

# Quantification of nucleic acid quality in postmortem tissues from a cancer research autopsy program

Jun Fan<sup>1</sup>, Raya Khanin<sup>2</sup>, Hitomi Sakamoto<sup>1</sup>, Yi Zhong<sup>1</sup>, Chelsea Michael<sup>3</sup>, Derwin Pena<sup>3</sup>, Breanna Javier<sup>1</sup>, Laura D. Wood<sup>6</sup>, Christine A. Iacobuzio-Donahue<sup>3,4,5</sup>

<sup>1</sup>Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

<sup>2</sup>Bioinformatics Core, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

<sup>3</sup>Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

<sup>4</sup>Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

<sup>5</sup>David M. Rubenstein Center for Pancreatic Cancer Research, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

<sup>6</sup>Department of Pathology, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins University School of Medicine, Baltimore, MD, 21231, USA

**Correspondence to:** Christine A. Iacobuzio-Donahue, **email:** iacobuzc@mskcc.org

**Keywords:** *autopsy, RNA, post-mortem, RNA sequencing, metastasis*

**Received:** July 22, 2016

**Accepted:** August 31, 2016

**Published:** September 02, 2016

## ABSTRACT

The last decade has seen a marked rise in the use of cancer tissues obtained from research autopsies. Such resources have been invaluable for studying cancer evolution or the mechanisms of therapeutic resistance to targeted therapies. Degradation of biomolecules is a potential challenge to usage of cancer tissues obtained in the post-mortem setting and remains incompletely studied. We analysed the nucleic acid quality in 371 different frozen tissue samples collected from 80 patients who underwent a research autopsy, including eight normal tissue types, primary and metastatic tumors. Our results indicate that RNA integrity number (RIN) of normal tissues decline with the elongation of post-mortem interval (PMI) in a tissue-type specific manner. Unlike normal tissues, the RNA quality of cancer tissues is highly variable with respect to post-mortem interval. The kinetics of DNA damage also has tissue type-specific features. Moreover, while DNA degradation is an indicator of low RNA quality, the converse is not true. Finally, we show that despite RIN values as low as 5.0, robust data can be obtained by RNA sequencing that reliably discriminates expression signatures.

## INTRODUCTION

Autopsy, derived from the Greek word *autopsia* meaning “to see for oneself”, is a method used since the 17th century to learn about disease and determine the cause of death [1]. Autopsy was a main form of understanding disease until the mid-20th century when medical imaging developed and allowed a view of the internal organs in a living patient [2]. In turn, the growth of laboratory medicine further diminished the need for autopsy as a diagnostic tool [3]. While postmortem exam has remained fundamental to improving knowledge of brain diseases, particularly neurodegenerative disorders, a renewed interest in its use for studying human cancer has gained traction in the past decade [4].

Sequencing of the human genome has led to a revolution in understanding of cancer etiology by revealing the genetic alterations characteristic of human tumors [5–7], the genetic features that underlie subtypes within a primary tumor type [8–11], or the mechanisms of therapeutic resistance [12–14]. With these advances has come a revival of interest in postmortem tissue collection because advanced stage disease is typically not accessible for study by next-generation methods in samples from living patients. As a result, research autopsy programs have emerged as a critical tool towards understanding the biology of lethal cancer and in many instances have led to significant insights into cancer progression and treatment resistance not possible with small tumor biopsies [14–16].

Despite the emergence of and implementation of research autopsy programs at a variety of institutions for obtaining cancer tissues, to date there are few formal evaluations of the quality of biomolecules in postmortem materials. A challenge to performing such evaluations is limited tissue resources. In some instances carefully screened cases and selected tissue types have been used to establish the relationship between potential quality-controlling factors and tissue sample quality [17–20]. Alternatively, simulated postmortem scenarios are used to mimic the postmortem interval and natural environment [21, 22]. Therefore, towards the goal of fully understanding these issues we leveraged our experience and resources amassed while running a cancer research autopsy program spanning more than a decade to determine the quality of nucleic acids in relation to tissue of origin, postmortem interval, normal versus neoplastic histology, primary versus metastases, and performance in downstream next generation sequencing methodologies.

## RESULTS AND DISCUSSION

### Sample set characteristics

Nucleic acid quality was analyzed in 371 different frozen tissue samples collected from 80 autopsied patients, 81% of which had been diagnosed with pancreatic cancer. The remaining patients had been diagnosed with breast cancer, lung cancer, colorectal cancer, germ cell tumor or melanoma (Figure 1). The postmortem interval (PMI) of these 80 autopsies ranged from 2 hours to > 36 hours. Cases with short PMIs were typically for those patients who expired while at the hospital, whereas cases with very long PMIs were a result of many factors including transport from outside the hospital or consent in the postmortem period by the patients' legally authorized representative to the program. Among the 371 tissue samples, 287 were histologically confirmed normal tissues sampled from the liver, lung, kidney, pancreas, spleen, heart, skeletal muscle and skin with a median of 35 normal tissues per site (range 30 to 49). We also collected 84 tumor samples of which 52 were from primary tumors, 16 from liver metastases and 16 from lung metastases including 10 patient-matched primary-metastatic pairs. To facilitate analyses, samples were arbitrarily categorized into four groups based on PMI: Category I, PMI ≤ 5 hours; Category II, PMI 6–10 hours; Category III, PMI 11–20 hours; and Category IV, ≥ 21 hours. When tumor samples were included there were a total of 98 samples from PMI Category I, 110 samples from PMI Category II, 88 samples from Category III, and 74 samples from Category IV (Table 1).

### RNA integrity in normal tissues is tissue-type specific

We first determined the extent to which RNA could be extracted from this large set of postmortem normal

tissues. RNA was successfully extracted from 269 of 287 (94%) samples attempted. The overall average RIN for all 269 samples was  $5.94 \pm 2.5$ , with a median RIN of 6.4. By contrast, for 10 of 287 samples (6%) the RNA yield was exceedingly low leading to unreported or unreliable RIN, and these samples were assigned an RIN value of 0 (Figure 1). In general, higher overall RNA yields were found from tissue samples collected within 1 year compared to those with long-term storage (> 5 years). This was unrelated to the number of freeze/thaw cycles per sample as with rare exceptions all frozen normals analyzed were previously unused and continuously stored at  $-80^{\circ}\text{C}$ . Furthermore, low RNA yields (defined as 20 ug/ml total RNA) were unrelated to PMI nor were they related to a specific histology.

We next calculated the mean RIN values for each individual tissue type for which RNA was obtained (Table 2). Overall, mean RIN values showed little variability among the eight normal tissue types and most tissue types had RIN values between 5 to 6.5. The tissue with the lowest mean RIN value was the kidney ( $4.63 \pm 1.95$ ) whereas the highest values were noted for skeletal muscle ( $9.01 \pm 1.36$ ), suggesting tissue-type specific differences in RNA stability in the postmortem interval. This pattern did not change when the median RIN value in each category was alternatively considered. To determine the extent to which RIN values of normal tissues show intra-patient variability we evaluated a subset of patients from each PMI category for which multiple normal tissues were evaluated (Figure 2). In all patients we noted variability in RIN values among different tissues, ranging from as low as an RIN values of < 2 to > 9 in a single individual (for example, patient RA15-11 in Category III or A164 in Category IV). However, the overall variability was less in PMI Category I samples than for PMI Category IV samples. Thus, RNA quality of one normal tissue retrieved postmortem is not a reliable predictor of RNA quality in a second tissue from that same patient, and good quality samples can be obtained despite the length of the postmortem interval.

We then determined the relationship of RIN values to PMI interval in greater detail by performing correlation analyses. Statistically significant negative correlations were noted for the liver, lung, kidney, pancreas, spleen and skin (Figure 3). Liver and skin showed particularly strong negative correlations between RIN value and PMI with *r* values close to  $-0.5$  and *p* values < 0.01. By contrast, no correlations were found for the heart or skeletal muscle with the RNA showing remarkable stability and quality in patients with PMIs as long as 36 hours or greater. No correlation was noticed between RIN and the length of sample storage in any of the tissue types examined (Figure 4). Taken together, these results indicate that the RNA quality of normal tissues declines with the elongation of PMI but not storage time, and the extent of degradation is tissue-type specific.

