Repository of Human Blood Derivative Biospecimens in Biobank: Technical Implications

Ashraf Mohamadkhani¹, Hossein Poustchi¹*

1. Liver and Pancreatobiliary Diseases Research Center, Digestive Diseases Research Institute, Tehran University of Medical Sciences, Tehran, Iran

ABSTRACT

Human biorepositories are collection of biological samples and health information from a large number of participants generally in the cohort studies. The main purpose of established biobanks is organization of biomedical research for upgrading the knowledge of human disorders from cancer to infectious and rare disease. The studies of generation relationships and understanding the preclinical stages of ageing are also from the solution of bitobank.

This review overview the significance and storage condition of biospecimens including whole blood, red blood cells (RBC), buffy coat, plasma, serum, DNA and RNA that derived from blood in human biobanks. These biological samples provide valuable information on the prevalence of germline mutations, epigenetic modifications or interaction between genes and proteins in associated with the development of certain types of disease. The quality of biospecimen in biobanks is a powerful tool for valid identification of biomarkers. Therefore optimum qualities of human biological samples in long time storage that have been assessed in several studies also indicate in this review.

KEYWORDS
Whole blood; Buffy coat; Plasma; Serum; DNA; Biobank

Please cite this paper as:

INTRODUCTION

Biobanks are carefully characterized to arrange for real answers in scientific and clinical research for diagnosis and treatment of a range of life-threatening diseases. The establishment of biobanks most often offers the volunteer-based sample for conducting the research.¹ ² Development of this new field of biorepository and biospecimen science could increase upcoming issues related to human genome, transcriptome, and proteome.³ The public health policy has convinced the governments to generate population biobanks for diagnostic and etiopathogenesis studies over the past two decades.⁴ Along these lines, there are a growing number of national biobanks for collecting data and samples from a population or subpopulation for providing large-scale research resources.⁵ ⁶ Initially repositories had taken small rules in which biospecimens were kept through laboratory notebook and limited to a few freezers. Today’s technological developments support specimens’ annotation and establishment and therefore accelerate the arrangement of hands-on biobanks. Moreover using web services for computerized biobank expand communication with clients and opened collaborations between the public and private divisions. These services
also guarantee acquire of biological material and help international research for improving human health.\textsuperscript{1,7} Therefore creations of biobank lead to accumulation of repositories of biological samples from clinical sources into the central database.\textsuperscript{3,8}

Cohort studies encompass a prospective design with long-term follow-up with a large number of participants. Genomic sequencing approaches become routine in research laboratories that assist researches in human blood DNA or RNA and studies of similarities of three-generation and identification of preclinical stages of ageing at earlier age.\textsuperscript{3,5} Scientists have sought to integrate relationship of genetic and biomarkers data with lifestyle and environmental factors by obtaining biospecimens from participants in cohort studies to solve the important questions on human health.\textsuperscript{6,9} They also arrange opportunities for biomarkers discovery to outline subgroups of diseases and the therapeutic activities that anticipate developing patient management.\textsuperscript{10-12} Biomarkers represent the biological characteristic from physiological state, pathogenic processes, or pharmacologic responses to a therapeutic intervention.\textsuperscript{11} Evaluation of biomarkers is useful approaches for early detection of disease, and identification of predictive and prognostic factors in disease management. Recent developments in areas of gene-expression microarrays, proteomics, and metabolomics with combining molecular profiling data from multiple omics lead to the discovery and validation of molecular biomarkers in medical research. Studies of biomarkers are non-invasive strategies that are afforded by keeping biospecimens in biobank,\textsuperscript{11,12} Along these lines, large pharmaceutical companies, universities, and state organizations attempt funding for biobanks undertake comprehensive benefits in analysis of metabolite, biomarkers, minerals and relevant immunoprotective factors related to health and diseases.\textsuperscript{3,11}

Peripheral blood, urine, hair and nail and purified DNA as well as biofluids (breast milk, semen and saliva), are important samples in the human biobank. It should be careful that people who join in research studies related to biobank are usually not the same to those who do not take part and mostly are in healthier condition therefore to remain valid studies belong to biobanks caution need to considered in keeping eligible prevalence of participants with cohort studies.\textsuperscript{2,13-15} Therefore, this review looks at studies that investigated the status of biobanking for their potential in using of blood derivative biospecimen and the importance of high quality storage of these samples.

