Front. Biol. 2018, 13(4): 287–292
https://doi.org/10.1007/s11515-018-1505-3

REVIEW

Biobanking and omics

David T. Harris (✉)

Executive Director, Biorepository, University of Arizona, PO Box 245221, AHSC 6122, Tucson, AZ 85724, USA

BACKGROUND: The "Era of Big Data" and "Precision Medicine" is now upon us. That is, interrogation of large data sets obtained from groups of similar patients or from the patient themselves over time will now hypothetically permit therapies to be designed to provide maximal efficacy with minimal side effects. However, such discoveries depend upon recruitment of very large numbers of subjects (tens of thousands) along with their associated biospecimens and medical records. When considering the establishment of a biobank or the refocusing of an existing repository for the purpose of "omics" research (i.e., genomics, metabolomics, proteomics, microbiomics, etc.) and/or precision medicine, there are a number of considerations to ponder. Each of these facets is discussed.
OBJECTIVE: The objective of this review is to describe best practices for the establishment and operations of a biobank that will be used for omics (genomics, proteomics, metabolomics, microbiomics) analyses based on published literature and our own practical experiences.
METHODS: We describe the most commonly described approaches to a variety of biobanking issues, including our own practical experiences over the past 5 years.
RESULTS: Based on the particular biobanking situation and downstream application, we have described best practices based on the literature and own experience, taking into consideration ease of application and costs.
CONCLUSIONS: The banking of various types of clinical biospecimens has many valuable uses but often depends on overall costs versus sample utility. In addition, specimen flexibility is important but is influenced by the ease or difficulty of the application. It is always preferable to collect and stored a biospecimen in a format that allows for multiple types of downstream analyses, but that often requires additional expertise, equipment and reagents that can increase overall costs. We have described the methodologies most successfully applied to many situations.

Keywords biobanking, genomics, metabolomics, microbiomics, methodology, precision medicine, big data

Introduction

The "Era of Big Data" and "Precision Medicine" is now upon us. That is, interrogation of large data sets obtained from groups of similar patients or from the patient themselves over time will now hypothetically permit therapies to be designed to provide maximal efficacy with minimal side effects. However, such discoveries depend upon recruitment of very large numbers of subjects (tens of thousands) along with their associated biospecimens and medical records. Theoretically, such studies may be performed prospectively or retrospectively. Prospective studies are often difficult to conduct in that it's nearly impossible to determine beforehand all of the information that needs to be collected at the beginning of a study. That is why retrospective studies, which depend upon large banks of biospecimens, are often preferred in that it is possible at any time to examine new variables and correlate findings with more recent medical observations.

Big data and precision medicine require large and diverse (ethnically, pathologically, and otherwise) biobanks (i.e., biorepositories) containing as many different types of samples (e.g., whole blood, plasma, sera, cells, urine, biopsies, DNA, RNA, etc.) as possible from each subject/patient. In addition, each sample needs to be linked to a trove of medical information contained within searchable electronic medical records in order to ascertain the medical and biological significance of any correlation between patient disease status/outcome and biospecimen measures or findings. Not only must there be clinical annotation for each

Received May 1, 2018; accepted June 6, 2018
Correspondence: David T. Harris
E-mail: davidh@email.arizona.edu

subject/patient providing biospecimens to the biobank, but the annotation must be consistent across patients, doctors and collection sites. That is, samples need to be collected in the same manner using identical methodology in order to know if observations are spurious or not. In addition, specimens need to be classified in the same way so that discriminatory power is maintained, and so that large databases may be searched in a consistent manner. An agreed upon level of classification is also needed so that subjects are equally and sufficiently characterized, making the samples more valuable. That is, it is not enough to classify a subject as diabetic when it would be more informative and useful to classify the subject as diabetic with peripheral neuropathy that has been treated with drug X.

