CrossMark

Biobanking in Precision Medicine

Shanavaz Nasarabadi[1] · Michael Hogan[1] · James Nelson[1]

Abstract

Purpose of Review Precise treatment of a disease, with a concerted analysis at the genomic and phenotypic level, is a paradigm shift from the one-size-fits-all treatment practiced for centuries. As part of the Precision Medicine Initiative, a million cohort samples will be collected for research and development of unique genetic markers in the study of the relationship of genomic markers to cancer. The million cohort samples are only as good as the conditions under which the samples are collected and stored. The purpose of this review is to discuss an economically viable biobanking solution for tissue, blood, and nucleic acid storage for such an endeavor!

Recent Findings Tissue biopsy and whole blood are two common types of tissues collected for the Precision Medicine Initiative. Tissue samples can be stored as formalin-fixed paraffin-embedded (FFPE) blocks at ambient for decades, but fresh tissue samples although limited in sample size are better suited for downstream application. Postmortem collection of tissue is a good alternative to fresh tissue samples if the samples can be acquired in a timely manner before cold ischemia sets in. Blood is the preferred tissue sample for the Precision Medicine Initiative as it is easy to collect compared to other tissue types. Energy and space limitations are going to be crucial for storing a million samples for decades. Dry storage at ambient temperature is an alternative to the ultra-low-temperature storage of samples. Dry storage of whole blood samples as dried blood spots (DBS) or of the isolated components such as nucleic acids at ambient is ideal. In this review, we discuss the ambient temperature storage of blood samples and of nucleic acid.

Summary The million cohort biobanked blood and tissue samples will be crucial references for decades to come as new discoveries are made and new markers identified. Collection of blood samples at ambient as DBS and storage of the associated nucleic acid at ambient will be the key to the long-term success of biobanking of this large cohort.

Keywords Biobanking · Precision medicine · Biospecimen · FFPE · Tissue · Blood · Nucleic acid · Dry storage · Ambient storage · DNA · RNA

Introduction

The last decade has seen significant advances in sources of individual patient-relevant biological information such as genomics, proteomics, metabolomics, microbiomics, and molecular diagnostics. The Precision Medicine (earlier called and now

✉ Shanavaz Nasarabadi
shanavazn@gentegra.com

[1] GenTegra LLC, 1061 Serpentine Lane, Suite B, Pleasanton, CA 94566, USA

confused with personal medicine) Initiative kick started by President Obama under the auspices of the National Institute of Health (NIH) January 20, 2015 [1•] aims to coalesce this diverse dataset to better maintain individual health and improve the management of chronic disease and cancer. In the new age of interactive medicine, where the individual has the freedom and the ability to dictate to an increasing extent the treatment options, this initiative is indeed timely. The vast amount of electronic medical records (EMRs), combined with the precipitous drop in whole genome sequencing cost paired with the multitude of cohort studies presently underway, will allow an ever-improving determination of the interrelation between phenotypic and genotypic information.

Precision medicine is an integrative approach to treatment (Fig. 1a). Euan Ashley [2•] provides an elegant description of the differences between Precision medicine and personal medicine and the challenges of Precision medicine in the areas of sequencing technology, algorithm development, and data sharing

A

B

in context of adapting patient sequencing to the requirements of a clinical diagnostic. Precision medicine can apply the individuality of ones genotypic and phenotypic makeup to the treatment debilitating and life-threatening diseases such as cancer. The synergy between genetic markers and new therapies for cancer treatment is one powerful example of precision medicine where the favorable outcome is biased by treatment personalized to the individual rather than to the disease. Advances in non-invasive diagnostics for cancer where routine blood sample collection can

be used for tracking progression of the disease are much more affordable and less painful than a solid tissue biopsy alternative. For example, liquid biopsies [3•] are now routinely used for the analysis of cell-free DNA to look for progression/regression of cancer as well as additional genomic mutations. Opponents of precision medicine [4, 5•] argue that matching the genotypic and phenotypic makeup of the individual to the treatment will not work, but thus far an individual-refined approach to selecting treatments has yielded demonstrable if still limited success.

Fig. 1 **a** Precision Medicine initiative will bring advances in a wide range of "omics." Precision medicine caters to the precise treatment of a clinical condition where the treatment is targeted to the genetic makeup or genetic markers of the individual rather than a customized or personalized treatment regimen where the treatment is a one-of-a-kind approach would have a better clinical outcome for the patient. Precision medicine is a holistic approach to treatment where for the first time the treatment options will take into consideration the phenotypic, proteomic, metabolomic, and the genomic composition of the individual and possibly tied to the microbiome as well as the environmental and lifestyle of the individual. This approach is markedly different from the norm where the symptom is treated in isolation rather than a "whole." **b** Infrastructure needed to biobank samples in a laboratory (left to right). Blood, buccal swab or tissue biopsy/aspirate samples collected from donors are transported to the research facility if the collection site is remote to the sample processing laboratory. To maintain traceability and transparency should be mandatory to all human samples, and the samples are cataloged electronically when received and connected with the donors electronic health record (EHR) if available, along with other relevant information before either biobanking or processing the sample for nucleic acid or other biomolecule (DNA, RNA, buffy coat, proteins) extraction. Depending on the size of the donor pool, the most efficient means of processing the large cohort of samples is by automation. The biomolecule extracted is analyzed for quantity and quality before storage in a biobank at $-80\,°C$ or $-196\,°C$ (liquid nitrogen). Consequently, for an economical means of storage of a large cohort of samples, the extracted biomolecules can be stored in a chemistry for dry state or "glassy state" such as RNA/DNA stable (Biomatrica), RNAsecure (ThermoFisher), or GenTegra RNA/DNA (GenTegra LLC) or on treated paper such as Whatman FTA or GenSaver or untreated paper such as Whatman 903 or GenCollect. The choice of media for storage for biobanking is institute-dependent

A new concept in biobanking is bioinformatic biobanking where the sample is analyzed upfront for genotyping, gene expression profiling, whole genome sequencing, and RNA sequencing, and the information along with the relevant biodata are stored digitally for future reference. An individual's immutable genetic markers form a library to be queried over a lifetime for continuing patient management. Once a sample is analyzed and the data stored, it is then a simpler matter of bioinformatically querying the data when and as required. This approach is currently being implemented by Helix personal genomics with whole genome sequencing, the first step, however, once the genomic analysis is completed; any future test queries involve only a bioinformatic search as opposed to additional sample collection and repeated analysis.

Sample Types Required for Precision Medicine

The Million Donor cohort comes with a great responsibility for those institutions preserving these samples to answer future research questions [6]. Precision medicine's success will depend on the donor pool size and donor diversity connected to EMR data. Biobanking of the collected samples with a user-friendly network for analysis of big data is crucial not only to the near-term goals of the Precision Medicine Initiative but also to the future as

new methods and processes are developed. Sample availability and sample preservation via biobanking are keys to the future of the Precision Medicine Initiative and beyond. Access to a large number of samples is necessary to fulfill the vision of combining established clinicopathological parameters with emerging molecular profiling approaches to create diagnostic, prognostic, and therapeutic solutions that are precisely tailored to an individual patient's unique requirements.

