# Distant Metastasis Occurs Late during the Genetic Evolution of Pancreatic Cancer

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

[1]Department of Pathology, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins Medical Institutions, Baltimore Maryland 21231 USA

[2]Department of Oncology, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins Medical Institutions, Baltimore Maryland 21231 USA

[3]Department of Surgery, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins Medical Institutions, Baltimore Maryland 21231 USA

[4]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

[5]Program for Evolutionary Dynamics, Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138 USA

[6]School of Mathematics, University of Edinburgh, Edinburgh EH9-3JZ, UK

## Summary

Metastasis, the dissemination and growth of neoplastic cells in an organ distinct from that in which they originated [12], 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 chemo- or radiation therapy [3]. 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 [4].

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 [5]. In this previous study, a total of 426 somatic mutations in 388 different genes were identified among 220,884,033 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 tumors 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; 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,

though 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 [6]. Through combined analysis of high density 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 tumor suppressor genes such as *SMAD4* and *CDKN2A* and often occur in association with chromosomal instability [7]. In sum, the parental clones harbored 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, Supplementary Figs. 3 to 8). These maps revealed 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, i.e., 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 to 8).

To distinguish between the possibilities that clonal evolution occurred inside the primary cancer versus within secondary sites, we sectioned the primary tumors from two patients into numerous, three-dimensionally organized pieces (Fig. 2a,b) and examined the DNA 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 tumor, it was clear that subclones giving rise to each of these metastases were present in the primary tumor. 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 numbers of cells, could be identified through the analysis of other pieces of the same tumor. These subclones could be put into an ordered hierarchy establishing an evolutionary path for tumor progression (Fig. 2c). Analysis of multiple primary tumor pieces and metastatic lesions from patient Pa04 revealed a similar

clonal evolution, with distinct, large subclones within the primary tumors giving rise to the various metastases (Supplementary Fig. 8).

To further clarify 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 tumor used to define them (Fig. 2a, 2b). Samples representative of the parental clone were located throughout the primary carcinoma. By contrast, samples representing subclones were nonrandomly 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 dataset could also be used to infer the timing of the development of the various stages of pancreatic tumor progression [8]. We assume that the tumor 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 tumor. To estimate the timing, we first used Ki67 labeling 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. Ki67-positive nuclei constituted an average of 0.4% of normal ductal cells, while an average of 16.3% of cancer cells within the index metastasis lesions were Ki67-positive, consistent with prior estimates [9-10] (Supplementary Table 4). Based on these data plus that from sequencing of the index lesions, we derived estimates for three critical times in tumor evolution: $T_1$ - the time between tumor 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 tumor 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 tumor 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 Supplemental 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 tumor years before the metastases become clinically evident. The features of these metastatic subclones that promote metastasis formation have yet to be discerned, as no consistent genetic signature of metastatic subclones could be identified. We did identify several genes 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 play 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 metastagenic 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 [11], one possibility is that intratumoral 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 tumors 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 [12], offer hope for such non-invasive early detection.

## Methods Summary

Rapid autopsies were performed on seven individuals with Stage IV pancreatic cancer [13]. Genomic DNA was extracted from cell lines or xenografts established from one metastasis of each patient and used for exomic sequencing as previously described [5]. The Illumina Infinium II Whole Genome Genotyping Assay employing the BeadChip platform was also used to analyze each sample at 1,072,820 (1M) SNP loci as previously described [5]. Samples of snap frozen pancreatic cancer tissue were microdissected using a PALM MicroLaser System (Carl Zeiss MicroImaging, Oberkochen, Germany) and DNA extracted using QIAmp DNA Micro Kits (Qiagen, Valencia, CA, USA). 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, Piscataway, NJ, USA). Ki67 immunolabeling (Clone MIB-1, Dako Cytomation, Glostrup, Denmark) 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, Tucson, AZ, USA), and this information 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 Online Methods).