**Table 1: Sample cohort**

Postmortem Interval	Heart	Kidney	Liver	Lung	Pancreas	Skeletal Muscle	Skin	Spleen	Primary Tumor	Liver Metastases	Lung Metastases
Category I (1–5 h)	7	15	8	10	7	7	7	12	15	5	5
Category II (6–10 h)	7	13	11	11	15	8	8	10	18	5	5
Category III (11–20 h)	9	11	10	8	8	8	8	11	9	3	3
Category IV (> 21 h)	7	10	7	8	4	7	7	8	10	3	3
Total Samples	30	49	36	37	34	30	30	41	52	16	16

**Table 2: Mean and median RIN values in postmortem tissues<sup>a</sup>**

Category	Heart	Kidney	Liver	Lung	Pancreas	Skeletal muscle	Skin	Spleen	Primary	Liver Met	Lung Met
I	7.4 ± 1.9 (7.6)	5.7 ± 1.4 (6.2)	7.3 ± 2.4 (7.6)	7.1 ± 1.7 (6.8)	6.3 ± 2.2 (6.7)	9.3 ± 0.3 (9.3)	8.0 ± 0.5 (7.9)	6.1 ± 2.2 (6.7)	6.6 ± 2.0 (7.2)	6.4 ± 2.7 (6.6)	8.2 ± 1.9 (8.4)
II	6.8 ± 1.7 (6.3)	4.3 ± 2.5 (4.3)	6.2 ± 3.0 (6.7)	5.2 ± 2.2 (5.0)	5.9 ± 1.9 (6.5)	8.6 ± 2.6 (9.5)	6.0 ± 1.6 (6.6)	6.8 ± 2.0 (7.8)	4.5 ± 2.1 (4.5)	5.0 ± 2.9 (3.7)	6.1 ± 2.9 (6.3)
III	5.3 ± 2.4 (5.3)	4.3 ± 1.6 (4.3)	4.7 ± 2.7 (3.3)	6.0 ± 2.5 (7.1)	4.6 ± 2.7 (3.3)	9.2 ± 0.6 (9.4)	6.5 ± 2.2 (6.9)	3.5 ± 1.8 (2.7)	3.6 ± 2.2 (3.2)	3.5 ± 2.3 (2.2)	6.2 ± 3.0 (7.7)
IV	6.8 ± 2.7 (7.9)	3.8 ± 1.9 (2.9)	4.2 ± 3.3 (2.3)	4.3 ± 1.6 (3.7)	3.5 ± 2.8 (2.5)	9.0 ± 0.5 (8.8)	5.0 ± 1.7 (4.9)	4.3 ± 1.5 (4.4)	5.5 ± 2.8 (6.7)	4.7 ± 1.6 (4.4)	3.4 ± 1.2 (2.8)

<sup>a</sup>Mean values are expressed as mean ± SD. Median values are shown in brackets.

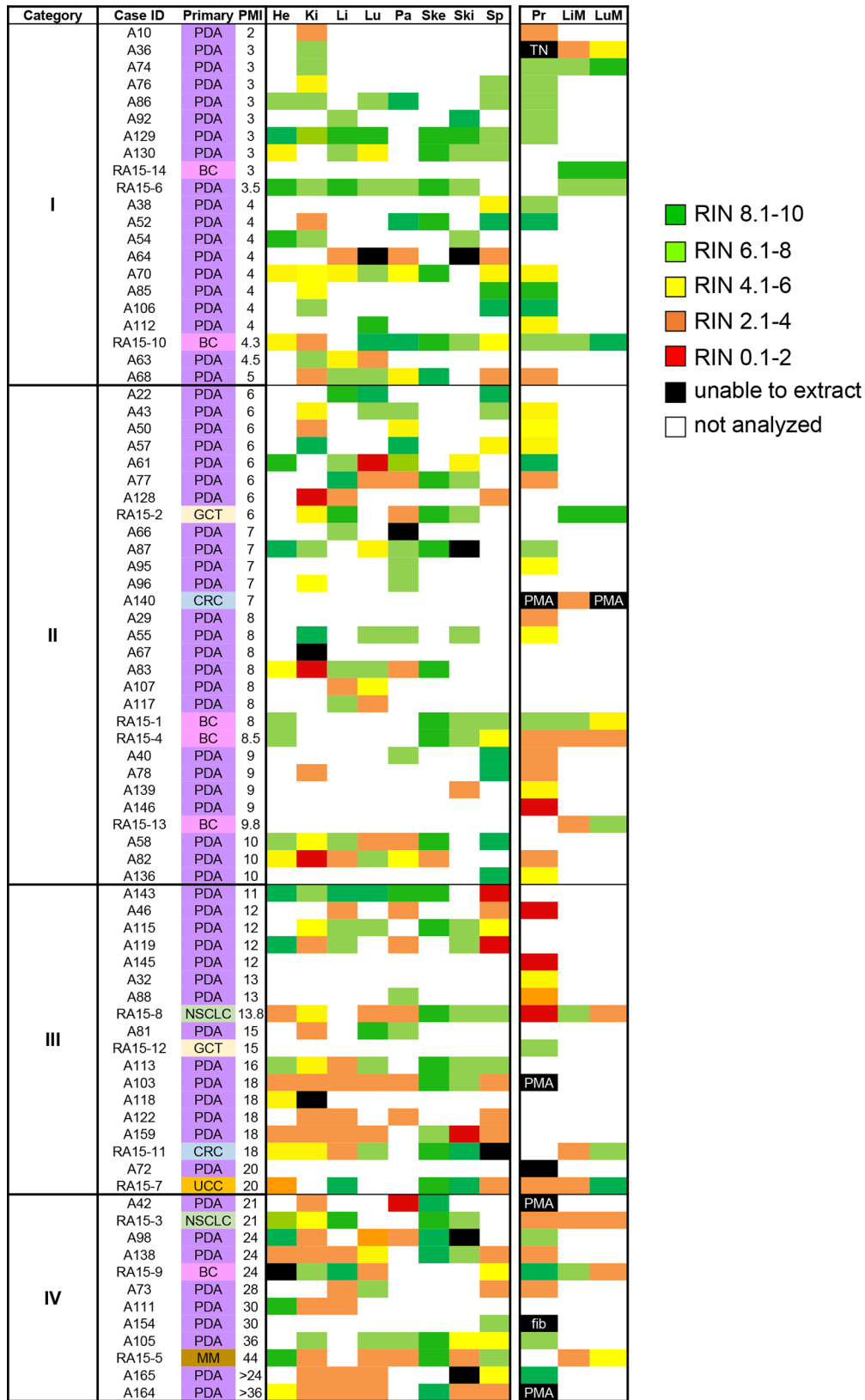
These results nonetheless demonstrate a clear negative correlation between RNA integrity and PMI in most tissue types indicating the importance of this variable. This correlation was established in the presence of several unavoidable confounding factors such as pyrexia, cachexia or prolonged hypoxia in the perimortem period, indirectly confirming that these factors may not be as influential as PMI in predicting RNA quality. We also observed a striking *lack* of correlation of PMI with RIN in normal skeletal muscle, and to a similar extent the heart, supporting the tissue type-specific nature of RNA degradation. In forensic settings, RNA has been shown to be stable in muscle up to 1 week after death [18]. While not addressed in this study, tissue-type specific RNA degradation has also been reported in ocular tissues with avascular structures having better RNA quality than vascularized structures such as the ciliary body [23]. Consistent with this notion, we found that normal kidney and liver, two highly vascularized organs, had among the lowest RIN values in each PMI Category. It may be reasonable to speculate that, when controlling for other factors, vascularized tissues are more sensitive to nutrient and oxygen deprivation resulting in a greater extent of sample degradation postmortem. However, given RNA decay is a precisely controlled process in living cells [24], such a process may also contribute to RNA quality in the postmortem period as suggested by Romero et al. [22].

### RNA integrity in cancer tissues

We next wondered if the integrity of ribonucleic acids in cancer tissues parallels that of normal tissues.

To address this question, we first analyzed RNA quality in 52 primary tumor tissues and 32 metastatic tumor tissues from the liver and lung. When stratified by PMI Category there were 15 primary tumors and 10 metastases in PMI category I, 18 primary tumors and 10 metastases in PMI category II, nine primary tumors and six metastases in PMI category III, and 10 primary tumors and six metastases in PMI category IV (Table 1). The mean RIN value in primary tumors was  $5.16 \pm 2.4$ , and for liver and lung metastases was  $5.07 \pm 2.51$  and  $6.29 \pm 2.72$ , respectively (Table 2). There was no correlation between RIN values and PMI in primary tumor tissues (Figure 3).

Forty-three of the 52 primary tumors (83%) analyzed were pancreatic ductal adenocarcinomas (PDA) (Figure 1), providing an opportunity to compare the RIN values in primary PDAs specifically to that of normal pancreatic tissues. The mean RNA integrity in PDA tissues was not significantly different from that of normal pancreas tissues when considering all samples (mean RIN  $5.4 \pm 2.4$  versus  $5.26 \pm 2.56$  respectively,  $p = \text{NS}$ ), or when limiting the comparison to 17 matched pairs of normal pancreas and primary PDA (mean  $6.16 \pm 1.96$  vs  $5.17 \pm 2.35$  respectively,  $p = \text{NS}$ ). This finding thus does not support the long-held “myth” that PDA tissues have worse quality than other tumor types. This may be partially explained by observations that PDA is characterized by a prominent desmoplastic/stromal reaction that is hypovascular compared to adjacent normal pancreas [25]. Nonetheless, we have found that screening multiple geographically distinct samples from different regions of the same neoplasm may be necessary to identify regions with preserved RNA quality, as we



**Figure 1: Heat map of RNA integrity numbers by tissue type.** RIN values from individual tissues were expressed with colored regions as defined by the accompanying legend. Abbreviations are PDA, pancreatic ductal adenocarcinoma; BC, breast cancer; GCT, germ cell tumor; CRC, colorectal cancer; NSCLC, non-small cell lung cancer; UCC, urothelial cell carcinoma; MM, melanoma; He, heart; Ki, kidney; Li, liver; Lu, Lung; Pa, pancreas; Ski, skin; Sp, spleen; Pr, primary tumor; LiM, liver metastasis; LuM, Lung metastasis. PMA, postmortem autolysis; TN, tumor-associated necrosis; fib, fibrosis.

have recently found in ongoing work in our laboratory (Figure 5). Finally, we next explored the relationships of metastatic tumor RNA quality between matched liver and lung metastases, i.e. from the same patient. There was no statistically significant correlation (Figure 6), indicating that RNA quality is highly variable among metastases, even within the same patient. Thus, unlike normal tissues that show fairly predictable and tissue-specific degradation in relation to PMI, cancer tissues derived from different organ sites appear less predictable with respect to RNA quality than that of normal tissues.

