



(86) **Date de dépôt PCT/PCT Filing Date:** 2014/06/20
(87) **Date publication PCT/PCT Publication Date:** 2014/12/24
(85) **Entrée phase nationale/National Entry:** 2016/12/14
(86) **N° demande PCT/PCT Application No.:** US 2014/043457
(87) **N° publication PCT/PCT Publication No.:** 2014/205374
(30) **Priorité/Priority:** 2013/06/20 (US61/837,358)

(51) **Cl.Int./Int.Cl.** *G01N 33/50* (2006.01)
(71) **Demandeur/Applicant:**
THE TRUSTEES OF THE UNIVERSITY OF
PENNSYLVANIA, US
(72) **Inventeurs/Inventors:**
ZARET, KENNETH S., US;
KIM, JUNG SUN, US
(74) **Agent:** SMART & BIGGAR

(54) **Titre :** METHODES PERMETTANT DE DIAGNOSTIQUER UN CANCER DU PANCREAS
(54) **Title:** METHODS FOR DIAGNOSING PANCREATIC CANCER

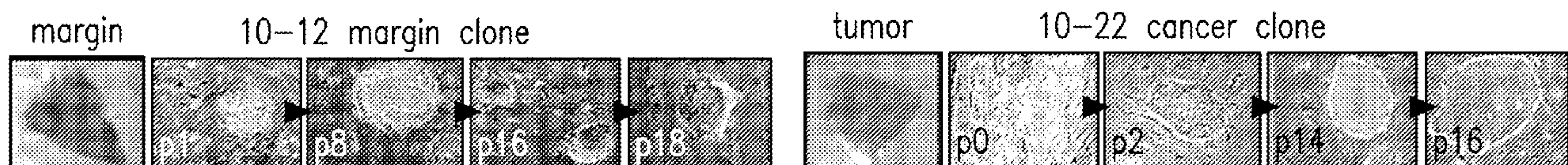


FIG. 1A

(57) **Abrégé/Abstract:**

The present disclosure relates to the use of one or more biomarkers to determine the presence of pancreatic cancer precursor lesions or pancreatic cancer in a subject. This disclosure is based, at least in part, on the discovery that early to invasive stages of pancreatic cancer release or secrete biomarkers that can be detected in biological samples of a subject. Accordingly, in certain non-limiting embodiments, the present disclosure provides for methods and kits for determining the presence of one or more biomarkers in a biological sample of a subject, and methods of using such determinations in selecting a therapeutic regimen for a cancer subject and in methods of treating cancer subjects.

(51) International Patent Classification:
G01N 33/50 (2006.01)

(21) International Application Number:
PCT/US2014/043457

(22) International Filing Date:
20 June 2014 (20.06.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/837,358 20 June 2013 (20.06.2013) US

(71) Applicant: **THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA** [US/US]; 3160 Chestnut Street, Suite 200, Philadelphia, PA 19104 (US).

(72) Inventors: **ZARET, Kenneth, S.**; 332 Waring Road, Elkins Park, PA 19027 (US). **KIM, Jungsun**; 3900 Chestnut Street, Chestnut Hall Apartments #612, Philadelphia, PA 19014 (US).

(74) Agents: **RAGUSA, Paul, A.** et al.; Baker Botts LLP, 30 Rockefeller Plaza, New York, NY 10112-4498 (US).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) Title: METHODS FOR DIAGNOSING PANCREATIC CANCER

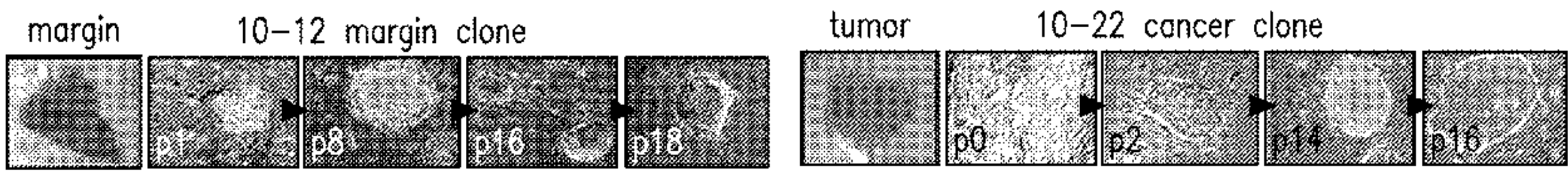


FIG. 1A

(57) **Abstract:** The present disclosure relates to the use of one or more biomarkers to determine the presence of pancreatic cancer precursor lesions or pancreatic cancer in a subject. This disclosure is based, at least in part, on the discovery that early to invasive stages of pancreatic cancer release or secrete biomarkers that can be detected in biological samples of a subject. Accordingly, in certain non-limiting embodiments, the present disclosure provides for methods and kits for determining the presence of one or more biomarkers in a biological sample of a subject, and methods of using such determinations in selecting a therapeutic regimen for a cancer subject and in methods of treating cancer subjects.

WO 2014/205374 A1

METHODS FOR DIAGNOSING PANCREATIC CANCER

PRIORITY CLAIM

This application claims priority to United States Provisional
5 Application No. 61/837,358, filed June 20, 2013, the contents of which are
incorporated by reference herein in its entirety.

GRANT INFORMATION

This invention was made with government support under Grant
10 Number R37GM36477 awarded by the National Institutes of Health. The government
has certain rights in the invention.

BACKGROUND

Pancreatic ductal adenocarcinoma (PDAC) carries a dismal prognosis,
15 with less than a 5% survival rate (Hezel et al., 2006; Maitra and Hruban, 2008; Rustgi,
2006). Diagnosis can be difficult because there are no noticeable symptoms in early
stages, and diagnosis is often determined when cancer has already disseminated to
other organs. In addition, there is a scarcity of biomarkers for early stage detection of
the disease, leading to PDAC usually detected at advanced stages, with limited
20 therapeutic options. In combination with late detection, pancreatic cancer displays a
poor response to chemotherapy, radiation therapy and surgery as conventionally used.

Many proteins secreted from pancreatic cancers (Harsha et al., 2009)
that may serve as biomarkers have been identified in advanced, invasive PDAC or cell
lines thereof, and thus may not represent markers for early stages of the disease. For
25 example, a number of protein biomarkers have been identified in the sera and
pancreatic juices from pancreatitis and pancreatic cancer patients. Serum biomarker
protein CA-19.9 is presently used to monitor pancreatic cancer but is not useful in
early diagnosis (Gattani et al., 1996). Therefore, markers have been sought for
precancerous lesions, such as PanINs and intraductal papillary mucinous neoplasms
30 (IPMNs) (Brat et al., 1998; Hruban et al., 2001), but the markers have typically been
intracellular or cell surface proteins (Harsha et al., 2009) rather than secreted or
released proteins that may provide an improved opportunity for diagnosis, especially
in its earlier and potentially curable stages.

The prominent animal model of PDAC is based upon inducing a G12D mutant allele of *Kras* in the mouse pancreatic epithelium (Hingorani et al., 2003), a mutation that frequently occurs in human PDAC. The mice develop pancreatic intra-epithelial neoplasias (PanINs) with prolonged latency and incomplete penetrance of PDAC. PDAC and related tumors develop much more rapidly when *Kras*G12D/+ mice also contain mutations of *Ink4a/Arf*, *Tgfbr2*, *p53*, or *PTEN* (Morris et al., 2010), although these mutations alone do not efficiently cause PDAC. In an effort to develop human models of early pancreatic cancer, PDAC cells have been grafted into immunodeficient mice either as tumor fragments (Rubio-Viqueira et al., 2006), dispersed cells (Kim et al., 2009) or cells sorted for cancer stem cell markers (Hermann et al., 2007; Ishizawa et al., 2010; Li et al., 2007). In these contexts, tumors rapidly arise that resemble the advanced PDAC stages from which the cells were derived and do not undergo the slow growing, early PanIN stages of PDAC (Ding et al., 2010). However, there is a need for a live human cellular model that undergoes the early stages of PDAC and disease progression.

SUMMARY

The present disclosure is based, at least in part, on the discovery that pancreatic cancer in its early to invasive stages releases or secretes biomarkers that can be detected in a biological sample of a subject. Accordingly, the disclosure provides methods for determining the presence of a biomarker indicative of pancreatic cancer in a biological sample of a subject. In certain embodiments, the detection of one or more biomarkers in a biological sample of a patient indicates the presence of early to invasive stages of pancreatic cancer. Thus, the methods of the disclosure also provide for early prognosis and diagnosis of cancer (e.g., identification of a biomarker prior to identification of a tumor by conventional means) and therapy monitoring in a subject.

In one aspect, the present disclosure provides methods of assessing whether a subject has pre-cancerous lesions and is at risk for developing advanced stage pancreatic cancer, including determining the presence of a biomarker in a biological sample obtained from the subject, wherein the presence of the biomarker is an indication that the patient is at increased risk for developing advanced stage cancer. In certain embodiments, the presence of a biomarker indicates early stage pancreatic cancer in the subject. In certain embodiments, the presence of a biomarker indicates

advanced stage, invasive and/or metastatic pancreatic cancer in the subject. In certain embodiments, the method is carried out prior to the identification of a primary tumor in the subject.

In yet another aspect, the present disclosure provides methods of assessing the efficacy of a therapeutic or prophylactic therapy for preventing, inhibiting or treating pancreatic cancer, in a subject, including determining the presence and/or level of a biomarker in a biological sample obtained from the subject prior to therapy; and determining the presence and/or level of a biomarker in a biological sample obtained from the subject at one of more time points during therapeutic or prophylactic therapy, wherein the therapy is efficacious for preventing, inhibiting or treating cancer in the subject when there is a lower level of the biomarker in the second or subsequent samples, relative to the first sample.

In certain non-limiting embodiments, the biological sample can be a blood sample and/or a plasma sample. In certain embodiments, the biological sample can be a stool sample. In certain embodiments, the biological sample can be fluid drained from a pancreatic cyst. In other embodiments, the biological sample can be a tissue, *e.g.*, a tissue biopsy. In certain embodiments, one or more biomarkers can be detected in one or more biological samples from a subject. The use of stool, plasma and/or blood as a biological sample makes it possible to eliminate invasiveness of the diagnostic or prognostic procedure, and dramatically improve the burden of the examination on the subject.

In certain embodiments, the biomarker is a protein and the presence of the protein is detected using a reagent which specifically binds with the protein. For example, the reagent can be selected from the group consisting of an antibody, an antibody derivative, an antigen-binding antibody fragment and a non-antibody peptide which specifically binds the protein. In certain embodiments, the antibody or antigen-binding antibody fragment is a monoclonal antibody or antigen-binding fragment thereof, or a polyclonal antibody or antigen-binding fragment thereof. In certain embodiments, the protein biomarker can be detected by biophysical techniques such as mass spectrometry. In certain embodiments, the protein biomarker can be detected by enzyme-linked immunosorbent assay (ELISA).

The biomarker can also be a transcribed polynucleotide or portion thereof, *e.g.*, a mRNA or a cDNA. In certain embodiments, detecting a transcribed polynucleotide includes amplifying the transcribed polynucleotide. In certain

embodiments, the nucleic acid biomarker can be detected by RNA in situ hybridization.

The disclosure also provides kits for monitoring, diagnosing or assessing whether a subject has pancreatic cancer, for monitoring the therapeutic treatment of a subject and for assessing the efficacy of a therapeutic treatment regime of a subject, where the kit contains reagents useful for detecting secreted or released biomarkers in a biological sample.

BRIEF DESCRIPTION OF THE DRAWINGS

10 **Figure 1A-I.** Establishing iPS-like lines from patient-matched margin and pancreatic ductal adenocarcinoma. A. Cells from pancreatic cancer and margin tissues were reprogrammed and different passages of the 10-12 margin and 10-22 cancer iPS-like clones from patient #10 are shown. B, C. Expression of pluripotency markers in 10-12 margin and 10-22 cancer iPS-like lines by immunostaining (B) and
15 RT-PCR (C). D. Three month teratomas in NSG mice from 10-12 margin and 10-22 cancer iPS-like lines showed differentiation into tissues of all three germ layer lineages (also see Figure 8A). E. Expression of differentiation markers for endoderm (*CXCR4*), mesoderm (*RUNX1*), and ectoderm (*SOX1*, *PAX6*) relative to *GAPDH* in embryoid bodies from H1 huES cells and 10-12 margin and 10-22 cancer iPS-like
20 lines cultured for 12-14 days. Error bars, mean $\pm$ SD. F, G. DNA sequence tracks revealing *KRAS* mutation and karyotype showing subtetraploidy in the 10-22 cancer iPS-like line (F) and wild type *KRAS* and normal karyotype in the 10-12 margin iPS-like line (G). H. PCR analysis of *CDKN2A* (*p16Ink4a*) exon 2 reveals a heterozygous deletion in 10-22 cells. I. Comparative genomic hybridization showing normal
25 profiles for the 10th primary margin cells (a) and the 10-12 margin iPS-like line (b). Gross chromosomal rearrangements are evident in the 10th primary cancer cell culture (c), which was mixed with normal stromal cells; the rearrangements are evident more clearly in the 10-22 cancer iPS-like line (d) (also see Figure 9B), demonstrating that the 10-22 line is clonally derived representative of the initial tumor
30 epithelial cell population. The locations of *PTEN* and *DPC4* (*SMAD4*) loci are shown in regions of chromosome loss in the 10th cancer and 10-22 cells.

Figure 2A-D. The 10-22 iPS-like cells preferentially generate ductal structures in teratomas. A. Summaries of teratoma tissue types arising in 3 months in immunodeficient mice based on the number of histologic structures seen. Teratomas

from huES cells are mostly ectodermal in appearance, whereas teratomas from 10-12 and 10-22 iPS-like lines are mostly endodermal/ductal. B. H&E staining of the margin and tumor tissue of patient #10 and of teratomas formed by 10-12 margin and 10-22 cancer iPS-like lines after 3 months. The latter lines create tubular, duct-like structures independent of passage number. Notably, the 10-22 cancer teratomas show a much a higher degree of differentiation than the primary tumor. C. The primary tumor shows an invasive phenotype (dotted line) and high nuclear to cytoplasmic ratio (arrow). D. 10-22 tumor iPS teratomas at 3 months show a low nuclear to cytoplasmic ratio, intracellular mucins (arrows), and a more differentiated phenotype.

Figure 3A-K'. Teratomas at three months from 10-22 cancer iPS-like cells exhibit PanIN-like structures and marker expression. A-C. No K19 staining in mouse subcutaneous tissue (negative control), weak K19 staining of 10-12 margin iPS-like teratomas at three months, and strong K19 staining of 10-22 cancer iPS-like teratomas at 3 months. D-I. Nuclear staining of PDX1 (arrow) and cytoplasmic staining of MUC5AC (arrow) only in teratomas of 10-22 cancer iPS-like teratomas at 3 months. J, K. Higher magnification shows uniform K19 staining (J, J', arrow) and heterogeneous MUC5AC staining (K, K') of teratomas from 10-22 cancer iPS-like cells.

Figure 4A-R. Teratomas at 6 and 9 months from 10-22 cancer iPS-like cells exhibit invasive stages of pancreatic cancer. A-B, H&E staining of 10-22 teratomas in NSG mice at 6 months (A, A') and 9 months (B, B'). Arrows indicate epithelial cells with a more dysmorphic phenotype than at 3 months; nuclei are heterotypic and epithelia with hypochromic nuclei are invading the stroma (A, B, arrows). C-R immunohistochemistry showing epithelial cells positive for K19, MUC5AC, PDX1, and SOX9 in tumors arising at 6 and 9 months in NSG mice injected with 10 22 cells. See Figure 11F for genetic confirmation of 10-22 cells in the tumors.

Figure 5A-F. Proteins secreted and HNF4 regulatory network within teratoma explants of 10-22 iPS-like cells at 3 months. Scheme for in vitro explants of three month teratomas from 10-22 cancer iPS-like lines. A-C. Whole mount view of teratoma explant from 10-22 iPS like cells (A), along with the explant sectioned and stained for K19 (B), and MUC5AC (C); arrows indicate positive domains. D. Proteomic analysis of conditioned media of 3 independent teratoma explants (#7761, 9223, and 9225) from 10-22 cancer iPS-like cells. Conditioned media was analyzed

by nanoLC/MS/MS and protein candidates were identified by a Sequest search in the IPI DB. Using UniProKB TrEMBL (downloaded on Oct, 2011), peptides were included in the list that either consists of uniquely human sequences or sequences that are common to humans and mice. Mouse-specific sequences were excluded. Protein I.D.'s of conditioned media from teratoma explants were compared to those in conditioned media from explants of contralateral control tissue and of media from undifferentiated 10-22 cells (left Venn diagram). The total 698 secreted or released proteins specific to the teratomas are indicated in the three different experiments (right Venn diagram). Proteins secreted from at least two teratoma explants are summarized in Tables 4 and 5 and were used for pathway analysis. E. Numerous proteins fall into linked pathways for TGF β and integrin signaling. Lines and arrows connecting molecules indicate direct interactions and dashed lines and arrows indicate indirect interactions. F. All proteins in bold are secreted or released specifically from the 10-22 teratoma explants and fall within a network controlled by the transcription factor HNF4 α (Table 9). Asterisks denote proteins whose genes are directly bound by HNF4 α .

Figure 6A-P'. Activation of HNF4 α in PanIN cells and well differentiated early pancreatic cancers in human clinical samples and a mouse model of human PDAC. A-D. Immunohistochemistry for HNF4 α showing the absence of staining in the main human pancreatic mass and ducts (A), strong nuclear staining in PanIN-like structures in a 3 month teratoma (B, arrow) and in invasive cells in a 9 month tumor (C, arrow) generated from NSG mice injected with 10-22 cells, and mostly cytoplasmic staining in small, moderately differentiated ducts (D, image section below the diagonal, brown, dashed arrow) and an absence of staining in undifferentiated ducts (D, image section above the diagonal) of the 10th patient's pancreatic tumor epithelium E-I. HNF4 α is weakly labeled in human PanIN-1 cells (E, arrow, N=8), strongly in PanIN2 (F, arrow, N=7), PanIN3 (G, arrow, N=3), and mucinous well differentiated human PDAC (H, arrow, N=8) but decreased or not in undifferentiated/or poorly differentiated human PDAC (I, N=8). J, Percentage of HNF4 α positive nuclei to total epithelial nuclei counted (See Table 14) in multiple PanIN and PDAC samples on human tissue microarrays. Error bars, SEM. The differences in HNF4 α expression is statistically significant between margin and PanIN-2 and -3 stages (* P<0.05); margin and well-differentiated PDAC (** P=5.46E-06); well differentiated PDAC and poorly/undifferentiated PDAC

(**P=3.05E-06), as per two-tailed Student's t-tests. K-P. In a mouse model of PDAC, HNF4 α is expressed weakly at the PanIN1 (K, K'), strongly at the PanIN2-3 stages (L, L', M, M') and in differentiated portions of tumors (PDAC) (O, O'), and weakly or not expressed in undifferentiated portions of the same tumor (P, P').

5 **Figure 7A-H.** A-B. Generating iPS-like lines from patient #14. Pancreatic margin and cancer iPS-like lines were generated from patient #14 and characterized by immunostaining for NANOG, OCT4, SSEA4 (A) and qRT-PCR for pluripotency genes (B). C-D. Pancreatic margin and cancer iPS-like lines were generated from patient #19, characterized by immunostaining for OCT4, NANOG,
10 and TRA-1-60 (C), RT-PCR for pluripotency (D). E. Embryoid bodies were generated from 14-24 margin iPS-like line for 15 days and checked for the expression of differentiation markers using qRT-PCR. Expression levels are relative to *Gapdh*. Error bars are the mean $\pm$ SD (E). F. Teratoma assays showed the differentiation of 14-24 margin and 14-27 cancer derived iPS-like lines into multilineage cells; but note
15 that in limited experiments with the 14-27 line, neural derivatives were not detected. G-H, Karyotype analysis showed that 14-24 margin and 14-27 cancer iPS-like lines had subdiploidy (G) and 19th lines showed diploidy (H).

Figure 8A-C. A. Immunohistochemistry (IHC) analysis of different germ layer tissues on 10-22 cancer iPS teratoma tissue after 3 months. H1 huES
20 teratoma tissue was used as a positive control. Glial Fibrillary Acidic protein (GFAP) and beta III tubulin were used for ectoderm, vimentin and MF20 were for mesoderm, and K19 was used for endoderm. B. Immunohistochemical staining for NANOG (upper panels) and OCT4 (lower panels), with the H1 huES line as a positive control and the primary tumor #10 experimental sample, analyzed at the same time. The data
25 indicate that NANOG is not expressed in the primary #10 tumor and that OCT4 is expressed weakly in sporadic cells. The absence of NANOG expression is consistent with the hypothesis that the primary tumor #10 was not composed largely of pancreatic cancer stem cells (Lonardo et al. 2011). C. Analysis of the CpG methylation state of 10-12 margin and 10-22 cancer iPS-like lines at the designated
30 CpG sites (ovals, below position with respect to the transcription start) in the 5' upstream of the human NANOG and OCT4 genes by bisulfite pyrosequencing. Genomic DNA from H1 huES cells was used for positive control and genomic DNA from parental primary tumor of patient #10 was used for a negative control. The schematic diagram of upstream region of NANOG and OCT4 examined in this study.

The regions amplified by PCR are enlarged. The graphs show the percent methylated C's in the designated CpG positions. At NANOG, all 9 CpG sites tested were methylated in the parental #10 tumor DNA and demethylated in the 10-22 iPS-like line, comparable to that seen in the H1 human ES control cells. At OCT4, C's at positions -497 and -532 exhibited marked demethylation in the 10-22 iPS-like line, while at position -161 there was modest demethylation. The H1 control was not markedly demethylated at the -497 and -532 sites, but was at other sites tested. Methylation of the NANOG locus provides further evidence that the primary tumor #10 was not composed of pancreatic cancer stem cells.

Figure 9A-B. A. Karyotype analysis of 10-12 margin and 10-22 cancer iPS-like clones. B. Comparative genomic hybridization assay of 10-12 margin, 10-22 cancer iPS-like clones. Of the 23 chromosomal aberrations seen in the epithelial cell cultures of the 10th primary cancer, 20 aberrations were present in the 10-22 cancer iPS-like line, as seen by the green and red double bars. Dashed and solid bars denote aberrations unique to the primary cancer or cell line, respectively. See Fig. 1I for ch10 and ch18. Note that the chromosomal aberrations seen in the 10th primary cancer cultures (solid CGH traces), which were composed mainly of the cancer epithelial cells but also contained stromal cells, are amplified in the 10-22 iPS line (dashed CGH traces), a clone from the cancer culture.

Figure 10A-I. A. 10-12 margin and 10-22 cancer teratoma ductal structures at three months expressed DBA lectin. Contralateral control subcutaneous fat tissue was used as a negative control and mouse pancreas was used as a positive control. B-G. The PanIN-like structures derived from teratomas from the 10-22 cancer iPS-like line at three months in independent NSG mice. Compared to the poorly differentiated structures in the parental primary tumor of patient #10 (B), the 10-22 cancer iPS-like lines generated PanIN stages in independent mice regardless passage number of initial 10-22 cancer iPS-like line (C-G). H. PCR for rt-TA in LCM-dissected 10-22 cancer iPS teratoma tissue at three months. PCR for rt-TA on laser capture microdissected PanIN-like duct, stroma, and mouse duct. All epithelial cells were removed from the region prior to removing stromal cells, to avoid possible cross-contamination. I. Immunostaining of rt-TA on PanIN-like epithelial and stroma cells derived from 10-22 cancer iPS teratoma tissue at three months. The fibroblast distal region was negative for rt-TA. The data show that at least part of the stroma

surrounding the PanIN-like epithelium in the teratoma is derived from the 10-22 human iPS-like cell line.

Figure 11A-F. A. Validation of antibodies against K19 and MUC5AC. Species-specificity of anti-human K19 was tested on mouse subcutaneous tissue, mouse and human pancreatic tissue at 1:2000 dilution ratios. The specificity of the MUC5AC antibody was tested in normal human pancreas and normal human stomach. B-E Immunostaining of SOX9 in PanIN-like ducts of 10-22 cancer iPS teratoma arising at 3 months in three different NSG mice. As a positive control, mouse pancreas was stained with SOX9. Subsets of duct (arrows) and centroacinar cells (arrowheads) in mouse pancreas moderately expressed SOX9 (B-B', brown). Islet cells didn't express SOX9 (B-B', dashed arrows). PanIN-like ducts expressed SOX9 (brown, arrows, C-E). F. Human origin of 9 month 10-22 teratoma arose from NSG mice injected with 10-22 cancer iPS-like lines was confirmed by PCR for rt-TA and *CDKN2A*. Four tumors derived from two different mice preserved rt-TA sequence and deletions of *CDKN2A*. DNA from contralateral control (CLC) tissue was used for negative control.

Figure 12A-D. A. Scheme for in vitro explants of three month teratomas from 10-22 cancer iPS-like lines. Three independent teratoma tissues were explanted for in vitro culture (#7761, 9223, and 9225). Contralateral control tissue and 10-22 teratoma tissue generated spheres. B. Human origin of 10-22 teratoma in vitro explants was confirmed by RT-PCR for rt-TA and *CDKN2A*, and sequencing for the *KRAS* G12D mutation in the 10-22 iPS-like cells (see Fig. 1F-H). C. Control tests of HNF4 α immunostaining on mouse liver tissue and normal pancreas. HNF4 α was expressed only in islets and not in acinar or ductal cells in the mouse normal pancreas. Solid arrow indicates the positive cells and dashed arrow indicates the negative cells. D. HNF4 α staining on additional 10-22 cancer iPS teratoma tissue derived from independent mice at 3 months. Solid arrow indicates the positive cells.

Figure 13. Statistical analysis of the biomarker TCHP in control and pancreatic cancer samples.

Figure 14. Statistical analysis of the biomarker ABCA13 in control and pancreatic cancer samples.

Figure 15. Statistical analysis of the biomarker STARD8 in control and pancreatic cancer samples.

Figure 16. Statistical analysis of the biomarker ATP2A1 in control and pancreatic cancer samples.

Figure 17. Statistical analysis of the biomarker FKBP10 in control and pancreatic cancer samples.

5 **Figure 18.** Statistical analysis of the biomarker SCN8A in control and pancreatic cancer samples.

Figure 19. Statistical analysis of the biomarker TCF20 in control and pancreatic cancer samples.

10 **Figure 20.** Statistical analysis of the biomarker SYNE1 in control and pancreatic cancer samples.

Figure 21. Statistical analysis of the biomarker UFD1L in control and pancreatic cancer samples.

Figure 22. Statistical analysis of the biomarker FLRT3 in control and pancreatic cancer samples.

15 **Figure 23.** Statistical analysis of the biomarker TOP2B in control and pancreatic cancer samples.

Figure 24. Statistical analysis of the biomarker ZHX2 in control and pancreatic cancer samples.

20 **Figure 25.** Statistical analysis of the biomarker LIMCH1 in control and pancreatic cancer samples.

Figure 26. Statistical analysis of the biomarker THBS2 in control and pancreatic cancer samples.

Figure 27. Statistical analysis of the biomarker SHROOM3 in control and pancreatic cancer samples.

25 **Figure 28.** Statistical analysis of the biomarker HMOX1 in control and pancreatic cancer samples.

Figure 29. Statistical analysis of the biomarker LOXL3 in control and pancreatic cancer samples.

30 **Figure 30.** Statistical analysis of the biomarker OBSCURIN in control and pancreatic cancer samples.

Figure 31. Statistical analysis of the biomarker MALECTIN in control and pancreatic cancer samples.

Figure 32. Statistical analysis of the biomarker DNAH5 in control and pancreatic cancer samples.

Figure 33. Statistical analysis of the biomarker CA19-9 in control and pancreatic cancer samples.

Figure 34. Statistical analysis of the biomarker AFP in control and pancreatic cancer samples.

5 **Figure 35.** Statistical analysis of the biomarker RTTN in control and pancreatic cancer samples.

Figure 36. Statistical analysis of the biomarker NLRX1 in control and pancreatic cancer samples.

10 **Figure 37.** Statistical analysis of the biomarker DNAH12 in control and pancreatic cancer samples.

Figure 38. Statistical analysis of the biomarker ODZ3 in control and pancreatic cancer samples.

Figure 39. Statistical analysis of the biomarker ADAMST9 in control and pancreatic cancer samples.

15 **Figure 40.** Statistical analysis of the biomarker TPM1 in control and pancreatic cancer samples.

Figure 41. Statistical analysis of the biomarker DNAH1 in control and pancreatic cancer samples.

20 **Figure 42.** Statistical analysis of the biomarker PMBP1 in control and pancreatic cancer samples.

Figure 43. Statistical analysis of the biomarker DNAH17 in control and pancreatic cancer samples.

Figure 44. Statistical analysis of the biomarker EPHB1 in control and pancreatic cancer samples.

25 **Figure 45.** Statistical analysis of the biomarker DOS in control and pancreatic cancer samples.

Figure 46. Statistical analysis of the biomarker MMP2 in control and pancreatic cancer samples.

30 **DETAILED DESCRIPTION**

The present disclosure relates to the use of one or more biomarkers identified herein to detect the presence of pancreatic cancer in a biological sample from a subject. It is based, at least in part, on the discovery that iPS cells created from a human pancreatic ductal adenocarcinoma (PDAC) sample provided a live cell

human model for studying early stages of pancreatic cancer. Furthermore, this disclosure is based, at least in part, on the discovery that early to invasive stages of pancreatic cancer released or secreted specific proteins that are detectable in biological samples, *e.g.*, blood, of a subject.

5 Accordingly, the disclosure provides for methods and kits for determining the presence of one or more biomarkers for pancreatic cancer in a biological sample of a subject, and methods of using the presence or level of such biomarkers to predict or diagnose pancreatic cancer in a subject, select a therapeutic regimen for a subject suffering from pancreatic cancer, and treat a subject suffering
10 from pancreatic cancer, wherein the presence of one or more biomarkers in a biological sample (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 25, 30 or more), or another defined minimum number depending on the subject, indicates the presence of pancreatic cancer in the subject. The biomarkers that can be used in the methods of the present disclosure are set forth in Tables 4, 5, 6, 7, 8, 9, 10,
15 11, 12, 13 and 15. In certain embodiments, the biomarkers that can be used in the methods of the present disclosure include RTTN, DNAH12, TPM1, DNAH1, STARD8, ATP2A1, TOP2B, LIMCH1, SYNE1, THBS2 and LOXL3.

 Based on the identification of a secreted or released biomarker of the present disclosure in a biological sample of a subject, a diagnosis of pancreatic cancer
20 in the subject can be made, even prior to the development, or identification of, tumor formation, thus allowing for prophylactic therapy in the subject. For example, further or more frequent monitoring, biopsy, surgical resection or other prophylactic measures to prevent tumor formation or identify cancer at a very early stage can be carried out based on the detection of one or more biomarkers in a biological sample.

25 Furthermore, the effectiveness of cancer therapy can be monitored by evaluating the presence and/or levels of the one or more biomarkers over the course of therapy, and decisions can be made regarding the type, duration and course of therapy based on these evaluations.

 In certain embodiments, the subject being tested for the presence of a
30 biomarker in a biological sample, as described herein, can be a subject who is at high risk for developing pancreatic cancer. In certain embodiments, a subject is at high risk for development of pancreatic cancer based on, for example, family history or determination of genetic predisposition. For example, these findings have implications for the management of individuals at high risk for pancreatic cancer,

including subjects with kindreds with inherited pancreatic cancer. Based on the identification that the biomarkers are secreted or released from early stage pancreatic cancer, a window of opportunity exists for prophylactic therapy in high-risk subjects in the time-period prior to detection of late-stage, invasive pancreatic cancer.

5 As exemplified herein, the present disclosure can be used to diagnose pancreatic ductal adenocarcinoma (PDAC), as the vast majority of patients with pancreatic cancer have metastatic disease at the time of diagnosis using current methods. More than 75% of patients who undergo surgical resection of small pancreatic tumors with clear surgical margins and no evidence of metastasis die from
10 metastatic disease within 5 years (Neoptolemos et al., 2004), a finding that is consistent with early spread. Moreover, metastatic PDAC has been documented in a cohort of patients who underwent pancreatectomy for chronic pancreatitis and in whom histologic analysis of the resected pancreas revealed only PanIN lesions (Sakorafas and Sarr, 2003). Accordingly, diagnosis and treatment at a very early
15 stage is important.

Definitions

As used herein, the use of the word “a” or “an” when used in conjunction with the term “comprising,” “comprise,” “includes” or “including” in the claims and/or the specification can mean “one,” but it is also consistent with the
20 meaning of “one or more,” “at least one,” and “one or more than one.” Certain embodiments of the present disclosure can consist of or consist essentially of one or more elements, method steps and/or methods of the invention. It is contemplated that any method or composition described herein can be implemented with respect to any other method or composition described herein.

25 As used herein, the term “biomarker” refers to a marker (*e.g.*, an expressed gene, including mRNA and/or protein) that allows detection of a disease in an individual, including detection of disease in its early stages. Biomarkers, as used herein, include nucleic acid and/or protein markers, set forth in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 15 or combinations thereof. In certain non-limiting embodiments,
30 a biomarker is a released and/or secreted protein that can be detected in a biological sample of a subject. In certain embodiments, the expression level of a biomarker as determined by mRNA and/or protein levels in tissue or biological sample from an individual to be tested is compared with respective levels in normal tissue or biological sample from the same individual or another healthy individual. In certain

embodiments, the presence of a biomarker as determined by mRNA and/or protein levels in a tissue or biological sample from an individual to be tested is compared with the respective presence or absence in normal tissue or biological sample from the same individual or another healthy individual. In certain embodiments, the presence of a biomarker as determined by mRNA and/or protein levels in a tissue or biological sample from an individual indicates that the individual has pancreatic cancer or is at an increased risk for developing late-stage pancreatic cancer.

As used herein, the term “biological sample” refers to a sample of biological material obtained from a subject, *e.g.*, a human subject, including tissue, a tissue sample, a cell sample, a tumor sample, a stool sample and a biological fluid, *e.g.*, plasma, serum, blood, urine, lymphatic fluid, ascites, pancreatic cyst fluid and a nipple aspirate. In certain embodiments, the presence of one or more biomarkers is determined in a peripheral blood sample obtained from a subject. In certain embodiments, the presence of one or more biomarkers is detected in a stool sample obtained from a subject. In certain embodiments, the presence of one or more biomarkers is detected in pancreatic cyst fluid obtained from a subject. In certain embodiments, the presence of one or more biomarkers is detected in one or more plasma samples obtained from a subject.

The term “patient” or “subject,” as used interchangeably herein, refers to any warm-blooded animal, *e.g.*, a human. Non-limiting examples of non-human subjects include non-human primates, dogs, cats, mice, rats, guinea pigs, rabbits, fowl, pigs, horses, cows, goats, sheep, etc.

The term “pancreatic cancer” as described herein refers to any type of cancerous or precancerous tissues arising from normal tissues of the pancreas, including, but not limited to, PanIN lesions, pancreatic ductal adenocarcinoma or pancreatic adenocarcinoma. Other types of pancreatic tumors include acinar-cell carcinoma, serous cystadenoma and pancreatic endocrine tumors. In certain embodiments, the biomarkers of the present disclosure can be used to detect cancers such as biliary cancer and liver cancer.

As used herein “resectable cancer” refers to a subset of cancers that are at an early stage and can be surgically excised. For example and not by way of limitation, stages IA, IB and IIA of pancreatic cancer are typically resectable.

The term “early stage cancer” as used herein refers to cancer prior to metastasis and/or organ extravasion. For example and not by way of limitation, with respect to pancreatic cancer, early stage cancer can include stages IA, IB and IIA.

Prognostic and Diagnostic Methods

5 Embodiments of the present disclosure relate to methods for diagnosing pancreatic cancer in a subject. In certain embodiments, a method for diagnosing prostate cancer in a subject is disclosed, where the method includes: obtaining a biological sample from the subject; determining the presence of one or more biomarkers in the biological sample; and diagnosing pancreatic cancer in the
10 subject, wherein the presence of the one or more biomarkers correlates to a positive diagnosis of pancreatic cancer in the subject. The biomarkers that can be used in the methods of the present disclosure are set forth in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 15.

 In certain embodiments, a method for diagnosing pancreatic cancer in
15 the subject includes obtaining at least one biological sample from the subject. In certain embodiments, the one or more biomarkers can be detected in blood (including plasma or serum) or in feces (*e.g.*, a stool sample), or alternatively at least one biomarker can detected in one sample, *e.g.*, the blood, plasma or serum, and at least one other biomarker is detected in another sample, *e.g.*, in feces. In certain
20 embodiments, the one or more biomarkers are detected in tissue samples. For example, the biological sample can be a tumor biopsy. In certain embodiments, the one or more biomarkers are detected in pancreatic cyst fluid. The step of collecting a biological sample can be carried out either directly or indirectly by any suitable technique. For example, a blood sample from a subject can be carried out by
25 phlebotomy or any other suitable technique, with the blood sample processed further to provide a serum sample or other suitable blood fraction, *e.g.*, plasma, for use in the methods of the presently disclosed subject matter.

 In certain embodiments, the methods for detection of one or more biomarkers can be used to monitor the response in a subject to prophylactic or
30 therapeutic treatment (for example, preventative cancer treatment or treatment of diagnosed cancer). In certain embodiments, the present disclosure further provides a method of treatment including measuring the presence of one or more biomarkers in a subject at a first timepoint, administering a therapeutic agent, re-measuring the one or more biomarkers at a second timepoint, comparing the results of the first and second

measurements and optionally modifying the treatment regimen based on the comparison. In certain embodiments, the one or more biomarkers are selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, 5 ZNF160, RTTN, ABCA13, DES, IMMT, TPM1, SNRPE, VCAM1, GRB2, SHROOM3, HMOX1, POSTN, MMP10, MMP-2, THBS2, EWSR1, NOD1, ADAMTS9, AFP, SYNE1, SYNE2, EPHB1, UFD1L, TEAD1, RYR3, CMYA5, MYLK, TOP2B, KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29, OBSCN, LOXL3, MLEC, DNAH1, DNAH5, DNAH12, DNAH17, SCYL2, FKBP10, 10 FLRT3, ZHX2 (AFR1), ZNF804A, ACTN2 or a combination thereof. In certain embodiments, the one or more biomarkers are selected from RTTN, DNAH12, TPM1, DNAH1, STARD8, ATP2A1, TOP2B, LIMCH1, SYNE1, THBS2, LOXL3 or a combination thereof.

In certain embodiments, the first timepoint is prior to an administration 15 of the therapeutic agent, and the second timepoint is after said administration of the therapeutic agent. In certain embodiments, the first timepoint is prior to the administration of the therapeutic agent to the subject for the first time. In certain embodiments, the dose (defined as the quantity of therapeutic agent administered at any one administration) is increased or decreased in response to the comparison. In 20 certain embodiments, the dosing interval (defined as the time between successive administrations) is increased or decreased in response to the comparison, including total discontinuation of treatment.

In certain embodiments of the present disclosure, the method of diagnosing, prognosing or screening for pancreatic cancer in a subject includes, (a) 25 obtaining a biological sample from the subject; (b) determining the level of one or more biomarkers in a biological sample of the subject selected from Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15 or a combination thereof; and (c) comparing the level of the one or more biomarkers to a reference sample, wherein an increase in the level of the one or more biomarkers indicates the presence of pancreatic cancer in the subject. In 30 certain embodiments, the reference sample can be obtained, for example, from a normal biological sample of the subject, *e.g.*, adjacent benign tissue, or from subjects that do not have pancreatic cancer.

In certain embodiments of the present disclosure, the method of diagnosing, prognosing or screening for pancreatic cancer in a subject includes (a)

obtaining a biological sample from the subject; (b) determining the presence and/or level of one or more biomarkers in a biological sample of the subject selected from Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15 or a combination thereof, wherein the detection of the one or more biomarkers indicates the presence of pancreatic cancer in the subject.

In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of one or more biomarkers in a biological sample from the subject, wherein the one or more biomarkers include one or more biomarkers of the TGF β /integrin signaling pathway.

In certain embodiments, the one or more TGF β /integrin signaling pathway biomarkers include, but are not limited to, DES, IMMT, TPM1, SNRPE, VCAM1, GRB2, SHROOM3, HMOX1, POSTN, MMP10, MMP-2, THBS2, EWSR1, NOD1, ADAMTS9, AFP, SYNE1, SYNE2, EPHB1, UFD1L, TEAD1, RYR3, CMYA5, MYLK, TOP2B or a combination thereof. In certain embodiments, the one or more TGF β /integrin signaling pathway biomarkers are SYNE1, THBS2, TOP2B, TPM1 or a combination thereof.

In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of one or more biomarkers in a biological sample from the subject, wherein the one or more biomarkers include one or more biomarkers of the HNF4 α transcription network pathway. In certain embodiments, the one or more HNF4 α transcription network pathway biomarkers include, but are not limited to, OBSCN, LOXL3, MLEC, DNAH1, DNAH5, DNAH12, DNAH17, SCYL2, FKBP10, FLRT3, ZHX2(AFR1), ZNF804A, ACTN2 or a combination thereof. In certain embodiments, the one or more HNF4 α transcription network pathway biomarkers can be LOXL3, DNAH12, DNAH1 or a combination thereof.

In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of one or more biomarkers in a biological sample from the subject, wherein the biomarker includes one or more biomarkers selected from the RAS/p53/JUN/CTNB1 signaling pathway. In certain embodiments, the one or more RAS/p53/JUN/CTNB1 signaling pathway biomarkers include, but are not limited to, KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29 or a combination thereof. In certain embodiments, the one or more RAS/p53/JUN/CTNB1 signaling pathway biomarkers is LIMCH1.

In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of one or more biomarkers in a biological sample from the subject, wherein the biomarker includes one or more biomarkers selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof. In certain embodiments, the one or more biomarkers are STARD8 (DLC3), ATP2A1, RTTN or a combination thereof.

In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of one or more biomarkers in a biological sample from the subject, wherein the biomarker includes one or more biomarkers selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13, DES, IMMT, TPM1, SNRPE, VCAM1, GRB2, SHROOM3, HMOX1, POSTN, MMP10, MMP-2, THBS2, EWSR1, NOD1, ADAMTS9, AFP, SYNE1, SYNE2, EPHB1, UFD1L, TEAD1, RYR3, CMYA5, MYLK, TOP2B, KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29, OBSCN, LOXL3, MLEC, DNAH1, DNAH5, DNAH12, DNAH17, SCYL2, FKBP10, FLRT3, ZHX2 (AFR1), ZNF804A, ACTN2 or a combination thereof.

In certain embodiments of the present disclosure, the method of diagnosing, prognosing or screening for pancreatic cancer in a subject includes, obtaining a biological sample from the subject and determining the presence of one or more biomarkers in a biological sample of the subject selected from RTTN, DNAH12, TPM1, DNAH1, STARD8, ATP2A1, TOP2B, LIMCH1, SYNE1, THBS2, LOXL3 or a combination thereof, wherein the detection of the one or more biomarkers indicates the presence of pancreatic cancer in the subject.

In certain embodiments of the present disclosure, the biomarkers described herein and shown in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 15 can be detected individually, as described above, or in panels including at least two biomarkers, at least three biomarkers, at least four biomarkers, at least five biomarkers, at least six biomarkers, at least seven biomarkers or at least eight biomarkers. For example, when used in a panel test, the levels of at least two biomarkers can be optionally tested from the same biological sample obtained from

the subject (*e.g.*, by detecting the quantities or amounts of various biomarkers in the same blood sample obtained from a patient) or in different biological samples from the subject. When combined in a panel test, the panel test can include determining the presence and/or levels for each of 2, 3, 4, 5, 6, 10, 15, 20, 25, 35 or more different
 5 biomarkers. The combination of multiple biomarkers in a panel test serves to reduce the number of false positives and false negatives should an aberrant value for one particular member of the panel be found.

In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of two
 10 or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway and the HNF4 α transcription factor network. Non-limiting examples of TGF β /integrin signaling pathway and HNF4 α transcription factor network biomarkers
 15 are described above and in Tables 4-13 and 15. In certain embodiments, at least one TGF β /integrin signaling pathway biomarker can be SYNE1, THBS2, TOP2B, TPM1 or a combination thereof. In certain embodiments, the at least one HNF4 α transcription network pathway biomarker can be LOXL3, DNAH12, DNAH1 or a combination thereof.

20 In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of two or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway
 25 and the RAS/p53/JUN/CTNB1 signaling pathway. Non-limiting examples of TGF β /integrin signaling pathway and RAS/p53/JUN/CTNB1 signaling pathway biomarkers are described above and in Tables 4-13 and 15. In certain embodiments, the at least one TGF β /integrin signaling pathway biomarker can be SYNE1, THBS2, TOP2B, TPM1 or a combination thereof. In certain embodiments, the at least one
 30 RAS/p53/JUN/CTNB1 signaling pathway biomarker can be LIMCH1.

In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of two or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from the

TGF β /integrin signaling pathway and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof. In certain embodiments, the at least one

5 TGF β /integrin signaling pathway biomarker can be SYNE1, THBS2, TOP2B, TPM1 or a combination thereof. In certain embodiments, the at least one biomarker can be STARD8 (DLC3), ATP2A1, RTTN or a combination thereof.

In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of two

10 or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from each of the following signaling pathways or networks: the HNF4 α transcription factor network and the RAS/p53/JUN/CTNB1 signaling pathway. Non-limiting examples of HNF4 α transcription factor network and RAS/p53/JUN/CTNB1 signaling pathway

15 biomarkers are described above and in Tables 4-13 and 15. In certain embodiments, the at least one HNF4 α transcription network pathway biomarker can be LOXL3, DNAH12, DNAH1 or a combination thereof. In certain embodiments, the at least one RAS/p53/JUN/CTNB1 signaling pathway biomarker can be LIMCH1.

In certain embodiments of the present disclosure, the method of

20 diagnosing a subject with pancreatic cancer includes determining the presence of two or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from the RAS/p53/JUN/CTNB1 signaling pathway and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8

25 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN and ABCA13. In certain embodiments, the at least one RAS/p53/JUN/CTNB1 signaling pathway biomarker can be LIMCH1. In certain embodiments, the at least one biomarker can be STARD8 (DLC3), ATP2A1, RTTN or a combination thereof.

30 In certain embodiments of the present disclosure, the method of diagnosing a subject with pancreatic cancer includes determining the presence of two or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from the HNF4 α transcription factor network and at least one biomarker selected from MANF,

ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN and ABCA13. In certain embodiments, the at least one HNF4 α transcription network pathway biomarker can be LOXL3, DNAH12, DNAH1 or a combination thereof. In certain embodiments, the at least one biomarker can be STARD8 (DLC3), ATP2A1, RTTN or a combination thereof.

In certain embodiments of the present disclosure, the method of diagnosing, prognosing or screening for pancreatic cancer in a subject includes determining the presence of three or more or four or more biomarkers in a sample from the subject. For example, but not by way of limitation, the method of diagnosing a subject with pancreatic cancer includes determining the presence of three or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway, the HNF4 α transcription factor network and the RAS/p53/JUN/CTNB1 signaling pathway. In certain embodiments, the at least one TGF β /integrin signaling pathway biomarker can be SYNE1, THBS2, TOP2B, TPM1 or a combination thereof. In certain embodiments, the at least one HNF4 α transcription network pathway biomarker can be LOXL3, DNAH12, DNAH1 or a combination thereof. In certain embodiments, the at least one RAS/p53/JUN/CTNB1 signaling pathway biomarker can be LIMCH1.

In certain embodiments, the method of diagnosing a subject with pancreatic cancer includes determining the presence of three or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway and the HNF4 α transcription factor network, and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN and ABCA13. In certain embodiments, the at least one TGF β /integrin signaling pathway biomarker can be SYNE1, THBS2, TOP2B, TPM1 or a combination thereof. In certain embodiments, the at least one HNF4 α transcription network pathway biomarker can be LOXL3, DNAH12, DNAH1 or a combination

thereof. In certain embodiments, the at least one biomarker can be STARD8 (DLC3), ATP2A1, RTTN or a combination thereof.

In certain embodiments, the method of diagnosing a subject with pancreatic cancer includes determining the presence of three or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from each of the following signaling pathways or networks: the HNF4 α transcription factor network and the RAS/p53/JUN/CTNB1 signaling pathway, and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN and ABCA13. In certain embodiments, the at least one RAS/p53/JUN/CTNB1 signaling pathway biomarker can be LIMCH1. In certain embodiments, the at least one HNF4 α transcription network pathway biomarker can be LOXL3, DNAH12, DNAH1 or a combination thereof. In certain embodiments, the at least one biomarker can be STARD8 (DLC3), ATP2A1, RTTN or a combination thereof.

In certain embodiments, the method of diagnosing a subject with pancreatic cancer includes determining the presence of three or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway and the RAS/p53/JUN/CTNB1 signaling pathway, and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN and ABCA13. In certain embodiments, the at least one RAS/p53/JUN/CTNB1 signaling pathway biomarker can be LIMCH1. In certain embodiments, the at least one TGF β /integrin signaling pathway biomarker can be SYNE1, THBS2, TOP2B, TPM1 or a combination thereof. In certain embodiments, the at least one biomarker can be STARD8 (DLC3), ATP2A1, RTTN or a combination thereof.

In certain embodiments, the method of diagnosing a subject with pancreatic cancer includes determining the presence of four or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least one biomarker selected from each of the following

signaling pathways or networks: the TGF β /integrin signaling pathway, the HNF4 α transcription factor network and the RAS/p53/JUN/CTNB1 signaling pathway, and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or combination thereof. In certain embodiments, the at least one RAS/p53/JUN/CTNB1 signaling pathway biomarker can be LIMCH1. In certain embodiments, the at least one TGF β /integrin signaling pathway biomarker can be SYNE1, THBS2, TOP2B, TPM1 or a combination thereof. In certain embodiments, the at least one HNF4 α transcription network pathway biomarker can be LOXL3, DNAH12, DNAH1 or a combination thereof. In certain embodiments, the at least one biomarker can be STARD8 (DLC3), ATP2A1, RTTN or a combination thereof.

In certain embodiments, the method of diagnosing a subject with pancreatic cancer includes determining the presence of six or more biomarkers in a panel of biomarkers in a biological sample from the subject. For example, and not by way of limitation, the six or more biomarkers can be selected from the following signaling pathways or networks: the TGF β /integrin signaling pathway, the HNF4 α transcription factor network, the RAS/p53/JUN/CTNB1 signaling pathway or a combination thereof, and/or selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof. In certain embodiments, the method includes determining the presence of six or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein the panel of biomarkers includes at least two biomarkers selected from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway, the HNF4 α transcription factor network and the RAS/p53/JUN/CTNB1 signaling pathway.

In certain embodiments, the method of diagnosing a subject with pancreatic cancer includes determining the presence of eight or more biomarkers in a panel of biomarkers in a biological sample from the subject. For example, and not by way of limitation, the eight or more biomarkers can be selected from the following signaling pathways or networks: the TGF β /integrin signaling pathway, the HNF4 α transcription factor network, the RAS/p53/JUN/CTNB1 signaling pathway or a combination thereof, and/or selected from MANF, ZNF485, IMPA1, SVEP1,

KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof. In certain embodiments, the method includes determining the presence of eight or more biomarkers in a panel of biomarkers in a biological sample from the subject, wherein
 5 the panel of biomarkers includes at least two biomarkers selected from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway, the HNF4 α transcription factor network and the RAS/p53/JUN/CTNBI signaling pathway, and at least two biomarkers selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP,
 10 RAD51C, ATP2A1, NLRX1, ZNF160, RTTN and ABCA13.

The present disclosure further provides methods for differentially diagnosing a subject with early stage (e.g., resectable cancer) or advanced stage (invasive and/or metastatic cancer) cancer. In certain embodiments, the method to diagnose a subject with advanced stage pancreatic cancer includes determining the
 15 presence of one or more biomarkers in a biological sample from the subject, wherein the detection of the one or more biomarkers is an indication that the subject has metastatic pancreatic cancer. For example, by not by way of limitation, the one or more biomarkers can include DNAH12, DNAH1, STARD8, ATP2A1, TOP2B, THBS2 or a combination thereof.

In certain embodiments, the method to diagnose a subject with early stage and/or resectable pancreatic cancer includes determining the presence of one or more biomarkers in a biological sample from the subject, wherein the detection of the one or more biomarkers is an indication that the subject has early stage and/or resectable pancreatic cancer. For example, by not by way of limitation, the one or
 20 more biomarkers can include RTTN, DNAH12, TPM1, DNAH1, STARD8, ATP2A1, TOP2B, SYNE1, THBS2, LOXL3 or a combination thereof.

In certain embodiments, the information provided by the methods described herein can be used by the physician in determining the most effective course of treatment (e.g., preventative or therapeutic). A course of treatment refers to
 30 the measures taken for a patient after the assessment of increased risk for development of pancreatic cancer is made. For example, when a subject is identified to have an increased risk of developing cancer, the physician can determine whether frequent monitoring for biomarker detection is required as a prophylactic measure. Also, when the subject is determined to have pancreatic cancer (e.g., based on the

presence of one or more biomarkers in a biological sample from a subject), it can be advantageous to follow such detection with a biopsy, surgical treatment, chemotherapy, radiation, immunotherapy, biological modifier therapy, gene therapy, vaccines and the like, or adjust the span of time during which the patient is treated.

5 Biomarker Detection

A biomarker used in the methods of the disclosure can be identified in a biological sample using any method known in the art. Determining the presence of a biomarker, protein or degradation product thereof, the presence of mRNA or pre-mRNA, or the presence of any biological molecule or product that is indicative of
10 biomarker expression, or degradation product thereof, can be carried out for use in the methods of the disclosure by any method described herein or known in the art.

Protein Detection Techniques

Methods for the detection of protein biomarkers are well known to those skilled in the art, and include but are not limited to mass spectrometry
15 techniques, 1-D or 2-D gel-based analysis systems, chromatography, enzyme linked immunosorbent assays (ELISAs), radioimmunoassays (RIA), enzyme immunoassays (EIA), Western Blotting, immunoprecipitation and immunohistochemistry. These methods use antibodies, or antibody equivalents, to detect protein, or use biophysical techniques. Antibody arrays or protein chips can also be employed, see for example
20 U.S. Patent Application Nos: 2003/0013208A1; 2002/0155493A1, 2003/0017515 and U.S. Pat. Nos. 6,329,209 and 6,365,418, herein incorporated by reference in their entireties.

ELISA and RIA procedures can be conducted such that a biomarker standard is labeled (with a radioisotope such as ^{125}I or ^{35}S , or an assayable enzyme,
25 such as horseradish peroxidase or alkaline phosphatase), and, together with the unlabeled sample, brought into contact with the corresponding antibody, whereon a second antibody is used to bind the first, and radioactivity or the immobilized enzyme assayed (competitive assay). Alternatively, the biomarker in the sample is allowed to react with the corresponding immobilized antibody, radioisotope or enzyme-labeled
30 anti-biomarker antibody is allowed to react with the system, and radioactivity or the enzyme assayed (ELISA-sandwich assay). Other conventional methods can also be employed as suitable.

