

# Distant metastasis occurs late during the genetic evolution of pancreatic cancer

Shinichi Yachida<sup>1\*</sup>, Siân Jones<sup>2\*</sup>, Ivana Bozic<sup>3</sup>, Tibor Antal<sup>3,4</sup>, Rebecca Leary<sup>2</sup>, Baojin Fu<sup>1</sup>, Mihoko Kamiyama<sup>1</sup>, Ralph H. Hruban<sup>1,5</sup>, James R. Eshleman<sup>1</sup>, Martin A. Nowak<sup>3</sup>, Victor E. Velculescu<sup>2</sup>, Kenneth W. Kinzler<sup>2</sup>, Bert Vogelstein<sup>2</sup> & Christine A. Iacobuzio-Donahue<sup>1,5,6</sup>

**Metastasis, the dissemination and growth of neoplastic cells in an organ distinct from that in which they originated<sup>1,2</sup>, is the most common cause of death in cancer patients. This is particularly true for pancreatic cancers, where most patients are diagnosed with metastatic disease and few show a sustained response to chemotherapy or radiation therapy<sup>3</sup>. Whether the dismal prognosis of patients with pancreatic cancer compared to patients with other types of cancer is a result of late diagnosis or early dissemination of disease to distant organs is not known. Here we rely on data generated by sequencing the genomes of seven pancreatic cancer metastases to evaluate the clonal relationships among primary and metastatic cancers. We find that clonal populations that give rise to distant metastases are represented within the primary carcinoma, but these clones are genetically evolved from the original parental, non-metastatic clone. Thus, genetic heterogeneity of metastases reflects that within the primary carcinoma. A quantitative analysis of the timing of the genetic evolution of pancreatic cancer was performed, indicating at least a decade between the occurrence of the initiating mutation and the birth of the parental, non-metastatic founder cell. At least five more years are required for the acquisition of metastatic ability and patients die an average of two years thereafter. These data provide novel insights into the genetic features underlying pancreatic cancer progression and define a broad time window of opportunity for early detection to prevent deaths from metastatic disease.**

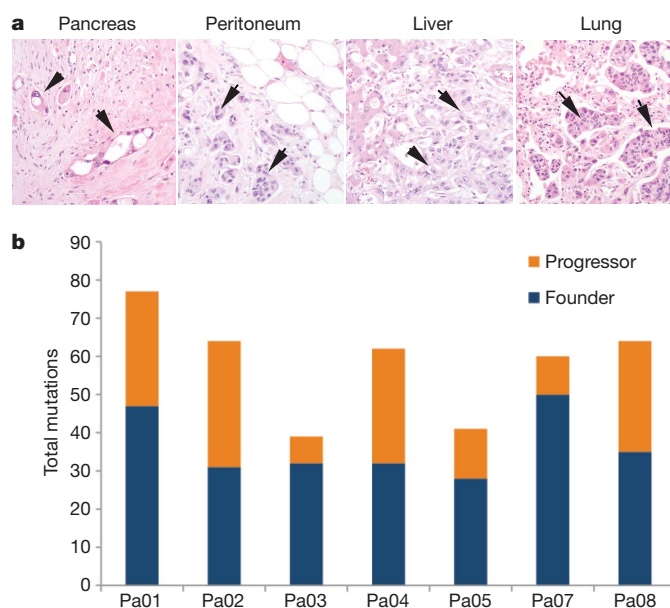
We performed rapid autopsies of seven individuals with end stage pancreatic cancer (Supplementary Table 1). In all patients, metastatic deposits were present within two or more anatomic sites in each patient, most often the liver, lung and peritoneum, as is typical for this form of neoplasia<sup>4</sup>.

Low passage cell lines (six patients) or first passage xenografts (one patient) were created from one of the metastases present at each patient's autopsy. These samples comprised seven of the 24 pancreatic cancers which previously underwent whole exome sequencing and copy number analysis, as described in a mutational survey of the pancreatic cancer genome<sup>5</sup>. In this earlier study, a total of 426 somatic mutations in 388 different genes were identified among 220,884,033 base pairs (bp) sequenced in the seven index metastatic lesions, corresponding to an average of 61 mutations per index metastatic lesion (range 41–77). In all samples, the vast majority of mutations were represented by missense or silent single base substitutions (Supplementary Fig. 1 and Supplementary Table 2).

For each of the somatic mutations identified in the seven index metastasis lesions, we determined whether the same somatic mutation was present in anatomically distinct metastases harvested at autopsy from the same patients. We also determined whether these mutations

were present in the primary pancreatic tumours from which the metastases arose. A small number of these samples of interest were cell lines or xenografts, similar to the index lesions, whereas the majority were fresh-frozen tissues that contained admixed neoplastic, stromal, inflammatory, endothelial and normal epithelial cells (Fig. 1a). Each tissue sample was therefore microdissected to minimize contaminating non-neoplastic elements before purifying DNA.

Two categories of mutations were identified (Fig. 1b). The first and largest category corresponded to those mutations present in all samples from a given patient ('founder' mutations, mean of 64%, range 48–83% of all mutations per patient; Fig. 1b, example in Supplementary Fig. 2a). These data indicate that the majority of somatically acquired mutations present in pancreatic cancers occur before the development of metastatic lesions. All other mutations were characterized as 'progressor' mutations (mean of 36%, range 17–52% of all mutations per patient;



**Figure 1 | Summary of somatic mutations in metastatic pancreatic cancers.** **a**, Histopathology of primary infiltrating pancreatic cancer and metastatic pancreatic cancer to the peritoneum, liver and lung. In addition to infiltrating cancer cells in each lesion (arrows), non-neoplastic cell types are abundant. **b**, Total mutations representing parental clones (founder mutations), and clonal evolution (progressor mutations) within the primary carcinoma based on comparative lesion sequencing. Mutations common to all samples analysed were the most common category identified.

<sup>1</sup>Department of Pathology, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins Medical Institutions, Baltimore, Maryland 21231, USA. <sup>2</sup>The Ludwig Center for Cancer Genetics and Therapeutics and The Howard Hughes Medical Institute at The Johns Hopkins Kimmel Cancer Center, Baltimore, Maryland 21231, USA. <sup>3</sup>Program for Evolutionary Dynamics, Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>4</sup>School of Mathematics, University of Edinburgh, Edinburgh EH9 3JZ, UK. <sup>5</sup>Department of Oncology, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins Medical Institutions, Baltimore, Maryland 21231, USA. <sup>6</sup>Department of Surgery, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins Medical Institutions, Baltimore, Maryland 21231, USA.

\*These authors contributed equally to this work.

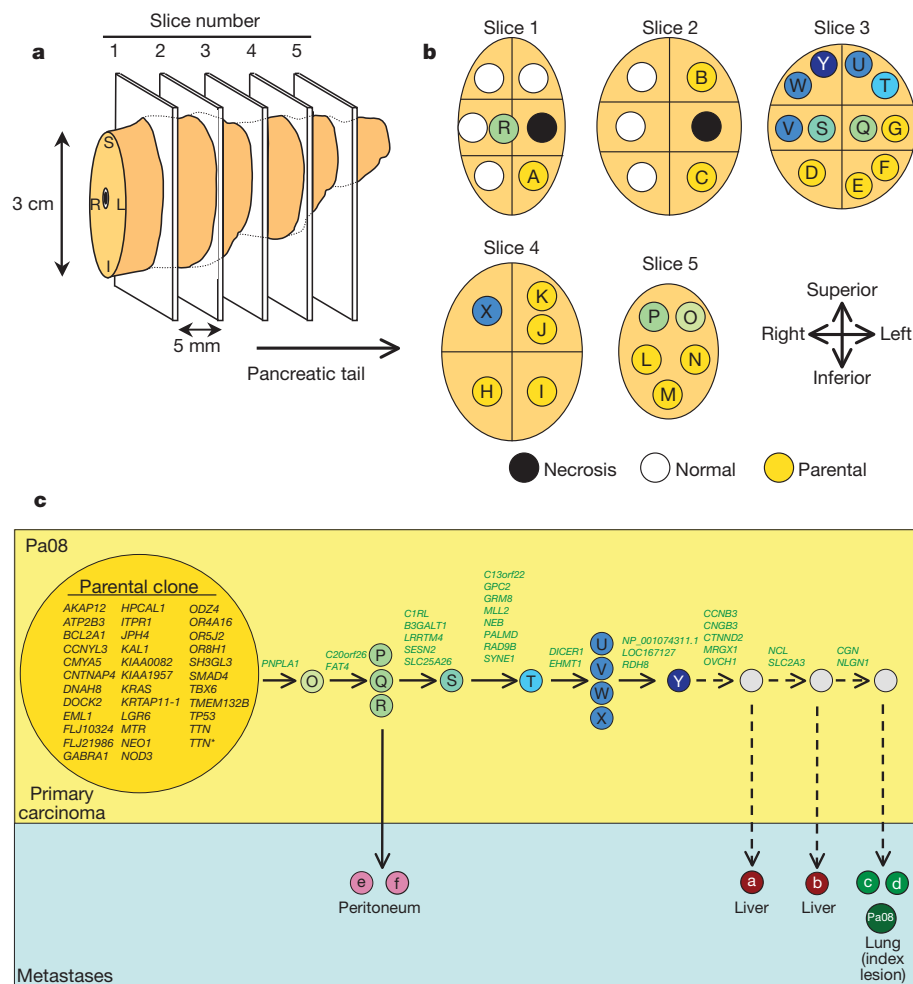
Fig. 1b, example in Supplementary Fig. 2b). These mutations were present in one or more of the metastases examined, including the index metastasis, but not the parental clone.

These mutation types were used to classify the lesions that contained them into parental clones (containing only founder mutations) and subclones (containing both founder and progressor mutations). By definition, there could be only one parental clone in a patient, although there could be many different subclones. Parental clones tended to contain more deleterious mutations (nonsense, splice site or frameshift mutations) than subclones (12.6% of the mutations in the parental clones versus 8.1% of the mutations in subclones, Supplementary Table 2). The parental clones had already accumulated mutations in all driver genes (*KRAS*, *TP53* and *SMAD4*) previously shown to drive pancreatic tumorigenesis<sup>6</sup>. Through combined analysis of high-density single nucleotide polymorphism (SNP) chip data on the index lesion (Supplementary Table 3) plus the sequencing data on all lesions (Supplementary Table 2) we found that the vast majority of homozygous mutations (51 mutations, representing 89% of all homozygous mutations) in the index lesion were already present in the parental clones. Homozygous mutations are characteristic of tumour suppressor genes such as *SMAD4* and *CDKN2A* and often occur in association with chromosomal instability<sup>7</sup>. In sum, the parental clones harboured the majority of deleterious genetic alterations and chromosomal instability,

upon which were superimposed an accumulation of progressor mutations associated with clonal evolution and metastasis.

Evolutionary maps were constructed for each patient's carcinoma based on the patterns of somatic mutation and allelic losses and the locations of individual metastatic deposits (Fig. 2 and Supplementary Figs 3–8). These maps showed that, despite the presence of numerous founder mutations within the parental clones, the cells giving rise to the metastatic lesions had a large number of progressor mutations. For example, in Pa01 the parental clone contained 49 founder mutations, yet a clonal expansion marked by the presence of mutations in six additional genes was present in the lung and peritoneal metastases (Supplementary Fig. 3). Moreover, 22 more mutations were found in the liver metastasis. Note that all mutations in the metastatic lesions were clonal, that is, present in the great majority if not all neoplastic cells of the metastasis, as assessed by Sanger sequencing. Thus, these mutations were present in the cell that clonally expanded to become the metastasis. Similarly, large numbers of progressor mutations were generally observed in the metastases from each of the seven cases examined (Fig. 2 and Supplementary Figs 3–8).

To distinguish between the possibilities that clonal evolution occurred inside the primary cancer versus within secondary sites, we sectioned the primary tumours from two patients into numerous, three-dimensionally organized pieces (Fig. 2a, b) and examined the DNA



**Figure 2 | Geographic mapping of metastatic clones within the primary carcinoma and proposed clonal evolution of Pa08.** **a**, Illustration of the pancreatic specimen removed from Pa08 at rapid autopsy, and the planes of sectioning of the specimen. **b**, Mapping of the parental clone and subclones identified by comparative lesion sequencing within serial sections of the infiltrating pancreatic carcinoma. Metastatic subclones giving rise to liver and lung metastases are non-randomly located within slice 3, indicated by blue

circles. These clones are both geographically and genetically distinct from clones giving rise to peritoneal metastases in this same patient, indicated in green. **c**, Proposed clonal evolution based on the sequencing data. In this model, after development of the parental clone, ongoing clonal evolution continues within the primary carcinoma (yellow rectangle), and these subclones seed metastases in distant sites. \*Two mutations were found in the *TTN* gene.

from each piece for each of the founder and progressor mutations. In Patient Pa08, there were three progressor mutations present in two independent peritoneal metastases (defining one subclone) and 23, 25 or 27 additional progressor mutations present in liver and lung metastases (defining three additional subclones; Fig. 2c). Through the analysis of distinct regions of the primary tumour, it was clear that subclones giving rise to each of these metastases were present in the primary tumour. Moreover, these subclones were not small; from the size of the pieces (Fig. 2a) and the amounts of DNA recovered, each subclone must have contained in excess of 100 million cells. In addition, more than four different subclones, each containing a similarly large number of cells, could be identified through the analysis of other pieces of the same tumour. These subclones could be put into an ordered hierarchy establishing an evolutionary path for tumour progression (Fig. 2c). Analysis of multiple primary tumour pieces and metastatic lesions from patient Pa04 revealed a similar clonal evolution, with distinct, large subclones within the primary tumours giving rise to the various metastases (Supplementary Fig. 8).

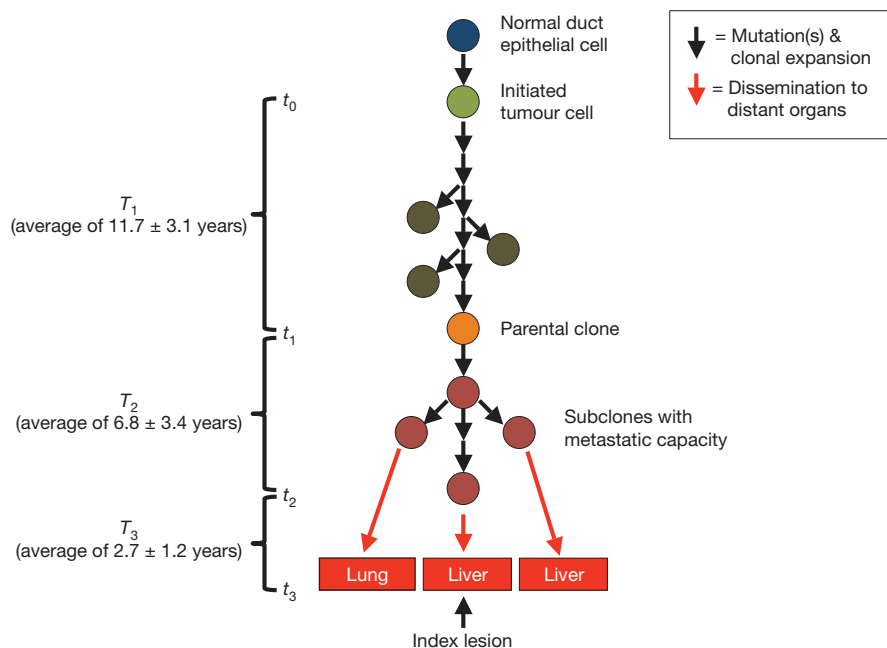
To clarify further clonal evolution within the primary site, we attempted to correlate the mutation signatures representing the subclones of Pa08 (Fig. 2c) with the geographic location of the pieces of the primary tumour used to define them (Fig. 2a, b). Samples representative of the parental clone were located throughout the primary carcinoma. By contrast, samples representing subclones were non-randomly located in proximity to each other, within which the subclones specifically giving rise to peritoneal versus distant metastases were seen. Thus, we conclude that the genetic heterogeneity of metastases reflects heterogeneity already existing within the primary carcinoma, and that the primary carcinoma is a mixture of numerous subclones, each of which has independently expanded to constitute a large number of cells.

This data set could also be used to infer the timing of the development of the various stages of pancreatic tumour progression<sup>8</sup>. We assume that the tumour is initiated by a genetic event that confers a selective growth advantage to the cell that goes on to become the

founder cell of the tumour. To estimate the timing, we first used Ki-67 labelling to determine the proliferation rate of seven samples of normal duct epithelium from surgically resected pancreata of individuals without pancreatic cancer as well as of each index metastasis. Ki-67-positive nuclei constituted an average of 0.4% of normal ductal cells, whereas an average of 16.3% of cancer cells within the index metastasis lesions were Ki-67-positive, consistent with prior estimates<sup>9,10</sup> (Supplementary Table 4). Based on these data plus that from sequencing of the index lesions, we derived estimates for three critical times in tumour evolution:  $T_1$ , the time between tumour initiation and the birth of the cell giving rise to the parental clone;  $T_2$ , the subsequent time required for the birth of the cell that gave rise to the index metastasis; and  $T_3$ , the time between the dissemination of this cell and the patients' death (Fig. 3). In other words, there is a time point,  $t_0$ , when the tumour was initiated, and a time point  $t_1$  when a cell is born that has all mutations that exist in the parental clone. Similarly, there is a time point in tumour evolution,  $t_2$ , when a cell is born that has all the mutations that exist in the index metastasis.  $T_1$  is given by  $t_1 - t_0$  and  $T_2$  is given by  $t_2 - t_1$ . If we denote  $t_3$  as the time of patient's death, then  $T_3 = t_3 - t_2$ .

Using the mathematical model described in the Methods, we were able to conservatively estimate an average of 11.7 years from the initiation of tumorigenesis until the birth of the cell giving rise to the parental clone, an average of 6.8 years from then until the birth of the cell giving rise to the index lesion, and an average of 2.7 years from then until the patients' death (see Supplementary Discussion and Supplementary Table 5).

We show, for the first time, that primary pancreatic cancers contain a mix of geographically distinct subclones, each containing large numbers (hundreds of millions) of cells that are present within the primary tumour years before the metastases become clinically evident. The features of these metastatic subclones that promote metastasis formation have yet to be discerned, because no consistent genetic signature of metastatic subclones could be identified. We did identify several genes



**Figure 3 | Schema of the genetic evolution of pancreatic cancer.**

Tumorigenesis begins with an initiating mutation in a normal cell that confers a selective growth advantage. Successive waves of clonal expansion occur in association with the acquisition of additional mutations, corresponding to the progression model of pancreatic intraepithelial neoplasia (PanIN) and time  $T_1$ . One founder cell within a PanIN lesion will seed the parental clone and hence initiate an infiltrating carcinoma (end of  $T_1$  and beginning of  $T_2$ ). Eventually,

the cell that will give rise to the index lesion will appear (end of  $T_2$  and beginning of  $T_3$ ). Unfortunately, most patients are not diagnosed until well into time interval  $T_3$  when cells of these metastatic subclones have already escaped the pancreas and started to grow within distant organs. The average time for intervals  $T_1$ ,  $T_2$  and  $T_3$  for all seven patients is indicated in the parentheses at left (see also Supplementary Table 6).

that were mutated in one or more of the index metastatic lesions from these seven patients with Stage IV disease, but not in the primary pancreatic index lesions from 17 patients with Stage II disease (Supplementary Table 2). These genes include those that may have a role in invasive or metastatic ability through heterotypic cell adhesion (*CNTN5*), motility (*DOCK2*), proteolysis (*MEP1A*) and tyrosine phosphorylation (*LMTK2*). However, these mutations were not metastasis-specific *per se* as all but one were present in the matched primary carcinoma of those same seven patients, and there is no evidence that the mutations we observed endowed these genes with metastatic activity. These data also do not reveal the selective pressures within the primary carcinoma that led to the formation of progressor mutations. In light of recent findings indicating that pancreatic cancers are poorly vascularized<sup>11</sup>, one possibility is that intratumoural hypoxia creates a fertile microenvironment for the formation of additional mutations beyond that of the parental clone.