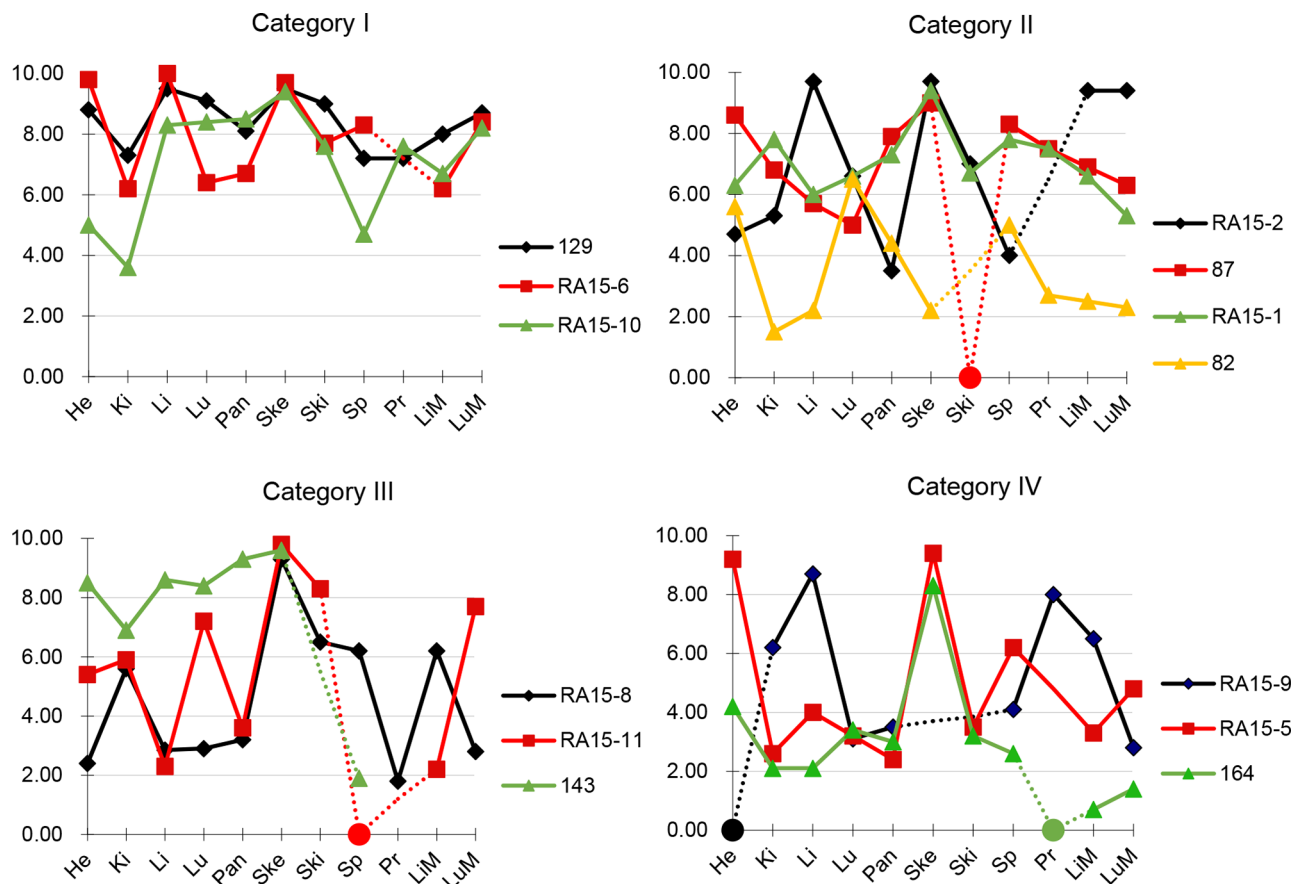
### DNA integrity in normal tissues is tissue-type specific

While generally more stable than RNA, DNA is also subject to degradation in the postmortem period [18]. Most methods to assess DNA degradation depend on examining selected target(s) to represent the overall sample quality with PCR-based methods among the most popular approach for this purpose [18, 26]. However, the extent to which such methods are reproducible or subject

to inter-experiment variation is unknown, as is the extent to which genomic DNA in postmortem tissues follows similar kinetics as RNA.

With these factors in mind we developed a semi-quantitative method to evaluate DNA quality to facilitate comparison among samples from independent experiments. In addition, unlike most studies detecting one locus, we simultaneously examined five chromosome loci of varying potential stability and susceptibility to DNA damage-inducing factors thereby achieving high sensitivity in detecting DNA damage in well-preserved samples (Table 3). DNA was extracted from 36 frozen autopsy samples that were collected from five patients in PMI Category I and five in PMI category IV (Table 4). To facilitate comparisons between RNA and DNA quality, samples with a wide spectrum of RIN values were selected from each PMI category, ranging from as low as 2.3 to as high as 9.4. These included samples from normal liver ( $n = 7$ ), normal kidney ( $n = 9$ ), primary tumors ( $n = 10$ ), liver metastases ( $n = 5$ ) and lung ( $n = 5$ ) metastases.

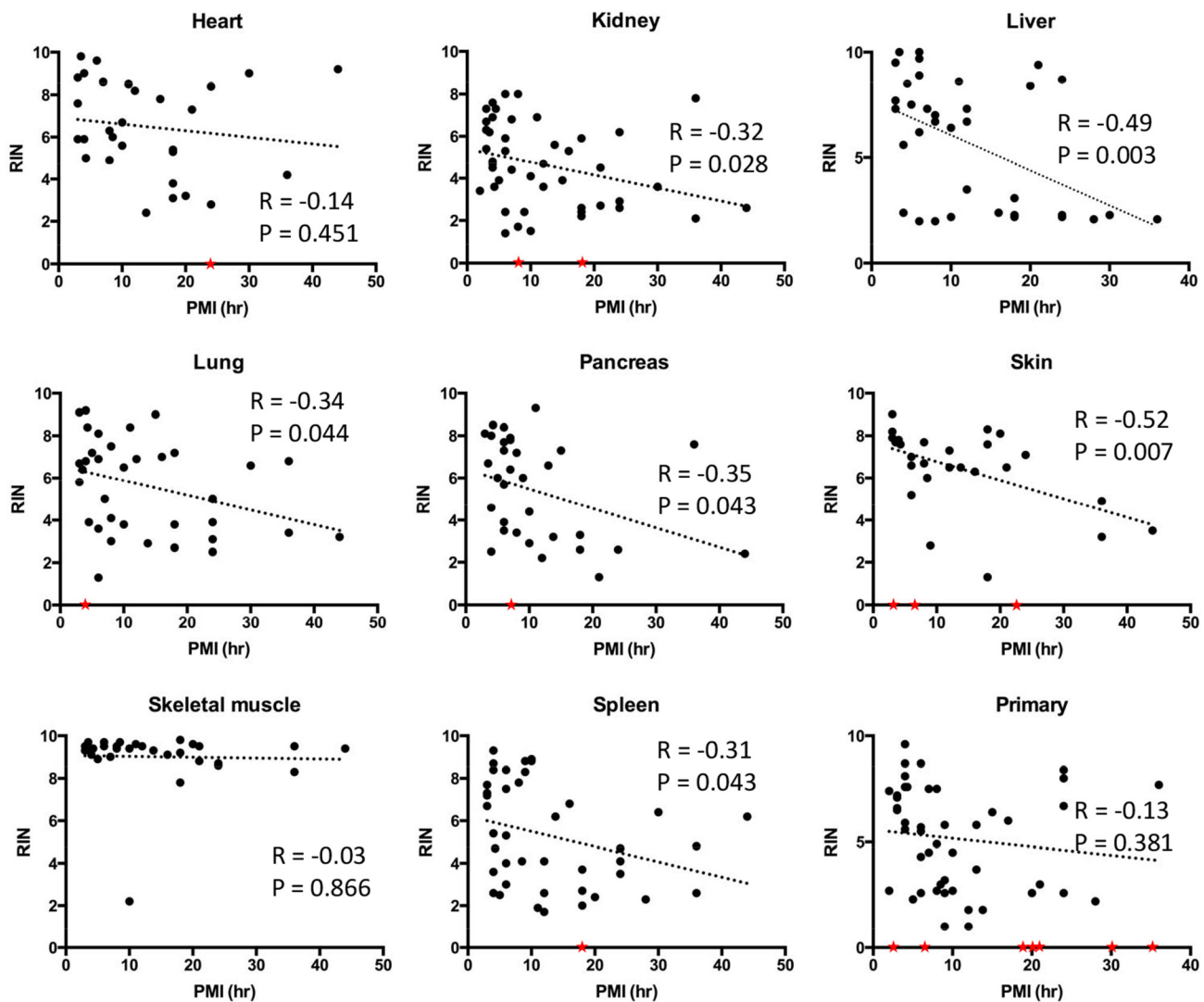
We successfully extracted genomic DNA from all 36 samples including one that failed in RNA extraction.