# Methods

## Patients and Tissue Samples

Tissue samples from 7 patients with pancreatic ductal adenocarcinoma were collected in association with the Gastrointestinal Cancer Rapid Medical Donation Program (GICRMDP). This program was approved by the Johns Hopkins institutional review board and deemed in accordance with the Health Insurance Portability and Accountability Act. Details of the program have been described in detail previously [13]. 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 Finetechnical, Tokyo, Japan) in liquid nitrogen and for snap-freezing in liquid nitrogen in 1.7 mL cryovials and storage at −80°C. Xenograft enriched or low passage cell lines were generated from the post mortem cancer tissues of these seven patients as previously described [13,14].

## Laser Capture Microdissection (LCM)

Frozen tissue sections of autopsy tissues were cut into 7 μm sections using a cryostat and embedded onto UV-treated PALM membrane slides (Carl Zeiss MicroImaging, Oberkochen, Germany) and the slides were stored immediately at −80 °C until subsequent fixation. Tissue sections that underwent LCM were defrosted, fixed in 100% methanol for 3 minutes, 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 in order to have 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 QIAmp DNA Micro Kit (Qiagen, Valencia, CA, USA) 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′-TGCTTTGAATGCGTCCCAGAG-3′ (reverse) was designed. The real-time PCR condition was 95°C for 10 minutes; 40 cycles of 94°C for 10 seconds, 58°C for 15 seconds and 70°C for 30 seconds. PCR was carried out using Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen, Carlsbad, CA, USA). To minimize sequencing bias from using low starting templates, only samples for which the measured concentration by LINE assay was 3.3 ng/ul (1000 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, Piscataway, NJ, USA), 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, Waltham, MA, USA) and diluted to 20 ng/μL 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 by Jones et al. [5]. 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 employing the BeadChip platform was used to analyze tumor samples at 1,072,820 (1M) SNP loci as previously described [5]. 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-color 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 $\log_2 (R_{experimental}/R_{training\ set})$. 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 [15]. 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 >2Mb with consecutive homozygous genotype calls (B Allele Frequency near 0 or 1). Smaller (<2Mb) 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 labeling 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 2 autopsied patients who died of causes other than pancreatic cancer and 5 patients who underwent distal pancreatectomy for a serous cystadenoma or a islet cell tumor at The Johns Hopkins Hospital. Paraffin blocks were cut into sections 4 μm thick for Ki-67 immunostaining with all staining processes from deparaffinization to counterstaining with hematoxylin being performed automatically with the Ventana Discovery staining system (Ventana Medical Systems, Tucson, AZ). An anti-human Ki-67 mouse monoclonal antibody (Clone MIB-1, Dako Cytomation, Glostrup, Denmark) was used. At least 12 randomly selected high-power fields containing a minimum of 2000 cells were evaluated for each case, and the labeling 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.

## Modeling Tumor 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 [5]. 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 [16].

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 tumor 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 tumor samples from one patient. Since we sequenced at least one sample from the primary tumor and at least three samples from different metastases from each patient, these specific N1 mutations had to be present in the founder cells of all three metastases and in cells in the primary tumor. Thus there was a cell in the tumor 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 tumor, these $N_1$ mutations are accumulated along the single lineage from the tumor 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). Since the number of base pairs sequenced in the study is $31.7 \cdot 10^6$, and the mutation rate per base pair per generation is estimated at $5 \cdot 10^{-10}$, $r = 31.7 \cdot 10^6 \cdot 5 \cdot 10^{-10} \approx 0.016$ per generation [8].

Using our measurements of Ki-67 labeling index of the 7 index lesions (average 16.3%), we were able to estimate the S-phase fraction of cells in the 7 index lesions (average LI=9.5%) [17]. Assuming a median value for the S-phase duration in human tissues and tumors, $T_s$, of 10 hours [18], and using the formula for the potential cell doubling time $T_{pot} = \lambda T_s / LI$, we get an estimate for $T_{pot}$ of 3.5 days. Here $\lambda$ is a correction factor for the non-linear age distribution of cells through the cell cycle, which was assumed to be 0.8 [19]. This estimate is consistent with the average cell doubling time in pancreatic cancer from Amikura et al. [20] of 2.3 days. We use this latter estimate in our analysis, as we believe it is more accurate for pancreatic cancer.