One of the major implications of these data is their implication for screening to prevent pancreatic cancer deaths. Quantitative analysis indicated a large window of opportunity for diagnosis while the disease was still in the curative stage—at least a decade. Our model also predicts an average of 6.8 years between the birth of the cell giving rise to the parental clone and the seeding of the index metastasis. Unfortunately, the great majority of patients are not diagnosed until the last 2 years of the entire tumorigenic process. The challenge is to detect these tumours during time  $T_1$ , or even after  $T_1$  but before seeding of metastases. Advanced imaging methods, as well as blood tests to detect cancer-specific proteins, transcripts or genes<sup>12</sup>, offer hope for such non-invasive early detection.

## METHODS SUMMARY

Rapid autopsies were performed on seven individuals with Stage IV pancreatic cancer<sup>13</sup>. Genomic DNA was extracted from cell lines or xenografts established from one metastasis of each patient and used for exomic sequencing as described previously<sup>5</sup>. The Illumina Infinium II Whole Genome Genotyping Assay using the BeadChip platform was also used to analyse each sample at 1,072,820 (1M) SNP loci as described previously<sup>5</sup>. Samples of snap-frozen pancreatic cancer tissue were microdissected using a PALM MicroLaser System (Carl Zeiss MicroImaging) and DNA extracted using QIAamp DNA Micro Kits (Qiagen). Genomic DNA was quantified by calculating long interspersed nuclear elements (LINE) by real-time PCR. Whole genome amplification (WGA) was performed using 10 ng total template DNA and an illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare). Ki-67 immunolabelling (Clone MIB-1, Dako Cytomation) was performed on formalin-fixed, paraffin-embedded sections of normal pancreatic ducts and metastatic pancreatic cancer tissues for each patient using the Ventana Discovery staining system (Ventana Medical Systems), and this information was used to inform computational models of the timing of clonal evolution of each patient's pancreatic cancer (full details of these models are available in Full Methods).

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 11 May; accepted 15 September 2010.

1. Fidler, I. J. Critical determinants of metastasis. *Semin. Cancer Biol.* **12**, 89–96 (2002).
2. Nguyen, D. X., Bos, P. D. & Massague, J. Metastasis: from dissemination to organ-specific colonization. *Nature Rev. Cancer* **9**, 274–284 (2009).

3. Stathis, A. & Moore, M. J. Advanced pancreatic carcinoma: current treatment and future challenges. *Nature Rev. Clin. Oncol.* **7**, 163–172 (2010).
4. Yachida, S. & Iacobuzio-Donahue, C. A. The pathology and genetics of metastatic pancreatic cancer. *Arch. Pathol. Lab. Med.* **133**, 413–422 (2009).
5. Jones, S. *et al.* Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* **321**, 1801–1806 (2008).
6. Maitra, A. & Hruban, R. H. pancreatic cancer. *Annu. Rev. Pathol.* **3**, 157–188 (2008).
7. Lengauer, C., Kinzler, K. W. & Vogelstein, B. Genetic instabilities in human cancers. *Nature* **396**, 643–649 (1998).
8. Jones, S. *et al.* Comparative lesion sequencing provides insights into tumor evolution. *Proc. Natl Acad. Sci. USA* **105**, 4283–4288 (2008).
9. Terada, T. *et al.* Cell proliferative activity in intraductal papillary-mucinous neoplasms and invasive ductal adenocarcinomas of the pancreas: an immunohistochemical study. *Arch. Pathol. Lab. Med.* **122**, 42–46 (1998).
10. Hisa, T. *et al.* Growth process of small pancreatic carcinoma: a case report with imaging observation for 22 months. *World J. Gastroenterol.* **14**, 1958–1960 (2008).
11. Olive, K. P. *et al.* Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science* **324**, 1457–1461 (2009).
12. Sidransky, D. Emerging molecular markers of cancer. *Nature Rev. Cancer* **2**, 210–219 (2002).
13. Embuscado, E. E. *et al.* Immortalizing the complexity of cancer metastasis: genetic features of lethal metastatic pancreatic cancer obtained from rapid autopsy. *Cancer Biol. Ther.* **4**, 548–554 (2005).
14. Fu, B. *et al.* Evaluation of GATA-4 and GATA-5 methylation profiles in human pancreatic cancers indicate promoter methylation patterns distinct from other human tumor types. *Cancer Biol. Ther.* **6**, 1546–1552 (2007).
15. Peiffer, D. A. *et al.* High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res.* **16**, 1136–1148 (2006).
16. Bignell, G. R. *et al.* Signatures of mutation and selection in the cancer genome. *Nature* **463**, 893–898 (2010).
17. Sasaki, K. *et al.* Relationship between labeling indices of Ki-67 and brdurd in human malignant tumors. *Cancer* **62**, 989–993 (1988).
18. Rew, D. A. & Wilson, G. D. Cell production rates in human tissues and tumours and their significance. Part II: clinical data. *Eur. J. Surg. Oncol.* **26**, 405–417 (2000).
19. Steel, G. G. *The Growth Kinetics of Tumours*. (Clarendon Press, Oxford, 1977).
20. Amikura, K., Kobari, M. & Matsuno, S. The time of occurrence of liver metastasis in carcinoma of the pancreas. *Int. J. Pancreatol.* **17**, 139–146 (1995).
21. Naumov, G. N. *et al.* A model of human tumor dormancy: an angiogenic switch from the nonangiogenic phenotype. *J. Natl. Cancer Inst.* **98**, 316–325 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by National Institutes of Health grants CA106610 (C.A.I.-D.), CA62924 (C.A.I.-D., M.A.N.), CA43460 (B.V.), CA57345 (K.W.K. and V.E.V.), CA121113 (V.E.V. and K.W.K.), GM078986 (M.A.N.), the Bill and Melinda Gates Foundation Grand Challenges Grant 37874, the Uehara Memorial Foundation (S.Y.), the AACR-Barletta Foundation (C.A.I.-D.), the John Templeton Foundation, Pilot Funding from the Sol Goldman Pancreatic Cancer Research Center, the Michael Rolfe Pancreatic Cancer Foundation, the George Rubis Endowment for Pancreatic Cancer Research, the Joseph C. Monastra Foundation for Pancreatic Cancer Research, the Alfredo Scatena Memorial Fund, the Virginia and the D.K. Ludwig Fund for Cancer Research, The Joint Program in Mathematical Biology and J. Epstein. We would like to acknowledge T. C. Cornish, C. Henderson, N. Omura and S.-M. Hong for their technical assistance in selected aspects of this work.

**Author Contributions** Sample collection and processing was performed by C.A.I.-D., S.Y., B.F. and M.K. Microdissection, DNA extractions and whole genome amplification reactions were performed by S.Y. Sequencing was performed by S.J. Copy number analyses were performed by R.L. Computational models and estimates of clonal evolution were performed by I.B., T.A. and M.A.N.; C.A.I.-D., S.Y., S.J., R.H.H., J.R.E., M.A.N., I.B., T.A., V.E.V., K.W.K. and B.V. directed the research. C.A.I.-D., B.V., S.Y., I.B. and T.A. wrote the manuscript, which all authors have approved.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to C.A.I.-D. ([ciacobu@jhmi.edu](mailto:ciacobu@jhmi.edu)).

## METHODS

**Patients and tissue samples.** Tissue samples from seven patients with pancreatic ductal adenocarcinoma were collected in association with the Gastrointestinal Cancer Rapid Medical Donation Program (GICRMDP). This programme was approved by the Johns Hopkins institutional review board and deemed in accordance with the Health Insurance Portability and Accountability Act. Details of the programme have been described in detail previously<sup>13</sup>. The tissue harvesting protocol consists of the following; after opening of the body cavity using standard techniques, the whole pancreas including the pancreatic cancer and each grossly identified metastasis were sampled using a sterile blade and forceps. The whole pancreas was sliced into  $1 \times 1 \times 0.4$  cm sections for overnight fixation in 10% buffered-formalin, for freezing in Tissue-Tek OCT compound (Sakura Finetech) in liquid nitrogen and for snap-freezing in liquid nitrogen in 1.7 ml cryovials and storage at  $-80^\circ\text{C}$ . Xenograft enriched or low passage cell lines were generated from the post mortem cancer tissues of these seven patients as described previously<sup>13,14</sup>.

**Laser capture microdissection (LCM).** Frozen tissue sections of autopsy tissues were cut into  $7\ \mu\text{m}$  sections using a cryostat and embedded onto UV-treated PALM membrane slides (Carl Zeiss MicroImaging) and the slides were stored immediately at  $-80^\circ\text{C}$  until subsequent fixation. Tissue sections that underwent LCM were defrosted, fixed in 100% methanol for 3 min, and stained with toluidine blue before microdissection to remove contaminating stromal elements. Sections were dissected using a PALM MicroLaser System (Carl Zeiss MicroImaging). Dissected tissues were catapulted into adhesive caps. Generally,  $>20,000$  cells were obtained from 5–10 serial sections by LCM to obtain sufficient quantity and quality of genomic DNA for subsequent amplification and sequencing.

**Genomic DNA extraction and whole genome amplification.** Genomic DNA from microdissected tissues was extracted using a QIAamp DNA Micro Kit (Qiagen) according to the manufacturer's protocol. Genomic DNA was quantified by calculating long interspersed nuclear elements (LINE) by real-time PCR. The LINE primer set 5'-AAAGCCGCTCAACTACATGG-3' (forward) and 5'-TGCTTTGAATGCGTCCAGAG-3' (reverse) was designed. The real-time PCR conditions were  $95^\circ\text{C}$  for 10 min; 40 cycles of  $94^\circ\text{C}$  for 10 s,  $58^\circ\text{C}$  for 15 s and  $70^\circ\text{C}$  for 30 s. PCR was carried out using Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen). To minimize sequencing bias from using low-copy starting templates, only samples for which the measured concentration by LINE assay was  $\geq 3.3\ \text{ng}\ \mu\text{l}^{-1}$  (1,000 genome equivalents) were used as a starting template for whole genome amplification (WGA). WGA was performed using 10 ng total template DNA and an illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare), following the manufacturer's protocol. WGA products were purified using a Microspin G-50 system (GE Healthcare). The purified WGA products were quantified by NanoDrop spectrophotometer (Thermo Fisher Scientific) and diluted to  $20\ \text{ng}\ \mu\text{l}^{-1}$  for sequencing analysis. Using these methods and quality controls, there was complete concordance in the mutational signatures obtained of cultured cell lines/xenografts versus WGA materials prepared from their matched frozen tissues.

**Sanger sequencing.** PCR amplification and sequencing was performed using the conditions and primers described previously<sup>5</sup>. A small number of sequencing reactions failed ( $<2\%$  of the total reactions) and these corresponding genes were not included in progression models or quantitative time estimates of clonal evolution.

**Genotyping.** The Illumina Infinium II Whole Genome Genotyping Assay using the BeadChip platform was used to analyse tumour samples at 1,072,820 (1M) SNP loci as previously described<sup>5</sup>. Briefly, all SNP positions were based on the hg18 (NCBI Build 36, March 2006) version of the human genome reference sequence. The genotyping assay begins with hybridization to a 50-nucleotide oligonucleotide, followed by a two-colour fluorescent single-base extension. Fluorescence intensity image files were processed using Illumina BeadStation software to provide normalized intensity values and allelic frequency for each SNP position. For each SNP, the normalized experimental intensity value ( $R$ ) was compared to the intensity values for that SNP from a training set of normal samples and represented as a ratio (called the 'log  $R$  ratio') of  $\log_2(R_{\text{experimental}}/R_{\text{training set}})$ . For each SNP, the normalized allele intensity ratio (theta) was used to estimate a quantitative allelic frequency value (called the 'B allele frequency') for that SNP<sup>15</sup>. Using Illumina BeadStudio software, log  $R$  ratio and B allele frequency values were plotted along chromosomal coordinates and examined visually. Regions of loss of heterozygosity (LOH) were identified as genomic regions  $>2$  megabases (Mb) with consecutive homozygous genotype calls (B allele frequency near 0 or 1). Smaller ( $<2$  Mb) regions of LOH were identified by requiring co-occurrence of decreased log  $R$  ratio scores in regions of consecutive homozygous genotype calls (B allele frequency near 0 or 1). Visual analysis of these data plotted along chromosomal coordinates was followed by manual analysis of the data for selected genes of interest.

**Estimations of proliferation rates.** To estimate the cell division rate, the Ki-67 labelling index (LI) in the proband lesion for each case was calculated. The Ki-67 LI on the pancreatic ducts in the histologically normal pancreas parenchyma was also

calculated. Normal pancreas was collected from two autopsied patients who died of causes other than pancreatic cancer and five patients who underwent distal pancreatectomy for a serous cystadenoma or an islet cell tumour at The Johns Hopkins Hospital. Paraffin blocks were cut into sections  $4\text{-}\mu\text{m}$  thick for Ki-67 immunostaining with all staining processes from deparaffinization to counterstaining with haematoxylin being performed automatically with the Ventana Discovery staining system (Ventana Medical Systems). An anti-human Ki-67 mouse monoclonal antibody (Clone MIB-1, Dako Cytomation) was used. At least 12 randomly selected high-power fields containing a minimum of 2,000 cells were evaluated for each case, and the labelling index (LI) was calculated as the percentage of positive cell nuclei. Reactive small lymphocytes in each case were regarded as internal positive controls for Ki-67. Equal or more intense nuclear staining in comparison with the internal positive controls was considered to indicate positivity.

**Modelling tumour evolution.** Passenger mutations were defined as those unlikely to drive tumorigenesis. To be conservative, we considered passenger mutations as those not included as candidate cancer genes in a recent study based on whole exome sequencing of 24 pancreatic cancers<sup>5</sup>. As the great majority of mutations identified in cancers are believed to be passengers, the results of the model are not highly dependent on the model used to estimate the relatively small number of drivers<sup>16</sup>.

Because passenger mutations are neutral and do not affect the evolution in any way, they are accumulated independently in each cell lineage. Following the lineage of the founder cell of the parental clone back in time, we can assume that it acquired a new neutral mutation with rate  $r$  at each cell division, with  $r$  being the product of the mutation rate per base pair per cell division and the number of base pairs sequenced. The accumulation of neutral mutations in a cell lineage can be well-described by a Poisson process with rate  $r$  per cell division. We are interested in the number of cell divisions in the single lineage between tumour initiation and birth of the founder cell of the parental clone during which  $N_1$  passenger mutations accumulate. On the other hand,  $N_1$  is also the number of mutations that are found in all tumour samples from one patient. Since we sequenced at least one sample from the primary tumour and at least three samples from different metastases from each patient, these specific  $N_1$  mutations had to be present in the founder cells of all three metastases and in cells in the primary tumour. Thus there was a cell in the tumour that had these  $N_1$  specific mutations for the first time, and that is, by definition, the founder cell of the parental clone. Since we can neglect the accumulation of mutations before the onset of the tumour, these  $N_1$  mutations are accumulated along the single lineage from the tumour initiator cell to the founder cell of the parental clone. As the number of cell divisions between two subsequent mutations is distributed according to an exponential distribution with mean  $1/r$ , the required number of cell divisions is the sum of  $N_1$  independent exponentially distributed random variables with mean  $1/r$ , and is distributed according to a Gamma distribution with shape parameter  $N_1$  and scale parameter  $1/r$ . The mean of this distribution is  $N_1/r$  and the standard deviation is  $\sqrt{N_1}/r$  (see Supplementary Table 6). Because the number of base pairs sequenced in the study is  $31.7 \times 10^6$ , and the mutation rate per base pair per generation is estimated at  $5 \times 10^{-10}$ ,  $r = 31.7 \times 10^6 \times 5 \times 10^{-10} \approx 0.016$  per generation<sup>8</sup>.

Using our measurements of Ki-67 labelling index of the seven index lesions (average 16.3%), we were able to estimate the S-phase fraction of cells in the seven index lesions (average LI = 9.5%)<sup>17</sup>. Assuming a median value for the S-phase duration in human tissues and tumours,  $T_s$ , of 10 h (ref. 18) and using the formula for the potential cell doubling time  $T_{\text{pot}} = \lambda T_s / \text{LI}$ , we get an estimate for  $T_{\text{pot}}$  of 3.5 days. Here  $\lambda$  is a correction factor for the nonlinear age distribution of cells through the cell cycle, which was assumed to be 0.8 (ref. 19). This estimate is consistent with the average cell doubling time in pancreatic cancer from ref. 20 of 2.3 days. We use this latter estimate in our analysis, as we believe it is more accurate for pancreatic cancer.

Our model works very well for estimating the number of cell divisions between discrete events in tumour evolution. In order to go from number of cell divisions to actual time we need to have an estimate for the average rate of cell division. The accuracy of our predictions regarding actual time therefore depends on the accuracy of that estimate. If we let  $T_{\text{gen}}$  denote the average time between subsequent cell divisions in a cell lineage, we arrive at the expression for time  $T_1$ :

$$T_1 = \frac{T_{\text{gen}}}{r} (N_1 \pm \sqrt{N_1}).$$

We therefore estimate the number of cell divisions, and hence the time  $T_1$  between tumour initiation and birth of the founder cell of the parental clone, to be proportional to the number of passenger mutations,  $N_1$ , that the tumour acquired during that time. In our calculations, we use the estimate for cell doubling time in pancreatic cancer from the literature<sup>20</sup> as the value of  $T_{\text{gen}}$ .

$T_2$  is determined analogously, with  $N_2$  defined as the number of passenger mutations present in the index lesion but not in the parental clone.  $T_3$  is

determined from literature-based estimates of the tumour and cell doubling times, and the size of the index lesions at autopsy<sup>20</sup>.

The median doubling time of pancreatic cancer metastases was reported as 56 days<sup>20</sup>. To estimate the age of the index metastasis, we used a two stage model.

We estimated the tumour doubling time was equal to the cell doubling time ( $T_{\text{gen}}$ ) until the tumour size reached 100  $\mu\text{m}$  in diameter at which time angiogenesis is required<sup>21</sup>. Thereafter, we used the median doubling time described above.

### Supplemental Discussion

Our estimate for  $T_1$  is based on the assumption that no passenger mutations were present prior to tumor initiation, as pancreatic ductal epithelium is not a proliferative tissue type and divides only rarely<sup>1,2</sup>. We further validated this assumption by evaluating whether mutations found in Pa08 were present in the matched precursor pancreatic intraepithelial neoplasia or acinar-ductal metaplastic lesions present within the pancreas of this same patient. Apart from the *KRAS* mutation, none of the other mutations in the infiltrating carcinoma or metastasis were identified in these precursor lesions, implying that all passenger mutations had accumulated after tumor initiation.