**Figure 2: RNA integrity numbers in patient-matched tissues.** RIN values from patient-matched tissues were plotted against tissue types. Selected cases from each of the four PMI categories were shown. Samples with low RNA yield leading to unreported or unreliable RIN were assigned to a RIN value of 0 and indicated by a circle. Dashed lines indicate tissues not analysed for that patient. Abbreviations are He, heart; Ki, kidney; Li, liver; Lu, Lung; Pa, pancreas; Ske, skeletal muscle; Ski, skin; Sp, spleen; Pr, primary tumor; LiM, liver mets; LuM, Lung mets.

**Table 3: Primers for DNA quality analysis and locus information**

Locus	UniSTS Number	Primers	Sequence	Size	Locus	Overlapped gene
Locus 1	SHGC105883	F R	5'-CCTGGCAAGTAATGGACAATGA-3' 5'-GCCTTTCCAGAGAACTGCAGAC-3'	980 bp	Chr13 q14.3	ATP7B
Locus 2	STB39J12.SP6	F R	5'-TTTCTAGAGCAGTGCAGAGTACTA GGAT-3' 5'-TCTTTCCCTCTACAACCCTCTAACC-3'	640 bp	Chr4 p15.33	/
Locus 3	STSG50529	F R	5'-TGAACAAGGGTTCAGGATG-3' 5'-GAGGTGGGCTTGACTTCGAG-3'	560 bp	Chr22 q13.32	/
Locus 4	SHGC147491	F R	5'-GGTAAACACACAATGGCCCAG-3' 5'-AAAAACGGAAGAAGTCTCTTGGC-3'	474 bp	Chr12 q13.13	/
Control Locus	CSNPHARP	F R	5'-CATGGCTCACTGGCTTACAA-3' 5'-TTGCCTCTTACAGAGGAGCAG-3'	196 bp	Chr2 q35	SMARCAL1



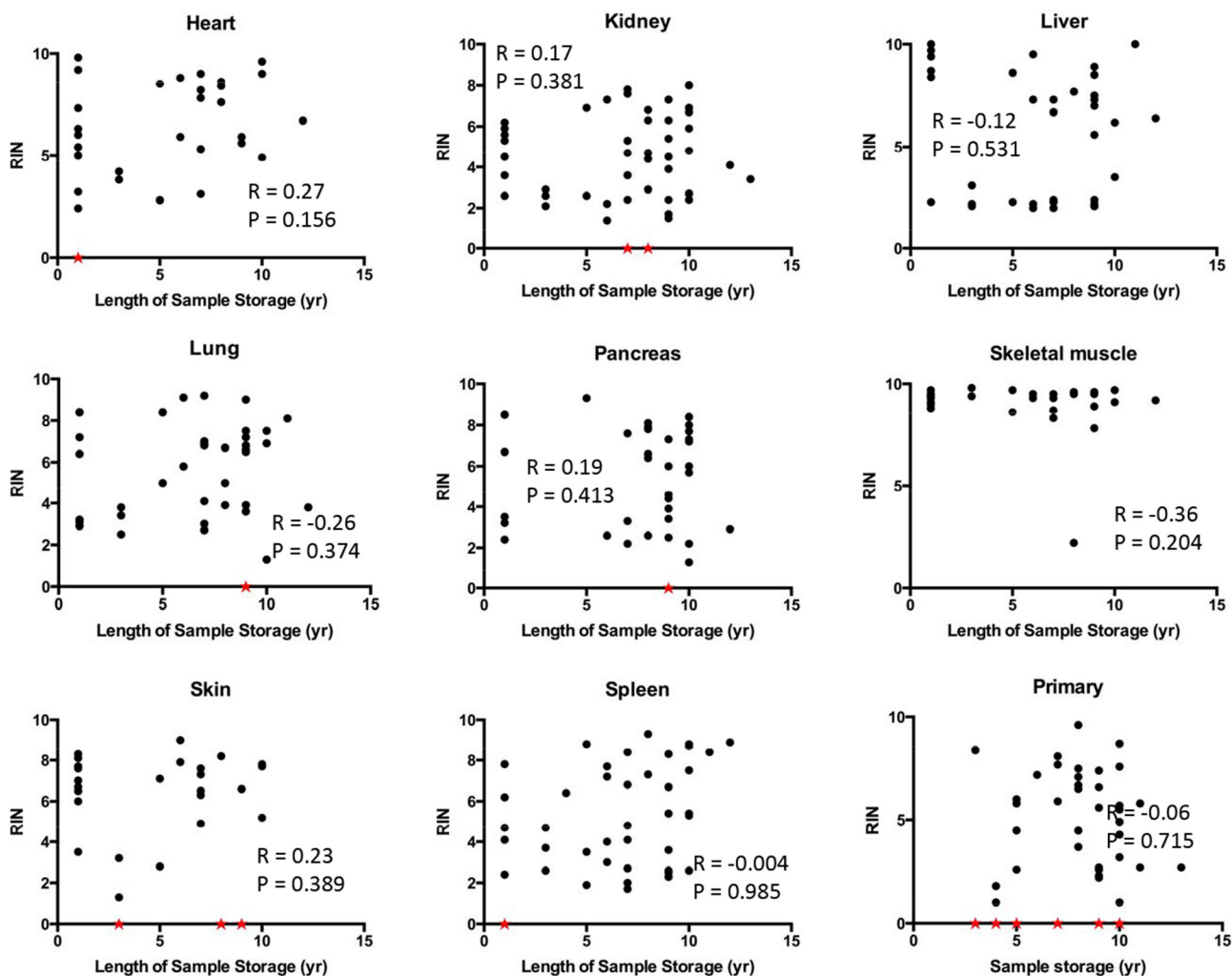
**Figure 3: Correlations between RNA integrity numbers and post mortem interval by tissue site.** Scatter plots were generated by plotting RIN values from each normal tissue type or primary tumors against PMI. Linear regression was performed to create curve fits. Samples with low RNA yield leading to unreported or unreliable RIN were assigned to a RIN value of 0 (red stars) and excluded from linear regression analysis.

No DNA damage was detected by our assay in 32/36 samples analyzed (89%) even though 17 of these 32 (53%) had RIN values less than 5. In the four samples with DNA damage three were from PMI category IV, two of which showed degradation at all four sensor loci (Figure 7). The remaining two samples showed only moderate damage as reflected by only two of the four chromosome loci affected. All four samples with DNA damage had concurrent RNA degradation (Table 4). Interestingly all four samples with DNA damage were from the liver, three of which were histologically normal, further supporting the greater susceptibility of vascularized tissues to postmortem degradation. All other samples except for these four patients did not show DNA damage, including all 10 primary tumors analyzed. Thus, while DNA damage in postmortem tissues may be an indicator of RNA quality the converse is not true. Moreover, in addition to PMI the kinetics of postmortem DNA degradation may also have tissue type-specific factors.

## RNA sequencing using postmortem RNA samples

Genomic DNA from postmortem tissues has been used successfully for next generation sequencing despite potential DNA degradation [14, 15, 27]. However, given that RNA is more sensitive than DNA during the postmortem interval, its utility in downstream sequencing applications is unknown. As a proof of principle study, we therefore performed RNA sequencing on five matched pairs of normal pancreas and pancreatic cancer tissues, all with an RIN value of 5 or greater.

Sequencing libraries were successfully generated from all samples using the poly-A enrichment method and used to generate 80 million reads per sample. Moreover, quality metrics of each library showed a sound distribution of coverage along transcripts and fragment lengths irrespective of RIN values (Figure 8A and 8B). One normal sample (patient A105), while showing

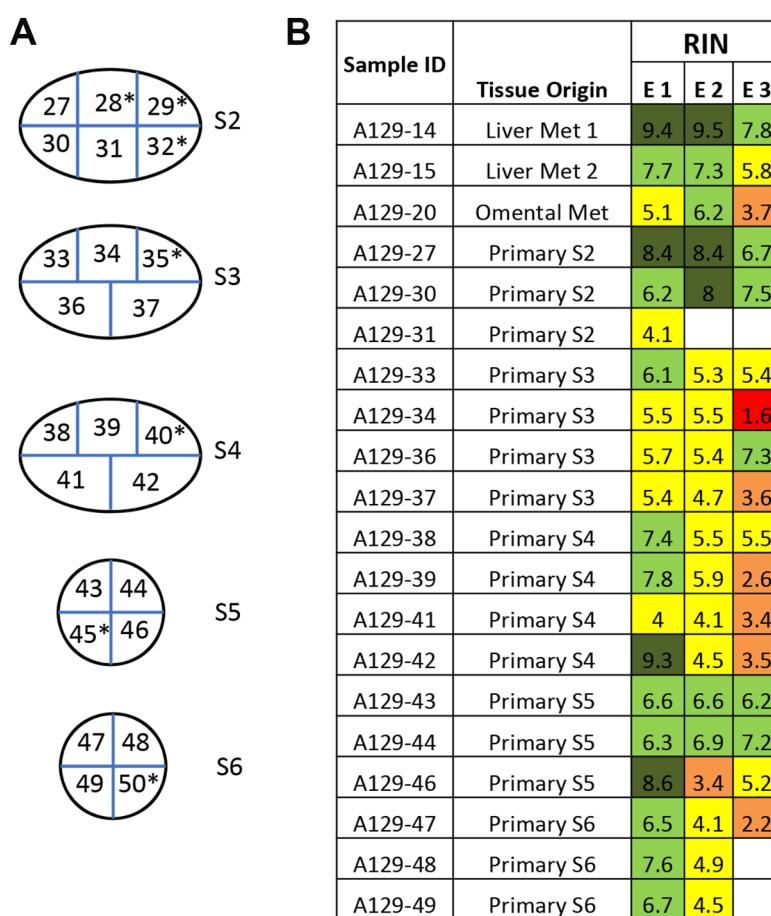


**Figure 4: Correlation between RIN and sample storage time.** Scatter plots were generated by plotting RIN values from each tissue types against sample storage time. Samples with low RNA yield leading to unreported or unreliable RIN were assigned to a RIN value of 0 and indicated by a red star.

