Multi-Omics Profiling for Health

Authors

Mohan Babu, and Michael Snyder
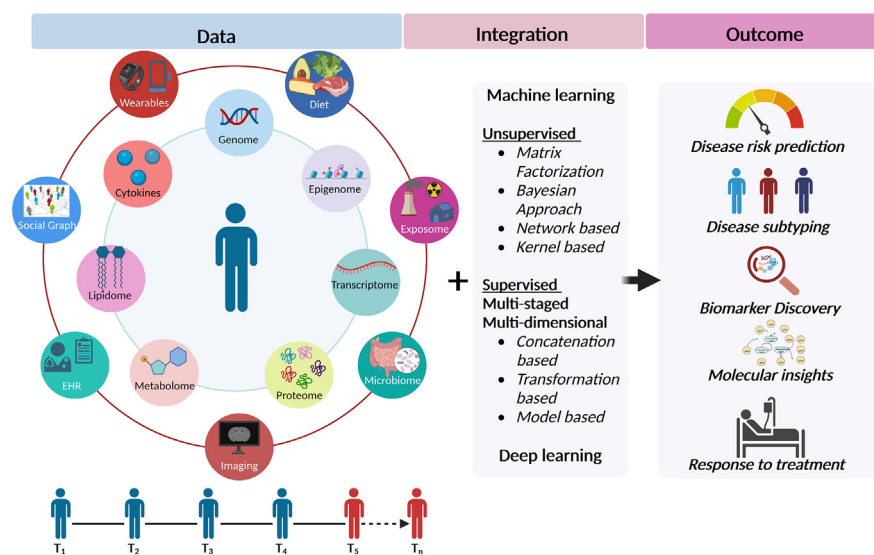
Correspondence

mpsnyder@stanford.edu

Graphical Abstract

In Brief

Current medical care focuses on treating people after they become patients rather than to preventing illness, leading to high costs in treating chronic and late-stage diseases. Additionally, a "one-size-fits all" approach to healthcare does not take into account individual differences in genetics, environment, or lifestyle factors, decreasing the number of people benefiting from interventions. Rapid advances in multi-omics enabled deep phenotyping, which profiles the interaction of multiple levels of biology over time, empowers precision health approaches and is poised to transform future healthcare.

Highlights

- Rapid advances in omics and computation are enabling multi-omics deep profiling.
- Multi-omics can capture the complex molecular interplay.
- Multi-omics is revealing insights for personalized health management.
- Multi-omics is poised to revolutionize the future healthcare.

Multi-Omics Profiling for Health

Mohan Babu[iD] and Michael Snyder*

The world has witnessed a steady rise in both non-infectious and infectious chronic diseases, prompting a cross-disciplinary approach to understand and treating disease. Current medical care focuses on treating people after they become patients rather than preventing illness, leading to high costs in treating chronic and late-stage diseases. Additionally, a "one-size-fits all" approach to health care does not take into account individual differences in genetics, environment, or lifestyle factors, decreasing the number of people benefiting from interventions. Rapid advances in omics technologies and progress in computational capabilities have led to the development of multi-omics deep phenotyping, which profiles the interaction of multiple levels of biology over time and empowers precision health approaches. This review highlights current and emerging multi-omics modalities for precision health and discusses applications in the following areas: genetic variation, cardio-metabolic diseases, cancer, infectious diseases, organ transplantation, pregnancy, and longevity/aging. We will briefly discuss the potential of multi-omics approaches in disentangling host-microbe and host-environmental interactions. We will touch on emerging areas of electronic health record and clinical imaging integration with muti-omics for precision health. Finally, we will briefly discuss the challenges in the clinical implementation of multi-omics and its future prospects.

The whole is greater than the sum of its parts.
  Aristotle

Current medical care primarily focuses on treating patients after the development of illness rather than preventing it, leading to high costs in treating chronic and late-stage diseases. Additionally, common "one-size-fits all" approaches do not take into account individual differences in genetics, environment, or lifestyle factors. This limits the number of people benefiting from known and new interventions. Omics techniques are comprehensive assessments of different classes of biological molecules, such as RNA or metabolites, that have revolutionized modern medicine by advancing our understanding of molecular complexity in health and disease (1). Individual omics approaches, such as genetic sequencing of cancers, are increasingly used in clinical settings and have greatly facilitated disease diagnosis and the identification of biomarkers to track disease or recommend effective treatments

(2–5). However, individual omics data for only one type of biology is largely correlative in nature and cannot capture the complexity of molecular events and their interactions. For example, genome-wide association studies (GWAS) have identified thousands of risk loci for several diseases, yet the causal gene is often not identified, limiting the clinical utility of such findings (6). Combining transcriptomics and proteomics or other omics can provide functional information that cannot be captured by genomics alone, enabling a new understanding of the molecular complexity underlying disease.