The above techniques can be conducted essentially as a “one-step” or “two-step” assay. A “one-step” assay involves contacting antigen with immobilized

antibody and, without washing, contacting the mixture with labeled antibody. A “two-step” assay involves washing before contacting the mixture with labeled antibody. Other conventional methods can also be employed as suitable.

In certain embodiments, a method for measuring biomarker expression includes the steps of: contacting a biological sample, *e.g.*, blood and/or plasma, with an antibody or variant (*e.g.*, fragment) thereof which selectively binds the biomarker, and detecting whether the antibody or variant thereof is bound to the sample. A method can further include contacting the sample with a second antibody, *e.g.*, a labeled antibody. The method can further include one or more steps of washing, *e.g.*,
10 to remove one or more reagents.

It can be desirable to immobilize one component of the assay system on a support, thereby allowing other components of the system to be brought into contact with the component and readily removed without laborious and time-consuming labor. It is possible for a second phase to be immobilized away from the first, but one phase is usually sufficient.
15

It is possible to immobilize the enzyme itself on a support, but if solid-phase enzyme is required, then this is generally best achieved by binding to antibody and affixing the antibody to a support, models and systems for which are well-known in the art. Simple polyethylene can provide a suitable support.

Enzymes employable for labeling are not particularly limited, but can be selected, for example, from the members of the oxidase group. These catalyze production of hydrogen peroxide by reaction with their substrates, and glucose oxidase is often used for its good stability, ease of availability and cheapness, as well as the ready availability of its substrate (glucose). Activity of the oxidase can be
20 assayed by measuring the concentration of hydrogen peroxide formed after reaction of the enzyme-labeled antibody with the substrate under controlled conditions well-known in the art.
25

Other techniques can be used to detect a biomarker according to a practitioner's preference based upon the present disclosure. One such technique that can be used for detecting and quantitating biomarker protein levels is Western blotting (Towbin et al., Proc. Nat. Acad. Sci. 76:4350 (1979)). Cells can be frozen,
30 homogenized in lysis buffer, and the lysates subjected to SDS-PAGE and blotting to a membrane, such as a nitrocellulose filter. Antibodies (unlabeled) are then brought into contact with the membrane and assayed by a secondary immunological reagent, such

as labeled protein A or anti-immunoglobulin (suitable labels including ¹²⁵I, horseradish peroxidase and alkaline phosphatase). Chromatographic detection can also be used. In certain embodiments, immunodetection can be performed with antibody to a biomarker using the enhanced chemiluminescence system (*e.g.*, from
5 PerkinElmer Life Sciences, Boston, Mass.). The membrane can then be stripped and re-blotted with a control antibody, *e.g.*, anti-actin (A-2066) polyclonal antibody from Sigma (St. Louis, Mo.).

Immunohistochemistry can be used to detect the expression and/
presence of a biomarker, *e.g.*, in a biopsy sample. A suitable antibody is brought into
10 contact with, for example, a thin layer of cells, followed by washing to remove unbound antibody, and then contacted with a second, labeled, antibody. Labeling can be by fluorescent markers, enzymes, such as peroxidase, avidin or radiolabeling. The assay is scored visually, using microscopy and the results can be quantitated.

Other machine or autoimaging systems can also be used to measure
15 immunostaining results for the biomarker. As used herein, “quantitative” immunohistochemistry refers to an automated method of scanning and scoring samples that have undergone immunohistochemistry, to identify and quantitate the presence of a specified biomarker, such as an antigen or other protein. The score given to the sample is a numerical representation of the intensity of the
20 immunohistochemical staining of the sample, and represents the amount of target biomarker present in the sample. As used herein, Optical Density (OD) is a numerical score that represents intensity of staining. As used herein, semi-quantitative immunohistochemistry refers to scoring of immunohistochemical results by human eye, where a trained operator ranks results numerically (*e.g.*, as 1, 2 or 3).

25 Various automated sample processing, scanning and analysis systems suitable for use with immunohistochemistry are available in the art. Such systems can include automated staining (see, *e.g.*, the Benchmark system, Ventana Medical Systems, Inc.) and microscopic scanning, computerized image analysis, serial section comparison (to control for variation in the orientation and size of a sample), digital
30 report generation, and archiving and tracking of samples (such as slides on which tissue sections are placed). Cellular imaging systems are commercially available that combine conventional light microscopes with digital image processing systems to perform quantitative analysis on cells and tissues, including immunostained samples. See, *e.g.*, the CAS-200 system (Becton, Dickinson & Co.).

Antibodies against biomarkers can also be used for imaging purposes, for example, to detect the presence of a biomarker in cells of a subject. Suitable labels include radioisotopes, iodine (^{125}I , ^{121}I), carbon (^{14}C), sulphur (^{35}S), tritium (^3H), indium (^{112}In), and technetium ($^{99\text{m}}\text{Tc}$), fluorescent labels, such as fluorescein and
5 rhodamine and biotin. Immunoenzymatic interactions can be visualized using different enzymes such as peroxidase, alkaline phosphatase, or different chromogens such as DAB, AEC or Fast Red.

For *in vivo* imaging purposes, antibodies are not detectable, as such, from outside the body, and so must be labeled, or otherwise modified, to permit
10 detection. Markers for this purpose can be any that do not substantially interfere with the antibody binding, but which allow external detection. Suitable markers can include those that can be detected by X-radiography, NMR or MRI. For X-radiographic techniques, suitable markers include any radioisotope that emits detectable radiation but that is not overtly harmful to the subject, such as barium or
15 caesium, for example. Suitable markers for NMR and MRI generally include those with a detectable characteristic spin, such as deuterium, which can be incorporated into the antibody by suitable labeling of nutrients for the relevant hybridoma, for example.

The size of the subject, and the imaging system used, will determine
20 the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of technetium-99m.

The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain a biomarker. The labeled antibody or
25 variant thereof, *e.g.*, antibody fragment, can then be detected using known techniques. Antibodies include any antibody, whether natural or synthetic, full length or a fragment thereof, monoclonal or polyclonal, that binds sufficiently strongly and specifically to the biomarker to be detected. An antibody can have a K_d of at most about 10^{-6}M , 10^{-7}M , 10^{-8}M , 10^{-9}M , 10^{-10}M , 10^{-11}M , 10^{-12}M . The phrase “specifically
30 binds” refers to binding of, for example, an antibody to an epitope or antigen or antigenic determinant in such a manner that binding can be displaced or competed with a second preparation of identical or similar epitope, antigen or antigenic determinant.

Antibodies and derivatives thereof that can be used encompasses polyclonal or monoclonal antibodies, chimeric, human, humanized, primatized (CDR-grafted), veneered or single-chain antibodies, phage produced antibodies (*e.g.*, from phage display libraries), as well as functional binding fragments, of antibodies. For example, antibody fragments capable of binding to a biomarker, or portions thereof, including, but not limited to Fv, Fab, Fab' and F(ab')₂ fragments can be used. Such fragments can be produced by enzymatic cleavage or by recombinant techniques. For example, papain or pepsin cleavage can generate Fab or F(ab')₂ fragments, respectively. Other proteases with the requisite substrate specificity can also be used to generate Fab or F(ab')₂ fragments. Antibodies can also be produced in a variety of truncated forms using antibody genes in which one or more stop codons have been introduced upstream of the natural stop site. For example, a chimeric gene encoding a F(ab')₂ heavy chain portion can be designed to include DNA sequences encoding the CH₁ domain and hinge region of the heavy chain.

Synthetic and engineered antibodies are described in, *e.g.*, Cabilly et al., U.S. Pat. No. 4,816,567 Cabilly et al., European Patent No. 0,125,023 B1; Boss et al., U.S. Pat. No. 4,816,397; Boss et al., European Patent No. 0,120,694 B1; Neuberger, M. S. et al., WO 86/01533; Neuberger, M. S. et al., European Patent No. 0,194,276 B1; Winter, U.S. Pat. No. 5,225,539; Winter, European Patent No. 0,239,400 B1; Queen et al., European Patent No. 0451216 B1; and Padlan, E. A. et al., EP 0519596 A1. See also, Newman, R. et al., *BioTechnology*, 10: 1455-1460 (1992), regarding primatized antibody, and Ladner et al., U.S. Pat. No. 4,946,778 and Bird, R. E. et al., *Science*, 242: 423-426 (1988)) regarding single-chain antibodies.

In certain embodiments, agents that specifically bind to a polypeptide other than antibodies are used, such as peptides. Peptides that specifically bind can be identified by any means known in the art, *e.g.*, peptide phage display libraries. Generally, an agent that is capable of detecting a biomarker polypeptide, such that the presence of a biomarker is detected and/or quantitated, can be used. As defined herein, an "agent" refers to a substance that is capable of identifying or detecting a biomarker in a biological sample (*e.g.*, identifies or detects the mRNA of a biomarker, the DNA of a biomarker, the protein of a biomarker). In certain embodiments, the agent is a labeled or labelable antibody which specifically binds to a biomarker polypeptide.

In addition, a biomarker can be detected using Mass Spectrometry such as MALDI/TOF (time-of-flight), SELDI/TOF, liquid chromatography-mass

spectrometry (LC-MS), gas chromatography-mass spectrometry (GC-MS), high performance liquid chromatography-mass spectrometry (HPLC-MS), capillary electrophoresis-mass spectrometry, nuclear magnetic resonance spectrometry, or tandem mass spectrometry (*e.g.*, MS/MS, MS/MS/MS, ESI-MS/MS, etc.). *See* for
5 example, U.S. Patent Application Nos: 20030199001, 20030134304, 20030077616, which are herein incorporated by reference.

Mass spectrometry methods are well known in the art and have been used to quantify and/or identify biomolecules, such as proteins (*see, e.g.*, Li et al. (2000) *Tibtech* 18:151-160; Rowley et al. (2000) *Methods* 20: 383-397; and Kuster
10 and Mann (1998) *Curr. Opin. Structural Biol.* 8: 393-400). Further, mass spectrometric techniques have been developed that permit at least partial *de novo* sequencing of isolated proteins. Chait et al., *Science* 262:89-92 (1993); Keough et al., *Proc. Natl. Acad. Sci. USA.* 96:7131-6 (1999); reviewed in Bergman, *EXS* 88:133-44 (2000).

15 In certain embodiments, a gas phase ion spectrophotometer is used. In other embodiments, laser-desorption/ionization mass spectrometry is used to analyze the sample. Modern laser desorption/ionization mass spectrometry ("LDI-MS") can be practiced in two main variations: matrix assisted laser desorption/ionization ("MALDI") mass spectrometry and surface-enhanced laser desorption/ionization
20 ("SELDI"). In MALDI, the analyte is mixed with a solution containing a matrix, and a drop of the liquid is placed on the surface of a substrate. The matrix solution then co-crystallizes with the biological molecules. The substrate is inserted into the mass spectrometer. Laser energy is directed to the substrate surface where it desorbs and ionizes the biological molecules without significantly fragmenting them. However,
25 MALDI has limitations as an analytical tool. It does not provide means for fractionating the sample, and the matrix material can interfere with detection, especially for low molecular weight analytes. *See, e.g.*, U.S. Pat. No. 5,118,937 (Hillenkamp et al.), and U.S. Pat. No. 5,045,694 (Beavis & Chait).

For additional information regarding mass spectrometers, *see, e.g.*,
30 *Principles of Instrumental Analysis*, 3rd edition. Skoog, Saunders College Publishing, Philadelphia, 1985; and *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed. Vol. 15 (John Wiley & Sons, New York 1995), pp. 1071-1094.

Detection of the presence of a marker or other substances will typically involve detection of signal intensity. This, in turn, can reflect the quantity and

character of a polypeptide bound to the substrate. For example, in certain embodiments, the signal strength of peak values from spectra of a first sample and a second sample can be compared (*e.g.*, visually, by computer analysis *etc.*), to determine the relative amounts of a particular biomarker. Software programs such as the Biomarker Wizard program (CIPHERGEN Biosystems, Inc., Fremont, Calif.) can be used to aid in analyzing mass spectra. The mass spectrometers and their techniques are well known to those of skill in the art.

Any person skilled in the art understands, any of the components of a mass spectrometer (*e.g.*, desorption source, mass analyzer, detect, *etc.*) and varied sample preparations can be combined with other suitable components or preparations described herein, or to those known in the art. For example, in certain embodiments a control sample can contain heavy atoms (*e.g.*, ^{13}C) thereby permitting the test sample to be mixed with the known control sample in the same mass spectrometry run.

In certain embodiments, a laser desorption time-of-flight (TOF) mass spectrometer is used. In laser desorption mass spectrometry, a substrate with a bound marker is introduced into an inlet system. The marker is desorbed and ionized into the gas phase by laser from the ionization source. The ions generated are collected by an ion optic assembly, and then in a time-of-flight mass analyzer, ions are accelerated through a short high voltage field and let drift into a high vacuum chamber. At the far end of the high vacuum chamber, the accelerated ions strike a sensitive detector surface at a different time. Since the time-of-flight is a function of the mass of the ions, the elapsed time between ion formation and ion detector impact can be used to identify the presence or absence of molecules of specific mass to charge ratio.

In certain embodiments, the relative amounts of one or more biomarkers present in a first or second sample is determined, in part, by executing an algorithm with a programmable digital computer. The algorithm identifies at least one peak value in the first mass spectrum and the second mass spectrum. The algorithm then compares the signal strength of the peak value of the first mass spectrum to the signal strength of the peak value of the second mass spectrum of the mass spectrum. The relative signal strengths are an indication of the amount of the biomarker that is present in the first and second samples. A standard containing a known amount of a biomarker can be analyzed as the second sample to better quantify the amount of the biomarker present in the first sample. In certain embodiments, the identity of the biomarkers in the first and second sample can also be determined.

RNA Detection Techniques

Any method for qualitatively or quantitatively detecting a nucleic acid biomarker can be used. Detection of RNA transcripts can be achieved, for example, by Northern blotting, wherein a preparation of RNA is run on a denaturing agarose gel, and transferred to a suitable support, such as activated cellulose, nitrocellulose or glass or nylon membranes. Radiolabeled cDNA or RNA is then hybridized to the preparation, washed and analyzed by autoradiography.

Detection of RNA transcripts can further be accomplished using amplification methods. For example, it is within the scope of the present disclosure to reverse transcribe mRNA into cDNA followed by polymerase chain reaction (RT-PCR); or, to use a single enzyme for both steps as described in U.S. Pat. No. 5,322,770, or reverse transcribe mRNA into cDNA followed by symmetric gap ligase chain reaction (RT-AGLCR) as described by R. L. Marshall, et al., PCR Methods and Applications 4: 80-84 (1994).

In certain embodiments, quantitative real-time polymerase chain reaction (qRT-PCR) is used to evaluate mRNA levels of biomarker. The levels of a biomarker and a control mRNA can be quantitated in cancer tissue or cells and adjacent benign tissues. In one specific embodiment, the levels of one or more biomarkers can be quantitated in a biological sample.

Other known amplification methods which can be utilized herein include but are not limited to the so-called "NASBA" or "3SR" technique described in PNAS USA 87: 1874-1878 (1990) and also described in Nature 350 (No. 6313): 91-92 (1991); Q-beta amplification as described in published European Patent Application (EPA) No. 4544610; strand displacement amplification (as described in G. T. Walker et al., Clin. Chem. 42: 9-13 (1996) and European Patent Application No. 684315; and target mediated amplification, as described by PCT Publication WO9322461.

In situ hybridization visualization can also be employed, wherein a radioactively labeled antisense RNA probe is hybridized with a thin section of a biopsy sample, washed, cleaved with RNase and exposed to a sensitive emulsion for autoradiography. The samples can be stained with haematoxylin to demonstrate the histological composition of the sample, and dark field imaging with a suitable light filter shows the developed emulsion. Non-radioactive labels such as digoxigenin can also be used.

Another method for evaluation of biomarker expression is to detect mRNA levels of a biomarker by fluorescent in situ hybridization (FISH). FISH is a technique that can directly identify a specific region of DNA or RNA in a cell and therefore enables to visual determination of the biomarker expression in tissue samples. The FISH method has the advantages of a more objective scoring system and the presence of a built-in internal control consisting of the biomarker gene signals present in all non-neoplastic cells in the same sample. Fluorescence in situ hybridization is a direct in situ technique that is relatively rapid and sensitive. FISH test also can be automated. Immunohistochemistry can be combined with a FISH method when the expression level of the biomarker is difficult to determine by immunohistochemistry alone.

Alternatively, mRNA expression can be detected on a DNA array, chip or a microarray. Oligonucleotides corresponding to the biomarker(s) are immobilized on a chip which is then hybridized with labeled nucleic acids of a test sample obtained from a subject. Positive hybridization signal is obtained with the sample containing biomarker transcripts. Methods of preparing DNA arrays and their use are well known in the art. (See, for example, U.S. Pat. Nos. 6,618,6796; 6,379,897; 6,664,377; 6,451,536; 548,257; U.S. 20030157485 and Schena et al. 1995 Science 20:467-470; Gerhold et al. 1999 Trends in Biochem. Sci. 24, 168-173; and Lennon et al. 2000 Drug discovery Today 5: 59-65, which are herein incorporated by reference in their entirety). Serial Analysis of Gene Expression (SAGE) can also be performed (See for example U.S. Patent Application 20030215858).

To monitor mRNA levels, for example, mRNA can be extracted from the biological sample to be tested, reverse transcribed and fluorescent-labeled cDNA probes are generated. The microarrays capable of hybridizing to a biomarker, cDNA can then probed with the labeled cDNA probes, the slides scanned and fluorescence intensity measured. This intensity correlates with the hybridization intensity and expression levels.

Types of probes for detection of RNA include cDNA, riboprobes, synthetic oligonucleotides and genomic probes. The type of probe used will generally be dictated by the particular situation, such as riboprobes for in situ hybridization, and cDNA for Northern blotting, for example. In certain embodiments, the probe is directed to nucleotide regions unique to the particular biomarker RNA. The probes can be as short as is required to differentially recognize the particular biomarker

mRNA transcripts, and can be as short as, for example, 15 bases; however, probes of at least 17 bases, at least 18 bases and at least 20 bases can be used. In certain embodiments, the primers and probes hybridize specifically under stringent conditions to a nucleic acid fragment having the nucleotide sequence corresponding to the target gene. As herein used, the term “stringent conditions” means hybridization will occur only if there is at least 95% or at least 97% identity between the sequences.

The form of labeling of the probes can be any that is appropriate, such as the use of radioisotopes, for example, ^{32}P and ^{35}S . Labeling with radioisotopes can be achieved, whether the probe is synthesized chemically or biologically, by the use of suitably labeled bases.

Kits

In certain non-limiting embodiments, the present disclosure provides for a kit for determining whether a subject has pancreatic cancer includes a means for detecting one or more biomarkers selected from the biomarkers set forth in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 15, or a combination thereof. The disclosure further provides for kits for determining the efficacy of a therapy for preventing or treating pancreatic cancer in a subject.

Types of kits include, but are not limited to, packaged probe and primer sets (e.g., TaqMan probe/primer sets), arrays/microarrays, biomarker-specific antibodies and beads, which further contain one or more probes, primers or other detection reagents for detecting one or more biomarkers of the present disclosure.

In a certain, non-limiting embodiment, a kit can include a pair of oligonucleotide primers suitable for polymerase chain reaction (PCR) or nucleic acid sequencing, for detecting one or more biomarker(s) to be identified. A pair of primers can include nucleotide sequences complementary to a biomarker set forth in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 15, and can be of sufficient length to selectively hybridize with said biomarker. Alternatively, the complementary nucleotides can selectively hybridize to a specific region in close enough proximity 5' and/or 3' to the biomarker position to perform PCR and/or sequencing. Multiple biomarker-specific primers can be included in the kit to simultaneously assay large number of biomarkers. The kit can also include one or more polymerases, reverse transcriptase and nucleotide bases, wherein the nucleotide bases can be further detectably labeled.

In certain embodiments, a primer can be at least about 10 nucleotides or at least about 15 nucleotides or at least about 20 nucleotides in length and/or up to

about 200 nucleotides or up to about 150 nucleotides or up to about 100 nucleotides or up to about 75 nucleotides or up to about 50 nucleotides in length.

In certain embodiments, the oligonucleotide primers can be immobilized on a solid surface or support, for example, on a nucleic acid microarray, wherein the position of each oligonucleotide primer bound to the solid surface or support is known and identifiable.

In a certain, non-limiting embodiment, a kit can include at least one nucleic acid probe, suitable for in situ hybridization or fluorescent in situ hybridization, for detecting the biomarker(s) to be identified. Such kits will generally include one or more oligonucleotide probes that have specificity for various biomarkers.

In certain non-limiting embodiments, a kit can include a primer for detection of a biomarker by primer extension.

In certain non-limiting embodiments, a kit can include at least one antibody for immunodetection of the biomarker(s) to be identified. Antibodies, both polyclonal and monoclonal, specific for a biomarker, can be prepared using conventional immunization techniques, as will be generally known to those of skill in the art. The immunodetection reagents of the kit can include detectable labels that are associated with, or linked to, the given antibody or antigen itself. Such detectable labels include, for example, chemiluminescent or fluorescent molecules (rhodamine, fluorescein, green fluorescent protein, luciferase, Cy3, Cy5 or ROX), radiolabels (^3H , ^{35}S , ^{32}P , ^{14}C , ^{131}I) or enzymes (alkaline phosphatase, horseradish peroxidase).

In a certain non-limiting embodiment, the biomarker-specific antibody can be provided bound to a solid support, such as a column matrix, an array, or well of a microtiter plate. Alternatively, the support can be provided as a separate element of the kit.

In certain non-limiting embodiments, a kit can include one or more primers, probes, microarrays, or antibodies suitable for detecting one or more biomarkers set forth in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15 or combinations thereof.

In certain non-limiting embodiments, a kit can include one or more primers, probes, microarrays, or antibodies suitable for detecting one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen or more of the following biomarkers: MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529,

GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13, DES, IMMT, TPM1, SNRPE, VCAM1, GRB2, SHROOM3, HMOX1, POSTN, MMP10, MMP-2, THBS2, EWSR1, NOD1, ADAMTS9, AFP, SYNE1, SYNE2, EPHB1, UFD1L, TEAD1, RYR3,
 5 CMYA5, MYLK, TOP2B, KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29, OBSCN, LOXL3, MLEC, DNAH1, DNAH5, DNAH12, DNAH17, SCYL2, FKBP10, FLRT3, ZHX2 (AFR1), ZNF804A and ACTN2.

In certain non-limiting embodiments, a kit can include one or more primers, probes, microarrays, or antibodies suitable for detecting one, two, three, four,
 10 five, six or seven of the following biomarkers: KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29.

In certain non-limiting embodiments, a kit can include one or more primers, probes, microarrays, or antibodies suitable for detecting one, two, three, four, five, six, seven, eight, nine, ten or eleven of the following biomarkers: RTTN, DNAH12, TPM1, DNAH1, STARD8, ATP2A1, TOP2B, LIMCH1, SYNE1, THBS2
 15 and LOXL3.

In certain embodiments, a kit can include one or more primers, probes, microarrays, or antibodies suitable for detecting one or more biomarkers of the TGF β /integrin signaling pathway, including but not limited to, DES, IMMT, TPM1, SNRPE, VCAM1, GRB2, SHROOM3, HMOX1, POSTN, MMP10, MMP-2, THBS2, EWSR1, NOD1, ADAMTS9, AFP, SYNE1, SYNE2, EPHB1, UFD1L, TEAD1, RYR3, CMYA5, MYLK, TOP2B or a combination thereof.
 20

In certain embodiments, a kit can include one or more primers, probes, microarrays, or antibodies suitable for detecting one or more biomarkers of the RAS/p53/JUN/CTNB1 signaling pathway, including but not limited to, KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29 or a combination thereof.
 25

In certain embodiments, a kit can include one or more primers, probes, microarrays, or antibodies suitable for detecting one or more biomarkers of the HNF4 α transcription network pathway, including but not limited to, OBSCN, LOXL3, MLEC, DNAH1, DNAH5, DNAH12, DNAH17, SCYL2, FKBP10, FLRT3, ZHX2(AFR1), ZNF804A, ACTN2 or a combination thereof.
 30

In certain embodiments, a kit can include one or more primers, probes, microarrays, or antibodies suitable for detecting one or more biomarkers selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS,

STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof.

In certain embodiments, a kit can include two or more primers, probes, microarrays, or antibodies suitable for detecting two or more biomarkers, where the
 5 kit includes at least one or more biomarkers from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway, *e.g.*, SYNE1, THBS2, TOP2B and/or TPM1, and the HNF4 α transcription factor network, *e.g.*, LOXL3, DNAH12 and/or DNAH1.

In certain embodiments, a kit can include two or more primers, probes,
 10 microarrays, or antibodies suitable for detecting two or more biomarkers, where the kit includes at least one or more biomarkers from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway, *e.g.*, SYNE1, THBS2, TOP2B and/or TPM1, and the RAS/p53/JUN/CTNB1 signaling pathway, *e.g.*, LIMCH1.

In certain embodiments, a kit can include two or more primers, probes, microarrays, or antibodies suitable for detecting two or more biomarkers, where the
 15 kit includes at least one or more biomarkers from each of the following signaling pathways or networks: the HNF4 α transcription factor network, *e.g.*, LOXL3, DNAH12 and/or DNAH1, and the RAS/p53/JUN/CTNB1 signaling pathway, *e.g.*,
 20 LIMCH1.

In certain embodiments, a kit can include two or more primers, probes, microarrays, or antibodies suitable for detecting two or more biomarkers, where the
 kit includes at least one or more biomarkers from the TGF β /integrin signaling pathway, *e.g.*, SYNE1, THBS2, TOP2B and/or TPM1, and at least one or more
 25 biomarkers selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof.

In certain embodiments, a kit can include two or more primers, probes, microarrays, or antibodies suitable for detecting two or more biomarkers, where the
 30 kit includes at least one or more biomarkers from the HNF4 α transcription factor network, *e.g.*, LOXL3, DNAH12 and/or DNAH1, and at least one or more biomarkers selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof.

In certain embodiments, a kit can include two or more primers, probes, microarrays, or antibodies suitable for detecting two or more biomarkers, where the kit includes at least one or more biomarkers from the RAS/p53/JUN/CTNB1 signaling pathway, *e.g.*, LIMCH1, and at least one or more biomarkers selected from
 5 MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof.

In certain embodiments, a kit can include three or more primers, probes, microarrays, or antibodies suitable for detecting three or more biomarkers,
 10 where the kit includes at least one or more biomarkers from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway, the HNF4 α transcription factor network and the RAS/p53/JUN/CTNB1 signaling pathway.

In certain embodiments, a kit can include three or more primers, probes, microarrays, or antibodies suitable for detecting three or more biomarkers,
 15 where the kit includes at least one or more biomarkers from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway and the HNF4 α transcription factor network, and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160,
 20 RTTN and ABCA13.

In certain embodiments, a kit can include three or more primers, probes, microarrays, or antibodies suitable for detecting three or more biomarkers, where the kit includes at least one or more biomarkers from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway and the
 25 RAS/p53/JUN/CTNB1 signaling pathway, and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN and ABCA13.

In certain embodiments, a kit can include three or more primers,
 30 probes, microarrays, or antibodies suitable for detecting three or more biomarkers, where the kit includes at least one or more biomarkers from each of the following signaling pathways or networks: the HNF4 α transcription factor network and the RAS/p53/JUN/CTNB1 signaling pathway, and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8

(DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof.

In certain embodiments, a kit can include four or more primers, probes, microarrays, or antibodies suitable for detecting four or more biomarkers, where the kit includes at least one or more biomarkers from each of the following signaling pathways or networks: the TGF β /integrin signaling pathway, the HNF4 α transcription factor network and the RAS/p53/JUN/CTNB1 signaling pathway, and at least one biomarker selected from MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 or a combination thereof.

In certain non-limiting embodiments, where the measurement means in the kit employs an array, the set of biomarkers set forth above can constitute at least 10 percent or at least 20 percent or at least 30 percent or at least 40 percent or at least 50 percent or at least 60 percent or at least 70 percent or at least 80 percent of the species of markers represented on the microarray. In certain non-limiting embodiments, a biomarker detection kit can include one or more detection reagents and other components (*e.g.*, a buffer, enzymes such as DNA polymerases or ligases, chain extension nucleotides such as deoxynucleotide triphosphates, and in the case of Sanger-type DNA sequencing reactions, chain terminating nucleotides, positive control sequences, negative control sequences, and the like) necessary to carry out an assay or reaction to detect a biomarker. A kit can also include additional components or reagents necessary for the detection of a biomarker, such as secondary antibodies for use in western blotting immunohistochemistry. A kit can further include one or more other biomarkers or reagents for evaluating other prognostic factors, *e.g.*, tumor stage.

A kit can further contain means for comparing the biomarker with a standard, and can include instructions for using the kit to detect the biomarker of interest. For example, the instructions can describe that the presence of a biomarker, set forth herein, is indicative that the subject has or will develop pancreatic cancer.

30 Reports, Programmed Computers and Systems

The results of a test (*e.g.*, an individual's risk for cancer, such as pancreatic cancer), or an individual's predicted drug responsiveness (*e.g.*, response to chemotherapy), based on assaying one or more biomarkers set forth in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 15, and/or any other information pertaining to a test, can be

referred to herein as a “report.” A tangible report can optionally be generated as part of a testing process (which can be interchangeably referred to herein as “reporting,” or as “providing” a report, “producing” a report or “generating” a report).

Examples of tangible reports can include, but are not limited to, reports
5 in paper (such as computer-generated printouts of test results) or equivalent formats and reports stored on computer readable medium (such as a CD, USB flash drive or other removable storage device, computer hard drive, or computer network server, etc.). Reports, particularly those stored on computer readable medium, can be part of
10 a database, which can optionally be accessible via the internet (such as a database of patient records or genetic information stored on a computer network server, which can be a “secure database” that has security features that limit access to the report, such as to allow only the patient and the patient's medical practitioners to view the report while preventing other unauthorized individuals from viewing the report, for example). In addition to, or as an alternative to, generating a tangible report, reports
15 can also be displayed on a computer screen (or the display of another electronic device or instrument).

A report can include, for example, an individual's risk for cancer, such as pancreatic cancer, or can just include presence, absence or levels of one or more biomarkers set forth in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 15 (for example, a
20 report on computer readable medium such as a network server can include hyperlink(s) to one or more journal publications or websites that describe the medical/biological implications, such as increased or decreased disease risk, for individuals having certain biomarkers or levels of certain biomarkers). Thus, for example, the report can include disease risk or other medical/biological significance
25 (e.g., drug responsiveness, suggested prophylactic treatment, etc.) as well as optionally also including the biomarker information, or the report can just include biomarker information without including disease risk or other medical/biological significance (such that an individual viewing the report can use the biomarker information to determine the associated disease risk or other medical/biological
30 significance from a source outside of the report itself, such as from a medical practitioner, publication, website, etc., which can optionally be linked to the report such as by a hyperlink).

A report can further be “transmitted” or “communicated” (these terms can be used herein interchangeably), such as to the individual who was tested, a

medical practitioner (*e.g.*, a doctor, nurse, clinical laboratory practitioner, genetic counselor, etc.), a healthcare organization, a clinical laboratory and/or any other party or requester intended to view or possess the report. The act of “transmitting” or “communicating” a report can be by any means known in the art, based on the format of the report. Furthermore, “transmitting” or “communicating” a report can include delivering a report (“pushing”) and/or retrieving (“pulling”) a report. For example, reports can be transmitted/communicated by various means, including being physically transferred between parties (such as for reports in paper format) such as by being physically delivered from one party to another, or by being transmitted electronically or in signal form (*e.g.*, via e-mail or over the internet, by facsimile and/or by any wired or wireless communication methods known in the art) such as by being retrieved from a database stored on a computer network server, etc.

In certain exemplary embodiments, the disclosed subject matter provides computers (or other apparatus/devices such as biomedical devices or laboratory instrumentation) programmed to carry out the methods described herein. For example, in certain embodiments, the disclosed subject matter provides a computer programmed to receive (*i.e.*, as input) the identity of the one or more biomarkers disclosed herein, alone or in combination with other biomarkers, and provide (*i.e.*, as output) the disease risk (*e.g.*, risk of pancreatic cancer) or other result (*e.g.*, disease diagnosis or prognosis, drug responsiveness, etc.) based on the level or identity of the biomarker(s). Such output (*e.g.*, communication of disease risk, disease diagnosis or prognosis, drug responsiveness, etc.) can be, for example, in the form of a report on computer readable medium, printed in paper form, and/or displayed on a computer screen or other display.

Certain further embodiments of the disclosed subject matter provide a system for determining an individual's cancer risk, or whether an individual will benefit from chemotherapy treatment (or other therapy) or prophylactic treatment. Certain exemplary systems include an integrated “loop” in which an individual (or their medical practitioner) requests a determination of such individual's cancer risk (or drug response), this determination is carried out by testing a sample from the individual, and then the results of this determination are provided back to the requester. For example, in certain systems, a sample (*e.g.*, stool, blood, etc.) is obtained from an individual for testing (the sample can be obtained by the individual or, for example, by a medical practitioner), the sample is submitted to a laboratory (or

other facility) for testing (*e.g.*, determining the biomarker(s) disclosed herein, alone or in combination with one or more other biomarkers), and then the results of the testing are sent to the patient (which optionally can be done by first sending the results to an intermediary, such as a medical practitioner, who then provides or otherwise conveys the results to the individual and/or acts on the results), thereby forming an integrated loop system for determining an individual's cancer risk (or drug response, etc.). The portions of the system in which the results are transmitted (*e.g.*, between any of a testing facility, a medical practitioner, and/or the individual) can be carried out by way of electronic or signal transmission (*e.g.*, by computer such as via e-mail or the internet, by providing the results on a website or computer network server which can optionally be a secure database, by phone or fax, or by any other wired or wireless transmission methods known in the art).

In certain embodiments, the system is controlled by the individual and/or their medical practitioner in that the individual and/or their medical practitioner requests the test, receives the test results back, and (optionally) acts on the test results to reduce the individual's disease risk, such as by implementing a disease management system.

The various methods described herein, such as correlating the presence or absence or level of a biomarker with an altered (*e.g.*, increased or decreased) risk (or no altered risk) for cancer, *e.g.*, pancreatic cancer, can be carried out by automated methods such as by using a computer (or other apparatus/devices such as biomedical devices, laboratory instrumentation, or other apparatus/devices having a computer processor) programmed to carry out any of the methods described herein. For example, computer software (which can be interchangeably referred to herein as a computer program) can perform correlating the presence or absence of a biomarker in an individual with an altered (*e.g.*, increased or decreased) risk (or no altered risk) for cancer, *e.g.*, pancreatic cancer for the individual. Accordingly, certain embodiments of the disclosed subject matter provide a computer (or other apparatus/device) programmed to carry out any of the methods described herein.

30

The following Examples are offered to more fully illustrate the disclosure, but are not to be construed as limiting the scope thereof.

EXAMPLE 1: An iPS cell line from human pancreatic ductal adenocarcinoma undergoes early to invasive stages of pancreatic cancer progression.

MATERIALS AND METHODS

5 General cell culture

293T cells and human pancreatic ductal carcinoma cell lines PanC-1 and MIAPaCa-2 were maintained in 90% Dulbecco's modified essential medium (Invitrogen, Carlsbad, CA) supplemented with 10% FBS (Hyclone, Logan, UT) and fed every other day. Irradiated mouse embryonic fibroblast (MEF) cells were
 10 purchased from R&D systems (Minneapolis, MN) and maintained in 85% DMEM (Invitrogen) supplemented with 15% FBS (Hyclone) on 0.1% gelatin (Millipore, Billerica, MA) pre-treated tissue culture dishes. Plated irradiated MEFs were used within 5 days.

H1 huES/ human iPS culture

15 H1 huES (Thomson et al., 1998) and iPS-like clones were maintained in 80% Dulbecco's modified essential medium (DMEM) /F12 supplemented with 20% KNOCKOUT serum replacement, 0.1 mM nonessential amino acids (Invitrogen), 0.1 mM β -mercaptoethanol (Sigma, St. Louis, MO), and 10 ng/ml human basic fibroblast growth factor (bFGF) (Invitrogen). Human ES/iPS-like clones were grown
 20 onto irradiated MEFs in 0.1% gelatinized tissue culture dishes. Cells were fed every day and passaged once a week. For passaging, human ES cells were detached by treatment with 1 mg/ml collagenase IV (Invitrogen) for 3 min at 37 °C, centrifuged, resuspended with human ES media supplemented with 10 μ M Y27632 (Calbiochem, Darmstadt, Germany), and then seeded onto irradiated MEFs. The iPS-like lines were
 25 passaged mechanically with needles every 5-7 days. For RNA purification and differentiation, to remove MEFs, human ES and iPS-like lines were cultured on hES-qualified Matrigel (BD Bioscience, San Jose, CA) coated tissue culture dishes under mTeSR1 media (Stem Cell Technologies, BC, Canada).

Lentivirus production and titration

30 The mouse tet-Oct4, -Sox2, -Klf4, and C-MyC lentiviral vectors are donated from the Jaenisch lab (Brambrink et al., 2008). The pWPT rtTA vector was generated by ligating the rtTA2-M2 gene, isolated from pUHRt 62-1 vector (Urlinger et al., 2000), into the PWT-GFP backbone vector, with GFP removed. 293T cells were plated at a density of 8 x10⁵ cells per 100 mm dish. The next day, cells were

transfected with 2.5 µg vector (PWPT-rtTA or TetO-four factor), 1.7 µg psPAX2 packaging vector, and 0.8 µg PMDG envelope vector with 30 µl Eugene6 (Roche, Basel, Switzerland), according to supplier's instructions. Sixteen hours post-transfection, medium was removed and fresh media was added, and cells were further
 5 cultured for 60 h. Finally, the virus supernatant was collected, and centrifuged at 4° C for 10 min. The supernatant was filtered through 0.45 µm, aliquoted, directly tested for virus titer, or kept until use. Tet0-GFP lentivirus was concomitantly produced to monitor the virus titer. The titer of lentivirus was checked by flow cytometry and microscopy 2-3 days post-infection. Generally 5-7 MOI of each lentivirus was used
 10 for infection. The lentivirus average titer of PanC-1 and MIA PaCa-2 control PDAC cells were 3 x10⁸ infection units (IU) /ml and 3x10⁹ IU/ml, respectively. The lentivirus titer on HT1080 fibroblasts was 2.5 x10⁸ IU/ml.

Isolation and culture of human pancreas epithelial tumor cells

Human pancreatic ductal adenocarcinoma and pancreatic margin tissue
 15 was obtained by surgical dissection under the patient's informed consent by Fox Chase Cancer Center. Patients signed consent and sample acquisition was approved by the Institutional Review Board in accordance to institutional sample procurement procedures. Tissue specimens were obtained from the operating room immediately after resection. 1-2 cc of tissue were taken from the center of the cancer and 1-2 cc of
 20 tissue were taken from the margin furthest from the cancer and immediately placed into sterile F12 media/or Leibovitz's L-15 media (Invitrogen) supplemented with 100 U/ml Penicillin, 100 µg/ml Streptomycin, 10 µg/ml Gentamycin, 2.5 µg/ml Fungizone, 10 µg/ml Ciprofloxacin, 100 U/ml Nystatin (Invitrogen) until dissociation. Tissue was dissociated in 0.7 mg/ml liberase HI (Roche, Switzerland) as a supplier's protocol.
 25 Briefly, after rinsing twice with Hank's balanced buffered saline solution (HBSS, Invitrogen), tissue was transferred to liberase working solution (0.7 mg/ml liberase HI, DNase I 100 µg/ml, 25mM HEPES in HBSS) and minced with a scalpel. The minced tissue was transferred to a glass vial and incubated at 37° C (time varied depending on tissue size, maximum 1 hour). The liberase activity was then inhibited in quenching
 30 buffer (HBSS supplemented with 10% FBS) and passed through a 380 µm filter (Sigma) to remove tissue debris. The dissociated cells were rinsed with liberase quenching buffer followed by centrifugation. After washing twice with quenching buffer, dissociated cells were resuspended in completed defined KSFM (Invitrogen) supplemented with 5 ng/ml human EGF (BD biosciences) and 50 ng/ml cholera toxin

(Sigma), 50 µg/ml bovine pituitary extract (Invitrogen), and seeded onto 5 µg/cm² rat collagen I (BD biosciences) pre-coated tissue culture dishes, and cultured in a 37° C CO₂ incubator. Cells were fed every other day until infection. In cases of a long delivery time for the tissue, to remove autolyzed cells, minced tissue was digested in mild conditions (liberase HI 0.5 mg/ml approximately for 30 min) first and then passed through a 380 µm filter. In this condition, the dissociated single cells usually included autolyzed cells as well as blood cells but not many cancer cells. Therefore, to remove the autolyzed cells, dissociated cells were collected separately or discarded. Tissue trunk retained on top of 380 µm filter after 1st digestion was collected and digested with 0.7 mg/ml liberase HI working solution at 37° C for 30-40 min (depending on tissue size), quenched, washed and then cultured as described above.

Generation of iPS-like clones from human primary pancreatic ductal cancer and margin

epithelial cells

Cultured pancreatic cancer and margin epithelial cells were infected with Tet0-Oct4, Tet0-Sox2, TetO-Klf4, TetO-Myc viruses, along with the PWPT-rtTA virus, 3-4 days after plating. Twenty four hours later, the cells were re-infected with the same viruses. After 48 hours post-2nd infection, the infected cells were detached and counted to reveal that approximately 100,000 cells survived from each margin or tumor sample. The cells were resuspended in human ES media (80% DMEM/F12 supplemented with 20% Knockout Serum Replacer, 1 mM L-glutamine, 0.1 mM non-essential amino acid, 0.1 mM beta-mercapto ethanol (BME), 10 ng/ml basic Fgf (Invitrogen) and plated onto irradiated mouse embryonic fibroblasts on 0.1% gelatinized tissue culture dishes. Cells were fed every day. ES-like flat colonies were picked with 22 gauge needles from days 12 to 36 postsecondary infection, deposited onto irradiated MEFs, fed every day, and passaged mechanically with needles every 5-7 days. The colonies were frozen down around passage 3-4, and the stable clones were frozen down after passage 10 (See above for ES/iPS cell culture).

Pyrosequencing for KRAS codon 12 mutation

The *KRAS* codon12 mutation in primary tumor tissue was examined by pyrosequencing (Kanda et al., 2012). Twenty five nanograms of genomic DNAs isolated from paraffin slides were PCR amplified with PyroMark polymerase chain reaction kit (Qiagen) according to manufacturer's protocol. After amplification, 3 µl of reaction was loaded onto agarose gel to check PCR product. Ten microliters of

biotinylated PCR product were immobilized onto Streptavidin Sepharose HP beads (GE Healthcare Bio-Sciences) and annealed with sequencing primer designed with PyroMark Assay Design Software (Qiagen) (See Table 1 for primers) as described above. Pyrosequencing for *KRAS* codon 12 was done in a PyroMark MD (Qiagen).

- 5 All experiments were repeated three times with different batch of PCR products. Genomic DNAs from 10-12 margin iPSlike line and 10th margin primary epithelial cells cultured were used for biological negative control (2% base line-for GAT G12D, 0.5% base line for GGT G12V mutation). No DNA amplified PCR product and no primer annealed PCR product were used as technical negative controls.

10 RT-PCR of RNA

- For RT-PCR/qRT-PCR, H1 and iPS-like clones were cultured onto irradiated MEF cells. For CGH analysis, H1/H9 human ES cells and iPS like clones were cultured onto Matrigel under mTeSR media (Stem Cell Technologies) without mouse feeder layers. Total RNA was isolated by RNeasy Micro Kit (Qiagen, Valencia, CA), and 100 ng total RNA were reverse transcribed by using the iScript cDNA Synthesis kit (Bio rad, Hercules, CA). The cDNA (2 ng) was subjected to real time PCR on an iCycler (Bio-rad) with either SYBR green primers set or Taqman probes. For gene expression levels with Taqman probes, the delta Ct method was used with either *Gapdh* or *beta-actin* as a reference. The RT-PCR products of SYBR green primers were visualized on 2% agarose gels after staining with ethidium bromide to be sure of proper target amplification. The primer pairs are shown in Table 1.

Immunostaining/Immunohistochemistry (IH)

- For immunostaining on H1/iPS, cells were with fixed with 4% paraformaldehyde/PBS for 15 minutes at room temperature (RT) and permeabilized with 0.1% Triton X-100/PBS for 10 min at RT. After washing with PBS twice, a blocking solution (4% normal goat serum (Sigma) in PBS) was applied for 30 min at RT. Primary antibodies including SSEA4 (Millipore, 1:100), NANOG (Abcam, 1:100), and OCT4 (Abcam, 1:100) were applied on the fixed cells and incubated at 4° C for overnight. After rinsing, the Alexa 488-conjugated goat anti mouse IgG or goat anti rabbit IgG antibodies were applied for 1 hour at RT. DAPI was applied at the final wash to stain nuclei. For immunohistochemistry, paraffin sectioned slides were antigen retrieved by boiling in 10 mM citric acid buffer (pH 6) in a microwave oven for 15 min. Next, the endogenous peroxidase activity in tissue slides was quenched in hydrogen peroxide solution for 15 min at RT. Tissues were blocked with protein

blocker (Thermo Scientific) for 10 min, followed by avidin/biotin blocking (Vector lab, Burlingame, CA) for 15 min. Primary antibodies including K19 (Abcam, 1:2000), MUC5AC (Vector lab, 1:100), PDX-1 (Santa Cruz, 1:1000), SOX9 (Santa Cruz 1:500), HNF-4 α (Santa Cruz, 1:250-1:500), MF20 (Hybridoma Bank 1:40), Vimentin
 5 (Stemgent, Cambridge, MA, 1:500), Beta III tubulin (Abcam, 1:2000), GFAP (Cell Signaling, 1:200), diluted in PBS supplemented with 0.1% BSA and 0.2% Triton-X 100, were applied and incubated for 12-16 hours at 4° C. After washing twice, tissues were incubated with biotinylated anti-mouse IgG (Vector lab) at 37° C for 30 min. Tissue sections were conjugated with avidin-Horseradish peroxidase (HRP) by using
 10 VectaStain Elite ABC kit (vector lab) at 37° C for 30 min, followed by developing with DAB peroxidase substrate kit (Vector Lab) for peroxidase for 1-2 min. Developed tissue sections were stained with hematoxylin for nucleus, dehydrated, and mounted. For rt-TA (MoBiTec, German, 1:50) staining, fresh frozen tissue was embedded in OCT and sectioned 8 μ M thick. The sections were fixed with acetone for
 15 5 min, washed with PBS, and processed for IHC as described above. For immunostaining of NANOG (Abcam 1:1000) and OCT4 (Abcam 1:2000) on the 10th primary tumor tissue, frozen tissue was embedded in OCT and sectioned 8 μ M thick. For positive controls, H1 human ES cells were embedded in 4% agarose, snap frozen in OCT, and sectioned 10 μ M thick. All sections were fixed with 4% PFA and
 20 processed for IHC as above. For immunostaining of NANOG (Abcam 1:1000) and OCT4 (Abcam 1:2000) on 10th primary tumor tissue, frozen tissue was embedded OCT and sectioned 8 μ M thick. For positive control, H1 human ES cells were embedded onto 4% agarose, snap frozen in OCT, and sectioned 10 μ M thick. All sections were fixed with 4% PFA and processed for IHC as above.

25 **Teratoma assays**

Female 4-6 week old NOD-SCID-IL2Rgc null (NSG) mice (University of Pennsylvania, Xenograft core) (Shultz et al., 2005) were used for subcutaneous injection of human pancreatic iPS-like lines. Briefly, 24h before injection, doxycycline was withdrawn from the culture media. Confluent cells from all wells of
 30 a six well plate were detached by rinsing the cells with DMEM/F12, adding collagenase (1 mg/ml), incubating 3 min at 37° C, and collected by centrifugation. The cells were resuspended in 420 μ l of complete human ES media and injected subcutaneously in a female NSG mouse. In some cases, 300 μ l of the cell pellet was mixed with 120 μ l of Matrigel prior to injection; this had no effect on whether the 10-

22 cells generated PanINs, but it did cause the lesions to be less dispersed. Twelve weeks (3 months) or 273 days (9 months) after injection, the injection area was collected, fixed with 4% paraformaldehyde, and embedded in paraffin. Paraffin blocks were sectioned and stained with hematoxylin and eosin for analysis.

5 **Differentiation of H1/iPS-like clones into embryoid bodies**

Briefly, cells were harvested from one sub-confluent six well plate using Accutase solution (Innovative Cell Technologies, San Diego, CA) and rinsed with huES media. Clumps of cells were dissociated cells by passing through a 70 μ m filter. Dissociated cells were cultured onto ultra-low attachment 6 well plate (Corning, Oneonta, NY) under huES media without Fgf2 and KSR but supplemented with 20% FBS or aggrewell media (Stem Cell Technologies) for 2-3 weeks, then collected for RNA for qRT-PCR.

Comparative genomic hybridization

Total genomic DNA was isolated from cultured primary cancer and margin epithelial cells and cultures of iPS-like clones by proteinase K/phenol-chloroform. Genomic DNA was amplified using Agilent Oligonucleotide Array-Based CGH for Genomic DNA Analysis enzymatic labeling kit (Agilent, Santa Clara CA), according to the manufacturer. Labeled genomic DNA was cohybridized with human genome CGH Microarray kit 44K (Agilent) by the Penn Microarray facility. Arrays were scanned with an Agilent Scanner System. Data were analyzed by using Partek Genomics Suite (Partek, Saint Louis, MO).

CpG methylation sequencing

Total genomic DNA was purified from H1 huES cells, 10-12 margin iPS-like cells, and 10-22 cancer iPS-like cells by phenol extraction. Parental primary cancer genomic DNA was isolated from the tissue embedded into paraffin or OCT blocks. Bisulfite conversion with 1 μ g genomic DNA was carried out using CpGenome™ DNA Modification Kit (Millipore) as described by the manufacturer. *NANOG* and *OCT4* upstream regions were amplified with 30 ng converted DNA using the primers previously published or designed with PyroMark Assay Design Software (Qiagen, Valencia, CA) (See Table 1 for primers). Amplified PCR products were subjected to bisulfite pyrosequencing as described (Tost and Gut, 2007). Briefly, 10 μ l of PCR products amplified with biotin conjugated primers immobilized to Streptavidin Sepharose HP beads (GE Healthcare Bio-Sciences AB, Sweden) for 10 min. The immobilized PCR product was separated into single strand, released onto the

Pyromark plate containing sequencing primers in a PyroMark® Q96 Vacuum Workstation (Qiagen) and annealed with sequencing primer by cooling to room temperatures after incubation at 80° C. Pyrosequencing for corresponding CpG sites was done in a PSQ 96 instrument (Qiagen). The percentage methylation at each CpG site was determined using the Q-CpG methylation software (Qiagen).

Table 1. Primer list.

gene name	type	use	oligo sequence
Endogenous OCT4	Forward	pluripotency (RT-PCR)	GAC AGG GGG AGG GGA GGA GCT AGG
	Reverse	pluripotency (RT-PCR)	CTT CCC TCG AAC CAG TTG CCC CAA AC
Endogenous SOX2	Forward	pluripotency (RT-PCR)	GGG AAA TGG GAG GGG TGC AAA AGA GG
	Reverse	pluripotency (RT-PCR)	TTG CGT GAG TGT GGA TGG GAT TGG TG
Endogenous KLF4	Forward	pluripotency (RT-PCR)	TAT GAC CCA CAC TGC CAG AA
	Reverse	pluripotency (RT-PCR)	TGG GAA CTT GAC CAT GAT TG
Endogenous c-MYC	Forward	pluripotency (RT-PCR)	GCG TCC TGG GAA GGG AGA TCC GGA GC
	Reverse	pluripotency (RT-PCR)	TTG AGG GGC ATC GTC GCG GGA GGC TG
NANOG	Forward	pluripotency (RT-PCR)	CAG CCC CGA TTC TTC CAG CAG TCC C
	Reverse	pluripotency (RT-PCR)	CGG AAG ATT CCC AGT CGG GTT CAC C
REX-1/ZFP42	Forward	pluripotency (RT-PCR)	CAG ATC CTA AAC AGC TCG CAG AAT
	Reverse	pluripotency (RT-PCR)	GCG TAC GCA AAT TAA AGT CCA GA
TERT	Forward	pluripotency (RT-PCR)	TGT GCA CCA ACA TGT ACA AG
	Reverse	pluripotency (RT-PCR)	GCG TTC TTG GCT TTC AG GAT
GDF3	Forward	pluripotency (RT-PCR)	AAATGTTTGTGTTGCGGTCA
	Reverse	pluripotency (RT-PCR)	TCTGGCACAGGTGTCTTCAG
GAPDH	Forward	housekeeping (RT-PCR)	TGT TGC CAT CAA TGA CCC CTT
	Reverse	housekeeping (RT-PCR)	CTC CAC GAC GTA CTC AGC G
ribosomal protein L19 (RPL19)	Forward	housekeeping (RT-PCR)	CGA ATG CCA GAG AAG GTC AC
	Reverse	housekeeping (RT-PCR)	CCA TGA GAA TCC GCT TGT TT
NCAM	Forward	differentiation; ectoderm (RT-PCR)	ATG GAA ACT CTA TTA AAG TGA ACC TG
	Reverse	differentiation; ectoderm (RT-PCR)	TAG ACC TCA TAC TCA GCA TTC CAG T
keratin associated protein 5-9 (KRTAP5-9)	Forward	differentiation; ectoderm (RT-PCR)	CAG ATG CTG TGT CCC TGT GT
	Reverse	differentiation; ectoderm (RT-PCR)	TTC AGA TCC AGA AGG GGA TG
PAX6	Forward	differentiation; ecto/endo (RT-PCR)	CCA GAA AGG ATG CCT CAT AAA GG
	Reverse	differentiation; ecto/endo (RT-PCR)	TCT GCG CGC CCC TAG TTA
SOX1	Forward	differentiation; ecto/endo (RT-PCR)	ATG CAC CGC TAC GAC ATG G
	Reverse	differentiation; ecto/endo (RT-PCR)	CTC ATG TAG CCC TGC GAG TTG
RUNX1	Forward	differentiation; mesoderm (RT-PCR)	CCC TAG GCG ATG TTC CAG AT
	Reverse	differentiation; mesoderm (RT-PCR)	TGA AGC TTT TCC CTC TTC CA
T, brachyury homolog	Forward	differentiation; mesoderm (RT-PCR)	A CCC AGT TCA TAG CCG GTG AC
	Reverse	differentiation; mesoderm (RT-PCR)	CCATTGGGAGTACCCAGGTT
Flk1	Forward	differentiation; mesoderm (RT-PCR)	AAGGTGACAGGAAAGAGCAACT
	Reverse	differentiation; mesoderm (RT-PCR)	TCCCTCCATTGGCCCGCTTAAC
Cartilage (COMP)	Forward	differentiation; mesoderm (RT-PCR)	CCG GAG CCA GGA CTA CAT TA
	Reverse	differentiation; mesoderm (RT-PCR)	GGT CTT GAA GTC AGC CGT GT
AFP	Forward	differentiation; endoderm (RT-PCR)	AGC TTG GTG GTG GAT GAA AC
	Reverse	differentiation; endoderm (RT-PCR)	CCC TCT TCA GC AAA GCA GAC
amylase, alpha 1C (salivary) (AMY1C)	Forward	differentiation; endoderm (RT-PCR)	CAT CTG TTT GAA TGG CGA TG
	Reverse	differentiation; endoderm (RT-PCR)	TTC CCA CCA AGG TCT GAA AG
CXCR4	Forward	differentiation; endoderm (RT-PCR)	CAC CGC ATC TGG AGA ACC A
	Reverse	differentiation; endoderm (RT-PCR)	GCG CAT TTC CTC GGT GTA GTT
FOXA2	Tagman	differentiation; endoderm (RT-PCR)	Applied Biosystems HS00232764_m1
SOX17	Tagman	differentiation; endoderm (RT-PCR)	Applied Biosystems HS00751752_s1
PDX1	Tagman	differentiation; endoderm (RT-PCR)	Applied Biosystems HS00426216_m1
ACTB	Tagman	housekeeping (RT-PCR)	
h-TA	Forward	subcloning	AAAGGATCCACCATGTCTAGA CTG GAC
	Reverse	subcloning	AAAGTCGACITACCGGGGAGCATGTG
GFP	Forward	subcloning	AAAGAAATTCGCCACCAATGGTGAACAAAGGC
	Reverse	subcloning	AAAGAAATTCCTAGCTACTAGCTAGTGG

Table 1, continued.