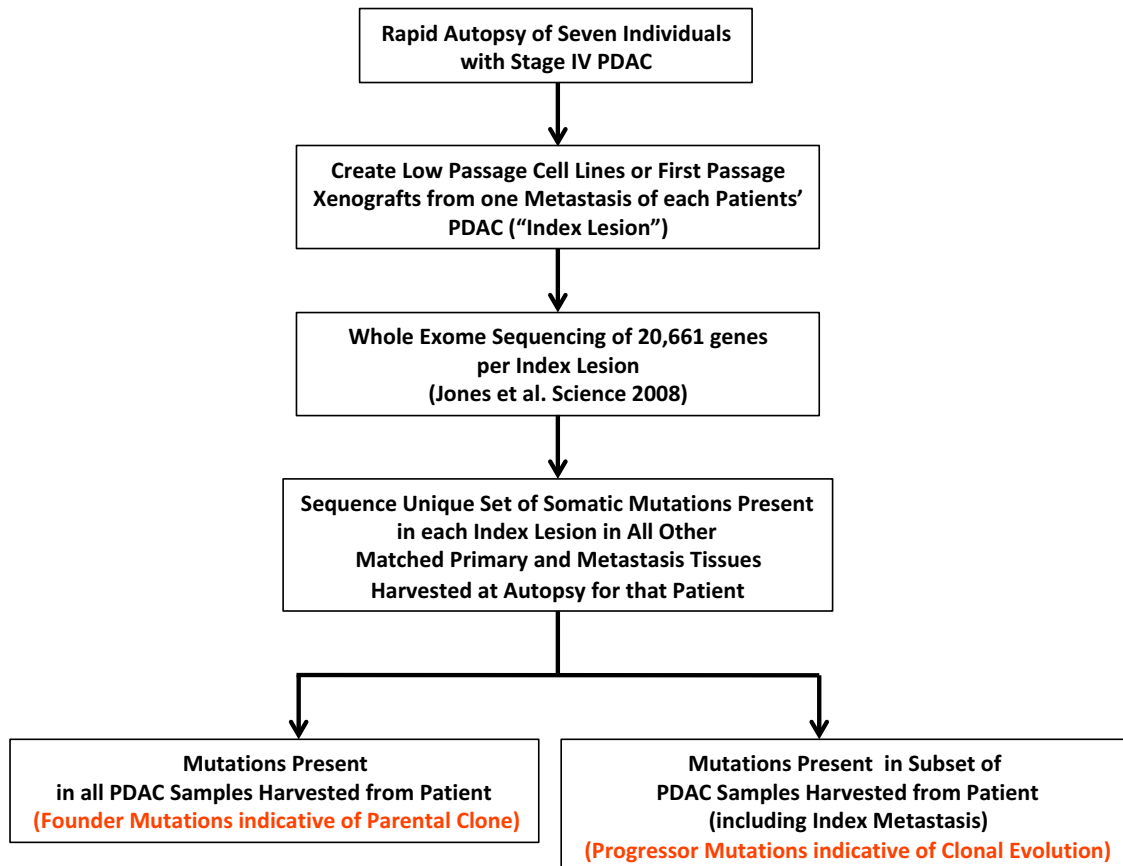
Previously, Meza et al.<sup>3</sup> reported the average sojourn time of premalignant pancreatic lesions to be 50-60 years based on the analysis of pancreatic cancer incidence data, which is significantly longer than the times we are reporting. We believe that the difference between the times that we report and those from Meza et al. is due to the fact that our times are based on data from patients that developed detectable pancreatic cancers, whereas Meza et al. reported average sojourn times for all premalignant pancreatic lesions, including those that never become malignant during an individual's lifetime.

We note that in estimating times  $T_1$  and  $T_2$  we made the assumption that the average cell doubling times are roughly the same in the two stages (tumor initiation to parental clone and parental clone to metastasis). We believe that the value we assigned to the average cell doubling time during the second stage (parental clone to metastasis) is accurate, as this has been measured in pancreatic cancers rather than in early stage

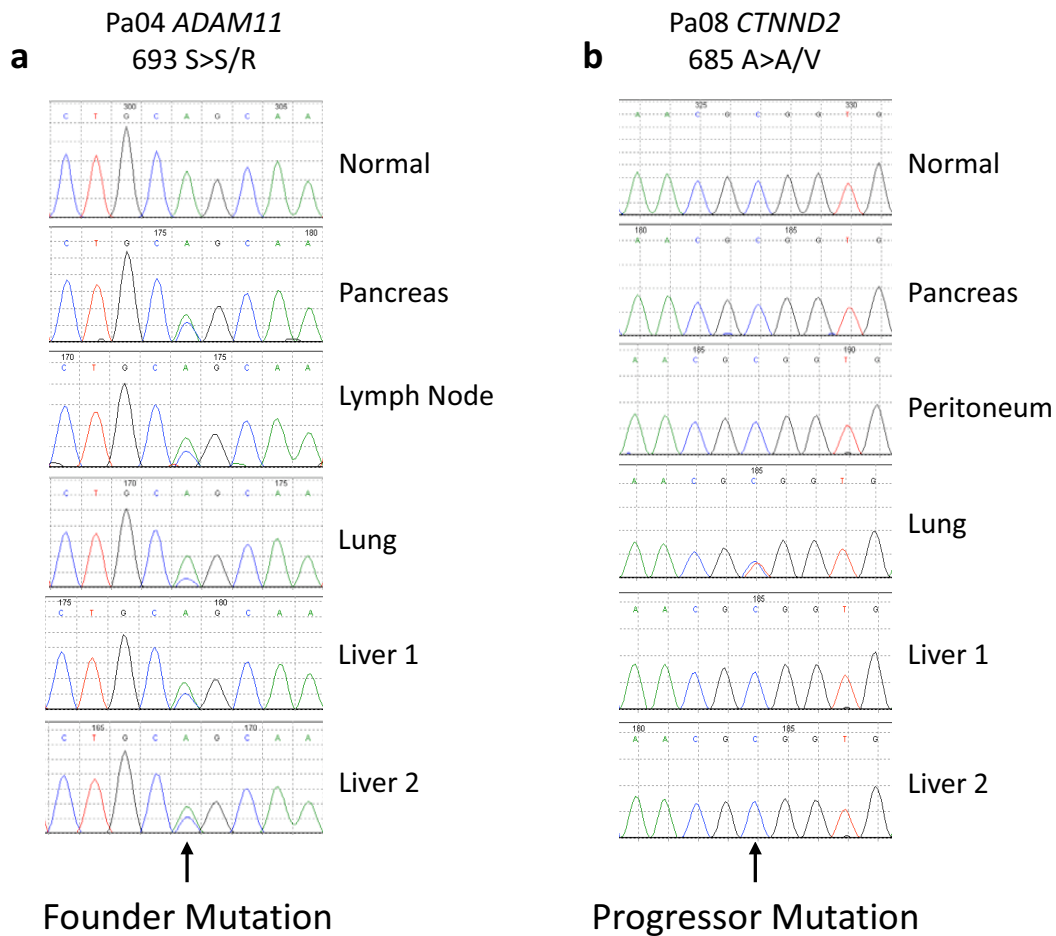
tumors. The actual cell doubling time during the first stage (tumor initiation to parental clone) may be somewhat longer than the value we assumed. The actual cell doubling time may also be somewhat longer than our estimate if extinct lineages have shorter cell cycle times. However, a larger value would make an even stronger case for the long window of opportunity for early detection of pancreatic cancer that we identified, and therefore our conclusion on this point is conservative. It is also possible that the rate of cell division is different in different patients, depending on genomic instability, treatment and other factors. For this reason we note that the averages for  $T_1$ ,  $T_2$  and  $T_3$  are more reliable than the corresponding times we report for individual patients.

#### REFERENCES

- <sup>1</sup> W. M. Klein, R. H. Hruban, A. J. Klein-Szanto et al., Direct Correlation between Proliferative Activity and Dysplasia in Pancreatic Intraepithelial Neoplasia (Panin): Additional Evidence for a Recently Proposed Model of Progression, *Mod Pathol* **15** (4), 441 (2002).
- <sup>2</sup> T. Terada, T. Ohta, Y. Kitamura et al., Cell Proliferative Activity in Intraductal Papillary-Mucinous Neoplasms and Invasive Ductal Adenocarcinomas of the Pancreas: An Immunohistochemical Study, *Archives of Pathology & Laboratory Medicine* **122** (1), 42 (1998).
- <sup>3</sup> R. Meza, J. Jeon, S. H. Moolgavkar et al., Age-Specific Incidence of Cancer: Phases, Transitions, and Biological Implications, *Proceedings of the National Academy of Sciences of the United States of America* **105** (42), 16284 (2008).

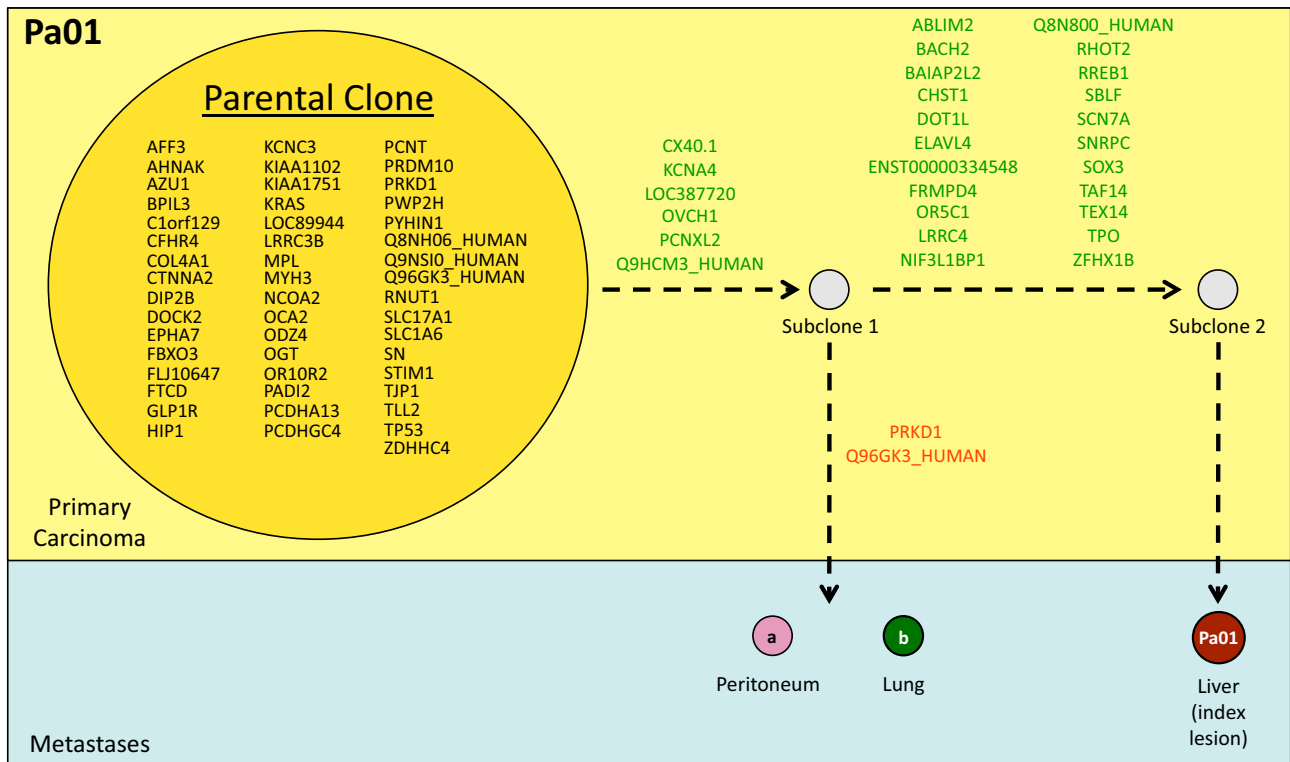


Supplementary Fig. 1. Schematic of methodology and analysis of data in the current study.

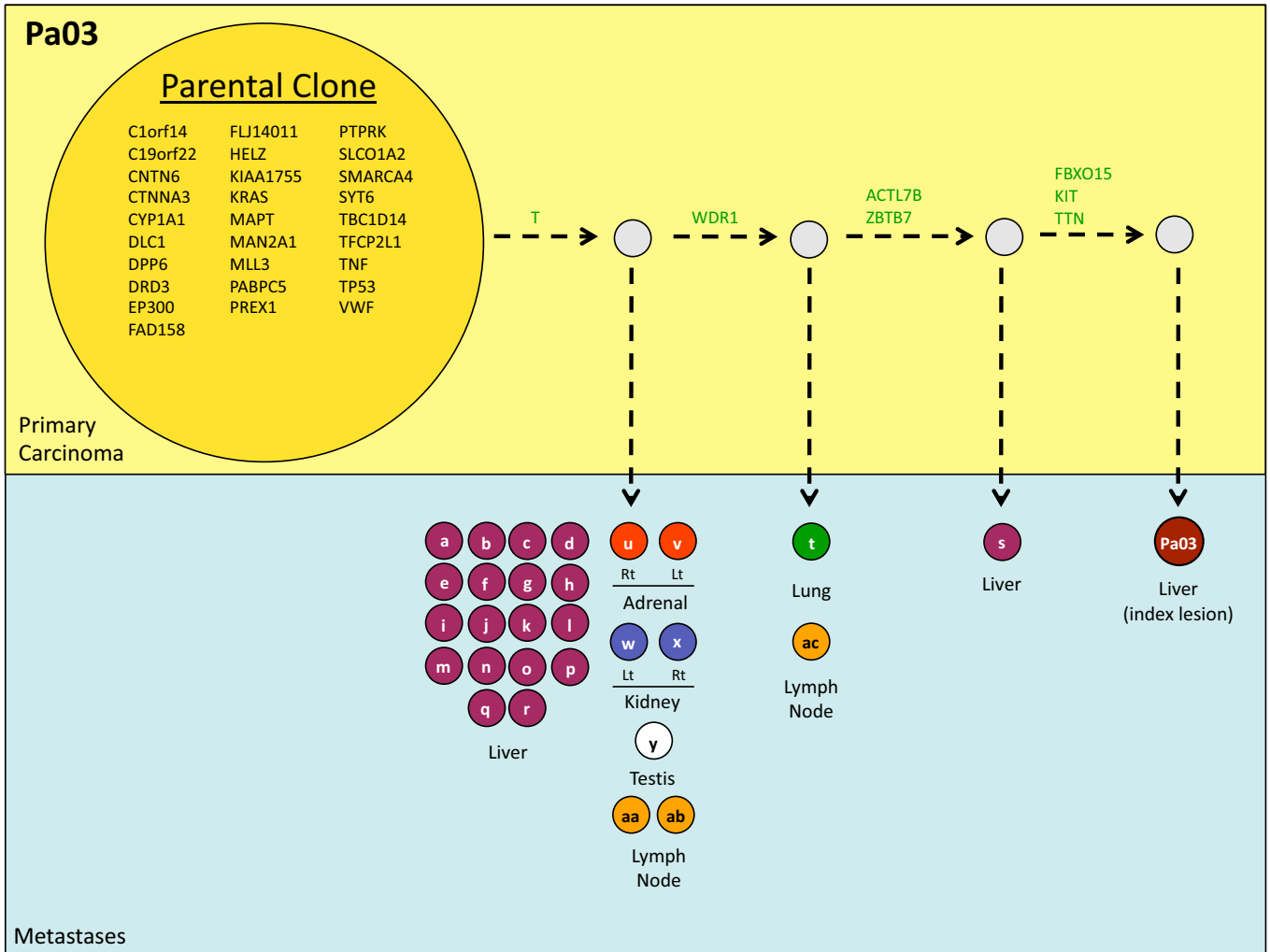


**Supplementary Fig. 2. Representative sequencing chromatograms of somatic mutations in Pa04 and Pa08.** a. *ADAM11* mutations were present in the primary carcinoma and all metastasis samples, indicating it arose prior to the development of metastasis. DNA from a normal tissue of the same patient (representing the germ-line) was negative. b. Mutations in *CTNND2* were present in one of the metastasis samples only (lung).

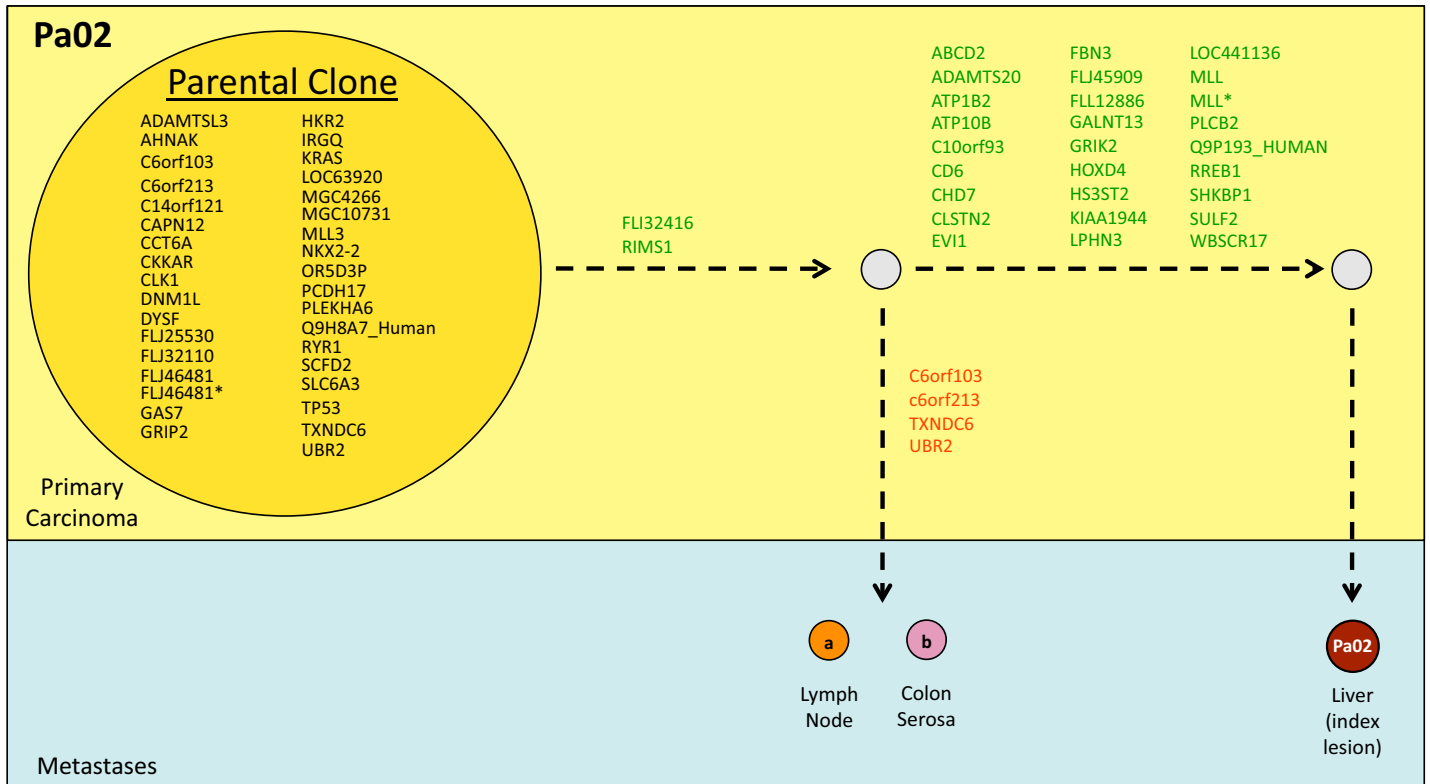
Supplementary Fig. 3. Proposed clonal evolution of Pa01.