Advances in different omics technologies, such as proteomics and metabolomics, and computing capabilities, have recently enabled novel integration of different omics data, called multi-omics, to capture the complex molecular interplay of health and disease by combining the power of individual data types (7). Since complex diseases often develop gradually over time and show incredible heterogeneity between individuals, longitudinal sampling and integrative multi-omics analysis enable deep phenotyping of individuals across the health-to-disease trajectory to unlock precision health approaches for earlier and/or more effective intervention. Precision health aims to predict, prevent, and cure disease more precisely by taking into account each individual's genetics, environment, and lifestyle factors in contrast to the "one-size-fits all" and reactive approach of traditional medicine. Additionally, by enabling more effective treatment for each patient based on their precise subtyping, this approach could also improve healthcare efficiency and quality. Here, we provide a brief overview of emerging omics technologies and their multi-omics applications in precision health areas including cardio-metabolic diseases, cancer, pregnancy, and longevity. Throughout, we discuss the advantages of multi-omics, where one omics technology can complement the shortcomings of another to provide a holistic view of molecular complexity. Finally, we briefly review emerging omics frontiers, discuss challenges in the clinical implementation of multi-omics, and highlight future prospects.

OMICS (R)EVOLUTION

Modern biology and medicine are being propelled by ongoing advancements in DNA sequencing, mass

From the Department of Genetics, Stanford University School of Medicine, Stanford, California, USA
*For correspondence: Michael Snyder, mpsnyder@stanford.edu.

spectrometry, wearable technologies, and big data computational approaches (Fig. 1). Each of these techniques has been progressing rapidly in the past 20 years:

Genomics

Genomics is the most mature of all the omics technologies and refers to the study of whole genome sequences and DNA sequence variants therein, including single nucleotide variations, insertion-deletions, structural variations, and copy number alterations. Genomics analysis has seen dramatic progress since the discovery of "Sanger sequencing" of DNA in 1977 (8). With the advent of next-generation sequencing