Biobanking has been practiced for a very long time as evidenced by the existence of well-known repositories such as the American Type Cell Collection (ATCC, established 1925), the Coriell Institute (established 1953) and the NCI Cancer Bank (established 2007). Generally these biobanks tend to be narrowly focused and offer limited access to outside investigators. That is, the biobanks tend to collect biospecimens from one type of patient or disease category (e.g., cancer) and for economic reasons to limit the types of biospecimens collected and stored (e.g., plasma versus DNA but not both), with one type of downstream analysis in mind (e.g., SNP chips analysis). In our experience it is rare to find a biobank that collects multiple types of biospecimens in multiple formats (e.g., viable cells paired with other biospecimen types) and is available to all interested parties on both the receiving and dispensing ends of the process. The biospecimens themselves are generally only available to those investigators who established the biobank and are not freely available to outside scientists, academic institutions and biotechnology companies (at least not without some compensation). Finally, when specimens are made available it is generally only with limited demographic information and without appreciable clinical annotation that might be obtained from a subject's electronic medical record. Longitudinal medical annotation is rarely available although it is unquestionably the most valuable. These observations tend to be true regardless of the size of the biobank, and may reflect both the economic and scientific limitations of the biobank. In those uncommon circumstances where biospecimens are clinically well annotated and contain multiple biospecimen types that are stored in various formats, bringing utility to the collections depends upon bioinformatics expertise and novel analytic algorithms being run on powerful computers. Otherwise, the collections are a waste of time and money. The existence of such resources is limited, and as a result the full potential of the vast majority of these biobanks is rarely realized.

When considering the establishment of a biobank or the refocusing of an existing repository for the purpose of "omics" research (i.e., genomics, metabolomics, proteomics, microbiomics, etc.) and/or precision medicine, there are a number of considerations to ponder. Each of these facets is discussed below. Due to space constraints a full discussion of each topic may not have been possible, and there may be topics that are not discussed that others may have considered significant. However, at the end of the review one should be more informed as to the logistics of biobanking and have at least a framework for establishing such a facility if one has that intention (and resources).

Logistical approaches

The most important decision that a biobank must make is to decide what type of biospecimen(s) to collect and how to store it so that it will be useful for as many types of analyses as possible, while costing as little as possible by taking as little time as possible to collect and process, while maximizing the number of samples per allocated space in the storage equipment. Unfortunately, most investigators, while realizing that they should collect, store and clinically annotate specimens from their patients do not really have a good idea of how to accomplish this task, including which specimens to collect and what downstream assays will be of value for those patients, which ultimately affects both processing and banking methodology decisions.

Blood, bodily fluids and other collections

The simplest approach to large scale biobanking is to merely obtain an additional vial of blood from patients either during their pre-surgery/clinic visit or at the time of surgery itself (provided that informed consent has been obtained). However, before making the collection one must decide what type of anti-coagulant to employ (EDTA, heparin or citrate) as that choice may negatively impact downstream assays (such as PCR which can be inhibited by heparin while Ca^{2+} chelation from EDTA can be toxic for cells), particularly for proteomics assays which often requires EDTA-plasma or clotted serum. Then, one must decide how much blood to collect (4cc or 8cc vacutainers or possibly blood spot cards requiring less than 1cc of blood) as that will impact the processing methodology to be utilized. When collecting whole blood one should realize it is composed of various leukocyte populations that can be separated and stored apart from the red blood cells (RBCs), or that the leukocytes found in banked whole blood can be sorted into individual subsets for analysis at later times. In addition, blood can be used as a source of cells to be stored viably in liquid nitrogen (LN2) allowing for decisions to be made at a later time when one's research has become more focused. Whole blood is also a valuable source of genomic material for analysis such as DNA for whole genome sequencing (WGS) or SNP assays, and RNA for exome sequencing or RNAseq applications. Histological slides can also be made from blood for later analysis, including immunohistochemical and FISH assays. Many biobanks collect whole blood and merely store DNA and/or

RNA for genetic analyses as its relatively simple, economical and requires little freezer space. Genetic materials can be collected and stored as "DNA/RNA-ready" by collecting blood into tubes filled with RNAlater or DNA lysis buffer, to be purified later when requested. Alternatively, purified genetic materials can be obtained from a biospecimen (such as blood) at the time of collection for frozen storage which requires much less freezer space (often involving only 96 well microtiter plates) and much lower storage temperatures (e.g., DNA at −20°C), but requires additional monies and personnel time.