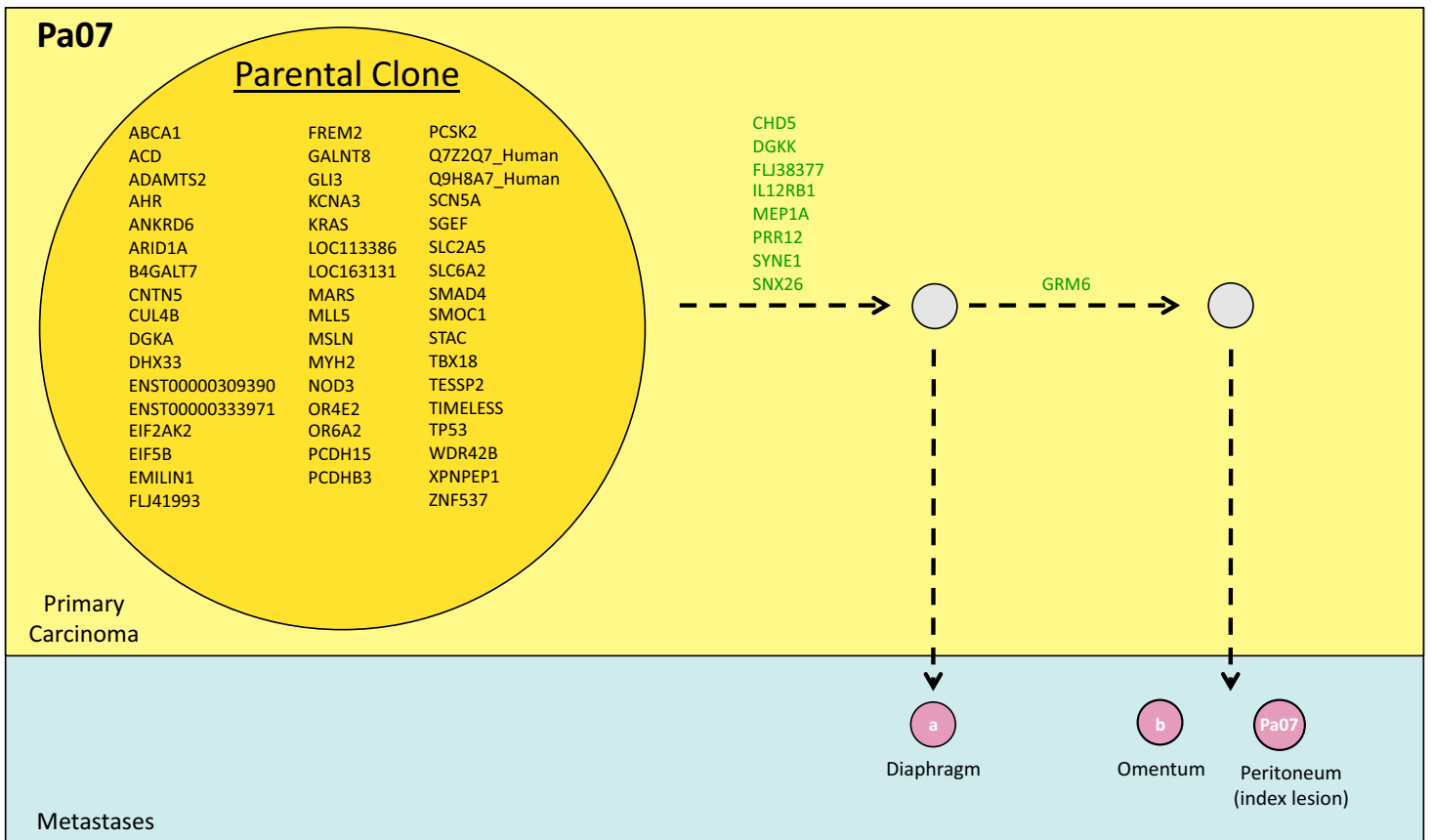
Supplementary Fig. 4. Proposed clonal evolution of Pa03.



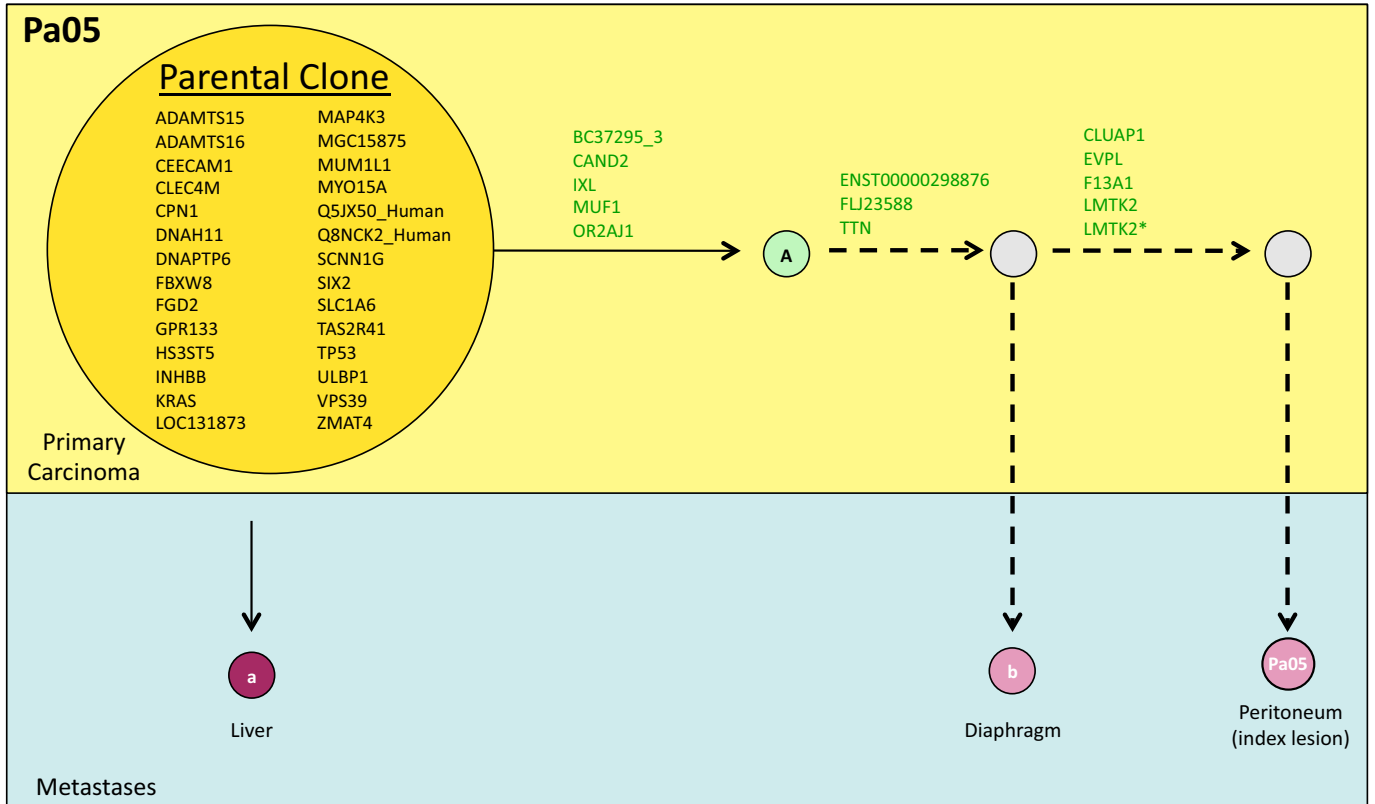
Supplementary Fig. 5. Proposed clonal evolution of Pa02.



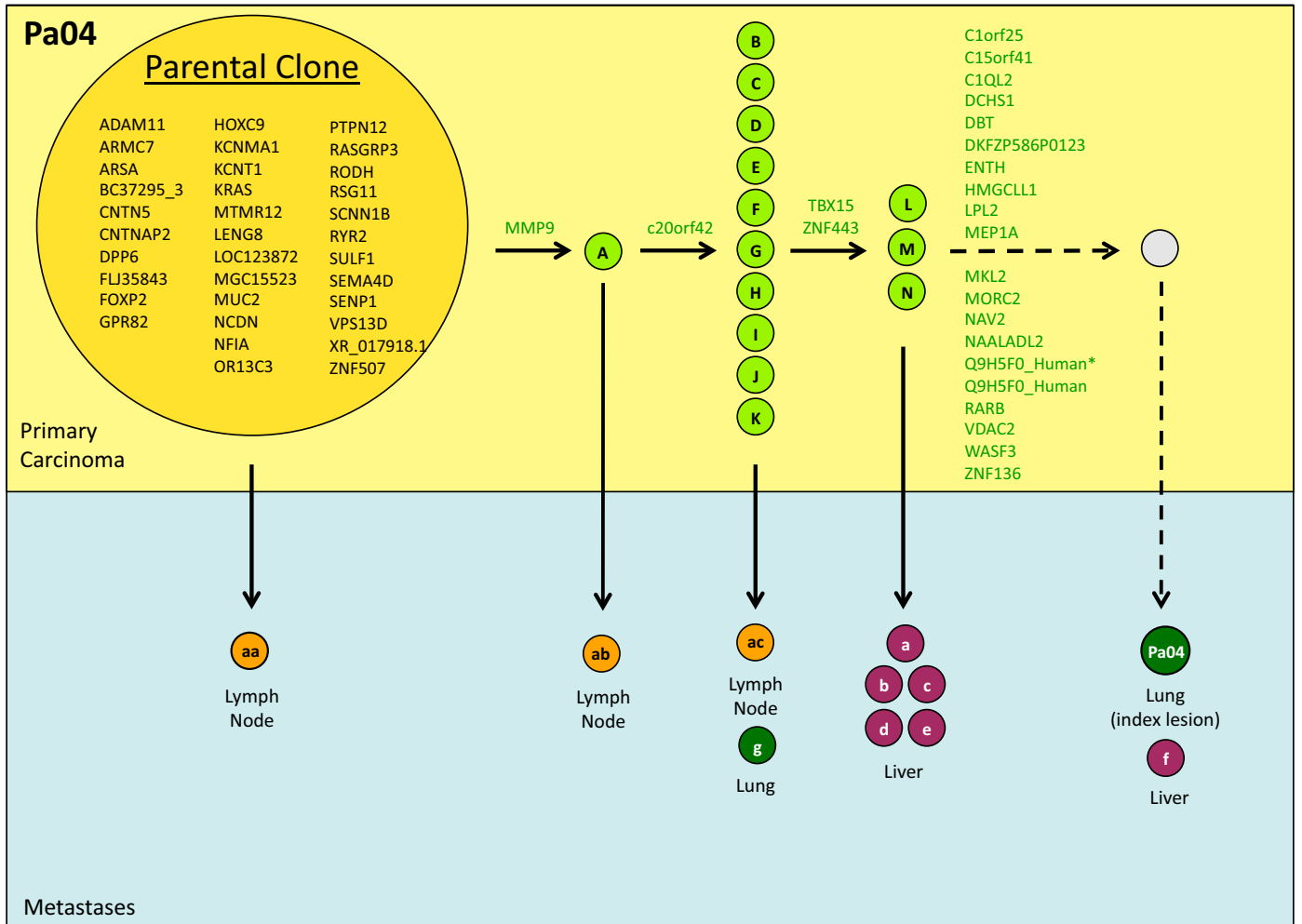
Supplementary Fig. 6. Proposed clonal evolution of Pa07.



Supplementary Fig. 7. Proposed clonal evolution of Pa05.



Supplementary Fig. 8. Proposed clonal evolution of Pa04.



**Supplementary Figs 3-8. Proposed clonal evolution based on the sequencing data.**

Somatic mutations identified in all samples (primary and metastases) from each patient are listed within the parental clone (indicated by an orange circle). Genes listed in green are acquired mutations (progressors) and genes listed in red were lost in the clones giving rise to specific metastases, indicated by their lower case alphabetical identifier. The index lesion for each patient is also indicated. Grey circles represent theoretical clones, and dashed arrows represent theoretical relationships between clones.

**Supplementary Table 1.** Clinical features of Seven Patients with Metastatic Pancreatic Ductal Adenocarcinoma

Patient no	Jones <i>et al.</i> <sup>7</sup> ID no.	Age/ Sex	TNM stage (UICC) at diagnosis	Therapy			Survival from diagnosis	Sites of Disease at Autopsy	Site of Index lesion
				Surgery	Adjuvant treatment	Treatment for advanced or recurrent disease			
A2	Pa01	62/M	cTxNxM1 (Stage IV)	No		Taxoprexin (4 cycles)→PD→Gemcitabine (1 cycle)	6 mo	Liver, Lung, Peritoneum	Liver
A6	Pa02	57/M	cT3NxM1 (Stage IV)	No		Gemcitabine (4 cycles)→PD→Troxacitabine (1 cycle)	8 mo	Liver, Lymph Node, Peritoneum	Liver
A10	Pa03	60/M	cT4NxM1 (Stage IV)	No			1 mo	Liver, Lung, Lymph Node, and Other	Liver
A13	Pa04	59/M	cTxNxM1 (Stage IV)	No			7mo	Liver, Lung, Lymph Node	Lung
A44	Pa05	50/M	pT3N1M0 (Stage IIB)	Yes	5-FU/XRT	Vaccine/Cyclophosphamide/Cetuximab	10 mo	Liver, Peritoneum	Peritoneum
A32	Pa07	81/M	cT4NxM0 (Stage III)	No		5-FU (2 cycles)/XRT	3 mo	Lung, Lymph Node, Peritoneum	Peritoneum
A38	Pa08	51/M	cTxNxM1 (Stage IV)	No		Gemcitabine/Bevacizumab (9 cycles)→PD→Rapamycin (1 cycle)	15 mo	Liver, Lung, Peritoneum	Lung

5-FU, 5-fluorouracil; XRT, radiotherapy; PD, progression of disease.

**Supplementary Table 2.** Somatic mutations identified in 7 Stage IV pancreatic cancers by whole exome sequencing<sup>a</sup>.

Gene	Transcript Accession ID	Tumor	Nucleotide (genomic)*	Nucleotide (cDNA)**	Amino acid (protein)**	Mutation Type	Mutated in Stage IV Index Lesions Only	Candidate Cancer Gene <sup>a</sup>
ABLIM2	NM_032432	Pa01C	g.chr4:8173987A>T	c.821A>T	p.K274M	Missense		
AFF3	NM_001025108	Pa01C	g.chr2:99668708G>A	c.2008G>A	p.V670M	Missense		
AHNAK	NM_024060	Pa01C	g.chr11:62059122G>A	c.271G>A	p.D91N	Missense		
AZU1	CCDS12044.1	Pa01C	g.chr19:781714C>T	c.367C>T	p.R123C	Missense		
BACH2	CCDS5026.1	Pa01C	g.chr6:90716990C>T	c.1556C>T	p.T519I	Missense		
BAIAP2L2	NM_025045	Pa01C	g.chr22:36818680C>T	c.355C>T	p.R119C	Missense		
BPIL3	CCDS13211.1	Pa01C	g.chr20:31084455C>G	c.102C>G	p.V34V	Synonymous		
C1orf129	NM_025063	Pa01C	g.chr1:167690800C>T	c.1026C>T	p.D342D	Synonymous		
CFHR4	NM_006684	Pa01C	g.chr1:193603315G>A	c.169G>A	p.D57N	Missense		
CHST1	CCDS7913.1	Pa01C	g.chr11:45628689C>T	c.361C>T	p.R121C	Missense		
COL4A1	CCDS9511.1	Pa01C	g.chr13:109623068G>T	c.3556G>T	p.G1186C	Missense		
CTNNA2	NM_004389	Pa01C	g.chr2:80008655G>A	c.524G>A	p.R175H	Missense		
CX40.1	CCDS7191.1	Pa01C	g.chr10:35937467G>A	c.1020G>A	p.A340A	Synonymous		
DIP2B	NM_173602	Pa01C	g.chr12:49407842G>A (homozygous)	c.3490G>A	p.G1164S	Missense		
DOCK2	CCDS4371.1	Pa01C	g.chr5:169059679G>T	c.1216G>T	p.V406L	Missense	Y	
DOT1L	NM_032482	Pa01C	g.chr19:2161814C>T	c.1311C>T	p.H437H	Synonymous		
ELAVL4	CCDS553.1	Pa01C	g.chr1:50373404C>G	c.660C>G	p.A220A	Synonymous		
ENST00000334548	ENST00000334548	Pa01C	g.chr3:132029537G>T	c.335G>T	p.C112F	Missense		
EPHA7	CCDS5031.1	Pa01C	g.chr6:94177446G>T	c.326G>T	p.C109F	Missense		
FBXO3	CCDS7887.1	Pa01C	g.chr11:33734042G>A	c.529G>A	p.V177I	Missense		
FLJ10647	CCDS406.1	Pa01C	g.chr1:36457669C>A	c.63C>A	p.I21I	Synonymous		
FRMPD4	NM_014728	Pa01C	g.chrX:12495340C>A	c.2738C>A	p.S913Y	Missense		
FTCD	CCDS13731.1	Pa01C	g.chr21:46382957C>T	c.1336C>T	p.R446W	Missense		
GLP1R	CCDS4839.1	Pa01C	g.chr6:39154880C>T	c.969C>T	p.I323I	Synonymous		

HIP1	NM_005338	Pa01C	g.chr7:74816871C>T	c.2840C>T	p.A947V	Missense		
KCNA4	NM_002233	Pa01C	g.chr11:29989035C>G	c.1767C>G	p.L589L	Synonymous		
KCNC3	CCDS12793.1	Pa01C	g.chr19:55515339C>T	c.2250C>T	p.N750N	Synonymous		
KIAA1102	NM_014988	Pa01C	g.chr4:41462176G>T	c.726G>T	p.L242L	Synonymous		
KIAA1751	ENST00000270720	Pa01C	g.chr1:1928626G>A	c.1448G>A	p.G480R	Missense		
KRAS	CCDS8702.1	Pa01C	g.chr12:25289551G>A	c.35G>A	p.G12D	Missense		Y
LOC387720	NM_001013633	Pa01C	g.chr10:129424833G>A	c.127G>A	p.G43R	Missense		
LOC89944	NM_138342	Pa01C	g.chr11:133720188C>T	c.404C>T	p.P135L	Missense		
LRRC3B	CCDS2644.1	Pa01C	g.chr3:26726196G>A (homozygous)	c.29G>A	p.R10H	Missense		
LRRC4	CCDS5799.1	Pa01C	g.chr7:127263383T>A	c.1262T>A	p.V421E	Missense		
MPL	CCDS483.1	Pa01C	g.chr1:43481660C>T	c.1270C>T	p.Q424X	Nonsense		
MYH3	CCDS11157.1	Pa01C	g.chr17:10483088G>C (homozygous)	c.3246G>C	p.K1082N	Missense		
NCOA2	NM_006540	Pa01C	g.chr8:71199554A>G (homozygous)	c.4017A>G	p.Q1339Q	Synonymous		
NIF3L1BP1	CCDS2900.1	Pa01C	g.chr3:63795942C>A	IVS6-3C>A	Splice Site	Splice Site		
OCA2	CCDS10020.1	Pa01C	g.chr15:25905345G>A	c.1222G>A	p.D408N	Missense		
ODZ4	ENST00000334175	Pa01C	g.chr11:78251725C>T	c.1104C>T	p.D368D	Synonymous		
OGT	CCDS14414.1	Pa01C	g.chrX:70540845T>A (homozygous)	c.364T>A	p.F122I	Missense		
OR10R2	NM_001004472	Pa01C	g.chr1:155263624C>T	c.884C>T	p.T295M	Missense	Y	
OR5C1	NM_001001923	Pa01C	g.chr9:122631133C>A	c.368C>A	p.A123D	Missense		
OVCH1	NM_183378	Pa01C	g.chr12:29519302G>A	c.1559G>A	p.R520H	Missense		Y
PADI2	CCDS177.1	Pa01C	g.chr1:17147557G>A	c.1378G>A	p.V460M	Missense		
PCDHA13	NM_031864	Pa01C	g.chr5:140235442G>A	c.201G>A	p.A67A	Synonymous		
PCDHGC4	CCDS4260.1	Pa01C	g.chr5:140791793delG	c.1283delG	fs	INDEL		
PCNT	NM_006031	Pa01C	g.chr21:46660418A>C	c.6158A>C	p.E2053A	Missense		
PCNXL2	ENST00000344698	Pa01C	g.chr1:229428820C>T	c.1949C>T	p.P650L	Missense		
PRDM10	CCDS8484.1	Pa01C	g.chr11:129298492G>A	c.1895G>A	p.R632H	Missense		
PRKD1	CCDS9637.1	Pa01C	g.chr14:29169824A>G	c.1547A>G	p.N516S	Missense		
PWP2H	NM_005049	Pa01C	g.chr21:44365318G>A	c.1543G>A	p.V515I	Missense		
PYHIN1	CCDS1178.1	Pa01C	g.chr1:155719945G>A	c.172G>A	p.G58S	Missense		
Q8N800_HUMAN	ENST00000322516	Pa01C	g.chr13:102185532G>A	c.1629G>A	p.T543T	Synonymous		
Q8NH06_HUMAN	ENST00000324144	Pa01C	g.chr17:3004230C>T (homozygous)	c.616C>T	p.R206C	Missense		
Q96GK3_HUMAN	ENST00000315264	Pa01C	g.chr14:103715712G>A	c.4575G>A	p.P1525P	Synonymous		
Q9HCM3_HUMAN	ENST00000242365	Pa01C	g.chr7:138003239C>T	c.4941C>T	p.P1647P	Synonymous		
Q9NSIO_HUMAN	ENST00000328881	Pa01C	g.chr21:45359175C>T	c.127C>T	p.R43C	Missense		

