# BANKING ON BIOLOGY

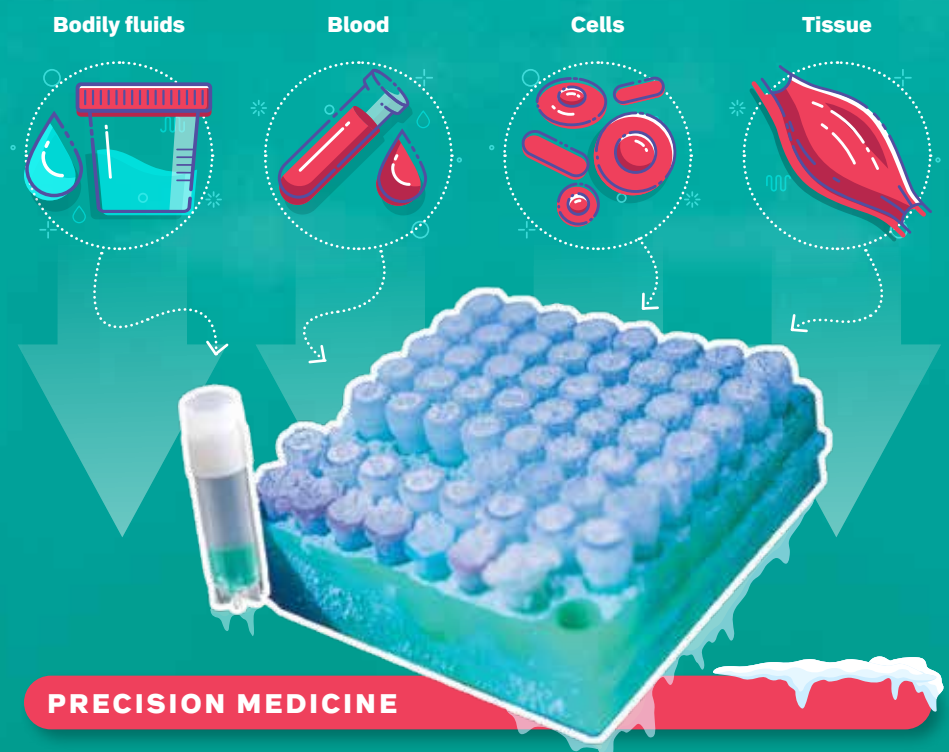
**Biobanking gives scientists access to thousands of biological samples, moving precision medicine one step closer to reality.**

## Biobanking

First coined in 1996, the term biobanking refers to collecting and storing human biological material.[1] Over 30 years, biobanks have evolved from mere storage facilities to valuable resources for population-level multifactorial data. Today, there are more than 120 biobanks worldwide[2] that house thousands of biological samples from patients with different ethnicities, lifestyles, and medical histories. Biobanks provide an unmatched resource for researchers to interrogate disease mechanisms, genetic predisposition and susceptibility to disease, drug tolerance, the influence of the environment on gene expression, and more.

## Sample Collection

Scientists have been biobanking for centuries. Every time a scientist collects a biological specimen and stores the excess unused sample in sub-zero temperatures, they are biobanking. In fact, the first biobanks started as university-based repositories, established for specific research projects. Biobanked tissues are often collected from autopsies, clinical trials, surgeries, or pathology tests.[1]




## PRECISION MEDICINE

The evolution of biobanking from small university-based repositories to autonomous organizations coincided with the completion of the human genome project in the 2000s. Physicians and scientists began to recognize the potential of biobanks for untapped insight into rare genetic variants that could only be identified with access to population-level data. The potential of biobanks to provide insight into human disease grew with the dawn of -omics-based science.[1]

## Types of Banks[1]

1 **Disease-based biobanks** collect and store biological samples from patients with specific diseases, most often from clinical trials or laboratory tests. The goal is to understand the etiology behind human illness and facilitate the discovery of therapeutic strategies.

2 **Population-based biobanks** accumulate biological specimens donated by volunteers without specific inclusion or exclusion criteria. These banks provide a population database for assessing genetic variability.