<i>KRAS</i> codon 12	Forward Reverse Reverse sequencing	gDNA PCR and pyrosequencing gDNA PCR biotinylated for pyrosequencing pyrosequencing	TTTGAGAGCCTTTAGCGGCGCG CGTCAAGGCACTCTTGCTACGC CGTCAAGGCACTCTTGCTACGC AAACTTGTGGTAGTTGGA
<i>KRAS</i> codon 61	Forward Reverse	gDNA PCR gDNA PCR	TCAAGTCTTTTGCGGATT TGCATGGCAATGCAAGAC
<i>MTS1(CDK2NA/p16^{INK4a})</i> exon1	Forward Reverse	gDNA PCR gDNA PCR	GAA GAA AGA GGA GGGGCTG GCG CTA CCT GAT TCC AAT TC
<i>MTS1(CDK2NA/p16^{INK4a})</i> exon 2	Forward Reverse	gDNA PCR gDNA PCR	ACA CAA GCT TCC TTT CCG TC TCT GAG CTT TGG AAG CTC TC
<i>BRAF</i> exon11	Forward Reverse	gDNA PCR gDNA PCR	TCCCTCTCAGGCATAAGGTAA CGAACAGTGAATATTCTTTGAT
<i>BRAF</i> exon15	Forward Reverse	gDNA PCR gDNA PCR	TCATAATGCTTGGCTGTGATAGGA GGCCAAAATTTAATCAGTGA
<i>NANOG</i> promoter site #1	Forward Reverse Sequencing	Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing	AGAGATAGGAGGGTAAATTTTTT CCAAAACAACCTAATTTACTCCACAA TGTTGAAGTATGATGTTATTA
<i>NANOG</i> promoter site #2	Forward Reverse Sequencing 1 Sequencing 2 Sequencing 3	Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing	GAGTTAAAGAGTTTGTTTTAAAAATTAT TCCAAATCTAATAATTTATCATATCTTC CATATCTTCAAACTACTCC ACACTAACATTTATAAACAATACTC CTTCTCTTTTCAATCT
<i>NANOG</i> promoter site #3	Forward Reverse Sequencing	Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing	AATTTATTGGGATTATAGGGG AACAAACAAAACCTAAACAAAC ACAAACCAACAACAAT
<i>NANOG</i> promoter #4	Forward Reverse Sequencing	Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing	AAGTTGGAAAAGTGGTGAATTTAGAAGTA AACAACTAAATCCACATCATATT AAAAGTGGTGAATTTAGAAGTAT
<i>OCT4</i> promoter site #3	Forward Reverse Sequencing 1 Sequencing 2	Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing	GTGTTTTTTGGGTTAAAGGT CCAATATCTTCATCTTAATAACATCC GTGAATATAGTTGTAATTTATTGT GATTTGTGGTAGGTATTG
<i>OCT4</i> promoter #7	Forward Reverse Sequencing 1 Sequencing 2	Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing	TAGTTGGGATGTGTAGAGTTTGAGA AAACCAAAACAATCCTTCTACT ATGTTGTAGAGTTTGAGAGA AAAGTTGGGTGTGGT
<i>OCT4</i> promoter site #8	Forward Reverse Sequencing	Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing	AAGTTTTGTGGGGGATTTGTAT CCACCACTAACCTTAACCTCTA TGGGGGATTTGTATTGA
<i>OCT4</i> promoter site #9	Forward Reverse Sequencing 1 Sequencing 2	Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing Bisulfite pyrosequencing	GTAGAGGTTAAGGTTAGTGGGTG AAACCTTAAAACTTAACCAATCC CATCACTCCACAC ATCCCAAAATCAACCAAC

In vitro culture of 10-22 cancer iPS-like line teratomas

- 5 After thirteen to fourteen weeks, teratomas were harvested from NSG mice harboring 10-22 cells. Dissected teratoma tissue was minced and dissociated in liberase T- flex (1.3W/ml) at 37 °C for 30 min. Subsequently, the reaction was quenched and washed with 15% FBS/DMEM. Dissociated cells were plated as either whole cell culture or single cell, passed through 40 µm filter. Dissociated teratoma
- 10 tissue was embedded as described previously (Lee et al., 2007). Briefly, prechilled 4 well plates (Nunc, Rochester, NY) were coated with a thin layer of 150 µl Matrigel (BD Biosciences) for 15 min at 37 °C incubator. Dissociated cells were pelleted by centrifugation at 1200 rpm for 5 min at 4 °C, resuspended into 200 µl of Matrigel, and incubated at 37 °C for 30 min to allow the cell-Matrigel complex to gel. Explants

were fed with 500 μ l of serum-free culture media. Six days postplating, serum free-conditioned media was collected, centrifuged to remove cell clumps, supernatant was collected, and kept at -80 °C for proteomic analysis. Explants were fed every 4-5 days for further culturing. To prepare a negative control for proteomic analysis, 10-22 (p10
 5 and p27) iPS-like lines collected mechanically using stem cell passaging tool (Invitrogen) and cultured onto a plate pre-coated with 5 μ g/cm² rat collagen in serum free DMEM media for 2 days. Media was collected, filtered through 0.45 μ m filter, and kept at -80 °C for proteomic analysis.

Proteomic analysis

10 Frozen media was divided into three tubes; each contained proteins corresponding to 30-60 μ g protein. Each sample was precipitated with acetone to concentrate proteins and remove salts and lipid soluble contaminants. Briefly, four volumes of chilled acetone were added to a sample and incubated sample at -20° C for overnight. The samples were pelleted at 16,000 g for 10 min at 4° C followed by
 15 washing with a solution of acetone and water (4:1). The pellet was air-dried for 10 min and denaturized by boiling in NuPage LDS sample buffer including reducing agent (Invitrogen) for 5 min. Each sample was subjected into Nupage 10% Bis-Tris Gel and run at 100 mV with Nupage MOPS SDS Running buffer (Invitrogen) for almost 2 hours. The gel was stained with Simplyblue safestain (Invitrogen) according
 20 to manufacturer's recommendation for MS analysis. The 5x5 mm pieces were excised and stored in 2% acetic acid solution. The excised gel samples were digested with trypsin (Strader et al., 2006). 5 μ l trypsin digested samples were injected with autosampler (Eksigent technologies, Dublin, CA) and a 10 cm C18 column was used to separate the digested peptides. Nano LC (Eksigent) was run at 200 nl/min flow rate
 25 for 100 min gradient. Online nanospray was used to spray the separated peptides into LTQ (Thermo Electron) and the raw data was acquired with Xcalibur. Sequest software was used to search database Uniprot and IPI and generated srf files for each sample. Scaffold 3.3 was used to combine and analyze the Sequest generated srf files quantitatively based on spectrum count. Cutoffs was peptide p-value >95% and
 30 protein p-value > 99%. To further sort human proteins out of the mouse and bovine, the peptide sequences were blasted using UniProKB TrEMBL (downloaded on Oct, 2011) allowing no mismatch. The human peptides and human and mouse common peptides were used to identify proteins secreted from the teratoma and 10-22 iPS-like line. To distinguish the proteins that have peptides only common in both human and

mouse from the protein derived from mouse background, all proteins were subtracted with the proteins secreted from contralateral control. A hierarchical clustering was applied using MeV (using Pearson Correlation as a metric) (data not shown). Proteins secreted from at least two teratoma explants were used for Ingenuity Pathway Analysis (<http://www.ingenuity.com>).

PanIN and PDAC mouse model

Pdx-Cre; LSL-KrasG12D; p53fl/+; RosaLSL-YFP mice (Rhim et al., 2012) aged 2-2.5 and 4-6 months were analyzed for HNF4 α expression at the PanIN and tumor stages, respectively. After euthanasia, pancreases were washed and incubated in Zinc-formalin (Polysciences, Warrington, PA) at 4° C for two hours. After washes and dehydration with ethanol, tissues were paraffin embedded and sectioned. After deparaffinization with xylene, tissue sections were antigen retrieved in R-buffer A using a steam retriever (both from Electron Microscopy Sciences, Hatfield, PA). Slides were blocked in 5% donkey serum in 0.1% Triton-X100 in PBS for one hour then incubated with primary antibodies (Chicken α GFP (Abcam, 1:500) and goat α HNF4 α (1:200) for one hour at RT. Slides were then incubated with secondary antibodies (FITC conjugated donkey α chicken (1:200; Life Technologies, Grand Island, NY); Alexa Fluor 594 conjugated donkey α goat (1:200; Jackson ImmunoResearch, West Grove, PA); DAPI (1:1000)) in block solution for 1h at RT, followed by washes and mounting of coverslips. Images were visualized using an Olympus IX71 inverted fluorescence microscope and captured and merged using Olympus DP Manager software (v.3.1.1).

INTRODUCTION

Cancer phenotypes can be suppressed in certain medulloblastoma cells, RAS-induced melanoma cells, and embryonal carcinoma cells and renal tumor cells when they are reprogrammed to pluripotency by nuclear transfer (Blelloch et al., 2004; Hochedlinger et al., 2004; Li et al., 2003; McKinnell et al., 1969). The resultant pluripotent cells can then differentiate into multiple early developmental cell types of the embryo. Such embryos die partly through organogenesis, presumably due to re-expression of the cancer phenotype. It is remarkable that, in certain circumstances, the pluripotency network can suppress the cancer phenotype sufficiently to allow early tissue differentiation. Using iPS cell technology (Takahashi and Yamanaka, 2006), cancer cell lines have been made into iPS cells (Carette et al., 2010; Miyoshi et al.,

2010). However, no iPS cell lines from solid primary human cancers have been reported. Not limited to one theory, creating iPS cells from an epithelial tumor would allow the cells to be propagated indefinitely in the pluripotent state and that, upon differentiation, a subset of the cells would undergo early developmental stages of the human cancer, providing a live cell human model of early stages of the disease.

RESULTS

Creating iPS-like cell lines from human pancreatic ductal adenocarcinoma.

Human pancreatic ductal adenocarcinoma samples were obtained immediately after resection (Table 2). Histologically normal pancreatic tissues at the margin of the specimens were used as controls. Epithelial cells were isolated and cultured in serum-free medium with cholera toxin to impair the growth of fibroblasts. Two successive infections of the pancreatic cancer and margin cells were performed with 5 lentiviruses separately encoding doxycycline-inducible mouse *Oct4*, *Sox2*, *Klf4*, and *c-Myc*, and the rt-TA transactivator, while genomic DNA was isolated from the pancreatic specimen margin and cancer epithelial cells that had been cultured separately. Pyrosequencing analysis revealed *KRAS* mutant cells in 7 of 9 of the initial tumor samples (Tables 2, 3). ES-like colonies began to arise after 7 days, with fewer colonies arising from cancer epithelial cells. ES-like clones from cancer or margins started to differentiate and disappear 4 days after withdrawal of doxycycline. Thus, colonies were maintained in the presence of a low doxycycline concentration (50 ng/ml) and called the resulting lines “iPS-like” because of their dependency upon doxycycline. iPS-like lines were established from 4 of the 9 tumor epithelial specimens and corresponding margin iPS-like lines from 3 of the specimens (Figure 1A, Table 2).

The cells' pluripotency were characterized by RT-PCR, immunostaining, embryoid body formation, teratoma assays, karyotyping, and a subset by CpG methylation analysis. Most iPS-like lines expressed endogenous pluripotency marker RNAs and protein (Figure 1B, C; Figure 7A, B, C and D). Teratoma assays in immunodeficient NSG mice and embryoid body assays revealed that the 10-12 pancreatic margin and 10-22 cancer iPS-like lines from the 10th patient (Figure 1D, E; Figure 8A) and the 14-24 pancreatic margin and 14-27 cancer iPS-like clones from the 14th patient (Figure 7E, F) could generate tissues of multiple germ layers, demonstrating pluripotency. The original tumor #10 was negative for NANOG

expression and exhibited sporadic expression of POU5F1/OCT4 (Figure 8B). Detailed pyrosequencing analysis of CpG methylation showed that the 10-22 iPS-like line exhibited demethylation at 9 of 9 sites at the NANOG promoter, compared to the primary tumor, and at 3 of 6 sites at the OCT4 promoter; by comparison, the H1 huES
 5 line was similarly demethylated at NANOG and demethylated at 4 of 6 sites at POU5F1 (Figure 8C). Thus, the iPS-like lines exhibit diverse characteristics of reprogrammed pluripotent cells.

Table 2. Characteristics of pancreatic cancer tissues used for this example.

patient #	type of cancer (diagnosis at the time operation)	location	greatest dimension of tumor	TNM stage	adjuvant treatment	KRAS codon 12 allele frequencies by pyrosequencing from paraffin slide	iPS results	established lines
10	PDAC recurrent invasive poorly differentiated, chronic pancreatitis	from head to tail	9.5 cm	T2NDMx	chemo. & rad. (1st cancer), then chemo (recurrent cancer)	G12D (22%)	10-12, 10-22	3
13	PDAC, moderately differentiated	head & neck	unknown	T3N1Mx	no	wild (G)	margin cells died (note 2), one iPS-like line from cancer (13-D3)	1
14	PDAC, moderately differentiated	head & neck	multiple foci (largest foci, 1 cm)	T1N1Mx	no	wild (G)	14-24, 14-27	5
16	PDAC	head & neck	2 cm	T3NDMx	chemo & rad.	G12V (3%)	no colonies	note 1
18	PDAC, poorly differentiated, PanIN3	head & neck	4.5 cm	T3NDMx (grad II)	chemo. & rad.	G12D (31%)	margin primary cells died (note 2), no colonies from cancer	note 1/2
19	PDAC, invasive	head & neck	4.5 cm	T3N1Mx (grad II)	no	G12D (14.4%)	characterized (19-D1, 19-17, 19-24, 19-32, 19-35)	5
20	PDAC, moderately differentiated	head & neck	unknown	T3N1Mx	chemo. & rad.	G12D (6.9%)	no colonies	note 1
21	PDAC	body	3.7 cm	T3NDMx	chemo. & rad.	G12A (5.4%) G12D (2.8%)	no colonies	note 1
22	pancreatic adenosquamous carcinoma, invasive, poorly differentiated	body	7.5 cm	T3NDMx	no	G12V (15.2%)	margin primary cells died (note 2), colonies from cancer were not expanded	0

10 ¹ No colonies were obtained from tissue that received irradiation or radiation plus chemotherapy.

² Margin tissue autolysis due to lengthy time for transfer to lab or tumor tissue autolysis derived from radiation/chemotherapy-treated patients.

All iPS-like lines were screened for mutations in *KRAS*, *CDKN2A*,
 5 and *BRAF*, which are common genetic alterations in pancreatic cancer (Moskaluk et al., 1997) (Table 3). The 10-22 iPS-like line, derived from the recurrent, invasive, and poorly differentiated PDAC of the 10th patient, harbors the same *KRAS* G12D mutation seen in the initial tumor epithelial population (Figure 1F; Table 2, 3). 10-22 cells also possess a *CDKN2A* heterozygous deletion (Figure 1H; Table 3) and a
 10 comparative genomic hybridization (CGH) pattern with numerous chromosomal aberrations, more exaggerated than that of the primary cancer culture, and a correspondingly aberrant karyotype (Figure 1I; Figure 3). The exaggeration of the primary cancer CGH pattern in the 10-22 iPS-like line is expected because the primary cancer culture contained some stromal cells and thus was contaminated by
 15 cells of a normal genotype. Specifically, 23 gross chromosomal aberrations were detected in the PDAC epithelial cell population of the 10th tumor and 20 were represented in the 10-22 line (Fig. 1I, Fig. 9B). Decreased CGH signals in the primary tumor cells and the 10-22 line spanned *PTEN* and *DPC4* (*SMAD4*), consistent with observations of allelic loss of these loci in human pancreatic cancers (Hahn et al.,
 20 1996; Hahn et al., 1995). By contrast, the isogenic 10-12 pancreatic margin iPS-like line had wild type *KRAS*, a flat CGH chromosomal profile, similar to the parental margin cell culture, and a normal karyotype (Figure 1G, H, I; Figure 9; Table 3). The 10-12 and 10-22 iPS-like lines were derived from pancreatic margin and cancer epithelial cells, respectively, with the 10-22 line harboring the marked genomic
 25 rearrangements seen in the initial advanced tumor epithelial population. The iPS-like lines from the 14th patient (Figure 7A, B), with moderately differentiated PDAC containing scattered 1 mm foci, showed point mutations in *CDKN2A* and *BRAF* but a wild type *KRAS*, as with the parental epithelial culture (Tables 2, 3), and flat CGH profiles (data not shown). A pair of margin and cancer derived iPS-like lines from the
 30 19th patient (Figure 7C, D) also had somatic mutations in *CDKN2A*, with wild type *KRAS* and *BRAF* (Table 3). Because the 10-22 cancer iPS-like line from the 10th patient contained the *KRAS* mutant background, typical of PDAC, a detailed study was performed of that line and its isogenic 10-12, *KRAS* wt margin iPS-like clone.

Table 3. Characteristics of iPS-like lines from human pancreatic cancer and margin samples.

patient #	label	source	karyotype (aver.)	CGH	point mutation					
					KRAS		CDKN2A		BRAF	
					codon 12 (GGT; gly)	codon 61 (CAA; gin)	exon 1	exon 2	exon 11	exon 15
10	10-12	margin iPS-like line	46	normal	WT	WT	no mutation observed	no mutation observed	WT	WT
	10-22	cancer iPS-like line	69	dislocated	hetero. G12D	WT	hetero. mutation	deletion	WT	WT
14	primary tissue/cells	margin tissue	ND	ND	ND	ND	hetero. 5' UTR-49 C>T	no mutation observed	ND	ND
		PanIN2/3 tissue	ND	ND	WT	ND	hetero. 5' UTR-49 C>T	no mutation observed	IVS10 - 18T>G	ND
		PDAC tissue	ND	ND	WT	ND	hetero. 5' UTR-49 C>T	no mutation observed	IVS10 - 18T>G	ND
		primary cancer epithelial culture	ND	normal	WT	WT	hetero. 5' UTR-49 C>T	no mutation observed	IVS10 - 18T>G	no mutation observed
	14-24	margin iPS-like line	45	X-chromosome depletion	WT	WT	hetero. 5' UTR-49 C>T	no mutation observed	IVS10 - 18T>G	no mutation observed
	14-27	cancer iPS-like line	44	normal	WT	WT	hetero. 5' UTR-49 C>T	no mutation observed	IVS10 - 18T>G	no mutation observed
19	primary tissue	margin	ND	ND	ND	ND	ND	ND	no mutation observed	no mutation observed
		PDAC	ND	ND	WT	ND	ND	ND	no mutation observed	no mutation observed
	19-01	margin iPS-like line	45	ND	WT	WT	Aa28Aa IVS1-89A>G, IVS1-31A>T	Aa139Thr	no mutation observed	no mutation observed
	19-17	margin iPS-like line	ND	ND	WT	WT	ND	Gly127Trp, nonsense mutation (Aa 132 stop codon)	no mutation observed	no mutation observed
	19-24	margin iPS-like line	ND	ND	WT	ND	ND	no mutation observed	ND	no mutation observed
	19-35	cancer iPS-like line	46	ND	WT	ND	ND	Gly127Trp, nonsense mutation (Aa 132 stop codon)	no mutation observed	no mutation observed

10-22 iPS-like cells from advanced human PDAC, upon differentiation, undergo the early stages of pancreatic cancer.

Despite being derived from cancer, the 10-22 cancer iPS-like line as well as its companion 10-12 margin line gave rise to much smaller subcutaneous teratomas, at three months, compared to those seen from the control H1 human ES cells. More striking, there was a much higher proportion of endodermal, DBA lectin-positive ductal structures in teratomas from the 10-12 and 10-22 iPS-like lines (Figures 2A; Figure 10A), whereas H1 cells generated mostly neuronal lineages, as seen previously (Bock et al., 2011). These results are in accordance with reports on

the propensity of other iPS cell lines to differentiate into the lineage from which they were derived (Bar-Nur et al., 2011; Kim et al., 2011).

The endodermal teratomas arising from the margin (10-12) and tumor (10-22) iPS lines with the original tumor of the 10th patient were compared. The original tumor exhibited many areas of poorly differentiated foci and infiltration, although occasionally showed a more organized epithelium, but no PanINs (Figure 2B, C; Figure 10B). The tumor epithelium exhibited irregularly shaped and hyperchromatic nuclei and a high nuclear to cytoplasmic ratio (Figure 2C, arrow), along with cytoplasmic protrusions indicative of an invasive phenotype (dotted line in Figure 2C). By contrast, teratoma ductal tissues at 3 months from the 10-22 cancer iPS-like line had a high level of architectural organization, with abundant gland formation and a more differentiated cytology compared to the primary tumor (Figure 2B). This included an increased cytoplasm and smaller, less hyperchromatic nuclei with even nuclear membrane contours, a decreased nuclear to cytoplasmic ratio, and increased cell polarity, indicative of PanIN stages. Abundant mucin was present in the apical aspect of the cells (arrows in Figure 2D). There was no evidence of acinar differentiation or neuroendocrine neoplasia. PanINs are graded by the extent of dysmorphic structures compared to normal ducts (Hruban et al., 2001) and a range of PanIN1-, PanIN2-, and PanIN3-like structures was observed in the histology of the endodermal teratomas from the 10-22 iPS-like line at 3 months; though predominantly structures resembling the PanIN2 and PanIN3 stages (Figure 10C-G). Not limited to a particular theory, these findings suggested that the 10-22 cancer iPS-like line generates ductal structures that resemble PanIN (Maitra and Hruban, 2008).

PanIN-like teratomas formed in 9 of 10 teratoma experiments at 3 months, regardless of the passage number of the 10-22 cells. By contrast, the 10-12 pancreatic margin iPS-like line did not generate PanIN-like structures in teratomas (n=4) (Figure 2B). Also, iPS-like lines from the 14th and 19th tumors, containing predisposing mutations, but not of *KRAS*, did not generate PanIN-like lesions (n=5, data not shown), and therefore were not studied further. PCR of DNA obtained by laser capture microdissection showed that the stromal cells surrounding the ductal epithelium in the PanIN-like teratomas of the 10-22 line contained the rt-TA lentiviral DNA, and thus were derived from the starting 10-22 cells (Figure 10H). This was confirmed by detecting rT-TA expression in the PanIN-like epithelium as well as throughout the local stroma, but not in the distal stromal portions of subcutaneous

tissue (Figure 10I). Recently, stromal-like cells surrounding PDAC have been found to be derived by EMT from the pancreatic cancer epithelium in a mouse model (Rhim et al., 2012). However, the study observed very few EMT-derived stromal cells at the PanIN stages. Taken together, it seems that most but perhaps not all of the human
 5 stromal cells in the PanIN regions of the teratomas were derived by local co-differentiation of the pluripotent 10-22 cells into epithelial and mesenchymal tissues.

Characterizing the ductal structures in teratomas of the 10-22 cancer iPS-like line in further detail, they were found to express keratin 19 (K19) as well as nuclear PDX1, the pancreatic determination factor, and nuclear SOX9 (Figures 3A-F,J; Figure 11A-E). PDX1 is expressed at very low levels in adult pancreatic duct
 10 cells, it is not expressed in adult exocrine cells, and it is up-regulated in pancreatitis, PanIN, and PDAC (Miyatsuka et al., 2006). SOX9 is a coordinate effector, with mutant *Kras*, of precursor pancreatic lesions in a mouse model (Kopp et al., 2012). Teratomas from the 10-12 pancreatic margin iPS-like cells did not express the gastric
 15 mucin MUC5AC, whereas the PanIN-like structures of the teratomas from the 10-22 cells expressed abundant MUC5AC (Figures 3G-I, K, brown staining; Figure 11A), as observed in PanIN lesions and well differentiated PDAC (Kim et al., 2002). Considering the histology and CGH genomic profile of the parental pancreatic cancer tissue from which the 10-22 iPS-like line was derived, the data indicate that the 10-22
 20 iPS-like line from poorly differentiated, late stage PDAC can differentiate into PanIN lesions associated with the early stage of the disease.

The 10-22 iPS-like line PDAC progresses to the invasive stage of human pancreatic cancer.

To assess whether PanIN-like structures from the 10-22 line could
 25 progress to later stages of PDAC, teratomas grown for 6-9 months in NSG mice were investigated. By 9 months, two solid, palpable tumors arose in each of two injected mice (3-6 mm diameter), all with a genotype characteristic of 10-22 cells (Figure 11F). While palpable tumors were not evident at 6 months, histological analysis showed highly glandular structures with nuclear heterotypia and hypochromia at both 6 and 9
 30 months (Figure 4A, B). The epithelial cells were positive for K19, MUC5AC, PDX1, and SOX9 (Figure 4C-R, arrows). Notably, structures indicative of a locally invasive phenotype were observed (Figure 4A, B; arrows). The PanIN-like stage of teratomas from the 10-22 iPS-like line is succeeded in vivo by the invasive stage of PDAC,

indicating that the 10-22 iPS-like model undergoes a spectrum of pancreatic carcinogenesis.

**Cultured organoids of PanIN-like cells from 10-22 line teratomas secrete or
5 release proteins indicative of early stage pancreatic cancer.**

A system where the PanIN-like structures occurring within teratomas from the 10-22 cells could be studied as a live, in vitro model of early stage human pancreatic cancer was generated. Accordingly, tissues from teratomas 3 months after injection, along with contralateral control tissue were harvested, and conditions were
10 established where the tissues were embedded separately into Matrigel and cultured in vitro (Figure 12A). PCR analysis of DNA from the resulting sphere-like organoids confirmed the 10-22 line genotype, which was absent from contralateral control explants (Figure 12B). The organoids from the 10-22 cancer iPS-like cells retained the expression of human K19 and MUC5AC (Figure 5A-C, sections).

Given the poor prognosis of PDAC, the organoid system was used to identify biomarkers and pathways that could permit early detection and facilitate disease monitoring after therapy. NanoLC/MS/MS was used to examine the proteins that were secreted or released from explants of three independent teratomas that were cultured for 6 days (Figure 5D, left). The data were compared to proteins secreted or
20 released from explants of contralateral control tissue cultured similarly and from the parent 10-22 iPS-like line cultured in the undifferentiated state. Peptides with perfect matches to human peptides and that were specific to the human PanIN-like teratoma explants were selected (Figure 5D, right). Of the 25 proteins that were secreted or released from all three 10-22 teratoma explants (Figure 5D, right; Table 4), 8 were
25 previously reported to be expressed at the RNA or protein levels within PanIN, IPMN, and/or PDAC, whereas the remaining proteins have not been reported (Harsha et al., 2009) (Table 5). Of the 82 proteins that were identified from the intersection of pairs of 10-22 teratoma explants (Figure 5D, right; Tables 6-8), 15 proteins have been previously reported in PanIN/IPMN/PDAC (Harsha et al., 2009) (Table 5). While
30 certain mRNAs were previously identified in PanIN, such as PGAM-M and VWF (Buchholz et al., 2005), the corresponding proteins were determined to be secreted or released from cells and remained stable in the medium.

Of the total 107 proteins secreted or released from at least two 10-22 teratoma explants (Figure 5), Ingenuity pathway analysis revealed that 42 fall into

interconnected TGF β 1 and Integrin signaling networks (Figure 5E, for 38 proteins, and Table 13). These pathways have been previously reported in PanIN, IPMN, and PDAC (Bardeesy et al., 2006; Jones et al., 2008). Additionally, 7 proteins that were secreted or released from the teratoma explants were identified as falling into the interconnected Ras/p53/JUN/CTNB1 signaling pathway (Table 12). In summary, numerous proteins were discovered that are secreted or released from the 10-22 live cell model of early stage human pancreatic cancer, as well as evidence of well-documented pathways involved in early cancer progression.

10 Table 4. Protein secreted from the three tetratomas of 10-22 iPS-like cells.

	ID	Description	gene	iPS									Teratoma								
				iPS#1			iPS#2			iPS#3			9223 (#2)			9225 (#3)			7761 (#1)		
				H	M	C	H	M	C	H	M	C	H	M	C	H	M	C	H	M	C
Proteins identified by human specific peptides (not mouse peptides)	P00022443	Alpha-fetoprotein	AFP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P00009305	Glucosamine-6-phosphate isomerase 1	GNPDA1	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0
	P00000980	Isoform 1 of Mitochondrial inner membrane protein	IMMT	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	P00028933	Zinc finger protein 160	ZNF160	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	P00004671	Golgin subfamily B member 1	GOLGB1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	P00045947	Isoform 1 of Rotatin	RTTN	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	P00028762	Isoform 1 of ATP-binding cassette subfamily A member 13	ABCA13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	P00028940	Isoform 1 of Obscurn	OBSCN	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	P000251101	Putative uncharacterized protein KIAA1109	KIAA1109	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	P00046594	Desmin	DES	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	P00021715	Isoform 1 of Collagen alpha-5(V) chain	COL4A5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0
	P00032259	Complement C4-A	C4A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
	P00024911	Endoplasmic reticulum resident protein 29	ERP29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
	P00007990	Isoform 1 of Periostin	POSTN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0
	P00059322	Isoform 1 of Transcription factor 20	TCF20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Proteins identified by peptides that are matched in both human and mouse (not specific human)	P000218474	Isoform 1 of Beta-enolase	ENO3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P000218282	Isoform Short of Ubiquitin fusion-degradation protein1 homolog	UFDY1L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0
	P00019864	Alpha-actinin-2	ACTN2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0
	P00013405	Stromelysin-2	MMP10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
	P00002901	Transcriptional enhancer factor TEF-1	TEAD1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
	P000288984	Talin-1	TLN1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
	P000171603	Isoform 1 of Heterogeneous nuclear ribonucleoprotein M	HNRNPM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
	P00021327	Isoform 1 of Growth factor receptor-bound protein 2	GRB2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	P00026266	Small nuclear ribonucleoprotein E	SNRPE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	P000171199	Isoform 2 of Proteasome subunit alpha type-3	PSMA3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0

Numbers highlighted indicate the number of peptides providing a protein identification only in conditioned medium from explants cultured in serum-free medium from three independent tetratomas of 10-22 iPS-like cells (mouse #9223,

9225 and 7761), and not in media from explants of contralateral control tissue or in media from the 10-22 iPS-like line cultured under pluripotency conditions. Total is 25 proteins. H-human specific peptide, M-mouse specific peptide, C-human and mouse peptide, B-bovine specific, U-species other than human/mouse/bovine.

5

HNF4 α network activated in early to intermediate, but not late stage pancreatic cancer. Further Ingenuity analysis revealed that another 25 of the secreted or released proteins were in a network observed to be centered on the transcription factor HNF4 α , including numerous direct gene targets (Odom et al., 2004) (Figure 5F, Table 9 and 10). Nine of the proteins in the HNF4 α network were also secreted into the plasma during the PanIN stage of a mouse PDAC model (Table 9) (Taguchi et al., 2011). Although HNF4 α has not been reported in the development or progression of PDAC, by searching databases, it was noted the amplification of the HNF4 α locus (Maser et al., 2007) and the up-regulation of HNF4 α mRNA (Iacobuzio-Donahue et al., 2003; Logsdon et al., 2003) in human PDAC. Although some reports show HNF4 α in pancreatic acinar cells as well as islets in adult mice (Gupta et al., 2007), HNF4 α was found to be predominantly expressed in islets, with very low or no expression elsewhere in the normal pancreas (Figure 12C). Strikingly, while HNF4 α was not expressed in normal human or mouse pancreatic ducts (Figure 6A; Figure 12C), it was expressed in the nucleus of 10-22 teratoma PanIN-like cells at 3 months and invasive stage cells at 9 months (Figure 6B,C; Figure 12D). HNF4 α was detected cytoplasmically in moderately differentiated domains of the original tumor of the 10th patient, from which 10-22 cells were derived, but not in undifferentiated portions of the tumor (Figure 6D).

To more quantitatively assess HNF4 α expression at different stages of human pancreatic cancer, a tissue microarray was used to assess multiple samples of human PanINs at different stages as well as samples of pancreatic cancer. HNF4 α was barely detected in the nuclei of normal pancreatic ducts and very weakly in the samples of PanIN-1 cells, but exhibited a statistically significant increase in nuclear expression in the samples of PanIN-2 ($p < 0.05$) and stronger and more uniform expression in PanIN-3 epithelia ($p < 0.05$) (Figure 6E, G, J; Table 10). HNF4 α was most frequently detected in well differentiated mucinous sections of PDAC ($p < 0.01$) and barely or not detectable in undifferentiated or poorly differentiated epithelial structures of PDAC (Figure 6H-J), as seen in the original patient #10 tumor (Figure

6D). It was also note that in the human protein atlas (Uhlen et al., 2010), HNF4 α appeared positive in well differentiated epithelial structures of PDAC but either cytoplasmic or not expressed in poorly differentiated or undifferentiated epithelial structures of PDAC, and not expressed in metastatic PDAC.

5

Table 5. List of secreted or released proteins reported in PDAC, pancreatitis or PanIN RNA or proteomics or newly discovered here (*i.e.*, previously unreported)

secreted proteins from 10-22 IPS teratoma culture #1, 2, 3 (mouse 7767, 9223, 9225)	ID	gene	type of source identified (PDAC/PanIN/IPMN/Pancreatitis)
reported in PDAC, pancreatitis, or PanIN RNA or proteomics	IP100022443	AFP	PDAC (cancer stem cell)
	IP100032258	C4A	PDAC-DIGE (protein), Pancreatic juice (protein)
	IP100465084	DES	PDAC
	IP100007960	POSTN	Chronic pancreatitis, PDAC, stroma cells (RNA-northern blot)
	IP100218292 IP100013405 IP100029256 IP100002901	UFD1L MMP10 CNRPE TEAD1	PDAC (RNA microarray) PDAC-RNA (SAGE) PDAC (RNA microarray) PDAC, MSLN
previously unreported in PanIN, IPMN, or PDAC	IP100171199	PSMA3	
	IP100019884	ACTN2	
	IP100009305	GNPDA1	
	IP100009960	IMMT	
	IP100028833	ZNF160	
	IP100004671	GOLGB1	
	IP100645947	RITN	
	IP100328762	ABCA13	
	IP100251161	KIAA1109	
	IP100021715	CCL44S	
	IP100024911	ERP29	
	IP100159322	TCF20	
	IP100171903	HNRNPM	
	IP100021327	GRB2	
	IP100298994	TLN1	
	IP100218474	ENO3	
	IP100288940	OBSCN	
secreted proteins from 10-22 IPS teratoma culture #1, 2 (mouse 7767, 9223)	ID	gene	type of source identified (PDAC/PanIN/IPMN/Pancreatitis)
reported in PDAC, pancreatitis, or PanIN RNA or proteomics	IP100014581	TPM1	PDAC (RNA-microarray)
	IP100216171	ENO2	small cell carcinoma, microcystic serous cystadenoma, microcystic serous cystadenoma; von Hippel-Lindau (VHL)-associated serous cystic neoplasm; solid serous Adenoma; serous cystadenocarcinoma
	IP100018769	THBS2	mucinous cystic neoplasms (RNA-microarray), chronic pancreatitis (RNA, protein), PDAC (RNA, protein)
previously unreported in PanIN, IPMN, or PDAC	IP100176193	COL14A1	PDAC, pancreatitis (protein ICAT)
	IP100247295	GYNE1	
	IP100152653	DNAH5	
	IP100002127	DNAH1	
	IP100375560	ZNF804A	
	IP100005776	NOD1	
	IP100396218	SCYL2	
	IP100183041	SCN8A	
	IP100647504	FAM184A	
	IP100168442	C9orf79	
	IP100335009	HMCN2	
	IP100026665	CDNA	
	IP100328748	FLJ77177	
	IP100293887	STARD6 (DLC3)	
	IP100334410	MYO18A	
secreted proteins from 10-22 IPS teratoma culture #1, 3 (mouse 7767, 9225)	ID	gene	type of source identified (PDAC/PanIN/IPMN/Pancreatitis)
reported in PanIN and IPMN (RNA)	IP100218570	PGAM2	PanIN-3 (RNA-microarray)
reported in PDAC, pancreatitis, or PanIN RNA or proteomics	IP100328113	FBN1	PDAC; Chronic Pancreatitis; Endocrine Neoplasms (RNA-microarray, Protein-ICAT)
	IP100019439	FBN2	5-UTR of FBN2 are aberrantly methylated in PDAC BxPC3 cell line
	IP100022200	COL6A3	PDAC, Pancreatitis (RNA-SAGE, microarray)
	IP100004233	MKI67	IPMN; PanIN-1A and PanIN-1B; PanIN-2; PanIN-3; PDAC (protein-IF, RNA-microarray)
	IP100062047 IP100027760 IP100221384	KIF12 MMP-2 COL12A1	PDAC, Pancreatic islet cell, fetal liver (mRNA in silico) PDAC, Pancreatitis (RNA, Protein) PDAC-RNA (microarray)
previously unreported in PanIN, IPMN, or PDAC	IP100166612	CMYA5	
	IP100235481	PMFBP1	
	IP100847543	TRANK1	
	IP100031104	TCHP	
	IP100012629	RAD51C	
	IP100024804	ATP2A1	
	IP100888430	DNAH17	
	IP100289776	MYCBP2	
	IP100177498	LIMCH1	
	IP100377214	NLRX1	
	IP100019090	COL19A1	
	IP100013914	RAG1	
	IP100217185	RYR3	
	IP100453476		

Table 5, continued.

Secreted proteins from 10-22 iPS teratoma culture # 2, 3 (mouse 9223, 9225)	ID	Gene	Type of source identified (PDAC/PanIN/IPMN/Pancreatitis)
Reported in PanIN and IPMN RNA	IP00023014	VWF	IPMN(14982844)-RNA, Endocrine Neoplasms(14676124), plasma(3488775, 16525410, 12588351)
Reported in PDAC, pancreatitis, or PanIN RNA or	IP00289329 IP00018136	EPHB3 VCAM1	PDAC cell lines(microarray) PDAC(RNA, protein-17652277)
Previously unreported in PanIN, IPMN, or PDAC	IP00847609	SVEP1	
	IP00152881	SHROOM3	
	IP00239405	SYNE2	
	IP00412106	DNAH12	
	IP00896378	GNN	
	IP00420019	DOS	
	IP00396634	KIAA1671	
	IP00412024	ZNF485	
	IP00220986	ADAMTS9	
	IP00002320	FLRT3	
	IP00007256	ZHX2(AFR1)	
	IP00221255	MYLK	
	IP00030153	TDRD6	
	IP00031589	ZNF354A	
	IP00383970	MUM1L1	
	IP00902614	USP24	
	IP00936051	DNHD1	
	IP00741537	PRRC2B	
	IP00743955	URGCP	
	IP00455316	FRAS1	
	IP00011063	PRSS12	
	IP00102575	ATAD5	
	IP00384122	DLST	
	IP00165506	POLDIP2	
	IP00011385	LOXL3	
	IP00020906	IMPA1	
	IP00292836	KIAA1529	
	IP00027280	TOP2B	
	IP00009841	EWSR1	
	IP00143753	U2SURP	
	IP00008315	EPHB1	
	IP00029046	MLEC	
	IP00013881	HNRNPB1	
	IP00431645	HP	
	IP00398020	ODZ3	
	IP00411680	PCMT1	
	IP00215893	HMOX1	
	IP00303300	FKBP10	
	IP00221325	RANBP2	

HNF4 α was assessed in a mouse model of PDAC arising in a *KrasG12D;p53L/+; Pdx1-Cre; RosaLSL-YFP* background (Rhim et al., 2012). As in humans, HNF4 α was sporadically expressed at the PanIN-1 stage (Figure 6K), it was expressed in most nuclei of PanIN-2 (Figure 6L) and PanIN-3 lesions (Figure 6M), and observed in nuclei of the more differentiated portions of the murine tumors but not in the undifferentiated portions (Figure 6O, P; white arrows). In conclusion, the ability of the 10-22 iPS-like cells to undergo early stages of human pancreatic cancer were validated by their morphology, histology, secreted or released proteins, and

using them to discover a previously unappreciated network associated with early to invasive stage pathology in human clinical samples and a mouse model of the disease.

Table 6

	ID	Description	Gene	iPS#1			iPS#2			iPS#3			9223 (#2)			Tetratoma 9225 (#3)			T761 (#1)		
				H	M	C	H	M	C	H	M	C	H	M	C	H	M	C	H	M	C
Proteins identified by human specific peptides (not mouse peptides)	IP00252838	isoform 1 of Uncharacterized protein KIAA1529	KIAA1529	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	IP00347609	isoform 1 of Swiss von Willebrand factor type A, EGF and pentrich domain-containing protein 1	EVGP1	0	0	0	0	0	0	0	0	0	2	0	1	0	4	0	0	0	0
	IP00152881	isoform 1 of Protein Shroom3	SHROOM3	0	0	0	0	0	0	0	0	0	2	0	0	0	3	0	0	0	0
	IP00239408	isoform 1 of Nesprin-2	SYNE2	0	0	0	0	0	0	0	0	0	2	0	1	0	2	0	0	0	0
	IP00018136	isoform 1 of Vascular cell adhesion protein 1	VCAM1	0	0	0	0	0	0	0	0	0	2	0	1	0	1	0	1	0	0
	IP00412105	isoform 1 of Dynein heavy chain 12, axonemal	DNAH12	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0
	IP00856378	isoform 2 of Tetrapeptide repeat protein (DNN)	DNN	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0
	IP00420015	isoform 1 of Protein_Dos	DOS	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0
	IP00386634	isoform 1 of Uncharacterized protein KIAA1671	KIAA1671	0	0	0	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0
	IP00412024	Zinc finger protein 485	ZNF485	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	0	0	0
	IP00220986	isoform 3 of A disintegrin and metalloproteinase with thrombospondin motifs 9	ADAMTS9	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0
	IP00002320	Leucine-rich repeat transmembrane protein FLRT3	FLRT3	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0
	IP00007256	Zinc fingers and homeoboxes protein2	ZHFX2(AFR1)	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0
	IP00211255	isoform 2 of Myosin light chain kinase, smooth muscle	MYLK	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0
	IP00301953	Tudor domain-containing protein 6	TDRD6	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00231589	Zinc finger, protein 354A	ZNF354A	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00363970	PWWP domain-containing protein MUM1L1	MUM1L1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00902614	Ubiquitin carboxyl-terminal hydrolase 24	USP24	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00936051	Dynein heavy chain domain-containing protein 1	DNRD1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00741937	protein PATZ-like 1	PRRC2L	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00743955	up-regulated gene 4 isoform 2	URG4CP	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00485316	extracellular matrix protein FRAS1 isoform 1 precursor	FRAS1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00211053	Neurotrophin	PRG13	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00100575	isoform 1 of ATPase family AAA domain-containing protein 5	ATAD5	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
	IP00354122	cDNA FLJ55034 highly similar to: Dihydroxylysine residue succinyltransferase component of 2 oxoglutarate dehydrogenase complex, mitochondrial	DLAT	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
	IP00203014	von Willebrand factor	VWF	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
Proteins identified by peptides that are matched in both human and mouse (not specific human)	IP00165506	Polymerase delta interacting protein 2	POLDIP2	0	0	0	0	0	0	0	0	0	0	0	5	0	0	4	0	0	0
	IP00111366	Lysyl oxidase homolog 3	LOXL3	0	0	0	0	0	0	0	0	0	0	0	5	0	0	4	0	0	0
	IP00289328	Ephrin type-A receptor 3	EPHA3	0	0	0	0	0	0	0	0	0	0	0	5	0	0	3	0	0	0
	IP00020905	Inositol monophosphatase 1	IMPA1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	3	0	0	0
	IP00037280	isoform Beta-2 of DNA topoisomerase 2-beta	TOP2B	0	0	0	0	0	0	0	0	0	0	0	3	0	0	3	0	0	0
	IP00009841	RNA-binding protein EWSR1 isoform 1	EWSR1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	2	0	0	0
	IP00143752	isoform 1 of U2-associated protein BP140	UBS1URP	0	0	0	0	0	0	0	0	0	0	0	2	0	0	3	0	0	0
	IP00008315	Ephrin type-B receptor 1	EPHB1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	3	0	0	0
	IP00029046	Moleclin	MLEC	0	0	0	0	0	0	0	0	0	0	0	2	0	0	3	0	0	0
	IP00013881	Heterogeneous nuclear ribonucleoprotein H	HNRNPH1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0
	IP00431645	HP protein	HP	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0
	IP00198020	Teneurin-3	ODZ3	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0
	IP00411680	isoform 1 of Protein-L-lysine/aspartate (O ⁶ -aspartate) O-methyltransferase	PCMT1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0
	IP00215893	Heme oxygenase 1	HMOX1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0
	IP00303300	Peptidyl prolyl isomerase FKBP10	FKBP10	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	0
	IP00221325	E3 SUMO protein ligase RanBP2	RANBP2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0

- 5 Numbers highlighted indicate the number of peptides providing a protein identification only in conditioned medium from explants cultured in serum-free medium from two independent tetratomas of 10-22 iPS-like cells (mouse #9223 and 9225), and not in media from explants of contralateral control tissue or in media from the 10-22 iPS-like line cultured under pluripotency conditions. Total is 42 proteins.

H-human specific peptide, M-mouse specific peptide, C-human and mouse peptide, B-bovine specific, U-species other than human/mouse/bovine.

DISCUSSION

5 There has been an absence of live human cell models of PDAC progression and consequently little information about proteins that could serve as released biomarkers and pathway indicators for early stages of the disease. When human PDAC or pancreatic cancer stem cells are grafted into immunodeficient mice, tumors rapidly arise that resemble the advanced PDAC stages from which the cells
10 were derived and they do not undergo the slow growing phenotype of PDAC precursors. Based on the ability of certain cancer cells to be reprogrammed to pluripotency by nuclear transfer and then to undergo early mammalian development, it was hypothesized that pluripotent stem cell lines from human pancreatic tumors might have the capacity to progress through early developmental stages of the cancer.
15 This would provide an opportunity for discovering intrinsic processes and secreted protein biomarkers of live, early-stage human cells for a devastating cancer. Indeed a rare, single pancreatic cancer iPS-like line, 10-22 cells, can provide novel insights into human cancer progression.

 The ectopic expression of Oct4, Sox2, Klf4, and c-Myc and a
20 pluripotent-like state suppressed the cancer phenotype. As previously observed, the pluripotency epigenetic environment can dominate over certain oncogenic states (Lonardo E, 2011). In nuclear transfer studies, only certain cancer cells are amenable to reprogramming (Blelloch et al., 2004; Hochedlinger et al., 2004; Li et al., 2003) and, similarly, one iPS-like line was obtained from pancreatic cancer harboring a
25 *KRAS* mutation, the predominant driver of PDAC. While the *KRAS* mutation induces MAPK signaling, which can trigger mouse ES cells to differentiate (Kunath et al., 2007), in human ES cells, MAPK signaling can promote self-renewal (Eiselleova et al., 2009). Oncogenic *RAS* induces cellular senescence by the accumulation of p53 or CDKN2A (Serrano et al., 1997) and the expression of the four reprogramming factors
30 also triggers senescence by inducing p53 and CDKN2A, thereby impairing reprogramming (Banito et al., 2009). Only patient #10 had a deletion in exon2 of CDKN2A, possibly explaining how the 10-22 cells could escape a senescent phenotype. Additional mutations could have arisen in the 10-22 cells that made the cells particularly amenable to iPS formation.

Table 7

	ID	Description	gene	iPS			Teratoma		
				iPS#1	iPS#2	iPS#3	9223 (#2)	9225 (#3)	7761 (#1)
				H M C B U	H M C B U	H M C B U	H M C B U	H M C B U	H M C B U
Proteins identified by human specific peptides	IP00247295	Isoform 4 of Nesprin-1	SYNE1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	2 0 0 0 0	0 0 0 0 0	3 0 0 0 0
	IP00176193	Isoform 3 of Collagen alpha-1(XIV) chain	COL14A1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 0 0 0	1 0 4 0 0
	IP00152653	Dynein heavy chain 5, axonemal	DNAH5	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	2 0 0 0 0	0 0 0 0 0	1 0 2 0 0
	IP00002127	Isoform 2 of Dynein heavy chain 1, axonemal	DNAH1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 3 0 0	0 0 0 0 0	1 0 1 0 0
	IP00375560	Zinc finger protein 804A	ZNF804A	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	2 0 0 0 0	0 0 0 0 0	1 0 0 0 0
	IP00005776	Nucleotide-binding oligomerization domain-containing protein 3	NOD1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	2 0 0 0 0	0 0 0 0 0	1 0 0 0 0
	IP00396218	SCY1-like protein 2	SCYL2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 1 0 0	0 0 0 0 0	1 0 0 0 0
	IP00183041	Isoform 1 of Sodium channel protein type 8 subunit alpha	SCN8A	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 1 0 0	0 0 0 0 0	0 0 1 0 0
	IP00647504	Isoform 2 of Protein FAM184A	FAM184A	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	0 0 0 0 0	1 0 0 0 0
	IP00166442	FAM75-like protein C9orf79	C9orf79	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	0 0 0 0 0	1 0 0 0 0
	IP00335009	Isoform 3 of Hemicentin-2	HMCN2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	0 0 0 0 0	1 0 0 0 0
Proteins identified by peptides that are matched in both human and mouse (not specific human)	IP00014581	Isoform 1 of Troponin alpha-1 chain	TPM1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 0 0 0	0 0 7 0 0
	IP00216171	Galactose-6-epimerase	ENO2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 0 0 0	0 0 7 0 0
	IP00018769	Thrombospondin-2	THBS2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 0 0 0	0 0 5 0 0
	IP00026665	cDNA FLJ75085, highly similar to Homo sapiens glutathione S-transferase (GARS), mRNA		0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 2 0 0	0 0 0 0 0	0 0 2 0 0
	IP00328748	mesencephalic astrocyte-derived neurotrophic factor		0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 0 0 0	0 0 2 0 0
	IP00293887	Isoform 2 of STAR-related lipid transfer protein 8	STAR8S	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 2 0 0	0 0 0 0 0	0 0 1 0 0
	IP00334410	Isoform 2 of Myosin-XVIIIa	MYO18A	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 0 0 0	0 0 1 0 0

Numbers highlighted indicate the number of peptides providing a protein identification only in conditioned medium from explants cultured in serum-free medium from two independent teratomas of 10-22 iPS-like cells (mouse #9223 and 7661), and not in media from explants of contralateral control tissue or in media from the 10-22 iPS-like line cultured under pluripotency conditions. Total is 18 proteins. H-human specific peptide, M-mouse specific peptide, C-human and mouse peptide, B-bovine specific, U-species other than human/mouse/bovine.

Table 8

	ID	Description	Gene	iPS			Teratoma		
				iPS#1 H M C B U	iPS#2 H M C B U	iPS#3 H M C B U	9223(#2) H M C B U	9225(#3) H M C B U	7761(#1) H M C B U
Proteins identified by human specific peptides	P00186612	Cardiomyopathy-associated protein 5	CMYA5	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	3 0 0 0 0
	P00235481	Isoform 1 of Polyamine-modulated factor 1-binding protein 1	PMFBP1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 1 0 0	2 0 0 0 0
	P00847543	TPR and ankyrin repeat-containing protein 1	TRANK1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	2 0 0 0 0
	P00021104	Trichoplein keratin filament-binding protein	TCHP	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	2 0 0 0 0
	P00012820	DNA repair protein RAD51 homolog 3	RAD51C	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	2 0 0 0 0
	P00024804	Isoform SERCA1B of Sarcoplasmic/endoplasmic reticulum calcium ATPase 1	ATP2A1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 2 0 0	1 0 0 0 0
	P00328113	Fibrillin-1	FBN1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 1 0 0	1 0 0 0 0
	P00022200	Isoform 1 of Collagen alpha-3(VI) chain	COL6A3	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 4 0 0	1 0 5 0 0
	P00888430	Isoform 1 of Dynein heavy chain 17, axonemal	DNAH17	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 1 0 0	1 0 1 0 0
	P00289778	probable E3 ubiquitin-protein ligase MYCBP2	MYCBP2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	1 0 1 0 0
	P00177496	Isoform 3 of LIM and calponin homology domains-containing protein 1	LMCH1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	2 0 0 0 0	1 0 0 0 0
	P00377214	Isoform 2 of NLR family member X1	NLRX1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	2 0 0 0 0	1 0 0 0 0
	P00004233	Isoform Long of Antigen Ki-67	MKI67	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	1 0 0 0 0
	P00012090	Collagen alpha-1(XIX) chain	COL19A1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 1	2 0 0 0 0
	P00013614	V(D)J recombination-activating protein 1	RAG1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	1 0 0 0 0
	P00062047	Kinesin-like protein KIF12	KIF12	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 0 0 0	1 0 0 0 0
	P00217185	Putative uncharacterized protein RYR3	RYR3	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 0 1 0 0	0 0 1 0 0
Proteins identified by peptides that are matched in both human and mouse (not specific human)	P00218570	Phosphoglycerate mutase 2	PGAM2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 4 0 0	0 0 6 0 0
	P00453478	uncharacterized 28 kDa protein		0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 2 0 0	0 1 4 0 0
	P00027780	72 kDa type 1/ collagenase	MMP-2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 3 0 0
	P00221384	Isoform 2 of Collagen alpha-1(XII) chain	COL12A1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 3 0 0
	P00019439	Isoform 1 of Fibrillin-2	FBN2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 1 0 0

Numbers highlighted indicate the number of peptides providing a protein identification only in conditioned medium from explants cultured in serum-free medium from two independent teratomas of 10-22 iPS-like cells (mouse #9223 and 7661), and not in media from explants of contralateral control tissue or in media from the 10-22 iPS-like line cultured under pluripotency conditions. Total is 22 proteins. H-human specific peptide, M-mouse specific peptide, C-human and mouse peptide, B-bovine specific, U-species other than human/mouse/bovine.