RHOT2	CCDS10417.1	Pa01C	g.chr16:663574C>A	c.1824C>A	p.F608L	Missense		
RNUT1	CCDS10281.1	Pa01C	g.chr15:73700386C>T	c.62C>T	p.T21I	Missense		
RREB1	NM_001003699	Pa01C	g.chr6:7127158G>A (homozygous)	c.15G>A	p.S5S	Synonymous		
SBLF	CCDS1840.1	Pa01C	g.chr2:48808746G>T	c.3325G>T	p.D1109Y	Missense		
SCN7A	NM_002976	Pa01C	g.chr2:167123728G>T	c.1842G>T	p.Q614H	Missense		
SLC17A1	CCDS4565.1	Pa01C	g.chr6:25938747_25938748dupGG	c.17_18dupG G	fs	INDEL		
SLC1A6	CCDS12321.1	Pa01C	g.chr19:14924831C>A	c.1408C>A	p.Q470K	Missense	Y	Y
SN	CCDS13060.1	Pa01C	g.chr20:3626618G>A	c.1949G>A	p.G650E	Missense		
SNRPC	NM_003093	Pa01C	g.chr6:34833298C>T	UTR-1C>T	5'UTR	5'UTR		
SOX3	CCDS14669.1	Pa01C	g.chrX:139312591C>T	c.155C>T	p.T52I	Missense		
STIM1	CCDS7749.1	Pa01C	g.chr11:4069576A>G	c.2030A>G	p.K677R	Missense		
TAF1L	NM_153809	Pa01C	g.chr9:32622253A>G	c.3325A>G	p.I1109V	Missense		
TEX14	NM_198393	Pa01C	g.chr17:54062833A>G	c.218A>G	p.D73G	Missense		
TJP1	NM_175610	Pa01C	g.chr15:27798569C>T	c.2829C>T	p.A943A	Synonymous		
TLL2	CCDS7449.1	Pa01C	g.chr10:98163018C>T	c.969C>T	p.G323G	Synonymous		
TP53	CCDS11118.1	Pa01C	g.chr17:7519192A>C (homozygous)	c.463A>C	p.T155P	Missense		Y
TPO	CCDS1642.1	Pa01C	g.chr2:1439044C>G	c.802C>G	p.P268A	Missense		
ZDHH4	CCDS5352.1	Pa01C	g.chr7:6394549G>A	c.189G>A	p.T63T	Synonymous		
ZFH1B	CCDS2186.1	Pa01C	g.chr2:145021235C>A	c.164C>A	p.P55H	Missense		
ABCD2	CCDS8734.1	Pa02C	g.chr12:38298955C>A	c.730C>A	p.Q244K	Missense		
ADAMTS20	NM_175851	Pa02C	g.chr12:42182299A>G	c.790A>G	p.M264V	Missense		
ADAMTS13	CCDS10326.1	Pa02C	g.chr15:82352540T>A (homozygous)	c.1363T>A	p.L455M	Missense		
AHNAK	NM_001620	Pa02C	g.chr11:62051161A>C	c.730A>C	p.K2435T	Missense		
ATP10B	ENST00000327245	Pa02C	g.chr5:159963569C>G	c.3218C>G	p.S1073C	Missense		
ATP1B2	NM_001678	Pa02C	g.chr17:7499652A>T	c.692A>T	p.Y231F	Missense		
C10orf93	CCDS7672.1	Pa02C	g.chr10:134593079G>A	c.912G>A	p.V304V	Synonymous		
C14orf121	NM_138360	Pa02C	g.chr14:23592778A>C (homozygous)	c.61A>C	p.S21R	Missense		
C6orf103	ENST00000326916	Pa02C	g.chr6:147103409T>C	c.1273T>C	p.F425L	Missense		
C6orf213	NM_001010852	Pa02C	g.chr6:123360840G>A	c.219G>A	p.A73A	Synonymous		
CAPN12	CCDS12519.1	Pa02C	g.chr19:43924315G>A	c.502G>A	p.V168M	Missense		
CCKAR	CCDS3438.1	Pa02C	g.chr4:26163779G>C (homozygous)	c.375G>C	p.V125V	Synonymous		
CCT6A	CCDS5523.1	Pa02C	g.chr7:55902305delG	c.1201delG	fs	INDEL		
CD6	CCDS7999.1	Pa02C	g.chr11:60541907C>G	c.1683C>G	p.N561K	Missense		

CHD7	NM_017780	Pa02C	g.chr8:61940604A>G	c.8552A>G	p.E2851G	Missense		
CLK1	CCDS2331.1	Pa02C	g.chr2:201551699C>G	IVS1-4C>G	Splice Site	Splice Site		
CLSTN2	CCDS3112.1	Pa02C	g.chr3:141767761C>T	c.2836C>T	p.L946L	Synonymous		
DNM1L	CCDS8728.1	Pa02C	g.chr12:32775629G>C	c.1273G>C	p.E425Q	Missense		
DYSF	CCDS1918.1	Pa02C	g.chr2:71706789G>A	c.2565G>A	p.K855K	Synonymous		
EVI1	CCDS3205.1	Pa02C	g.chr3:170317123G>T	c.675G>T	p.M225I	Missense		
FBN3	CCDS12196.1	Pa02C	g.chr19:8082567G>A	c.4049G>A	p.R1350H	Missense		
FLJ12886	NM_019108	Pa02C	g.chr19:48943738C>T	c.377C>T	p.P126L	Missense		
FLJ25530	CCDS8456.1	Pa02C	g.chr11:124300140C>T (homozygous)	c.121C>T	p.R41C	Missense		
FLJ32110	CCDS5613.1	Pa02C	g.chr7:88610529C>T	c.3582C>T	p.T1194T	Synonymous		
FLJ32416	CCDS12086.1	Pa02C	g.chr19:2204762C>T	c.293C>T	p.A98V	Missense		
FLJ45909	CCDS12522.1	Pa02C	g.chr19:44052611C>A	c.812C>A	p.A271E	Missense		
FLJ46481	CCDS3384.1	Pa02C	g.chr4:6084945G>T (homozygous)	IVS5-1G>T	Splice Site	Splice Site		
FLJ46481	CCDS3384.1	Pa02C	g.chr4:6084863C>T (homozygous)	c.539C>T	p.P180L	Missense		
GALNT13	CCDS2199.1	Pa02C	g.chr2:155120619C>A	c.1403C>A	p.S468Y	Missense		
GAS7	CCDS11152.1	Pa02C	g.chr17:9813791C>T (homozygous)	c.399C>T	p.H133H	Synonymous		
GRIK2	CCDS5048.1	Pa02C	g.chr6:102609988G>T	c.2402G>T	p.G801V	Missense		
GRIP2	ENST00000273083	Pa02C	g.chr3:14530868G>A (homozygous)	c.1432G>A	p.A478T	Missense		
HKR2	CCDS12975.1	Pa02C	g.chr19:63538050G>C	c.70G>C	p.E24Q	Missense		
HOXD4	CCDS2269.1	Pa02C	g.chr2:176841961G>T	c.93G>T	p.E31D	Missense		
HS3ST2	CCDS10606.1	Pa02C	g.chr16:22834012A>G (homozygous)	c.732A>G	p.S244S	Synonymous		
IRGQ	NM_001007561	Pa02C	g.chr19:48791322G>A	c.9G>A	p.P3P	Synonymous		
KIAA1944	CCDS9266.1	Pa02C	g.chr12:128083339G>A	c.3261G>A	p.E1087E	Synonymous		
KRAS	CCDS8702.1	Pa02C	g.chr12:25271542A>C (homozygous)	c.183A>C	p.Q61H	Missense		Y
LOC441136	NM_001013719	Pa02C	g.chr6:27464731C>A	c.119C>A	p.A40E	Missense		
LOC63920	NM_022090	Pa02C	g.chr5:159754461C>G	c.615C>G	p.L205L	Synonymous		
LPHN3	NM_015236	Pa02C	g.chr4:62738940C>T	c.3517C>T	p.R1173X	Nonsense		
MGC10731	CCDS171.1	Pa02C	g.chr1:16304359G>A	c.479G>A	p.R160H	Missense		
MGC4266	CCDS8522.1	Pa02C	g.chr12:3659708T>C (homozygous)	c.297T>C	p.D99D	Synonymous		
MLL	NM_005933	Pa02C	g.chr11:117878646G>C	c.6820G>C	p.G2274R	Missense		
MLL	NM_005933	Pa02C	g.chr11:117879596T>G	c.7770T>G	p.N2590K	Missense		
MLL3	CCDS5931.1	Pa02C	g.chr7:151283363G>A	c.13297G>A	p.A4433T	Missense		Y
NKX2-2	CCDS13145.1	Pa02C	g.chr20:21442284G>A	c.24G>A	p.T8T	Synonymous		
OR5D3P	ENST00000333984	Pa02C	g.chr11:55250887A>T	c.505A>T	p.I169L	Missense		

PCDH17	NM_014459	Pa02C	g.chr13:57106030C>T	c.1349C>T	p.P450L	Missense		
PLCB2	NM_004573	Pa02C	g.chr15:38373206G>C	c.2061G>C	p.G687G	Synonymous		
PLEKHA6	CCDS1444.1	Pa02C	g.chr1:200942238G>A	c.2331G>A	p.S777S	Synonymous		
Q9H8A7_HUMAN	ENST00000053084	Pa02C	g.chr4:15211134C>T (homozygous)	c.1369C>T	p.L457F	Missense	Y	Y
Q9P193_HUMAN	ENST00000359406	Pa02C	g.chr5:60447362C>A	c.158C>A	p.P53Q	Missense		
RIMS1	NM_014989	Pa02C	g.chr6:72863449G>A	c.322G>A	p.E108K	Missense		
RREB1	NM_001003699	Pa02C	g.chr6:7175652C>A	c.2321C>A	p.T774K	Missense		
RYR1	NM_000540	Pa02C	g.chr19:43750265G>A	c.13527G>A	p.G4509G	Synonymous		
SCFD2	NM_152540	Pa02C	g.chr4:54059762C>T (homozygous)	c.938C>T	p.A313V	Missense		
SHKBP1	CCDS12560.1	Pa02C	g.chr19:45786501G>A	c.1468G>A	p.G490S	Missense		
SLC6A3	CCDS3863.1	Pa02C	g.chr5:1475122G>A	c.661G>A	p.V221M	Missense		
SULF2	CCDS13408.1	Pa02C	g.chr20:45738667C>T	c.1358C>T	p.T453M	Missense		
TP53	CCDS11118.1	Pa02C	g.chr17:7518236T>C (homozygous)	c.770T>C	p.L257P	Missense		Y
TXNDC6	CCDS3099.1	Pa02C	g.chr3:139505168C>T	c.624C>T	p.Y208Y	Synonymous		
UBR2	CCDS4870.1	Pa02C	g.chr6:42679418G>A	c.646G>A	p.A216T	Missense		
WBSCR17	CCDS5540.1	Pa02C	g.chr7:70330682G>T	c.902G>T	p.C301F	Missense		
ACTL7B	CCDS6771.1	Pa03C	g.chr9:108696542C>T	c.1224C>T	p.S408S	Synonymous		
C19orf22	CCDS12048.1	Pa03C	g.chr19:848479G>A (homozygous)	c.765G>A	p.P255P	Synonymous		
C1orf14	NM_030933	Pa03C	g.chr1:179641193T>C	c.698T>C	p.L233P	Missense		
CDH10	CCDS3892.1	Pa03C	g.chr5:24523720A>T	c.2176A>T	p.T726S	Missense		
CNTN6	CCDS2557.1	Pa03C	g.chr3:1393745G>A (homozygous)	c.2152G>A	p.V718I	Missense		
CTNNA3	CCDS7269.1	Pa03C	g.chr10:67710272A>C (homozygous)	c.1846A>C	p.T616P	Missense		
CYP1A1	CCDS10268.1	Pa03C	g.chr15:72802200C>T	c.292C>T	p.R98W	Missense		
DLC1	CCDS5989.1	Pa03C	g.chr8:12990385G>A	c.4274G>A	p.R1425Q	Missense		
DPP6	NM_130797	Pa03C	g.chr7:154033237G>A	c.1423G>A	p.D475N	Missense		Y
DRD3	CCDS2978.1	Pa03C	g.chr3:115332814C>T (homozygous)	c.847C>T	p.R283W	Missense		
EP300	CCDS14010.1	Pa03C	g.chr22:39881227G>A	IVS20+1G>A	Splice Site	Splice Site		
FAD158	CCDS725.1	Pa03C	g.chr1:89890904G>A	c.754G>A	p.V252M	Missense		
FBXO15	CCDS12002.1	Pa03C	g.chr18:69900151C>A	c.1026C>A	p.G342G	Synonymous		
FLJ14011	CCDS12944.1	Pa03C	g.chr19:61644429T>C	c.1747T>C	p.Y583H	Missense		
HELZ	NM_014877	Pa03C	g.chr17:62616878C>T	c.613C>T	p.Q205X	Nonsense		
HIST1H4J	CCDS4630.1	Pa03C	g.chr6:27899909G>C	c.28G>C	p.G10R	Missense		
ITGB4BP	CCDS13249.1	Pa03C	g.chr20:33331245_33331244delTT	c.460_461delTT	fs	INDEL		

KIAA1755	NM_001029864	Pa03C	g.chr20:36302825G>A	c.1122G>A	p.M374I	Missense		
KIT	CCDS3496.1	Pa03C	g.chr4:55405430C>T	c.390C>T	p.N130N	Synonymous		
KRAS	CCDS8702.1	Pa03C	g.chr12:25289551G>A	c.35G>A	p.G12D	Missense		Y
LRP1	CCDS8932.1	Pa03C	g.chr12:55842541C>T	c.2377C>T	p.R793X	Nonsense		
MAN2A1	NM_002372	Pa03C	g.chr5:109230526C>T (homozygous)	c.3363C>T	p.P1121P	Synonymous		
MAPT	CCDS11499.1	Pa03C	g.chr17:41451831G>T	c.998G>T	p.G333V	Missense		
MLL3	CCDS5931.1	Pa03C	g.chr7:151329239C>T	c.4441C>T	p.R1481X	Nonsense		Y
PABPC5	CCDS14460.1	Pa03C	g.chrX:90496873G>A (homozygous)	c.152G>A	p.R51H	Missense		
PREX1	CCDS13410.1	Pa03C	g.chr20:46677552G>T	c.4878G>T	p.R1626S	Missense		
PTPRK	CCDS5137.1	Pa03C	g.chr6:128339562A>T	c.3902A>T	p.E1301V	Missense		
SLCO1A2	CCDS8686.1	Pa03C	g.chr12:21348642C>T	c.575C>T	p.S192F	Missense		
SMARCA4	CCDS12253.1	Pa03C	g.chr19:11005005C>T (homozygous)	c.3586C>T	p.Q1196X	Nonsense		
SYT6	CCDS871.1	Pa03C	g.chr1:114392518C>T	c.457C>T	p.R153C	Missense		
T	CCDS5290.1	Pa03C	g.chr6:166550664C>T	c.298C>T	p.H100Y	Missense		
TBC1D14	CCDS3394.1	Pa03C	g.chr4:7043315G>A	c.82G>A	p.A28T	Missense		
TFCP2L1	CCDS2134.1	Pa03C	g.chr2:121716181C>A	c.752C>A	p.T251N	Missense		
TNF	CCDS4702.1	Pa03C	g.chr6:31653210C>T	c.619C>T	p.R207X	Nonsense		
TNR	CCDS1318.1	Pa03C	g.chr1:172104173G>A	c.736G>A	p.V246M	Missense		
TP53	CCDS11118.1	Pa03C	g.chr17:7514721T>C (homozygous)	c.1031T>C	p.L344P	Missense		Y
TTN	NM_133378	Pa03C	g.chr2:179236181A>G	c.87585A>G	p.P29195P	Synonymous		
VWF	CCDS8539.1	Pa03C	g.chr12:5929302C>G	c.8164C>G	p.P2722A	Missense		
WDR1	NM_005112	Pa03C	g.chr4:9794081G>C	c.63G>C	p.K21N	Missense		
XKR4	NM_052898	Pa03C	g.chr8:56177818C>T	c.216C>T	p.G72G	Synonymous		
ZBTB7	CCDS12119.1	Pa03C	g.chr19:4005888G>T	c.343G>T	p.E115X	Nonsense		
AARS	NM_001605	Pa04C	g.chr16:68857047C>T	c.1242C>T	p.L414L	Synonymous		
ADAM11	CCDS11486.1	Pa04C	g.chr17:40210852A>C	c.2077A>C	p.S693R	Missense		
ARMC7	CCDS11714.1	Pa04C	g.chr17:70618173_70618184delGCCA ACCTCGCC	c.112_123del GCCAACCTCG CC	fs	INDEL		
ARSA	CCDS14100.1	Pa04C	g.chr22:49355454C>T	c.630C>T	p.A210A	Synonymous		
BC37295_3	NM_001005850	Pa04C	g.chr19:61868151C>T	c.294C>T	p.D98D	Synonymous		
C15orf41	NM_032499	Pa04C	g.chr15:34887823G>C	c.429G>C	p.G143G	Synonymous		
C1orf25	CCDS1366.1	Pa04C	g.chr1:181838554G>C	c.1354G>C	p.V452L	Missense		
C1QL2	NM_182528	Pa04C	g.chr2:119631996G>C	c.80G>C	p.G27A	Missense		

