Supplementary Information

Identification of *TMEM115* as a tumor suppressor gene through *TERT* regulation in pancreatic cancer

Yu Sakano[1, 2], Haruka Ui[2], Takuki Yagyu[1], Takahito Ohira[2, 3], Tomohiko Sakabe[4], Teruhisa Sakamoto[1], Yoshihisa Umekita[4], Yoshiyuki Fujiwara[1], Hiroyuki Kugoh[2, 3] *

[1] Division of Gastrointestinal and Pediatric Surgery, Department of Surgery, School of Medicine, Faculty of Medicine, Tottori University, Yonago, Tottori, Japan

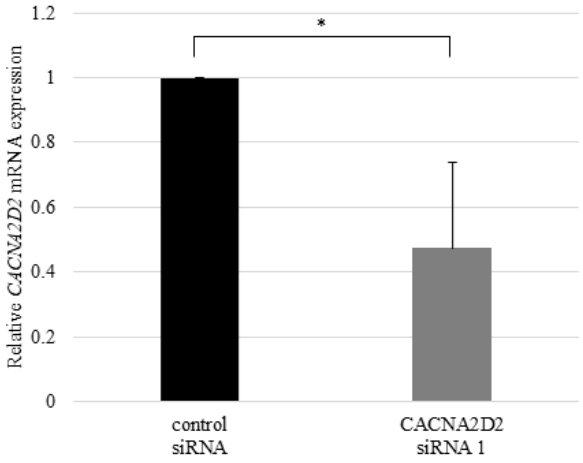
[2] Department of Molecular and Cellular Biology, Division of Genome and Cellular Function, Tottori University, Yonago, Tottori, Japan

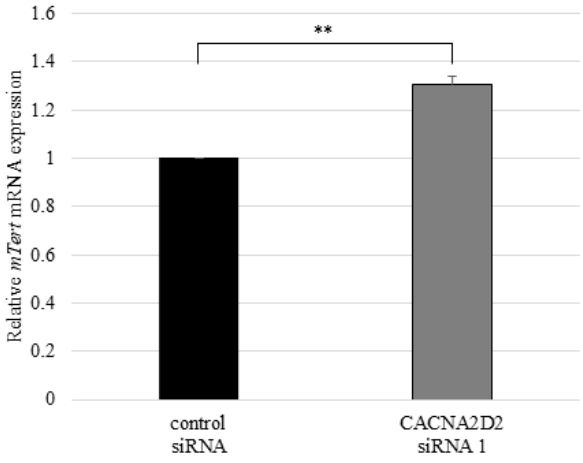
[3] Chromosome Engineering Research Center, Tottori University, Yonago, Tottori, Japan

[4] Department of Pathology, School of Medicine, Faculty of Medicine, Tottori University, Yonago, Japan

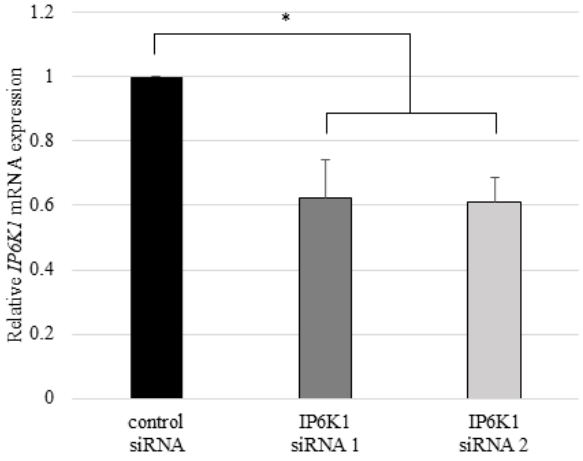
* Corresponding author: kugoh@tottori-u.ac.jp