**Table 4: Selected cases for DNA quality analyses<sup>a</sup>**

Case ID	PMI (hrs)	Normal Kidney	Normal Liver	Primary Tumor	Liver Metastasis	Lung Metastasis
A36	3	6.7		N/A	2.4	5.1
A74	3	6.3		7.4	6	9.3
A70	4	4.5	<b>5.6*</b>	5.6		
RA15-10	4.3	3.6		7.6	6.7	8.2
A68	5	3.9	7.5	2.3		
RA15-3	21	4.5	9.4	3	<b>4.4*</b>	2.7
A138	24	2.6	2.3	2.6		
RA15-9	24	6.2	8.7	8	6.5	2.8
A73	28		<b>2.1*</b>	2.2		
A165	> 24	2.1	<b>2.1*</b>	8.4		

<sup>a</sup>Numbers listed under each tissue type are the corresponding RIN value. Samples with DNA damage are in bold and marked with asterisks.



**Figure 5: Variation of RNA integrity in primary tumors and metastases in a metastatic cancer case A129.** (A) One primary pancreatic tumor was bread-loafed as indicated by primary tumor section S2 to S6. Each slice was further cut into equally sized small pieces (approximately 1 × 1 cm) and numbered accordingly to generate sample IDs. Samples with high tumor cellularity were used for RNA extraction and determination of RNA integrity. Samples indicated by an asterisk (\*) were excluded due to low tumor cellularity or necrosis. (B) Each piece of tissue was trisected and RNA extracted (E1,E2,E3). The color code for RIN value is the same as in Figure 1.



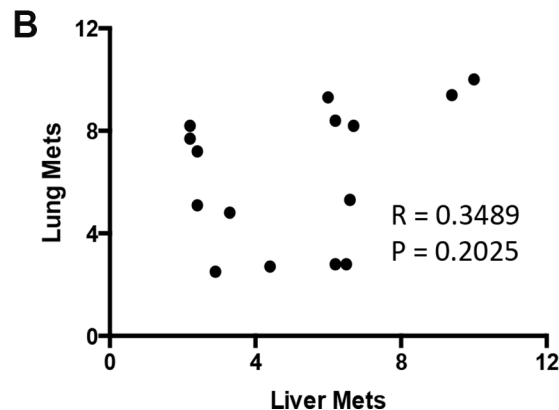
good quality sequencing data, was excluded because the resulting data indicated contamination by cancer cells. This was confirmed histologically. A heat map of the top 250 genes differentially expressed between five pancreatic cancers and the remaining four normal pancreata showed a pattern that clearly discriminated the two groups (Figure 8C). Genes transcripts overrepresented in normal tissues included PRSS1 (cationic trypsinogen), CPA2 (pancreatic specific carboxypeptidase), AMY2A (pancreatic amylase 2), and the pancreatic developmental transcription factor PTF1A consistent with the greater abundance of acinar cells or cells with stem-like properties within these samples [28–30]. Cancer samples showed greater heterogeneity with respect to the most

differentially expressed genes. Genes overrepresented in the cancer samples included PSCA [31], MMP3 [32], MMP11 [33] and SOX2 [34]. The small sample size otherwise precluded a more thorough classification of each carcinoma's subtype as recently described [11]. Finally, we leveraged our sequencing coverage to identify potential fusion events. We identified two recurrent fusion events, *AXGP1-GJC3* and *SIDT2-TAGLN*, in six of nine postmortem RNA samples (three normal and three tumor, including two normal-tumor matched pairs). These two fusions were recently reported in normal pancreatic tissues within the context of a larger pan-tissue analysis [35].

Collectively, these data are encouraging and suggest that despite being collected under postmortem conditions

**A**

Case ID	PMI (hrs)	Liver Mets	Lung Mets
A36	3	2.40	5.10
A74	3	6.00	9.30
RA15-14	3	10.00	10.00
RA15-6	3.5	6.20	8.40
RA15-10	4.3	6.70	8.20
RA15-2	6	9.40	9.40
A140	7	3.70	N/A
RA15-1	8	6.60	5.30
RA15-4	8.5	2.90	2.50
RA15-13	9.8	2.40	7.20
RA15-8	13.8	6.20	2.80
RA15-11	18	2.20	7.70
RA15-7	20	2.20	8.20
RA15-3	21	4.40	2.70
RA15-9	24	6.50	2.80
RA15-5	44	3.30	4.80

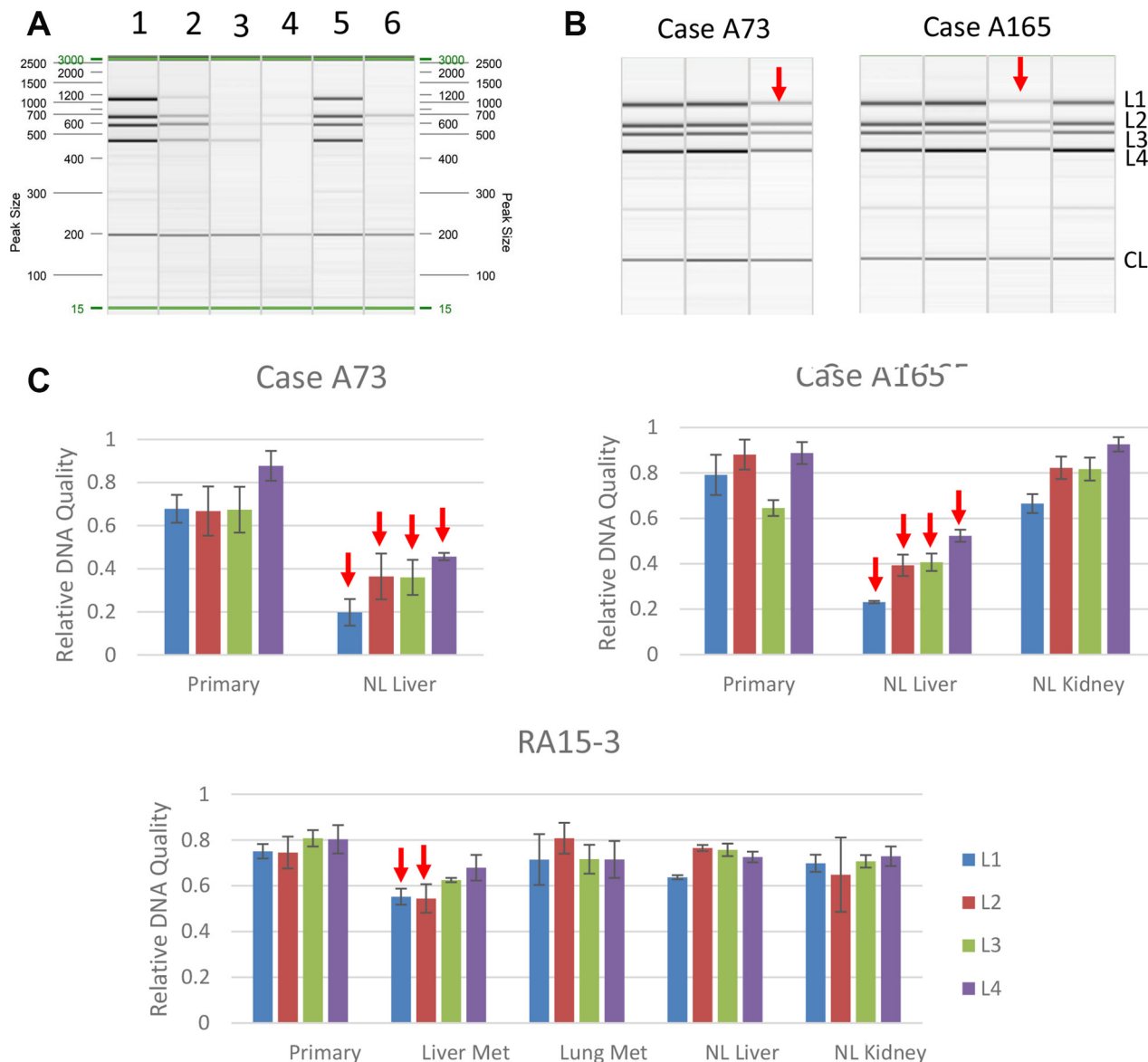


**Figure 6: RNA integrity in patient-matched liver metastases and lung metastases.** (A) Numbers listed are the corresponding RIN value for each individual tissue. N/A, data not available due to low RNA yield. (B) Scatter plot was generated by plotting RIN values from liver metastases against that from lung metastases. Pearson  $r$  and  $p$  values were from correlation analysis.