The release from pluripotency allowed the cancer genome to be expressed in a stage-specific fashion, as opposed to undergoing an immediate regression to the late stage phenotype. Release from pluripotency is normally

accompanied by the development of germ layer cells and then specialized tissues,
 which may continue to dominate, epigenetically, over the resident cancer genome
 (Blelloch et al., 2004; Hochedlinger et al., 2004; Li et al., 2003). The 10-22 cells from
 PDAC generated diverse tissue types in teratomas as well as pancreatic ductal tissue
 5 that exhibited PanIN lesions and later progression. The apparent preference for
 pluripotent cells to regenerate the cancer type from which they were derived reflects
 the tendency of iPS cell lines in general to preferentially differentiate into their
 lineages of origin (Bar-Nur et al., 2011; Kim et al., 2011). Several lines of evidence
 indicate that the 10-22 iPS-like line is derived from PDAC. First, the pathology of the
 10 original, recurrent tumor was that of PDAC and the CGH profile of the bulk
 population of cultured cells, which had a highly disrupted genome, was represented in
 the CGH profile of the 10-22 iPS-like line (Figs. 11I, 9B). While the tumor harbored
 pockets of more differentiated epithelial cells amidst a vast majority of
 undifferentiated cells, and therefore it is not certain which type of cell was
 15 immortalized in the 10-22 line, the 10-22 cells' disrupted genome does reflect that of
 a typical epithelial cell in the recurrent tumor. Second, the PanIN-like structures from
 the 10-22 cells' teratomas expressed SOX9 (Figure 11B-E), which is required for
 early, *KrasG12D*-dependent pancreatic precursor lesions in a mouse model (Kopp et
 al., 2012), as well as PDX1, a definitive pancreatic cancer epithelial marker. Thus the
 20 ductal lesions from the 10-22 cells are of a pancreatic type. Third, teratomas at 9
 months from 10-22 cells progressed to the histology and locally invasive
 characteristics of later stage PDAC (Fig. 4). Thus, the 10-22 line was not from an
 early stage cell that would solely undergo an early stage phenotype. Not limited to a
 particular theory, the evidence indicates that the 10-22 iPS-like line is from PDAC
 25 cells in the original tumor and that, upon re-differentiation in teratomas, it undergoes
 progression of the disease. This is unlike other human PDAC lines, which exhibit late
 stages of cancer (Lieber et al., 1975; Yunis et al., 1977).

Pluripotency genes such as *NANOG* are expressed in sphere cultures of
 pancreatic cancer stem cells (CSCs), suggesting that such cells might be more
 30 susceptible to reprogramming (Lonardo et al., 2011). However, CD133+CXCR4+
 pancreatic CSCs are not enriched and the expression of pluripotent genes is not
 observed in the adherent culture conditions used to derive the 10-22 cells (Hermann et
 al., 2007). Also, the OCT4 and *NANOG* pluripotency genes were highly methylated
 in the parental tumor #10 epithelium cultures, in contrast to the 10-22 cell line and the

huES H1 control, and NANOG itself was not expressed in the primary tumor, although OCT4 was expressed sporadically (Fig. 8B, C). Not limited to a particular theory, it seems unlikely that 10-22 cells were derived from pancreatic CSCs. In addition, pancreatic CSCs (Hermann et al., 2007; Ishizawa et al., 2010; Li et al., 5 2007) rapidly generate aggressive tumors that represent the primary tumors; whereas the 10-22 cells generate slow growing PanINs (Figs. 3, 10). Finally, tumors generated with pancreatic CSCs give rise to both cytokeratin negative and positive cells in the resultant tumors, in contrast to the homogenous K19 positive staining in PanIN-like ducts at 3 month teratomas (Figs. 3, 8A). Thus, 10-22 cells appear not to exhibit 10 properties of pancreatic cancer stem cells.

The proteins released or secreted from the PanIN-like teratomas fell into at least 3 major networks, including inter-connected networks for TGF β and integrin signaling that suppress PDAC progression (Hezel et al., 2012). For the first time, the activation of an HNF4 α network distinctive for the late PanIN to invasive 15 stages was discovered. HNF4 α is not or barely expressed in normal pancreatic ductal cells, poorly expressed in the PanIN1 stage, but is activated in PanIN2 and PanIN3 stages, invasive stages, and in early well-differentiated human pancreatic cancer. HNF4 α levels then decrease markedly in advanced or undifferentiated PDAC. It was found that these expression states also occur in a mouse model of PDAC progression. 20 Dynamics in HNF4 α expression affect the oncogenic transformation of liver cells (Hatzia Apostolou et al., 2011). It remains to be determined whether the expression of HNF4 α and its target genes is a cause or consequence of pancreatic cancer progression. Yet considering that pancreatic cancer is typically discovered in advanced or metastatic stages, activation of HNF4 α and the release or secretion of 25 proteins from the factor's target genes specifically in the late PanIN stages should provide useful diagnostics.

Table 9. Teratoma-secreted proteins that function in the HNF4 α pathway.

Gene	description
ATAD5*	Isoform 1 of ATPase family AAA domain-containing protein 5
ACTN2	Alpha-actinin-2
C9orf79	FAM75-like protein C9orf79
DNAH1	Isoform 2 of Dynein heavy chain 1, axonemal
DNAH5*	Dynein heavy chain 5, axonemal
DNAH12	Isoform 1 of Dynein heavy chain 12, axonemal
DNAH17*	Isoform 1 of Dynein heavy chain 17, axonemal
FBN1	Fibrillin-1
FBN2	Isoform 1 of Fibrillin-2
FKBP10	Peptidyl prolyl cis trans isomerase FKBP10
FLRT3	Leucine-rich repeat transmembrane protein FLRT3
FRAS1*	extracellular matrix protein FRAS1 isoform 1 precursor
OBSCN*	Isoform 1 of Obscurin
LOXL3*	Lysyl oxidase homolog 3
MLEC	Malectin
PCMT1	Isoform 1 of Protein-L-isoaspartate (D-aspartate) O-methyltransferase
PGAM2*	Phosphoglycerate mutase 2
PRRC2B	protein BAT2-like 1
RAG1	V(D)J recombination-activating protein 1
RANBP2*	E3 SUMO protein ligase RanBP2
SCYL2	SCY1-like protein 2
TCHP	Trichoplein keratin filament-binding protein
USP24*	Ubiquitin carboxyl-terminal hydrolase 24
ZHX2	Zinc fingers and homeoboxes protein2
ZNF804A	Zinc finger protein 804A

Proteins that are directly or indirectly regulated by HNF4 α , out of the secreted proteins from at least two 10-22 teratoma explants embedded into Matrigel.

- 5 Of 107 proteins reproducibly released or secreted from PanIN-like cells derived from the 10-22 line, a total of 68 proteins overlap with genes, proteins and networks expressed in human PanIN and PDAC (Tables 10, 11, 12 and 13), further validating the origin of the cells. In addition, it is shown that such proteins are released from the cells and stable, thus serving as biomarkers of early PDAC. A
- 10 subset of the secreted proteins could be from locally activated stromal cells in the explants. Not limited to a particular theory, the combined detection of released proteins that are the products of the HNF4 α , TGF β and integrin networks within PanIN and invasive PDAC cells provide means for noninvasively detecting the progression of pancreatic cancer in humans.

Table 10. Teratoma-secreted proteins that function in the HNF4 α pathway.

ID	Gene	Teratoma															Network	functions...etc.	Identified in plasma - based on plasma protein DB	Identified in plasma (MOPED, eg HP-1000 nmol)	RNA level (BioGPS DB) Genome Biology 2008, 10:R130 doi:10.1186/gb-2008-10-11-r130
		S223(#2)					S225(#3)					7761(#1)									
		H	M	C	B	U	H	M	C	B	U	H	M	C	B	U					
IP10028940	CEBPN	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	HNF4a	Identified as one of the genes frequently mutated in breast/colorectal cancer, implicating it in cancer formation, tumor suppressor	yes	0.1nmol	most of tissue-pancreas less than average(14.68)
IP100011365	LOXL3	0	0	5	0	0	0	0	4	0	0	0	0	0	0	0	HNF4a	Involved in tumor suppressor, catalyzes the first step in the formation of cross linking in collagen and elastin, expressed in placenta, small intestine, testis, heart, ovary and spleen (decreased in pancreas)	yes	0 nmol	most of tissue-small amount of RNA (pancreas less than average 17.6)
IP100028346	MLEC	0	0	2	0	0	0	0	3	0	0	0	0	0	0	0	HNF4a	conserved ER-resident lectin, mRNA level was upregulated in PDAC compared to pancreas(MOPED)	yes	no plasma, 7/8-1nmol	no information there
IP100002127	DNAH1	0	0	3	0	0	0	0	0	0	0	1	0	1	0	0	HNF4a	axon guidance	yes	0 nmol	most of tissue-small amount of RNA (pancreas less than average 8.5)
IP100152653	DNAH5	2	0	0	0	0	0	0	0	0	0	1	0	2	0	0	HNF4a	axon guidance	yes	1(0.1 nmol) (panc)	most of tissue-small amount (7.8)
IP100412106	DNAH12	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	HNF4a	axon guidance	yes	1(5nmol)	most of tissue small amount, testis high, pancreas below aver(7.69)
IP10088430	DNAH17	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0	HNF4a	axon guidance	no	0 (0.2 nmol)	testis highly(77), other tissue small amount, pancreas less than aver(6.3)
IP100066218	SCYL5	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	HNF4a	Component of AP2-containing clathrin coated structures at the plasma membrane or of endocytic coated, RNA expressed in normal pancreas (but more in PDAC)	yes	0nmol	most of tissue low level, pancreas below than Av (3.68)
IP100303300	FKBP10	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	HNF4a	RNA expressed in normal pancreas (but more in cancer)	no	0nmol	smooth muscle most high level(60), Other tissue low level, pancreas lower than aver(6.3)
IP100002320	FLRT3	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	HNF4a	downregulated in PDAC	yes	no plasma, Tcell-0.01nmol (panc)	bronchial epithelial high, other tissue low, pancreas below than aver(9.3)
IP100007286	ZHX2(AFR1)	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	HNF4a	Interact with AFP, if select AFP, I have to choose this one too.	yes	0nmol	CD19+8 cells the highest(212.65), other tissue intermediate/or low, pancreas low level less than aver(20)
IP100375580	ZNF804A	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	HNF4a	Interact with ATXN1 which has been interact with GAPDH	no	0.1 nmol	expressed in most tissue low levels, Pancreas less than aver(7.5)
IP100018884	ACTN2	0	0	4	0	0	0	0	3	0	0	0	0	24	0	0	HNF4a	interact with CMYA5 (in TGF)	yes	0.1 nmol	the highest level in skeletal muscle and Heart(1635.96), intermediate-Tongue(30-115), Thyroid, other tissue below aver 52.5

Table 11

ID	Gene	Teratoma															Network	functions, etc.	Identified in plasma-based on plasma protein DB	Identified in plasma (MOPED, eg HPLC-1000 nmol)	RNA level (BioGPS DB) Genome Biology 2006, 10.R130 doi 10.1186/gb-2006-10-11-r130
		2273(#2)					2273(#3)					7761(#1)									
		H	M	C	B	U	H	M	C	B	U	H	M	C	B	U					
IP00847606	SVEP1	2	0	1	0	0	4	0	0	0	0	0	0	0	0	no network	Serologically defined breast cancer antigen NY-BR-382, putative Integrin ligand, may play a role in the cell attachment. Present in mesenchymal primary cultured cell lysates (at protein level). Highly expressed in placenta. Also expressed in marrow stromal cell. Weakly or not expressed in other antibodies against KIAA1871 are present in sera from breast cancer patient	Yes	0.1 nmol	not at all (only placenta)	
IP00866834	KIAA1871	1	0	0	0	0	3	0	0	0	0	0	0	0	0	no network		no	0nmol	no data available	
IP00292536	KIAA1522	5	0	0	0	0	3	0	0	0	0	0	0	0	0	no network	in chr9 (interferon cluster)	yes	no plasma, Toes-0.0nmol	no data available	
IP00325748	MANF	0	0	1	0	0	0	0	0	0	0	0	0	2	0	no network	cDNA FLJ77177, highly similar to Homo sapiens arginine-rich, mutated in early stage tumors (ARIMET). mRNA originally identified from testis, mesencephalic trophoblast-derived neurotrophic factor	no information	10 nmol	express highest level in liver(~1022), pancreas exp average level (~220)	
IP00453476	26kDa proteins	0	0	0	0	0	0	0	2	0	0	0	1	4	0		Uncharacterized protein ENSP00000248237, phosphoglycerate mutase 1, upregulated in most of cancer	no information	50 nmol	most of tissue express, but more in brain tissue(4345), pancreas below aver(1444)	
IP00026665		0	0	1	0	0	0	0	0	0	0	0	0	2	0	no network	CDNA FLJ75085, highly similar to human glutamyl-tRNA synthetase (QARS), mRNA	no information	searched based on QARS (no plasma, monocyte 10nmol)	blood cell express highest(2612), most of tissue express, pancreas-moderately below aver(441)	
IP00865378	GNN	2	0	0	0	0	1	0	0	0	0	0	0	0	0	no network		no	no information	most of tissue (all tissue highly)	
IP00420018	DOS	2	0	0	0	0	1	0	0	0	0	0	0	0	0	no network		no	0.5 nmol	most of tissue(pancreas too)	
IP00293457	STARD8 (DLC3)	0	0	2	0	0	0	0	0	0	0	0	0	1	0	no network		yes	no plasma, Boehr-0.1nmol	most of tissue express similar level along the aver(5.95), pancreas (4.45)	
IP00183041	SCN8A	1	0	1	0	0	0	0	0	0	0	0	0	1	0	no network		yes	0.5 nmol	most of tissue (abundant)	
IP00143753	U2SURF	0	0	2	0	0	0	0	3	0	0	0	0	0	0	no network		no	no plasma, Boehr-1 nmol	blood cell express highest, most of tissue express, pancreas below aver(53)	
IP00031104	TGIF	0	0	0	0	0	1	0	0	0	0	2	0	0	0	no network		yes	1nmol	most of tissue express similar level along the aver(5.61)	
IP00012820	RADE1C	0	0	0	0	0	1	0	0	0	0	2	0	0	0	no network		yes	0nmol (panc)		
IP00024804	ATF2A1	0	0	0	0	0	0	0	2	0	0	1	2	8	0	no network		yes	no plasma, Toes-0.1nmol	tongue, skeletal muscle highest (7545), other tissue undetectable	
IP00377214	NLRX1	0	0	0	0	0	2	0	0	0	0	1	0	0	0	no network		no	0nmol	most of tissue express similar level along the aver(55)	
IP00028693	ZNF160	1	0	0	0	0	1	0	0	0	0	3	0	0	0	no network		yes	1 nmol	most of tissue (but not much in pancreas)	
IP00045947	RYTN	1	0	0	0	0	1	0	0	0	0	1	0	0	0	no network		no	0 nmol	bronchial epithelial cells highest (54)most of tissue express similar level along the aver(7.5), pancreas(6.85)	
IP00326762	ABCA13	1	0	0	0	0	1	0	0	0	0	1	0	0	0	no network		yes	0nmol	most of tissue express similar level along the aver(60)	
IP00412324	ZNF485	1	0	0	0	0	2	0	0	0	0	0	0	0	0	no network		no	5 nmol	most of tissue express similar level along the aver(7.21)	
IP00220906	IMPA1	0	0	3	0	0	0	0	3	0	0	0	0	0	0	no network		yes	no plasma, Boehr-0nmol	most of tissue express but selective level, pancreas express below aver(88)	

Table 12. Teratoma-secreted proteins that function in the Ras/p53/JUN/CTNB1 pathway.

ID	Gene	Teratoma															Network	functions...etc..	Identified in plasma - based on plasma protein DB	Identified in plasma (MOPED, eg HP-1000 nmol)	RNA level (BioGPS DB) Genome Biology 2009, 10:R130 doi:10.1186/gb-2009-10-11-r130	
		9223(#2)					9225(#3)					7781(#1)										
		H	M	C	B	U	H	M	C	B	U	H	M	C	B	U						
IP100251101	KIAA1109	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	RAS/p53/JUN/CTNB1	fragile site associated protein/testis, ovary, 2C	yes	0.1 nmol	most of tissue (not much in pancreas)	
IP100398020	QDZ3	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	RAS/p53/JUN/CTNB1	plasma protein (can be cleavage) expressed in adult/fetal brain, slightly lower levels in testis, ovary, intermediate level all other peripheral tissue	yes	0 nmol (pan)	fetal brain, brain tissue only, other tissue very small	
IP100235481	PMFBP1	0	0	0	0	0	1	0	1	0	0	2	0	0	0	0	RAS/p53/JUN/CTNB1	May play a role in sperm morphology especially the sperm tail and consequently affect fertility	yes	0.1 nmol (plasma/ser secret proteins moderately)	most of tissue	
IP100289329	EPH83	0	0	5	0	0	0	0	3	0	0	0	0	0	0	0	RAS/p53/JUN/CTNB1	most of tissue express RNA	yes	0 nmol (pan) (plasma secret proteins moderately)	most of tissue express similar level, pancreas express similar level to aver(4.41)	
IP100177499	UMCH1	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	RAS/p53/JUN/CTNB1		yes	0nmol (pan)	olfactory bulb highest level(74), most of tissue express, pancreas below aver(9.8)	
IP100024011	ERP29	0	0	2	0	0	1	0	1	0	0	0	0	0	4	0	RAS/p53/JUN/CTNB1		yes	10 nmol (pancreas)	most of tissue express, pancreas similar level to aver(0.8)	
IP100150322	TCF20	0	0	1	0	0	0	0	2	0	0	1	0	0	0	0	RAS/p53/JUN/CTNB1	SPRE-binding protein2, Interacts with RNF4 and JUN	yes	0.1 nmol	most of tissue express similar level, pancreas epress similar level to aver(0.9)	

Table 13. Teratoma-secreted proteins that function in the TGFβ/Integrin pathway.

ID	Gene	Teratoma															Network	functions, etc.	Identified in plasma-based on plasma protein DB	Identified in plasma (MOPEd, eg HP-1000 nmol)	RNA level (BioGPS DB) Genome Biology 2009 10 R130 doi:10.1186/gb-2009-10-11-r130
		9223(#2)					9225(#3)					7761(#1)									
		H	M	C	B	U	H	M	C	B	U	H	M	C	B	U					
IP100152881	SHROOM3	2	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0.1 nmols	Yes	No data for gene 57619 in the "GeneAtlas U133A, gcma" data set.
IP100215993	HMOX1	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0 nmols	Yes	heart, blood cell (150-200), other tissue nearly no or less than aver (12.9)
IP100007960	POSTN	0	0	2	0	0	0	0	0	3	0	0	2	0	4	0	0	1 nmols (panc)	Yes	not all	
IP100013405	MMP10	0	0	2	0	0	0	0	0	4	0	0	0	0	2	0	0	0nmol	no	not at all(only uterus)	
IP100027790	MMP-2	0	0	0	0	0	0	0	0	1	0	0	0	0	3	0	0	2 nmol	no	muscle, adipocyte(not much pancreas)	
IP100018769	THBS2	0	0	1	0	0	0	0	0	0	0	0	0	0	5	0	0	2nmol (panc)	Yes	muscle, adipocyte(not much pancreas)	
IP100009841	EWSR1	0	0	3	0	0	0	0	0	2	0	0	0	0	0	0	0	1(1 nmol- TIF/monocyte cell) (panc)	Yes	most of tissue, pancreas below aver(221)	
IP100005776	NOD1	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1(0nmol) (panc)	Yes	most of tissue express, pancreas below aver(8.3)	
IP100220386	ADAMTS9	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1(0nmol) (panc)	Yes	not much(a little bit)	
IP100022443	AFP	3	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1(0nmol)	Yes	not at all (only liver)	
IP100465084	DES	1	0	3	0	0	0	0	0	1	0	0	2	0	6	0	0	no plasma, neutrophil-8nmol	Yes	highest expression in Uterus, Testis, Heart, but other tissue nearly no	
IP100247295	SYNE1	2	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0.2nmol	yes	blood cells, cortex highest(910), other tissue below aver(108)	
IP100239405	SYNE2	2	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0.1 nmol (panc)	Yes	thyroid(not much in pancreas)	
IP100008315	EPHB1	0	0	2	0	0	0	0	0	3	0	0	0	0	0	0	0	no plasma, monocyte-1nmol	Yes	Cerebellum most highest(10.5), most tissue express, pancreas below aver(3.9)	
IP100218292	UFD1L	0	0	4	0	0	0	0	0	5	0	0	0	0	1	0	0	0 (no plasma, monocyte/TIF cell-2nmol)	no	the highest in leukemia tissue(871), most of tissue express, pancreas below aver(131.7)	
IP100002901	TEAD1	0	0	2	0	0	0	0	0	2	0	0	0	0	2	0	0	0nmol	no	most of tissue, pancreas below aver (7.11)	

Table 13, continued.

ID	Gene	Teratoma															Network	functions, etc.	Identified in plasma-based on plasma protein DB	Identified in plasma (MOPED, eg HP-1000 nmol)	RNA level (Biogps DB) Genome Biology 2009, 10:R130 doi:10.1186/gb-2009-10-11-r130
		9223(#2)					9225(#3)					7761(#1)									
		H	M	C	B	U	H	M	C	B	U	H	M	C	B	U					
IP100217185	RVR3	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	yes	0.8nmol	the highest in pineal tissue(90-127), other tissue low level, pancreas less than aver (14.5)	
IP100166612	CMYA5	0	0	0	0	0	1	0	0	0	0	3	0	0	0	0	0	yes	0nmol	most of tissue evenly express along the average value(9.75), other transkeletal muscle(18.35)	
IP100009960	INMT	1	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	yes	no plasma, monocyte/TB cell-8nmol	B lymphoblast highest level (860), most of tissue express low level, pancreas similar level to aver(86.7), uterus and prostate express highest (4699) then, testis, small intestine, retina express (500-2000), other tissue nearly no, less than aver(292) heart(9507), muscle, tongue highest level, other tissue undetectable, less than aver (297)	
IP100221255	MYLK	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	yes	0.2 nmol	blood cells highest level (3046), other tissue detectable, pancreas below aver (504)	
IP100014581	TPM1	0	0	1	0	0	0	0	0	0	0	0	0	0	7	0	0	yes	10nmol	endothelium, fetal liver, lympho node-highest level, most of tissue below aver(59.6), pancreatic islet twice more than aver	
IP100029266	SNRPE	0	0	1	0	0	0	0	0	1	0	0	0	0	2	0	0	yes	no plasma, Tcell-10nmol	most of tissue express similar level other than blood cells, pancreas below aver (103)	
IP100027280	TOP2B	0	0	3	0	0	0	0	0	3	0	0	0	0	0	0	0	yes	0.1-1 nmo	most of tissue express similar level other than blood cells, pancreas below aver (19)	
IP100018136	VCAM1	2	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	yes	10 nmol	most of tissue express similar level other than blood cells, pancreas below aver (19)	
IP100021327	GRB2	0	0	1	0	0	0	0	0	2	0	0	0	0	5	0	0	yes	12 nmol	most of tissue express similar level other than blood cells, pancreas below aver (19)	
IP100171903	HNRNPM	0	0	2	0	0	0	0	0	1	0	0	0	0	2	0	0	yes	no plasma, Tcell-80nmol	most of tissue express similar level other than blood cells, pancreas below aver (19)	
IP100013881	HNRNPH1	0	0	2	0	0	0	0	0	2	0	0	0	0	0	0	0	yes	no plasma, Tcell-80nmol	most of tissue express, pancreatic islet express(587), but pancreas below aver(587)	

Table 14. Summary of HNF4a immunohistochemistry on human tissues.

	margin of tumor	PanIN-1	PanIN-2	PanIN-3	well diff. PDAC	poorly diff. PDAC
# samples	4	8	7	3	8	8
sections analyzed	9	18	17	17	25	21
HNF4 α + nuclei	11	248	867	1107	3099	60
total nuclei counted	209	1271	2008	1981	4585	1,677
%HNF4 α + nuclei	4.91	16.32	38.64	59.32	71.22	3.76

EXAMPLE 2: Detection of biomarkers in subjects with pancreatic cancer by ELISA.

In this Example, validation of the disclosed biomarkers for pancreatic cancer was performed in plasma samples of subjects that had pancreatic cancer using enzyme-linked immunosorbent assays (ELISA).

Validation of the disclosed biomarkers was performed by analyzing plasma samples of 10 patients with pancreatic cancer at various stages and analyzing plasma samples of 10 control subjects that did not have pancreatic cancer. Of the 10 subjects with pancreatic cancer, 7 had resectable and locally advanced cancer and 5 had resectable pancreatic cancer that was not as far advanced (Table 15).

Thirty three of the identified biomarkers described in Example 1, and one control, were analyzed in the plasma of the subjects by ELISA (Table 15). Human plasma samples were prepared from blood drawn from patients or controls by treating with EDTA or heparin as an anticoagulant. The mixture was centrifuged for 15 min at 1000 x g at 2-8 degrees centigrade within 30 min of blood collection. The material was frozen in aliquots and repeated freeze-thaw cycles were avoided. ELISA assays were performed per the instructions of the manufacturer of the ELISA kit. In brief, 100 μ l of a standard or 10 μ l of a sample were added per well of the microtiter plate; 90 μ l of diluent was added to the sample well. The wells were covered with the provided adhesive strip and incubated for 2 hours at 37°C. The liquid present in each well was removed and 100 μ l of manufacturer's Biotin-antibody (1x) was added to each well and covered with a new adhesive strip, followed by incubation at 37°C for 1 hour. Each well was aspirated and washed with 200 μ l of the manufacturer's wash

buffer repeatedly for a total of three washes. After the last wash, any remaining wash buffer was removed by aspirating or decanting and the microtiter plate was inverted and blotted to further remove any remaining fluid in each well. 100 μ l of HRP-avidin (1x) was added to each well, and the microtiter plate was covered with a new adhesive strip and incubated for 1 hour at 37°C. The aspiration/wash process was repeated for five times as performed in the previous wash step, followed by the addition of 90 μ l of the manufacturer's TMB Substrate to each well and incubation for 15-30 minutes at 37°C. After the incubation at 37°C, 50 μ l of manufacturer's Stop Solution was added to each well and the plate was gently tapped to ensure thorough mixing. The optical density of each well was determined within 5 minutes using a microplate reader set to 450 nm. The signals in each well were analyzed by a 4 parameter logistic nonlinear regression model for curve-fitting to the standard curve, using Soft Max Pro, allowing the determination of antigen concentration in each plasma sample.

To compare antigen concentrations between samples and between patient (cases) plasma and plasma of control individuals, a logistic regression analysis was performed to estimate the area under the receiver operating characteristic (ROC) curve (*i.e.*, c-statistic). The ROC curve was estimated by plotting the true positive rate (sensitivity) vs. the false positive rate (1-specificity) across the range of values observed for the biomarker of interest in the dataset. A c-statistic of 0.5 coincides with the marker predicting case/control status correctly 50% of the time. A c-statistic higher than 0.7 would indicate that the biomarker is reliable as a diagnostic for pancreatic cancer.

Table 15

Biomarker	Network	c-statistic from logistic regression (area under ROC curve)		
		All Cases (N=10)	Resectable & Locally Advanced (N=7)	Resectable (N=5)
AFP	TGF β /integrin	0.480	0.464	0.420
RTTN	No network	0.595	0.579	0.730
NLRX1	No network	0.490	0.529	0.580
DNAH12	HNF4a	1.00	1.00	1.00
ODZ3	Ras/p53/JUN/CTN B1	0.560	0.529	0.560
ADAMST9	TGF β /integrin	0.610	0.600	0.540
TPM1	TGF β /integrin	0.645	0.686	0.800

DNAH1	HNF4a	0.740	0.764	0.780
PMFBP1	Ras/p53/JUN/CTN B1	0.550	0.55	0.550
DNAH17	HNF4a	0.630	0.486	0.550
EPHB1	TGFβ/integrin	0.330	0.357	0.430
DOS	No network	0.545	0.543	0.640
MMP2	TGFβ/integrin	0.570	0.557	0.600
STARD8	no network	0.700	0.729	0.800
ATP2A1	no network	0.725	0.736	0.700
FKBP10	HNF4a	0.700	0.657	0.640
TCHP	no network	0.605	0.393	0.470
TCF20	Ras/p53/JUN/CTN B1	0.630	0.507	0.500
ABCA13	no network	0.550	0.550	0.550
SCN8A	no network	0.550	0.550	0.550
TOP2B	TGFβ/integrin	0.860	0.829	0.760
LIMCH1	Ras/p53/JUN/CTN B1	0.725	0.607	0.560
UFD1L	TGFβ/integrin	0.695	0.607	0.500
FLRT3	HNF4a	0.400	0.386	0.400
ZHX2	HNF4a	0.570	0.514	0.610
SYNE1	TGFβ/integrin	0.560	0.586	0.780
THBS2	TGFβ/integrin	0.760	0.886	0.840
HMOX1	TGFβ/integrin	0.610	0.600	0.620
Obscurin	HNF4a	0.550	0.486	0.550
DNAH5	HNF4a	0.550	0.593	0.540
Shroom3	TGFβ/integrin	0.450	0.471	0.500
LOXL3	HNF4a	0.480	0.614	0.700
Malectin	HNF4a	0.510	0.557	0.460
Ca199	N/A	1.000	1.000	1.000

The thirty three biomarkers tested were AFP, RTTN, NLRX1, DNAH12, ODZ3, ADAMST9, TPM1, DNAH1, PMFBP1, DNAH17, EPHB1, DOS, MMP2, STARD8, ATP2A1, FKBP10, TCHP, TCF20, ABCA13, SCN8A, TOP2B, LIMCH1, UFD1L, FLRT3, ZHX2, SYNE1, THBS2, HMOX1, Obscurin, DNAH5, Shroom3, LOXL3, Malectin (Table 15 and Figures 13-46). Cancer antigen 19-9 (Ca199), which is a reliable tumor marker for pancreatic cancer, was used as positive control and resulted in a c-statistic value of 1.0 (Table 15 and Figure 33). Of the thirty three biomarkers tested, RTTN, DNAH12, TPM1, DNAH1, STARD8, AP2A1, TOP2B, LIMCH1, SYNE1, THBS2 and LOXL3 were identified as reliable biomarkers for indicating the presence of pancreatic cancer in a subject as they exhibited a c-statistic value greater than 0.7 (Table 15 and Figures 13-46). For example, TOP2B, a member of the TGFβ/integrin pathway, exhibited a c-statistic

value of 0.860 in all pancreatic cancers, indicating that TOP2B is a reliable biomarker for early and advanced stage pancreatic cancer (Table 15 and Figure 23). Similar results were observed for DNAH1, STARD8, ATP2A1, LIMCH1 and THBS2 (Table 15 and Figures 15, 16, 25, 26 and 41). RTTN, TPM1, LOXL3 and SYNE1 exhibited a c-statistic value greater than 0.7 in the plasma samples of subjects with resectable cancer indicating that RTTN, TPM1, LOXL3 and SYNE1 are reliable biomarkers for early stage, resectable pancreatic cancer (Table 15 and Figures 20, 29, 35 and 40). DNAH12, a member of the HNFa pathway exhibited a c-statistic value of 1.0 in plasma samples of all pancreatic cancer patients, *i.e.*, patients with locally advanced and resectable pancreatic cancers, indicating that DNAH12 can be a reliable biomarker for pancreatic cancer (Table 15 and Figure 37).

EXAMPLE 3: Detection of biomarkers in subjects with pancreatic cancer by mass spectrometry.

In this Example, detection of the disclosed biomarkers for pancreatic cancer was performed in plasma samples of subjects that had pancreatic cancer using mass spectrometry.

Two control plasmas (M1, M2) and two metastatic PDAC case plasmas (M3, M4) were selected and coded to allow blind testing. Each plasma sample was subjected to a serum albumin removal gel ("pull-down"), and the supernatants were collected. The supernatant of each plasma sample was subjected to a mixture of protein A beads and protein G beads to remove IgG, and adjusted with 6M Urea, 10 mM DTT and 50 mM IAA to block free sulfhydryl groups. The supernatants were subsequently treated with trypsin at 1:50 and subjected to a C18 cartridge desalting step. 30 µg were aliquoted from each supernatant sample for each LC-MS/MS run. pFind was used to analyze the data and to identify peptide IDs. The reverse peptide decoys were compared with the forward peptides, a 5% FDR filter was applied and a list was generated that provided the peptide hits. Each plasma sample gave about 22,000 peptide hits (MS-1: 21294 peptide hits; MS-2: 23621 peptide hits; MS-3: 22342 peptide hits; and MS-4: 21991 peptide hits), which allowed comparisons between samples. As shown in Table 16, peptide hits for some of the 64 biomarker candidates were observed in the PDAC samples.

The most abundant proteins that were identified from the samples were Alpha-2-macroglobulin with 301 peptide hits, A2M with 2700 peptides hits,

Ceruloplasmin with 1495 peptides hits and Albumin with 1398 peptides hits. These results indicate that there were a number of proteins received the greatest number of peptide hits, as noted above, and dominated the spectrum leading to the masking of low abundance proteins in the samples. This can be due to the insufficient depletion
 5 of plasma proteins from the plasma samples.

Of the 64 candidate proteins, the following were detected in the two tested PDAC samples (MS3 and MS4): DNAH1, VCAM1 and MYLK (Table 16). Of the 64 candidate proteins, the following were detected in one of the PDAC case samples (MS3 or MS4): ACTN2, ATP2A1, DNAH5, TCF20, SYNE2, SYNE1,
 10 KIAA1109, ABCA13, ZNF804A, SCYL2 and KIAA1671 (Table 16). Of the four candidates that were identified to have c-statistics > 0.7 by ELISA, as described in Example 2 and Table 15, DNAH1, ATP2A1, SYNE1 were detected in at least one of the PDAC plasma samples (MS3 or MS4) (Table 16). These data provide independent experimental results indicating that the identified proteins can serve as
 15 biomarkers for pancreatic cancer.

Table 16

Gene	Pathway	Proteins	Peptide hits				c ≥ 0.7	c < 0.7	not tested
			MS1	MS2	MS3	MS4			
DNAH1	HNF4a	IPI00002127	0	1	2	9	X		
VCAM1	TGFb/integrin	IPI00018136	3	5	2	1			X
ACTN2	HNF4a	IPI00019884	0	0	0	1			X
ATP2A1	no network	IPI00024804	1	0	1	0	X		
DNAH5	HNF4a	IPI00152653	2	0	1	0		X	
TCF20	RAS/P53/JUN	IPI00159322	0	0	1	0		X	
MYLK	TGFb/integrin	IPI00221255	1	0	1	1			X
PMFBP1	RAS/P53/JUN	IPI00235481	1	0	0	0		X	
SYNE2	TGFb/integrin	IPI00239405	0	1	0	2			X
SYNE1	TGFb/integrin	IPI00247295	2	2	0	1	X		
KIAA1109	RAS/P53/JUN	IPI00251161	0	0	0	1			X
ABCA13	no network	IPI00328762	1	0	0	1			X
ZNF804A	HNF4a	IPI00375560	2	0	0	1			X
SCYL2	HNF4a	IPI00396218	0	3	2	0			X
KIAA1671	no network	IPI00396634	0	0	0	1			X
ODZ3	RAS/P53/JUN	IPI00398020	1	0	0	0		X	
DNAH12	HNF4a	IPI00412106	0	1	0	0	X		
DES	TGFb/integrin	IPI00465084	2	0	0	0			X
CA19-9		IPI00007671	0	0	0	0			
CA19-9		IPI00910296	0	0	0	0			

REFERENCES

- Banito, A., Rashid, S.T., Acosta, J.C., Li, S., Pereira, C.F., Geti, I., Pinho, S., Silva, J.C., Azuara, V., Walsh, M., *et al.* (2009). Senescence impairs successful reprogramming to pluripotent stem cells. *Genes Dev* 23, 2134-2139.
- 5 Bar-Nur, O., Russ, H.A., Efrat, S., and Benvenisty, N. (2011). Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet Beta cells. *Cell Stem Cell* 9, 17-23.
- Bardeesy, N., Cheng, K.H., Berger, J.H., Chu, G.C., Pahler, J., Olson, P., Hezel, A.F., Horner, J., Lauwers, G.Y., Hanahan, D., *et al.* (2006). Smad4 is
 10 dispensable for normal pancreas development yet critical in progression and tumor biology of pancreas cancer. *Genes Dev* 20, 3130-3146.
- Blelloch, R.H., Hochedlinger, K., Yamada, Y., Brennan, C., Kim, M., Mintz, B., Chin, L., and Jaenisch, R. (2004). Nuclear cloning of embryonal carcinoma cells. *Proc Natl Acad Sci U S A* 101, 13985-13990.
- 15 Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H., *et al.* (2011). Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* 144, 439-452.
- Brambrink et al.(2008) *Cell Stem Cell* 2, 151.
- 20 Brat, D.J., Lillemoe, K.D., Yeo, C.J., Warfield, P.B., and Hruban, R.H. (1998). Progression of pancreatic intraductal neoplasias to infiltrating adenocarcinoma of the pancreas. *Am J Surg Pathol* 22, 163-169.
- Buchholz, M., Braun, M., Heidenblut, A., Kestler, H.A., Kloppel, G., Schmiegel, W., Hahn, S.A., Luttges, J., and Gress, T.M. (2005). Transcriptome
 25 analysis of microdissected pancreatic intraepithelial neoplastic lesions. *Oncogene* 24, 6626-6636.
- Carette, J.E., Pruszk, J., Varadarajan, M., Blomen, V.A., Gokhale, S., Camargo, F.D., Wernig, M., Jaenisch, R., and Brummelkamp, T.R. (2010). Generation of iPSCs from cultured human malignant cells. *Blood* 115, 4039-4042.
- 30 Ding, Y., Cravero, J.D., Adrian, K., and Grippo, P. (2010). Modeling pancreatic cancer in vivo: from xenograft and carcinogen-induced systems to genetically engineered mice. *Pancreas* 39, 283-292.
- Eiselleova, L., Matulka, K., Kriz, V., Kunova, M., Schmidtova, Z., Neradil, J., Tichy, B., Dvorakova, D., Pospisilova, S., Hampl, A., *et al.* (2009). A complex role

for FGF-2 in self renewal, survival, and adhesion of human embryonic stem cells. *Stem Cells* 27, 1847-1857.

Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* 8, 286-298.

- 5 Gattani A M, Mandeli J, Bruckner H W. (1996). Tumor markers in patients with pancreatic carcinoma. *Cancer* 78: 57-62.

- Gupta, R.K., Gao, N., Gorski, R.K., White, P., Hardy, O.T., Rafiq, K., Brestelli, J.E., Chen, G., Stoeckert, C.J., Jr., and Kaestner, K.H. (2007). Expansion of adult beta-cell mass in response to increased metabolic demand is dependent on HNF-
10 4alpha. *Genes Dev* 21, 756-769.

Hahn, S.A., Schutte, M., Hoque, A.T., Moskaluk, C.A., da Costa, L.T., Rozenblum, E., Weinstein, C.L., Fischer, A., Yeo, C.J., Hruban, R.H., *et al.* (1996). DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* 271, 350-353.

- 15 Hahn, S.A., Seymour, A.B., Hoque, A.T., Schutte, M., da Costa, L.T., Redston, M.S., Caldas, C., Weinstein, C.L., Fischer, A., Yeo, C.J., *et al.* (1995). Allelotype of pancreatic adenocarcinoma using xenograft enrichment. *Cancer Res* 55, 4670-4675.

- Harsha, H.C., Kandasamy, K., Ranganathan, P., Rani, S., Ramabadran, S.,
20 Gollapudi, S., Balakrishnan, L., Dwivedi, S.B., Telikicherla, D., Selvan, L.D., *et al.* (2009). A compendium of potential biomarkers of pancreatic cancer. *PLoS Med* 6, e1000046.

- Hatziapostolou, M., Polytarchou, C., Aggelidou, E., Drakaki, A., Poultsides, G.A., Jaeger, S.A., Ogata, H., Karin, M., Struhl, K., Hadzopoulou-Cladaras, M., *et al.*
25 (2011). An HNF4alpha-miRNA inflammatory feedback circuit regulates hepatocellular oncogenesis. *Cell* 147, 1233-1247.

- Hermann, P.C., Huber, S.L., Herrler, T., Aicher, A., Ellwart, J.W., Guba, M., Bruns, C.J., and Heeschen, C. (2007). Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell*
30 *Stem Cell* 1, 313-323.

Hezel, A.F., Deshpande, V., Zimmerman, S.M., Contino, G., Alagesan, B., O'Dell, M.R., Rivera, L.B., Harper, J., Lonning, S., Brekken, R.A., *et al.* (2012). TGF-beta and alphavbeta6 Integrin Act in a Common Pathway to Suppress Pancreatic Cancer Progression. *Cancer Res* 72, 4840-4845.

Hezel, A.F., Kimmelman, A.C., Stanger, B.Z., Bardeesy, N., and Depinho, R.A. (2006). Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev* 20, 1218-1249.

Hingorani, S.R., Petricoin, E.F., Maitra, A., Rajapakse, V., King, C., Jacobetz, M.A., Ross, S., Conrads, T.P., Veenstra, T.D., Hitt, B.A., et al. (2003). Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell* 4, 437-450.

Hochedlinger, K., Belloch, R., Brennan, C., Yamada, Y., Kim, M., Chin, L., and Jaenisch, R. (2004). Reprogramming of a melanoma genome by nuclear transplantation. *Genes Dev* 18, 1875-1885.

Hruban, R.H., Adsay, N.V., Albores-Saavedra, J., Compton, C., Garrett, E.S., Goodman, S.N., Kern, S.E., Klimstra, D.S., Kloppel, G., Longnecker, D.S., et al. (2001). Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions. *Am J Surg Pathol* 25, 579-586.

Iacobuzio-Donahue, C.A., Maitra, A., Olsen, M., Lowe, A.W., van Heek, N.T., Rosty, C., Walter, K., Sato, N., Parker, A., Ashfaq, R., et al. (2003). Exploration of global gene expression patterns in pancreatic adenocarcinoma using cDNA microarrays. *Am J Pathol* 162, 1151-1162.

Ishizawa, K., Rasheed, Z.A., Karisch, R., Wang, Q., Kowalski, J., Susky, E., Pereira, K., Karamboulas, C., Moghal, N., Rajeshkumar, N.V., et al. (2010). Tumor-initiating cells are rare in many human tumors. *Cell Stem Cell* 7, 279-282.

Jones, S., Zhang, X., Parsons, D.W., Lin, J.C., Leary, R.J., Angenendt, P., Mankoo, P., Carter, H., Kamiyama, H., Jimeno, A., et al. (2008). Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* 321, 1801-1806.

Kanda et al. (2012) *Gastroenterology* 142, 730.

Kim, G.E., Bae, H.I., Park, H.U., Kuan, S.F., Crawley, S.C., Ho, J.J., and Kim, Y.S. (2002). Aberrant expression of MUC5AC and MUC6 gastric mucins and sialyl Tn antigen in intraepithelial neoplasms of the pancreas. *Gastroenterology* 123, 1052-1060.

Kim, K., Zhao, R., Doi, A., Ng, K., Unternaehrer, J., Cahan, P., Huo, H., Loh, Y.H., Aryee, M.J., Lensch, M.W., et al. (2011). Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat Biotechnol* 29, 1117-1119.

Kim, M.P., Evans, D.B., Wang, H., Abbruzzese, J.L., Fleming, J.B., and Gallick, G.E. (2009). Generation of orthotopic and heterotopic human pancreatic cancer xenografts in immunodeficient mice. *Nat Protoc* 4, 1670-1680.

Kopp, J.L., von Figura, G., Mayes, E., Liu, F.F., Dubois, C.L., Morris, J.P.t.,
 5 Pan, F.C.,
 Akiyama, H., Wright, C.V., Jensen, K., *et al.* (2012). Identification of Sox9-Dependent Acinar-to-Ductal Reprogramming as the Principal Mechanism for Initiation of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* 22, 737-750.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S.,
 10 and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* 134, 2895-2902.

Lee et al. (2007) *Nature Methods* 4, 359.

Li, C., Heidt, D.G., Dalerba, P., Burant, C.F., Zhang, L., Adsay, V., Wicha, M.,
 15 Clarke, M.F., and Simeone, D.M. (2007). Identification of pancreatic cancer stem cells. *Cancer Res* 67, 1030-1037.

Li, L., Connelly, M.C., Wetmore, C., Curran, T., and Morgan, J.I. (2003). Mouse embryos cloned from brain tumors. *Cancer Res* 63, 2733-2736.

Lieber, M., Mazzeita, J., Nelson-Rees, W., Kaplan, M., and Todaro, G. (1975).
 20 Establishment of a continuous tumor-cell line (panc-1) from a human carcinoma of the exocrine pancreas. *Int J Cancer* 15, 741-747.

Logsdon, C.D., Simeone, D.M., Binkley, C., Arumugam, T., Greenson, J.K., Giordano, T.J., Misek, D.E., Kuick, R., and Hanash, S. (2003). Molecular profiling of pancreatic adenocarcinoma and chronic pancreatitis identifies multiple genes
 25 differentially regulated in pancreatic cancer. *Cancer Res* 63, 2649-2657.

Lonardo, E., Hermann, P.C., Mueller, M.T., Huber, S., Balic, A., Miranda-Lorenzo, I., Zagorac, S., Alcala, S., Rodriguez-Arabaolaza, I., Ramirez, J.C., *et al.*
 (2011). Nodal/Activin signaling drives self-renewal and tumorigenicity of pancreatic cancer stem cells and provides a target for combined drug therapy. *Cell Stem Cell* 9,
 30 433-446.

Lonardo E, H.P., Mueller MT, Huber S, Balic A, Miranda-Lorenzo I, Zagorac S, Alcala S, Rodriguez-Arabaolaza I, Ramirez JC, Torres-Ruiz R, Garcia E, Hidalgo M, Cebrián DÁ, Heuchel R, Löhr M, Berger F, Bartenstein P, Aicher A, Heeschen C. (2011). Nodal/Activin signaling drives self-renewal and tumorigenicity of pancreatic

cancer stem cells and provides a target for combined drug therapy. *Cell Stem Cell* 9, 433-446.

Maitra, A., and Hruban, R.H. (2008). Pancreatic cancer. *Annu Rev Pathol* 3, 157-188.

5 Maser, R.S., Choudhury, B., Campbell, P.J., Feng, B., Wong, K.K., Protopopov, A., O'Neil, J., Gutierrez, A., Ivanova, E., Perna, I., *et al.* (2007). Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. *Nature* 447, 966-971.

10 McKinnell, R.G., Deggins, B.A., and Labat, D.D. (1969). Transplantation of pluripotential nuclei from triploid frog tumors. *Science* 165, 394-396.

Miyatsuka, T., Kaneto, H., Shiraiwa, T., Matsuoka, T.A., Yamamoto, K., Kato, K., Nakamura, Y., Akira, S., Takeda, K., Kajimoto, Y., *et al.* (2006). Persistent expression of PDX-1 in the pancreas causes acinar-to-ductal metaplasia through Stat3 activation. *Genes Dev* 20, 1435-1440.

15 Miyoshi, N., Ishii, H., Nagai, K., Hoshino, H., Mimori, K., Tanaka, F., Nagano, H., Sekimoto, M., Doki, Y., and Mori, M. (2010). Defined factors induce reprogramming of gastrointestinal cancer cells. *Proc Natl Acad Sci U S A* 107, 40-45.

20 Morris, J.P.t., Wang, S.C., and Hebrik, M. (2010). KRAS, Hedgehog, Wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma. *Nat Rev Cancer* 10, 683-695.

Moskaluk, C.A., Hruban, R.H., and Kern, S.E. (1997). p16 and K-ras gene mutations in the intraductal precursors of human pancreatic adenocarcinoma. *Cancer Res* 57, 2140-2143.

25 Odom, D.T., Zizlsperger, N., Gordon, D.B., Bell, G.W., Rinaldi, N.J., Murray, H.L., Volkert, T.L., Schreiber, J., Rolfe, P.A., Gifford, D.K., *et al.* (2004). Control of pancreas and liver gene expression by HNF transcription factors. *Science* 303, 1378-1381.

Park *et al.* (2008) *Cell* 134, 877.

30 Rhim, A.D., Mirek, E.T., Aiello, N.M., Maitra, A., Bailey, J.M., McAllister, F., Reichert, M., Beatty, G.L., Rustgi, A.K., Vonderheide, R.H., *et al.* (2012). EMT and dissemination precede pancreatic tumor formation. *Cell* 148, 349-361.

Rubio-Viqueira, B., Jimeno, A., Cusatis, G., Zhang, X., Iacobuzio-Donahue, C., Karikari, C., Shi, C., Danenberg, K., Danenberg, P.V., Kuramochi, H., *et al.*

(2006). An in vivo platform for translational drug development in pancreatic cancer. *Clin Cancer Res* 12, 4652-4661.

Rustgi, A.K. (2006). The molecular pathogenesis of pancreatic cancer: clarifying a complex circuitry. *Genes Dev* 20, 3049-3053.

5 Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D., and Lowe, S.W. (1997). Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* 88, 593-602.

Shultz, L.D., Lyons, B.L., Burzenski, L.M., Gott, B., Chen, X., Chaleff, S., Kotb, M., Gillies, S.D., King, M., Mangada, J., *et al.* (2005). Human lymphoid and
10 myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *J Immunol* 174, 6477-6489.

Strader (2006) *Anal. Chem* 78, 125.

Taguchi, A., Politi, K., Pitteri, S.J., Lockwood, W.W., Faca, V.M., Kelly-Spratt, K., Wong, C.H., Zhang, Q., Chin, A., Park, K.S., *et al.* (2011). Lung cancer
15 signatures in plasma based on proteome profiling of mouse tumor models. *Cancer Cell* 20, 289-299.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.

20 Thomson et al.(1998) *Science* 282, 1145.

Tost and Gut (2007) *Nat Protocols* 2, 2265.

Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., *et al.* (2010). Towards a knowledge-based Human Protein Atlas. *Nat Biotechnol* 28, 1248-1250.

25 Urlinger et al. (2000) *Proc Natl Acad Sci U S A* 97, 7963.

Yunis, A.A., Arimura, G.K., and Russin, D.J. (1977). Human pancreatic carcinoma (MIA PaCa-2) in continuous culture: sensitivity to asparaginase. *Int J Cancer* 19, 128-135.

30 Various publications, patents and patent applications are cited herein, the contents of which are hereby incorporated by reference herein in their entireties.

WHAT IS CLAIMED IS:

1. A method of determining whether a subject has pancreatic cancer, comprising obtaining one or more biological samples from the subject and detecting, in the one or more biological samples, one or more biomarkers selected from the group consisting of MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13, DES, IMMT, TPM1, SNRPE, VCAM1, GRB2, SHROOM3, HMOX1, POSTN, MMP10, MMP-2, THBS2, EWSR1, NOD1, ADAMTS9, AFP, SYNE1, SYNE2, EPHB1, UFD1L, TEAD1, RYR3, CMYA5, MYLK, TOP2B, KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29, OBSCN, LOXL3, MLEC, DNAH1, DNAH5, DNAH12, DNAH17, SCYL2, FKBP10, FLRT3, ZHX2 (AFR1), ZNF804A, ACTN2 and combinations thereof, wherein the detection of the one or more biomarkers is an indication that the subject has pancreatic cancer.
2. A method of determining whether a subject has pancreatic cancer, comprising obtaining one or more biological samples from the subject and detecting, in the one or more biological samples, one or more biomarkers selected from the group consisting of RTTN, DNAH12, TPM1, DNAH1, STARD8, ATP2A1, TOP2B, LIMCH1, SYNE1, THBS2, LOXL3 and combinations thereof, wherein the presence of the one or more biomarkers is an indication that the subject has pancreatic cancer.
3. A method of determining whether a subject has pancreatic cancer, the method comprising obtaining a biological sample from the subject and determining if the biological sample from the subject contains at least three biomarkers, wherein one of the at least three biomarkers is a member of the TGF β /integrin signaling pathway, one of the at least three biomarkers is a member of the HNF4 α transcription factor network and the one of the at least three biomarkers is a member of the Ras/p53/JUN/CTNB1 signaling pathway, and wherein the presence of the at least three biomarkers in the sample is an indication that the subject has pancreatic cancer.
4. The method of claim 1, 2 or 3, wherein the subject is human.
5. The method of claim 1, 2 or 3, wherein the biological sample is selected from the group consisting of a stool sample, a blood sample, a plasma sample, a pancreatic cyst fluid sample and combinations thereof.
6. The method of claim 1, 2 or 3, wherein the biomarker is a protein.

7. The method of claim 6, wherein the presence of the protein biomarker is detected using a reagent which specifically binds to the protein biomarker.

8. The method of claim 7, wherein the reagent is a monoclonal antibody or antigen-binding fragment thereof, or a polyclonal antibody or antigen-binding fragment thereof.

9. The method of claim 6, wherein the presence of the protein is detected by enzyme-linked immunosorbent assay.

10. The method of claims 1, 2 or 3, wherein the biomarker comprises a transcribed polynucleotide or portion thereof.

11. A kit for diagnosing whether a subject has pancreatic cancer or assessing the efficacy of a therapeutic treatment, comprising reagents useful for detecting one or more biomarkers selected from the group consisting of MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13, DES, IMMT, TPM1, SNRPE, VCAM1, GRB2, SHROOM3, HMOX1, POSTN, MMP10, MMP-2, THBS2, EWSR1, NOD1, ADAMTS9, AFP, SYNE1, SYNE2, EPHB1, UFD1L, TEAD1, RYR3, CMYA5, MYLK, TOP2B, KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29, OBSCN, LOXL3, MLEC, DNAH1, DNAH5, DNAH12, DNAH17, SCYL2, FKBP10, FLRT3, ZHX2 (AFR1), ZNF804A, ACTN2 and combinations thereof, in one or more biological samples from the subject.

12. The kit of claim 11, comprising one or more of packaged probe and primer sets, arrays/microarrays, biomarker-specific antibodies or beads for detecting the one or more biomarkers.

13. The kit of claim 11, comprising at least one monoclonal antibody or antigen-binding fragment thereof, or a polyclonal antibody or antigen-binding fragment thereof, for detecting the one or more biomarkers to be identified.