The most common sample types collected for precision medicine are tissue samples and whole blood. Tissue samples can be further subdivided into liquid biopsy samples for circulating tumor cells (CTC), tissue biopsy samples such as formalin-fixed paraffin-embedded (FFPE) tissue, fresh frozen tissue samples, or wet mount tissue slides. Whole blood samples can be either peripheral blood mononuclear cells (PBMCs), serum, or plasma. Additional, albeit less common, sample types collected are cerebrospinal fluid (CSF), urine, and fecal material. When collecting these samples, it will be imperative to have "True Control" samples from surrounding disease-free tissue and corresponding known disease-state samples for comparisons which are not always practical for tissue biopsy samples or CSF. In such instances, "external" matched controls must serve as acceptable substitutes for True Control samples.

Sample collection is an intricate process that involves having many facets of the process coming together such as participant willingness, maintenance of anonymity of the sample source, and collecting samples in an ethically appropriate manner. Primarily, sample collection depends on the individual's willingness to participate in the clinical study and their trust in the collecting institution. Higher participation rates may be anticipated if the need for the study results is focused on the greater good of the community [7]. Often, the most problematic when collecting large number of donor samples is maintaining donor anonymity. The Kaiser foundation admits that donor personal information is vulnerable and has placed controls to mitigate the issue such as educating the donor as well as assigning an alternate identification to donor samples that are for the Precision Medicine Initiative [8].

Both private and federally funded institutions have set up repositories to collect and archive biological samples to be then made available to researchers globally through for-profit, paid services, or for free through not-for-profit organizations. One such globally recognized institute, the Kaiser Foundation, launched their initiative in October of 2015 and has thus far accumulated over 220,000 samples through volunteers with an end goal of collecting a total of 500,000 samples [8, 9•]. The Kaiser Foundation has the added advantage of possession of the patient's lifestyle and EMRs that can be integrated along with the sample.

A typical workflow (Fig. 1b) for biobanking of donor samples starts with collection of samples either at the medical institution (such as biopsy samples), off-site clinics for whole blood and plasma collections, or by at-home kits provided for collection of dried blood spots (DBS), buccal, urine, or fecal samples by

the donors. Depending on the number of samples being processed by a laboratory, the nucleic acid extraction (DNA and/or RNA) from the samples can be either automated or manual. The quantification and qualification of the extracted nucleic acid by optical density or fluorescence-based methods are typically followed by storage of the nucleic acid. Most nucleic acid samples are analyzed for gene expression by quantitative reverse transcription polymerase chain reaction (q-RTPCR), polymerase chain reaction (PCR), or next-generation sequencing (NGS). These techniques generally require less than 10 ng to 1 μg of total nucleic acid per test. Of prime importance is protection of the integrity of the sample during transit to the processing site since this will determine the quality of nucleic acid isolated and thus impacts all downstream applications. Nucleic acid samples are typically stored frozen at − 20 or − 80 °C but can also now be stored dried at ambient in a stabilization media.

There are many factors that affect the integrity of the isolated nucleic acid during storage. Long-term storage via cryopreservation of extracted DNA and RNA samples at − 80 or at − 196 °C (liquid nitrogen) is often described as the ultimate standard, but cryopreservation puts the specimen at risk of dehydration by cold sublimation, heat damage upon loss of power during natural disasters, and transient damage due to repeated use-thaw cycles. Chemical and biological processes can both affect the quality of the stored nucleic acid. Moreover, preservation methods must be cost-effective and able to deliver quality specimens that will yield both short-term and long-term samples obtained either in the discovery phase or clinical trials. Cryopreservation and ambient storage of extracted nucleic acid samples are both now acceptable modes of preservation. Presently, biological samples are cryopreserved either by private repositories (e.g. BioBanking without Borders LLC, NC, USA) or by public repositories (e.g. American Type Culture Collection, Manassas, USA, and Coriell Institute, NJ, USA). The cost of cryopreservation may become prohibitive when the goal is a million cohort samples for the Precision Medicine Initiative.

Blood samples collected in ethylenediaminetetraacetic acid (EDTA) tubes are the most economical and convenient form of sample collection for the Million Cohort Precision Medicine and the Million Cohort Veteran initiative [10•]. Many specialized blood collection tubes are commercially available for stabilization of transcripts such as Tempus Blood RNA tubes and PAXgene Blood RNA tubes [11]. The approach for storing samples for short-term use and long-term biobanking needs can differ but will always determine the quality of the sample. Acceptable short-term storage of weeks to months of blood and blood components such as serum, plasma, and peripheral blood mononuclear cells (PBMC's) is at 4 to − 20 °C. Long-term storage for many decades requires storage at − 196 °C under liquid nitrogen or at ambient on treated paper such as Ahlstrom-Munksjö GenSaver™ or Whatman® FTA cards. For short-term storage, untreated paper such as Whatman 903 paper, Ahlstrom-Munksjö GenCollect™ paper, may be used. Since the 1960s, DBS stored on treated paper have been used for metabolic screening, but DBS can also be used for virology, immunology, toxicology, molecular epidemiology, biomarker analysis, drug monitoring, health and wellness, and other diagnostic applications [12]. Both untreated and treated paper products function by drawing water out of the sample causing localized dehydration. In addition, treated papers such as Whatman FTA and Ahlstrom-Munksjö GenSaver cards further stabilize the sample by either lysis of the cells and/or by prevention of different kinds of oxidative damage to the sample. DNA [12] and proteins [13•] can be isolated from DBS, after decades of storage without compromise to the quality of biomolecules. Björkesten et al. [13•] demonstrated that a 1.2-mm diameter punch of DBS can give enough sample for a 92-plex protein extension assay (PEA). When analyzed from up to 30-year-old DBS samples, 36% of proteins stored at + 4 °C and 73% of proteins stored at − 24 °C were not affected with 99% reproducibility using the PEA, whereas half-lives of remaining proteins in DBS samples were between 10 and 50 years. Long-term storage temperature did affect proteins in DBS, as those stored at − 24 °C were better preserved than DBS stored at + 4 °C. DNA stored for decades on treated paper is superior to that stored on untreated paper. Viral RNA [14] and messenger RNA [15] in blood can be successfully stored in DBS for extended periods of time. Ribosomal RNA (18S and 28S rRNA) is more labile when stored in DBS yielding a less than ideal RNA integrity number (RIN), a RIN of 6.5 to 10 being ideal. To date, there is no commercially available solution for dry storage of RNA in blood other than storing at − 196 °C.

Another common sample type collected for Precision Medicine Initiative will be tissue samples from biopsies. Although core biopsy samples yield a healthy amount of tissue, tissue biopsy procedure is a painful process for the patient and can potentially cause considerable trauma to the surrounding tissue. Fine needle aspirate (FNA) biopsy with a 21-gauge needle to remove tissue samples for pathology is less traumatic to the patient and to the surrounding tissue. Compared to core biopsy samples that are typically about 17 mg or more, the FNA samples are just 2 to 10 mg and the amount of sample that is donated to research is often less than 1 mm, as priority for testing of the biopsy sample is to perform cytopathology. The best outcome for nucleic acid-based testing from tissue samples is to isolate nucleic acid from fresh or fresh-frozen at − 196 °C tissue samples. There is no ambient temperature method available to preserve tissue samples for extracting good quality nucleic acid.