The utility of blood spot cards (BSC) in this regard should not be overlooked for many reasons. BSC are generally stable at room temperature, often for weeks to months and easy to collect using a simple finger stick as opposed to performing a venipuncture which requires a phlebotomist and specialized vacutainers for collection. BSC can be stored long-term at lower temperatures and often do not need to be kept at LN2 temperatures. We have found it easiest to store hundreds of BSC in a recipe box on a shelf at −20°C (or −80°C) temperatures. Newer versions of BCS also stabilize RNA (not just DNA) and can be kept in the field for up to 6 months at a time (according to the manufacturer, Gentegra). BSC can be used for PCR analysis (genomic and RT) and via the Cytof methodology for phenotypic analyses such as marker expression and proteomics (Yao et al., 2014).

If one desires to bank the acellular portion of a blood collection one must determine whether to perform serum collections using red-top clot tubes versus plasma collected with heparin, citrate or EDTA vacutainer tubes. The choice of anti-coagulant depends on what is later being done with the sample, and if any of the additives inhibit the proposed measure to be performed (e.g., if heparin or EDTA inhibits an ELISA or other analytic measure). It seems that it is unwise to believe that the only real difference between these types of samples is that serum does not contain clotting factors and/or fibrinogen while plasma does, as other compounds may be present that may alter downstream measures. Finally, a sample may need heat inactivation if used in some assays (e.g., mAb screening) in that complement may still be active and alter the observed results.

Similar constraints apply when considering the collection and frozen storage of biopsy material, whether from normal or diseased tissue sources. Biopsy material can be fixed for fresh frozen paraffin blocks (FFPB), dissociated into viable cells for frozen storage in liquid nitrogen, or "dissolved" into RNAlater or Trizol to obtain genetic materials. Two points to remember. First, one must be careful to well characterize the biopsy so as to ascertain if it contains a mixture of cell types (normal and malignant, for example). Second, biopsy materials of any kind are more useful/valuable if accompanied by subject whole blood specimens for comparison.

What once was considered waste is now also a source of biobanking specimens; i.e., urine and feces. One can collect urine into a sterile specimen cup to analyze metabolites as hopefully it contains neither bacteria nor cells (under most circumstances). Urine should be centrifuged at high speed immediately upon collection to remove any solids (or cells) and the supernatant frozen at −80°C in aliquots for later utilization. One concern may be that urine can be so dilute that future analyses become difficult; so it's worthwhile noting anything significant in this regard in the patient medical record. One can collect feces to perform microbiome analysis, as well as examine for normal and malignant cells shed by the gastrointestinal tissues. One may store genomic material originating from either the prokaryotic or eukaryotic source. Feces should be collected into a stabilizing solution (e.g., RNAlater or lysis buffer) if genomic analyses are envisioned. Isolation of viable cells from feces is problematic in that it's a tedious process that requires immediate attention after collection to prevent cell degradation and cell death, and generally is not practiced. If one does indeed wish to isolate viable cells then feces may be collected into a nutrient buffer and kept at cold temperatures for later isolation and analyses.

The importance of deciding in advance how biospecimens will be collected, processed and used downstream cannot be overstated. If this decision is not known in advance then the construction and use of pre-made collection kits that cover all possible types of biospecimen collections becomes extremely important. Further, use of biobank provided collection kits insures against accidental oversight when collecting biospecimens, and provides some cold chain documentation that can be critical.