14. A method of determining whether a subject has pancreatic cancer comprising, obtaining a biological sample from the subject and determining if the biological sample from the subject contains at least two biomarkers, wherein one of the at least two biomarkers is a member of the TGF β /integrin signaling pathway and the other biomarker of the at least two biomarkers is a member of the HNF4 α transcription factor network, and wherein the presence of the at least two biomarkers in the sample is an indication that the subject has pancreatic cancer.

15. A method of determining whether a subject has pancreatic cancer, the method comprising obtaining a biological sample from the subject and determining if the biological sample from the subject contains at least two biomarkers, wherein one of the at least two biomarkers is a member of the Ras/p53/JUN/CTNB1 signaling pathway and the other biomarker of the at least two biomarkers is a member of the HNF4 α transcription factor network, and wherein the presence of the at least two biomarkers in the sample is an indication that the subject has pancreatic cancer.

16. A method of determining whether a subject has pancreatic cancer, the method comprising obtaining a biological sample from the subject and determining if the biological sample from the subject contains at least two biomarkers, wherein one of the at least two biomarkers is a member of a pathway selected from the group consisting of the Ras/p53/JUN/CTNB1 signaling pathway, the HNF4 α transcription factor network, the TGF β /integrin signaling pathway and combinations thereof and the other biomarker of the at least two biomarkers is selected from the group consisting of MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13 and combinations thereof, wherein the presence of the at least two biomarkers in the sample is an indication that the subject has pancreatic cancer.

17. A method of assessing the efficacy of a therapy for preventing or treating pancreatic cancer in a subject, comprising:

(a) determining the level of one or more biomarkers in a biological sample obtained from the subject, wherein the biomarkers are selected from the group consisting of MANF, ZNF485, IMPA1, SVEP1, KIAA1671, KIAA1529, GNN, DOS, STARD8 (DLC3), SCN8A, U2SURP, TCHP, IPI00026665, RAD51C, ATP2A1, NLRX1, ZNF160, RTTN, ABCA13, DES, IMMT, TPM1, SNRPE, VCAM1, GRB2, SHROOM3, HMOX1, POSTN, MMP10, MMP-2, THBS2, EWSR1, NOD1, ADAMTS9, AFP, SYNE1, SYNE2, EPHB1, UFD1L, TEAD1, RYR3, CMYA5, MYLK, TOP2B, KIAA1109, ODZ3, PMFBP1, EPHB3, LIMCH1, TCF20, ERP29, OBSCN, LOXL3, MLEC, DNAH1, DNAH5, DNAH12, DNAH17, SCYL2, FKBP10, FLRT3, ZHX2 (AFR1), ZNF804A, ACTN2 and combinations thereof, prior to therapy; and

(b) determining the level of the one or more biomarkers in a biological sample obtained from the subject, at one of more time points during therapy,

wherein the therapy is efficacious for preventing or treating the cancer in the subject when there is a lower level of the one or more biomarkers in the second or subsequent samples, relative to the first sample.

18. A method for producing a pancreatic tumor cell model system comprising:

- (a) isolating pancreatic ductal adenocarcinoma cells from a pancreatic ductal adenocarcinoma sample;
- (b) overexpressing Klf4, Sox2, Oct4 and c-Myc in the isolated pancreatic ductal adenocarcinoma cells to produce induced pluripotent stem cells;
- (c) injecting the induced pluripotent stem cells into an immunocompromised animal to produce a teratoma in the animal;
- (d) isolating the teratoma from the animal; and
- (e) culturing the teratoma to obtain the pancreatic tumor cell model system.

19. A pancreatic tumor cell model system generated by the method of claim 18.

20. The use of the pancreatic tumor cell model system of claim 19 to identify one or more biomarkers for pancreatic cancer.

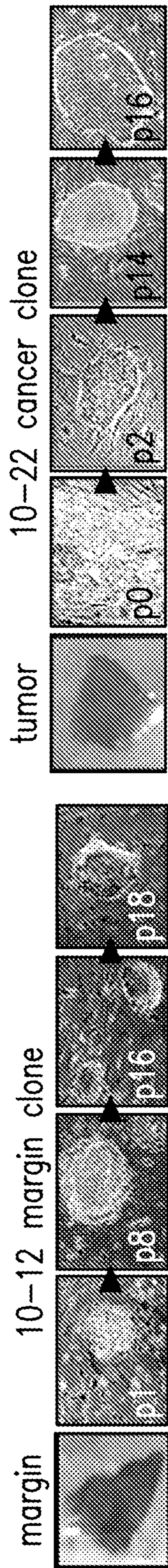


FIG. 1A

1/100

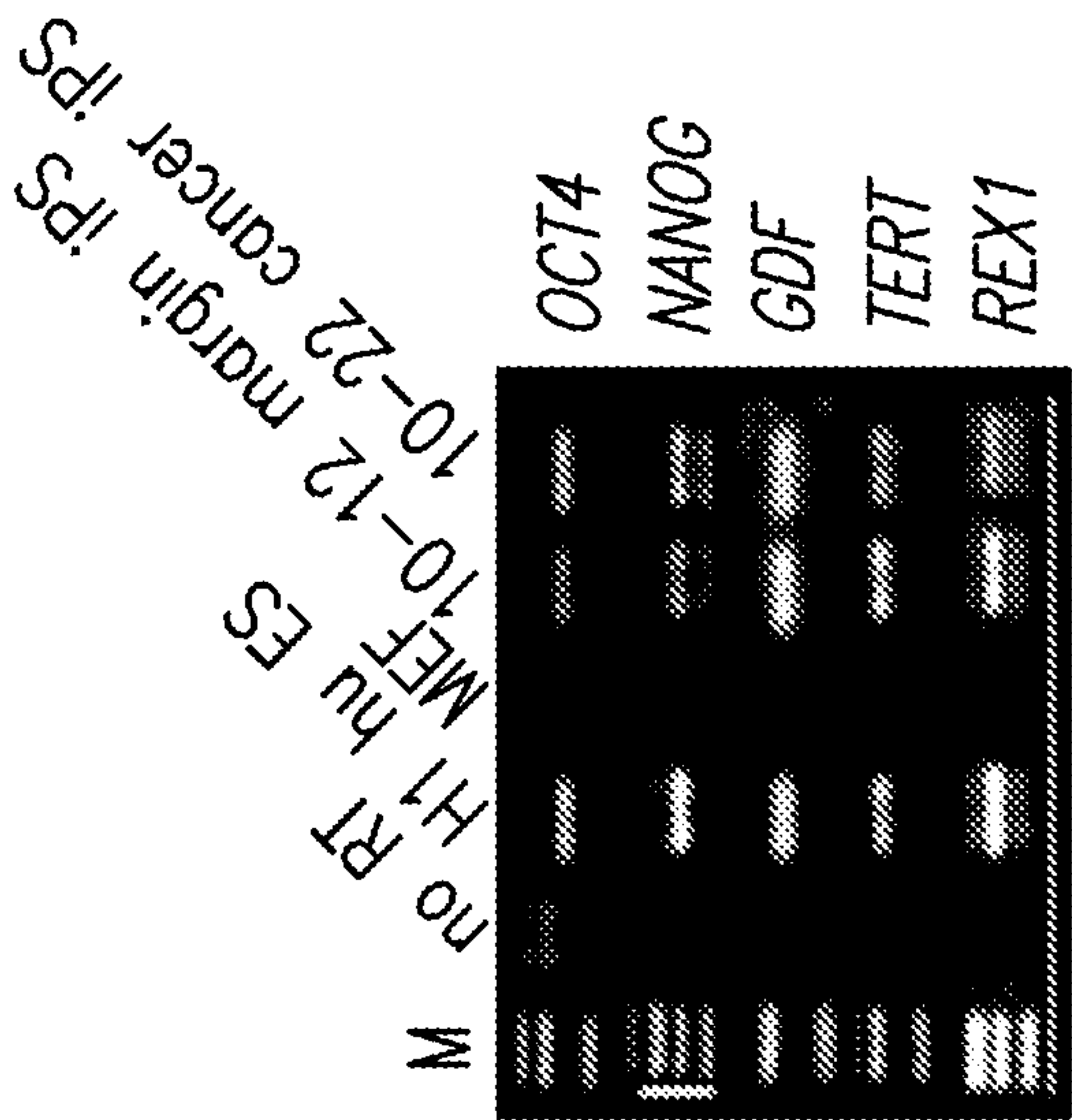


FIG. 1C

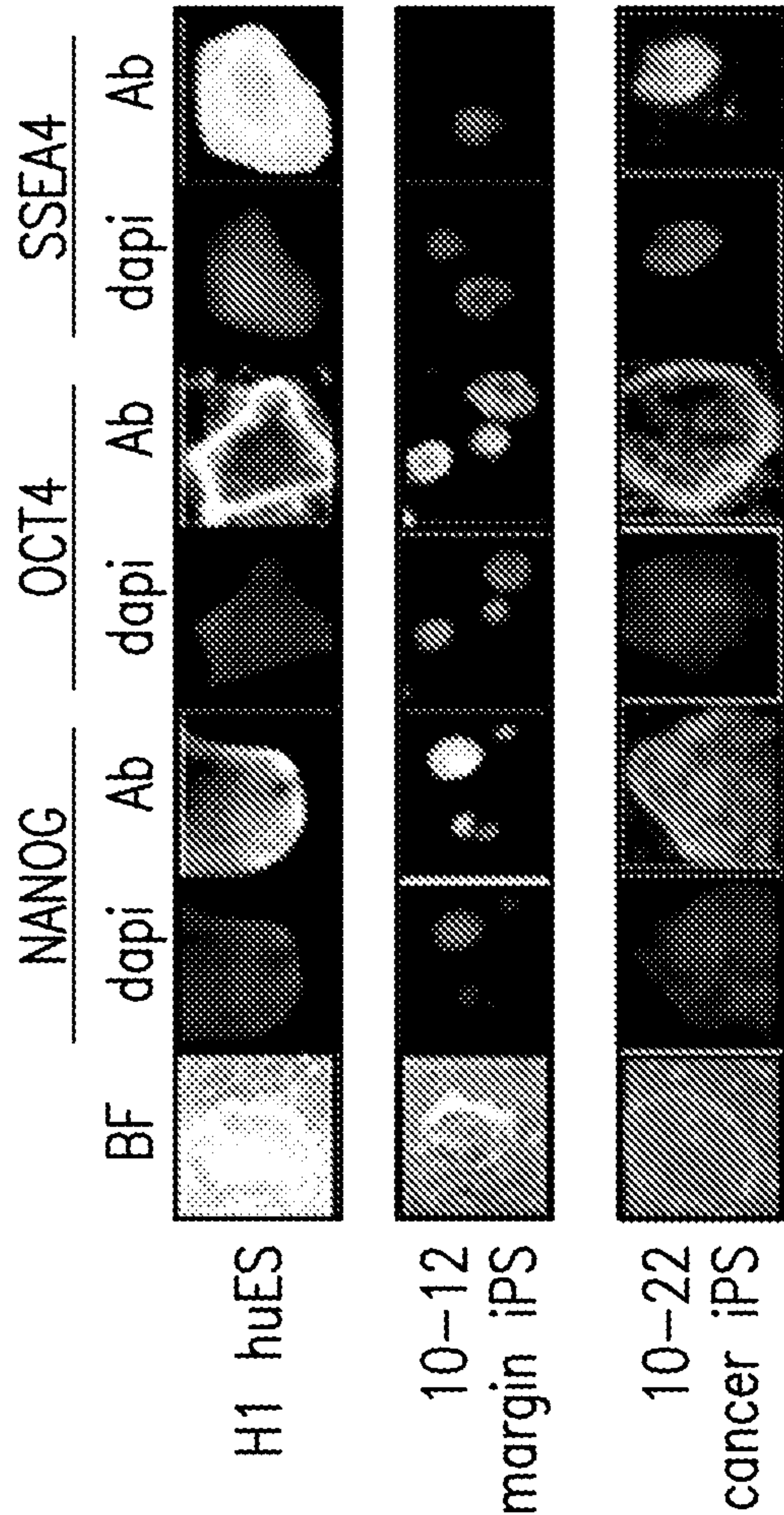


FIG. 1B

2/100

teratoma assays

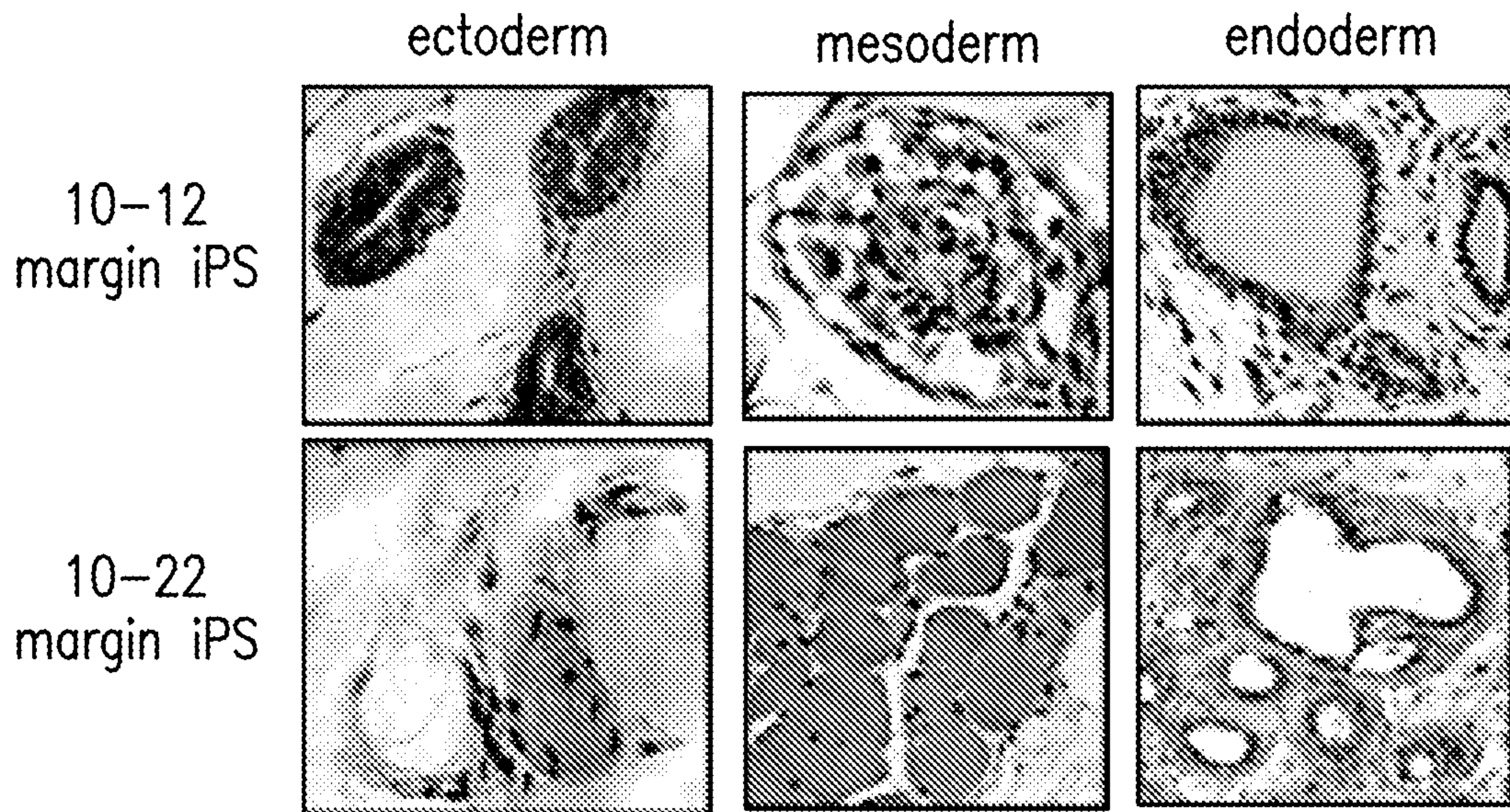


FIG. 1D

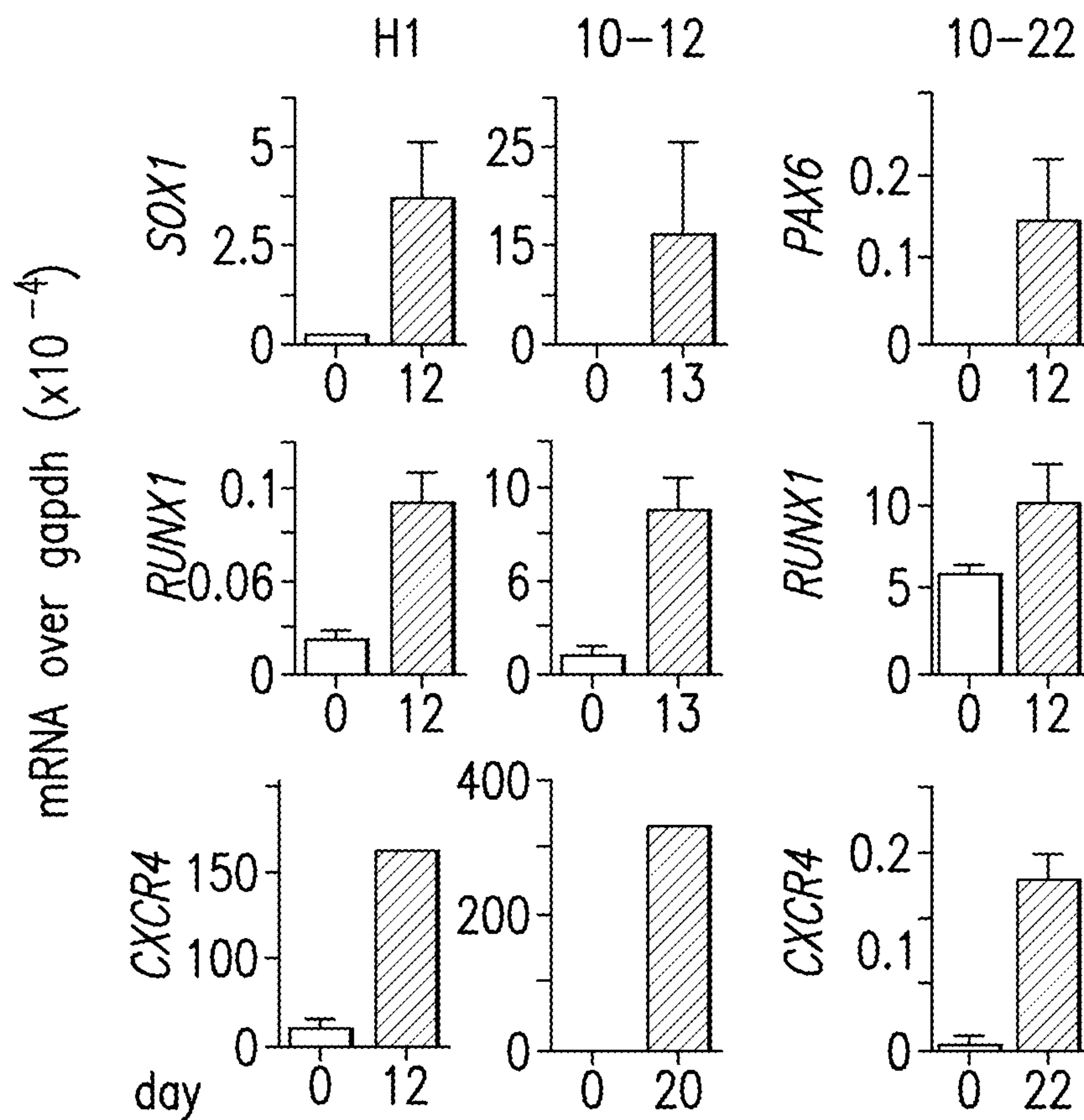
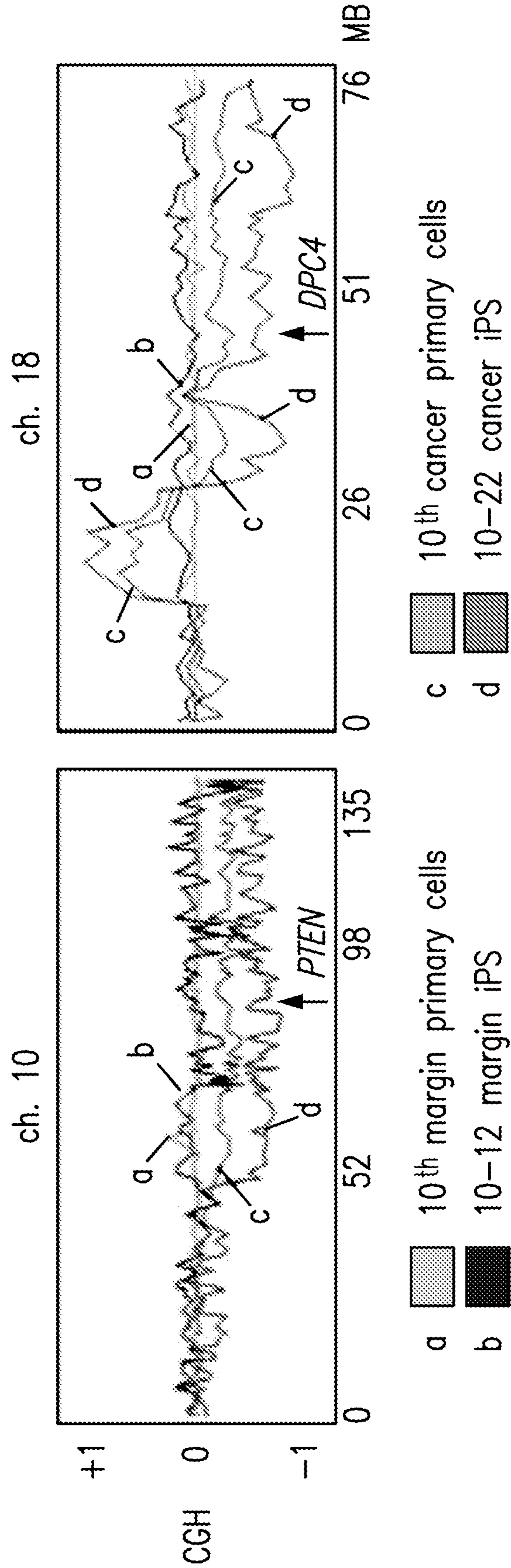
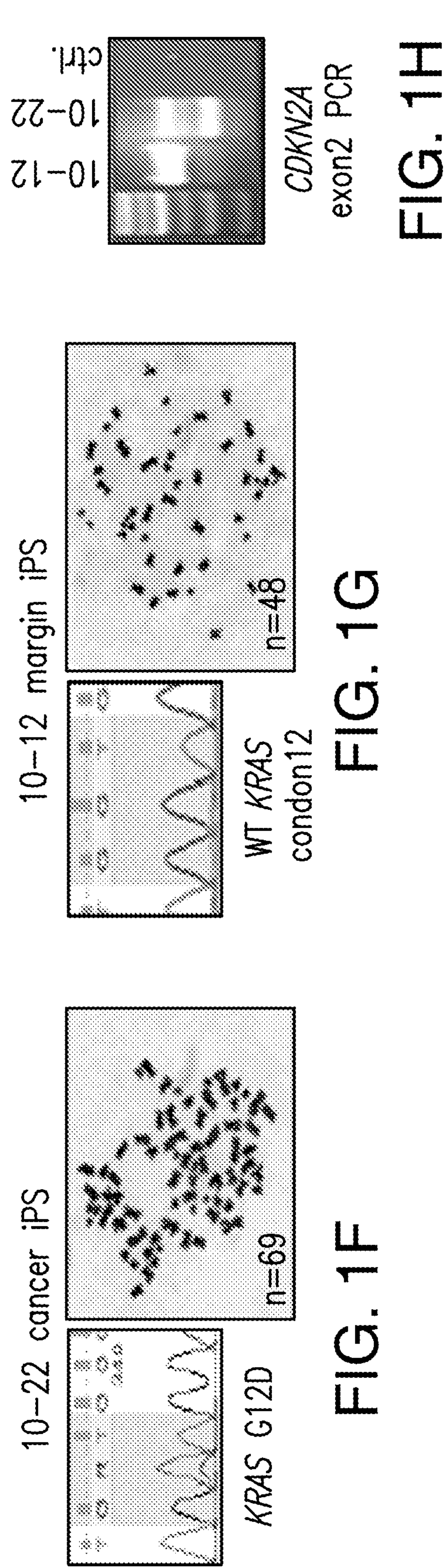


FIG. 1E



4/100

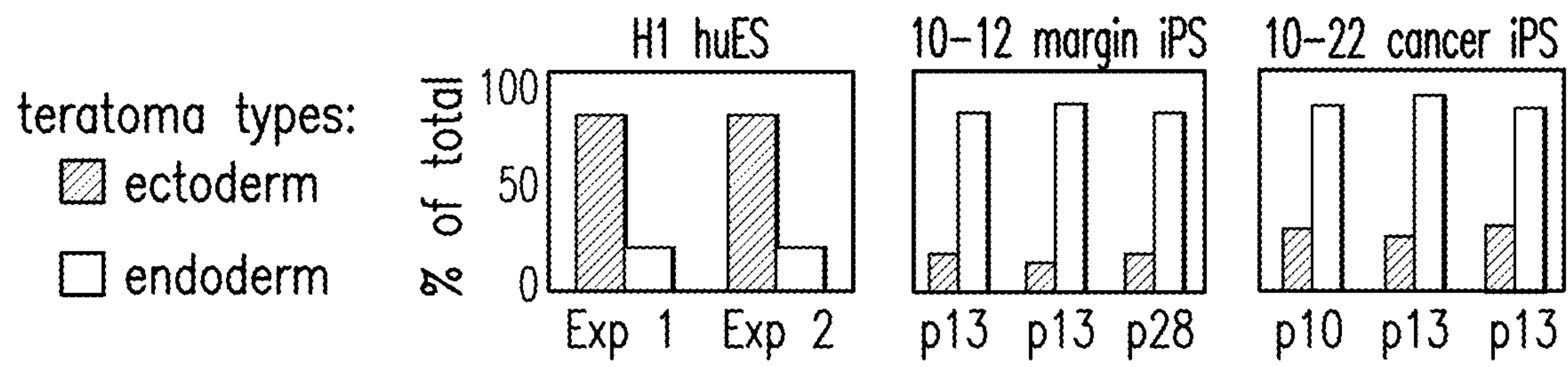
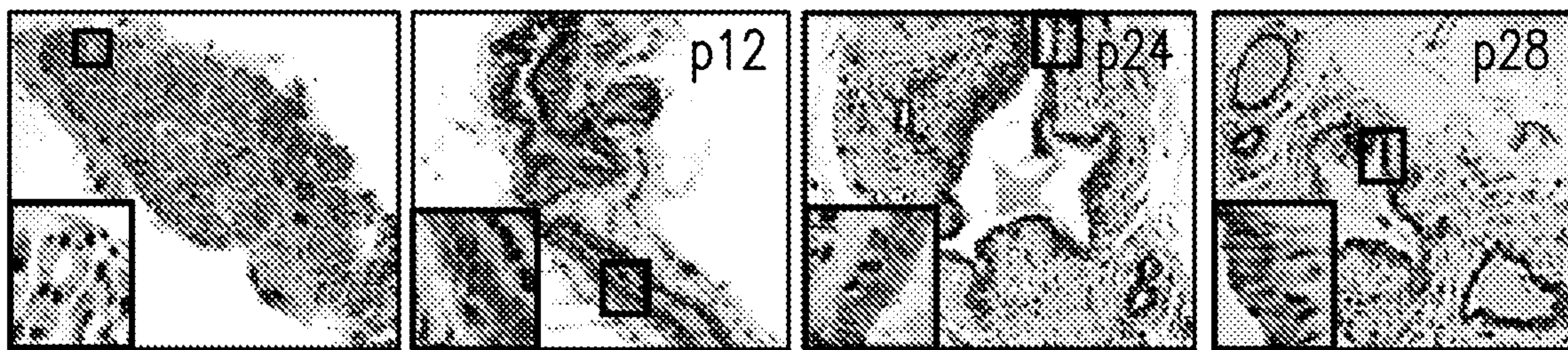


FIG. 2A

10th primary margin independent 10-12 margin iPS teratomas (3 mo.)



10th primary tumor independent 10-22 cancer iPS teratomas (3 mo.)

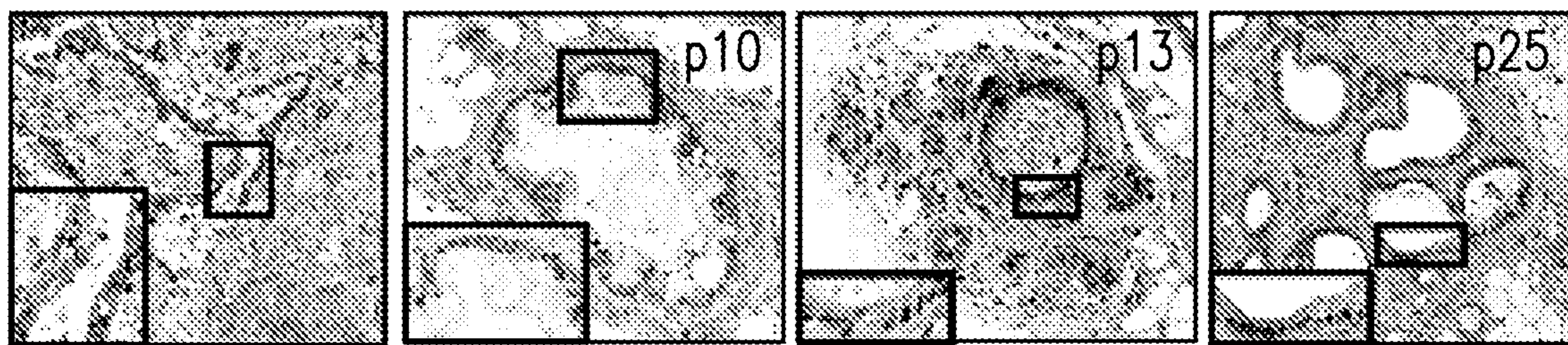


FIG. 2B

10th primary tumor



FIG. 2C

10-22 cancer iPS teratomas (3 mo.)

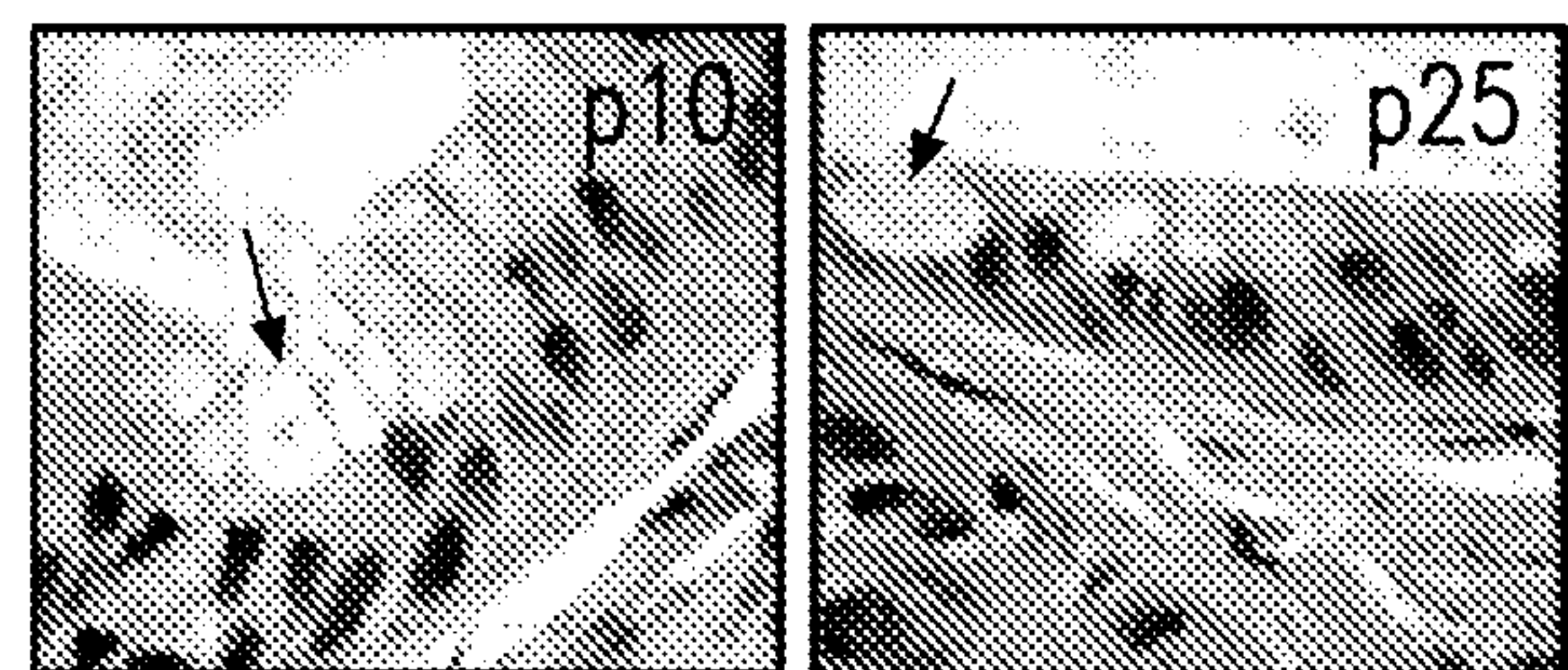
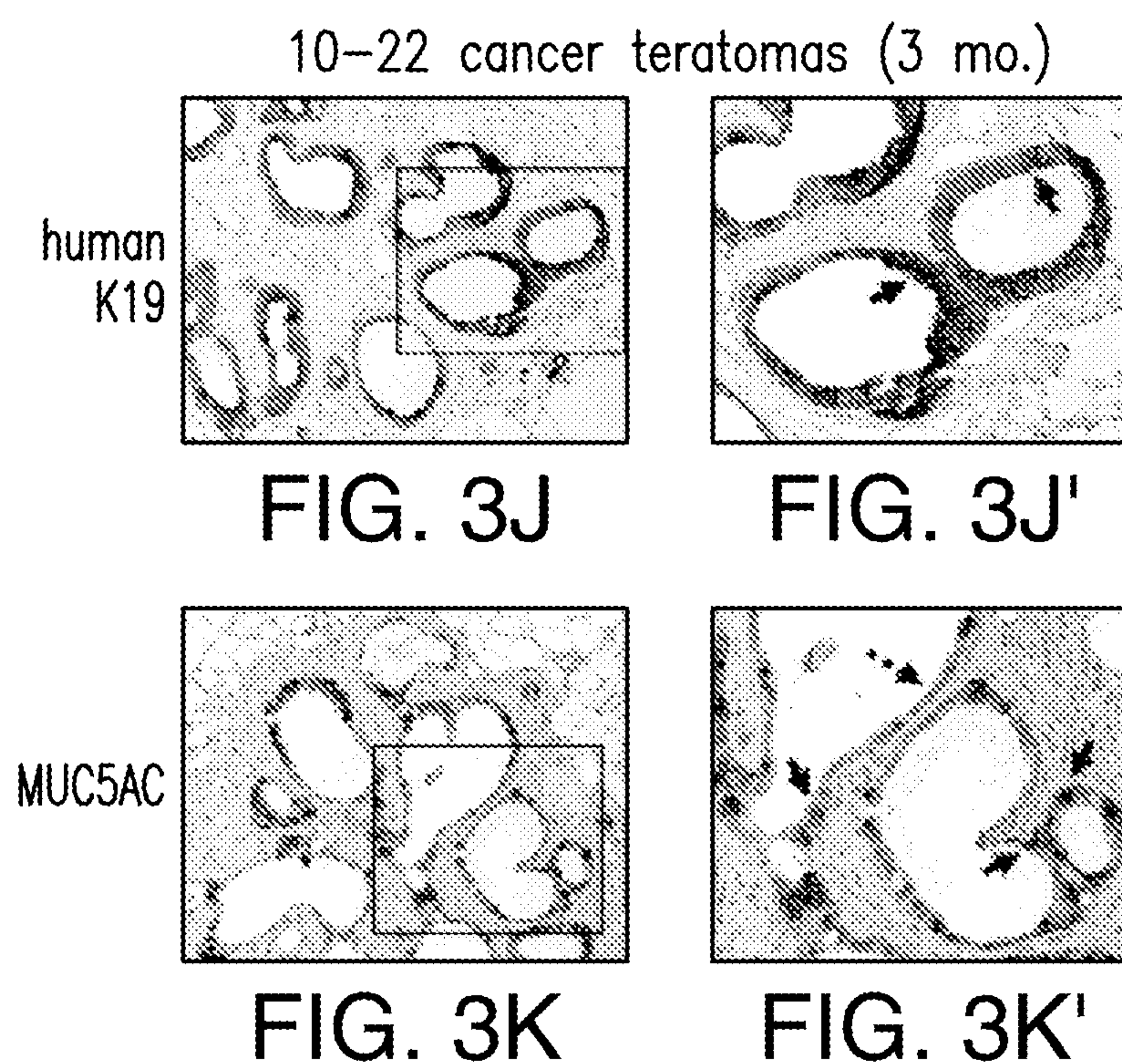
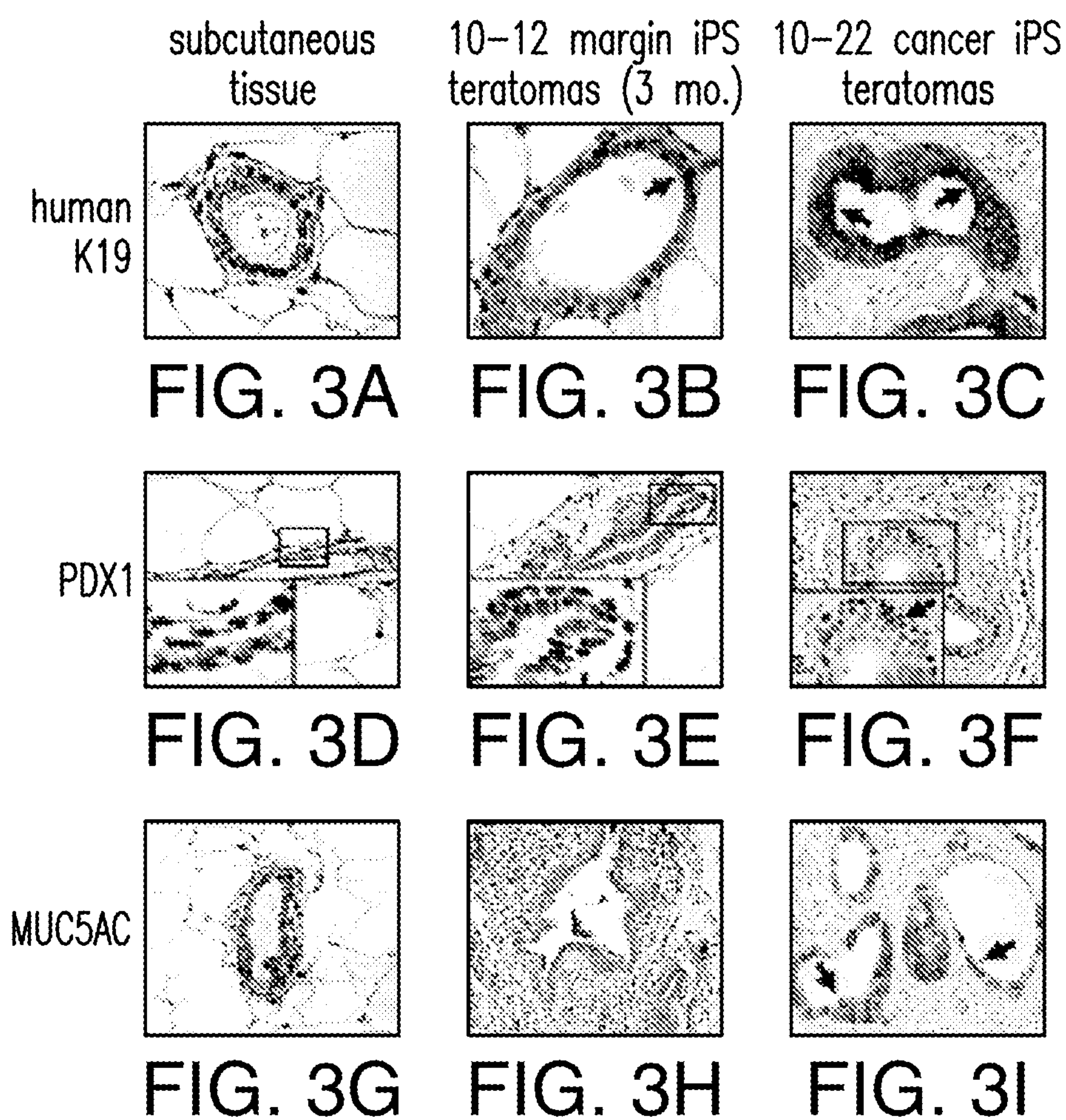


FIG. 2D

5/100



6/100

H&E

10-22 cancer iPS
teratoma
at 6 months
(mouse 9354)

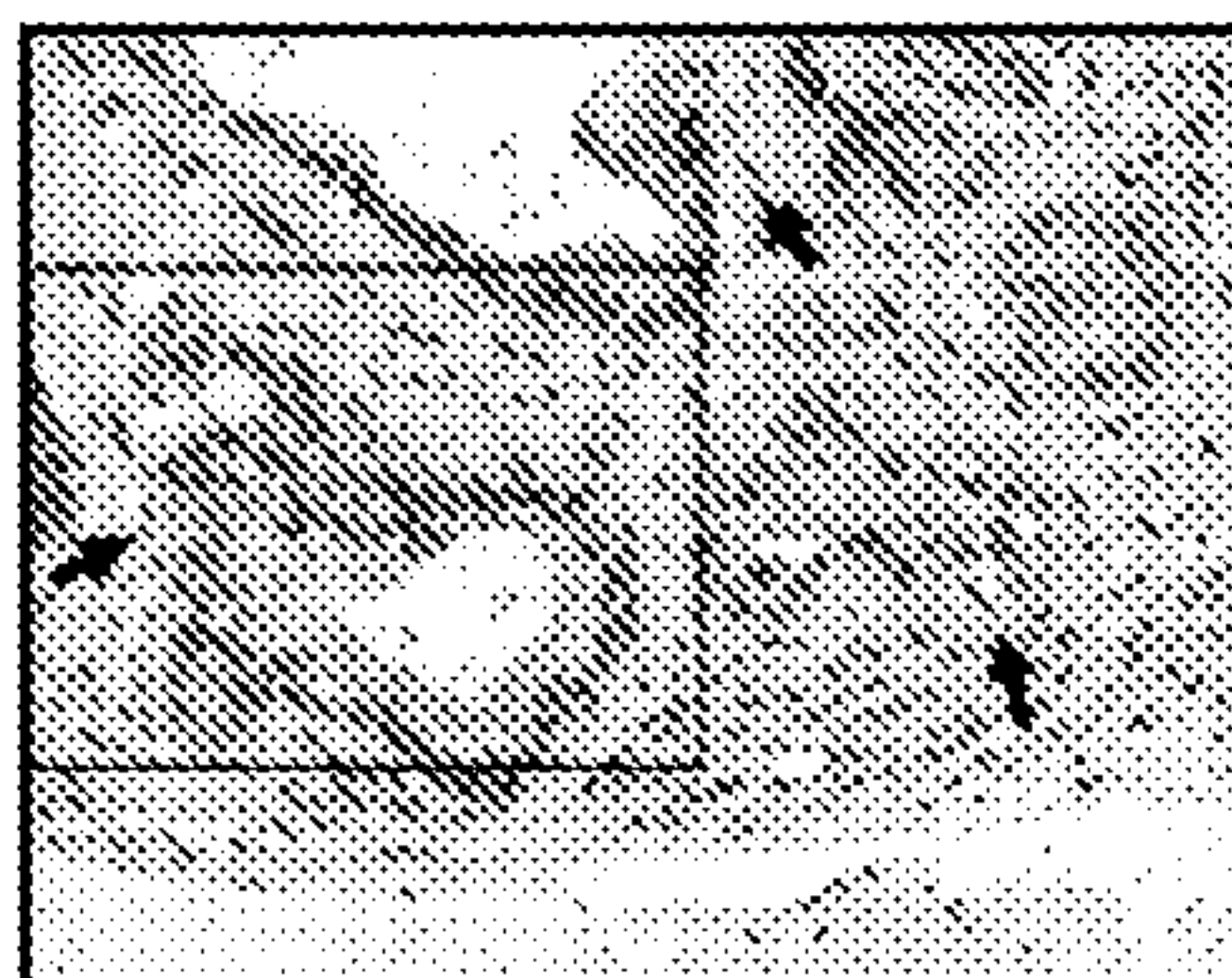


FIG. 4A



FIG. 4A'

10-22 cancer iPS
teratoma
at 9 months
(mouse 9351)
(tumor #1)

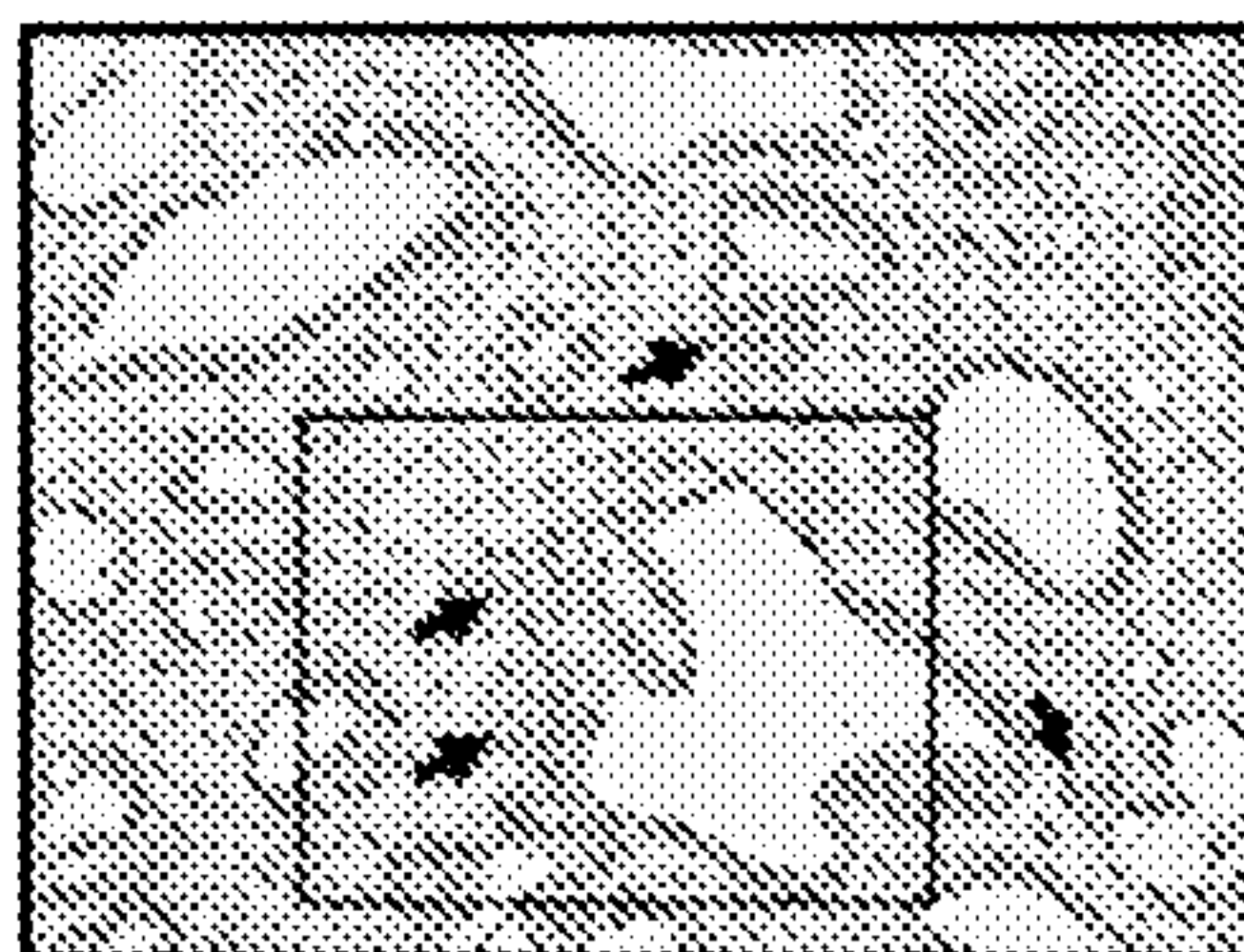
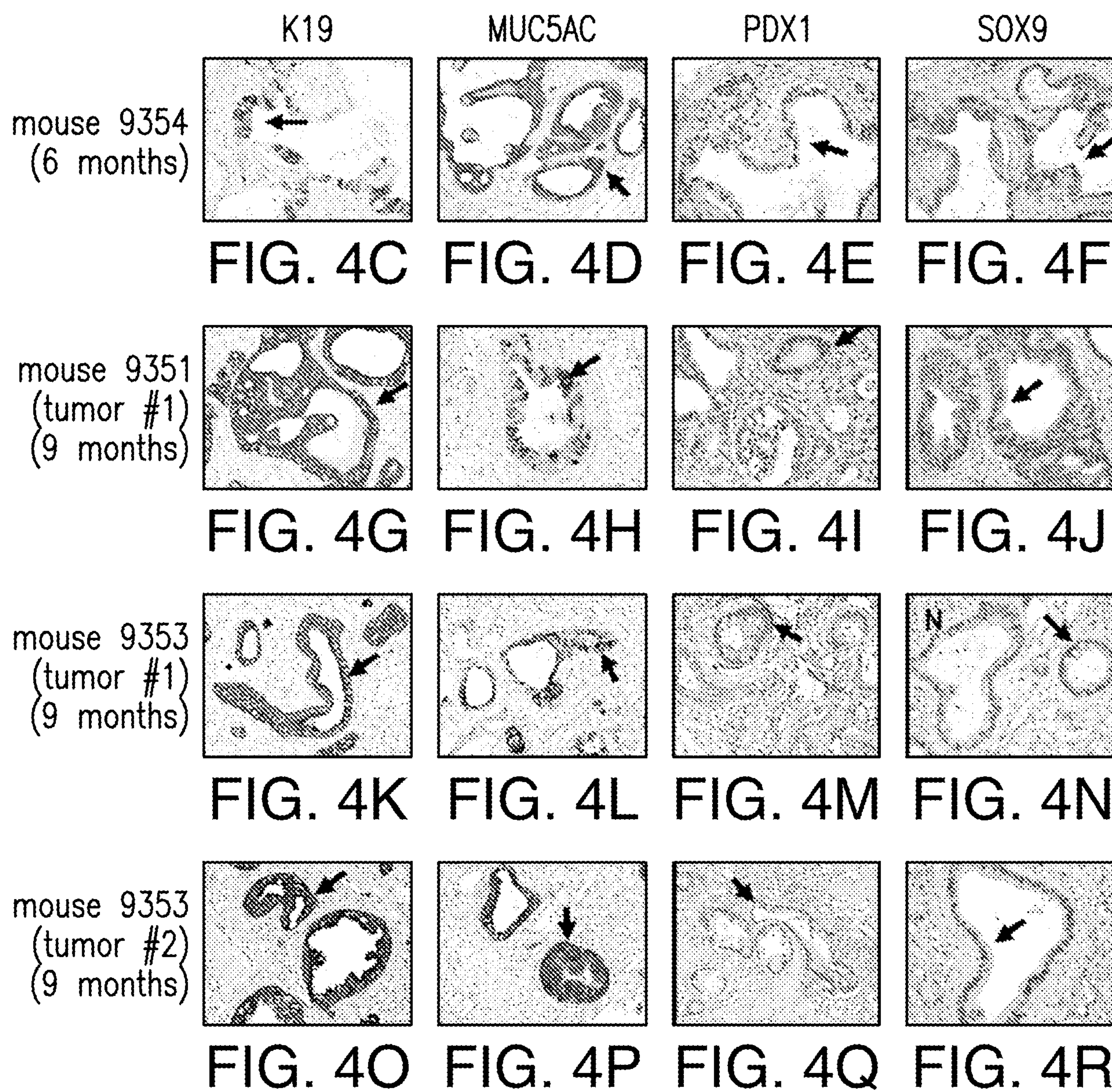


FIG. 4B



FIG. 4B'

7/100



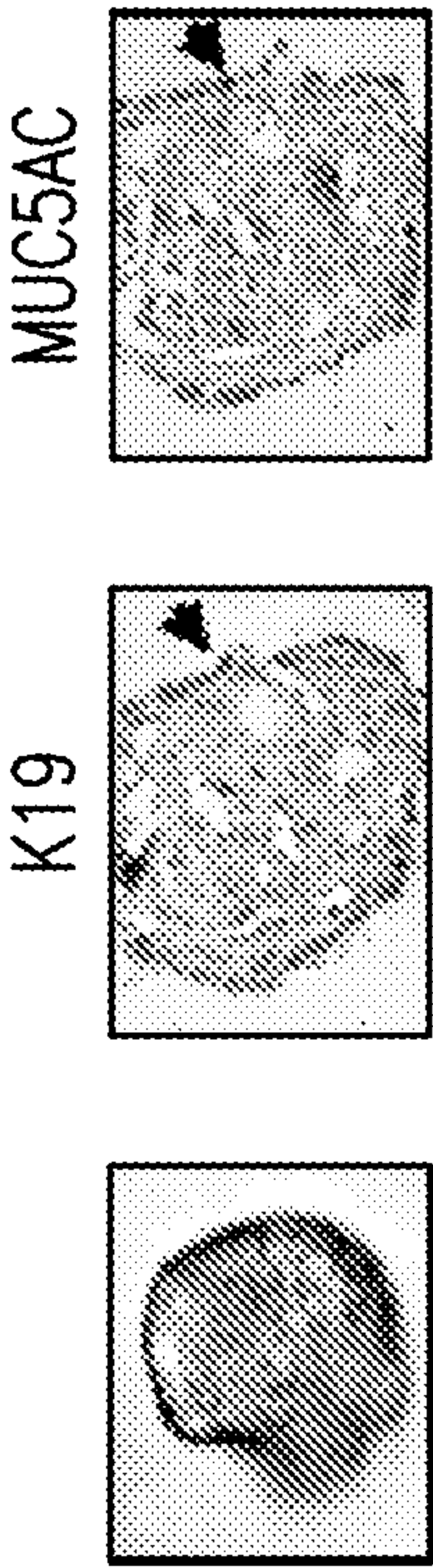
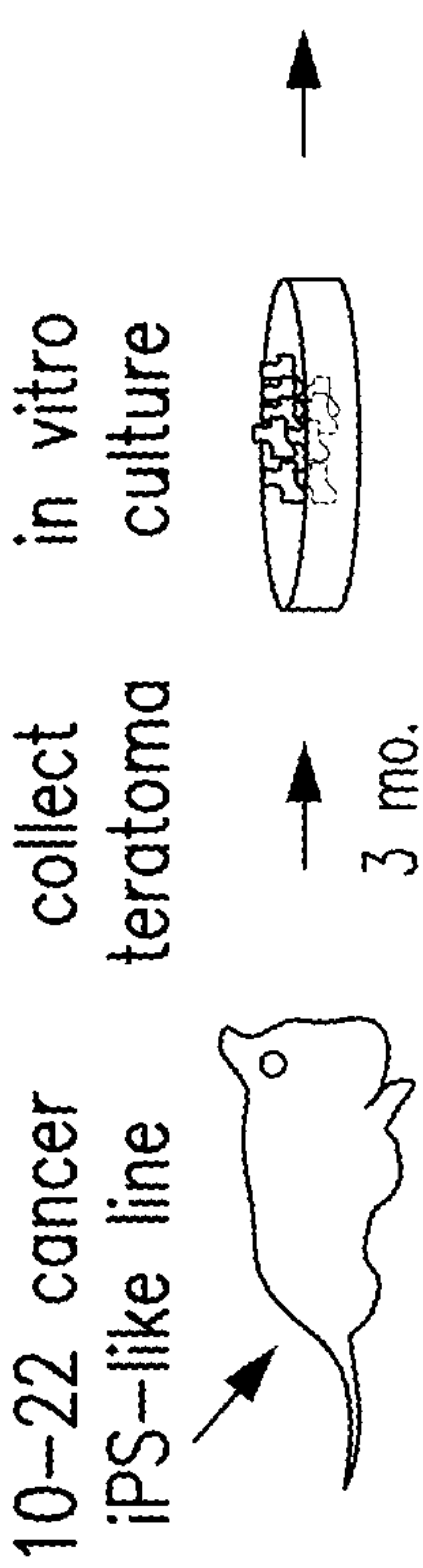