FFPE is the most common method of preserving tissue samples at ambient. FFPE tissue storage has been used for 3 decades [16•] as the means of keeping tissue samples at ambient temperature for future research [17, 18]. This has created a large resource of pathologically interesting human and animal samples. Fixing tissue samples with formalin and embedding in paraffin preserves the pathology of the tissue. But, formalin fixation can cause both inter and intra protein crosslinking [19–21] as well as cross-linking of histones to DNA [22]. Other factors affecting the

quality of nucleic acid from FFPE samples include buffering formalin, time and temperature of fixation, and penetration of formalin into the tissue by stasis or by ultrasound or microwave irradiation. The nucleic acid and protein quality are additionally dependent on the time of collection of tissue following post-mortem interval and cold ischemia. Acceptable time for collection of tissue samples is between 4 h postmortem interval and 12 h after cold ischemia has set in. Acceptable time for formalin fixation of tissue postmortem is < 48 h for RNA [23, 24], < 24 h for proteins [12, 25–29], and < 72 h for DNA [30–33]. It would be best to isolate the nucleic acids from FFPE samples on or before the acceptable time to ensure the best outcome for the quality of the nucleic acid isolated. The isolated nucleic acid can be further stored at ambient by removing the aqueous media from the nucleic acid sample or by adding some commercially available stabilizers for ambient storage of nucleic acid.

Although cross-linking of nucleic acid is of concern with aged FFPE samples [16•, 30], nucleic acid extracted from FFPE samples has been successfully used for amplification, single cell analysis, and methylation studies. Decalcification of the FFPE sample using EDTA allows for longer PCR product [34], stronger fluorescence in situ hybridization (FISH) signals, lower background staining [35], and superior comparative genomic hybridization [36] results as compared to other methods.

Nucleic Acid Extraction and Storage

The quality and quantity of nucleic acid (DNA and RNA) depends on the quality of the nucleic acid in the starting material and the extraction method. There are many commercially available kits for extracting nucleic acid from varied sample types
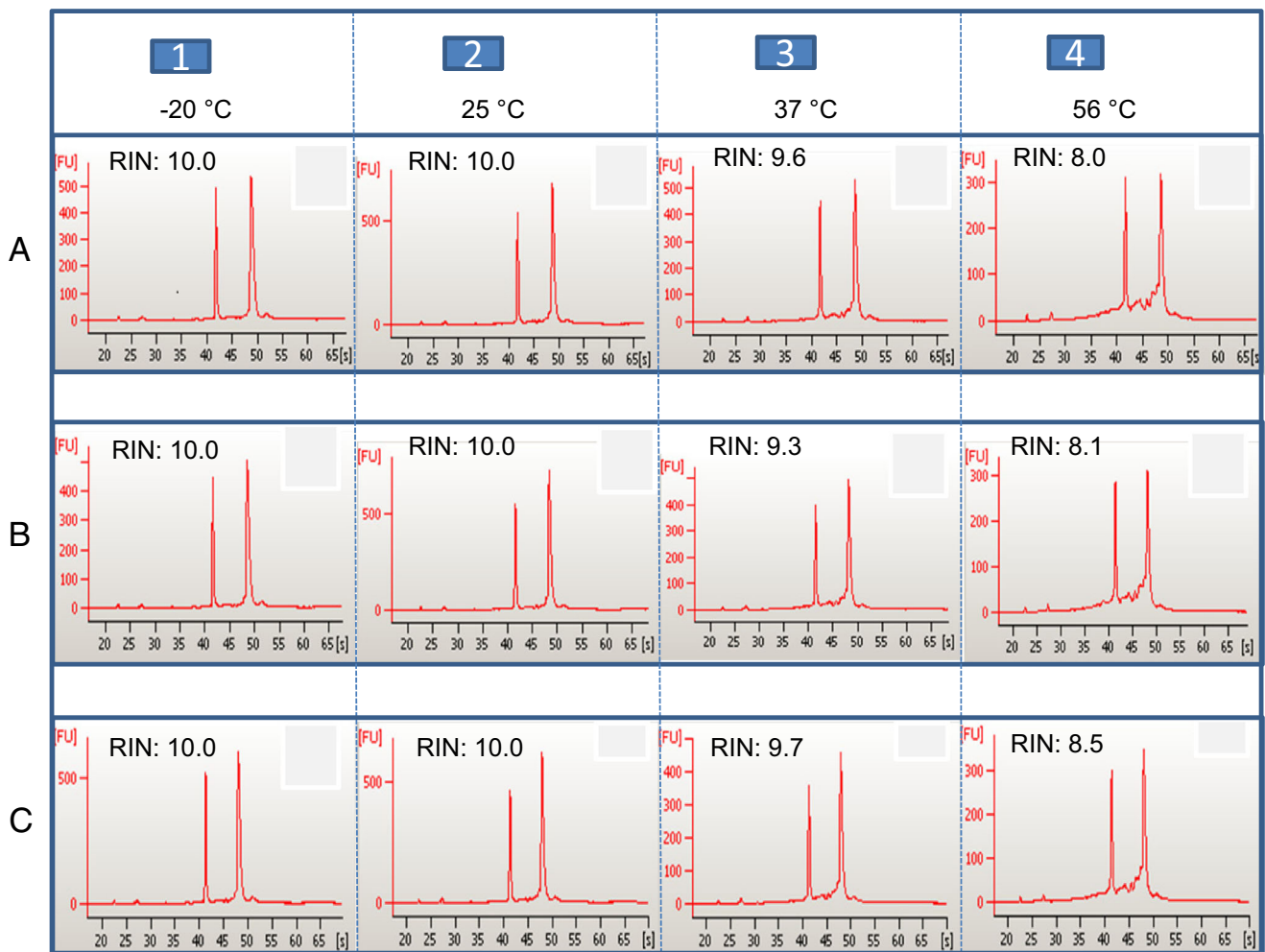
Fig. 2 Dry vitreous state or "glassy state" storage of total RNA isolated from HeLa cells diluted to 0.2 µg/µL in water (**a**), 10 mM Citrate buffer at pH 7.0 (**b**), or 1 mM EDTA (**c**) was added to GenTegra-RNA as per manufacturer's instructions and air dried before freezing at -20 °C (unaged) or aging the samples by heating at 25 °C (column 2), 37 °C (column 3), or 56 °C (column 4) for 7 months. Upon completion of the 7-month incubation, the samples were rehydrated and analyzed by electrophoresis in the Agilent Bioanalyzer. RIN scores and strand breakage frequencies for the aged samples were calculated (Table 1) from the ratios of the peak heights of 18S and 28S rRNA of HeLa RNA in GenTegra-RNA stored in dry state at -20 °C to HeLa RNA stored in GenTegra-RNA for 7 months at 25 or 37 °C (equivalent to 1 year, 8 months at ambient) or 56 °C (equivalent to 6 years, 2 months at ambient or 25 °C)