C20orf42	CCDS13098.1	Pa04C	g.chr20:6005895G>T	c.1959G>T	p.L653F	Missense		
CNTN5	NM_014361	Pa04C	g.chr11:99195617G>T	c.188G>T	p.S63I	Missense	Y	Y
CNTNAP2	CCDS5889.1	Pa04C	g.chr7:146530473G>T	c.1623G>T	p.P541P	Synonymous		
DBT	CCDS767.1	Pa04C	g.chr7:100388309G>A	c.979G>A	p.A327T	Missense		
DCHS1	CCDS7771.1	Pa04C	g.chr11:6601843G>T	c.7640G>T	p.G2547V	Missense		
DKFZP586P0123	NM_015531	Pa04C	g.chr11:73466993C>G	c.4418C>G	p.A1473G	Missense		
DPP6	NM_130797	Pa04C	g.chr7:154118657G>A	c.2332G>A	p.A778T	Missense		Y
EBF3	NM_001005463	Pa04C	g.chr10:131561760C>T	c.727C>T	p.R243W	Missense		
ENTH	NM_014666	Pa04C	g.chr5:157154482T>C	IVS9+2T>C	Splice Site	Splice Site		
FLJ35843	CCDS9151.1	Pa04C	g.chr12:109780527A>G	c.587A>G	p.Y196C	Missense		
FOXP2	CCDS5760.1	Pa04C	g.chr7:113862679A>G	c.392A>G	p.Q131R	Missense		
GPR82	CCDS14259.1	Pa04C	g.chrX:41343413C>A (homozygous)	c.880C>A	p.L294I	Missense		
HMGCLL1	NM_019036	Pa04C	g.chr6:55489347C>T	c.401C>T	p.T134I	Missense		
HOXC9	CCDS8869.1	Pa04C	g.chr12:52680669G>T (homozygous)	c.430G>T	p.G144C	Missense		
KCNMA1	CCDS7352.1	Pa04C	g.chr10:78538318G>A	c.1150G>A	p.V384I	Missense		
KCNT1	NM_020822	Pa04C	g.chr9:135890629C>T	c.2132C>T	p.A711V	Missense		
KRAS	CCDS8702.1	Pa04C	g.chr12:25289551G>T (homozygous)	c.35G>T	p.G12V	Missense		Y
LENG8	CCDS12894.1	Pa04C	g.chr19:59658439C>T	c.906C>T	p.T302T	Synonymous		
LOC123872	CCDS10943.1	Pa04C	g.chr16:82750807G>A	c.507G>A	p.P169P	Synonymous		
LPAL2	ENST00000342479	Pa04C	g.chr6:160858399G>A	c.300G>A	p.T100T	Synonymous		
MEP1A	CCDS4918.1	Pa04C	g.chr6:46905073T>C	c.950T>C	p.F317S	Missense	Y	Y
MGC15523	CCDS11780.1	Pa04C	g.chr17:76840601A>G	c.1934A>G	p.D645G	Missense		
MGC20806	CCDS11797.1	Pa04C	g.chr17:77576240C>T	c.429C>T	p.N143N	Synonymous		
MKL2	NM_014048	Pa04C	g.chr16:14247913C>A	c.1295C>A	p.P432H	Missense		
MMP9	CCDS13390.1	Pa04C	g.chr20:44073223G>A	c.684G>A	p.A228A	Synonymous		
MORC2	NM_014941	Pa04C	g.chr22:29662658A>G	c.395A>G	p.D132G	Missense		
MTMR12	NM_019061	Pa04C	g.chr5:32269694_32269695delCT	c.1622_1623 delCT	fs	INDEL		
MUC2	NM_002457	Pa04C	g.chr11:1093892G>A	c.8179G>A	p.V2727I	Missense		
NAALADL2	NM_207015	Pa04C	g.chr3:176647739C>T	c.1111C>T	p.R371X	Nonsense		
NAV2	CCDS7850.1	Pa04C	g.chr11:19917888A>C	c.2139A>C	p.E713D	Missense		
NCDN	CCDS392.1	Pa04C	g.chr1:35698033C>T	c.1523C>T	p.S508L	Missense		
NFIA	CCDS615.1	Pa04C	g.chr1:61265954G>A	c.140G>A	p.R47H	Missense		
OBSCN	CCDS1570.1	Pa04C	g.chr1:224774783G>A	c.7832G>A	p.R2611H	Missense		

OR10R2	NM_001004472	Pa04C	g.chr1:155263066C>G	c.326C>G	p.A109G	Missense	Y	
OR13C3	NM_001001961	Pa04C	g.chr9:104378262C>A	c.388C>A	p.Q130K	Missense		
PER3	CCDS89.1	Pa04C	g.chr1:7821471_7821474delCGGC	c.2192_2195 delCGGC	fs	INDEL		
PIK3CG	CCDS5739.1	Pa04C	g.chr7:106103628G>A	c.1671G>A	p.A557A	Synonymous		
PTPN12	CCDS5592.1	Pa04C	g.chr7:76900982G>C	c.1335G>C	p.E445D	Missense		
Q8N800_HUMAN	ENST00000322516	Pa04C	g.chr13:102185624G>T	c.1537G>T	p.E513X	Nonsense		
Q9H5F0_HUMAN	ENST00000360484	Pa04C	g.chr1:241614343C>A	c.101C>A	p.P34H	Missense		
Q9H5F0_HUMAN	ENST00000360484	Pa04C	g.chr1:241614381C>A	c.139C>A	p.P47T	Missense		
RARB	CCDS2642.1	Pa04C	g.chr3:25517765G>A	c.395G>A	p.R132Q	Missense		
RASGRP3	NM_170672	Pa04C	g.chr2:33660595A>G	IVS5-2A>G	Splice Site	Splice Site		
RGS11	CCDS10403.1	Pa04C	g.chr16:259383C>T	c.1227C>T	p.R409R	Synonymous		
RODH	CCDS8925.1	Pa04C	g.chr12:55464963C>T	c.632C>T	p.T211M	Missense		
RYR2	NM_001035	Pa04C	g.chr1:233906412G>A	c.796G>A	p.A266T	Missense		
SCNN1B	CCDS10609.1	Pa04C	g.chr16:23299368G>A	c.1668G>A	p.V556V	Synonymous		
SEMA4D	CCDS6685.1	Pa04C	g.chr9:89223296C>T	c.2466C>T	p.I822I	Synonymous		
SENP1	NM_014554	Pa04C	g.chr12:46745163dupA	c.1227dupA	fs	INDEL		
STAMPB	CCDS1929.1	Pa04C	g.chr2:73986261A>T	c.468A>T	p.Q156H	Missense		
SULF1	CCDS6204.1	Pa04C	g.chr8:70695902C>T (homozygous)	c.1456C>T	p.R486W	Missense		
TBX15	NM_152380	Pa04C	g.chr1:119139811C>T	c.1077C>T	p.Y359Y	Synonymous		
VDAC2	CCDS7348.1	Pa04C	g.chr10:76648965G>A	c.289G>A	p.A97T	Missense		
VPS13D	NM_018156	Pa04C	g.chr1:12287936G>A	c.5942G>A	p.R1981H	Missense		
WASF3	CCDS9318.1	Pa04C	g.chr13:26157827A>G	c.1354A>G	p.I452V	Missense		
XR_017918.1	ENST00000258651	Pa04C	g.chr13:53913399G>A	c.53G>A	p.R18H	Missense		
ZNF136	NM_003437	Pa04C	g.chr19:12158475G>C	c.282G>C	p.K94N	Missense		
ZNF443	NM_005815	Pa04C	g.chr19:12402716A>C	c.1270A>C	p.K424Q	Missense		
ZNF507	NM_014910	Pa04C	g.chr19:37536920T>C	c.1344T>C	p.T448T	Synonymous		
ADAMTS15	CCDS8488.1	Pa05X	g.chr11:129837197G>A	c.1096G>A	p.V366M	Missense		
ADAMTS16	NM_139056	Pa05X	g.chr5:5292308C>T	c.2199C>T	p.D733D	Synonymous		
BC37295_3	NM_001005850	Pa05X	g.chr19:61867445_61867409delGGC GCGCTCTTCAGCCAGAGCGCCTCTCTG GCCGAGC	c.1000_1037 delGGCGCGC TCTTCAGCCA GAGCGCCTCT CTGCCGAG C	indel	INDEL		

CAND2	ENST00000295989	Pa05X	g.chr3:12833092C>T	c.1382C>T	p.P461L	Missense		
CEECAM1	CCDS6901.1	Pa05X	g.chr9:128266306C>T (homozygous)	c.391C>T	p.R131C	Missense		
CLEC4M	CCDS12187.1	Pa05X	g.chr19:7739763C>T	c.1089C>T	p.S363S	Synonymous		
CLUAP1	NM_015041	Pa05X	g.chr16:3509954A>C	c.630A>C	p.E210D	Missense		
CPN1	CCDS7486.1	Pa05X	g.chr10:101806789C>T (homozygous)	c.982C>T	p.R328W	Missense		
DNAH11	NM_003777	Pa05X	g.chr7:21432055C>T	c.4367C>T	p.A1456V	Missense		
DNAPTP6	NM_015535	Pa05X	g.chr2:201109563C>G	c.283C>G	p.P95A	Missense		
ENST00000298876	ENST00000298876	Pa05X	g.chr12:27514906G>T	c.75G>T	p.L25L	Synonymous		
EVPL	CCDS11737.1	Pa05X	g.chr17:71517727C>T	c.3154C>T	p.R1052W	Missense		
F13A1	CCDS4496.1	Pa05X	g.chr6:6119855A>G	c.1704A>G	p.E568E	Synonymous		
FBXW8	CCDS9182.1	Pa05X	g.chr12:115928634C>G	c.1734C>G	p.S578S	Synonymous		
FGD2	CCDS4829.1	Pa05X	g.chr6:37103268G>A	c.1691G>A	p.R564Q	Missense		
FLJ23588	CCDS14049.1	Pa05X	g.chr22:42297206C>G	c.3292C>G	p.L1098V	Missense		
GPR133	CCDS9272.1	Pa05X	g.chr12:130141219C>T	c.2245C>T	p.H749Y	Missense		
HS3ST5	NM_153612	Pa05X	g.chr6:114490631C>T (homozygous)	c.72C>T	p.L24L	Synonymous		
INHBB	CCDS2132.1	Pa05X	g.chr2:120823305C>T	c.849C>T	p.G283G	Synonymous		
IXL	NM_017592	Pa05X	g.chr19:44573866C>T	c.27C>T	p.R9R	Synonymous		
KRAS	CCDS8702.1	Pa05X	g.chr12:25289551G>A	c.35G>A	p.G12D	Missense		Y
LMTK2	CCDS5654.1	Pa05X	g.chr7:97467809C>T	c.3381C>T	p.A1127A	Synonymous	Y	
LMTK2	CCDS5654.1	Pa05X	g.chr7:97467810T>C	c.3382T>C	p.L1128L	Synonymous	Y	
LOC131873	ENST00000358511	Pa05X	g.chr3:131769843C>G	c.1999C>G	p.L667V	Missense		
MAP4K3	CCDS1803.1	Pa05X	g.chr2:39448258G>A	c.1051G>A	p.E351K	Missense		
MGC15875	CCDS4434.1	Pa05X	g.chr5:177582153G>T	c.736G>T	p.D246Y	Missense		
MUF1	CCDS533.1	Pa05X	g.chr1:46476057C>A	c.205C>A	p.P69T	Missense		
MUM1L1	NM_152423	Pa05X	g.chrX:105255948C>T (homozygous)	c.378C>T	p.S126S	Synonymous		
MYO15A	NM_016239	Pa05X	g.chr17:17982902C>T	c.5060C>T	p.P1687L	Missense		
OR2AJ1	ENST00000318244	Pa05X	g.chr1:244423744C>A	c.633C>A	p.F211L	Missense		
Q5JX50_HUMAN	ENST00000325076	Pa05X	g.chr20:33653256C>T	c.672C>T	p.R224R	Synonymous		
Q8NCK2_HUMAN	ENST00000325720	Pa05X	g.chr17:71097866G>A	c.606G>A	p.K202K	Synonymous		
SCNN1G	CCDS10608.1	Pa05X	g.chr16:23134233G>A	c.1892G>A	p.R631H	Missense		
SIX2	CCDS1822.1	Pa05X	g.chr2:45145131G>A	c.705G>A	p.P235P	Synonymous		
SLC1A6	CCDS12321.1	Pa05X	g.chr19:14928457G>A	c.1000G>A	p.V334I	Missense	Y	Y
TAS2R41	NM_176883	Pa05X	g.chr7:142692636C>T	c.834C>T	p.V278V	Synonymous		
TP53	CCDS11118.1	Pa05X	g.chr17:7518284C>T (homozygous)	c.722C>T	p.S241F	Missense		Y

TTN	NM_133378	Pa05X	g.chr2:179311902G>A	c.37451G>A	p.C12484Y	Missense		
ULBP1	CCDS5223.1	Pa05X	g.chr6:150382336G>A (homozygous)	c.351G>A	p.E117E	Synonymous		
VPS39	CCDS10083.1	Pa05X	g.chr15:40264103G>A	c.622G>A	p.A208T	Missense		
ZMAT4	NM_024645	Pa05X	g.chr8:40802278C>T	c.75C>T	p.S25S	Synonymous		
ABCA1	CCDS6762.1	Pa07C	g.chr9:104672874C>T	c.1779C>T	p.F593F	Synonymous		
ACD	CCDS10842.1	Pa07C	g.chr16:66249088G>A	c.1548G>A	p.R516R	Synonymous		
ADAM21	CCDS9804.1	Pa07C	g.chr14:69994312G>A	c.343G>A	p.V115M	Missense		
ADAMTS2	CCDS4444.1	Pa07C	g.chr5:178487615C>T	c.2568C>T	p.Y856Y	Synonymous		
AHR	CCDS5366.1	Pa07C	g.chr7:17152357G>T	c.1668G>T	p.M556I	Missense		
ALDH1A3	CCDS10389.1	Pa07C	g.chr15:99237656C>T	c.21C>T	p.A7A	Synonymous		
ANKRD6	NM_014942	Pa07C	g.chr6:90388378G>A	c.809G>A	p.R270H	Missense		
ARID1A	CCDS285.1	Pa07C	g.chr1:26784089C>T	c.3826C>T	p.R1276X	Nonsense		
B4GALT7	CCDS4429.1	Pa07C	g.chr5:176964045C>T	c.310C>T	p.R104C	Missense		
CHD5	CCDS57.1	Pa07C	g.chr1:6101062C>T	c.5622C>T	p.D1874D	Synonymous		
CNTN5	NM_014361	Pa07C	g.chr11:99631845C>T	c.2149C>T	p.Q717X	Nonsense	Y	Y
CUL4B	NM_003588	Pa07C	g.chrX:119476101C>T	c.329C>T	p.S110F	Missense		
DGKA	CCDS8896.1	Pa07C	g.chr12:54621336G>A	c.1135G>A	p.V379I	Missense		
DGKK	NM_001013742	Pa07C	g.chrX:50046399_50046376delTGCCA CAGAGCCGCCCAAGAAC	c.315_338del TGCCACAGA GCCGGCCCC AGAACC	indel	INDEL		
DHX33	CCDS11072.1	Pa07C	g.chr17:5312806T>C (homozygous)	c.98T>C	p.V33A	Missense		
EIF2AK2	CCDS1786.1	Pa07C	g.chr2:37261448C>T	c.919C>T	p.R307C	Missense		
EIF5B	NM_015904	Pa07C	g.chr2:99469722G>A	c.3112G>A	p.D1038N	Missense		
EMILIN1	CCDS1733.1	Pa07C	g.chr2:27219041G>A	c.2548G>A	p.G850R	Missense		
ENST00000309390	ENST00000309390	Pa07C	g.chr6:7933061C>T	c.1293C>T	p.H431H	Synonymous		
ENST00000333971	ENST00000333971	Pa07C	g.chr18:70208268G>A (homozygous)	c.257G>A	p.R86H	Missense		
FLJ38377	CCDS2164.1	Pa07C	g.chr2:131354922G>A	c.1538G>A	p.R513H	Missense		
FLJ41993	NM_001001694	Pa07C	g.chr22:48739118T>C	c.679T>C	p.C227R	Missense		
FREM2	NM_207361	Pa07C	g.chr13:38164611C>T	c.5130C>T	p.L1710L	Synonymous		
GALNT8	CCDS8533.1	Pa07C	g.chr12:4740462C>T	c.1251C>T	p.Y417Y	Synonymous		
GLI3	CCDS5465.1	Pa07C	g.chr7:41961112C>T	c.320C>T	p.T107M	Missense		
GRM6	CCDS4442.1	Pa07C	g.chr5:178354480G>A	c.72G>A	p.A24A	Synonymous		
IL12RB1	NM_153701	Pa07C	g.chr19:18048135C>T	c.552C>T	p.G184G	Synonymous		

KCNA3	CCDS828.1	Pa07C	g.chr1:110928628G>A	c.690G>A	p.S230S	Synonymous		
KIN	CCDS7080.1	Pa07C	g.chr10:7838072G>A	c.1159G>A	p.E387K	Missense		
KRAS	CCDS8702.1	Pa07C	g.chr12:25289551G>A	c.35G>A	p.G12D	Missense		Y
LOC113386	NM_138781	Pa07C	g.chr19:63515554G>T	c.208G>T	p.E70X	Nonsense		
LOC163131	NM_001005851	Pa07C	g.chr19:45232382C>T	c.2224C>T	p.Q742X	Nonsense		
MARS	CCDS8942.1	Pa07C	g.chr12:56170671G>A	c.747G>A	p.P249P	Synonymous		
MCTP2	NM_018349	Pa07C	g.chr15:92689362C>T	c.970C>T	p.R324C	Missense		
MEP1A	CCDS4918.1	Pa07C	g.chr6:46905163delT	c.1037delT	fs	INDEL	Y	Y
MLL5	NM_182931	Pa07C	g.chr7:104325651delC	c.1661delC	fs	INDEL		
MSLN	NM_013404	Pa07C	g.chr16:755236G>A	c.636G>A	p.P212P	Synonymous		
MYH2	CCDS11156.1	Pa07C	g.chr17:10369514C>T	c.4516C>T	p.R1506X	Nonsense		
NOD3	NM_178844	Pa07C	g.chr16:3567189dupC	c.27dupC	fs	INDEL		
OR4E2	NM_001001912	Pa07C	g.chr14:21203968G>A	c.832G>A	p.V278I	Missense		
OR6A2	CCDS7772.1	Pa07C	g.chr11:6773137C>T	c.379C>T	p.R127C	Missense		
PCDH15	CCDS7248.1	Pa07C	g.chr10:55666676G>T	c.898G>T	p.V300F	Missense		
PCDH8B3	CCDS4245.1	Pa07C	g.chr5:140462670C>T	c.2253C>T	p.Y751Y	Synonymous		
PCSK2	CCDS13125.1	Pa07C	g.chr20:17382461C>T (homozygous)	c.960C>T	p.D320D	Synonymous		
PFAS	CCDS11136.1	Pa07C	g.chr17:8111125C>A	c.3022C>A	p.R1008R	Synonymous		
PRR12	ENST00000246798	Pa07C	g.chr19:54794366G>A	c.1244G>A	p.G415D	Missense		
Q7Z2Q7_HUMAN	ENST00000334994	Pa07C	g.chr5:61912677T>C	c.1653T>C	p.F551F	Synonymous		
Q9H8A7_HUMAN	ENST00000053084	Pa07C	g.chr4:15264100_15264101delTC	c.3762_3763 delTC	fs	INDEL	Y	Y
SCN5A	NM_198056	Pa07C	g.chr3:38567493G>A	c.5374G>A	p.D1792N	Missense		
SGEF	NM_015595	Pa07C	g.chr3:155324933G>T	c.1120G>T	p.E374X	Nonsense		
SLC2A5	CCDS99.1	Pa07C	g.chr1:9034296C>T	c.714C>T	p.R238R	Synonymous		
SLC6A2	CCDS10754.1	Pa07C	g.chr16:54285456dupT	c.952dupT	fs	INDEL		
SMAD4	CCDS11950.1	Pa07C	g.chr18:46857030C>T (homozygous)	c.1333C>T	p.R445X	Nonsense		Y
SMOC1	CCDS9798.1	Pa07C	g.chr14:69489919C>T	c.295C>T	p.R99W	Missense		
SNX26	CCDS12477.1	Pa07C	g.chr19:40968054C>T	c.1845C>T	p.T615T	Synonymous		
STAC	CCDS2662.1	Pa07C	g.chr3:36502658C>T	c.600C>T	p.Y200Y	Synonymous		
SYNE1	CCDS5236.1	Pa07C	g.chr6:152743723dupA	c.14208dupA	fs	INDEL		
TBX18	ENST00000330469	Pa07C	g.chr6:85529113C>T	c.110C>T	p.P37L	Missense		
TESSP2	NM_182702	Pa07C	g.chr3:46849470G>T	c.602G>T	p.W201L	Missense		
TIMELESS	CCDS8918.1	Pa07C	g.chr12:55113483T>G	c.378T>G	p.S126R	Missense		