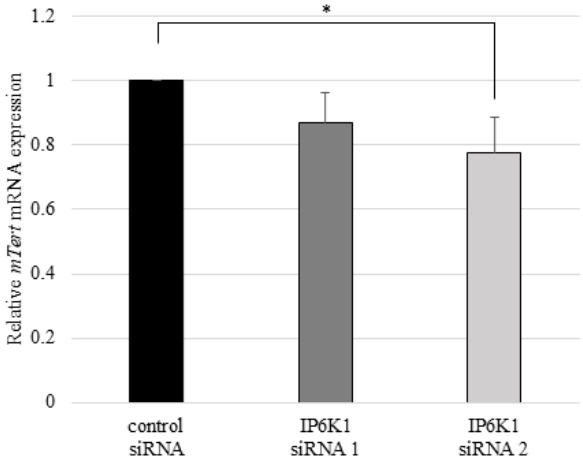
CACNA2D2



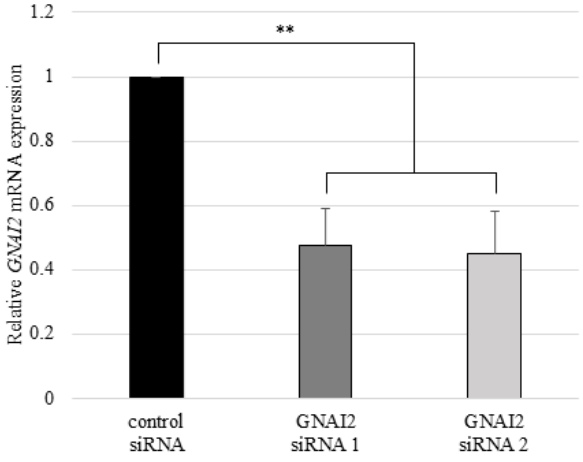


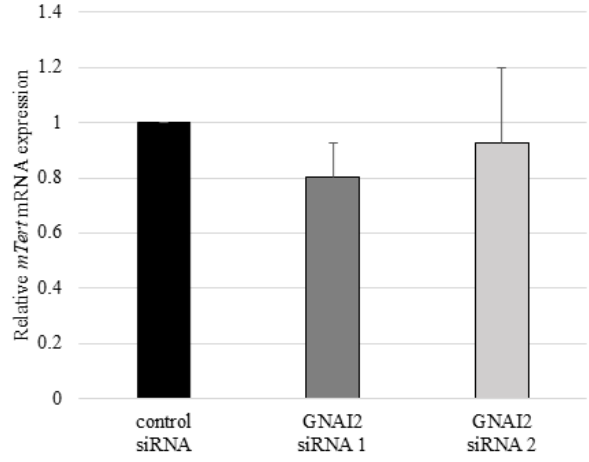
IP6K1

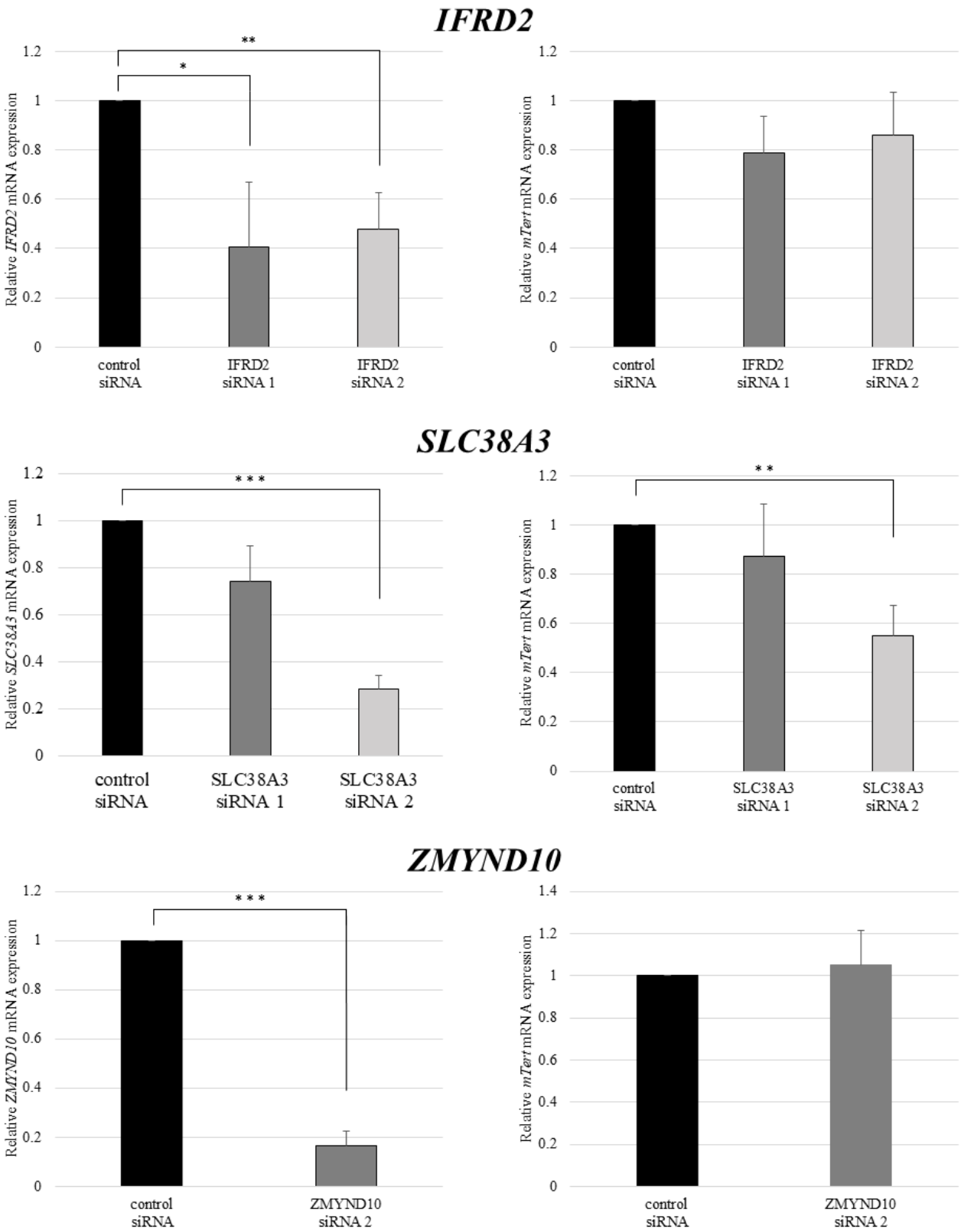




GNAI2

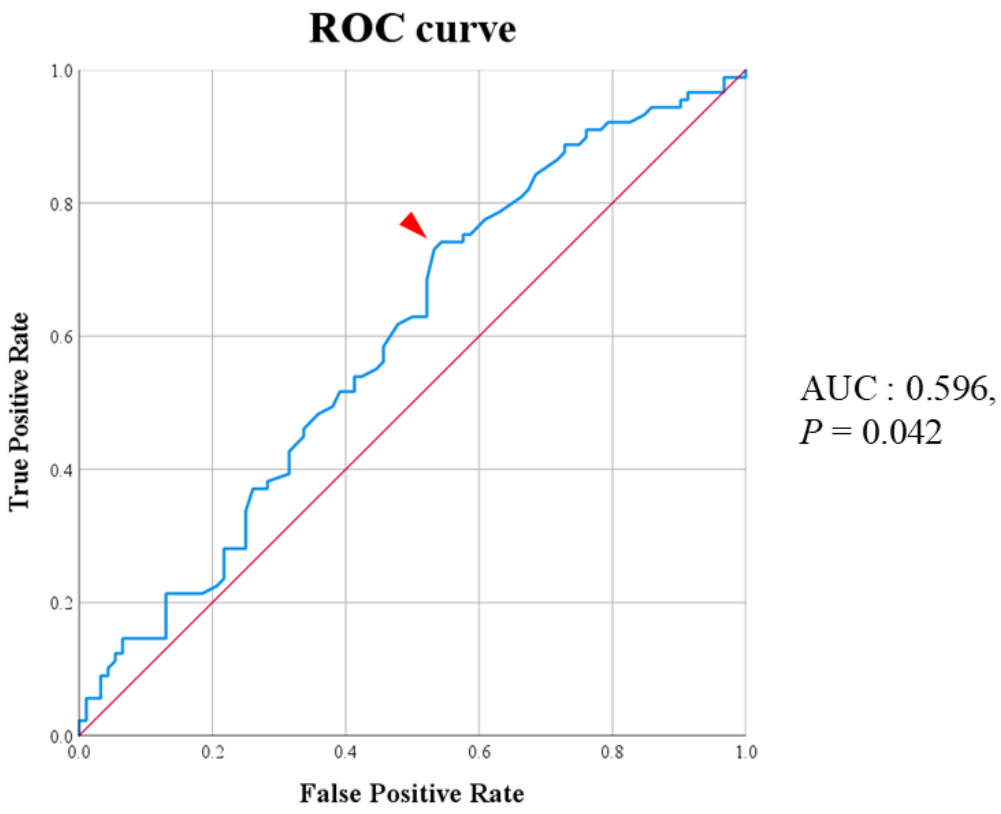





IFRD2
Relative IFRD2 mRNA expression
1.2
1
0.8
0.6
0.4
0.2
0
**
*
control siRNA
IFRD2 siRNA 1
IFRD2 siRNA 2
Relative mTert mRNA expression
1.2
1
0.8
0.6
0.4
0.2
0
control siRNA
IFRD2 siRNA 1
IFRD2 siRNA 2
SLC38A3
Relative SLC38A3 mRNA expression
1.2
1
0.8
0.6
0.4
0.2
0
***
control siRNA
SLC38A3 siRNA 1
SLC38A3 siRNA 2
Relative mTert mRNA expression
1.2
1
0.8
0.6
0.4
0.2
0
**
control siRNA
SLC38A3 siRNA 1
SLC38A3 siRNA 2
ZMYND10
Relative ZMYND10 mRNA expression
1.2
1
0.8
0.6
0.4
0.2
0
***
control siRNA
ZMYND10 siRNA 2
Relative mTert mRNA expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
control siRNA
ZMYND10 siRNA 2

Supplementary Figure S1. Knockdown analysis of the candidate genes other than transmembrane protein 115 (*TMEM115*) and subsequent examination of the effects on mouse telomerase reverse transcriptase (*mTert*) transcription.