such as blood, PBMC's, serum, DBS and fresh or frozen tissue samples. Decalcified FFPE samples are treated in the same manner as tissue samples. The common mechanism by which most nucleic acid extraction kits work is through lysis of the cells to release the nucleic acid followed by capture of the nucleic acid in chaotropic agents such as guanidinium salt on paramagnetic silica beads or on glass fiber filters. The silica beads or glass fiber filters are then washed to remove the proteins and cellular debris leaving relatively clean nucleic acid on the beads/filter. The nucleic acid is then released with a buffered elution solution most commonly Tris-EDTA at pH 8.3. But, extraction methods or kits must be chosen that are well suited to the sample type [37–40]. In the literature, opinions vary regarding the choice of nucleic acid extraction kit to use for different sample types. Molteni et al. [38] determined that the efficiency of extracting DNA from DBS on plain paper and Whatman FTA paper was better with Masterpure kit than with Qiagen's QiaAmp Blood mini kit and GenSolve DNA extraction kit (GenTegra LLC) next best. In contrast, Daniels et al. of Broad institute [40] determined that the efficiency of extracting DNA from DBS on Whatman FTA paper was superior with GenSolve DNA extraction kit than the Qiagen's QiaAmp Blood mini kit. Mathew et al. [41] report comparable quality of DNA extracted from DBS on Whatman FTA paper with the GenSolve DNA extraction kit to that extracted from whole blood samples. These DNA samples were compared on the Illumina BovineSNP50 iSelect BeadChip which requires unbound, relatively intact (fragment sizes ≥ 2 kb), and high-quality DNA.

Superior quality total RNA can be extracted with the time-tested phenol extraction using the commercially available Tri Reagent available from Molecular Research Center and other suppliers, although good quality total RNA can also be obtained by using commercially available RNA extraction kits (Zymo Research, Qiagen, and ThermoFisher).

Clearly, the choice of method of nucleic acid extraction is dependent on prior sample expertise and analysis methods to be used for the study. A distinction between the quantity of nucleic acid extracted versus the quality of nucleic acid is crucial, as having large quantity of compromised nucleic acid will still result in an unsatisfactory outcome. A good check for the quality of DNA and RNA is by calculating the DNA integrity number (DIN) [42] or the RNA integrity number (RIN) [43, 44] of the sample by electrophoresis in the Agilent TapeStation or similar device. The quality of DNA can also be assessed by amplification of a 3 to 7 kb fragment of a low copy housekeeping gene such as glyceraldehyde 3-phosphate dehydrogenase (GAPDH) [45] and for RNA a > 0.9 kb fragment of a low copy gene such as RNase P [39].

Nucleic acid (both DNA and RNA) extracted from samples can be stored either at very low temperatures of − 20, − 80, or − 196 °C or in a dry state by spray drying, lyophilization, air drying in the presence of commercially available protective chemistries such as RNAstable, DNAstable (Biomatrica Inc., San Diego, CA), GenTegra-DNA [46], or GenTegra-RNA (GenTegra LLC, Pleasanton, CA), or by spotting on paper. The ribose-phosphate backbone in RNA molecules make them susceptible to degradation. RNA consequently needs to

Table 1 Total RNA extracted from HeLa cells were diluted to 0.2 µg/µL in GenTegra-RNA combined with water, 10 mM citrate buffer at pH 7.0, or 1 mM EDTA respectively and air-dried as per the manufacturer's instructions

| RNA	Buffer in GenTegra-RNA	Storage	Duration	Recovery %	RIN score	28S:18S	Normalized ratio "R_n"	Strand break per fragment "$-\ln(R_n)$"	Strand break per kb "X"
HeLa	Water	− 20 °C	7 months	100	10.0	1.10	1.00	0.00	0.00
		Ambient		100	10.0	1.22	1.11	− 0.10	− 0.03
		37 °C		101	9.6	1.16	1.05	− 0.05	− 0.02
		56 °C		101	8.0	1.00	0.91	0.10	0.03
	Citrate Buffer	− 20 °C	7 months	100	10.0	1.15	1.00	0.00	0.00
		Ambient		100	10.0	1.35	1.18	− 0.16	− 0.05
		37 °C		101	9.3	1.22	1.06	− 0.06	− 0.02
		56 °C		101	8.1	1.00	0.87	0.14	0.04
	1 mM EDTA	− 20 °C	7 months	100	10.0	1.15	1.00	0.00	0.00
		Ambient		100	10.0	1.29	1.13	− 0.12	− 0.04
		37 °C		101	9.7	1.29	1.13	− 0.12	− 0.04
		56 °C		101	8.5	1.11	0.96	0.04	0.01

The RNA samples were stored at − 20 °C, ambient (25 °C), 37 °C, or 56 °C for the duration of 7 months. The samples were rehydrated to the original volume with RNase free water. The % recovery of the samples on hydration, the RIN scores, and the strand break per fragment and per kilobase (kb) was calculated from the Agilent Bioanalyzer output. Samples diluted in either water, citrate buffer, or EDTA have similar outcomes with less than 0.03 breaks per kb. Seven months of stability at 56 °C is equivalent to 6 years, 2 months at ambient indicating that total RNA stored in GenTegra-RNA can be kept unharmed for that period of time at ambient

be stored short term at − 80 °C or long term at ultra-low temperature of − 196 °C or in a precipitated form under ethanol. It is also possible to store RNA vitrified in the dry state at ambient in the presence of protectants (GenTegra-RNA) that form the "glassy state" to prevent oxidative, hydrolytic, or RNase damage to the ribonucleic acid. The protective chemistries also allow the dry nucleic acid to re-dissolve easily because the chemistries prevent the formation of the gel-state that pure nucleic acids often form at high concentration. The gel-state makes solubilization very difficult without using mechanical forces that will also break the nucleic acid strands.

Oxygen and water are essential components in the generation of reactive molecules with the degradation process accelerating with increased temperature, reduced ionic strength of storage solution, increased concentration of divalent cations (greater than 5 ppb), or nucleolytic enzymes. In aqueous solutions (a convenient format for storage), nucleic acids are sensitive to depurination, depyrimidination, deamination, and hydrolytic cleavage. To inhibit this acid-catalyzed degradation of DNA, sample storage solutions for DNA need to be slightly alkaline-buffered solutions such as tri-buffered to pH of 8.3. Nucleic acid

extracted from clinical samples likely contains up to 30 to 40 ppb of iron (from heme or haem). Presence of even trace amounts of divalent cations (greater than 5 ppb) increases the oxidative degradation of nucleic acid due to the formation of highly reactive free radicals via Fenton reaction [47]. Adding chelating agents such as EDTA, EGTA to a concentration of 500 mM to the nucleic acid storage solution would ensure that the intrinsic divalent cations present in the clinical samples are chelated.