Amikura et al [20] also reported the median doubling time of pancreatic cancer metastases as 56 days. To estimate the age of the index metastasis, we used a two stage model. We estimated the tumor doubling time was equal to the cell doubling time ($T_{gen}$) until the tumor size reached 100 microns in diameter at which time angiogenesis is required [21]. Thereafter, we used the median doubling time described above.

Our model works very well for estimating the number of cell divisions between discrete events in tumor 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 $T_{gen}$ denote the average time between subsequent cell divisions in a cell lineage, we arrive at the expression for time $T_1$:

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

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

$T_2$ 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 tumor and cell doubling times, and the size of the index lesions at autopsy [20].

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

## REFERENCES

1. Fidler IJ. Critical Determinants of Metastasis. Seminars in Cancer Biology. 2002; 12(2):89. [PubMed: 12027580]

2. Nguyen DX, Bos PD, Massague J. Metastasis: From Dissemination to Organ-Specific Colonization. Nature reviews. 2009; 9(4):274.

3. Stathis A, Moore MJ. Advanced Pancreatic Carcinoma: Current Treatment and Future Challenges. Nat Rev Clin Oncol. 7(3):163. [PubMed: 20101258]

4. Yachida S, Iacobuzio-Donahue CA. The Pathology and Genetics of Metastatic Pancreatic Cancer. Archives of Pathology & Laboratory Medicine. 2009; 133(3):413. [PubMed: 19260747]

5. Jones S, Zhang X, Parsons DW, et al. Core Signaling Pathways in Human Pancreatic Cancers Revealed by Global Genomic Analyses. Science (New York, N.Y. 2008; 321(5897):1801.

6. Maitra A, Hruban RH. Pancreatic Cancer. Annual Review of Pathology. 2008; 3:157.

7. Lengauer C, Kinzler KW, Vogelstein B. Genetic Instabilities in Human Cancers. Nature. 1998; 396(6712):643. [PubMed: 9872311]

8. Jones S, Chen WD, Parmigiani G, et al. Comparative Lesion Sequencing Provides Insights into Tumor Evolution. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105(11):4283. [PubMed: 18337506]

9. Terada T, Ohta T, Kitamura Y, 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. 1998; 122(1):42. [PubMed: 9448015]

10. Hisa T, Ohkubo H, Shiozawa S, et al. Growth Process of Small Pancreatic Carcinoma: A Case Report with Imaging Observation for 22 Months. World J Gastroenterol. 2008; 14(12):1958. [PubMed: 18350642]

11. Olive KP, Jacobetz MA, Davidson CJ, et al. Inhibition of Hedgehog Signaling Enhances Delivery of Chemotherapy in a Mouse Model of Pancreatic Cancer. Science (New York, N.Y. 2009; 324(5933):1457.

12. Leary RJ, Kinde I, Diehl F, et al. Development of Personalized Tumor Biomarkers Using Massively Parallel Sequencing. Science Translational Medicine. 2(20):20ra14.

13. Embuscado EE, Laheru D, Ricci F, et al. Immortalizing the Complexity of Cancer Metastasis: Genetic Features of Lethal Metastatic Pancreatic Cancer Obtained from Rapid Autopsy. Cancer Biology & Therapy. 2005; 4(5):548. [PubMed: 15846069]

14. Fu B, Guo M, Wang S, 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 Biology & Therapy. 2007; 6(10):1546. [PubMed: 17912029]

15. Peiffer DA, Le JM, Steemers FJ, et al. High-Resolution Genomic Profiling of Chromosomal Aberrations Using Infinium Whole-Genome Genotyping. Genome Research. 2006; 16(9):1136. [PubMed: 16899659]

16. Bignell GR, Greenman CD, Davies H, et al. Signatures of Mutation and Selection in the Cancer Genome. Nature. 463(7283):893. [PubMed: 20164919]

17. Sasaki K, Matsumura K, Tsuji T, et al. Relationship between Labeling Indices of Ki-67 and Brdurd in Human Malignant Tumors. Cancer. 1988; 62(5):989. [PubMed: 3409179]