**Types of Biospecimen in Human Biobank**

Blood, urine and tumour tissue are common biological samples in biobanks that donate by participant under a separate consent form and are critical for biomedical research. However more progressively wide-ranging of biospecimen types including hair, nail clippings, saliva, faeces and breast milk have also might be collected and processed according to guidelines.\textsuperscript{6,13,16} Demographic associated data from life style, medical history and medications, and possible causes of exposure data manage by trained research staff to create a resource of scientific database. Samples do not labeled with individual names and they are identifiable by only a unique code.\textsuperscript{6} Sometimes collected samples link to the clinical trials or public health studies therefore the privacy of general public is crucial concern in management of any biobank. Maintain of samples quality and protections of individual rights are very critical thus biobank managers are responsible for potential staff with perfect understanding of the facts of given biological samples in to a biobank. This helps the researchers’ access to biobanks when they are in need of specimen for their research studies. Biobanks work as vital resource for worldwide scientists and facilitates to find particular collections of biospecimens.\textsuperscript{15} In following we talk about the potential signficance of human blood derivative biospecimens in biobank such as whole blood, red blood cells (RBC), buffy coat, plasma, serum, and purified nucleic acids.

**Significance of Whole Blood and Blood Cells in Long Term Human Biobanks**

From human biospecimens, the whole blood and buffy coat are essential to biobanking efforts and
are willing to share among international biobank. Whole blood is typically collective with an anticoagulant and consists of RBCs, white blood cells (WBCs), and platelets which suspend in plasma. Equally, the part of anticoagulated blood sample that consist of the white blood cells and platelets termed buffy coat. Both are the main source for cellular nucleic acids and achieving optimal quality and quantity of germline DNA and constructing DNA biobank.\textsuperscript{3,17} DNA/RNA biorepositories are potentially required for the comprehensive analysis of functional genomics that give details for characteristic and expression of whole sets of human genes.\textsuperscript{7,17} Prominent request of whole blood and buffy coat in biobanks are exploration of copy-number variations (CNVs) or chromosomal regions in association with non-Mendelian diseases\textsuperscript{18} as well as molecular mechanisms including epigenetic modifications or interaction between genes and proteins that play a major role in the risk of developing some diseases.\textsuperscript{19} DNA is the main informative molecule that leading the protein synthesis activity in a particular cell. Crystal structures of coiled coils DNA indicate that their geometries can vary according to sequence or the presence of stabilizers such as proteins or small molecules.\textsuperscript{20} Likewise, protein-RNA interaction is required for the stability of the supercoiled RNA.\textsuperscript{21} One important concern in long term storage of buffy coat specimens is stage-specific transcriptome studies by using gene expression array and RNA sequencing analyses in a great number of infectious and non-infectious diseases.\textsuperscript{3} Epigenomics related studies are another field of genomics approach in biobank medicine. DNA methylations play an important role in silencing tumor suppressor genes during cancer development.\textsuperscript{18} Furthermore, red blood cells (RBCs) reveal unique protein expression profile in the states of health and disease. During maturation they lost all cellular organelles and consequently are not capable for proteins substitute. RBCs inherited disorders with heterogeneous clinical presentation either in membrane structural combination or function of membrane transporters that characterize them for comprehensive membrane proteomics research from transfusion medicine to infectious diseases.\textsuperscript{22,23}

**Storage-Keeping Quality for Human Blood Cells**

Blood is the most collective biospecimens in human biobank as a source of DNA and RNA. Anticoagulated blood is prerequisite in purification of genomic or mitochondrial DNA and RNA molecules (tRNA, mRNA, miRNA) as well as plasma derived cell-free circulating nucleic acid molecules. Among the most commonly used anticoagulants, EDTA used for a variety of DNA based studies and protein assays, citrate is more appropriate in WBC culture, however heparin despite the advantage for RNA and DNA yields is not recommended as it could inhibit PCR.\textsuperscript{24,25} The yields of extracted DNA and RNA from whole blood or buffy coat specimens are determined by blood storage condition and the keeping quality of biospecimens in long term. RNA is the most easily degradable part of a biological sample in biological contexts by abundant and ubiquitous RNAses that make it a major problem in biobank.\textsuperscript{21,26} WHO-IARC biobank suggests nitrogen storage below -130°C as optimal temperature for long term preservation of biomolecules such as RNA.\textsuperscript{27} Samples storage at -140°C by liquid nitrogen keep the RNA in intact and functional state in a period of more than 50 months as biochemical activity that influence intracellular contents are assumed to be inactive.\textsuperscript{28-30} Currently numerous centers use −80 °C temperature freezing for simple and safe storage compare to liquid nitrogen storage.\textsuperscript{6,31} Although some studies have indicated for 5’ transcript tags RNA fragmentation after five years storage at −80 °C\textsuperscript{32} however, recent study showed that long-term freezing in -80°C does not adversely affect to the RNA extracted value from the deposited biospecimen.\textsuperscript{33} The impact of thawing on RNA integrity and gene expression in fresh frozen biospecimens has also been evaluated. Minimal RNA degradation was detected after half hour of thawing in unfixed samples.\textsuperscript{34} RNA integrity is kept with RNase-free handling however, addition of commercially available RNase inhibitor such as RNAshell or RNAstable give accurate real-time PCR results
and protect the value of the biospecimens for other assays. Tissues incubated in RNAlater efficiently prevent RNA degradation.