Quantitative RT-PCR analysis of the mRNA expression levels of the six candidate genes in the LTPA 3p21.3-HAC cells transfected with the control small interfering RNA (siRNA) or siRNA 1 or 2 targeting each gene. Because of the strong cytotoxicity caused by transfection of the *CACNA2D2* and *ZMYND10* siRNAs, only one type of siRNA was used in the knockdown analysis of these two genes. Quantitative RT-PCR analysis was also performed for the relative *mTert* mRNA expression levels in these cells. The *mGapdh* mRNA expression levels were used as the internal control. Data are presented as the mean ± standard deviation of three independent experiments (***$P < 0.001$, **$P < 0.01$, *$P < 0.05$).



Supplementary Figure S2. Receiver operating characteristic (ROC) curve of the transmembrane protein 115 (*TMEM115*) mRNA expression levels in pancreatic cancer (PC) patients from The Cancer Genome Atlas (TCGA) database for overall survival (OS).

ROC curve of the *TMEM115* mRNA expression levels in PC patients from the TCGA database for OS. The arrow indicates the optimal cutoff value. AUC, area under the ROC curve.

Supplementary Table S1. The list of 26 genes encoded in the 3p21.3 region that were significantly upregulated in the LTPA 3p21.3-HAC cells.

	Gene ID	Gene Symbol	*P* -value
1	ENSG00000012171	SEMA3B	$9.75\acute{}10^{-70}$