Non-invasive genomics approaches

An alternative method to the collection of blood to obtain patient genomic (DNA) material is the collection and banking of cheek swabs or saliva. Both approaches provide ample material for SNP analysis or WGS. Generally, 1-2cc of saliva or 2 Q-tips (one for each check) provides microgram quantities of DNA (both eukaryotic and prokaryotic). In addition, the microbiome distributions obtained using either swabs or saliva approaches are very similar. However, the microbiome obtained from feces, although somewhat overlapping, is very different from both saliva and cheek swabs (Segata et al., 2012). Changes in the oral microbiome can be informative and sometimes predictive of changes in intestinal microbiome populations (Zarco et al., 2012). If DNA is preferred the check swab or saliva may be collected directly into RNAlater or into DNA lysis buffer, both of which can remain at room temperatures (RT) for a week or more before processing is needed. If RNA is needed, then collection into RNAlater and processing within 24 h is more optimal. A word of caution is in order here. Although saliva and cheek swabs should be identical (at least highly representative) of genomic material obtained elsewhere in the body (e.g., blood) there may be instances where the two sources could differ (e.g., cancers of the blood or certain inflammatory states). Overall, the saliva method is generally easier and more user friendly to

implement, especially out in the field, although the cheek swab approach is not terribly onerous as long as one is careful to vigorously perform the swabbing technique. Both methods can be collected directly into stabilizing agents that allow for field collections over prolonged periods of time without access to either cold storage or freezers. An important note to remember is that with either method (saliva or swab) the collections should be performed no sooner than 30min after the last meal or drink.

If biospecimens are to be collected to bank genomic material, one has many options as how to proceed. One can collect and bank either DNA or RNA for later analysis (i.e., "genomics ready"). Biospecimens for use in DNA analyses can be collected directly into DNA lysis buffer, while RNA biospecimens (being more labile) can be collected directly into RNAlater (Mutter et al., 2004) or Trizol (Amanda et al., 2007) buffers, or even into PAX gene tubes (Qiagen, Rainen et al., 2002). To provide the greatest flexibility we recommend that biospecimens be collected into RNAlater which is suitable for both downstream DNA and RNA isolation and analyses. Further, we recommend that biospecimens be collected into RNAlater buffer, rather than isolate the genomic material (DNA or RNA) itself which is more economical and less time intensive. As most biospecimens may never be used, and as one may not know if DNA or RNA will be required, it seems more efficacious to collect into "genome-ready" tubes and isolate the genomic material later on when the downstream application if known. At that time costs may be less or it may be possible to pass costs on to another investigator through collaborative endeavors. This approach also requires fewer personnel and thus can be very economical. The main advantage of processing a biospecimen to its final genomic content is that it requires less frozen storage space and may (or may not) be less labile over prolonged periods of time. We avoid the use of Trizol as it makes later DNA isolation difficult (if not impossible). We also avoid the use of PAX gene tubes as it requires even more frozen storage space (upright 8cc tubes vs. 2cc cryovials), requires additional physical manipulations, and are up to 10X more expensive than other methods. If one knows in advance that only DNA banking will be required it is possible to construct homemade collection kits using vacutainers filled with DNA lysis buffer Strauss, 1998), which is very inexpensive (pennies per tube). DNA requires no rush to process as it is very stable at most temperatures over weeks to months, but RNA should really be stabilized quickly as degradation can occur rapidly and may skew one's results. In addition, some thought needs to be given to the method of collection of the biospecimen so that when the time does come for genomic processing and analysis, the processing to obtain the genetic material may be performed with the use of robots (much less expensive) in a format where the genomic concentration is sufficiently high for downstream analysis. Downstream analyses may include microchips, exome analysis, SNP analysis, RNAseq or WGS. SNP analysis generally is the least expensive often costing only $50-$100/sample for an analysis of greater than 750,000 SNPs. WGS and Exome analyses often costs more than 10X that amount per sample which often limits the scope of the analyses.

