Schimmack S, Svejda B, Lawrence B, Kidd M, Modlin IM. The diversity and commonalities of gastroenteropancreatic neuroendocrine tumors. Langenbecks Arch Surg (2011) 396:273–9. doi:10.1007/s00423-011-0739-1
5. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. Science (2008) 320:1643–7. doi:10.1126/science.1155390
6. Ji Z, Tian B. Reprogramming of 3′ untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. PLoS One (2009) 4:e4619. doi:10.1371/journal.pone.0008419
7. Derti A, Garrett-Engele P, Maciasac KD, Stevens RC, Siriam S, Chen R, et al. A quantitative atlas of polyadenylation in five mammal genomes. Genome Res (2012) 22:1173–83. doi:10.1101/gr.132563.111
8. Danckwardt S, Hentez MW, Kaluzik AE. 3′ End mRNA processing: molecular mechanisms and implications for health and disease. EMBO J (2008) 27:482–98. doi:10.1038/mbio.2007.1932
9. Rehfeld A, Plass M, Krogh A, Friis–Hansen L. Alterations in polyadenylation and its implications for endocrine disease. Front Endocrinol (Lausanne) (2013) 4:53. doi:10.3389/fendo.2013.00053
10. Elkouk R, Drost J, van Haften G, Jenal M, Schrier M, Vrielink JAO, et al. E2F regulation of target gene expression by a modified U1 snRNA. *Nucleic Acids Res* (2012) 40:4860–71. doi:10.1093/nar/gks467

11. Lin Y, Li Z, Ozsolak F, Kim SW, Arango-Argoty G, Liu TT, et al. Global changes in processing of mRNA 3′ untranslated regions characterize clinically distinct cancer subtypes. *Cancer Res* (2009) 69:9422–30. doi:10.1158/0008-5472.CAN-09-2236

12. Mayr C, Bartel DP. Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncomed genes in cancer cells. *Cell* (2009) 138:673–84. doi:10.1016/j.cell.2009.06.016

13. Morris AR, Bos A, Diosdado B, Rooijers K, Elkouk R, Bolijn AS, et al. Alternative cleavage and polyadenylation during colorectal cancer development. *Clin Cancer Res* (2012) 18:5256–66. doi:10.1158/1078-0432.CCR-12-0543

14. Lin Y, Li Z, Ozsolak F, Kim SW, Arango-Argoty G, Liu TT, et al. An in-depth map of polyadenylation sites in cancer. *Nucleic Acids Res* (2012) 40:8460–71. doi:10.1093/nar/gks467

15. Fu Y, Sun X, Li Y, Li J, Rao X, Chen C, et al. Differential genome-wide profiling of tandem 3′ UTRs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res* (2011) 21:741–7. doi:10.1101/gr.115295.110

16. Lembo A, Di Cunto F, Provero P. Shortening of 3′UTRs correlates with poor prognosis in breast and lung cancer. *PloS One* (2012) 7:e31129. doi:10.1371/journal.pone.0031129

17. Ozsolak F, Platt AR, Jones DR, Reifenberger JG, Sass LE, McInerney P, et al. Alternative polyadenylation in SI-NETs. *Genome Res* (2011) 21:749–57. doi:10.1101/gr.115033.110

18. Singh P, Alley TL, Wright SM, Kamdar S, Schott W, Wilpan RY, et al. Global genetic alterations in cervical cancer. *BMC Genomics* (2007) 8:53. doi:10.1186/1471-2164-8-53

19. Duman-Scheel M. Deleted in colorectal cancer (DCC) pathfinding: axon guidance gene finally turned tumor suppressor. *Curr Drug Targets* (2012) 13:1445–53. doi:10.2174/13894501280350215

20. Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S, et al. Ensembl 2013. *Nucleic Acids Res* (2013) 41:D48–55. doi:10.1093/nar/gks1236

21. Sherstnev A, Duc C, Cole C, Zacharaki V, Hornyik C, Ozsolak F, et al. Alternative polyadenylation. *Cell* (2001) 104:1018–29. doi:10.1016/S0017-0215(01)00368-9

22. Abrahamsen E, Cui J, Naka H, Faidi M, Wadleigh M, Thakker RV, et al. Identification of MAGI1 as a tumor-suppressor protein induced by cyclooxygenase-2 inhibitors in colorectal cancer cells. *Oncogene* (2012) 31:48–59. doi:10.1038/onc.2011.218

23. Leja J, Essaghir A, Essand M, Wester K, Oberg K, Tötterman TH, et al. Novel conserved targets of microRNAs. *Nature* (2009) 459:141–7. doi:10.1038/nature08130

24. Friedman RC, Farh KK-H, Burge CB, Bartel DP. Most mammalian mRNAs are targets of microRNAs. *Cell* (2009) 144:677–86. doi:10.1016/j.cell.2011.06.012