Dry storage of nucleic acids in the presence of protective chemistries causes the molecules to lose the ability to diffuse as the sample undergoes a non-crystalline amorphous phase or a glassy state. In this dry glassy state, the movement of protons is expected to be approximately one atomic diameter in 200 years, thus preventing both oxidative and nucleolytic degradation of the nucleic acid. Storage at ultralow temperatures of − 196 °C also vitrified as the water becomes solid ice and the molecules lose their ability to move. If moisture is added to the dry sample or the temperature is raised in ultra-cooled samples above the glass transition temperature of water, DNA/RNA damage can occur as the proton movement and reactivity resume [48]. Trace amounts of RNase
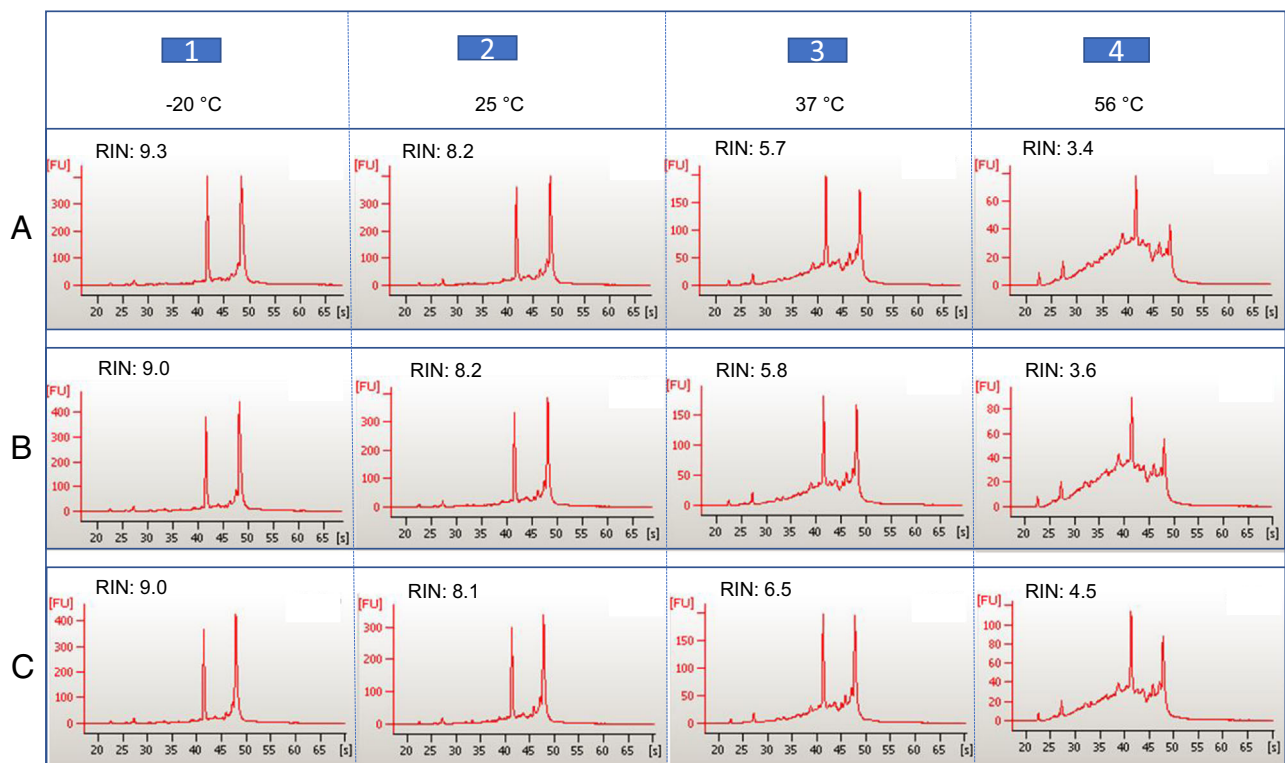
Fig. 3 Rat liver RNA diluted to 0.2 μg/μL in water (a), 10 mM citrate buffer at pH 7.0 (b), or 1 mM EDTA (c) was added to GenTegra-RNA as per manufacturer's instructions and air dried before freezing at -20 °C (unaged) or aging the samples by heating at 25 °C (column 2), 37 °C (column 3), or 56 °C (column 4) for 7 months. Upon completion of the 7-month incubation, the samples were rehydrated and analyzed by electrophoresis in the Agilent Bioanalyzer. RIN scores and strand breakage frequencies for the aged samples were calculated (Table 2) from the ratios of the peak heights of 18S and 28S rRNA. The high content of residual cellular constituents in rat liver could be responsible for the poor quality of the RNA as evident by the RIN score of the treated sample. Eighty percent of RNA was protected when stored with GenTegra-RNA for 7 months at 25 °C , 66% protected when stored in GenTegra-RNA for 7 months at 37 °C (equivalent to 1 year, 8 months at ambient or 25 °C). Heating the GenTegra-RNA treated RNA for 7 months at 56 °C (equivalent to 6 years, 2 months at ambient), RIN values dropped to RIN3.5 (4% protected), indicative of about 0.2 strand breaks per kilobase (Table 2)

Table 2 Total RNA extracted from rat liver was diluted to 0.2 μg/μL in GenTegra-RNA combined with water, 10 mM citrate buffer at pH 7.0, or 1 mM EDTA respectively and air-dried as per the manufacturer's instructions

RNA	Buffer in GenTegra-RNA	Storage	Duration	Recovery %	RIN score	28S:18S	Normalized ratio "R_n"	Strand break per fragment "$-\ln(R_n)$"	Strand break per kb "X"
Rat liver	Water	− 20 °C	7 months	100	9.3	1.00	1.00	0.00	0.00
		Ambient		104	8.2	1.16	1.16	− 0.15	− 0.05
		37 °C		105	5.7	0.81	0.81	0.21	0.07
		56 °C		104	3.4	0.57	0.57	0.56	0.17
	Citrate buffer	− 20 °C	7 months	100	9.0	1.16	1.00	0.00	0.00
		Ambient		104	8.2	1.16	1.00	0.00	0.00
		37 °C		105	5.8	0.95	0.82	0.20	0.06
		56 °C		104	3.6	0.64	0.55	0.59	0.18
	1 mM EDTA	− 20 °C	7 months	100	9.0	1.16	1.00	0.00	0.00
		Ambient		104	8.1	1.16	1.00	0.00	0.00
		37 °C		105	6.5	1.05	0.91	0.10	0.03
		56 °C		104	4.5	0.81	0.70	0.36	0.11

The RNA samples were stored at − 20 °C, Ambient (25°C), 37°C or 56°C for the duration of 7 months. The samples were rehydrated to the original volume with RNase free water. The % recovery of the samples on hydration, the RIN scores and the strand break per fragment and per kilobase (kb) was calculated from the Agilent Bioanalyzer output. The RIN score of the RNA is affected due to perhaps an amount of nucleases that surpasses the capacity of the preservation effects of GenTegra-RNA. Dry state storage for 7 months at 25°C or 7 months at 37°C (equivalent to 1 year, 8 months at ambient) still gave a RIN of greater than 8.0, but upon heating for 7 months to 56°C (equivalent to 6 years, 2 months at ambient), RIN values dropped to RIN 3.5, indicative of about 0.2 strand breaks per kilobase

would also become active upon rehydration causing RNA damage.