2	ENSG00000164078	MST1R	$6.32\acute{}10^{-62}$
3	ENSG00000114353	GNAI2	$3.49\acute{}10^{-56}$
4	ENSG00000214706	IFRD2	$6.58\acute{}10^{-50}$
5	ENSG00000126062	TMEM115	$4.21\acute{}10^{-44}$
6	ENSG00000114383	TUSC2	$7.78\acute{}10^{-40}$
7	ENSG00000243477	NAA80	$1.33\acute{}10^{-39}$
8	ENSG00000173540	GMPPB	$1.93\acute{}10^{-39}$
9	ENSG00000068028	RASSF1	$5.37\acute{}10^{-38}$
10	ENSG00000183763	TRAIP	$3.28\acute{}10^{-36}$
11	ENSG00000068001	HYAL2	$4.08\acute{}10^{-36}$
12	ENSG00000114388	NPRL2	$3.64\acute{}10^{-35}$
13	ENSG00000114395	CYB561D2	$9.96\acute{}10^{-33}$
14	ENSG00000164077	MON1A	$1.09\acute{}10^{-28}$
15	ENSG00000004534	RBM6	$2.65\acute{}10^{-28}$
16	ENSG00000281358	RASSF1-AS1	$6.25\acute{}10^{-24}$
17	ENSG00000186792	HYAL3	$7.09\acute{}10^{-23}$
18	ENSG00000003756	RBM5	$8.26\acute{}10^{-17}$
19	ENSG00000004838	ZMYND10	$6.06\acute{}10^{-13}$
20	ENSG00000007402	CACNA2D2	$1.63\acute{}10^{-11}$
21	ENSG00000114378	HYAL1	$5.30\acute{}10^{-11}$
22	ENSG00000176095	IP6K1	$5.51\acute{}10^{-10}$
23	ENSG00000188338	SLC38A3	$2.00\acute{}10^{-8}$
24	ENSG00000271858	CYB561D2	$4.53\acute{}10^{-8}$
25	ENSG00000164068	RNF123	$1.80\acute{}10^{-6}$
26	ENSG00000232352	SEMA3B-AS1	$3.87\acute{}10^{-6}$