RNA samples can provide biologically meaningful information in downstream analyses. These findings are particularly exciting considering recent reports of improved methods to directly assess mRNA integrity and control for it in analyses of RNAseq data [22, 36, 37]. While we did not utilize these methods for our pilot comparison, we nonetheless were able to discern gene expression signatures of known biologic importance in normal and pancreatic cancer tissues. With larger sample sizes and use of these methods it is reasonable to expect novel observations to be made in cancer tissues, for example with respect to treatment resistance or subclonal evolution. It is also important to note that our data do

not indicate the RIN threshold value below which RNA sequencing cannot be performed, and in light of the pilot data shown it may be worth studying samples with low RIN values as well particularly as 39% of our samples had an RIN value of < 5.

Previous evaluations of the quality of biomolecules in postmortem tissues have been in the context of brain banking, forensic analyses or minimally invasive autopsies [19, 38–40]. However, such an analysis using tissues obtained from a research autopsy program for cancer research has not been performed. Such information is critical to know with the growing interest in creation of biobanks from postmortem tissues of cancer patients, and

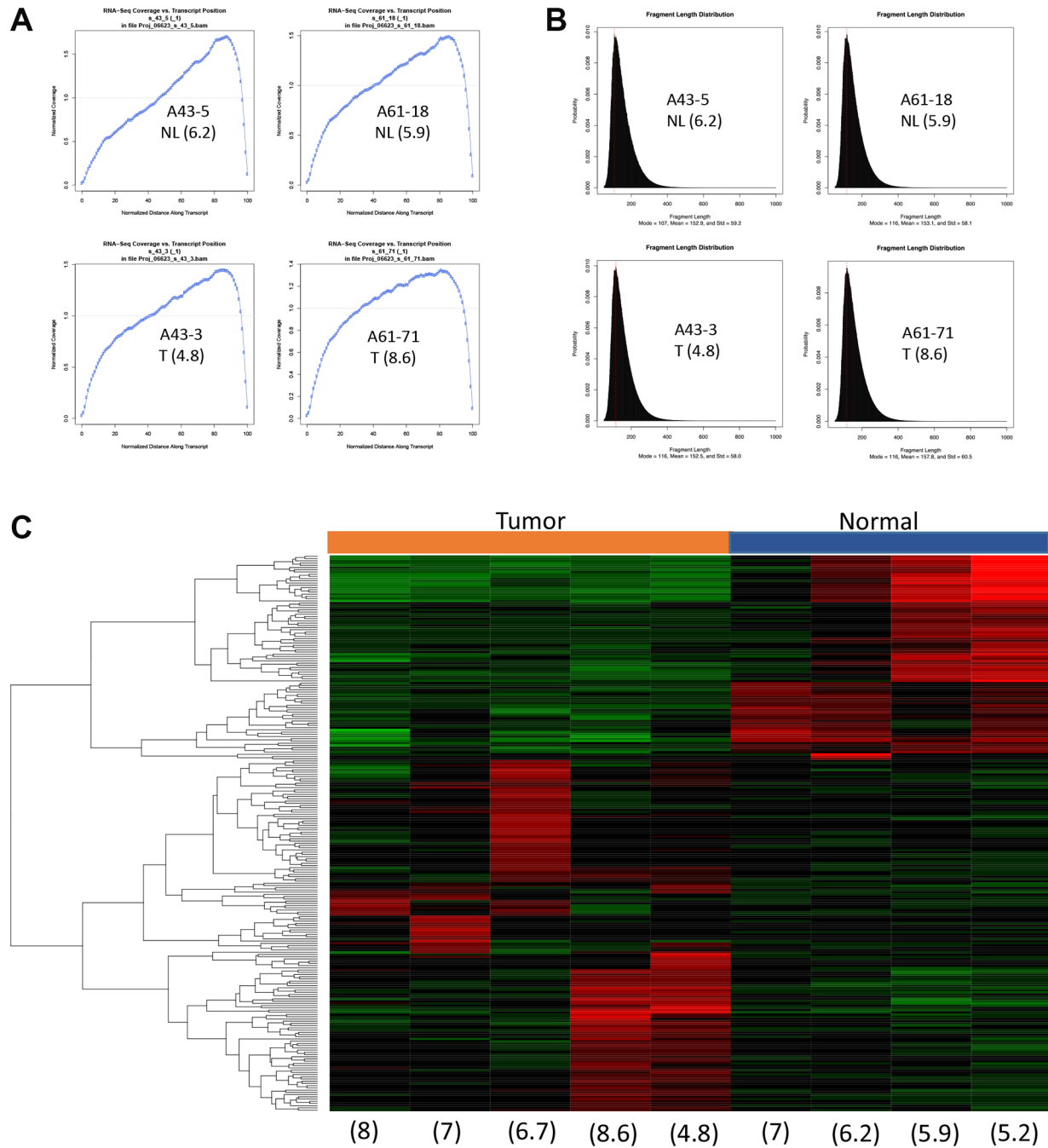


**Figure 7: Detection of DNA damage in autopsy samples.** (A) PCR band pattern using control DNA. Lane 1, Intact genomic DNA; Lane 2–4, Genomic DNA sonicated to 500 bp, 300 bp or 150 bp; Lane 5, DNA from a representative frozen autopsy tissue; Lane 6, DNA from a representative formalin-fixed paraffin-embedded (FFPE) tissue sample. (B) Representative PCR band patterns of samples collected from two autopsy cases. In each panel, the first lane is intact genomic DNA and the remaining lanes correspond to the tissue types shown in (C) Arrows indicate samples with DNA damage. L1, L2, L3, L4 and CL represent locus 1, 2, 3, 4 and control locus, respectively. (C) Relative DNA quality calculated as described in Materials and Methods. Met, metastasis; NL, normal.

use of these materials for ongoing scientific discovery and collaborations. While the scope of our dataset parallels that of another recent study [41], it differs in that we also studied DNA integrity as well and cancer tissues derived from different primary or metastatic sites. Nonetheless, our data is consistent with these prior studies that indicate nucleic acid quality of normal tissues is affected by a considerable number of factors in addition to the postmortem interval. At least one of these is likely the

cause of death as we noticed several patients with a short PMI (< 3 hours) with exceedingly poor RNA quality, all of whom died of sepsis (personal observations, C.I.D.).

Despite an increasing demand for research autopsy samples in the cancer research community, collecting high quality tissue samples is challenging due to numerous factors that can be legal, ethnic, emotional or social in nature. Exclusion of samples with a degree of poor quality is not always realistic, particularly when not all



**Figure 8: RNAseq clearly separates a tumor signature from the normal counterpart.** (A) Normalized RNA-seq coverage is plotted against transcript position. (B) Fragment length distribution. Representative samples with wide spectrum of RIN are shown in A and B. (C) Heat map shows expression profile of top 250 differentially expressed genes. Numbers in parenthesis indicate RIN values. Tumor, tissues from primary pancreatic cancer. Normal, tissue from normal pancreas.

downstream applications are equally sensitive to sample quality. Thus, analyses that can incorporate these variables would be expected to improve comparisons across patients, tissue types and sample sets. Table 5 summarizes these variables and their effects on RNA quality. Our hope is that these findings provide insight on the sample variability expected from research autopsy resources and ultimately facilitate data interpretation.

## **MATERIALS AND METHODS**

### **Tissues**

Autopsy samples were collected in association with the Johns Hopkins Gastrointestinal Cancer Rapid Medical Donation Program (GICRMDP) or the Memorial Sloan Kettering Cancer Center Medical Donation Program (MDP). Both programs were approved by the IRB at their respective institutions and in accordance with an assurance filed with and approved by the U.S. Department of Health and Human Services. Details of the program have been described previously [42]. Briefly, the tissue harvesting protocol consists of opening of the body cavity using standard techniques and sterile sampling of a variety of normal tissues, the primary tumor if present and each grossly identified metastasis using a sterile blade and forceps. For snap-freezing, tissues were collected in 1.7 ml cryovials and immediately placed in liquid nitrogen before transferring to  $-80^{\circ}\text{C}$  for long-term storage. Information regarding patient characteristics were recorded including the postmortem interval (PMI), defined as the time from death to the time of first incision. In all instances the time from the start to end of tissue sampling was  $\leq 2$  hours.