Successful storage of biomolecules including nucleic acid is ultimately dictated by the purity of the extracted material. Dry storage of highly pure total RNA samples from HeLa cells with a RIN of 10 can be stored dry for up to 6 years and 2 months at room temperature in GenTegra-RNA (Fig. 2) without appreciable loss of RNA integrity or strand breakage (Table 1), but rat liver RNA (Fig. 3) that carries along cellular impurities in the extracted total RNA shows degradation of up to 0.2 strand breaks per kilobase, deterioration in RIN from 9.0 to 4.0, and a short storage life of 1 year and 8 months (Table 2). Human blood lymphocyte RNA, like rat liver RNA (at an intermediate level of residual purity), displays more damage after 4 years as assessed by RIN analysis, suggestive of 0.4 to 0.5 RNA breaks per kilobase after 4 years of ambient temperature dry-state storage in GenTegra-RNA. WBC RNA, like rat liver RNA samples stored with additional 1 mm of EDTA, incurred much less damage upon 7 months storage at 56 °C (only about 0.1 break/kb) (data not shown). RNA strand breakage (X) in Tables 1 and 2 is determined from the calculated RIN value of the aged RNA to the RIN value of the unaged RNA stored at − 20 °C. To estimate the strand break density (X) from Bioanalyzer data for each sample, 18S/28S ratio of ribosomal RNA (rRNA) was calculated by comparing measured 18S rRNA peak height to the corresponding 28S rRNA peak height. That measured ratio (28S:18S$_{test}$) was then normalized to the ribosomal RNA ratio measured for the matched unaged − 20 °C sample (28S:18S$_{Time\ Zero}$) to generate R_n, where R_n =

28S:18S$_{test}$/28S:18S$_{Time\ Zero}$. Strand breaks accumulate more rapidly on the longer 28S target (5 kb) than for the shorter 18S target (1.9 kb) as a function of increasing overall strand break density. Following the assumption of Poisson statistics for calculating the frequency of strand breakage in nucleic acids [49, 50], we calculate the strand breakage density X as the natural log of R_n; the ratio of peak heights of 28S:18S rRNA test sample to peak heights of 28S:18S rRNA of control unaged − 20 °C sample as X = $[-\ln(R_n)]/3.1$ (see supplemental information).

Conclusion

Despite the tremendous strides in treatment made in the past century, cancer is still a confounding disease showing extreme heterogeneity even within those cancer cells from the same individual [51]. The aim of precision medicine is to more methodically and systematically find a way to defeat cancer [1, 52]. Creative new approaches drawing on molecular, genomic, cellular, clinical, lifestyle, environmental, behavioral, and physiological parameters are needed for detecting, measuring, and analyzing this disease [53•]. The Precision Medicine Initiative aimed at precisely and rapidly analyzing many more cancer genomes will bring about a deeper understanding of cancers fueled by discoveries of molecular diagnostic methods. The first fruits of precision medicine are already apparent, as a wide range of nucleic acid and antibody/protein-based drugs have been optimized for individuals with favorable genetic makeup. Future of drug discoveries and diagnostics will likely leverage

the twenty-first century connectivity to the masses of patients via mobile devices to provide real-time monitoring of glucose, blood pressure, cardiac rhythm, and other biological data. Genotyping may reveal new genetic variations that confer protection against specific diseases or unrecognized diversity in fecal microbiota that might identify patterns that contribute to, e.g., Parkinson's disease or to obesity [54, 55].

Living donors contribute tissue samples only if it is a medical necessity. A possible viable source of large quantities of tissue samples is through postmortem collection of whole organs and tissues from consenting families. This avenue incurs a whole host of new challenges such as the donation consent process, recovery of organs and tissues in a limited time frame postmortem, impact of donation on the donor families, and steps necessary for creating a postmortem biobank such as IRB and registries. Carthiers et al. [56•] describe development of eligibility requirements aligned with scientific needs of the project and implementation of a successful infrastructure for biospecimen procurement to support the prospective collection, annotation and distribution of blood, tissue, and cell lines, and associated clinical data from postmortem samples. The development of donor eligibility criteria is crucial since limited donor history is available within the timeframe needed for the collection of potential donor samples as degradation of biomolecules starts immediately with death.

Biobanks are key infrastructure for research and development. Located in hospitals, universities, non-profit organizations, and pharmaceutical companies, biobanks play a crucial role in health care. Biobanks are cataloged libraries of a wide range of samples like hair, bone, tissue, sperm, saliva, blood, and purified nucleic acid (DNA and RNA). Biobanking biospecimens is an expensive endeavor both in terms of manpower and natural resources used. For example, a single upright $-80\ °C$ freezer capable of holding approximately five hundred 96-well freezer boxes consumes as much energy as a small studio apartment. Approximately twenty $-80\ °C$ freezers would be needed to store the biospecimen samples for the million-sample cohort. Besides, the biospecimen sample needs to be stored for 10 years as per CLIA and CAP guidelines. Many institutions store samples for longer than a decade for research, test development, and validation purposes. With a goal of collecting a million samples for the precision medicine initiative, storing the samples such as blood at -80 or $-196\ °C$ for prolonged period (decades) could easily become impractical at some point. Consideration needs to be given to space and energy requirements for such an undertaking. A more practical approach is to consider dry ambient storage of those biomolecules that have a commercially available solution for storage. Ambient storage is economical, environmentally friendly, has a low carbon footprint, and is a practical way of storage when storage of samples for decades is needed. In addition to reducing molecular mobility, drying the samples removes water that can participate in hydrolytic degradation reactions. Furthermore, storing samples in a dry state at ambient is

independent of environmental factors such as electrical supply, temperature, and humidity. Although dry storage of nucleic acid and DBS at ambient is an economical alternative, adoption of this concept by the research community would be a paradigm shift from the time-tested method of preservation by cryogenics. This could be due simply to availability of freezers for storing other sample types that yet do not have an ambient storage method. A new technology introduced to the marketplace has a 30-year adoption cycle, and dry storage is only a couple decades into that cycle with increasing number of research facilities converting to ambient storage over time.

Biobanking of human samples has many ramifications that go beyond the science and technology of their storage. There are national, state, and even local regulations that must be met to ensure the protection of individual rights and individual privacy. Educating donors on the purpose of biospecimen collection and assurance of maintaining the privacy of the donor has favorable outcome. Perhaps, it is reasonable to consider in a future review these legal and privacy issues.

Standardization of samples is the key to successful biobanking. Reliability of samples collected in an ethical and legal manner with the oversight of the Institutional Review Board (IRB) or equivalent ethics committee for the biobanking institutions in their respective countries is crucial to ensuring reproducibility of results. International Standards are being established by both the European Union [57] and International Society of Biological and Environmental Repositories (ISBER) in the USA [58] to establish standardization metrics for biobanked samples. ISBER coordinated the launch of the International Repository Locator (IRL) website in early 2015. This centralized locator, analogous to a "repository directory," was created to increase the profile of individual repositories including ISBER, researchers, funding bodies, governments, and private industry.