Microbiomics collections

Recently, more and more interest has focused on the role of the microbiome in overall health. Studies implicate the microbiome as playing a critical role in inflammatory processes, in patient response to chemotherapy and immunotherapy, infectious disease as well as a determining factor in the development of autoimmune and cardiovascular diseases (Cho and Blaser, 2012). The microbiome may be assessed by sequencing the 16S RNA gene unique to prokaryotic organisms (Langille et al., 2013), and commercial methods/devices are currently available (e.g., MiSeq from Illumina), generally for less than $30/sample. The microbiome may be assessed by collection of biospecimens originating from a variety of anatomical locations, including feces, various mucosal surfaces, the skin and the oral cavity. Collection and banking of feces was discussed above as well as the collection of the oral microbiome. The latter approach is non-invasive and very amenable to implementation for wide-scale, in-the-field collections. The former approach (GI microbiome) is a source with more literature documentation in terms of its effects on health but is problematic in terms of field collections. Although the microbiomes from these two locations overlap, the results are very distinct and may have different relationships with overall health and disease. 1-2cc of saliva or 2 buccal swabs generally provides more than sufficient amounts of DNA and RNA for any downstream microbiome analysis. Fecal collections generally are not limiting but it is unclear if similar results might be obtained using rectal swabs (Q-tips) independent of a bowel movement compared to intact feces. Our preference is to collect saliva into RNAlater which also may be applied to buccal swabs and feces. Downstream processing of the banked microbiome samples is somewhat different than for eukaryotic biospecimens in that the tough outer cell wall of the prokaryote needs to be taken into consideration, often requiring the use of "beater-balls" before DNA can be isolated (Griffiths et al., 2000).

Metabolomics collections

A more recent development in the "omics" field is that of metabolomics in which one assesses the health of an organism via an analysis of its metabolism (Gowda and Djukovic, 2014). These analyses can be performed at the single cell level or at the level of the intact/entire organism. Biospecimen banking usually consists of the collection of blood or other bodily fluid (e.g., saliva) that should be immediately stabilized followed by immediate processing to halt any underlying metabolic processes. Generally, this could

mean specimen collection into EDTA or RNAlater, followed by processing into 80% methanol: 20% water which is then stored at −80°C. Two issues need to be considered when banking for later downstream metabolomics assays. One, analyses can be performed on the intact cell (reflecting internal metabolomics) or on cell supernatants (reflecting excreted metabolomics). The decision as to which to assess seems to be a personal one. And two, any relationships between metabolomics results obtained with blood versus any other source (e.g., saliva) is unclear at the present time. Finally, depending on the source material the results may reflect contributions from prokaryotic organisms (e.g., saliva, mucosal tissues or skin).

"OMICS" applications

Once samples are collected, processed and banked, and annotated with clinical data (possibly extracted over varying periods of time), what applications can then be performed? The majority of past and even present studies have concentrated on genomics. However, genomics may imply NGS or SNPs or PCR analysis of selected genes. The analysis of choice often depends on budget, time and ingenuity. Genomics analysis can reveal differences in gene copy number or genetic variants, but only when examining a limited number of genes. WGS and SNPs can correlate genetics with diseases including predisposition and susceptibility. Annotated clinical specimens of any type can be used for big data analyses (i.e., data mining) with an eye toward precision medicine (personalized medical decisions). Note that all but the simple (and limited) gene number analyses require the assistance of trained IT personnel including bioinformaticians.

Although not as commonly applied, but of more and more interest these days, is the application of microbiomics and metabolomics to banked biospecimens, particularly if clinical annotation can be applied to the samples and correlated with health status. Differences in microbiomics are now known to influence levels of (chronic) inflammation, changes in immune function, and development of autoimmune and cardiovascular disease (Börnigen et al., 2013; Young, 2017). In addition, microbial shifts in composition may influence obesity independent of diet and may even determine if chemotherapy is successful (Parekh et al., 2015; Gopalakrishnan et al., 2018). Finally, if performed correctly, one can also measure metabolomics from the same samples used for genomics analyses. Results from those analyses can be correlated to general subject health and genomic results, as well as disease (Gu et al., 2015), and may indicate if there are molecular or post-transcriptional alterations that could be responsible for the observed correlations. In addition, clinical annotation of the metabolic specimens may reveal novel biomarkers that may be predictive of disease or response to therapy.