### **RNA extractions**

For each sample, approximately 30 mg tissue was carefully harvested on ice and homogenized using the Fastprep-24 system with ceramic beads (MP Biomedicals). Total RNA was then extracted using RNeasy mini kits or Fibrous tissue mini kits (Qiagen) following the manufacturer's instruction. The RNA extraction step was carefully monitored by simultaneously extracting RNA from freshly sacrificed snap frozen mouse tissues of the same tissue type as a positive control. Total RNA was eluted in DNase/RNase free water. RNA was quantified using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific) and subsequently stored at  $-80^{\circ}\text{C}$ . No more than 10 samples were extracted at one time. To confirm reproducibility of our extraction procedure, 10 samples were randomly selected and RNA re-extracted from the same tissue. In all cases similar yields were obtained from the first and second extraction. Any sample for which RNA could not be extracted was independently extracted at least one more time to rule out technical errors during the extraction procedure.

### **RNA integrity analysis**

The RNA integrity number (RIN) was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies) with an RNA nano-kit system as described in the manufacturer's instructions. For each chip analysis, a commercially available tissue-matched control human RNA with high RNA quality (ZYAGEN) was run together with autopsy RNA samples to monitor the whole analysis procedure. Data reproducibility was confirmed by repeating the RNA chip analysis in 10 randomly selected samples at least two days apart from the first analysis. In all cases the results were highly reproducible with an overall difference in RIN value between the two chip assays of  $< 1$ .

### **Genomic DNA extractions and PCR amplification**

Genomic DNA was extracted using DNeasy Mini Kits (Qiagen) according to the manufacturer's instructions. Genomic DNA was quantified by Qubit fluorometer (Invitrogen) and diluted to 20 ng/ul. 20 ng diluted DNA was subjected to PCR amplification in a total volume of 20 ul. PCR was carried out using a *Taq* PCR Core Kit (Qiagen). PCR conditions were  $94^{\circ}\text{C}$  for 2 min; 35 cycles of  $94^{\circ}\text{C}$  for 1 min,  $60^{\circ}\text{C}$  for 1 min and  $72^{\circ}\text{C}$  for 1 min 30 sec followed by 1 cycle of  $72^{\circ}\text{C}$  for 7 min. 3 ul of PCR products amplified from each locus were pooled in a new PCR tube and run on DNA screening gel cartridges (Cartridge ID C15D4D3A11) on a QIAxcel advanced system (Qiagen) according to manufacturer instructions.

### **DNA quality analysis using two-step data normalization**

We adapted the qualitative multiplex PCR assay developed by Sigma-Aldrich to detect DNA damage in formalin-fixed paraffin embedded (FFPE) tissues (<http://www.sigmaaldrich.com/technical-documents/articles/life-science-innovations/qualitative-multiplex.html>) for analysis of snap frozen postmortem tissues. The original assay consists of five primer sets derived from the NCBI UniSTS database that amplify products ranging from 132 bp to 295 bp; some or all of these products will fail to amplify when DNA damage is present. PCR primers were modified to increase the amplicon sizes from 474 to 980 bp (Table 3). One loci, a 196 bp amplicon, was used as an internal control to normalize PCR template input and amplification efficiency. The remaining four amplicons are located within known chromosome fragile sites that are relatively more susceptible to hydrolytic DNA damage and therefore serve as sensors of DNA quality [43]. To avoid amplification bias that may be introduced during multiplex PCR each loci was amplified independently and then pooled for visualization and band quantification

**Table 5: Considerations when performing RNA sequencing from postmortem tissues**

Factor <sup>a</sup>	Examples	Potential Effect on RNA Yield, Quality or Data <sup>a</sup>
<i>Ultimate Cause of Death of Patient</i>	Sepsis, Prolonged Hypotension	Lower RNA yield and quality of the affected tissues
	Acute Myocardial Infarction, Stroke	Minimal effect
<i>Postmortem Interval</i>	Prolonged	Lower RNA yield and quality that is tissue type dependent
<i>Postmortem Refrigeration</i>	Immediate delivery of patient to morgue after death	Can negate effects of prolonged postmortem interval
<i>Tissue type</i>	Normal tissue	RNA yields and quality are tissue type dependent and related to cause of death and PMI
	Cancer tissue	RNA yields and quality may not be tissue type dependent and related to extent of necrosis or autolysis in the sample
<i>Tissue storage</i>	Prolonged storage (e.g. >10 years)	Lower RNA yields
<i>Tissue handling</i>	Multiple freeze/thaw cycles	Lower RNA quality
	Prolongation between death and tissue processing	Lower RNA quality
	Prolongation between time of processing of different samples during autopsy	Potential inpatient variability in RNA quality
<i>RNA handling</i>	Multiple freeze/thaw cycles	Lower RIN values
<i>Method of RNA Extraction</i>	RNeasy mini plus kit	RNA > 200 bp
<i>Method of Library Preparation</i>	PolyA enrichment (requires relatively high quality RNA (RIN > 5))	Higher transcriptome coverage
	Ribosomal depletion (best for degraded RNA)	Lower transcriptome coverage

<sup>a</sup>These potential effects are based on the assumption that all other factors for the patient and tissues are optimal. Should more than one adverse factor exist an even greater loss of RNA yield or quality could be expected.

as described above. Commercially available human genomic DNA was used as an intact control and sonication fragmented DNA as damaged DNA control. Intact and fragmented DNA controls were analyzed in parallel with all human samples.

Data were analyzed using a two-step data normalization to acquire a relative DNA quality of each sensor loci amplicon compared to the control amplicon. First, for each sample the band Intensity from each of the four sensor loci were normalized to that of the control locus. A standard band intensity was established from the intact control DNA by using mean values from three independent amplifications. Second, the relative band intensity from all samples was subsequently normalized

to the standard. Based on this metric samples with perfect DNA quality have a value of 1 and samples with complete DNA degradation will have a value of 0. Values below the threshold of 0.6 were arbitrarily defined as having DNA damage.

### RNA sequencing

Selected RNA samples from postmortem tissues with RIN value above 5 were used for RNA sequencing. RNA sequencing was performed in the MSK Genomics Core using the Illumina Truseq RNA sample prep protocol. Briefly, RNA sequencing libraries were generated with poly-A-enrichment method and sequencing was performed

on an Illumina HiSeq2000 following standard protocols. Reads were paired-end 50 bp in length with a total of 80 millions of reads per sample. All sequenced libraries were mapped to the human genome (hg19) using rnaSTAR aligner [44]. After mapping the expression count matrix was computed from the mapped reads using HTSeq ([www-huber.embl.de/users/anders/HTSeq](http://www-huber.embl.de/users/anders/HTSeq)). The raw count matrix generated by HTSeq was then processed using the R/Bioconductor package DESeq ([www-huber.embl.de/users/anders/DESeq](http://www-huber.embl.de/users/anders/DESeq)) [45], which is used to both normalize counts and to identify differentially expressed genes between two conditions. A gene was declared differentially expressed if the fold-change was greater than 2 and the adjusted *p*-value was less than 0.05. Normalized counts were log<sub>2</sub> transformed after addition of 1 to all values. Hierarchical clustering was performed using the R *hclust* function with the Euclidean distance measure on normalized log<sub>2</sub> transformed counts after addition of 1 to all values. A heatmap was generated using the *heatmap.2* function from the *gplots* R package. The data plot shows the mean centered normalized log<sub>2</sub> expression of the top 250 genes differentially expressed between tumor and normal tissues. To detect fusion chimeras from RNA-seq data, meta-analysis that runs four fusion detection algorithms (ChimeraScan, FusionCatcher, MapSplice and DeFuse) was applied. The pipeline computes a meta-score for each detected fusion thus alleviating a problem of high numbers of false positives in each method taken independently.

## Statistics

Statistical analysis was performed using GraphPad Prism Version 6.0 (GraphPad Software, Inc. La Jolla, CA). To determine the relationship between RIN and PMI, correlations were performed to determine the *R*<sup>2</sup> value and *p* value. Curve fits were added to scatterplots by performing linear regressions. Patient-matched tumor-metastasis comparisons were compared by a two-tailed Student *t* test. A *p* value of 0.05 or less was considered statistically significant.

## CONFLICTS OF INTEREST

The authors have declared no potential conflicts of interest.

## FUNDING

This work was supported by the National Institutes of Health grants CA140599, CA179991, P30 CA008748 and a Melanoma Research Alliance grant to CID.