TP53	CCDS11118.1	Pa07C	g.chr17:7519119A>G (homozygous)	c.536A>G	p.H179R	Missense		Y
WDR42B	ENST00000329763	Pa07C	g.chrX:27524862G>A (homozygous)	c.193G>A	p.D65N	Missense		
XPNPEP1	CCDS7560.1	Pa07C	g.chr10:111657509C>A	c.47C>A	p.A16D	Missense		
ZNF537	CCDS12421.1	Pa07C	g.chr19:36461398C>T	c.592C>T	p.R198W	Missense		
AKAP12	CCDS5229.1	Pa08C	g.chr6:151719101G>A	c.268G>A	p.G90S	Missense		
ATP2B3	CCDS14722.1	Pa08C	g.chrX:152336396C>T (homozygous)	c.1628C>T	p.T543M	Missense		
B3GALT1	CCDS2227.1	Pa08C	g.chr2:168551355C>T	c.299C>T	p.T100M	Missense		
BCL2A1	CCDS10312.1	Pa08C	g.chr15:78040564G>C (homozygous)	c.428G>C	p.G143A	Missense		
C13orf22	CCDS9336.1	Pa08C	g.chr13:30130832C>G	c.2618C>G	p.S873C	Missense		
C1RL	CCDS8573.1	Pa08C	g.chr12:7140801C>T (homozygous)	c.792C>T	p.I264I	Synonymous		
C20orf26	NM_015585	Pa08C	g.chr20:20180359G>C	c.2280G>C	p.E760D	Missense		
CCNB3	CCDS14331.1	Pa08C	g.chrX:49885977A>C	c.1772A>C	p.K591T	Missense		
CCNYL3	ENST00000332505	Pa08C	g.chr16:34132052T>G	c.248T>G	p.L83R	Missense		
CGN	CCDS999.1	Pa08C	g.chr1:148309864A>C	c.1358A>C	p.K453T	Missense		
CMYA5	NM_153610	Pa08C	g.chr5:79062356C>G	c.2012C>G	p.T671R	Missense		
CNGB3	CCDS6244.1	Pa08C	g.chr8:87821047A>C	c.163A>C	p.T55P	Missense		
CNTNAP4	CCDS10924.1	Pa08C	g.chr16:74946796G>A	c.202G>A	p.A68T	Missense		
CTNND2	CCDS3881.1	Pa08C	g.chr5:11212793C>T	c.2054C>T	p.A685V	Missense		
DICER1	CCDS9931.1	Pa08C	g.chr14:94668662C>G	c.250C>G	p.Q84E	Missense		
DNAH8	CCDS4838.1	Pa08C	g.chr6:38985291C>G	c.8882C>G	p.A2961G	Missense		
DOCK2	CCDS4371.1	Pa08C	g.chr5:169041455A>C	c.600A>C	p.E200D	Missense	Y	
EHMT1	CCDS7050.1	Pa08C	g.chr9:137948184C>G	c.1939C>G	p.P647A	Missense		
EML1	NM_004434	Pa08C	g.chr14:99475370G>A	c.2275G>A	p.D759N	Missense		
FAT4	CCDS3732.1	Pa08C	g.chr4:126770404G>A	c.9545G>A	p.G3182D	Missense		
FLJ10324	NM_018059	Pa08C	g.chr7:4690651G>A	c.361G>A	p.D121N	Missense		
FLJ21986	NM_024913	Pa08C	g.chr7:120362439A>G	c.1355A>G	p.N452S	Missense		
GABRA1	CCDS4357.1	Pa08C	g.chr5:161256762C>A	c.1127C>A	p.T376N	Missense		
GPC2	CCDS5689.1	Pa08C	g.chr7:99414178C>A	c.1045C>A	p.P349T	Missense		
GRM8	CCDS5794.1	Pa08C	g.chr7:125767758T>G	c.1629T>G	p.G543G	Synonymous		
HPCAL1	CCDS1671.1	Pa08C	g.chr2:10517488C>T	c.525C>T	p.S175S	Synonymous		
HSGT1	CCDS7321.1	Pa08C	g.chr10:74590241G>A	c.280G>A	p.V94I	Missense		
ITPR1	NM_002222	Pa08C	g.chr3:4658919G>A (homozygous)	c.509G>A	p.R170Q	Missense		
JPH4	CCDS9603.1	Pa08C	g.chr14:23110276dupG	c.1504dupG	fs	INDEL		
KAL1	CCDS14130.1	Pa08C	g.chrX:8374868A>C (homozygous)	c.484A>C	p.K162Q	Missense		

KIAA0082	CCDS4835.1	Pa08C	g.chr6:37537819T>C	c.1288T>C	p.F430L	Missense		
KIAA1957	ENST00000332235	Pa08C	g.chr19:358514G>A	c.1028G>A	p.R343Q	Missense		
KRAS	CCDS8702.1	Pa08C	g.chr12:25289551G>T	c.35G>T	p.G12V	Missense		Y
KRTAP11-1	CCDS13608.1	Pa08C	g.chr21:31175392C>G	c.323C>G	p.S108C	Missense		
LGR6	CCDS1424.1	Pa08C	g.chr1:199019324G>A	c.2080G>A	p.V694I	Missense		
LOC167127	CCDS3914.1	Pa08C	g.chr5:36085237G>A	c.354G>A	p.A118A	Synonymous		
LRRTM4	NM_024993	Pa08C	g.chr2:77657374A>T	c.1276A>T	p.I426F	Missense		
MLL2	NM_003482	Pa08C	g.chr12:47730476G>A	c.2337G>A	p.P779P	Synonymous		
MRGX1	CCDS7846.1	Pa08C	g.chr11:18912631A>C (homozygous)	c.277A>C	p.T93P	Missense		
MRGX1	CCDS7846.1	Pa08C	g.chr11:18912623T>C	c.285T>C	p.S95S	Synonymous		
MTR	CCDS1614.1	Pa08C	g.chr1:233363178G>A	c.2470G>A	p.A824T	Missense		
NCL	NM_005381	Pa08C	g.chr2:232146931A>C	c.1621A>C	p.N541H	Missense		
NEB	NM_004543	Pa08C	g.chr2:152252196C>G	c.12198C>G	p.D4066E	Missense		
NEO1	CCDS10247.1	Pa08C	g.chr15:71357558C>G (homozygous)	c.3227C>G	p.P1076R	Missense		
NLGN1	CCDS3222.1	Pa08C	g.chr3:175479470A>C	c.977A>C	p.D326A	Missense		
NOD3	NM_178844	Pa08C	g.chr16:3554351G>A	c.588G>A	p.S196S	Synonymous		
NP_001074311.1	ENST00000326096	Pa08C	g.chr6:29339624A>T	c.212A>T	p.Y71F	Missense		
ODZ4	ENST00000278550	Pa08C	g.chr11:78058287G>A	c.6757G>A	p.V2253M	Missense		
OR4A16	NM_001005274	Pa08C	g.chr11:54867358A>G	c.106A>G	p.T36A	Missense		
OR5J2	NM_001005492	Pa08C	g.chr11:55701580C>A	c.911C>A	p.A304D	Missense		
OR8H1	NM_001005199	Pa08C	g.chr11:55814357C>A	c.758C>A	p.T253N	Missense		
OVCH1	NM_183378	Pa08C	g.chr12:29471803G>T	c.3373G>T	p.V1125L	Missense		Y
PALMD	CCDS758.1	Pa08C	g.chr1:99866975G>A (homozygous)	c.1138G>A	p.E380K	Missense		
PNPLA1	NM_001039725	Pa08C	g.chr6:36383451C>G	c.1582C>G	p.P528A	Missense		
RAD9B	CCDS9148.1	Pa08C	g.chr12:109431108C>G	c.710C>G	p.P237R	Missense		
RDH8	CCDS12223.1	Pa08C	g.chr19:9985233A>C	c.60A>C	p.E20D	Missense		
SESN2	CCDS321.1	Pa08C	g.chr1:28283231C>T (homozygous)	IVS4-3C>T	Splice Site	Splice Site		
SH3GL3	CCDS10325.1	Pa08C	g.chr15:82024888A>C (homozygous)	IVS2-2A>C	Splice Site	Splice Site		
SLC25A26	CCDS2905.1	Pa08C	g.chr3:66503677C>G (homozygous)	c.404C>G	p.S135C	Missense		
SLC2A3	CCDS8586.1	Pa08C	g.chr12:7976860C>T	c.259C>T	p.R87C	Missense		
SMAD4	CCDS11950.1	Pa08C	g.chr18:46838724delG (homozygous)	c.804delG	fs	INDEL		Y
SNX16	CCDS6234.1	Pa08C	g.chr8:82914771A>C	c.6A>C	p.A2A	Synonymous		
SYNE1	CCDS5237.1	Pa08C	g.chr6:152568219T>A	c.7623T>A	p.S2541S	Synonymous		
TBX6	CCDS10670.1	Pa08C	g.chr16:30007920C>T	c.466C>T	p.R156W	Missense		

TMEM132B	NM_052907	Pa08C	g.chr12:124663469A>C	c.2570A>C	p.K857T	Missense		
TP53	CCDS11118.1	Pa08C	g.chr17:7518972delT (homozygous)	c.602delT	fs	INDEL		Y
TTK	CCDS4993.1	Pa08C	g.chr6:80804442G>T	c.2089G>T	p.D697Y	Missense		
TTN	NM_133379	Pa08C	g.chr2:179435973G>A	c.16660G>A	p.V5554I	Missense		
TTN	NM_133378	Pa08C	g.chr2:179299693G>A	c.44146G>A	p.D14716N	Missense		
UBE2M	CCDS12987.1	Pa08C	g.chr19:63759348C>T	c.472C>T	p.R158W	Missense		
XKR4	NM_052898	Pa08C	g.chr8:56178292C>T	c.690C>T	p.G230G	Synonymous		
ZNF253	ENST00000327867	Pa08C	g.chr19:23110681C>A	c.219C>A	p.P73P	Synonymous		

<sup>a</sup> Data from cited reference 5 (Jones et al. Core Signaling Pathways in Human Pancreatic Cancers Revealed by Global Genomic Analyses. Science 2008 321:5897).

<sup>\*</sup> Genomic positions are coordinates in the May 2004, hg17 35.1 UCSC Santa Cruz release of the human genome. Genomic coordinates and sequences of mutations are on the coding strand. All changes are heterozygous unless marked as homozygous. g., genomic sequence; c., cDNA sequence; p., protein sequence; del, deletion; dup, duplication; ins, insertion.

<sup>†</sup> Mutations in intronic sequences are annotated by intron number preceded by "IVS", with positive numbers starting from the G of the GT splice donor site and negative numbers starting from the G of the AG splice acceptor site. Mutations in untranslated regions by "UTR", with positive numbers starting at the first nucleotide after the translation start and negative numbers starting at the first nucleotide before the translation initiation.

<sup>‡</sup> fs, frameshift mutation; UTR, mutation in 5' or 3' untranslated region; indel, in frame insertion, deletion or duplication change affecting more than a single codon. The amino acid change resulting from mutations in the translation initiating methionine are indicated as "unknown".

**Supplementary Table 3.** Allelic Status of Index Lesions based on Illumina 1M SNP Arrays

Chr	Pa01	Pa02	Pa03	Pa04	Pa05	Pa07	Pa08
1	2 copies 0-24Mb; LOH 24-27Mb; 2 copies 27-30Mb; >2 copies 30-120Mb; 2 copies 142Mb-246Mb	2 copies 0-145Mb; LOH 145-147Mb; 2 copies 147-246Mb	>2 copies	>2 copies 0Mb-228Mb, >>2 copies 228-247Mb	>2 copies 0-48Mb; LOH 48-52Mb; >2 copies 52-246Mb	> 2 copies whole chromosome	0Mb-20Mb >2 copies, 20Mb-118Mb LOH, 118Mb-190Mb >2 copies, 190Mb-247Mb >>2 copies
2	whole chromosome >2 copies; very small region ~142Mb >3 copies	2 copies 0-12Mb; LOH 12-42Mb; 2 copies 42-242Mb	>2 copies	LOH 0-15Mb, 2 copies 15-90Mb, >2 copies 90Mb-245Mb	LOH 0-35Mb; 2 copies 35-76Mb; LOH 76-78Mb; >2 copies 78-242Mb	2 copies whole chromosome	2 copies
3	LOH 0-91Mb; 2 copies 95-162Mb; LOH 162-165Mb; 2 copies 165-200Mb	LOH 0-90Mb; >2 copies 90-198Mb	whole chromosome LOH	p arm LOH, q arm >2 copies	>2 copies 0-4Mb; LOH 4-7Mb; >2 copies 7-17Mb; LOH 17-69Mb; >2 copies 69-198Mb	>2 copies 0-168Mb; 2 copies 168-187Mb; >2 copies 187-198Mb	p arm LOH, q arm >2 copies
4	whole chromosome 2 copies	whole chromosome LOH	>2 copies	p arm >2 copies, q arm LOH	>2 copies 0-173Mb; LOH 173-175Mb; >2copies 175-191Mb	>2 copies whole chromosome	2 copies
5	>>2 copies p arm; >2 copies q arm (it looks like p arm might be AAABB and q arm is AAB)	>2 copies 0-68Mb; LOH 68-70Mb; >2 copies 70-180Mb	p arm >2 copies, q arm LOH	p arm 2 copies, q arm LOH	>2 copies 0-122Mb; 2 copies 122-125Mb; >2 copies 125-128Mb; LOH 128-132Mb; >2 copies 132-151Mb; 2 copies 151-153Mb; >2 copies 153-180Mb	>2 copies 0-172Mb; 2 copies 172-181Mb	2 copies, small region around 55Mb >2 copies

6	LOH 0-23Mb, >2 copies 24-65Mb; 2 copies 65-149Mb; LOH 149-151Mb; 2 copies 151-152Mb; LOH 152-153Mb; 2 copies 153-171Mb	>2 copies 0-10Mb; >2 copies/complex 10-48Mb; 2 copies 48-57Mb; >2 copies 57-67Mb; 2 copies 67-170Mb	2 copies	>2 copies	LOH 0-29Mb; 2 copies 29-52Mb; LOH 52-92Mb; 2 copies 92-93Mb; >2 copies 93-94Mb; LOH 94-99Mb; 2 copies 99-102Mb; LOH 102-105Mb; 2 copies 105-106Mb; LOH 106-108Mb; 2 copies 108-112Mb; LOH 112-138Mb; 2 copies 138-141Mb; LOH 141-171Mb	>2 copies whole chromosome; very small amplified region ~163Mb (likely the whole chromosome is AAB and the amplified region is AABB)	0Mb-26Mb LOH, 26Mb-27Mb >2 copies, 27Mb-29Mb LOH, 29Mb-171Mb >2 copies
7	>2 copies p arm; 2 copies q arm	>2 copies whole chromosome	>2 copies	p arm 2 copies, q arm > 2 copies	2 copies 0-7Mb; LOH 7-10Mb; 2 copies 10-149Mb; LOH 149-150Mb; 2 copies 150-158Mb	> 2 copies whole chromosome	2 copies
8	whole chromosome LOH	LOH 0-33Mb; >2 copies 33-146Mb	>2 copies	LOH 0-128Mb, >2 copies 128-129Mb, LOH 129-146Mb	>2 copies whole chromosome; 88-146Mb appears to be at very high copy (genotype scores approach homozygous, indicative of a AAAAAAB genotype)	>2 copies whole chromosome; >2 copies, but very complex fragmenting 36-146Mb	p arm LOH, q arm 2 copies
9	whole chromosome LOH	whole chromosome LOH	whole chromosome LOH	whole chromosome LOH	LOH whole chromosome	>2 copies 0-14Mb; LOH 14-19Mb; >2 copies 19-27Mb; LOH 27-	whole chromosome LOH, except for small regions of >2 copies at

						33Mb; >2 copies 33-140Mb	35Mb, 37Mb and 136-140Mb
10	LOH 0-30Mb; >2 copies 30-135Mb	>2 copies whole chromosome	whole chromosome LOH	LOH 0-40Mb, >2 copies 40-122Mb, LOH 122-126Mb, >2 copies 126-135Mb	LOH 0-118Mb; 2 copies 118-135Mb	>2 copies whole chromosome	2 copies
11	>2 copies whole chromosome	>2 copies 0-97Mb; 2 copies 97-98Mb; LOH 98-134Mb	LOH 0-20Mb, >2 copies 20-30Mb, 2 copies 30-135Mb	p arm LOH, q arm 2 copies	2 copies 0-2Mb; >2 copies 2-5Mb; 2 copies 5-127Mb; >2 copies 127-134Mb	>2 copies whole chromosome	0Mb-28Mb LOH, 28Mb-38Mb >2 copies, 38Mb-48Mb LOH, 48Mb-135Mb 2 copies
12	>2 copies 0-33Mb; LOH 33-51Mb; >2 copies 51-56Mb; LOH 56-107Mb; >2 copies 107-133Mb	LOH 0-25Mb; >2 copies 25-132mb	2 copies 0-18Mb, >2 copies 18-35Mb, LOH 35-63Mb, 2 copies 63-134Mb	whole chromosome LOH	>2 copies whole chromosome; noisy, but small regions of LOH at ~88Mb, 99Mb, 104Mb and 117Mb	>2 copies whole chromosome; p arm (0-35Mb) is complex and appears to be at very high copy (genotype scores approach homozygous, indicative of a AAAAAAB genotype)	0Mb-23Mb LOH, 23Mb-32Mb complex >2 copies, 32Mb-36Mb LOH, 36Mb-82Mb 2 copies, 82Mb-86Mb LOH, 86Mb-132Mb 2 copies
13	2 copies 17-62Mb; LOH 62-64Mb; 2 copies 64-94Mb; LOH 94-96Mb; 2 copies 96-114mb	whole chromosome >2 copies	whole chromosome LOH	2 copies	2 copies 0-79Mb; >2 copies 79-80Mb; 20 copies 80-91Mb; LOH 91-93Mb; 2 copies 93-114Mb	>2 copies whole chromosome	17Mb-21Mb LOH, 21Mb-114Mb 2 copies
14	2 copies 19-24Mb; >2 copies 24-33Mb; 2 copies 33-78Mb;	LOH 18-30Mb; >2 copies 30-35Mb, 2 copies 35-39Mb; >2 copies	>2 copies	>2 copies	2 copies whole chromosome	2 copies whole chromosome	2 copies

	>2 copies 78-81Mb; 2 copies 81-101Mb; >2 copies 101-107Mb	39-50Mb; LOH 50-106mb					
<b>15</b>	>2 copies 18-26Mb; 2 copies 26-86Mb; >2 copies 86-100Mb	whole chromosome LOH	whole chromosome LOH	>2 copies 18-24Mb, 2 copies 24-100Mb	2 copies 18-33Mb, LOH 33-40Mb; 2 copies 40-59Mb; LOH 59-67Mb; 2 copies 67-100Mb	>2 copies whole chromosome	LOH whole chromosome
<b>16</b>	>2 copies whole chromosome; q arm is complex	whole chromosome LOH	>2 copies	>2 copies 0-2Mb, 2 copies 2-3Mb, >2 copies 3-35Mb, 2 copies 35-87Mb	>2 copies whole chromosome	>2 copies whole chromosome	0Mb-75Mb 2 copies, 75Mb-78Mb complex >2 copies, 78Mb-88Mb LOH
<b>17</b>	LOH 0-23Mb; >2 copies 23-78Mb	LOH 0-22Mb; >2 copies 22-53Mb; LOH 53-72Mb; >2 copies 72-78Mb	LOH 0-19Mb, 2 copies 19-21Mb, >2 copies 21-78Mb	LOH 0-16Mb, >2 copies/complex 16-22Mb, >2 copies 22-77Mb	LOH 0-14Mb; >2 copies/complex 14-22Mb; >2 copies 22-27Mb; 2 copies 27-79Mb	LOH 0-22Mb, >2 copies 22-79Mb	0Mb-30Mb LOH, 30Mb-46Mb 2 copies, 46Mb-78Mb LOH
<b>18</b>	>2 copies 0-17Mb; LOH 17-76Mb; very small region ~31Mb at >2 copies	2 copies 0-2Mb; LOH 2-3Mb; >2 copies 3-7Mb; 2 copies 7-24Mb; LOH 24-76Mb	whole chromosome LOH	p arm >>2 copies, q arm 2 copies	>2 copies/complex 0-21Mb; LOH 21-77Mb	LOH whole chromosome	LOH whole chromosome
<b>19</b>	2 copies 0-44Mb; >2 copies 44-45Mb; LOH 45-52Mb; >2 copies 52-64Mb	2 copies 0-10Mb; >2 copies/complex 10-16Mb; >2 copies 16-64Mb	LOH p arm, 2 copies q arm	>2 copies	>2 copies whole chromosome	>2 copies 0-45Mb; 2 copies 45-64Mb	>2 copies
<b>20</b>	2 copies 0-26mb; >2 copies 26-60Mb	LOH 0-18Mb; >2 copies 18-33Mb; LOH 33-35Mb; >2 copies 35-63Mb	2 copies	>2 copies 0-15Mb, 2 copies 15-16Mb, >2 copies 16-27Mb, 2 copies 27-62Mb	LOH 0-21Mb; 2 copies 21-22Mb; LOH 22-27Mb; 2 copies 27-63Mb	>2 copies 0-2Mb; LOH 2-13Mb; >2 copies 13-15Mb; LOH 15-18Mb; >2 copies 18-	0Mb-10Mb LOH, 10Mb-20Mb 2 copies, 20Mb-63Mb complex >2 copies