18. Rew DA, Wilson GD. Cell Production Rates in Human Tissues and Tumours and Their Significance. Part Ii: Clinical Data. Eur J Surg Oncol. 2000; 26(4):405. [PubMed: 10873364]

19. Steel, GG. 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. 1995; 17(2):139. [PubMed: 7622937]

21. Naumov GN, Bender E, Zurakowski D, et al. A Model of Human Tumor Dormancy: An Angiogenic Switch from the Nonangiogenic Phenotype. Journal of the National Cancer Institute. 2006; 98(5):316. [PubMed: 16507828]
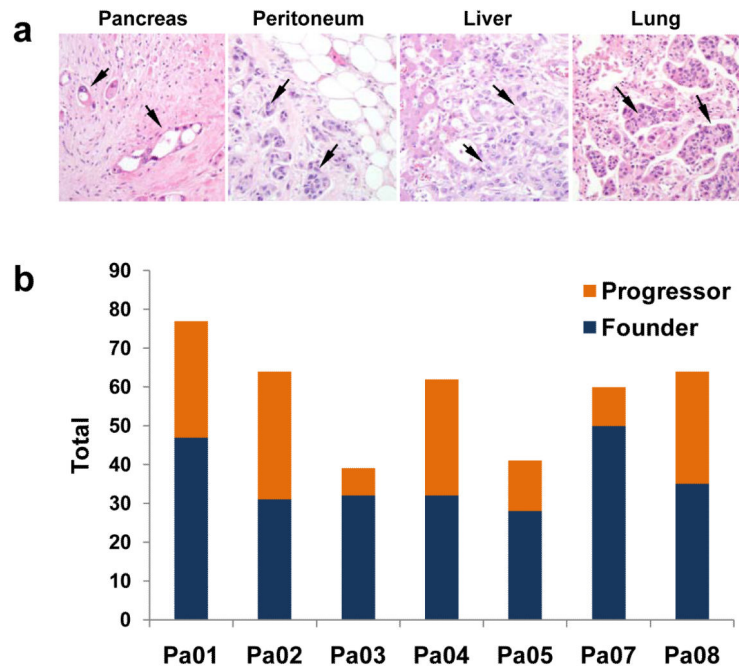
**Fig. 1. Summary of somatic mutations in metastatic pancreatic cancers**
a. Histopathology of primary infiltrating pancreatic cancer and metastatic pancreatic cancer to the peritoneum, liver and lung. In addition to infiltrating cancer cells in each lesion (arrows), non-neoplastic cell types are abundant. 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 analyzed were the most common category identified.
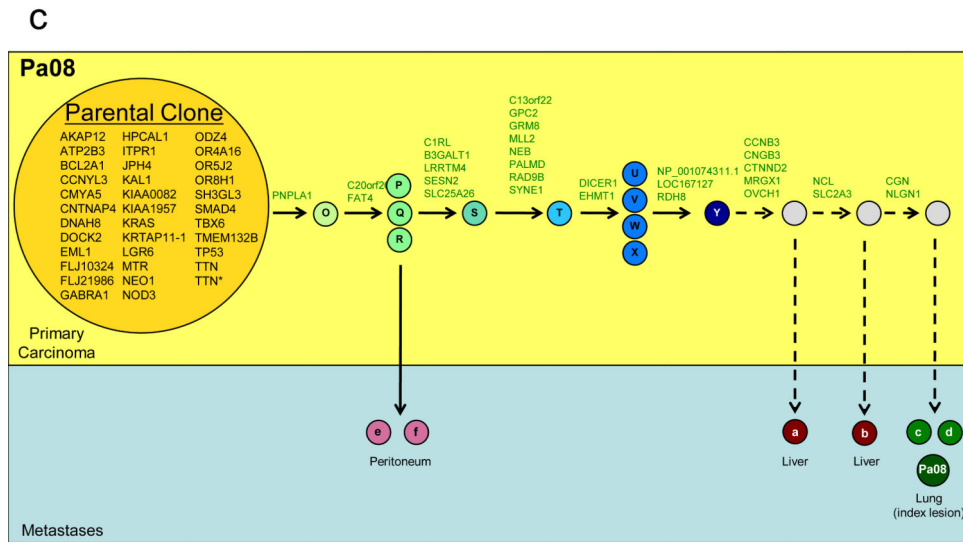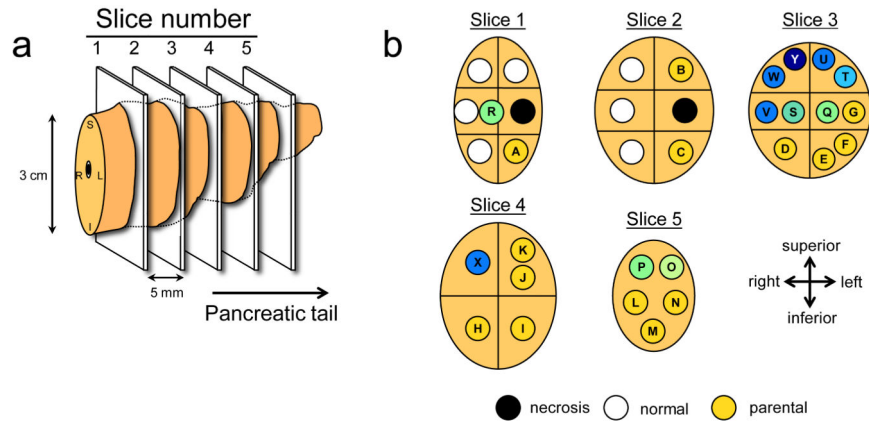
**Fig. 2. Geographic mapping of metastatic clones within the primary carcinoma and proposed clonal evolution of Pa08**