FIG. 5A FIG. 5B FIG. 5C

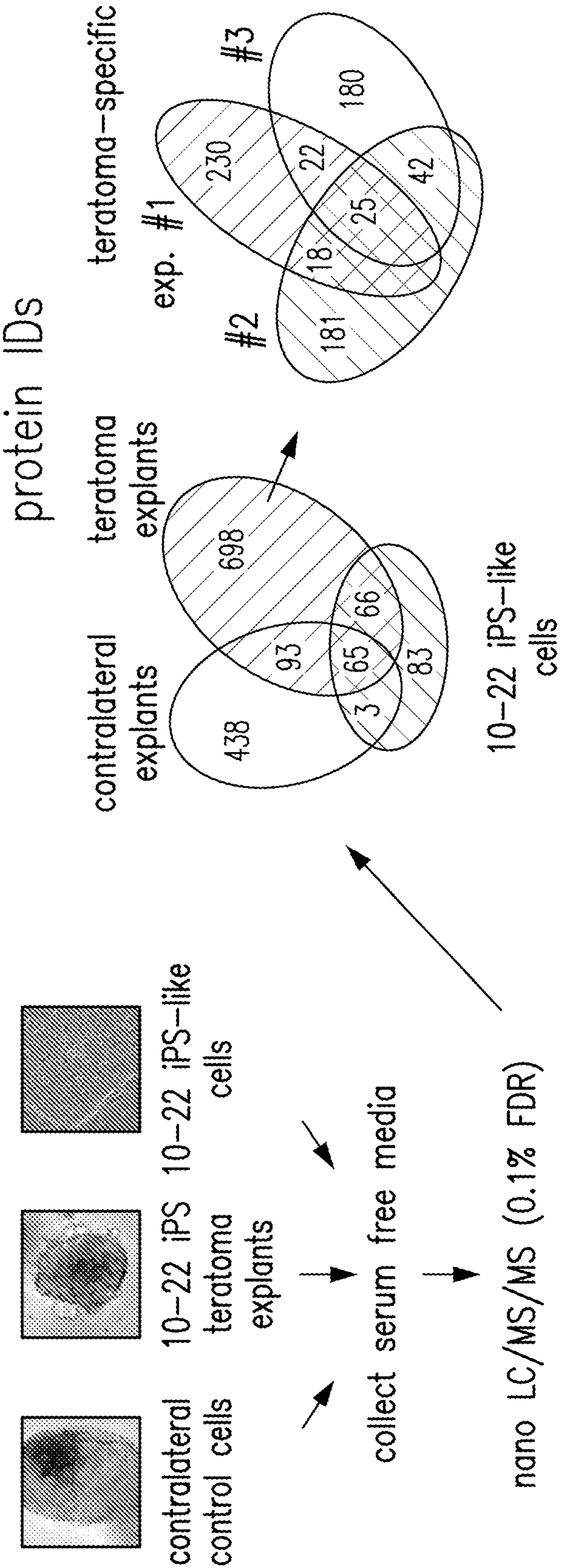


FIG. 5D

9/100

Linked TGF β and integrin pathways

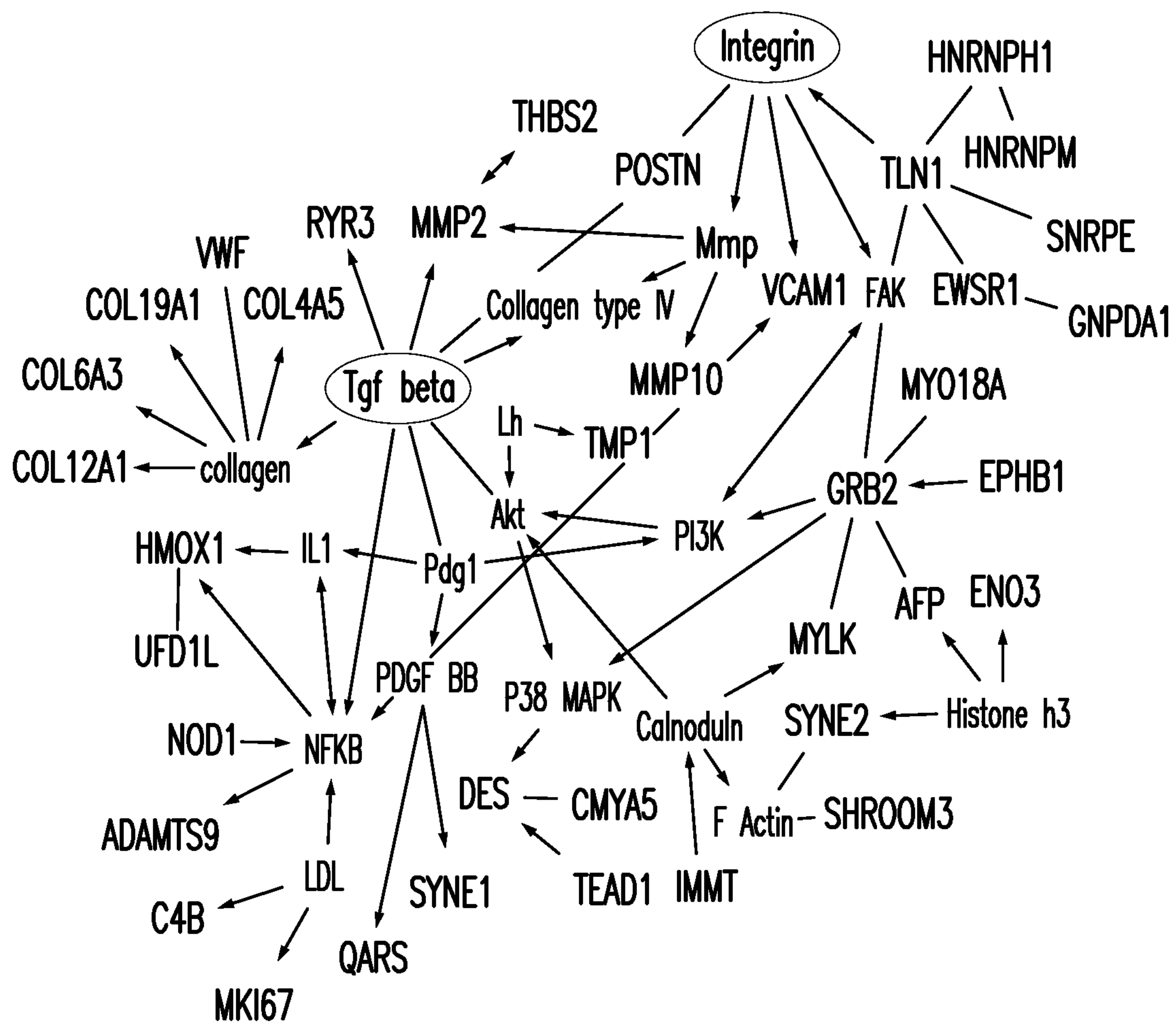


FIG. 5E

10/100

HNF4 α network

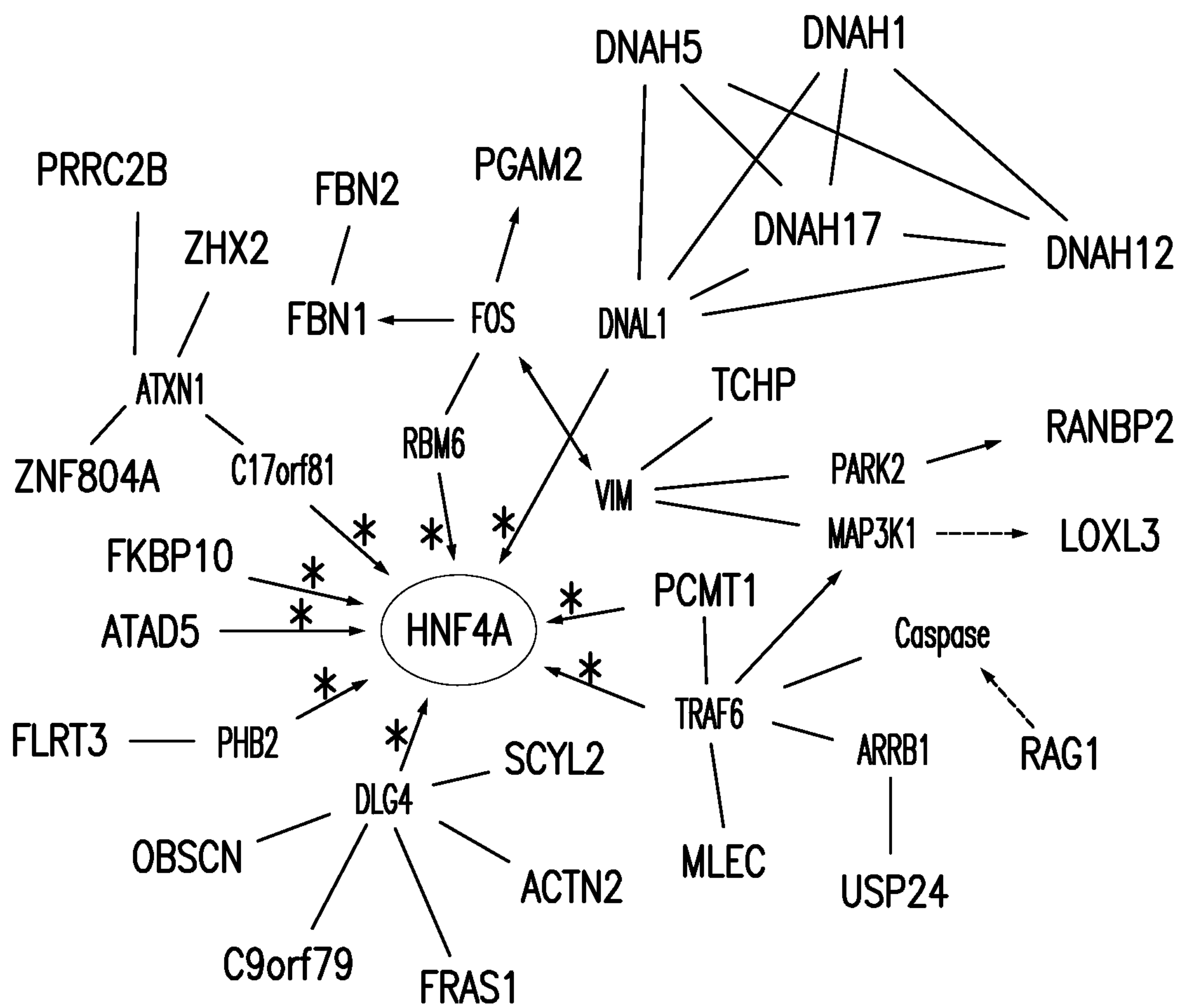


FIG. 5F

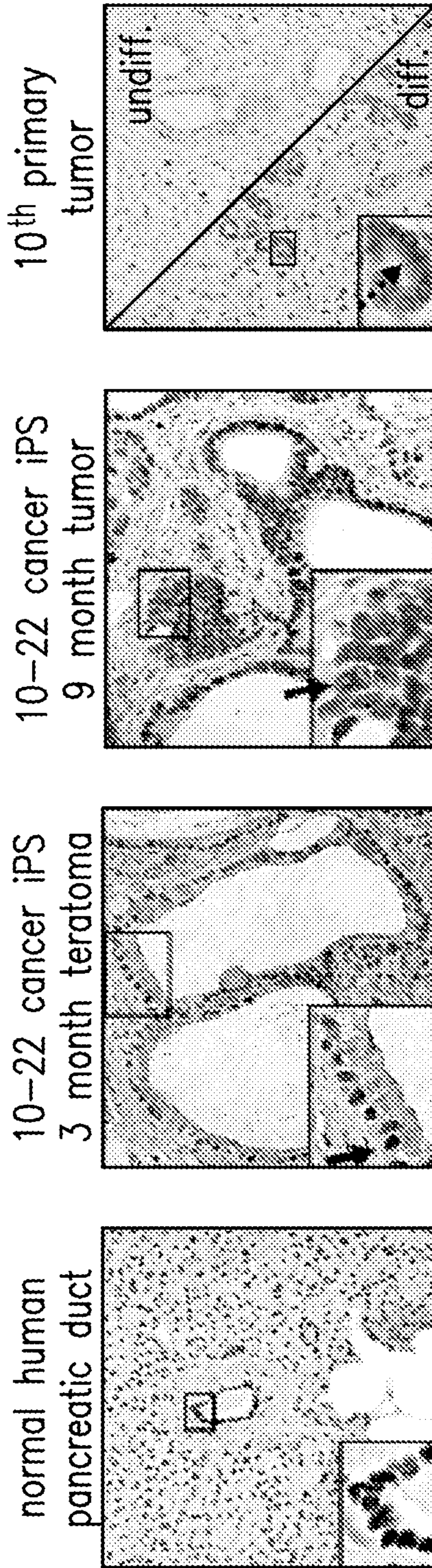


FIG. 6A

FIG. 6B

FIG. 6C

FIG. 6D

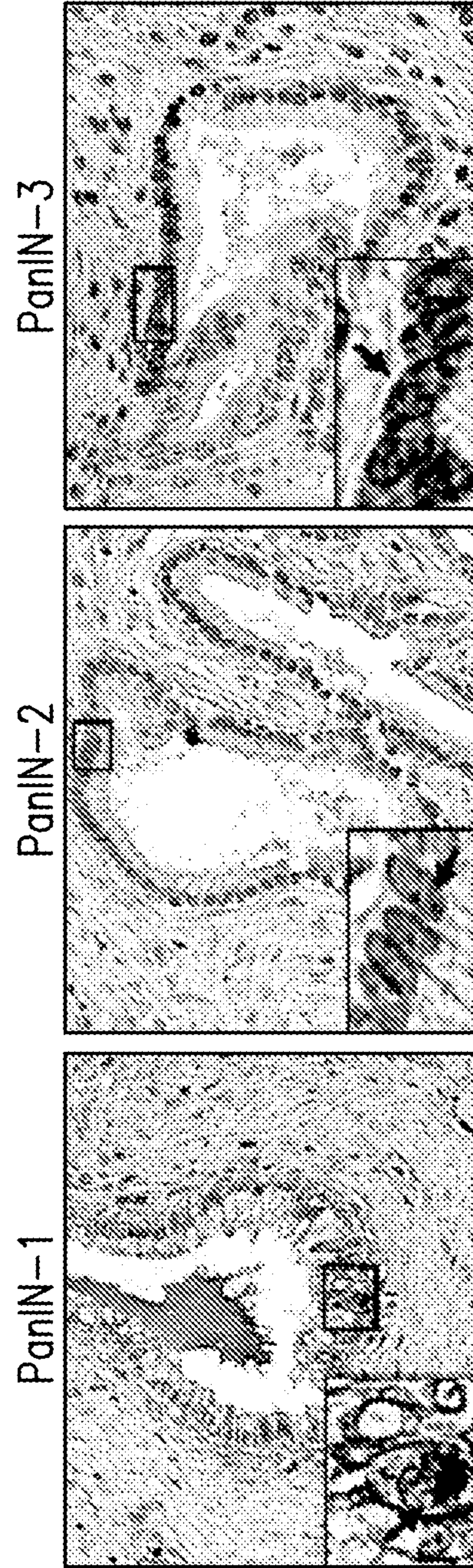


FIG. 6E

FIG. 6F

FIG. 6G

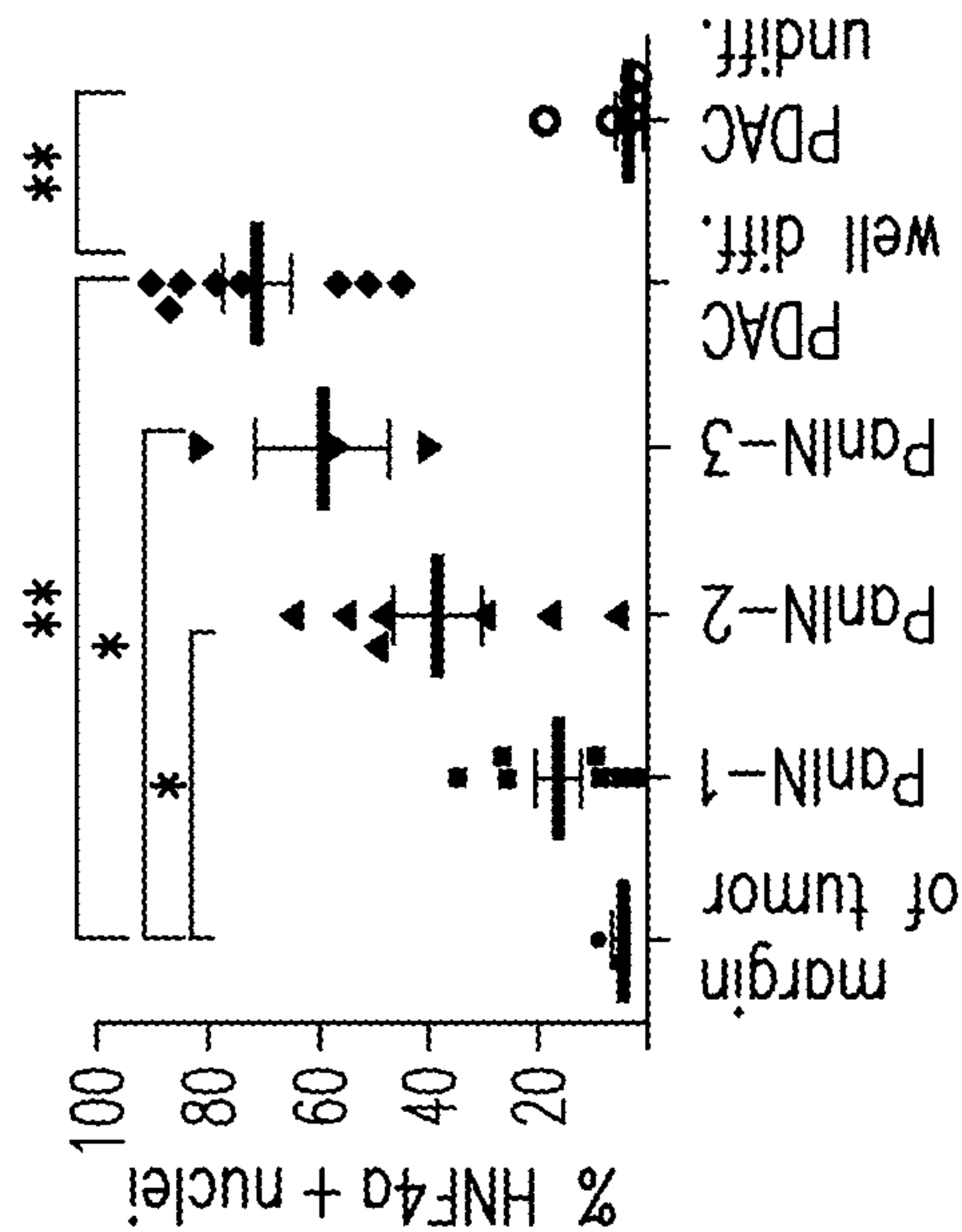


FIG. 6J

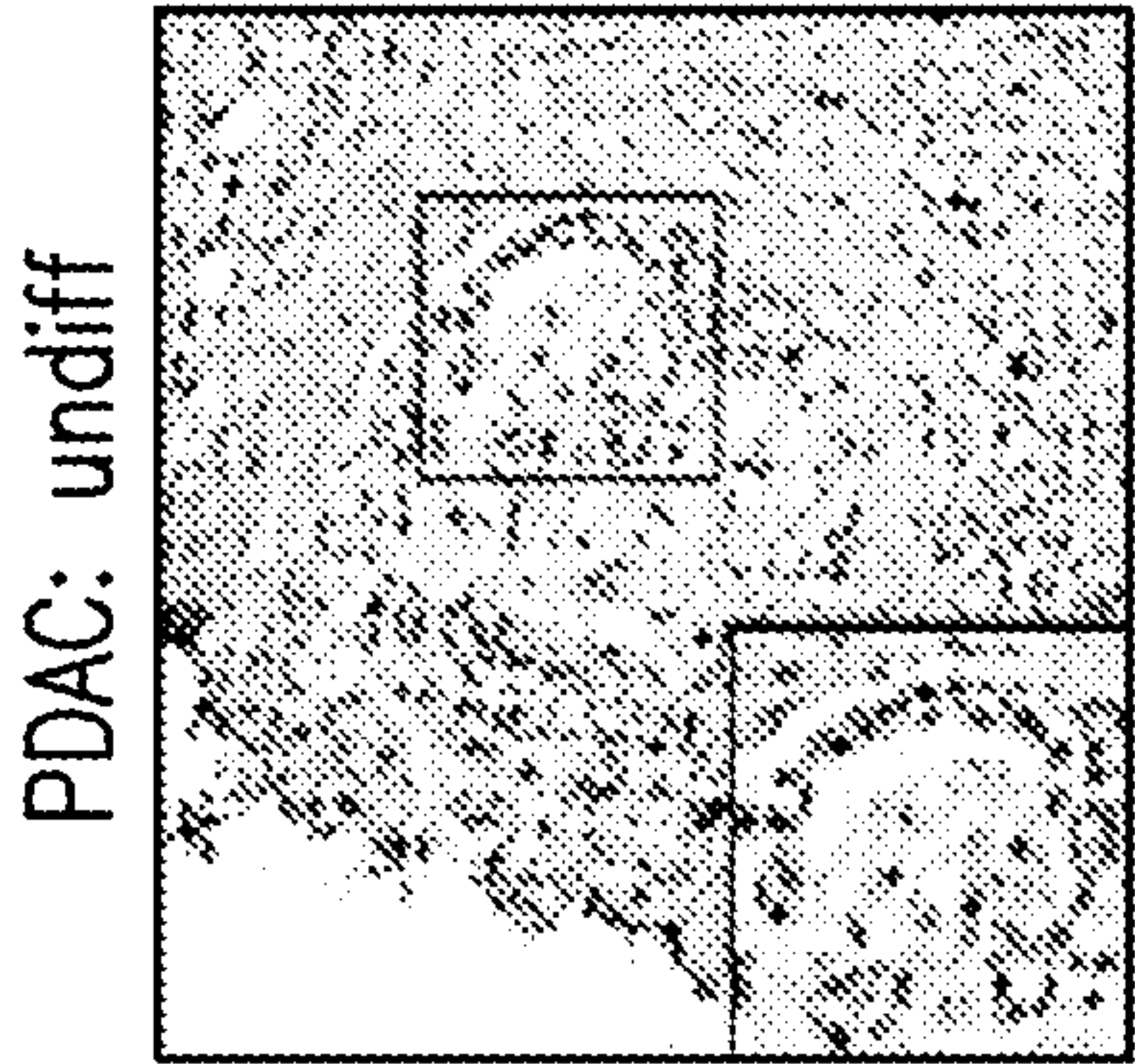


FIG. 6I



FIG. 6H

mouse HNF4 α /Pdx1-cre;RosaYFP/DAPI

PanIN-1



FIG. 6K



FIG. 6K'

PanIN-2



FIG. 6L



FIG. 6L'

PanIN-3

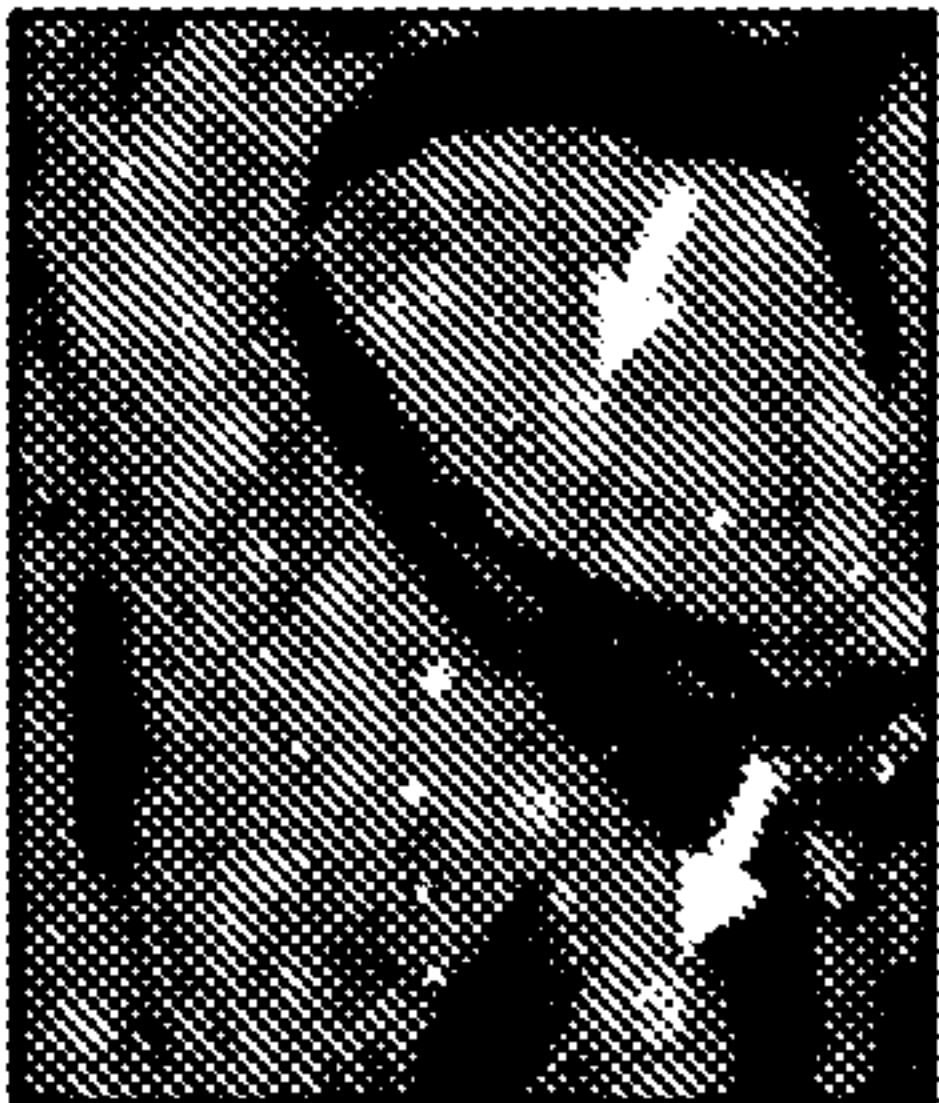


FIG. 6M



FIG. 6M'

PDAC: diff



FIG. 6O

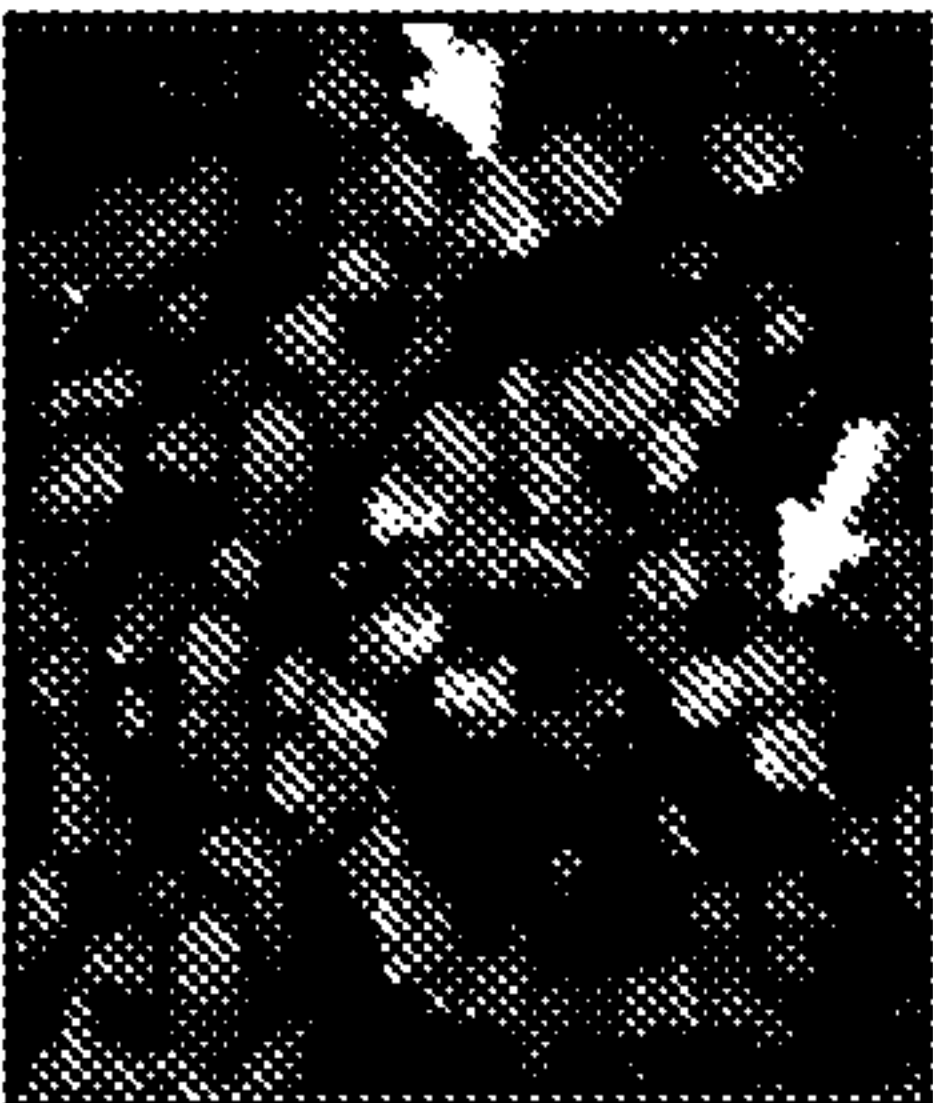


FIG. 6O'

PDAC: undiff

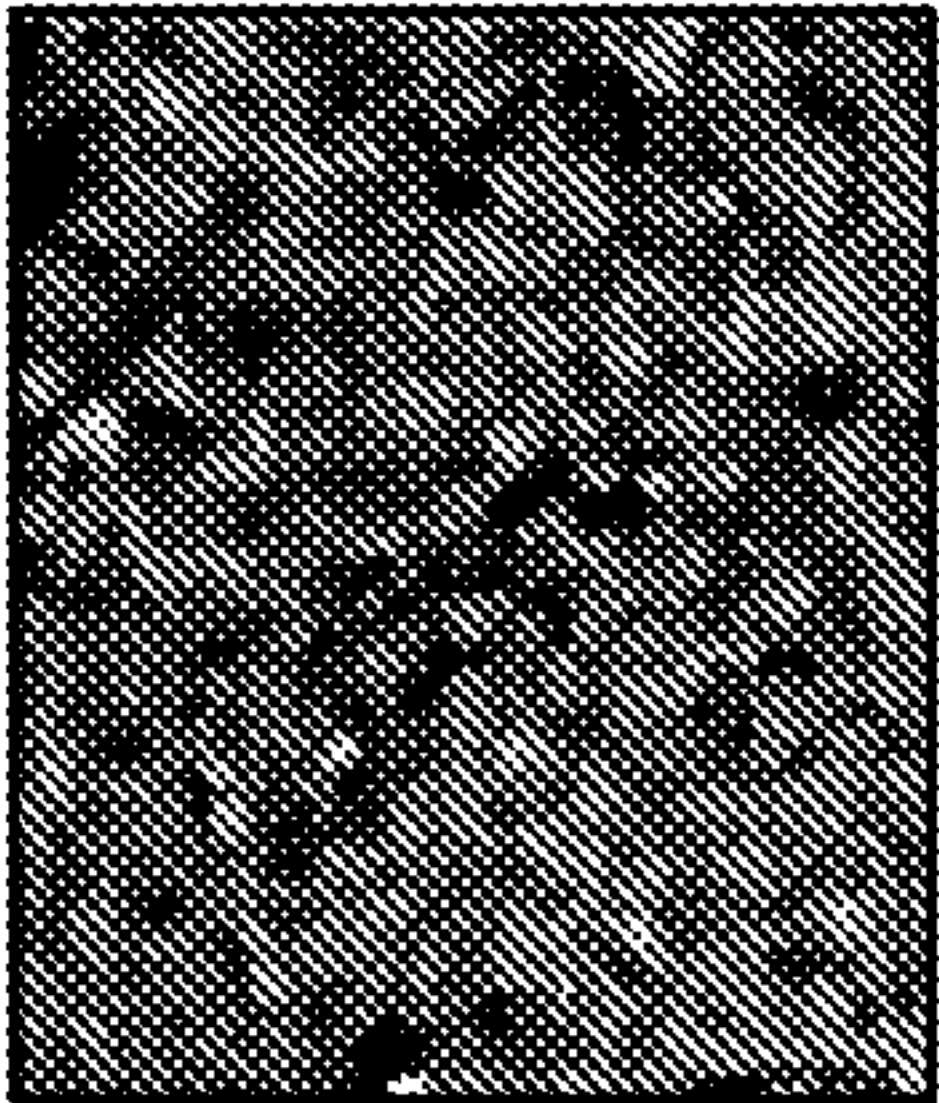


FIG. 6P



FIG. 6P'