						22Mb; LOH 22-26Mb; >2 copies 26-63Mb	
<b>21</b>	>2 copies 0-16Mb; 2 copies 16-23Mb; >2 copies 23-24Mb; 2 copies 24-28Mb; >2 copies 28-37Mb; 2 copies 37-40Mb; >2 copies 40-47Mb	2 copies 13-23Mb; LOH 23-25Mb; 2 copies 25-47Mb	whole chromosome LOH	>2 copies 10-15Mb, 2 copies 15-47Mb	LOH whole chromosome	>2 copies whole chromosome	>2 copies
<b>22</b>	> 2 copies whole chromosome	LOH whole chromosome	>2 copies	LOH 14-15Mb, 2 copies 15-16Mb, LOH 16-18Mb, >2 copies/complex 18-28Mb, LOH 28-39Mb, complex 39-41Mb, LOH 41-48Mb	2 copies whole chromosome	>2 copies whole chromosome	whole chromosome LOH
<b>X</b>	LOH whole chromosome/patient is male	LOH whole chromosome/patient is male	LOH whole chromosome/patient is male	2 copies 0-2Mb, LOH 2-88Mb, 2 copies 88-93Mb, LOH 93-154Mb	LOH whole chromosome/patient is male	LOH whole chromosome/patient is male	LOH whole chromosome/patient is male

**Supplementary Table 4.** Ki67 labeling indices in normal and neoplastic pancreatic tissues.

Patient no.	Organ	Ki-67 IHC (raw data)			Ki-67 labeling index (%)
		Positive nuclei	Negative nuclei	Total nuclei	
Pa01	liver mets	247	2457	2704	10.05
Pa02	liver mets	442	2497	2939	15.04
Pa03	liver mets	450	1967	2417	18.62
Pa04	lung mets	601	2596	3197	18.8
Pa05	peritoneal mets	1352	2484	3836	35.25
Pa07	peritoneal mets	119	2804	2923	4.07
Pa08	lung mets	484	3311	3795	12.75
	<b>TOTALS</b>	3695	18116	21811	
	<b>AVERAGE</b>	527.8571429	2588	3115.857143	16.36857143
N1	normal pancreas	2	1866	1868	0.11
N2	normal pancreas	9	1312	1321	0.68
N3	normal pancreas	5	881	886	0.56
N4	normal pancreas	3	960	963	0.31
N5	normal pancreas	16	2108	2124	0.75
N6	normal pancreas	5	1485	1490	0.34
N7	normal pancreas	1	919	920	0.11
	<b>TOTALS</b>	41	9531	9572	
	<b>AVERAGE</b>	5.857142857	1361.571429	1367.428571	0.41

**Supplementary Table 5:** Estimates of Time in the Clonal Evolution of Metastatic Pancreatic Cancer.

Patient	$T_1$	$T_2$	$T_3$	Total tumor time	Patient age at tumor initiation
Pa01C	16.1 ± 2.5	9.8 ± 2.0	2.9 ± 1.2	28.8 ± 3.4	33.7 ± 3.4
Pa02C	10.6 ± 2.0	9.4 ± 1.9	2.7 ± 1.2	22.7 ± 3.0	34.8 ± 3.0
Pa03C	7.9 ± 1.8	2.4 ± 1.0	2.7 ± 1.2	13.0 ± 2.4	47.5 ± 2.4
Pa04C	11.4 ± 2.1	7.9 ± 1.8	2.7 ± 1.2	22.0 ± 3.0	37.5 ± 3.0
Pa05X	9.1 ± 1.9	4.3 ± 1.30	2.3 ± 1.2	15.7 ± 2.6	34.8 ± 2.6
Pa07C	15.7 ± 2.5	3.1 ± 1.1	2.7 ± 1.2	21.5 ± 3.0	60.0 ± 3.0
Pa08C	11.4 ± 2.1	10.6 ± 2.0	2.7 ± 1.2	24.7 ± 3.1	26.8 ± 3.1
<b><u>Average</u></b>	<b>11.7 ± 3.1</b>	<b>6.8 ± 3.4</b>	<b>2.7 ± 1.2</b>	<b>21.2 ± 4.8</b>	<b>39.3 ± 4.8</b>

**Supplementary Table 6:** Estimates of Number of Cell Divisions in Different Stages of Clonal Evolution of Metastatic Pancreatic Cancer

<b>Patient</b>	<b><i>Number of divisions during <math>T_1</math></i></b>	<b><i>Number of divisions during <math>T_2</math></i></b>
Pa01C	2557 ± 397	1556 ± 318
Pa02C	1683 ± 318	1493 ± 302
Pa03C	1255 ± 285	381 ± 159
Pa04C	1810 ± 333	1255 ± 286
Pa05X	1445 ± 302	683 ± 206
Pa07C	2493 ± 397	492 ± 175
Pa08C	1810 ± 333	1683 ± 318

its current level around 700 million years ago, coinciding with the cessation of snowball Earth conditions. This peak reflects a large increase in the marine phosphorus inventory that was sustained over tens of millions of years. Given that 'modern' marine phosphorus has a residence time of tens of thousands of years, it is hard to imagine ocean conditions in which sustained high levels of phosphorus were not driven by a step-change in the global phosphorus mass balance.

The thinking about the consequences of such high phosphorus levels then runs as follows. The result of the phosphorus-driven marine productivity was sustained algal blooms in the ocean, much like those found today in ponds and streams near areas of heavy fertilizer application. The death and settling of these blooms caused long-term, enhanced organic-carbon burial, which (via the mass-balance relationship between carbon and oxygen<sup>9</sup>) resulted in the addition of oxygen to the ocean-atmosphere system (Fig. 1). This increase in atmospheric oxygen controlled the evolutionary patterns of oxygen-dependent metazoans. Such a scenario provides a plausible link between the roles of snowball Earth glaciations and late Proterozoic oxygenation in leading to the explosion in metazoan diversity.

The implications of the new results<sup>4</sup> for understanding glacially induced phosphorus weathering on landscapes are also interesting. We are beginning to appreciate how glacial dynamics affects phosphorus weathering on land and transport to the oceans<sup>10</sup>. In the modern 'rooted' world, in which soil development is generally mediated by plants, most of the weathered phosphorus is mobilized and transported from landscapes in a narrow time window after a glacier retreats. Continental records<sup>11</sup> indicate that this large flux of phosphorus occurs in about 10,000 years in most landscapes (probably faster in low-relief/high-rainfall landscapes and slower in high-relief/low-rainfall landscapes). In a modern landscape with its considerable plant coverage, maintenance of a sustained increase in phosphorus delivery to the oceans would require periodic removal of weathered and phosphorus-depleted soils, and exposure of fresh material for further soil development.

In late Proterozoic time, however, the absence of land plants meant that there would have been little soil development to stabilize landscapes, or to convert mineral-based phosphorus forms on pristine mineral surfaces to the organically and oxide-bound phosphorus found in modern soils. Presumably, phosphorus stripping from rocks was much more extensive, as the landscapes themselves were much less stabilized because of the lack of flora. Thus, in this case, the present is not a key to the distant past. Without the systems to stabilize phosphorus, the phosphorus cycle in the late Proterozoic would have been

permanently revved up — that is, until rootedness came into play several hundred million years later. A test of this hypothesized mechanism of enhanced phosphorus stripping from landscapes would be to identify deltaic or other sedimentary-basin environments from the late Proterozoic, and to use proxy estimates of phosphorus loss (for example, the phosphorus/aluminium ratio) to determine whether values for this time interval are lower than expected given the source material. Such analyses would provide independent evidence of high phosphorus weathering rates.

Meanwhile, with this paper<sup>4</sup> and use of the phosphorus/iron proxy, there is now another way to look at nutrient variations in the ancient oceans. No proxy is perfect, however. In this instance, the weaknesses are both obvious (the limited spatial and temporal range of the proper rock types for analysis) and less obvious (variations in original iron-oxide composition that are now masked by mineral maturation). Nevertheless, thanks to Planavsky and colleagues<sup>4</sup>, we have a picture of the marine phosphorus cycle through deep time. We can

begin to develop informed hypotheses about how variations in the phosphorus cycle are driven, and what impact they have on the global carbon cycle, oxygen levels and the evolution of marine ecosystems. ■

**Gabriel M. Filippelli** is in the Department of Earth Sciences, Indiana University–Purdue University Indianapolis (IUPUI), Indianapolis, Indiana 46202, USA. e-mail: [gfilippe@iupui.edu](mailto:gfilippe@iupui.edu)

1. Catling, D. C., Glein, C. R., Zahnle, K. J. & McKay, C. P. *Astrobiology* **5**, 415–438 (2005).
2. Hoffman, P. F., Kaufmann, A. J., Halverson, G. P. & Schrag, D. P. *Science* **281**, 1342–1346 (1998).
3. Hoffman, P. F. & Schrag, D. P. *Terra Nova* **14**, 129–155 (2002).
4. Planavsky, N. J. *et al. Nature* **467**, 1088–1090 (2010).
5. Tyrell, T. *Nature* **400**, 525–531 (1999).
6. Mort, H. P. *et al. Geology* **35**, 483–486 (2007).
7. Paytan, A. & McLaughlin, K. *Chem. Rev.* **107**, 563–576 (2007).
8. Diaz, J. *et al. Science* **320**, 652–655 (2008).
9. Berner, R. A. *Nature* **426**, 323–326 (2003).
10. Föllmi, K. B., Hosein, R., Arn, K. & Steinmann, P. *Geochim. Cosmochim. Acta* **73**, 2252–2282 (2009).
11. Filippelli, G. M. *et al. Quat. Res.* **66**, 158–166 (2006).

## CANCER

## Genomic evolution of metastasis

**Prognosis for patients with pancreatic cancer is bleak, often owing to late diagnosis. The estimate that at least 15 years pass from tumour initiation to malignancy offers hope for early detection and prevention. [SEE LETTERS P.1109](#) & [P.1114](#)**

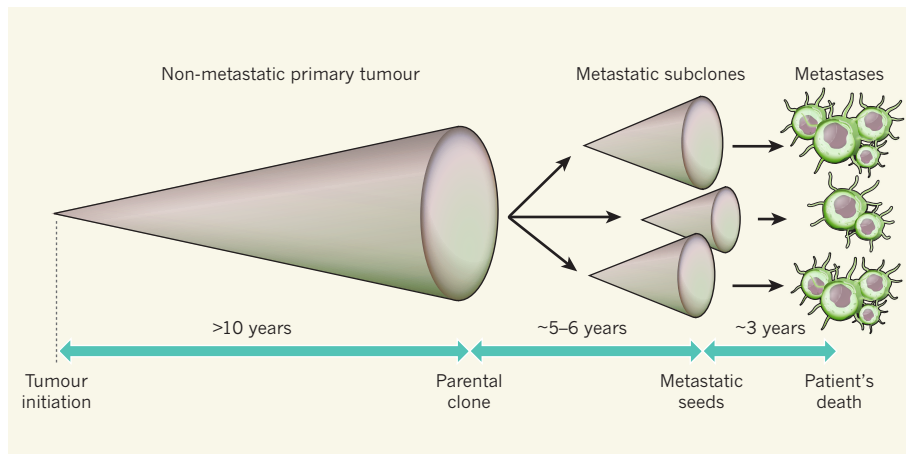
**E. GEORG LUEBECK**

**R**adiocarbon dating and comparative analyses of skeletal anatomy have informed the theory of human evolution; similarly, DNA sequencing of tumour cells coupled with dissection of the molecular anatomy of chromosomal aberrations is beginning to yield deeper insight into the evolution of cancer. In this issue, two papers<sup>1,2</sup> present findings from sequencing the protein-coding regions (exons) of more than 20,000 genes from the genomes of patients with metastatic (stage IV) pancreatic cancer. The findings are unprecedented, providing the first high-resolution image (at the level of single base pairs) of the non-germline mutational spectrum of pancreatic tumours and their metastatic descendants.

It has long been recognized that genomic instability is a hallmark of cancer. However, its significance in cancer progression has been the subject of debate for just as long. To shed light on this, Campbell *et al.*<sup>1</sup> (page 1109) performed a DNA-sequence-based study of chromosomal

rearrangements. They find that specific chromosomal rearrangements known as fold-back inversions occur in almost all of a patient's metastatic lesions. What's more, unlike other chromosomal rearrangements, which seem to occur either in the primary, parental tumour or in the metastatic lesion, Campbell *et al.* detect fold-back inversions in both primary and metastatic tumours. The authors therefore argue that fold-back inversions occur early in tumorigenesis and are probably a crucial driver of pancreatic-cancer progression.

The precise origin of fold-back inversions is unknown. It could be that DNA-replication-related erosion of the telomeres (the chromosome ends) — potentially because of suppressed or dysfunctional activity of the enzyme telomerase — triggers recurrent breakage–fusion–bridge (B/F/B) cycles<sup>3,4</sup>, which, in turn, cause progressive gains and losses of genetic material and so genomic instability. Intriguingly, telomerase activity seems to be restored in the invasive tumours, which might have a stabilizing effect on the abundance of B/F/B-induced rearrangements,



**Figure 1 | The pancreatic-cancer timeline.** Mathematical analyses of tumour-DNA sequence data from two collaborative studies<sup>1,2</sup> suggest that it probably takes more than 10 years from the initiation of a pancreatic tumour to the birth of the parental clone that results in pancreatic cancer. However, this clone does not have metastatic potential, and the subclones with the ability to spread to other tissues develop over an additional 5–6 years. The metastases, which are soon followed by the patient's death, occur over roughly the next 3 years.

but not on other rearrangements.

Campbell and colleagues' results affirm the presence of genomic instability in the development of pancreatic cancer. But because of extensive differences in the number, type and position of the rearrangements among patients — and even between the metastatic deposits in the same organ of a single patient — the functional consequences of this instability remain unclear. Studies using next-generation sequencing technologies on a larger number of patients are likely to fill in the missing pieces and pinpoint the driving forces in tumour progression and metastatic dissemination across different types of cancer.

In a separate study, Yachida *et al.*<sup>2</sup> (page 1114) address the clinically relevant issue of the timescales associated with tumour progression. These authors also carry out genomic sequencing of pancreatic-cancer metastases and examine their phylogenetic relationship with their respective, previously sequenced, primary tumours in seven patients. They thus derive estimates of three timescales: the time from tumour initiation to the birth of the founder cell of the parental (non-metastatic) clone; the sojourn time between the parental clone arising and its acquisition of metastatic potential; and the time from metastatic dissemination to the patient's death (Fig. 1).

Remarkably, the authors estimate that the time from tumour initiation to metastatic dissemination is at least a decade — a conclusion that suggests that there is a window of opportunity for medical intervention before the cancer spreads to distant organs. This finding is not inconsistent with that inferred from quantitative analyses of the age-specific incidence of pancreatic cancer in the general population<sup>5</sup>. On the basis of a general mathematical description that recognizes the random nature of both mutation accumulation and clonal expansion in pancreatic cancer, the

earlier, population-based analysis<sup>5</sup> estimated that the mean sojourn time from the tumour-initiating mutation to clinical diagnosis may be as much as five to six decades.

On the surface, the population-based estimate seems much longer than Yachida *et al.* conclude. It should be kept in mind, however, that the present sequence-based time estimates<sup>2</sup> are not general: they do not refer to the average of all pancreatic lesions with cancerous and metastatic potential in the tissue, but rather refer to the one lesion in the tissue which, by chance, leads to the first primary tumour in

#### STEM CELLS

## The intestinal-crypt casino

**Stem cells can renew themselves indefinitely — a feature that is often attributed to asymmetrical cell division. Fresh experimental and mathematical models of the intestine provide evidence that begs to differ.**

MICHAEL P. VERZI & RAMESH A. SHIVDASANI

Certain tissues, such as the skin, blood and intestinal lining, replenish millions of lost cells every day. The burden of renewal falls on small populations of stem cells, which can make exact copies of themselves, as well as generate all the resident cell types that differentiate and eventually die. This rare dual ability, a defining property of all stem cells, is exemplified by the asymmetrical division of germ cells<sup>1</sup> and neuronal precursors<sup>2</sup> in the fruitfly. Nonetheless, as long as the total stem-cell pool in a tissue remains roughly constant,

that tissue. Thus, Yachida and colleagues' estimate must be considered a lower bound for the mean sojourn time of pancreatic lesions, such as pancreatic intraepithelial neoplasia, that have the potential to cause invasive and metastatic cancer. From a clinical perspective, what matters is the prospective disease risk, which may involve multiple lesions individually evolving towards cancer. Thus, the time estimates of Yachida *et al.* are conservative and so clinically relevant.

These two studies<sup>1,2</sup> are a bellwether, and are among the first to explore the biological and clinical implications of sequence data for individual tumours. As the sequencing technology moves forward — and it does so at a blinding speed — more exciting details of the evolutionary processes involved in tumour progression are likely to be unearthed. It is to be hoped that such information will not only deepen our understanding of the cancer process, but also lead to new approaches to early cancer detection, better prognosis and, ultimately, prevention. ■

**E. Georg Luebeck** is at the Fred Hutchinson Cancer Research Center, Program in Computational Biology, Seattle, Washington 98109-1024, USA.  
e-mail: gluebeck@fhcrc.org

1. Campbell, P. J. *et al.* *Nature* **467**, 1109–1113 (2010).
2. Yachida, S. *et al.* *Nature* **467**, 1114–1117 (2010).
3. McClintock, B. *Genetics* **26**, 234–282 (1941).
4. Lo, A. W. I. *et al.* *Neoplasia* **4**, 531–538 (2002).
5. Meza, R., Jeon, J., Moolgavkar, S. H. & Luebeck, E. G. *Proc. Natl Acad. Sci. USA* **105**, 16284–16289 (2008).

in principle there is no reason why individual stem cells should not divide symmetrically, to generate either two identical stem cells or two daughters that exit the pool to differentiate. Indeed, two reports published in *Science*<sup>3</sup> and *Cell*<sup>4</sup> demonstrate that, in the normal course of tissue renewal, intestinal stem cells divide symmetrically.

In the small intestine, stem cells lie at visually identifiable positions within pocket-like crypts, and their progeny migrate in predictable streams (Fig. 1a, overleaf). In mice, two cell populations manifest the capacity for both prolonged self-renewal and multi-lineage