3 **Virtual biobanks** serve as electronic information repositories. Facilitated by the rise in documenting clinical data electronically, scientists can access a range of patient medical information, including date of disease onset, comorbidities, and gene-expression data without touching a biological sample.

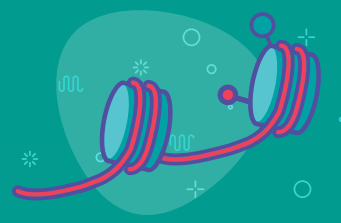
4 **Commercial biobanks** store biological samples upon request from patients for a price. These biobanks store and reserve samples in the event that the patient or a close family member becomes ill with a disease, such as leukemia, which may be treated with a closely matched genetic sample.

**Genomics / Transcriptomics**
Large-scale genomic sequencing of biobanked tumor tissue facilitated the development of the Cancer Genome Atlas, a comprehensive reference of cancer-causing mutations and genomic profiles.[3]

**Proteomics**
Using electronic health records, researchers identified patients with and without heart failure and collected their biobanked plasma samples. They then used these discarded clinical specimens to identify two protein biomarkers that robustly predicted heart failure.[4]

**Metabolomics**
Researchers analyzed more than 100,000 blood samples biobanked prior to the COVID-19 pandemic. Using high-throughput metabolomics, they identified a molecular signature that predicted which patients would develop severe COVID-19 symptoms.[5]

**Protein-protein interactions**
Using a population-based biobank, researchers identified core genes based on protein-protein interactions for a number of diseases including breast cancer, Alzheimer's disease, and type 2 diabetes.[6]



**Epigenetics**
Researchers discovered a distinct pattern of PSEN1 gene methylation in patients with Alzheimer's disease using biobanked post-mortem brain tissue. This biomarker may help assess the influence of environmental triggers such as lifestyle and nutrition on neurodegeneration.[7]

## From A Dark Past To Bright Future

One of the earliest documented examples of biobanking occurred in 1951 when medical staff at Johns Hopkins Hospital collected cervical cancer cells from Henrietta Lacks post-surgery without her consent and used them for research purposes. Today, biobanks require signed permission from patients to store and conduct research on biological material.

The majority of samples stored in biobanks are from Caucasian donors. In population-based or commercial biobanks, donors are disproportionally wealthier and healthier.[2] Lack of sample diversity inadequately represents the population and impedes scientists' ability to find genetic variants among all ethnic groups. In one example, scientists who used more diverse biobanks identified a significant association between a genetic variant in the transthyretin gene and a treatable form of heart failure.[8]

## References


1. L. Coppola, et. al., "Biobanking in health care: evolution and future directions," *Journal of Translational Medicine*, 17:172, 2019.
2. Nature, "Population matters: Biobanks accelerate geno-pheno discoveries," *Nature Research*, accessed Dec. 2, 2020, https://go.nature.com/2VrsRAj.
3. J. N. Weinstein, et. al., "The Cancer Genome Atlas Pan-Cancer analysis project," *Nature Genetics*, 45:1113–1120, 2013.
4. Q. S. Wells, et. al., "Accelerating biomarker discovery through electronic health records, automated biobanking, and proteomics," *Journal of the American College of Cardiology*, 73(17):2195-2205, 2019.
5. Nightingale Health UK Biobank Initiative, et. al., "Blood biomarker score identifies individuals at high risk for severe COVID-19 a decade prior to diagnosis: metabolic profiling of 105,000 adults in the UK Biobank," *medRxiv*, 2020.
6. A. Ratnakumar, et. al., "Protein-Protein interactions uncover candidate 'core genes' within omnigenic disease networks," *PLOS Genetics*, 2020.
7. N. Monti, et. al., "CpG and non-CpG Presenilin1 methylation pattern in course of neurodevelopment and neurodegeneration is associated with gene expression in human and murine brain," *Epigenetics*, 15(8):781-799, 2020.
8. S. M. Damrauer, et. al., "Association of the V122I hereditary transthyretin amyloidosis genetic variant with heart failure among individuals of African or Hispanic/Latino ancestry," *JAMA*, 322(22):2191-2202, 2019.