Patient #14

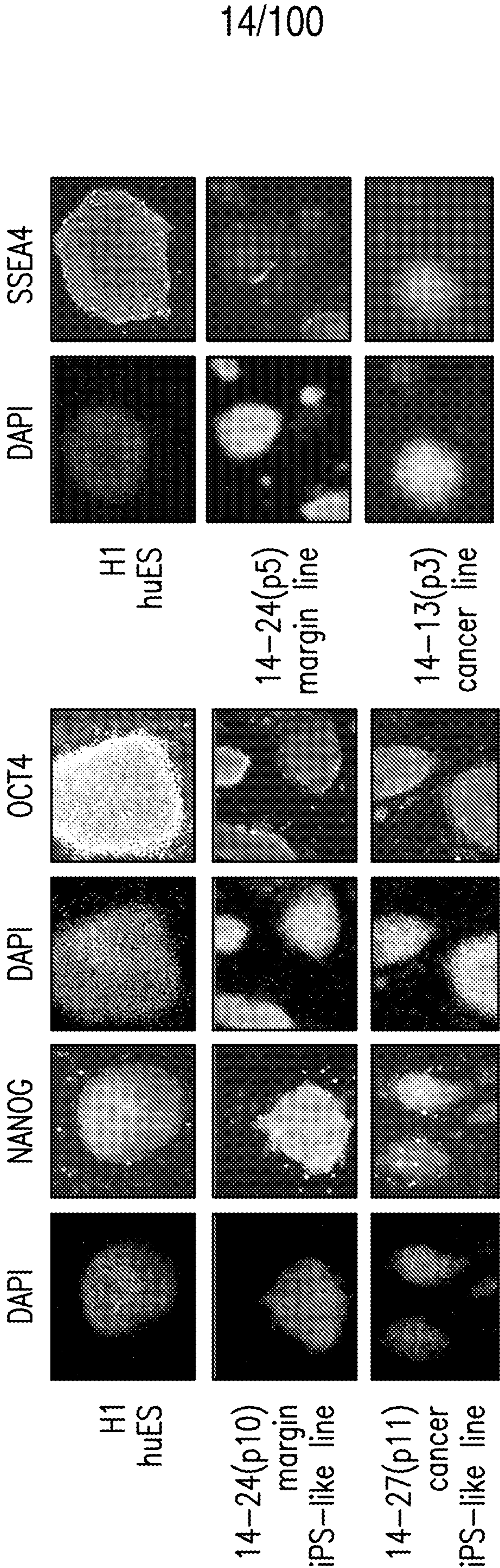


FIG. 7A

15/100

Patient #14

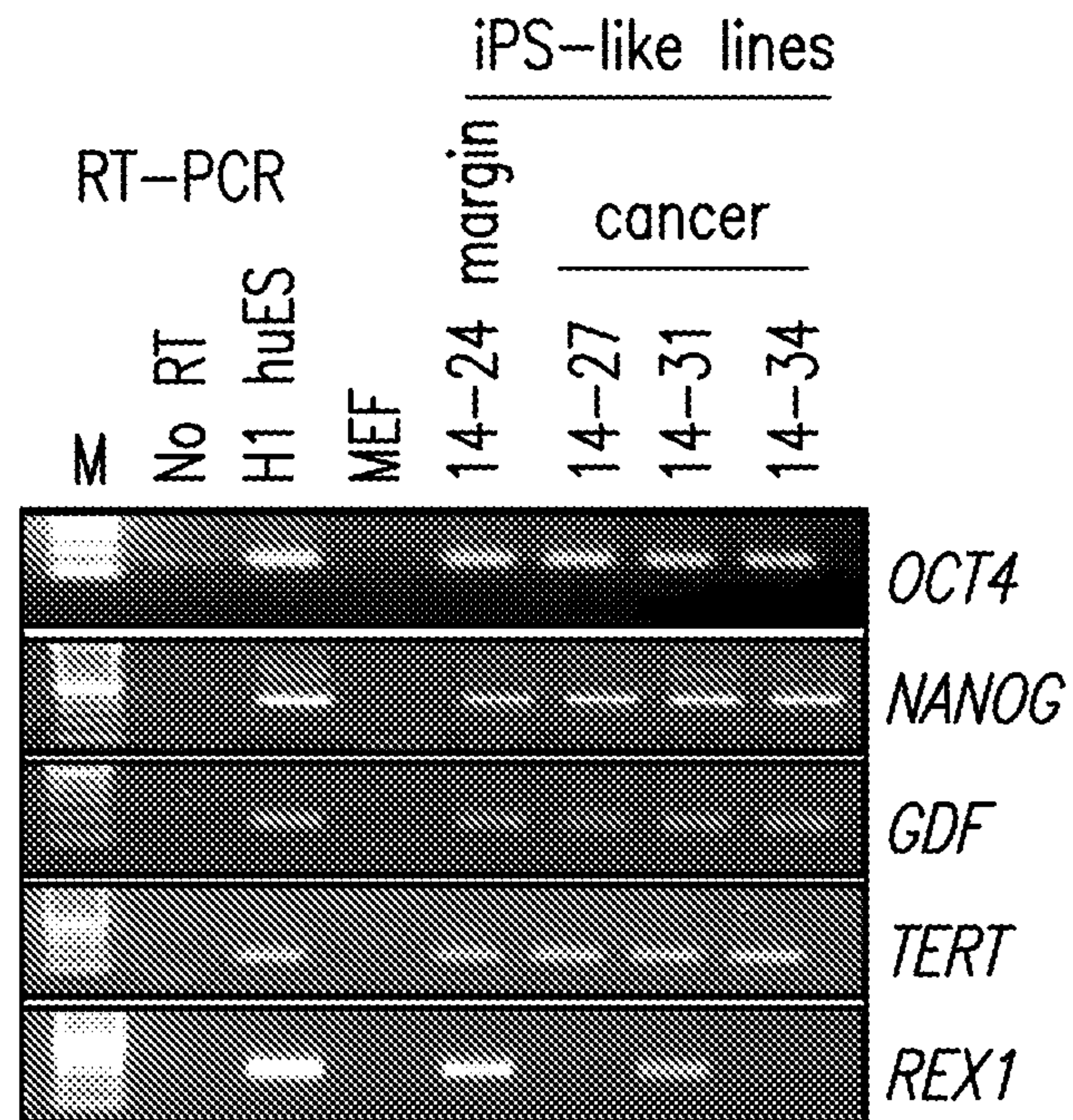


FIG. 7B

Patient #19

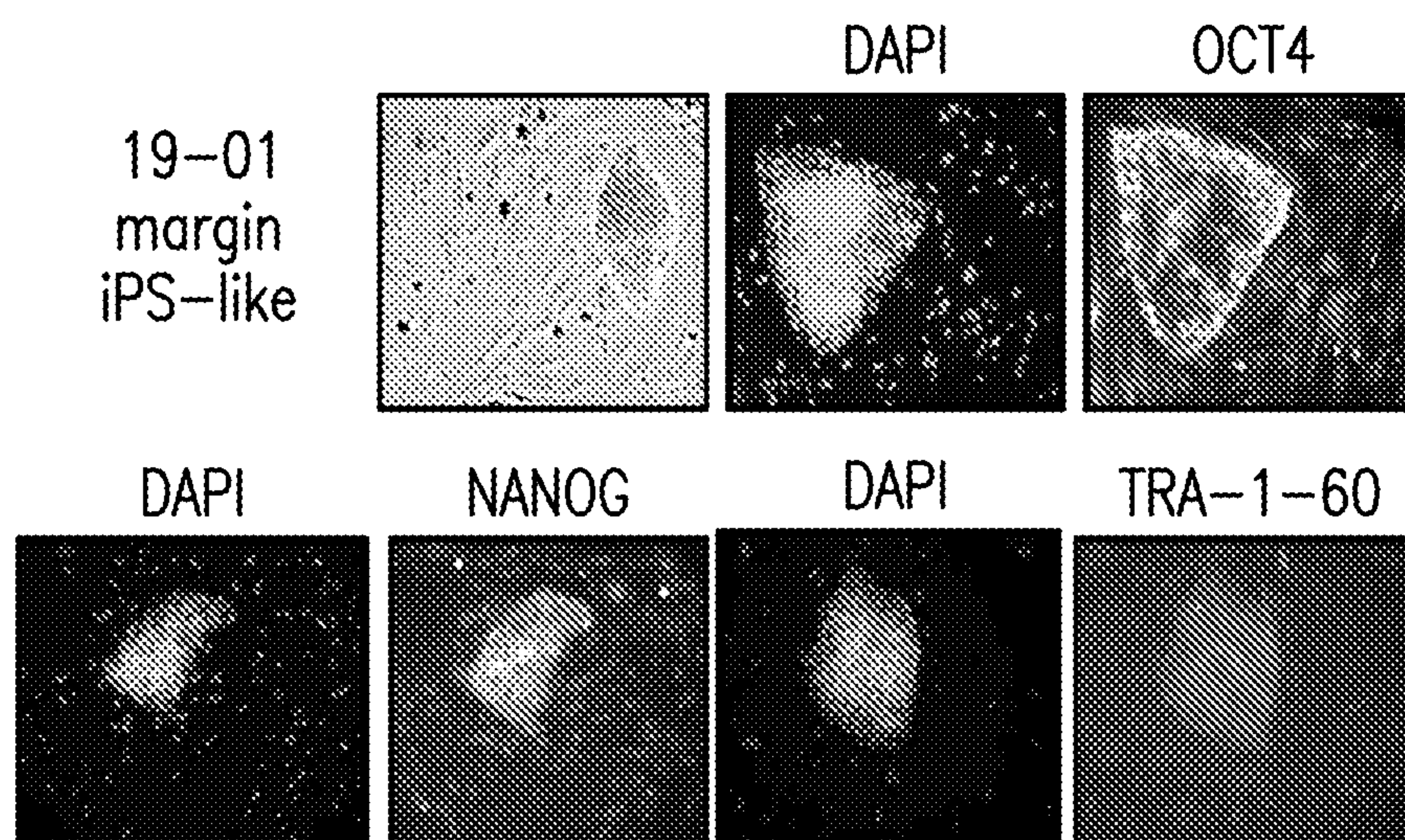


FIG. 7C

16/100

Patient #19

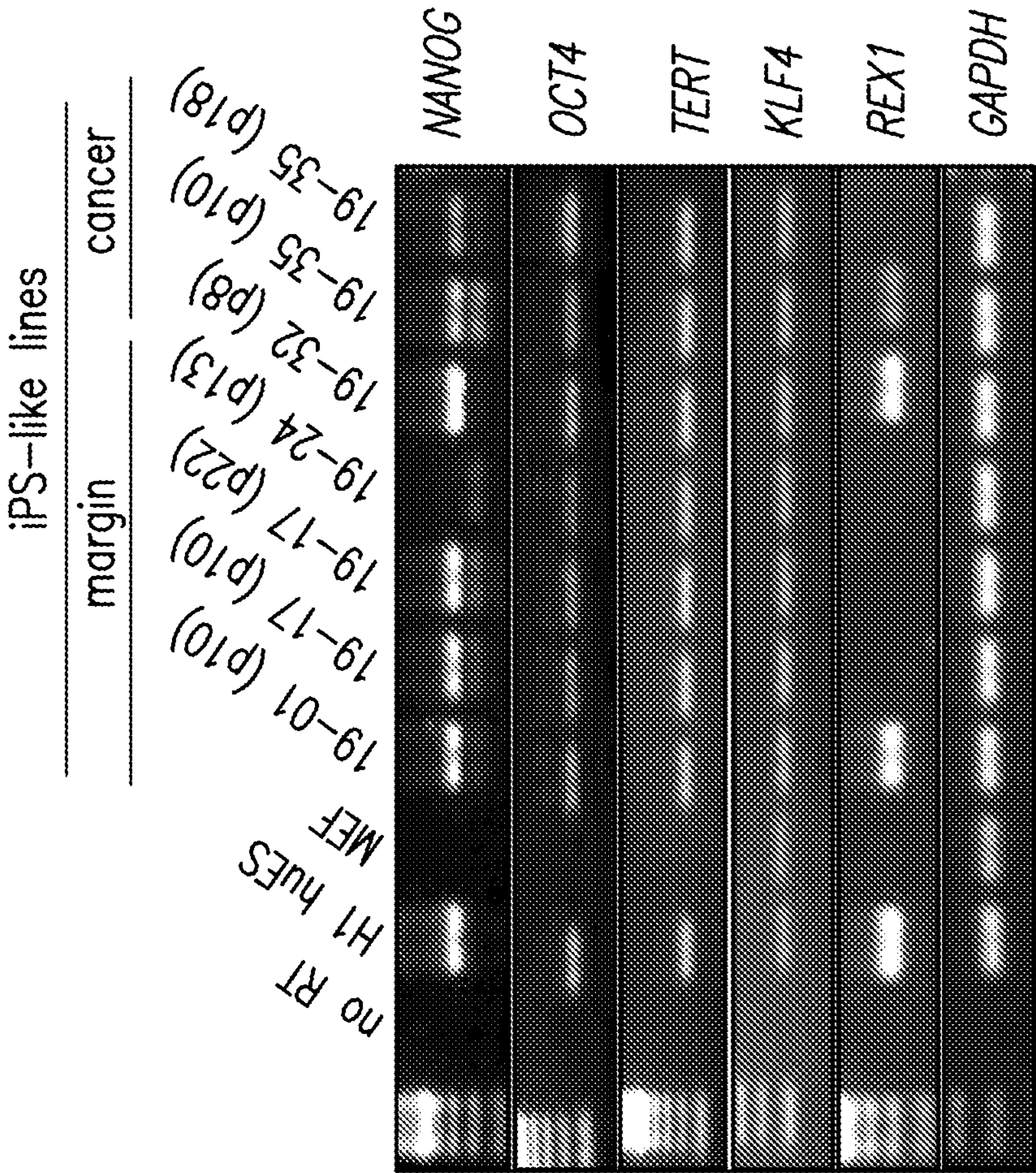


FIG. 7D

17/100

RT-qPCR EB's 14-24 margin line

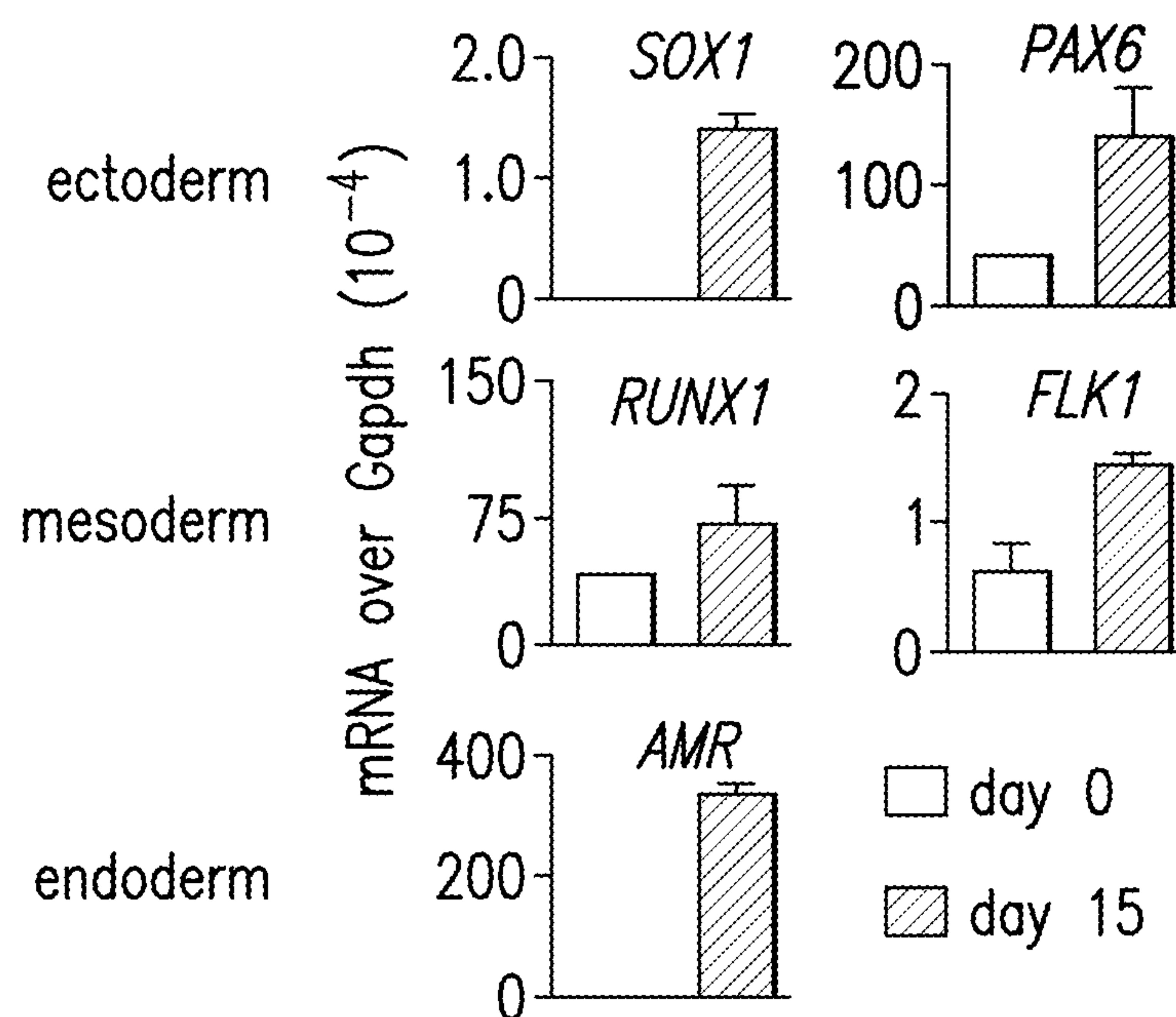


FIG. 7E

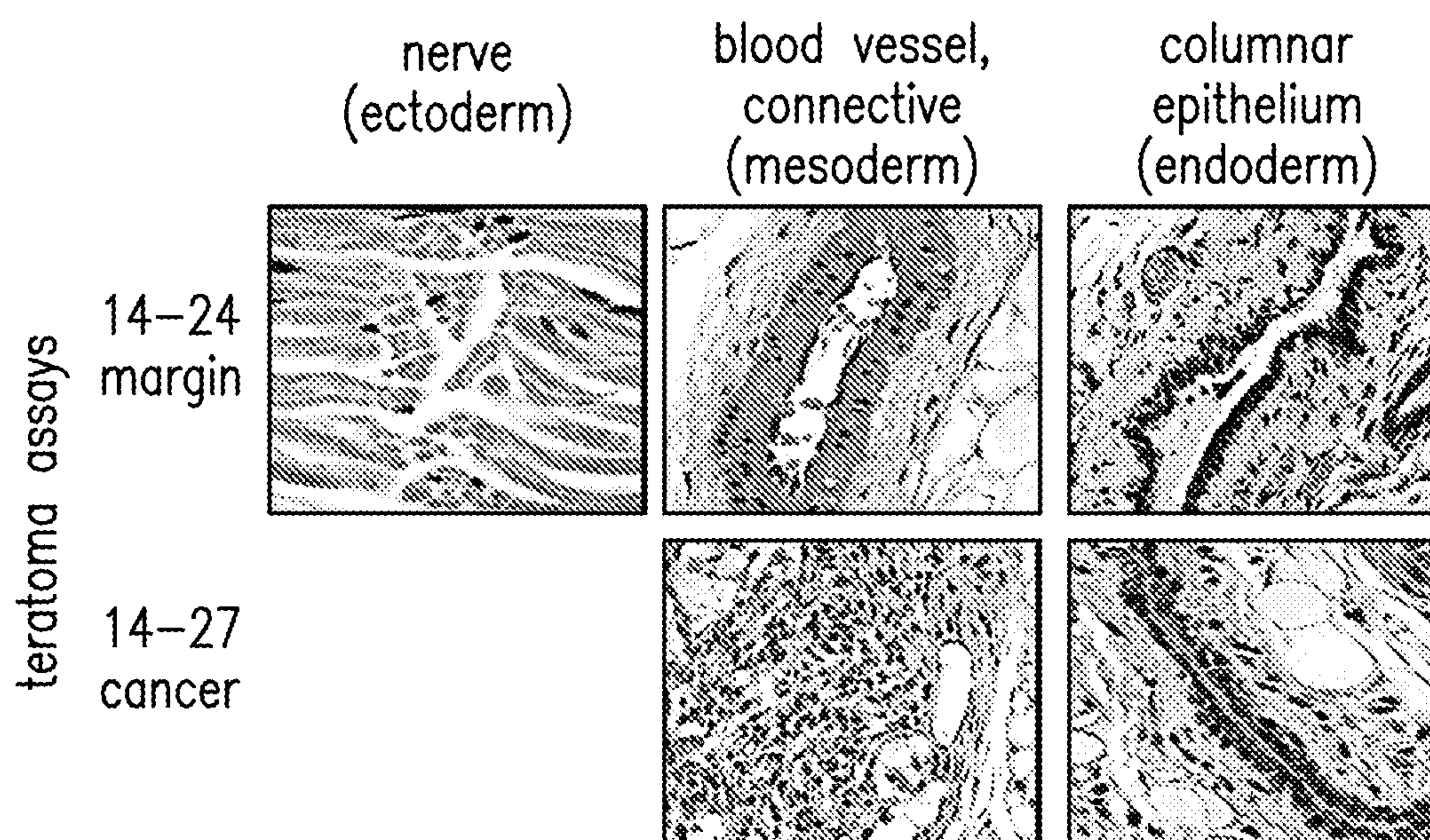


FIG. 7F

18/100

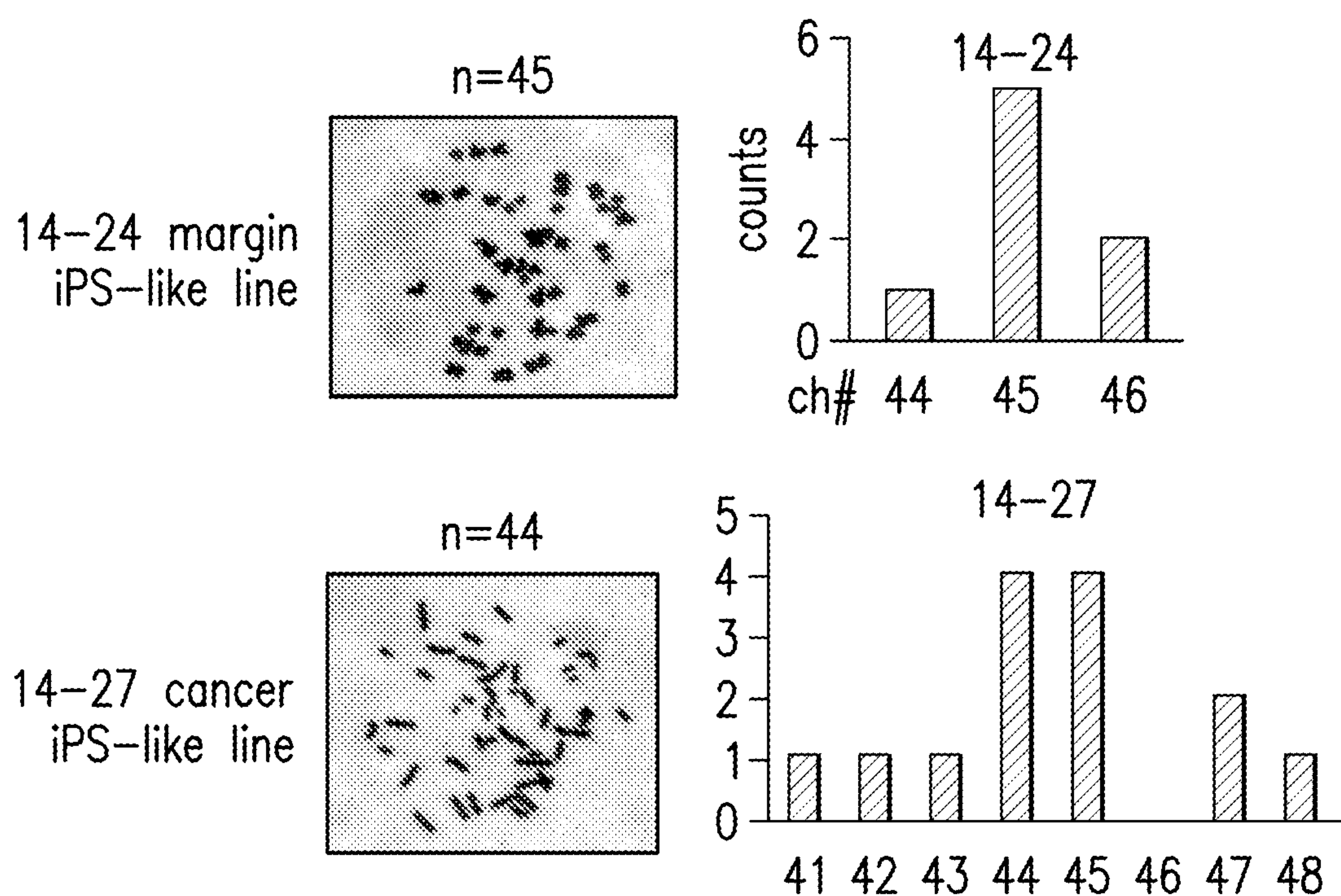


FIG. 7G

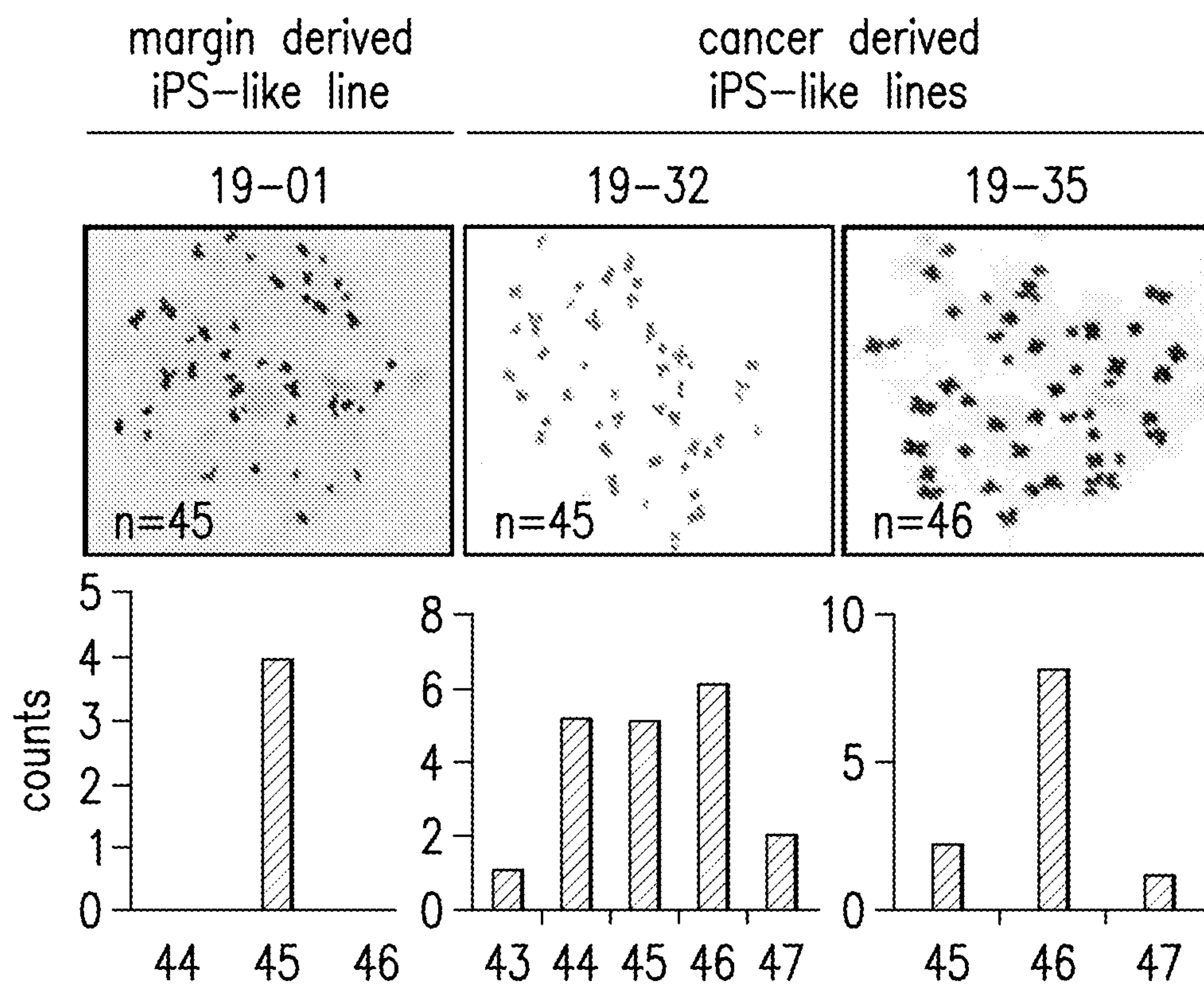


FIG. 7H

Immunohistochemistry (IHC) analysis of different germ layer tissues
in teratomas from H1 huES (positive control) and 10-22 cancer iPS cells

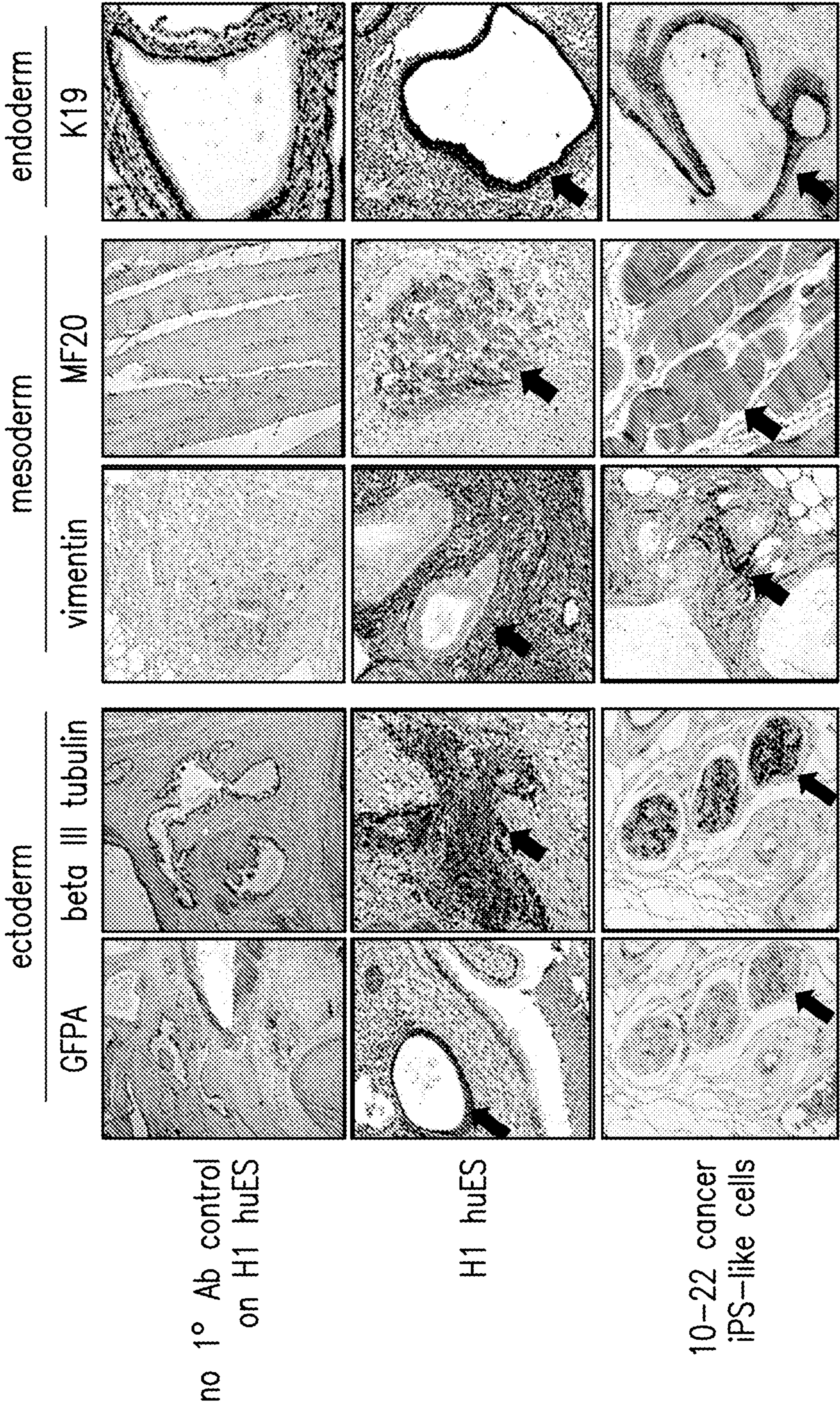
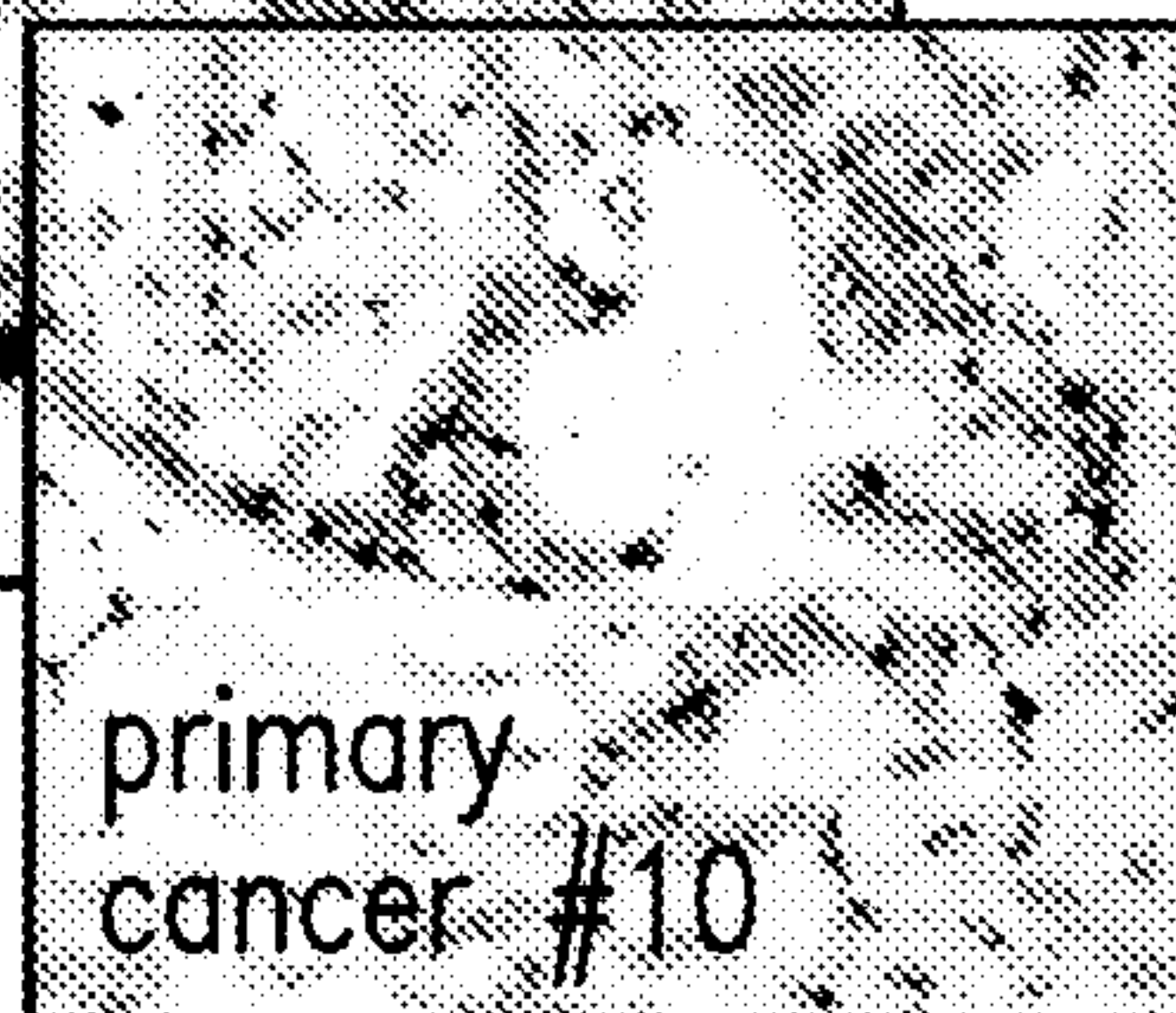


FIG. 8A

20/100

NANOG IHC

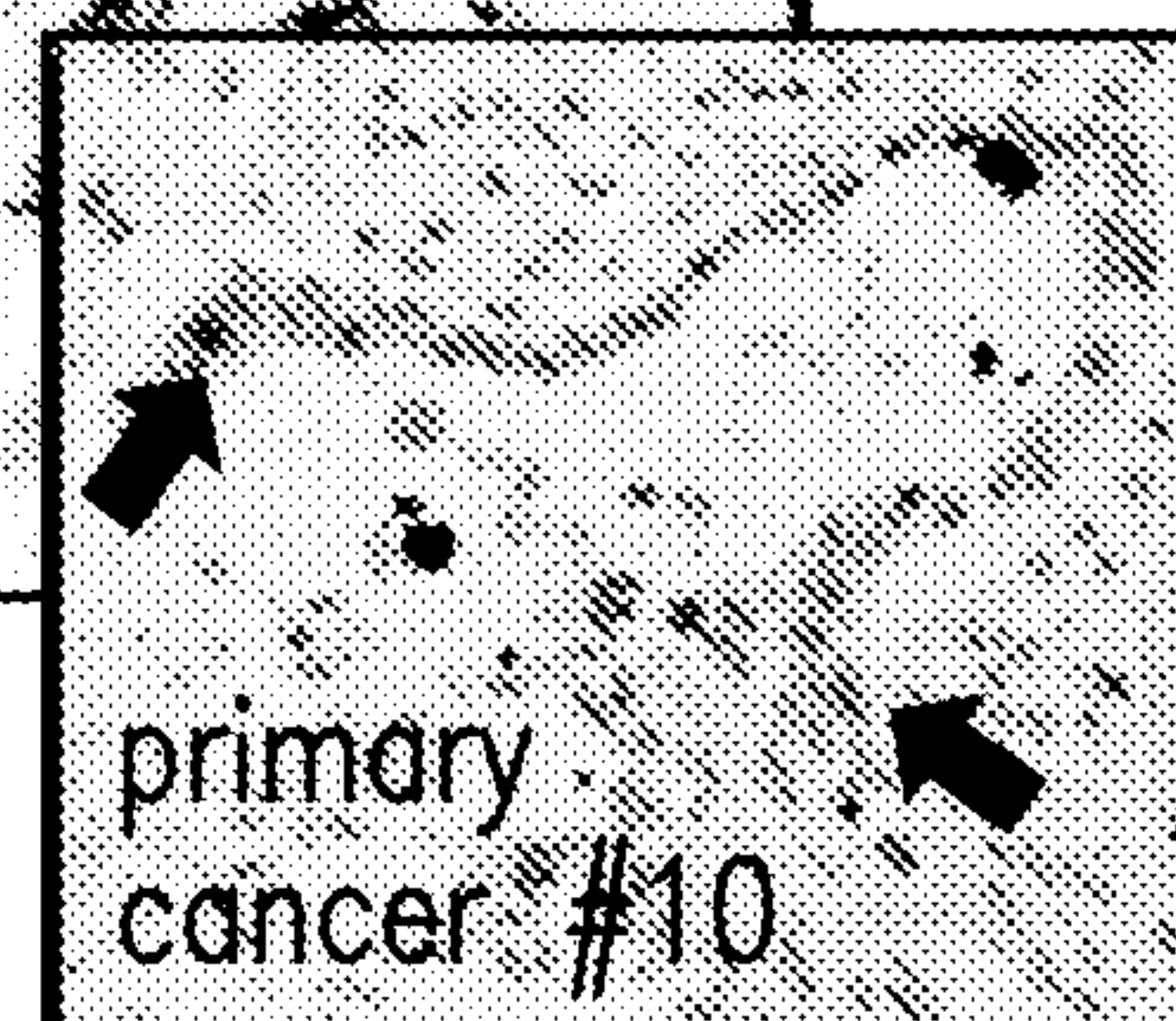
H1 huES



(negative)

OCT4 IHC

H1 huES



(sporadic)

FIG. 8B

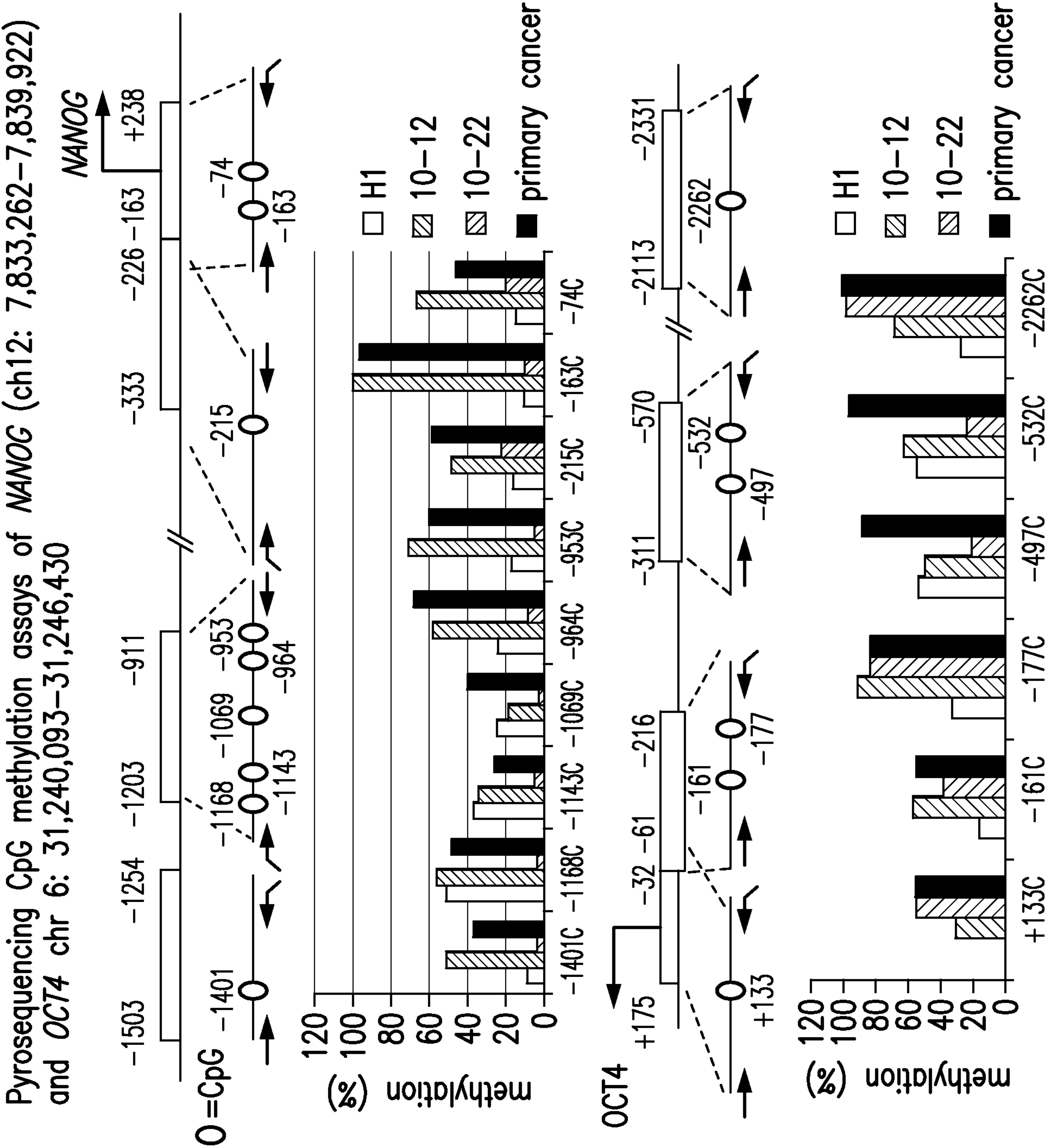


FIG. 8C

22/100

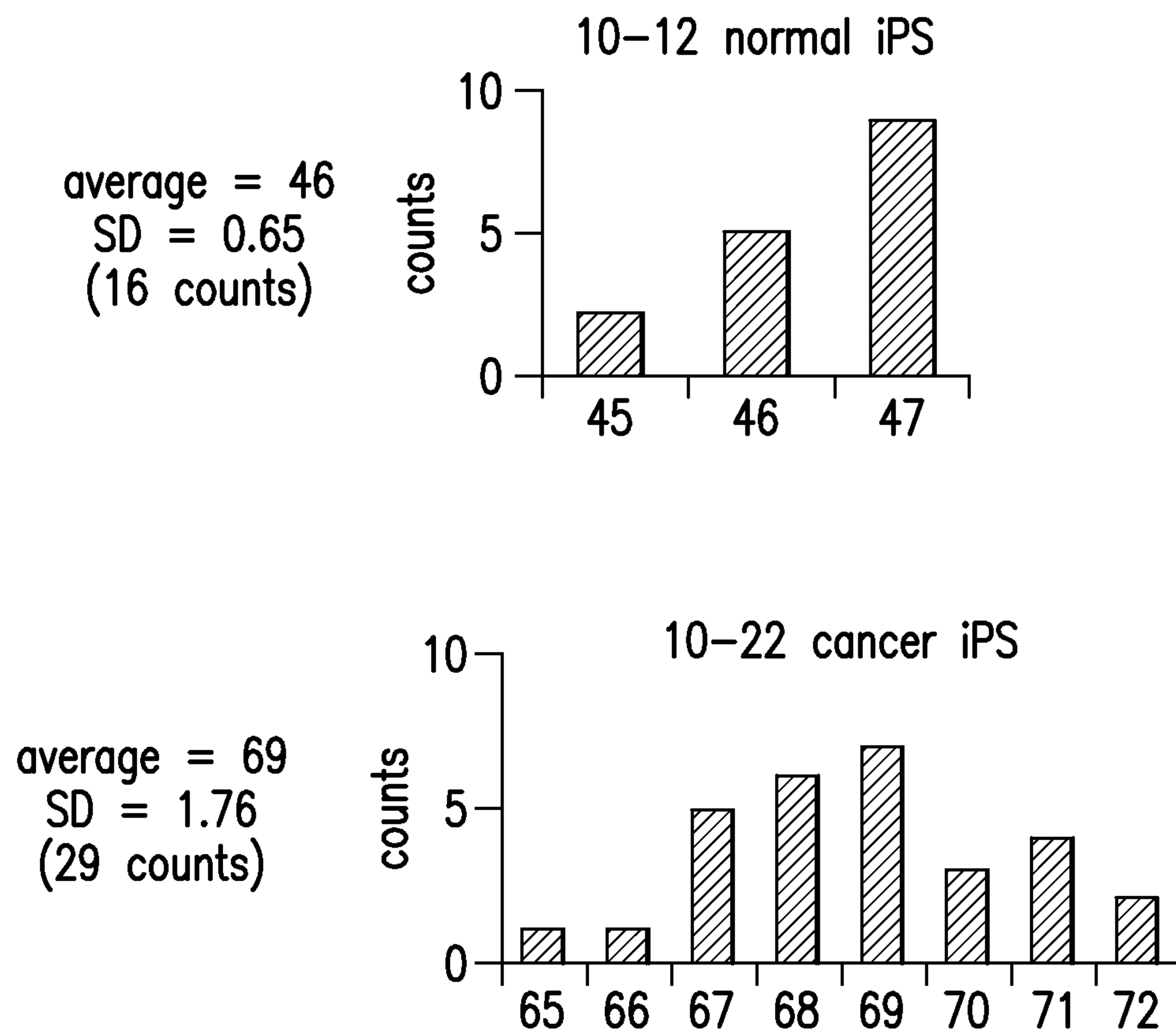
A Karyotyping of 10th experiment clones

FIG. 9A

Comparative Genomic Hybridization (CGH) of 10th clones

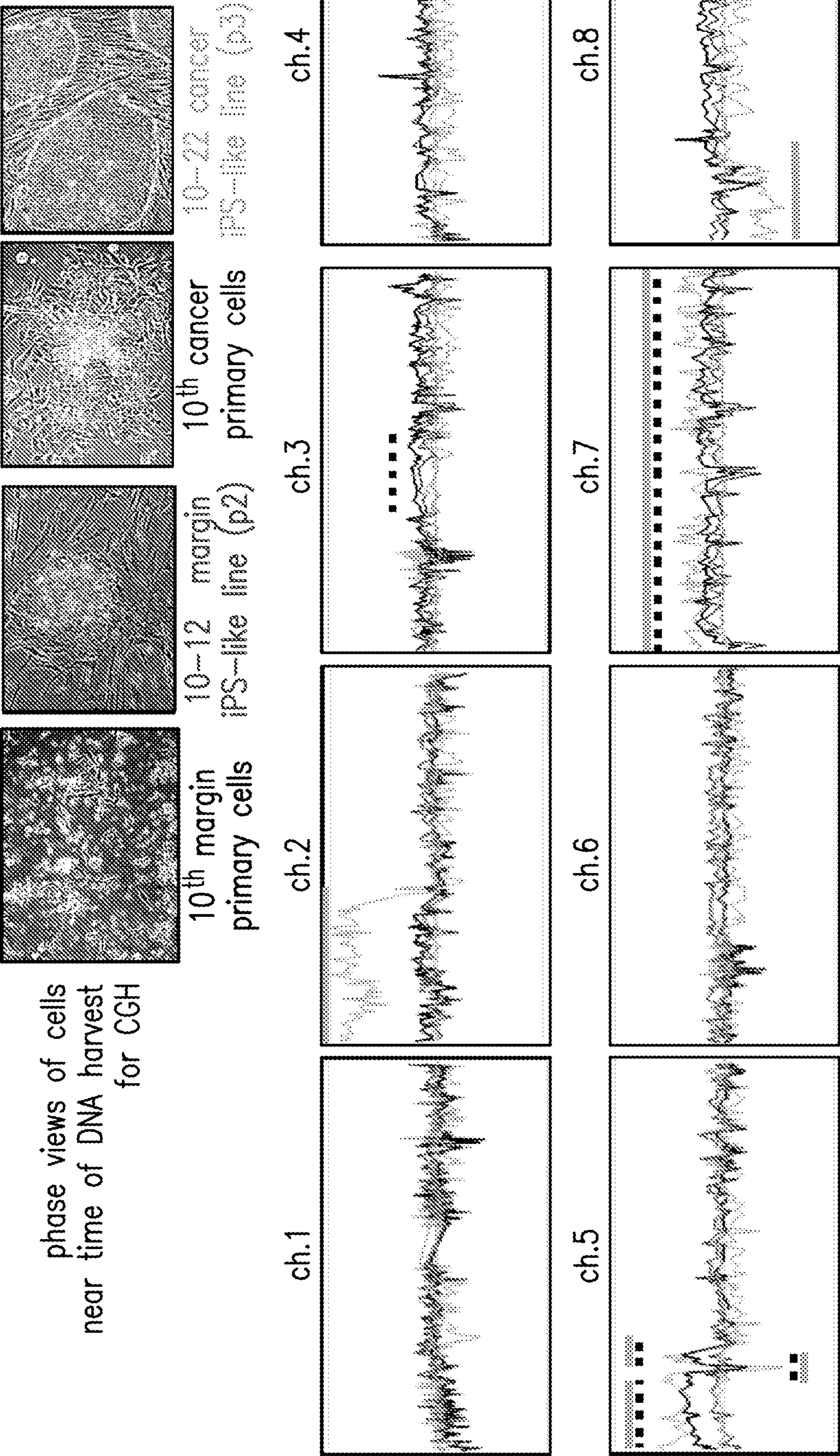


FIG. 9B

24/100

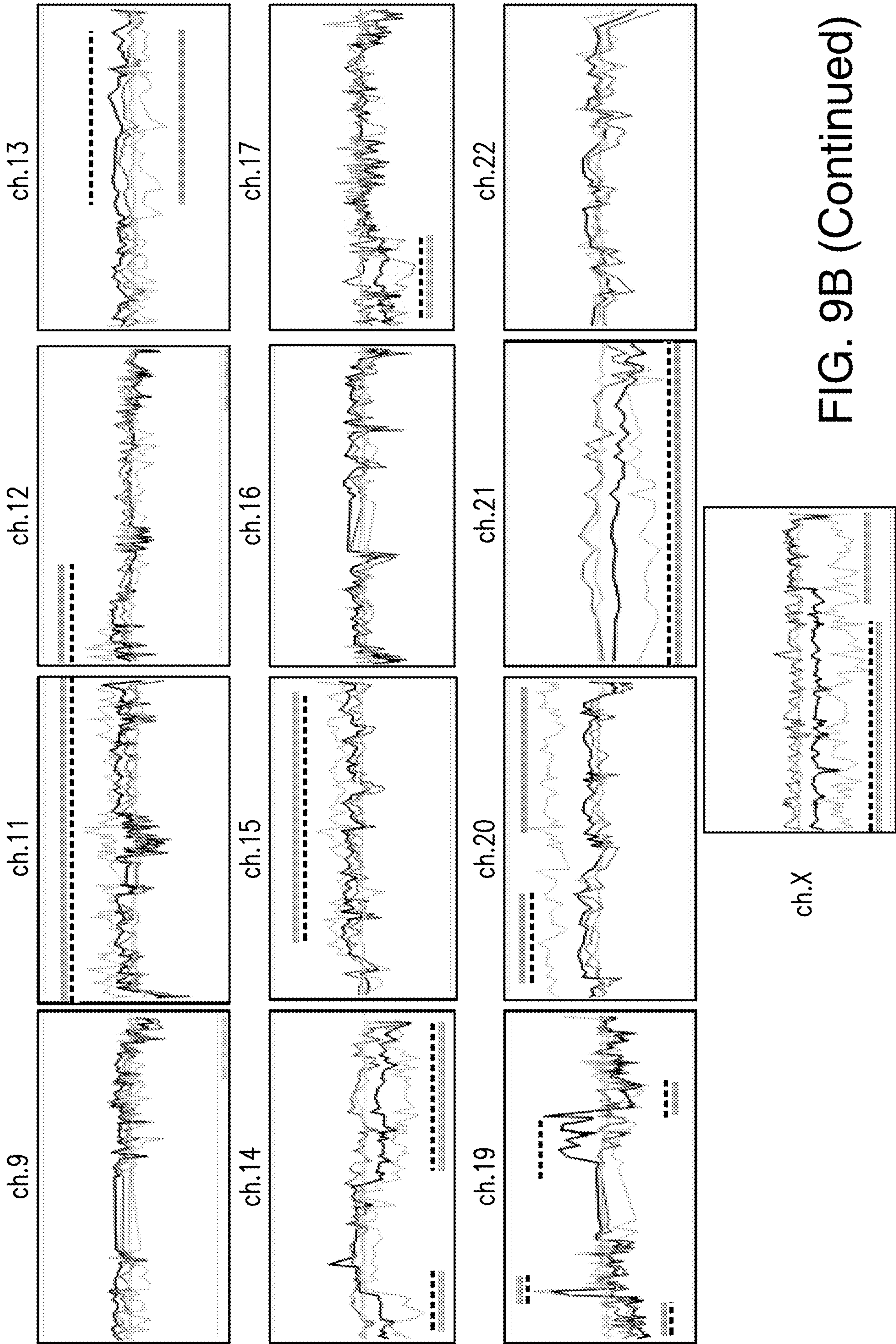


FIG. 9B (Continued)

25/100

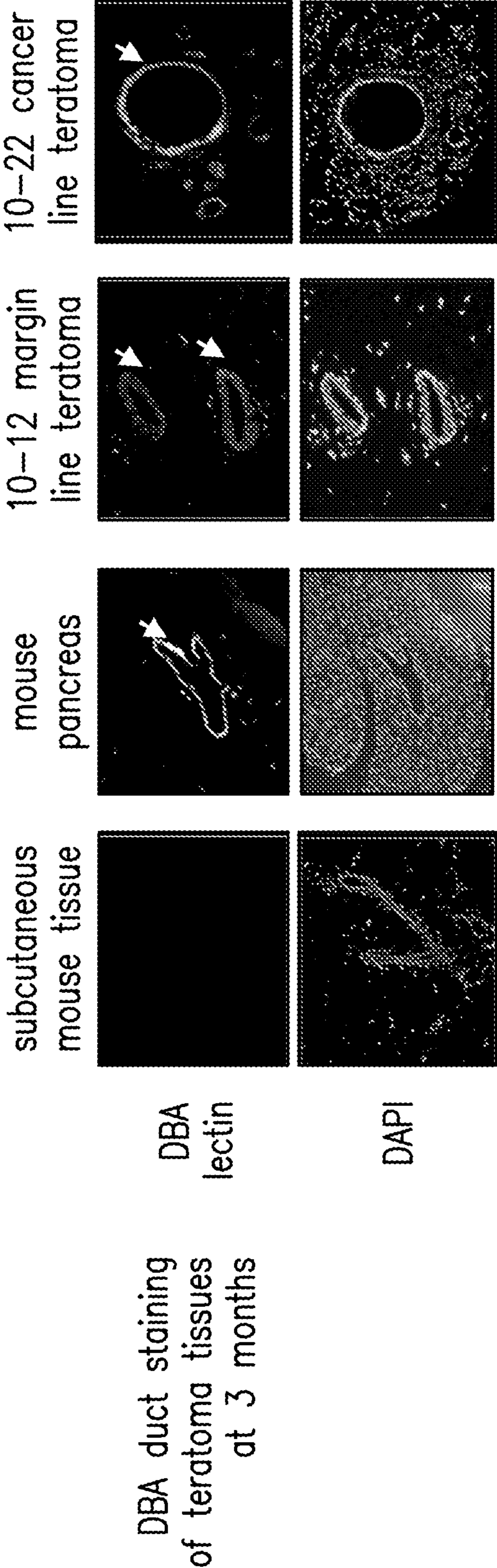


FIG. 10A

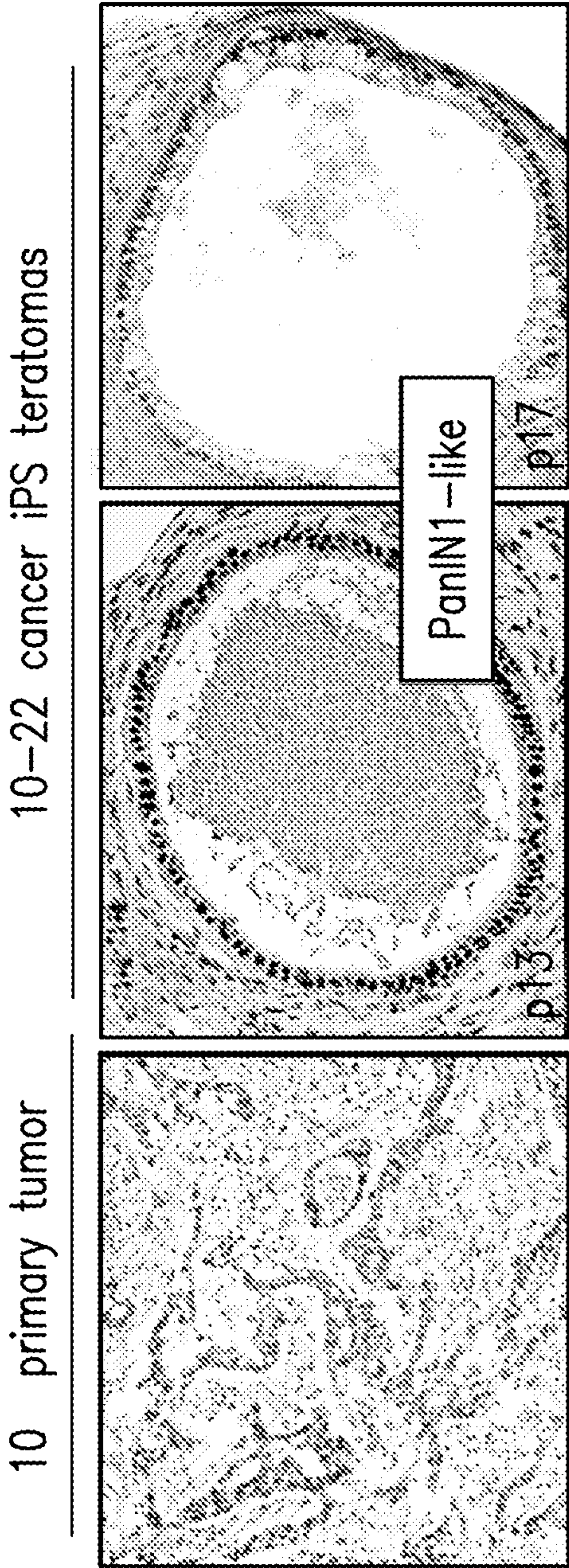


FIG. 10B

FIG. 10C

FIG. 10D

10-22 cancer iPS
teratoma PanIN-like
structures in
independent mice
at 3 months

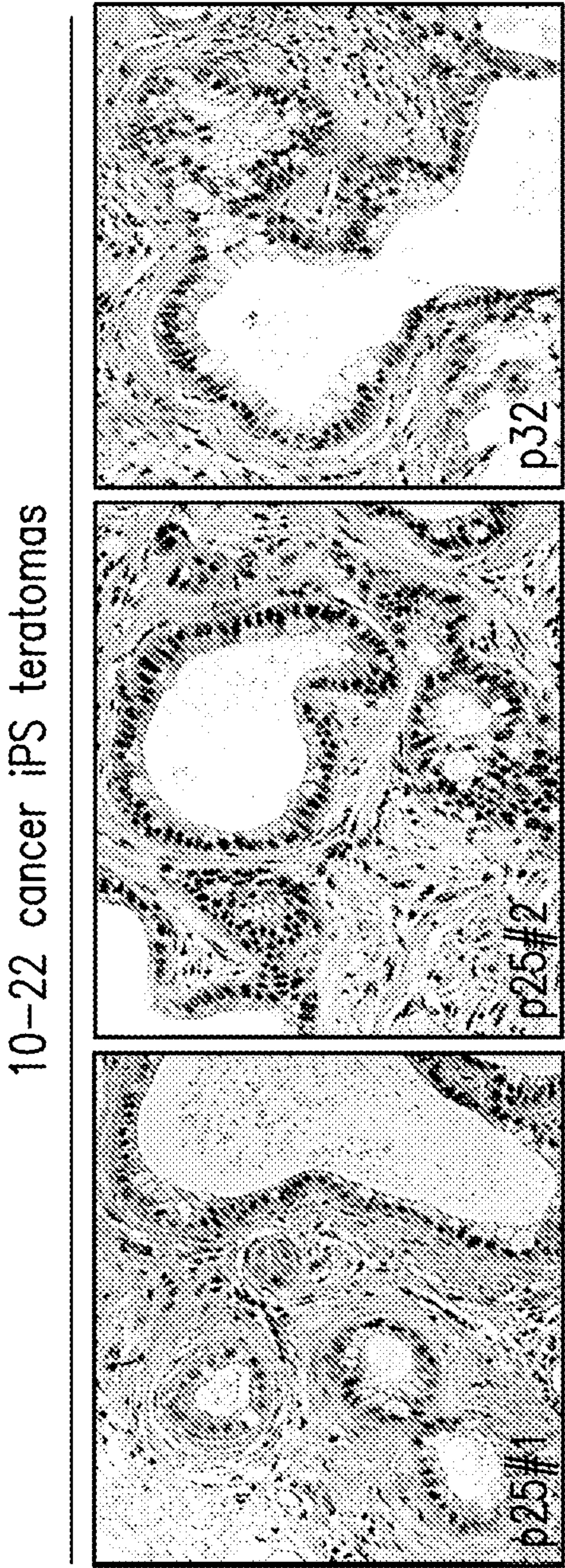


FIG. 10E

FIG. 10F

FIG. 10G

PanIN-2/3-like
structures

27/100

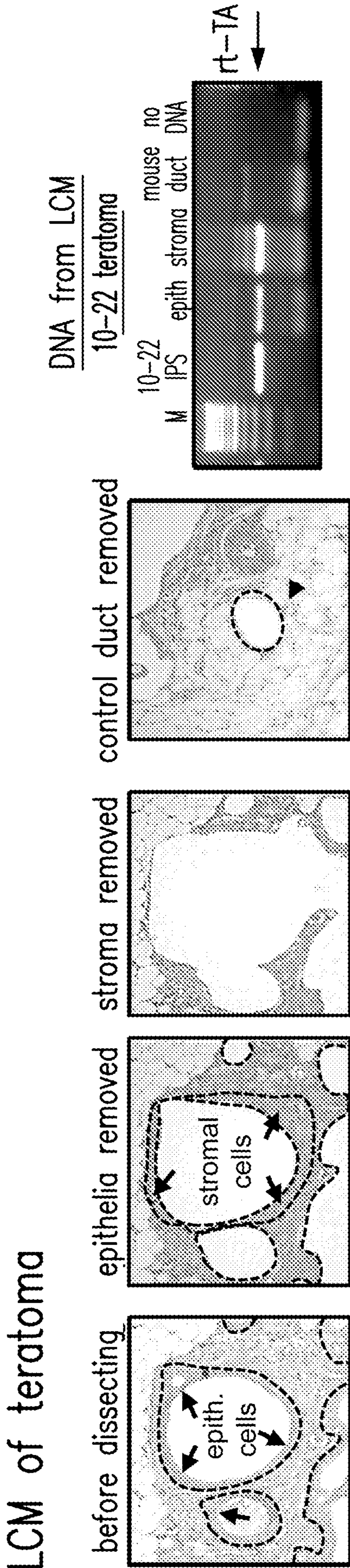


FIG. 10H

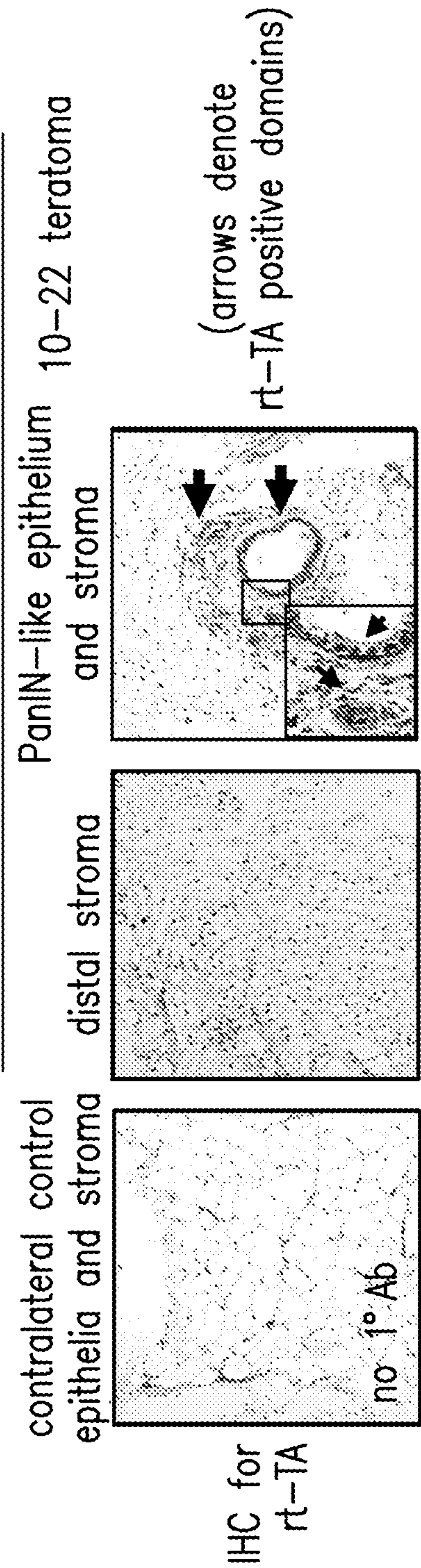


FIG. 10I

Control tests of human K19 and MUC5AC antibodies

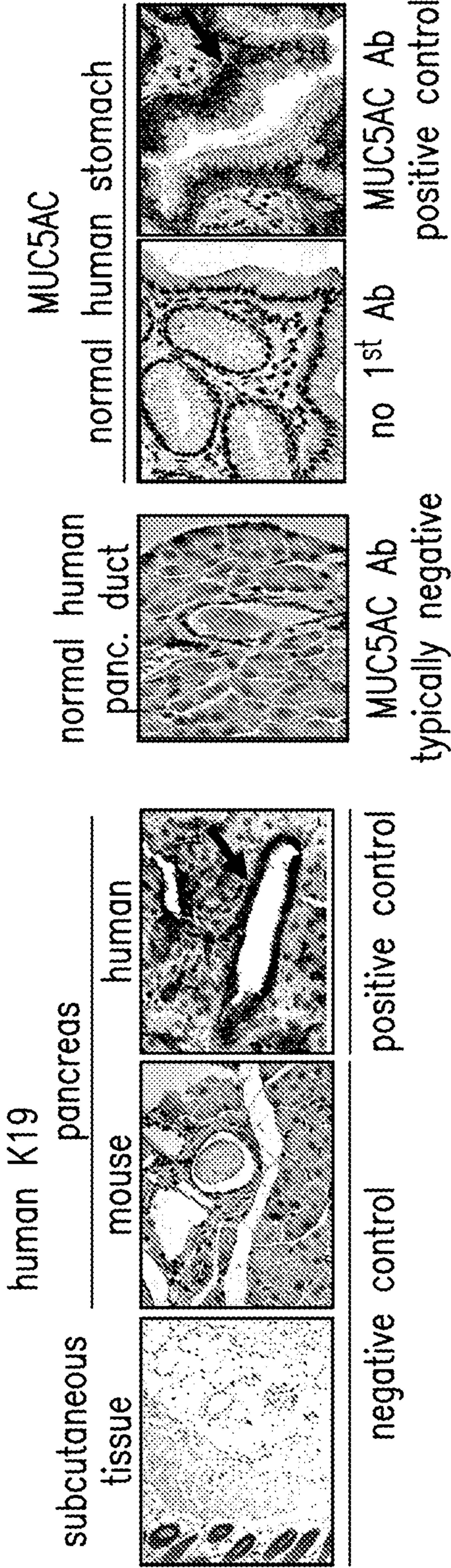
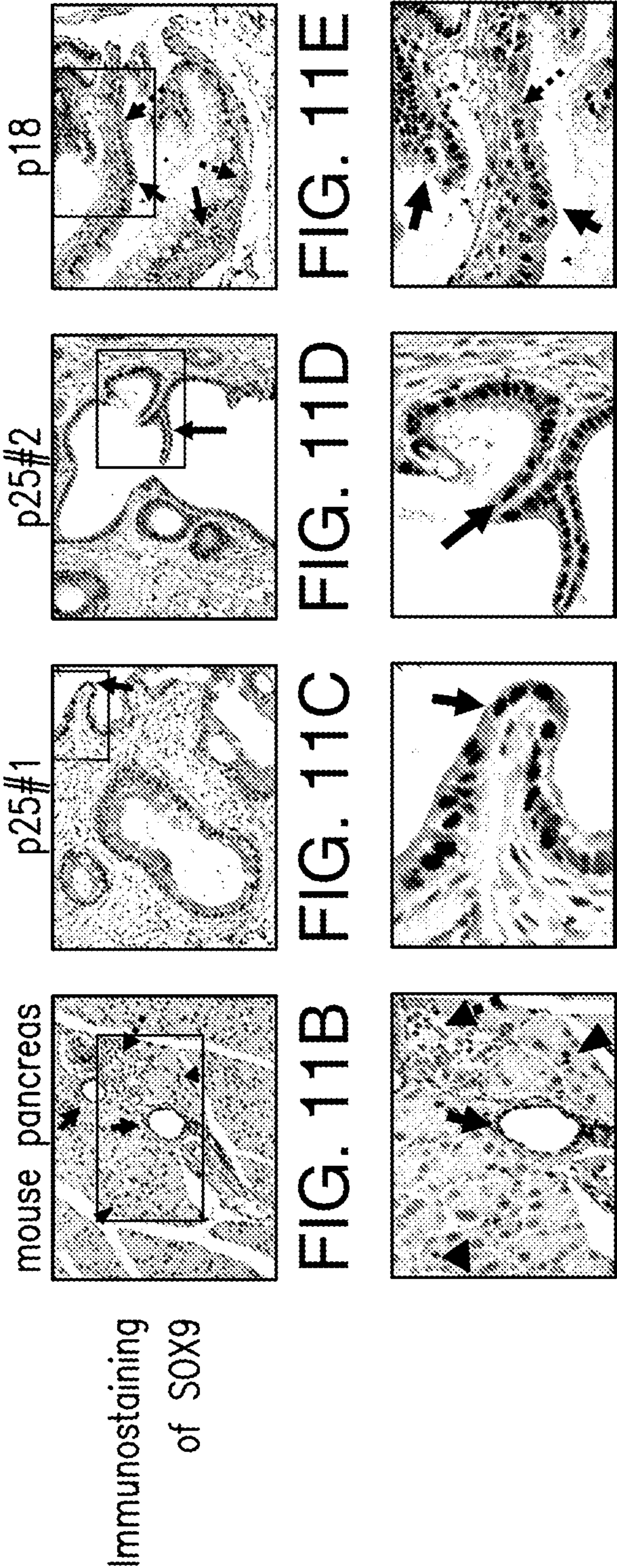


FIG. 11A

28/100

10-22 cancer iPS teratoma (3 month)