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance

1.• Collins FS, Varmus H. A New Initiative on Precision Medicine. N Engl J Med. 2015;372:793–5. **Principle article describing the Precision Medicine initiative**

2.• Ashley EA. Towards precision medicine. Nat Rev. 2016;17(9): 507–22. **Highlights differences between precision medicine and personal medicine**. https://doi.org/10.1038/nrg.2016.86.

3.• Alix-Panabie C, Pantel K. Circulating Tumor Cells: Liquid Biopsy of Cancer. Clinical Chemistry. 2013;59(1):110–8. **Review article on methods developed for assessing CTC's.**

4. Prasad V. Perspective: the precision-oncology illusion. Nature. 2016;537(7619):S63. https://doi.org/10.1038/537S63a.

5.• Tannock IF, Hickman JA. Limits to personalized cancer medicine. N Engl J Med. 2016;**375**:1289–94. **Argues limitations of the precision medicine approach to treatment.**

6. Simeon-dubach D, Perren A. Better provenance for biobank samples. Nature. 2011;475:454–5.

7. Gillespie K, Luft H, Hernandez Y, Lee S, Cho M, Craft S. Patient views on the use of personal health information and biological samples for biobank research. J Patient Cent Res Rev. 2017;4:17.

8. https://healthitanalytics.com/news/kaiser-permanente-expands-precision-medicine-biobanking-effort

9.• https://researchbank-econsent.kaiserpermanente.org/ConsentForm/ConsentForLocalPrint/. **Kaiser Foundation's efforts to collect and store the biological samples.**

10.• Gazianoa JM, Concatoc J, d, Brophya M, Fiorea L, Pyarajana S, Breelinga J, et al. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. J Clin Epidemiol. 2016;70:214e223. **The Million Veteran Program initiative to look for genetic markers, a prequel to Precision Medicine Initiative.**

11. Hantzsch M, Tolios A, Beutner F, Nagel D, Thiery J, Teupser D, et al. Comparison of whole blood RNA preservation tubes and novel generation RNA extraction kits for analysis of mRNA and MiRNA profiles. PLoS One. 2014;9(12):e113298. https://doi.org/10.1371/journal.pone.0113298.

12. von Wasielewski R, Mengel M, Nolte M, Werner M. Influence of fixation, antibody clones, and signal amplification on steroid receptor analysis. Breast J. 1998;4(1):33–40. https://doi.org/10.1046/j.1524-4741.1998.410033.x.

13.• Björkesten J, Enroth S, Shen Q, Wik L, Hougaard DM, Colen AS, et al. Stability of Proteins in Dried Blood Spot Biobanks. Mol Cell Proteomics. 2017;16(7):1286–96. **Report isolation of stable protein from DBS stored for up to 30 years at +4C or -24C. Drying blood on paper was 99% reproducible with half-life in the range of 10 to 50 years.**

14. Bertagnolio, S., Parkin, N., T., Jordan, M., Brooks, J., Garcia-Lerma, J., G. Dried Blood spots for HIV-1 drug resistance and viral load testing: a review of current knowledge and WHO efforts for global HIV drug resistance surveillance. AIDS Rev 2010 Oct; 12(4):195–208.

15. Gauffin F, Nordgren A, Barbany G, Gustafsson B, Karlsson H. Quantitation of RNA decay in dried blood spots during 20 years of storage. Clin Chem Lab Med. 2009;47(12):1467–9. https://doi.org/10.1515/CCLM.2009.351.

16.• Fox CH, Johnson FB, Whiting J, Roller PP. Formaldehyde fixation. J Histochem Cytochem. 1985;33:845–53. **Review article on FFPE**

17. B. Paige Bass, PhD; Kelly B. Engel, PhD; Sarah R. Greytak, PhD; Helen M. Moore, PhD A review of preanalytical factors affecting molecular, protein, and morphological analysis of formalin-fixed, paraffin-embedded (FFPE) tissue: how well do you know your FFPE specimen? Arch Pathol Lab Med 2014 Nov; 138:1520–1530, 11, DOI: https://doi.org/10.5858/arpa.2013-0691-RA.

18. Hood BL, Darfler§ MM, Guiel TG, Furusato B, Lucas DA, Ringeisen BR, et al. Proteomic analysis of formalin-fixed prostate cancer tissue. Mol Cell Proteomics. 2005;4(11):1741–53. https://doi.org/10.1074/mcp.M500102-MCP200.

19. Fraenkel-Conrat H, Olcott HS. The reaction of formaldehyde with proteins; crosslinking between amino and primary amide or guanidyl groups. J Am Chem Soc. 1948;70(8):2673–84. https://doi.org/10.1021/ja01188a018.

20. Metz B, Kersten GF, Hoogerhout P, Brugghe HF, Timmermans HA, de Jong A, et al. Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. J Biol Chem. 2004;279(8):6235–43. https://doi.org/10.1074/jbc.M310752200.

21. Middlebrook WR, Phillips H. The action of formaldehyde on the cystine disulphide linkage of wool; the conversion of subfraction a of the combined cystine into combined lanthionine and djenkolic acid and subfraction B into combined thiazolidine-4-carboxylic acid. Biochem J. 1947;41:218223.

22. Jackson V. Studies on histone organization in the nucleosome using formaldehyde as a reversible cross-linking agent. Cell. November 1978;15(3):945–54. https://doi.org/10.1016/0092-8674(78)90278-7.

23. Beaulieu M, Desaulniers M, Bertrand N, et al. Analytical performance of a qRT-PCR assay to detect guanylyl cyclase C in FFPE lymph nodes of patients with colon cancer. Diagn Mol Pathol. 2010;19(1):20–7.

24. Chu WS, Furusato B, Wong K, et al. Ultrasound-accelerated formalin fixation of tissue improves morphology, antigen and mRNA preservation. Mod Pathol. 2005;18(6):850–63.

25. Boenisch T. Effect of heat-induced antigen retrieval following inconsistent formalin fixation. Appl Immunohistochem Mol Morphol. 2005;13(3):283–6. https://doi.org/10.1097/01.0000146524.74402.a4.

26. Ibarra JA, Rogers LW, Kyshtoobayeva A, Bloom K. Fixation time does not affect the expression of estrogen receptor. Am J Clin Pathol. 2010;133(5):747–55. https://doi.org/10.1309/AJCPPIUHS4GVAR0I.

27. Leong AS, Milios J, Duncis CG. Antigen preservation in microwaveirradiated tissues: a comparison with formaldehyde fixation. J Pathol. 1988;156(4):275–82. https://doi.org/10.1002/path.1711560402.

28. O'Rourke MB, Padula MP. Analysis of formalin-fixed, paraffin-embedded (FFPE) tissue via proteomic techniques and misconceptions of antigen retrieval. BioTechniques. May 2016;60(5):229–38. https://doi.org/10.2144/000114414.

29. Williams JH, Mepham BL, Wright DH. Tissue preparation for immunocytochemistry. J Clin Pathol. 1997;50(5):422–8. https://doi.org/10.1136/jcp.50.5.422.