Conclusions/best practices

The banking of various types of clinical biospecimens has many valuable uses but often depends on overall costs versus sample utility. In addition, specimen flexibility is important but is influenced by the ease or difficulty of the application. It is always preferable to collect and stored a biospecimen in a format that allows for multiple types of downstream analyses, but that often requires additional expertise, equipment and reagents that can increase overall costs. In our experience we have found the following recommendations to be successfully applied to many situations.

Banking genomic material is the most often practiced form of biospecimen storage. It is simple, inexpensive and there are now robotic platforms designed just for this purpose. Biospecimen collection can be done cheaply without special needs for sample handling before processing, and sample storage can be carried out at −180°C (or even −20/-40°C as a specimen on a blood spot card) instead of using LN2 dewers, which allows for more samples stored per dollar with less physical space commitment. Banked DNA samples can be used for PCR, WGS or SNP chip analyses with good reproducibility between samples and from one time of collection to a later collection time. Cells (e.g., blood or biopsy material) collected in either DNA lysis buffer or RNAlater can be banked for less than $1/sample spent on consumables. For pennies per sample one can spot 1ml of blood onto BSC as an alternative, making using of standard FTA cards (used for baby heel sticks) or newer treated paper cards (Gentegra). Although the costs of performing WGS is still relatively expensive ($500 or more per sample) even if conducted on a large scale, the expense of limited PCR analyses is quite reasonable (approx.. $200 for 80 genes in a 96 well kit format). Furthermore, analysis of 720k SNPs is available from multiple commercial entities for approx.. $50-100/sample.

In our experience, collection of biosamples in lysis buffer (Tris-EDTA-sucrose-SDS) provides the greatest flexibility for downstream analyses aside from exome analyses (in which case collection in RNAlater may be preferred). Not only are the samples amenable to the above described genomic analyses but the samples are also suitable for microbiome analyses, although it's worth noting that such samples may need additional processing to lyse bacterial cell walls (e.g., use of "beater-balls" and heating to 70°C). It is important to emphasize that one must be consistent in whatever methodology is chosen in order to compare results (such as bacterial family frequency) across different subjects. Generally we have found that basic microbiome analyses can be performed for $10-30/sample depending on whether one has an academic core facility onsite or uses a commercial entity. Most microbiome analyses are performed using feces as the source, although the approach can also be successfully applied to saliva and other bodily fluids. Blood by nature should be sterile and only be representative of the eukaryotic

subject while other source material may include prokaryotic contributions.

Although it may seem counter-intuitive, collection of biospecimens in DNA lysis buffer or RNAlater is amenable to downstream metabolomics analyses. However, these samples should be place on ice and processed within 4h to prevent metabolite degradation. Processing involves either addition of 80:20 methanol:water to the cell pellet or diluting the lysis buffer to achieve an 80:20 methanol:water ratio, followed by freezing at −80°C. Metabolomics analyses generally require very specialized equipment with increased costs. We have found that such analyses cost from $10 for 10 metabolites to $100 for a complete assay of 200 or so metabolites. Metabolomics assays are amenable to saliva, blood or cellular samples although the inter-relationship between the source material is not yet known.

As a final note it is extremely important to save all data and associated genomic sequences (e.g., WGS, SNP, metabolomics, microbiomics) in an online format that can be utilized for future data mining and analyses so as to eliminate the need for further sampling. Of course, as new assays become available it is also worthwhile to have additional aliquots of one's samples banked to also avoid resampling.

Compliance with ethics guidelines

Authors declare that they have no conflict of interest. This article contains studies with human subjects performed by the author under IRB approval.