The list of 26 genes encoded in the 3p21.3 region that were significantly upregulated in the LTPA 3p21.3-HAC cells by RNA sequencing analysis. The red text indicates the seven candidate genes that were narrowed down.

Supplementary Table S2. Sequences of the primers used in this study.

Gene name	Primer sequence	
	Forward (5'-3')	Reverse (5'-3')
mTert	ATGTCACGGAGAGCACATTC	CTGCAGATGGGCATGGCTA
mGapdh	TCATTGTCATACCAGGAAATGAGC	GTCTCCTGCGACTTCAACAG
TMEM115	GGAAGACCAGTCCATCTGGC	GAGCTGCCTCGAAGGTGATT
CACNA2D2	TGGAGAATGACACTGTGGGC	CGCACTTCTGAGGCTGGTC
IP6K1	CAGGATGACACAACAGAACGGG	ACCTTCGGACTCTTTGCCTCCT
GNAI2	CATCTTCTGCGTAGCCTTGAGC	GATGGACGTGTCTGTGAACCAC
IFRD2	AAGTACCGTGCCAAGGCTGATC	GTAGAGCACCTCAAAGCCGAAG
SLC38A3	GCCACTTGTCATACAGACCTTCC	GGCAGAATGATGGTGACAGAGAC
ZMYND10	GGAGCGAGAAAACAGAGGCAAG	AGCACATCCAGCCTGTAGGTCT
hTERT	GCCTTCAAGAGCCACGTC	CCACGAACTGTCGCATGT
GAPDH	AGCCACATCGCTCAGACAC	GCCCAATACGACCAAATCC

Supplementary Table S3. Sequences of the small interfering RNA (siRNA) duplexes for each candidate gene.

Gene name (siRNA ID)	Sense (5'-3')	Antisense (5'-3')
TMEM115 (11070-1)	GUGGAAGACCAGUCCAUCUtt	AGAUGGACUGGUCUUCCACtt
TMEM115 (11070-2)	UCUUGCUACCUGGUUCCAAtt	UUGGAACCAGGUAGCAAGAtt
CACNA2D2 (9254-1)	CAUAUACACCUGGAACAUUtt	AAUGUUCCAGGUGUAUAUGtt
IP6K1 (9807-1)	CACAUGCACGUUCCCUAGAtt	UCUAGGGAACGUGCAUGUGtt
IP6K1 (9807-2)	GAGAACCUCAUCAGCAUCAtt	UGAUGCUGAUGAGGUUCUCtt
GNAI2 (2771-1)	CACCAAGAACGUGCAGUUCtt	GAACUGCACGUUCUUGGUGtt
GNAI2 (2771-2)	CCAAGAACGUGCAGUUCGUtt	ACGAACUGCACGUUCUUGGtt
IFRD2 (7866-1)	GACCUUGAGGAAAAGCUGAtt	UCAGCUUUUCCUCAAGGUCtt
IFRD2 (7866-2)	GUCGUCUGGCCAAGUCUCAtt	UGAGACUUGGCCAGACGACtt
SLC38A3 (10991-1)	CAGAACAUCGGAGCCAUGUtt	ACAUGGCUCCGAUGUUCUGtt
SLC38A3 (10991-2)	CGGGCAUUAUCCUUUUCCUtt	AGGAAAAGGAUAAUGCCCGtt
ZMYND10 (51364-2)	CUCUUCCUCUAGCACAGUAtt	UACUGUGCUAGAGGAAGAGtt