Biosources of DNA in biobank is an integral part for clinical research. It has been shown that storing the blood sample at -80 °C does not adversely emotional impact on the quality of extracted DNA. To keep the biospecimens reliability from multiple freeze-thaw cycles, protection and stabilization of DNA at room temperature (RT) endorsed in biobanking. It also eliminates the costs associated with freezer storage and reduces the maintenance costs for biobanks. In this way, DNAshell and RNAshell make sure the safe storage of blood spots at room temperature for long periods of time and DNA molecules are appropriate for common downstream analysis. Stable dry-state storage of blood is desirable to reduce required storage space and to diminish electrical and the budgets for biospecimens shipping. DNA purified from various commercially available dry-state stabilization matrices run successfully in molecular applications. Formalin-free preservatives show approval but fluctuating resources to preserve DNA and RNA, though RNA is degraded during paraffin embedding process.

Purified DNA is stable at 4 °C for numerous weeks and for months at -20 °C. However, purified DNA and RNA are more stable at -80 °C in aqueous buffers or nuclease-free water for long-term storage. Blood cells without dimethylsulphoxide (DMSO) are unable to get along viable because cells will rupture if frozen. However storage of RBCs is recommended at -80 °C for cell membrane proteomics research. Complete descriptions of blood derived biospecimens in biobanks are explained in Table 1.

Technical Importance of Serum and Plasma in Human Biobanks

Plasma is the liquid part of normal unclotted blood that suspending red and white cells as well as platelets and need anticoagulants like EDTA for its separation. UK Biobank, anti-coagulants EDTA and lithium–heparin are used for high-quality plasma. Serum forms when blood cells and clotting proteins remove from blood by centrifugation. Plasma contributes up to about 55% of the body’s total blood volume. This difference in processing is the basis of some diversity such as metabolites between serum and plasma. Throughout the coagulation, platelets proteins such as proinflammatory cytokines and metabolites like sphingosine-1-phosphate diffuse within serum. Therefore the advantages of serum for some metabolite are on debate. Human plasma and serum are most common materials in biological and clinical studies and could practice interchangeably for a wide variety of substrates like enzymes, hormones and cell free DNA or RNA for diagnosis and monitoring of disease. However a number of studies have demonstrated the different level of metabolites between plasma and serum. More than 4000 compounds that are annotated in human metabolite databases and plasma is preferable matrices in clinical chemistry investigations for metabolic disease such as diabetes. Moreover higher serum levels of lysophosphatidylcholine and diacylglycerol releasing in coagulation process, prefer plasma in study of lipid metabolite. There are a wide range of proteins in size and structure with a concentration of 60–80 mg per ml that is highly correlated between serum and plasma. Nonetheless because of lacking coagulant protein, the assessment of proteins in serum is more valuable in clinical diagnosis and monitoring of a range of diseases. Recent advance in proteomics and bioinformatics procedures improved the number of up proteins to 10,000 in serum though most of which present at very low relative abundances. The study of serum protein profile supportively evaluates the state of health or disease and help to manufacture the design of proper drugs in interaction with host cells. Inflammation is the common cause in serum protein alterations mostly by change in expression of the acute phase response. Albumin, prealbumin, apo B, apo A-1, retinol-binding proteins (RBP) and transferrin are acute phase hepatic proteins that usually decreases while C-reactive protein (CRP) and serum amyloid A (SAA) with marked increases and complement factors C3 and C4, ceruloplas-
min, fibrinogen, plasminogen, haptoglobin, alpha 1-antitrypsin/acid glycoprotein/chymotrypsin, and ferritin with more modest increases relate to inflammation that could confound serum protein interpretations.

Any disequilibrium in serum proteins points to the presence of infectious diseases, inflammation, electrolyte disorders, autoimmune disease or tumors.

All known resources of proteins in plasma and serum based on Mass spectrometry derived data are available in Plasma Proteome Database (PPD) that allows users in identifying of proteins or peptides compared to known plasma proteins.

Tumor-associated DNA in patients with cancer divide into DNA from cells in the blood and circulating cell free DNA from the source of tumor cells. The studies of free DNA are sensitive and specific for cancer diagnosis and follow-up therapy and improve by great development in PCR-based methods, mutations detection and DNA methylation analysis. The advantages of serum or plasma for these kinds of studies remain to determine, however it was reported that serum has greater yield of free DNA or RNA due to the release of nuclear acids from damaged blood cells throughout coagulation.