References

Börnigen D, Morgan X C, Franzosa E A, Ren B, Xavier R J, Garrett W S, Huttenhower C (2013). Functional profiling of the gut microbiome in disease-associated inflammation. Genome Med, 5(7): 65

Cho I, Blaser M J (2012). The human microbiome: at the interface of health and disease. Nat Rev Genet, 13(4): 260–270

Gopalakrishnan V, Spencer C N, Nezi L, Reuben A, Andrews M C, Karpinets T V, Prieto P A, Vicente D, Hoffman K, Wei S C, Cogdill A P, Zhao L, Hudgens C W, Hutchinson D S, Manzo T, Petaccia de Macedo M, Cotechini T, Kumar T, Chen W S, Reddy S M, Szczepaniak Sloane R, Galloway-Pena J, Jiang H, Chen P L, Shpall E J, Rezvani K, Alousi A M, Chemaly R F, Shelburne S, Vence L M, Okhuysen P C, Jensen V B, Swennes A G, McAllister F, Marcelo Riquelme Sanchez E, Zhang Y, Le Chatelier E, Zitvogel L, Pons N, Austin-Breneman J L, Haydu L E, Burton E M, Gardner J M, Sirmans E, Hu J, Lazar A J, Tsujikawa T, Diab A, Tawbi H, Glitza I C, Hwu W J, Patel S P, Woodman S E, Amaria R N, Davies M A, Gershenwald J E, Hwu P, Lee J E, Zhang J, Coussens L M, Cooper Z A, Futreal P A, Daniel C R, Ajami N J, Petrosino J F, Tetzlaff M T, Sharma P, Allison J P, Jenq R R, Wargo J A (2018). Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. Science, 359(6371): 97–103

Gowda G A N, Djukovic D (2014). Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. In: Raftery D. (eds) Mass Spectrometry in Metabolomics. Methods in Molecular Biology (Methods and Protocols), vol 1198. Humana Press, New York, NY.

Griffiths R I, Whiteley A S, O'Donnell A G, Bailey M J (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. Appl Environ Microbiol, 66(12): 5488–5491

Gu H, Zhang P, Zhu J, Raftery D (2015). Globally optimized targeted mass spectrometry: Reliable metabolomics analysis with broad coverage. Anal Chem, 87(24): 12355–12362

Hummon A B, Lim S R, Difilippantonio M J, Ried T (2007). Isolation and solubilization of proteins after TRIzol extraction of RNA and DNA from patient material following prolonged storage. Biotechniques, 42(4): 467–470, 472

Langille M G I, Zaneveld J, Caporaso J G, McDonald D, Knights D, Reyes J A, Clemente J C, Burkepile D E, Vega Thurber R L, Knight R, Beiko R G, Huttenhower C (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol, 31(9): 814–821

Mutter G L, Zahrieh D, Liu C, Neuberg D, Finkelstein D, Baker H E, Warrington J A (2004). Comparison of frozen and RNALater solid tissue storage methods for use in RNA expression microarrays. BMC Genomics, 5(1): 88

Parekh P J, Balart L A, Johnson D A (2015). The Influence of the Gut Microbiome on Obesity, Metabolic Syndrome and Gastrointestinal Disease. Clin Transl Gastroenterol, 6(6): e91–e102

Rainen L, Oelmueller U, Jurgensen S, Wyrich R, Ballas C, Schram J, Herdman C, Bankaitis-Davis D, Nicholls N, Trollinger D, Tryon V (2002). Stabilization of mRNA expression in whole blood samples. Clin Chem, 48(11): 1883–1890

Segata N, Haake S K, Mannon P, Lemon K P, Waldron L, Gevers D, Huttenhower C, Izard J (2012). Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol, 13(6): R42

Strauss W M (1998) . Preparation of Genomic DNA from Mammalian Tissue. Curr Prot Mol Biol, 42: 2.2.1–2.2.3

Yao Y, Liu R, Shin M S, Trentalange M, Allore H, Nassar A, Kang I, Pober J S, Montgomery R R (2014). CyTOF supports efficient detection of immune cell subsets from small samples. J Immunol Methods, 415(15): 1–5

Young V B (2017). The role of the microbiome in human health and disease: an introduction for clinicians. BMJ, 356: j831

Zarco M F, Vess T J, Ginsburg G S (2012). The oral microbiome in health and disease and the potential impact on personalized dental medicine. Oral Dis, 18(2): 109–120