Supplementary Materials and Methods

RNA-seq analysis

Total RNA samples were quantified and qualified using the NanoDrop, Qubit RNA Assay (Thermo Fisher Scientific, Waltham, MA, USA), and TapeStation RNA ScreenTape (Agilent, Santa Clara, CA, USA). Then, 500 ng of the total RNA samples that met the quality guideline (RNA integrity number of 8 or higher) was treated with NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, Ipswich, MA, USA) to enrich poly-A mRNA and remove rRNA molecules. Next, cDNA synthesis followed by transcriptome library preparation were conducted using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs), where dUTP was incorporated in the 2nd strand cDNA synthesis process instead of dTTP to block PCR amplification against the 2nd strand templates. A 10-cycle PCR amplification was performed to increase the library yield and incorporate sample barcodes into the library fragments. The resulting transcriptome libraries were quantified using the Qubit DNA Assay (Thermo Fisher Scientific), with their fragment size distribution estimated using TapeStation D1000 ScreenTape (Agilent).

The individual libraries were then pooled/multiplexed together at an equimolar amount, followed by sequencing on a NovaSeq 6000 instrument (Illumina, San Diego, CA, USA) with v1.5 reagents and a 150 bp paired-end configuration. Demultiplexing and FASTQ generation were performed with a bcl2fastq pipeline. Approximately 20M paired-end reads with Q30 score > 90% were obtained for each sample.

In the data analysis process, Bcl2fastq (v2.20.0.422) was used to process the original image data for base calling and preliminary quality analysis. The filtered fastq format data were processed using Cutadapt (V1.9.1) with parameters set to ensure high-quality clean data. The clean data were then aligned to the human reference genome using the Hisat2 (version 2.2.1). For expression analysis,

the transcripts were first converted from known gff annotation files into fasta format, then properly indexed. HTSeq (v0.6.1) was used to estimate the gene and isoform expression levels from the pair-end clean data, with the gff annotation file serving as the reference gene file. Differential expression analysis was conducted using the DESeq2 Bioconductor package. GOSeq (v1.34.1) was employed to identify gene ontology (GO) terms with a list of enriched genes with a significant Padj less than or equal to 0.05. Additionally, topGO was used to plot the directed acyclic graph for the GO terms. For novel transcript prediction, we assembled the transcriptome from one or more samples, then used Cuffcompare, a tool from Cufflinks v2.2.1, to compare the assembly to known transcripts. Novel transcripts were predicted from the Cuffcompare results. Alternative splicing analysis was performed using rMATs v4.1.0. Single nucleotide variant analysis was conducted using Samtools v0.1.19, with the mpileup command and Bcftools v0.1.19 used for single nucleotide variant calling. The Bioconductor package DEXSeq (V 1.21.1) was used for differential exon usage analysis. Principal component analysis was conducted using R language with FPKM values as the input.

The transcriptomic analysis described above was conducted at GENEWIZ from Azenta Life Sciences (Tokyo, Japan). The RNA-seq data have been deposited to DNA Data Bank of Japan (DDBJ) (accession number: DRA018313).

IHC staining

IHC staining was performed with Histostainer-36A (Nichirei Biosciences, Tokyo, Japan). Sections (4 μm-thick) were deparaffinized and hydrated, then incubated in heat processor solution pH 9.0 (715291, Nichirei Biosciences) at 100°C for 15 min, blocked with 3% H_2O_2 (715242, Nichirei Biosciences) for 5 min, and incubated with a rabbit polyclonal anti-TMEM115 antibody (1:500; AssayGenie) for 60 min at room temperature. Subsequently, the slides were incubated with the secondary antibody MAX-PO (MULTI) (724152; Nichirei Biosciences) for 30 min at room temperature, visualized with

DAB (725191; Nichirei Biosciences), and counterstained with hematoxylin.