## REFERENCES

1. Clark MJ. Autopsy. *Lancet*. 2005; 366:1767.
2. Scatliff JH, Morris PJ. From Roentgen to magnetic resonance imaging: the history of medical imaging. *N C Med J*. 2014; 75:111–113.
3. Burke MD. Laboratory medicine in the 21st Century. *Am J Clin Pathol*. 2000; 114:841–846.
4. Samarasekera N, Al-Shahi Salman R, Huitinga I, Klioueva N, McLean CA, Kretzschmar H, Smith C, Ironside JW. Brain banking for neurological disorders. *Lancet Neurol*. 2013; 12:1096–1105.
5. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr, Kinzler KW. Cancer genome landscapes. *Science*. 2013; 339:1546–1558.
6. Stratton MR. Exploring the genomes of cancer cells: progress and promise. *Science*. 2011; 331:1553–1558.
7. Garraway LA, Lander ES. Lessons from the cancer genome. *Cell*. 2013; 153:17–37.
8. Berger MF, Lawrence MS, Demichelis F, Drier Y, Cibulskis K, Sivachenko AY, Sboner A, Esgueva R, Pflueger D, Sougnez C, Onofrio R, Carter SL, Park K, et al. The genomic complexity of primary human prostate cancer. *Nature*. 2011; 470:214–220.
9. Cancer Genome Atlas N. Genomic Classification of Cutaneous Melanoma. *Cell*. 2015; 161:1681–1696.
10. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C, Bowlby R, Shen H, Hayat S, et al. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell*. 2015; 163:506–519.
11. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, Miller DK, Christ AN, Bruxner TJ, Quinn MC, Nourse C, Murtaugh LC, Harliwong I, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*. 2016; 531:47–52.
12. Bertotti A, Papp E, Jones S, Adleff V, Anagnostou V, Lupo B, Sausen M, Phallen J, Hruban CA, Tokheim C, Niknafs N, Nesselbush M, Lytle K, et al. The genomic landscape of response to EGFR blockade in colorectal cancer. *Nature*. 2015; 526:263–267.
13. Van Allen EM, Wagle N, Sucker A, Treacy DJ, Johannessen CM, Goetz EM, Place CS, Taylor-Weiner A, Whittaker S, Kryukov GV, Hodis E, Rosenberg M, McKenna A, et al. The genetic landscape of clinical resistance to RAF inhibition in metastatic melanoma. *Cancer Discov*. 2014; 4:94–109.
14. Juric D, Castel P, Griffith M, Griffith OL, Won HH, Ellis H, Ebbesen SH, Ainscough BJ, Ramu A, Iyer G, Shah RH,

- Huynh T, Mino-Kenudson M, et al. Convergent loss of PTEN leads to clinical resistance to a PI(3)K inhibitor. *Nature*. 2015; 518:240–244.
15. Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B, Kamiyama M, Hruban RH, Eshleman JR, Nowak MA, Velculescu VE, Kinzler KW, Vogelstein B, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature*. 2010; 467:1114–1117.
  16. Liu W, Laitinen S, Khan S, Vihinen M, Kowalski J, Yu G, Chen L, Ewing CM, Eisenberger MA, Carducci MA, Nelson WG, Yegnasubramanian S, Luo J, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat Med*. 2009; 15:559–565.
  17. Stan AD, Ghose S, Gao XM, Roberts RC, Lewis-Amezcuea K, Hatanpaa KJ, Tamminga CA. Human postmortem tissue: what quality markers matter? *Brain Res*. 2006; 1123:1–11.
  18. Hansen J, Lesnikova I, Funder AM, Banner J. DNA, RNA analysis of blood and muscle from bodies with variable postmortem intervals. *Forensic Sci Med Pathol*. 2014; 10:322–328.
  19. van der Linden A, Blokker BM, Kap M, Weustink AC, Robertus JL, Riegman PH, Oosterhuis JW. Post-mortem tissue biopsies obtained at minimally invasive autopsy: an RNA-quality analysis. *PLoS One*. 2014; 9:e115675.
  20. Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, Compton CC, DeLuca DS, Peter-Demchok J, Gelfand ET, Guan P, Korzeniewski GE, Lockhart NC, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank*. 2015; 13:311–319.
  21. Gupta S, Halushka MK, Hilton GM, Arking DE. Postmortem cardiac tissue maintains gene expression profile even after late harvesting. *BMC Genomics*. 2012; 13:26.
  22. Gallego Romero I, Pai AA, Tung J, Gilad Y. RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biol*. 2014; 12:42.
  23. Kim BJ, Sprehe N, Morganti A, Wordinger RJ, Clark AF. The effect of postmortem time on the RNA quality of human ocular tissues. *Mol Vis*. 2013; 19:1290–1295.
  24. Schoenberg DR, Maquat LE. Regulation of cytoplasmic mRNA decay. *Nat Rev Genet*. 2012; 13:246–259.
  25. Olive KP, Jacobetz MA, Davidson CJ, Gopinathan A, McIntyre D, Honess D, Madhu B, Goldgraben MA, Caldwell ME, Allard D, Frese KK, Denicola G, Feig C, et al. Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science*. 2009; 324:1457–1461.
  26. Shved N, Haas C, Papageorgopoulou C, Akguel G, Paulsen K, Bouwman A, Warinner C, Ruhli F. Post mortem DNA degradation of human tissue experimentally mummified in salt. *PLoS One*. 2014; 9:e110753.
  27. Zhang P, Singh A, Yegnasubramanian S, Esopi D, Kombairaju P, Bodas M, Wu H, Bova SG, Biswal S. Loss of Kelch-like ECH-associated protein 1 function in prostate cancer cells causes chemoresistance and radioresistance and promotes tumor growth. *Mol Cancer Ther*. 2010; 9:336–346.
  28. Iacobuzio-Donahue CA, Maitra A, Olsen M, Lowe AW, van Heek NT, Rosty C, Walter K, Sato N, Parker A, Ashfaq R, Jaffee E, Ryu B, Jones J, et al. Exploration of global gene expression patterns in pancreatic adenocarcinoma using cDNA microarrays. *Am J Pathol*. 2003; 162:1151–1162.
  29. Han H, Bearss DJ, Browne LW, Calaluce R, Nagle RB, Von Hoff DD. Identification of differentially expressed genes in pancreatic cancer cells using cDNA microarray. *Cancer Res*. 2002; 62:2890–2896.
  30. Crnogorac-Jurcevic T, Efthimiou E, Capelli P, Blaveri E, Baron A, Terris B, Jones M, Tyson K, Bassi C, Scarpa A, Lemoine NR. Gene expression profiles of pancreatic cancer and stromal desmoplasia. *Oncogene*. 2001; 20:7437–7446.
  31. Argani P, Rosty C, Reiter RE, Wilentz RE, Murugesan SR, Leach SD, Ryu B, Skinner HG, Goggins M, Jaffee EM, Yeo CJ, Cameron JL, Kern SE, et al. Discovery of new markers of cancer through serial analysis of gene expression: prostate stem cell antigen is overexpressed in pancreatic adenocarcinoma. *Cancer Res*. 2001; 61:4320–4324.
  32. Mehner C, Miller E, Nassar A, Bamlet WR, Radisky ES, Radisky DC. Tumor cell expression of MMP3 as a prognostic factor for poor survival in pancreatic, pulmonary, and mammary carcinoma. *Genes Cancer*. 2015; 6:480–489. doi: 10.18632/genesandcancer.90.
  33. Ricci F, Kern SE, Hruban RH, Iacobuzio-Donahue CA. Stromal responses to carcinomas of the pancreas: juxtatumoral gene expression conforms to the infiltrating pattern and not the biologic subtype. *Cancer Biol Ther*. 2005; 4:302–307.
  34. Wuebben EL, Wilder PJ, Cox JL, Grunkemeyer JA, Caffrey T, Hollingsworth MA, Rizzino A. SOX2 functions as a molecular rheostat to control the growth, tumorigenicity and drug responses of pancreatic ductal adenocarcinoma cells. *Oncotarget*. 2016; 7:34890–906. doi: 10.18632/oncotarget.8994.
  35. Babiceanu M, Qin F, Xie Z, Jia Y, Lopez K, Janus N, Facemire L, Kumar S, Pang Y, Qi Y, Lazar IM, Li H. Recurrent chimeric fusion RNAs in non-cancer tissues and cells. *Nucleic Acids Res*. 2016; 44:2859–2872.
  36. Feng H, Zhang X, Zhang C. mRIN for direct assessment of genome-wide and gene-specific mRNA integrity from large-scale RNA-sequencing data. *Nat Commun*. 2015; 6:7816.
  37. Sigurgeirsson B, Emanuelsson O, Lundeberg J. Sequencing degraded RNA addressed by 3' tag counting. *PLoS One*. 2014; 9:e91851.
  38. Sheedy D, Harding A, Say M, Stevens J, Kril JJ. Histological assessment of cerebellar granule cell layer in postmortem brain; a useful marker of tissue integrity? *Cell Tissue Bank*. 2012; 13:521–527.
  39. Birdsill AC, Walker DG, Lue L, Sue LI, Beach TG. Postmortem interval effect on RNA and gene expression in human brain tissue. *Cell Tissue Bank*. 2011; 12:311–318.

40. Koppelkamm A, Vennemann B, Lutz-Bonengel S, Fracasso T, Vennemann M. RNA integrity in post-mortem samples: influencing parameters and implications on RT-qPCR assays. *Int J Legal Med.* 2011; 125:573–580.
41. Walker DG, Whetzel AM, Serrano G, Sue LI, Lue LF, Beach TG. Characterization of RNA isolated from eighteen different human tissues: results from a rapid human autopsy program. *Cell Tissue Bank.* 2016.
42. Embuscado EE, Laheru D, Ricci F, Yun KJ, de Boom Witzel S, Seigel A, Flickinger K, Hidalgo M, Bova GS, Iacobuzio-Donahue CA. Immortalizing the complexity of cancer metastasis: genetic features of lethal metastatic pancreatic cancer obtained from rapid autopsy. *Cancer Biol Ther.* 2005; 4:548–554.
43. Alaeddini R, Walsh SJ, Abbas A. Forensic implications of genetic analyses from degraded DNA—a review. *Forensic Sci Int Genet.* 2010; 4:148–157.
44. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013; 29:15–21.
45. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010; 11:R106.