25. Betel D, Wilson M, Gabow A, Marks DS, Sandër C. The microRNA.org resource: targets and expression. *Nucleic Acids Res* (2008) 36:D149–53. doi:10.1093/nar/gkm595

26. Ahlman H, Nilsson O. The gut as the largest endocrine organ in the body. *Ann Oncol* (2001) 12(Suppl 2):S63–8. doi:10.1093/annonc/12.suppl_2.S63

27. Akerfelt M, Lehtonen P, Leskela I, Salovaara R, Yli-Harja O, et al. Inhibiting expression of specific genes in mammalian cells with small interfering RNAs targeted to terminal exons of pre-mRNA. *Proc Natl Acad Sci U S A* (2003) 100:8264–9. doi:10.1073/pnas.1332691100

28. Vickers TA, Crooke ST. siRNAs targeted to certain polyadenylation sites promote inhibition. *Nucleic Acids Res* (2012) 40:2815–25. doi:10.1128/MCB.21.8.2815-2825.2011

29. Liu P, Pentlicky S, Jung SP, Fortes P, Vickers TA, Crooke ST. Inhibiting expression of specific genes in mammalian cells with small intronic polyA RNAs. *Nucleic Acids Res* (2012) 40:gk529. doi:10.1093/nar/gk529

30. Fortes P, Cuevas Y, Gafni G, Lefkowitz E, Provencher S, Jung SP, et al. Inhibiting expression of specific genes in mammalian cells with small nuclear small RNAs targeted to terminal exons of pre-mRNA. *Proc Natl Acad Sci U S A* (2003) 100:8264–9. doi:10.1073/pnas.1332691100

31. Modlin IM, Oberg K, Chung DC, Jensen RT, de Herder WW, Thakker RV, et al. The functional genomic profiling reveals gain of chromosome 14 as a predictor of poor outcome in classical Hodgkin lymphoma. *Clin Exp Immunol* (2010) 161:367–74. doi:10.1111/j.1365-2249.2010.04166.x

32. Rippe V, Drieschner M, Neimboos M, Murru Escobar H, Bonk U, Belge G, et al. Identification of a gene rearranged by 2p21 aberrations in thyroid adenomas. *Oncogene* (2003) 22:611–4. doi:10.1038/sj.onc.1206867

33. Schwerdtle M, Pfranger R, Eick GN, Champaneria MC. The functional characterization of normal and neoplastic human enterochromaffin cells. *J Clin Endocrinol Metab* (2006) 91:2340–8. doi:10.1210/jc.2005-0110

34. Hashemi J, Fotouhi O, Sulaiman L, Kjellman M, Höög A, Zedenius J, et al. Identification of antagonistic soluble decoy receptor tyrosine kinases by intronic polyA inclusion. *Nucleic Acids Res* (2012) 40:gk529. doi:10.1093/nar/gk529

35. Johnstone RW, Wang J, Tommerup N, Vissing H, Roberts T, Shi Y. Ciao 1 is a novel WD40 protein that interacts with the tumor suppressor protein WTI. *J Biol Chem* (1998) 273:10880–7. doi:10.1074/jbc.273.18.10880

36. Vickers TA, Crooke ST. siRNAs targeted to certain polyadenylation sites promote inhibition. *Nucleic Acids Res* (2012) 40:2815–25. doi:10.1128/MCB.21.8.2815-2825.2011

37. Blazquez L, Gonzalez-Rojas SJ, Abad A, Rasaquin N, Abad X, Fortes P. Increased in vivo inhibition of gene expression by combining RNA interference and U1 inhibition. *Nucleic Acids Res* (2012) 40:gk529. doi:10.1093/nar/gk529
51. Goraczniak R, Behlke MA, Gunderson SI. Gene silencing by synthetic U1 adaptors. Nat Biotechnol (2009) 27:257–63. doi:10.1038/nbt.1525

52. Vickers TA, Sabripour M, Crooke ST. U1 adaptors result in reduction of multiple pre-mRNA species principally by sequestering U1snRNP. Nucleic Acids Res (2011) 39:e71. doi:10.1093/nar/gkr150

53. Di Giammartino DC, Nishida K, Manley JL. Mechanisms and consequences of alternative polyadenylation. Mol Cell (2011) 43:853–66. doi:10.1016/j.molcel.2011.08.017

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 February 2014; accepted: 22 March 2014; published online: 15 April 2014.
Citation: Rehfeld A, Plass M, Døssing K, Knigge U, Kjær A, Krogh A and Friis-Hansen L (2014) Alternative polyadenylation of tumor suppressor genes in small intestinal neuroendocrine tumors. Front. Endocrinol. 5:46. doi: 10.3389/fendo.2014.00046
This article was submitted to Cancer Endocrinology, a section of the journal Frontiers in Endocrinology. Copyright © 2014 Rehfeld, Plass, Døssing, Knigge, Kjær, Krogh and Friis-Hansen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.