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 in a by comparative lesion sequencing within serial sections of the infiltrating pancreatic carcinoma. Metastatic subclones giving rise to liver and lung metastases are nonrandomly 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 by 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. * indicates two mutations were found in the *TTN* gene.
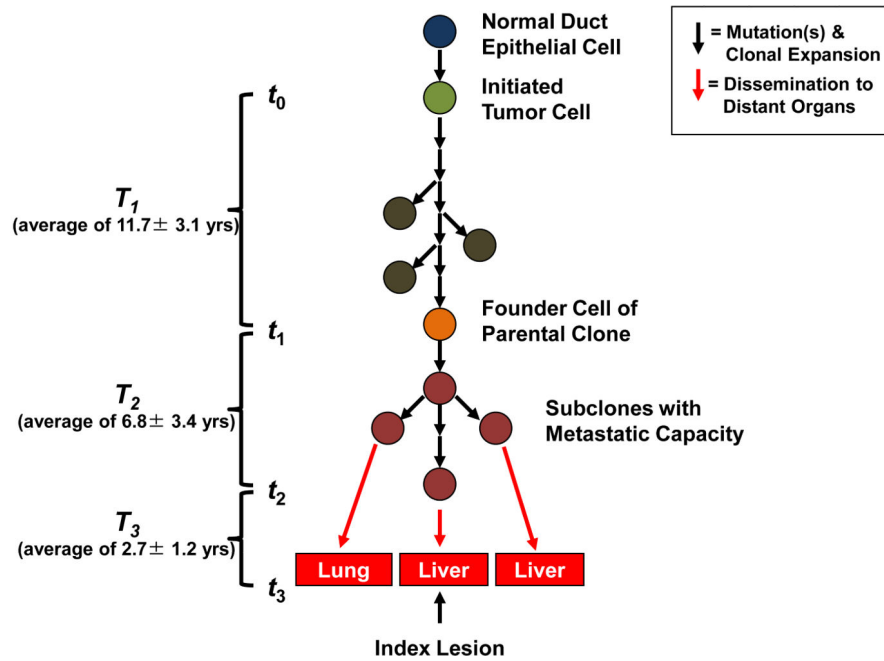
**Fig. 3. Schema of the genetic evolution of pancreatic cancer**

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).