### Serum or Plasma Stability in Long Time Storage

Previous studies have indicated that the values of most analytes are comparable in mammalian serum and plasma when the blood is taken in standardized way and separated from the cells in 2 hours. Hemolyzed samples convince significant differences in potassium, phosphorus, albumin, and lactate dehydrogenase. Appropriate blood collection tubes with and without anticoagulant are accessible commercially. Clotted blood need to collect either into glass or coated plastic tubes to inhibit sticking the clot to the container walls. Gel-barrier tubes are frequently used with no significant effect on metabolites concentration. They are preferable because of the easy administration of blood collection and separation of serum in the period of shipment. Samples for biochemical analysis need to separate as quickly as possible to prevent ongoing metabolism of cellular constituents as well as hemolysis and leakage of analytes between the plasma or serum and cellular compartments. Despite many research, the ideal storage of serum and plasma specimens for long time in human biobanks remains to determine. Previous report showed that time and temperature has the major effect on the analyte stability.

### Table 1: Characteristics of blood derived biospecimens in biobanks

<table>
<thead>
<tr>
<th>Bio-specimen</th>
<th>The characteristics and usage of biospecimen</th>
<th>Storage protocols</th>
<th>Effects of storage variations on specimen stability</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole blood</td>
<td>Collected in EDTA tubes</td>
<td>Have to be frozen at −80 °C if DNA extraction cannot be carry out immediately</td>
<td>Stable for years to at −80 °C</td>
<td>24,25</td>
</tr>
<tr>
<td></td>
<td>Main source for germline DNA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Use in cancer research</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buffy coat</td>
<td>WBC portion of anticoagulated blood</td>
<td>Ideally, storage at −150 °C for RNA stability</td>
<td>5′ RNA tags in late handling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Main source for mitochondrial and germline DNA, cellular RNAs and proteins.</td>
<td>Storage at −80 °C preserve DNA and protein for years</td>
<td>Stability of RNA at −80 °C for 5 years or more</td>
<td>6,20-32</td>
</tr>
<tr>
<td>Genomic DNA</td>
<td>Purified from whole blood or buffy coat</td>
<td>Should be quantified, aliquot and keep at 4 °C to avoid freeze/thaw</td>
<td>Stable at 4 °C for a number of weeks, at −20 °C for months, and for years at −80 °C</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Vital in SNP and Copy-Number variants analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Epigenetic research</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cellular RNAs</td>
<td>Purified from buffy coat in the forms of tRNA, mRNA, miRNA</td>
<td>Ought to be quantified, aliquot and keep at −80 °C</td>
<td>Labile and degrades rapidly at temperatures greater than −80 °C. microRNAs (miRNAs) is stable at −80 °C</td>
<td>30,33</td>
</tr>
<tr>
<td>RBC</td>
<td>Provided from anticoagulated blood</td>
<td>Should be aliquot and keep at −80 °C</td>
<td>Stable for years to decades at −80 °C</td>
<td>22</td>
</tr>
</tbody>
</table>

Mohamadkhani et al.
Freeze-drying increase maximum storage time and the numbers of time to brought samples to room temperature. However, storage at -80°C or liquid nitrogen showed the adequate stability of different biomolecules in serum and plasma. Repeated freeze/thaw essentially need to avoid because of degradation of proteins and nucleic acids, however sometimes is unavoidable due to a limited amount of aliquots. Most lipid metabolites concentration in serum or plasma decreases by freeze-thawing however recent study prove that in two freeze-thaw cycles from 159 investigated metabolites most are stable except for methionine sulfoxide.\textsuperscript{54,55} Table 2 describes briefly the best condition for storage and potential use of serum and plasma.

**CONCLUSIONS**

In summary, the scientific developments in human health that have been occurring over the past few years indicate the need for suitable biological material for the conduct of research. Biobanks support researchers’ access for knowledge generation by a large numbers of biological well documented samples that selected with appropriate scientific criteria. The worth of biobanks will depend on the management of data related to phenotyping of donors and processing and storage of biospecimen. The potential of blood for biobanking appears to be excellent. It includes a variety of blood derivative biospecimens such as serum, plasma, buffy coat and red blood cells. The obtaining, storage, and quality valuation of these biospecimens need the standards operating procedures for further studies that has been discussed in this review.

**CONFLICT OF INTEREST**

The authors declare no conflict of interest related to this work.

**REFERENCES**

12. Chaingneau C, Cabioch T, Beaumont K, Betsou F. Serum biobank certification and the establishment of quality controls for biological fluids: examples of serum biomarker