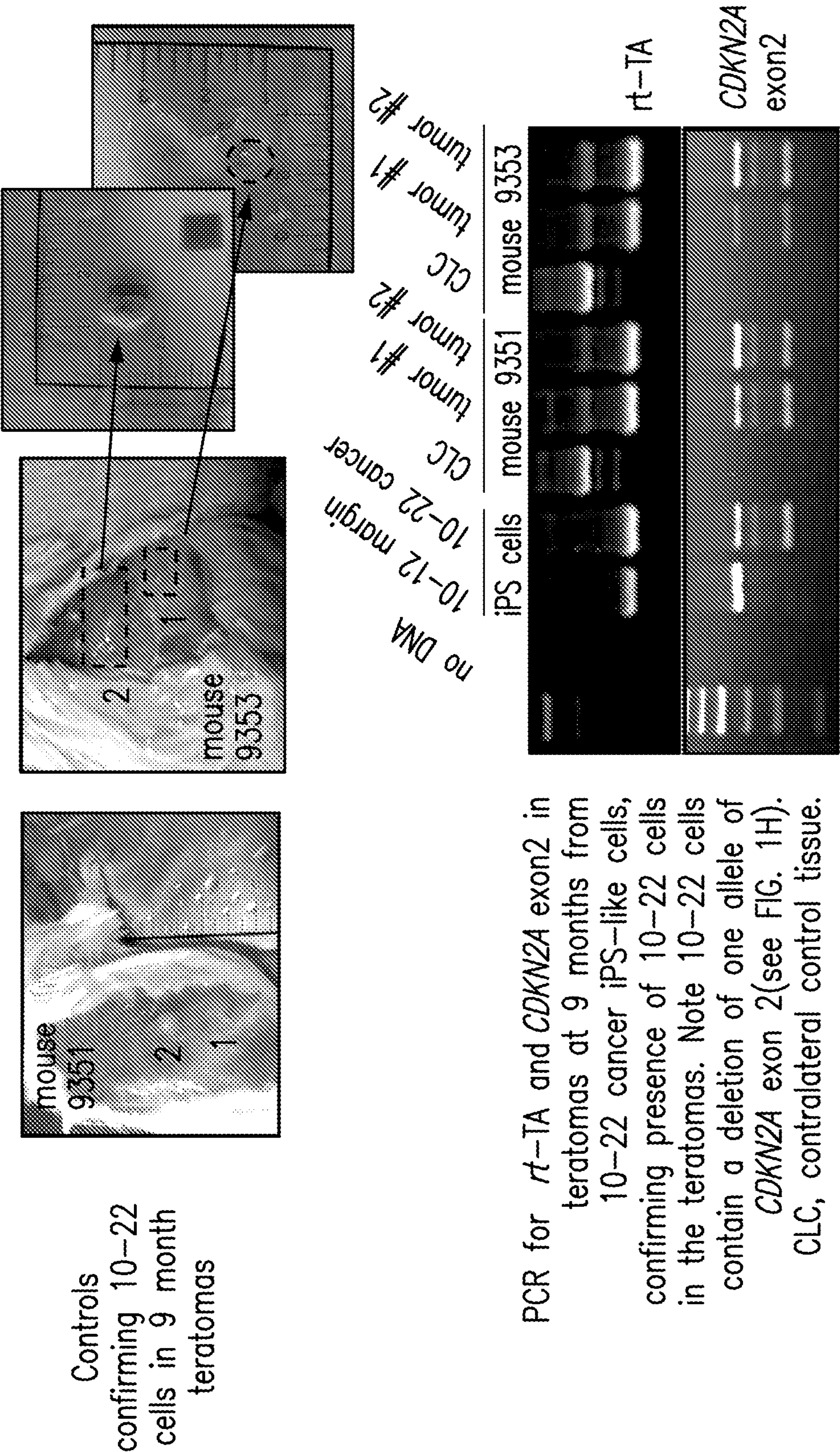
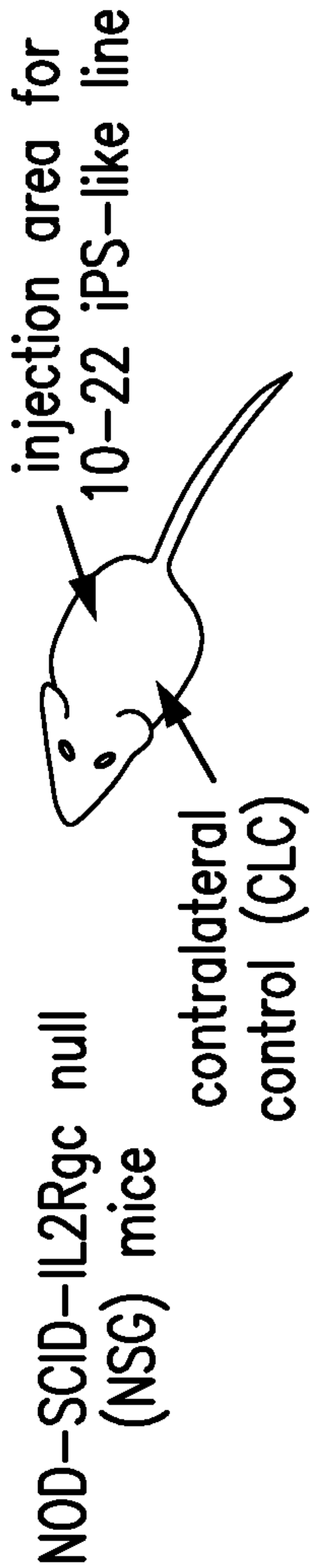


FIG. 11F

Scheme for generating organoid cultures from teratomas of 10-22 iPS-like cells



Day 0 Tissue was minced and dissociated in liberase T flex (1.3W/ml) at 37°C for 30 min, then cells embedded into Matrigel in serum-free medium

Day 6 Collect conditioned media for proteomic analysis

Day 9 To passage, cells were dissociated and re-embedded onto Matrigel

Day 19 Cells harvested for DNA and PCR analysis at 2nd passage (see B below)

Day 26 Collect organoids for IHC for MUC5AC and K19 staining (see FIG 5A-C)

FIG. 12A

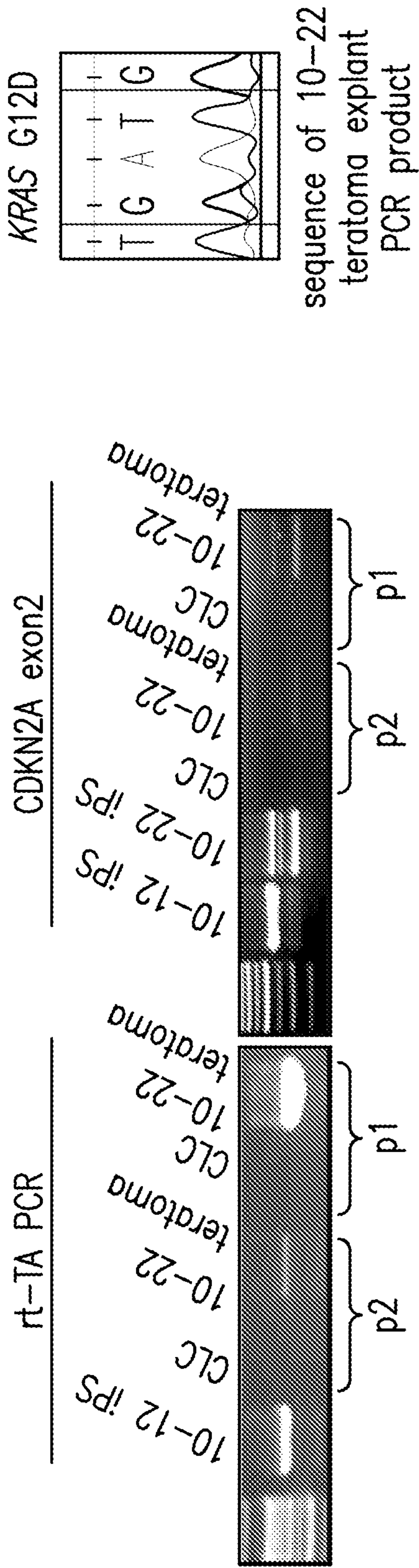


FIG. 12B

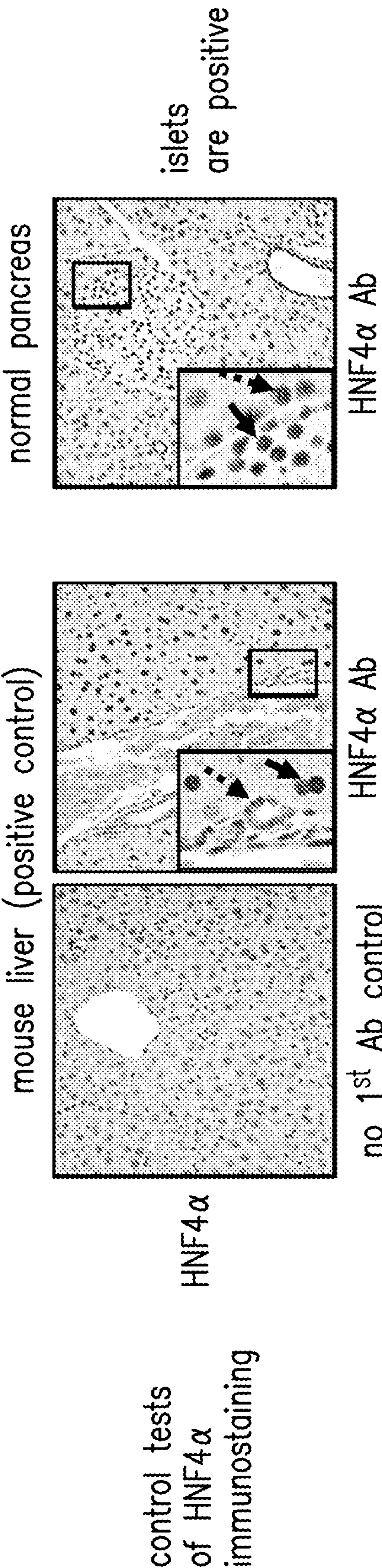


FIG. 12C

32/100

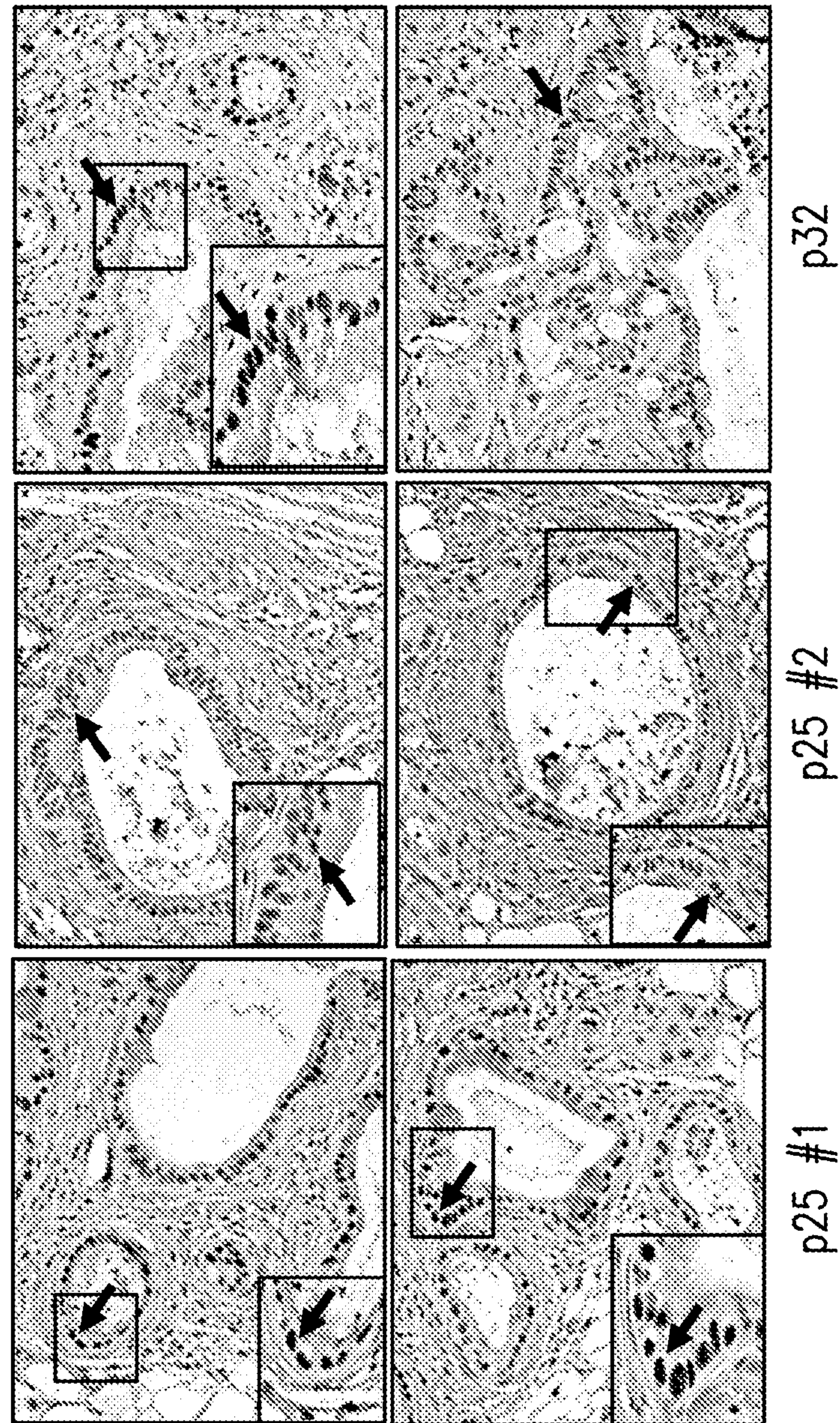


FIG. 12D

immunostaining
of HNF4α on
10-22 cancer
iPS-like teratomas
(3 months)

33/100

Pancreatic Cancer Cases vs. Controls – TCHP – 07MAY14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
TCHP	CASE=Control	10	26.62	0	0.000	173.657	0.394	0.98903
	CASE=Case	10	26.91	0	17.302	82.728		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
TCHP	Control	10	26.6214	0.000	0.000	0.000	2.077	173.657
TCHP	Resectable	5	6.9208	0.000	0.000	0.000	5.149	29.455
TCHP	Locally Advanced	2	46.6490	42.448	42.448	46.649	50.850	50.850
TCHP	Metastatic	3	47.0780	0.000	0.000	58.506	82.728	82.728

Kruskal-Wallis Test	
Chi-Square	3.6197
DF	3
Pr > Chi-Square	0.3056

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	40.0	Somers'D	0.210
Percent Discordant	19.0	Gamma	0.356
Percent Tied	41.0	Tau-a	0.111
Pairs	100	c	0.605

FIG. 13

34/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	24.3	Somers'D	-.214
Percent Discordant	45.7	Gamma	-.306
Percent Tied	30.0	Tau-a	-.110
Pairs	70	c	0.393

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	26.0	Somers'D	-.060
Percent Discordant	32.0	Gamma	-.103
Percent Tied	42.0	Tau-a	-.029
Pairs	50	c	0.470

FIG. 13 Continued

35/100

Pancreatic Cancer Cases vs. Controls – ABCA13 – 07MAY14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
ABCA13	CASE=Control	10	40.83	0	0	408.263	0.74	0.33056
	CASE=Case	10	0.00	0	0	0.000		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
ABCA13	Control	10	40.8263	0	0	0	0	408.263
ABCA13	Resectable	5	0.0000	0	0	0	0	0.000
ABCA13	Locally Advanced	2	0.0000	0	0	0	0	0.000
ABCA13	Metastatic	3	0.0000	0	0	0	0	0.000

Kruskal-Wallis Test	
Chi-Square	1.0000
DF	3
Pr > Chi-Square	0.8013

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.053
Pairs	100	c	0.550

FIG. 14

36/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.051
Pairs	70	c	0.550

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.048
Pairs	50	c	0.550

FIG. 14 Continued

37/100

Pancreatic Cancer Cases vs. Controls – STARD8 – 07MAY14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
STARD8	CASE=Control	10	566.56	289.178	610.702	869.45	0.144	0.10449
	CASE=Case	10	815.12	329.130	785.444	1668.62		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
STARD8	Control	10	566.558	289.178	335.352	610.702	747.99	869.45
STARD8	Resectable	5	826.075	374.418	678.076	962.615	965.58	1149.69
STARD8	Locally Advanced	2	610.971	329.130	329.130	610.971	892.81	892.81
STARD8	Metastatic	3	932.947	497.507	497.507	632.712	1668.62	1668.62

Kruskal-Wallis Test	
Chi-Square	3.4200
DF	3
Pr > Chi-Square	0.3313

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	70.0	Somers'D	0.400
Percent Discordant	30.0	Gamma	0.400
Percent Tied	0.0	Tau-a	0.211
Pairs	100	c	0.700

FIG. 15

38/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	72.9	Somers'D	0.457
Percent Discordant	27.1	Gamma	0.457
Percent Tied	0.0	Tau-a	0.235
Pairs	70	c	0.729

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	80.0	Somers'D	0.600
Percent Discordant	20.0	Gamma	0.600
Percent Tied	0.0	Tau-a	0.286
Pairs	50	c	0.800

FIG. 15 Continued

39/100

Pancreatic Cancer Cases vs. Controls – ATP2A1 – 07MAY14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
ATP2A1	CASE=Control	10	2609.71	1022.89	2158.12	6322.68	0.09	0.37134
	CASE=Case	10	3155.83	1986.38	3041.97	5669.19		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
ATP2A1	Control	10	2609.71	1022.89	1865.36	2158.12	2538.35	6322.68
ATP2A1	Resectable	5	2925.89	1986.38	2334.25	2701.02	3578.92	4028.89
ATP2A1	Locally Advanced	2	3394.06	3382.92	3382.92	3394.06	3405.19	3405.19
ATP2A1	Metastatic	3	3380.24	2087.22	2087.22	2384.31	5669.19	5669.19

Kruskal-Wallis Test	
Chi-Square	3.2295
DF	3
Pr > Chi-Square	0.3576

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	72.0	Somers'D	0.450
Percent Discordant	27.0	Gamma	0.455
Percent Tied	1.0	Tau-a	0.237
Pairs	100	c	0.725

FIG. 16

40/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	72.9	Somers'D	0.471
Percent Discordant	25.7	Gamma	0.478
Percent Tied	1.4	Tau-a	0.243
Pairs	70	c	0.736

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	68.0	Somers'D	0.400
Percent Discordant	28.0	Gamma	0.417
Percent Tied	4.0	Tau-a	0.190
Pairs	50	c	0.700

FIG. 16 Continued

41/100

Pancreatic Cancer Cases vs. Controls – FKBP10 – 07MAY14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
FKBP10	CASE=Control	10	2.93	0	0.000	15.235	0.144	0.017453
	CASE=Case	10	39.83	0	24.886	100.205		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
FKBP10	Control	10	2.9289	0	0	0.000	0.000	15.235
FKBP10	Resectable	5	31.0212	0	0	0.000	68.215	86.891
FKBP10	Locally Advanced	2	24.8860	0	0	24.886	49.772	49.772
FKBP10	Metastatic	3	64.4727	0	0	93.213	100.205	100.205

Kruskal-Wallis Test	
Chi-Square	4.5055
DF	3
Pr > Chi-Square	0.2118

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	50.0	Somers'D	0.400
Percent Discordant	10.0	Gamma	0.667
Percent Tied	40.0	Tau-a	0.211
Pairs	100	c	0.700

FIG. 17

42/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	42.9	Somers'D	0.314
Percent Discordant	11.4	Gamma	0.579
Percent Tied	45.7	Tau-a	0.162
Pairs	70	c	0.657

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	40.0	Somers'D	0.280
Percent Discordant	12.0	Gamma	0.538
Percent Tied	48.0	Tau-a	0.133
Pairs	50	c	0.640

FIG. 17 Continued

43/100

Pancreatic Cancer Cases vs. Controls – SCNA8 – 07MAY14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
SCNA8	CASE=Control	10	124.33	0	0	1243.34	0.74	0.33056
	CASE=Case	10	0.00	0	0	0.00		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
SCNA8	Control	10	124.33	0	0	0	0	1243.34
SCNA8	Resectable	5	0.000	0	0	0	0	0.000
SCNA8	Locally Advanced	2	0.000	0	0	0	0	0.000
SCNA8	Metastatic	3	0.000	0	0	0	0	0.000

Kruskal-Wallis Test	
Chi-Square	1.0000
DF	3
Pr > Chi-Square	0.8013

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.053
Pairs	100	c	0.550

FIG. 18

44/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.051
Pairs	70	c	0.550

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.048
Pairs	50	c	0.550

FIG. 18 Continued

45/100

Pancreatic Cancer Cases vs. Controls – TCF20 – 07MAY14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
TCF20	CASE=Control	10	792.5	387	707	1336	0.352	0.15492
	CASE=Case	10	1116.4	317	877	2014		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
TCF20	Control	10	792.50	387	600	707.0	977	1336
TCF20	Resectable	5	777.60	317	618	799.0	833	1321
TCF20	Locally Advanced	2	1189.50	545	545	1189.5	1834	1834
TCF20	Metastatic	3	1632.33	921	921	1962.0	2014	2014

Kruskal-Wallis Test	
Chi-Square	4.8067
DF	3
Pr > Chi-Square	0.1865

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	62.0	Somers'D	0.260
Percent Discordant	36.0	Gamma	0.265
Percent Tied	2.0	Tau-a	0.137
Pairs	100	c	0.630

FIG. 19

46/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	48.6	Somers'D	0.014
Percent Discordant	47.1	Gamma	0.015
Percent Tied	4.3	Tau-a	0.007
Pairs	70	c	0.507

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	46.0	Somers'D	0.000
Percent Discordant	46.0	Gamma	0.000
Percent Tied	8.0	Tau-a	0.000
Pairs	50	c	0.500

FIG. 19 Continued

47/100

Pancreatic Cancer Cases vs. Controls – SYNE1 – 07APR14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
SYNE1	CASE=Control	10	249.87	0	229.604	786.097	0.684	0.68925
	CASE=Case	10	208.74	0	149.883	680.775		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
SYNE1	Control	10	249.873	0.000	96.292	229.604	339.971	786.097
SYNE1	Resectable	5	69.689	0.000	0.000	48.682	126.089	173.676
SYNE1	Locally Advanced	2	549.428	418.080	418.080	549.428	680.775	680.775
SYNE1	Metastatic	3	213.368	19.183	19.183	240.622	380.298	380.298

Kruskal-Wallis Test	
Chi-Square	6.9581
DF	3
Pr > Chi-Square	0.0732

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	55.0	Somers'D	0.120
Percent Discordant	43.0	Gamma	0.122
Percent Tied	2.0	Tau-a	0.063
Pairs	100	c	0.560

FIG. 20

48/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	57.1	Somers'D	0.171
Percent Discordant	40.0	Gamma	0.176
Percent Tied	2.9	Tau-a	0.088
Pairs	70	c	0.586

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	76.0	Somers'D	0.560
Percent Discordant	20.0	Gamma	0.583
Percent Tied	4.0	Tau-a	0.267
Pairs	50	c	0.780

FIG. 20 Continued

49/100

Pancreatic Cancer Cases vs. Controls – UFD1L – 07APR14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
UFD1L	CASE=Control	10	3975.46	3208.68	3524.92	5675.10	0.166	0.088459
	CASE=Case	10	4865.56	2903.98	4504.49	6613.25		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
UFD1L	Control	10	3975.46	3208.68	3414.28	3524.92	4468.76	5675.10
UFD1L	Resectable	5	3917.36	2903.98	3394.29	4279.56	4468.52	4542.46
UFD1L	Locally Advanced	2	5977.96	5342.67	5342.67	5977.96	6613.25	6613.25
UFD1L	Metastatic	3	5704.29	4406.16	4406.16	6217.06	6489.65	6489.65

Kruskal-Wallis Test	
Chi-Square	7.1800
DF	3
Pr > Chi-Square	0.0664

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	69.0	Somers'D	0.390
Percent Discordant	30.0	Gamma	0.394
Percent Tied	1.0	Tau-a	0.205
Pairs	100	c	0.695

FIG. 21

50/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	60.0	Somers'D	0.214
Percent Discordant	38.6	Gamma	0.217
Percent Tied	1.4	Tau-a	0.110
Pairs	70	c	0.607

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	48.0	Somers'D	0.000
Percent Discordant	48.0	Gamma	0.000
Percent Tied	4.0	Tau-a	0.000
Pairs	50	c	0.500

FIG. 21 Continued

51/100

Pancreatic Cancer Cases vs. Controls – FLRT3 – 07APR14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
FLRT3	CASE=Control	10	134.22	0	0	1331.39	0.482	0.41303
	CASE=Case	10	22.40	0	0	84.86		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
FLRT3	Control	10	134.223	0	0	0.000	0.000	1331.39
FLRT3	Resectable	5	19.858	0	0	0.000	14.433	84.86
FLRT3	Locally Advanced	2	28.301	0	0	28.301	56.602	56.60
FLRT3	Metastatic	3	22.688	0	0	0.000	68.063	68.06

Kruskal-Wallis Test	
Chi-Square	0.9500
DF	3
Pr > Chi-Square	0.8133

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	16.0	Somers'D	-.200
Percent Discordant	36.0	Gamma	-.385
Percent Tied	48.0	Tau-a	-.105
Pairs	100	c	0.400

FIG. 22

52/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	15.7	Somers'D	-.229
Percent Discordant	38.6	Gamma	-.421
Percent Tied	45.7	Tau-a	-.118
Pairs	70	c	0.386

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	16.0	Somers'D	-.200
Percent Discordant	36.0	Gamma	-.385
Percent Tied	48.0	Tau-a	-.095
Pairs	50	c	0.400

FIG. 22 Continued

53/100

Pancreatic Cancer Cases vs. Controls – TOP2B – 07APR14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
TOP2B	CASE=Control	10	2200.57	1217.86	2229.61	3516.31	.006	0.017887
	CASE=Case	10	4756.27	1869.09	3257.13	11830.39		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
TOP2B	Control	10	2200.57	1217.86	1724.26	2229.61	2613.84	3516.31
TOP2B	Resectable	5	2803.83	1869.09	2709.30	2926.53	2970.36	3543.90
TOP2B	Locally Advanced	2	6370.55	6245.86	6245.86	6370.55	6495.25	6495.25
TOP2B	Metastatic	3	6934.14	2786.54	2786.54	6185.50	11830.39	11830.39

Kruskal-Wallis Test	
Chi-Square	10.0229
DF	3
Pr > Chi-Square	0.0184

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	86.0	Somers'D	0.720
Percent Discordant	14.0	Gamma	0.720
Percent Tied	0.0	Tau-a	0.379
Pairs	100	c	0.860

FIG. 23

54/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	82.9	Somers'D	0.657
Percent Discordant	17.1	Gamma	0.657
Percent Tied	0.0	Tau-a	0.338
Pairs	70	c	0.829

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	76.0	Somers'D	0.520
Percent Discordant	24.0	Gamma	0.520
Percent Tied	0.0	Tau-a	0.248
Pairs	50	c	0.760

FIG. 23 Continued

55/100

Pancreatic Cancer Cases vs. Controls – ZHX2 – 07APR14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
ZHX2	CASE=Control	10	7034.26	3508.07	5902.30	13307.83	0.63	0.27253
	CASE=Case	10	10214.79	4235.31	7230.32	30731.86		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
ZHX2	Control	10	7034.26	3508.07	4922.95	5902.30	7917.84	13307.83
ZHX2	Resectable	5	5902.83	4235.31	4614.13	5725.35	7235.74	7703.64
ZHX2	Locally Advanced	2	11690.08	7224.91	7224.91	11690.08	16155.25	16155.25
ZHX2	Metastatic	3	16417.87	4375.09	4375.09	14146.63	30731.88	30731.88

Kruskal-Wallis Test	
Chi-Square	2.9124
DF	3
Pr > Chi-Square	0.4053

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	56.0	Somers'D	0.140
Percent Discordant	42.0	Gamma	0.143
Percent Tied	2.0	Tau-a	0.074
Pairs	100	c	0.570

FIG. 24

56/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	50.0	Somers'D	0.029
Percent Discordant	47.1	Gamma	0.029
Percent Tied	2.9	Tau-a	0.015
Pairs	70	c	0.514

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	60.0	Somers'D	0.220
Percent Discordant	38.0	Gamma	0.224
Percent Tied	2.0	Tau-a	0.105
Pairs	50	c	0.610

FIG. 24 Continued

57/100

Pancreatic Cancer Cases vs. Controls – LIMCH1 – 07APR14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
LIMCH1	CASE=Control	10	2.79	0	0.0000	27.879	0.09	0.11660
	CASE=Case	10	105.18	0	27.6295	636.213		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
LIMCH1	Control	10	2.788	0.000	0.000	0.000	0.000	27.879
LIMCH1	Resectable	5	11.415	0.000	0.000	0.000	0.000	57.074
LIMCH1	Locally Advanced	2	78.854	0.000	0.000	78.854	157.707	157.707
LIMCH1	Metastatic	3	278.992	55.259	55.259	145.503	636.213	636.213

Kruskal-Wallis Test	
Chi-Square	10.4467
DF	3
Pr > Chi-Square	0.0151

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	50.0	Somers'D	0.450
Percent Discordant	5.0	Gamma	0.818
Percent Tied	45.0	Tau-a	0.237
Pairs	100	c	0.725

FIG. 25

58/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	28.6	Somers'D	0.214
Percent Discordant	7.1	Gamma	0.600
Percent Tied	64.3	Tau-a	0.110
Pairs	70	c	0.607

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	20.0	Somers'D	0.120
Percent Discordant	8.0	Gamma	0.429
Percent Tied	72.0	Tau-a	0.057
Pairs	50	c	0.560

FIG. 25 Continued

59/100

Pancreatic Cancer Cases vs. Controls – THBS2 – 24OCT13

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
THBS2	CASE=Control	10	24.85	16.557	25.6015	34.068	0.052	0.068339
	CASE=Case	10	49.69	17.978	38.5965	155.413		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
THBS2	Control	10	24.8493	16.557	22.529	25.6015	27.898	34.068
THBS2	Resectable	5	46.0672	22.228	37.249	49.9730	56.592	64.294
THBS2	Locally Advanced	2	97.6785	39.944	39.944	97.6785	155.413	155.413
THBS2	Metastatic	3	23.7517	17.978	17.978	24.7280	28.549	28.549

Kruskal-Wallis Test	
Chi-Square	8.4067
DF	3
Pr > Chi-Square	0.0383

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	76.0	Somers'D	0.520
Percent Discordant	24.0	Gamma	0.520
Percent Tied	0.0	Tau-a	0.274
Pairs	100	c	0.760

FIG. 26

60/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	88.6	Somers'D	0.771
Percent Discordant	11.4	Gamma	0.771
Percent Tied	0.0	Tau-a	0.397
Pairs	70	c	0.886

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	84.0	Somers'D	0.680
Percent Discordant	16.0	Gamma	0.680
Percent Tied	0.0	Tau-a	0.324
Pairs	50	c	0.840

FIG. 26 Continued

61/100

Pancreatic Cancer Cases vs. Controls – Shroom3 – 24OCT13

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
Shroom3	CASE=Control	10	13.82	0	0	129.028	0.528	0.71068
	CASE=Case	10	8.60	0	0	47.992		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
Shroom3	Control	10	13.8174	0	0	0.00	0.000	129.028
Shroom3	Resectable	5	2.1058	0	0	0.00	0.176	10.353
Shroom3	Locally Advanced	2	0.0000	0	0	0.00	0.000	0.000
Shroom3	Metastatic	3	25.1573	0	0	27.48	47.992	47.992

Kruskal-Wallis Test	
Chi-Square	3.6147
DF	3
Pr > Chi-Square	0.3062

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	17.0	Somers'D	-.100
Percent Discordant	27.0	Gamma	-.227
Percent Tied	56.0	Tau-a	-.053
Pairs	100	c	0.450

FIG. 27

62/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	18.6	Somers'D	-.057
Percent Discordant	24.3	Gamma	-.133
Percent Tied	57.1	Tau-a	-.029
Pairs	70	c	0.471

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	18.0	Somers'D	0.000
Percent Discordant	18.0	Gamma	0.000
Percent Tied	64.0	Tau-a	0.000
Pairs	50	c	0.500

FIG. 27 Continued

63/100

Pancreatic Cancer Cases vs. Controls – HMOX1 – 24OCT13

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
HMOX1	CASE=Control	10	73.47	0	0.0	328.5	0.436	0.48123
	CASE=Case	10	108.09	0	133.3	250.9		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
HMOX1	Control	10	73.470	0	0	0.00	111.9	328.5
HMOX1	Resectable	5	118.380	0	0	140.50	200.5	250.9
HMOX1	Locally Advanced	2	78.350	0	0	78.35	156.7	156.7
HMOX1	Metastatic	3	110.767	0	0	126.10	206.2	206.2

Kruskal-Wallis Test	
Chi-Square	0.9397
DF	3
Pr > Chi-Square	0.8158

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	49.0	Somers'D	0.220
Percent Discordant	27.0	Gamma	0.289
Percent Tied	24.0	Tau-a	0.116
Pairs	100	c	0.610

FIG. 28

64/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	47.1	Somers'D	0.200
Percent Discordant	27.1	Gamma	0.269
Percent Tied	25.7	Tau-a	0.103
Pairs	70	c	0.600

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	50.0	Somers'D	0.240
Percent Discordant	26.0	Gamma	0.316
Percent Tied	24.0	Tau-a	0.114
Pairs	50	c	0.620

FIG. 28 Continued

65/100

Pancreatic Cancer Cases vs. Controls – LOXL3 – 24OCT13

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
LOXL3	CASE=Control	10	925.44	0	0	7742.69	0.97	0.90599
	CASE=Case	10	1045.56	0	0	6315.23		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
LOXL3	Control	10	925.44	0.000	0.000	0.00	593.35	7742.59
LOXL3	Resectable	5	0.00	0.000	0.000	0.00	0.00	0.00
LOXL3	Locally Advanced	2	1335.56	13.116	13.116	1335.56	2658.01	2658.01
LOXL3	Metastatic	3	2594.82	0.000	0.000	1469.22	6315.23	6315.23

Kruskal-Wallis Test	
Chi-Square	6.2035
DF	3
Pr > Chi-Square	0.1021

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	27.0	Somers'D	-.040
Percent Discordant	31.0	Gamma	-.069
Percent Tied	42.0	Tau-a	-.021
Pairs	100	c	0.480

FIG. 29

66/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	35.7	Somers'D	0.229
Percent Discordant	12.9	Gamma	0.471
Percent Tied	51.4	Tau-a	0.118
Pairs	70	c	0.614

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	40.0	Somers'D	0.400
Percent Discordant	0.0	Gamma	1.000
Percent Tied	60.0	Tau-a	0.190
Pairs	50	c	0.700

FIG. 29 Continued

67/100

Pancreatic Cancer Cases vs. Controls – Obscurin – 24OCT13

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
Obscurin	CASE=Control	10	17.76	0	0	177.59	0.74	0.35712
	CASE=Case	10	137.23	0	0	1259.03		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
OBSCURIN	Control	10	17.759	0	0	0.000	0.00	177.59
OBSCURIN	Resectable	5	0.000	0	0	0.000	0.00	0.00
OBSCURIN	Locally Advanced	2	56.617	0	0	56.617	113.23	113.23
OBSCURIN	Metastatic	3	419.676	0	0	0.000	1259.03	1259.03

Kruskal-Wallis Test	
Chi-Square	3.3885
DF	3
Pr > Chi-Square	0.3355

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	19.0	Somers'D	0.100
Percent Discordant	9.0	Gamma	0.357
Percent Tied	72.0	Tau-a	0.053
Pairs	100	c	0.550

FIG. 30

68/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	-.029
Percent Discordant	12.9	Gamma	-.125
Percent Tied	77.1	Tau-a	-.015
Pairs	70	c	0.486

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.048
Pairs	50	c	0.550

FIG. 30 Continued

69/100

Pancreatic Cancer Cases vs. Controls – Malectin – 24OCT13

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
Malectin	CASE=Control	10	13.61	0	0	132.759	0.97	0.44841
	CASE=Case	10	46.61	0	0	406.892		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
MALECTIN	Control	10	13.613	0	0	0.000	0.00	132.759
MALECTIN	Resectable	5	11.836	0	0	0.000	0.00	59.180
MALECTIN	Locally Advanced	2	203.446	0	0	203.446	406.892	406.892
MALECTIN	Metastatic	3	0.000	0	0	0.000	0.000	0.000

Kruskal-Wallis Test	
Chi-Square	2.3545
DF	3
Pr > Chi-Square	0.5022

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	19.0	Somers'D	0.020
Percent Discordant	17.0	Gamma	0.056
Percent Tied	64.0	Tau-a	0.011
Pairs	100	c	0.510

FIG. 31

70/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	27.1	Somers'D	0.114
Percent Discordant	15.7	Gamma	0.267
Percent Tied	57.1	Tau-a	0.059
Pairs	70	c	0.557

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	-.080
Percent Discordant	18.0	Gamma	-.286
Percent Tied	72.0	Tau-a	-.038
Pairs	50	c	0.460

FIG. 31 Continued

71/100

Pancreatic Cancer Cases vs. Controls – DNAH5 – 24OCT13

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
DNAH5	CASE=Control	10	8.05	0	0	80.475	0.74	0.45071
	CASE=Case	10	23.47	0	0	181.223		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
DNAH5	Control	10	8.0475	0	0	0.0000	0.000	80.475
DNAH5	Resectable	5	10.7000	0	0	0.0000	0.000	53.500
DNAH5	Locally Advanced	2	90.6115	0	0	90.6115	181.223	181.223
DNAH5	Metastatic	3	0.0000	0	0	0.0000	0.000	0.000

Kruskal-Wallis Test	
Chi-Square	3.0829
DF	3
Pr > Chi-Square	0.3790

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	19.0	Somers'D	0.100
Percent Discordant	9.0	Gamma	0.357
Percent Tied	72.0	Tau-a	0.053
Pairs	100	c	0.550

FIG. 32

72/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	27.1	Somers'D	0.186
Percent Discordant	8.6	Gamma	0.520
Percent Tied	64.3	Tau-a	0.096
Pairs	70	c	0.593

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	18.0	Somers'D	0.080
Percent Discordant	10.0	Gamma	0.286
Percent Tied	72.0	Tau-a	0.038
Pairs	50	c	0.540

FIG. 32 Continued

73/100

Pancreatic Cancer Cases vs. Controls – ca199 – 24OCT13

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
ca199	CASE=Control	6	15.33	0	16.0	32	0	.002811233
	CASE=Case	8	849.25	127	907.5	1482		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
CA199	Control	6	15.33	0	9	16.0	19.0	32
CA199	Resectable	4	802.75	127	294	884.5	1311.5	1315
CA199	Locally Advanced	1	286.00	286	286	286.0	286.0	286
CA199	Metastatic	3	1099.00	577	577	1238.0	1482.0	1482

Kruskal-Wallis Test	
Chi-Square	10.2048
DF	3
Pr > Chi-Square	0.0169

Note: 6 observation were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	100.0	Somers'D	1.000
Percent Discordant	0.0	Gamma	1.000
Percent Tied	0.0	Tau-a	0.527
Pairs	48	c	1.000

FIG. 33

74/100

Note: 9 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	100.0	Somers'D	1.000
Percent Discordant	0.0	Gamma	1.000
Percent Tied	0.0	Tau-a	0.545
Pairs	30	c	1.000

Note: 10 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	100.0	Somers'D	1.000
Percent Discordant	0.0	Gamma	1.000
Percent Tied	0.0	Tau-a	0.533
Pairs	24	c	1.000

FIG. 33 Continued

75/100

Pancreatic Cancer Cases vs. Controls – AFP – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
AFP	CASE=Control	10	3.01	1.427	2.5590	6.672	0.912	0.67884
	CASE=Case	10	2.77	1.551	2.5465	4.101		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
AFP	Control	10	3.01010	1.427	2.092	2.559	3.652	6.672
AFP	Resectable	5	2.92640	2.105	2.399	2.484	3.815	3.829
AFP	Locally Advanced	2	2.39900	1.674	1.674	2.399	3.124	3.124
AFP	Metastatic	3	2.75367	1.551	1.551	2.609	4.101	4.101

Kruskal–Wallis Test	
Chi–Square	0.4152
DF	3
Pr > Chi–Square	0.9371

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	47.0	Somers'D	–.040
Percent Discordant	51.0	Gamma	–.041
Percent Tied	2.0	Tau–a	–.021
Pairs	100	c	0.480

FIG. 34

76/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	45.7	Somers'D	-.071
Percent Discordant	52.9	Gamma	-.072
Percent Tied	1.4	Tau-a	-.037
Pairs	70	c	0.464

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	40.0	Somers'D	-.160
Percent Discordant	56.0	Gamma	-.167
Percent Tied	4.0	Tau-a	-.076
Pairs	50	c	0.420

FIG. 34 Continued

77/100

Pancreatic Cancer Cases vs. Controls – RTTN – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
RTTN	CASE=Control	10	1491.75	1118.56	1425.83	1993.80	0.528	0.14254
	CASE=Case	10	1932.71	885.06	1657.87	3228.66		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
RTTN	Control	10	1491.75	1118.56	1332.58	1425.83	1600.31	1993.80
RTTN	Resectable	5	1280.60	885.06	1118.26	1164.60	1401.94	1833.15
RTTN	Locally Advanced	2	1916.31	1482.59	1482.59	1916.31	2350.03	2350.03
RTTN	Metastatic	3	3030.48	2704.08	2704.08	3158.69	3228.66	3228.66

Kruskal-Wallis Test	
Chi-Square	10.2800
DF	3
Pr > Chi-Square	0.0163

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	59.0	Somers'D	0.190
Percent Discordant	40.0	Gamma	0.192
Percent Tied	1.0	Tau-a	0.100
Pairs	100	c	0.595

FIG. 35

78/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	57.1	Somers'D	0.157
Percent Discordant	41.4	Gamma	0.159
Percent Tied	1.4	Tau-a	0.081
Pairs	70	c	0.579

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	72.0	Somers'D	0.460
Percent Discordant	26.0	Gamma	0.469
Percent Tied	2.0	Tau-a	0.219
Pairs	50	c	0.730

FIG. 35 Continued

79/100

Pancreatic Cancer Cases vs. Controls – NLRX1 – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
NLRX1	CASE=Control	10	1163.35	54.796	831.546	5215.26	0.912	0.63522
	CASE=Case	10	1568.24	33.396	527.514	6379.51		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
NLRX1	Control	10	1163.35	54.80	203.46	831.55	1105.06	5215.26
NLRX1	Resectable	5	729.16	33.40	398.39	466.76	588.27	2158.99
NLRX1	Locally Advanced	2	2757.14	1113.71	1113.71	2757.14	4400.57	4400.57
NLRX1	Metastatic	3	2174.11	69.35	69.35	73.46	6379.51	6379.51

Kruskal-Wallis Test	
Chi-Square	2.4495
DF	3
Pr > Chi-Square	0.4845

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	46.0	Somers'D	-.020
Percent Discordant	48.0	Gamma	-.021
Percent Tied	6.0	Tau-a	-.011
Pairs	100	c	0.490

FIG. 36

80/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	51.4	Somers'D	0.057
Percent Discordant	45.7	Gamma	0.059
Percent Tied	2.9	Tau-a	0.029
Pairs	70	c	0.529

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	56.0	Somers'D	0.160
Percent Discordant	40.0	Gamma	0.167
Percent Tied	4.0	Tau-a	0.076
Pairs	50	c	0.580

FIG. 36 Continued

81/100

Pancreatic Cancer Cases vs. Controls – DNAH12 – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
DNAH12	CASE=Control	10	128.15	0.00	151.752	238.82	0	.005846218
	CASE=Case	10	863.79	297.58	544.855	2600.00		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
DNAH12	Control	10	128.15	0.000	41.066	151.75	203.37	238.82
DNAH12	Resectable	5	656.44	297.580	299.700	531.03	558.68	1595.23
DNAH12	Locally Advanced	2	1001.03	963.268	963.268	1001.03	1038.80	1038.80
DNAH12	Metastatic	3	1117.86	329.149	329.149	424.43	2600.00	2600.00

Kruskal-Wallis Test	
Chi-Square	14.6324
DF	3
Pr > Chi-Square	0.0022

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	100.0	Somers'D	1.000
Percent Discordant	0.0	Gamma	1.000
Percent Tied	0.0	Tau-a	0.526
Pairs	100	c	1.000

FIG. 37

82/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	100.0	Somers'D	1.000
Percent Discordant	0.0	Gamma	1.000
Percent Tied	0.0	Tau-a	0.515
Pairs	70	c	1.000

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	100.0	Somers'D	1.000
Percent Discordant	0.0	Gamma	1.000
Percent Tied	0.0	Tau-a	0.476
Pairs	50	c	1.000

FIG. 37 Continued

83/100

Pancreatic Cancer Cases vs. Controls – ODZ3 – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
ODZ3	CASE=Control	10	1.19	0	0	11.886	0.684	0.20831
	CASE=Case	10	15.34	0	0	99.517		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
ODZ3	Control	10	1.1886	0	0	0	0.000	11.886
ODZ3	Resectable	5	10.7728	0	0	0	0.000	53.864
ODZ3	Locally Advanced	2	0.0000	0	0	0	0.000	0.000
ODZ3	Metastatic	3	33.1723	0	0	0	99.517	99.517

Kruskal-Wallis Test	
Chi-Square	1.7324
DF	3
Pr > Chi-Square	0.6297

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	20.0	Somers'D	0.120
Percent Discordant	8.0	Gamma	0.429
Percent Tied	72.0	Tau-a	0.063
Pairs	100	c	0.560

FIG. 38

84/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	14.3	Somers'D	0.057
Percent Discordant	8.6	Gamma	0.250
Percent Tied	77.1	Tau-a	0.029
Pairs	70	c	0.529

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	20.0	Somers'D	0.120
Percent Discordant	8.0	Gamma	0.429
Percent Tied	72.0	Tau-a	0.057
Pairs	50	c	0.560

FIG. 38 Continued

85/100

Pancreatic Cancer Cases vs. Controls – ADAMST9 – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
ADAMST9	CASE=Control	10	36.18	21.739	36.845	46.032	0.436	0.69448
	CASE=Case	10	34.15	16.615	32.647	62.069		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
ADAMST9	Control	10	36.1780	21.739	29.274	36.845	43.381	46.032
ADAMST9	Resectable	5	36.0142	25.063	31.884	33.410	36.885	52.829
ADAMST9	Locally Advanced	2	29.6940	24.661	24.661	29.694	34.727	34.727
ADAMST9	Metastatic	3	34.0233	16.615	16.615	23.386	62.069	62.069

Kruskal-Wallis Test	
Chi-Square	1.4314
DF	3
Pr > Chi-Square	0.6982

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	61.0	Somers'D	0.220
Percent Discordant	39.0	Gamma	0.220
Percent Tied	0.0	Tau-a	0.116
Pairs	100	c	0.610

FIG. 39

86/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	60.0	Somers'D	0.200
Percent Discordant	40.0	Gamma	0.200
Percent Tied	0.0	Tau-a	0.103
Pairs	70	c	0.600

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	50.0	Somers'D	0.080
Percent Discordant	42.0	Gamma	0.087
Percent Tied	8.0	Tau-a	0.038
Pairs	50	c	0.540

FIG. 39 Continued

87/100

Pancreatic Cancer Cases vs. Controls – TPM – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
TPM	CASE=Control	10	35.55	0	13.344	132.267	0.314	0.77063
	CASE=Case	10	28.50	0	0.000	151.577		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
TPM	Control	10	35.5519	0	0	13.3440	70.377	132.267
TPM	Resectable	5	0.0000	0	0	0.0000	0.000	0.000
TPM	Locally Advanced	2	75.7885	0	0	75.7885	151.577	151.577
TPM	Metastatic	3	44.4720	0	0	0.0000	133.416	133.416

Kruskal-Wallis Test	
Chi-Square	4.1788
DF	3
Pr > Chi-Square	0.2428

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	48.0	Somers'D	0.290
Percent Discordant	19.0	Gamma	0.433
Percent Tied	33.0	Tau-a	0.153
Pairs	100	c	0.645

FIG. 40

88/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	51.4	Somers'D	0.371
Percent Discordant	14.3	Gamma	0.565
Percent Tied	34.3	Tau-a	0.191
Pairs	70	c	0.686

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	60.0	Somers'D	0.600
Percent Discordant	0.0	Gamma	1.000
Percent Tied	40.0	Tau-a	0.286
Pairs	50	c	0.800

FIG. 40 Continued

89/100

Pancreatic Cancer Cases vs. Controls – DNAH1 – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
DNAH1	CASE=Control	10	57.52	0	40.618	162.678	0.076	0.024957
	CASE=Case	10	7.31	0	0.000	40.098		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
DNAH1	Control	10	57.5161	0	0	40.618	97.673	162.678
DNAH1	Resectable	5	8.0196	0	0	0.000	0.000	40.098
DNAH1	Locally Advanced	2	6.1570	0	0	6.157	12.314	12.314
DNAH1	Metastatic	3	6.9080	0	0	10.161	10.563	10.563

Kruskal-Wallis Test	
Chi-Square	4.0020
DF	3
Pr > Chi-Square	0.2612

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	65.0	Somers'D	0.480
Percent Discordant	17.0	Gamma	0.585
Percent Tied	18.0	Tau-a	0.253
Pairs	100	c	0.740

FIG. 41

90/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	65.7	Somers'D	0.529
Percent Discordant	12.9	Gamma	0.673
Percent Tied	21.4	Tau-a	0.272
Pairs	70	c	0.764

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	66.0	Somers'D	0.560
Percent Discordant	10.0	Gamma	0.737
Percent Tied	24.0	Tau-a	0.267
Pairs	50	c	0.780

FIG. 41 Continued

91/100

Pancreatic Cancer Cases vs. Controls – PMBP1 – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
PMBP1	CASE=Control	10	16.82	0	0	168.23	0.74	0.33056
	CASE=Case	10	0.00	0	0	0.00		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
PMBP1	Control	10	16.823	0	0	0	0	168.23
PMBP1	Resectable	5	0.000	0	0	0	0	0.00
PMBP1	Locally Advanced	2	0.000	0	0	0	0	0.00
PMBP1	Metastatic	3	0.000	0	0	0	0	0.00

Kruskal-Wallis Test	
Chi-Square	1.0000
DF	3
Pr > Chi-Square	0.8013

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.053
Pairs	100	c	0.550

FIG. 42

92/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.051
Pairs	70	c	0.550

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.048
Pairs	50	c	0.550

FIG. 42 Continued

93/100

Pancreatic Cancer Cases vs. Controls – DNAH17 – 17JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
DNAH17	CASE=Control	10	1.15	0	0	11.604	0.352	0.76837
	CASE=Case	10	1.59	0	0	7.356		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
DNAH17	Control	10	1.1604	0.000	0.000	0.000	0.000	11.604
DNAH17	Resectable	5	0.0000	0.000	0.000	0.000	0.000	0.000
DNAH17	Locally Advanced	2	3.7960	0.236	0.236	3.796	7.356	7.356
DNAH17	Metastatic	3	2.7780	0.000	0.000	2.875	5.459	5.459

Kruskal-Wallis Test	
Chi-Square	9.3881
DF	3
Pr > Chi-Square	0.0246

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	36.0	Somers'D	0.260
Percent Discordant	10.0	Gamma	0.565
Percent Tied	54.0	Tau-a	0.137
Pairs	100	c	0.630

FIG. 43

94/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	-.029
Percent Discordant	12.9	Gamma	-.125
Percent Tied	77.1	Tau-a	-.015
Pairs	70	c	0.486

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Warning: The maximum likelihood estimate may not exist.

Warning: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	0.100
Percent Discordant	0.0	Gamma	1.000
Percent Tied	90.0	Tau-a	0.048
Pairs	50	c	0.550

FIG. 43 Continued

95/100

Pancreatic Cancer Cases vs. Controls – EPHB1 – 02JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
EPHB1	CASE=Control	10	28.73	0	0.0000	275.285	0.144	0.82116
	CASE=Case	10	22.10	0	11.0265	85.685		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
EPHB1	Control	10	28.7284	0.000	0.000	0.0000	5.312	275.285
EPHB1	Resectable	5	24.9686	0.000	0.000	0.2700	38.888	85.685
EPHB1	Locally Advanced	2	35.6865	18.395	18.395	35.6865	52.978	52.978
EPHB1	Metastatic	3	8.2580	0.000	0.000	3.6580	21.116	21.116

Kruskal-Wallis Test	
Chi-Square	3.8296
DF	3
Pr > Chi-Square	0.2805

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	18.0	Somers'D	-.340
Percent Discordant	52.0	Gamma	-.486
Percent Tied	30.0	Tau-a	-.179
Pairs	100	c	0.330

FIG. 44

96/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	10.0	Somers'D	-.286
Percent Discordant	38.6	Gamma	-.588
Percent Tied	51.4	Tau-a	-.147
Pairs	70	c	0.357

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	22.0	Somers'D	-.140
Percent Discordant	36.0	Gamma	-.241
Percent Tied	42.0	Tau-a	-.067
Pairs	50	c	0.430

FIG. 44 Continued

97/100

Pancreatic Cancer Cases vs. Controls – DOS – 02JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
DOS	CASE=Control	10	279.3	68	230	562	0.74	0.72804
	CASE=Case	10	304.4	0	345	520		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
DOS	Control	10	279.3	68	123	230	416	562
DOS	Resectable	5	345.0	180	228	338	459	520
DOS	Locally Advanced	2	196.0	0	0	196	392	392
DOS	Metastatic	3	309.0	215	215	352	360	360

Kruskal-Wallis Test	
Chi-Square	1.0982
DF	3
Pr > Chi-Square	0.7775

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	54.0	Somers'D	0.090
Percent Discordant	45.0	Gamma	0.091
Percent Tied	1.0	Tau-a	0.047
Pairs	100	c	0.545

FIG. 45

98/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	54.3	Somers'D	0.086
Percent Discordant	45.7	Gamma	0.086
Percent Tied	0.0	Tau-a	0.044
Pairs	70	c	0.543

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	64.0	Somers'D	0.280
Percent Discordant	36.0	Gamma	0.280
Percent Tied	0.0	Tau-a	0.133
Pairs	50	c	0.640

FIG. 45 Continued

99/100

Pancreatic Cancer Cases vs. Controls – MMP2 – 02JUN14

Measurement Variable	Class Variable	n	mean	min	median	max	Rank Sum	T-Test EQ Var.
MMP2	CASE=Control	10	431.9	245	387.0	718	0.63	0.87457
	CASE=Case	10	440.5	326	438.5	571		

variable	z_pcstage	n	mean	min	P_25	median	P_75	max
MMP2	Control	10	431.900	245	339	387.0	561	718
MMP2	Resectable	5	457.000	387	390	423.0	514	571
MMP2	Locally Advanced	2	421.500	326	326	421.5	517	517
MMP2	Metastatic	3	425.667	344	344	454.0	479	479

Kruskal-Wallis Test	
Chi-Square	0.5400
DF	3
Pr > Chi-Square	0.9100

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	56.0	Somers'D	0.140
Percent Discordant	42.0	Gamma	0.143
Percent Tied	2.0	Tau-a	0.074
Pairs	100	c	0.570

FIG. 46

100/100

Note: 3 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	55.7	Somers'D	0.114
Percent Discordant	44.3	Gamma	0.114
Percent Tied	0.0	Tau-a	0.059
Pairs	70	c	0.557

Note: 5 observations were deleted due to missing values for the response or explanatory variables.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	60.0	Somers'D	0.200
Percent Discordant	40.0	Gamma	0.200
Percent Tied	0.0	Tau-a	0.095
Pairs	50	c	0.600

FIG. 46 Continued

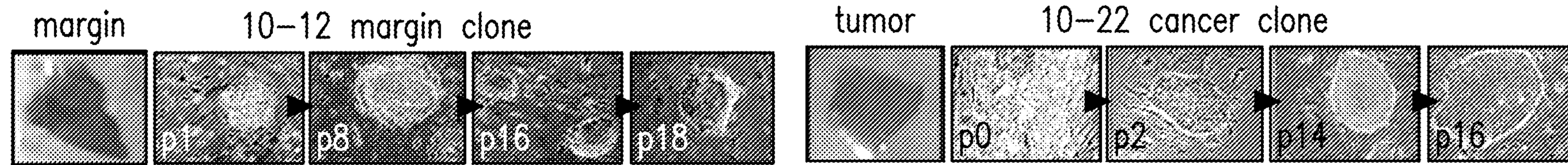


FIG. 1A