30. Ferrer I, Armstrong J, Capellari S, et al. Effects of formalin fixation, paraffin embedding, and time of storage on DNA preservation in brain tissue: a BrainNet Europe study. Brain Pathol. 2007;17(3):297–303.

31. Nuovo GJ, Richart RM. Buffered formalin is the superior fixative for the detection of HPV DNA by in situ hybridization analysis. Am J Pathol. 1989; 134(4):837–842.O'Leary JJ, Browne G, Landers RJ, et al. The importance of fixation procedures on DNA template and its suitability for solution-phase polymerase chain reaction and PCR in situ hybridization. Histochem J. 1994;26(4):337–46.

32. Reineke T, Jenni B, Abdou MT, et al. Ultrasonic decalcification offers new perspectives for rapid FISH, DNA, and RT-PCR analysis in bone marrow trephines. Am J Surg Pathol. 2006;30(7):892–6.

33. Zsikla V, Baumann M, Cathomas G. Effect of buffered formalin on amplification of DNA from paraffin wax embedded small biopsies using real-time PCR. J Clin Pathol. 2004;57(6):654–6. https://doi.org/10.1136/jcp.2003.013961.

34. Wickham CL, Sarsfield P, Joyner MV, Jones DB, Ellard S, Wilkins B. Formic acid decalcification of bone marrow trephines degrades DNA: alternative use of EDTA allows the amplification and sequencing of relatively long PCR products [published correction appears in Mol Pathol. 2001;54(2):120]. Mol Pathol. 2000; 53(6): e336.

35. Babic A, Loftin IR, Stanislaw S, et al. The impact of pre-analytical processing on staining quality for H&E, dual hapten, dual color in situ hybridization and fluorescent in situ hybridization assays. Methods. 2010;52(4): 287–300. Reineke T, Jenni B, Abdou MT, et al. ultrasonic decalcification offers new perspectives for rapid FISH, DNA, and RT-PCR analysis in bone marrow trephines. Am J Surg Pathol. 2006;30(7):892–6.

36. Alers JC, Krijtenburg PJ, Vissers KJ, van Dekken H. Effect of bone decalcification procedures on DNA in situ hybridization and comparative genomic hybridization: EDTA is highly preferable to a routinely used acid decalcifier. J Histochem Cytochem. 1999;47(5):703–10. https://doi.org/10.1177/002215549904700512.

37. Choi E-H, Lee SK, Ihm C, Sohn Y-H. Rapid DNA extraction from dried blood spots on filter paper: potential applications in biobanking. Osong Public Health Res Perspect. 2014 Dec;5(6): 351–7. https://doi.org/10.1016/j.phrp.2014.09.005.

38. Molteni CG, Terranova L, Zampiero A, Galeone C, Principi N, Esposito S. Comparison of manual methods of extracting genomic DNA from dried blood spots collected on different cards: implications for clinical practice. Int J Immunopathol Pharmacol. 2013 Jul-Sep;26(3):779–83. https://doi.org/10.1177/039463201302600324.

39. Olsvik PA, Lie KK, Jordal A-EO, Nilsen TO, Hordvik I. Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. BMC Mol Biol. 2005;6:21.

40. Daniels R, Volkman SK, Milner DA, Mahesh N, Neafsey DE, Park DJ, et al. A general SNP-based molecular barcode for plasmodium falciparum identification and tracking. Malar J. 2008;7(1):223. https://doi.org/10.1186/1475-2875-7-223.

41. McClure MC, McKay SD, Schnabel RD, Taylor JF. Assessment of DNA extracted from FTA®cards for use on the Illumina iSelect BeadChip. BMC Research Notes. December 2009;2(1):107. https://doi.org/10.1186/1756-0500-2-107.

42. Application note: Arunkumar Padmanaban. DNA Integrity Number (DIN) For the Assessment of Genomic DNA Samples in Real-Time Quantitative PCR (qPCR) Experiments. http://hpst.cz/sites/default/files/attachments/5991-6368en.pdf

43. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Grassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol. 2006;7(1):3. https://doi.org/10.1186/1471-2199-7-3.

44. Application note: Odilo Mueller Samar Lightfoot Andreas Schroeder. RNA Integrity Number (RIN) – Standardization of RNA Quality Control. http://gene-quantification.net/RIN.pdf

45. Kozera B, Marcin M. Reference genes in real-time PCR. J Appl Genetics. 2013;54(4):391–406. https://doi.org/10.1007/s13353-013-0173-x.

46. Shana L McDevitt, Michael E Hogan, Derek J Pappas, Lily Y Wong, and Janelle A Noble. DNA storage under high temperature conditions does not affect performance in human leukocyte antigen genotyping via next-generation sequencing (DNA integrity

47. Graf E, Mahoney J, Bryant R, Eaton J. Iron-catalyzed hydroxyl radical formation. Stringent requirement for free iron coordination site. J Biol Chem. 1984;259(6):3620–4.

48. Williams RJ, Carl Leopold A. The glassy state in corn embroys. Plant Physiol. 1989 Mar;89(3):977–81. https://doi.org/10.1104/pp.89.3.977.

49. Bohr VA, and Okumoto DS. (pp. 347–366): DNA Repair: A Laboratory Manual of Research Procedures. In:Friedberg, E.C., and Hanawalt, P.C., editors. Marcel Dekker, Inc., New York, 1988.

50. Fritz LK, Suquet C, Smerdon MJ. Strand breaks are repaired efficiently in human ribosomal genes. J Biol Chem. 1996;271(22): 12972–6. https://doi.org/10.1074/jbc.271.22.12972.

51. Alizadeh AA, Aranda V, Bardelli A, Blanpain C, Bock C, Borowski C, et al. Toward understanding and exploiting tumor heterogeneity. Nat Med. 2015;21(8):846–53. https://doi.org/10.1038/nm.3915.

52. Shin SH, Bode AM, Dong Z. Precision medicine: the foundation of future cancer therapeutics. npj Precision Oncology. 2017;1, 1

53.• Reza Mirnezami MRCS, Nicholson J, PD, Darzi A, Md. Preparing for Precision Medicine. N Engl J Med. 2012;366:489–91. **Brings together the perspective on the paradigm shift of how the holistic approach of precision medicine will aid in the treatment of cancer and potentially reduce the time to market of life saving drugs**

54. Parashar A, Udayabanu M. Gut microbiota: implications in Parkinson's disease. Parkinsonism Relat Disord. 2017 May;38:1–7. https://doi.org/10.1016/j.parkreldis.2017.02.002.

55. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science. 2013;341(6150):1241214. https://doi.org/10.1126/science.1241214.

56.• Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. Biopreserv Biobank. 2015, 13;(5):311–9. **Describes criteria to use for consent forms in the collection of post-mortem tissue for research purposes. The article also outlines implementation of a network for successful acquisition and dissemination of post-mortem tissue samples**

57. http://www.bbmri-eric.eu/wp-content/uploads/2017/02/2017_Newsletter6_7_WEB.pdf

58. Vaught Jim. The International Scene in 2016: Biopreservation and biobanking. **ISBER**.

maintained in extreme conditions). Biopreservation and Biobanking. Volume 12, Number 6, 2014.