(NGS) technologies in the past couple of decades, genomes can now be analyzed faster, cheaper, and in a high-throughput manner. Genome sequencing costs have plummeted steadily from billions of dollars to sequence the initial human genome in 2000 to just $100 per genome in 2022 (Ultima Genomics). While it took 13 years to sequence the first human genome, patients' genomes can now be sequenced in as few as 5 h with long-read sequencing techniques, speeding up genetic diagnosis and treatment (9). A prime example of NGS approaches to genomics is high-throughput and massive paired-end mapping, which was used to reveal that genomic structural variation among humans is much larger
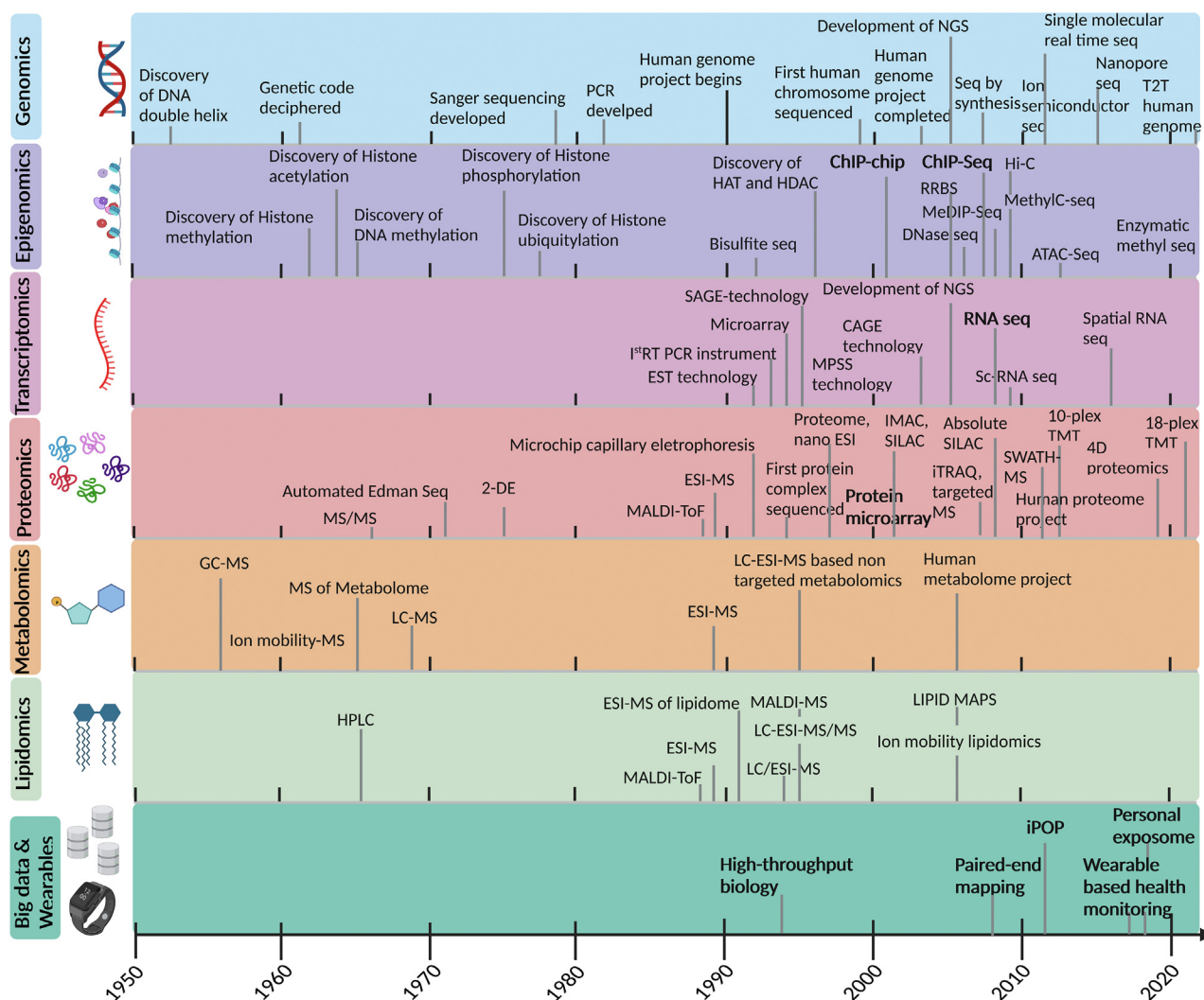

FIG. 1. Timeline of major technological developments and milestones in different Omics analysis. Technologies developed in the lab are highlighted in bold. 2-DE, 2-Dimensional Electrophoresis; ATAC, Assay for Transposase-Accessible Chromatin; CAGE, Cap Analysis Gene Expression; DNMT1, DNA Methyl Transferase one; ESI, Electrospray Ionization; EST, Expressed Sequence Tags; GC, Gas Chromatography; HAT, Histone Acetyl Transferase; HDAC, Histone Deacetylase; Hi-C, High-throughput Chromosome Conformation Capture; HPLC, High-Performance Liquid Chromatography; IMAC, Immobilized Metal Affinity Chromatography; iTRAQ, Isobaric Tags for Relative and Absolute Quantitation; LC, Liquid Chromatography; MALDI, Matrix-Assisted Laser Desorption/Ionization; MeDIP, Methylated DNA Immunoprecipitation; MPSS, Massively Parallel Signature Sequencing; MS, Mass spectrometry; NGS, Next Generation Sequencing; PCR, Polymerase Chain Reaction; SAGE, Serial Analysis of Gene Expression; Sc-RNA, Single cell RNA; SILAC, Stable Isotope Labeling with Amino acids in Cell culture; T2T, Telomere-To-Telomere.

than initially hypothesized, paving the way for improved understanding of phenotypic variation and genetic disease (10). Genome analyses have already had a major impact on medicine with applications in the diagnosis of diseases, response to treatment, and prognosis (11), such as the identification of treatment interactions and the development of cancer drugs such as erlotinib targeted to specific genetic mutations.

Epigenomics

Epigenomics refers to the complete cataloging of chemical modifications of DNA and the histones it is wound around. The field of epigenomics began with the discovery of DNA methylation (12) and histone modifications in the 1960s (13) and was accelerated by NGS technologies. Different NGS techniques such as DNA methylation through techniques, including bisulfite sequencing (14), reduced representation bisulfite sequencing (15), methyl-seq (16), methylated DNA immunoprecipitation (17), and enzymatic methyl sequencing (18) have enabled precise mapping of genome-wide methylation patterns and other epigenetic markers that affect gene regulation. The histones around which DNA is wound are composed of dimers made up of four basic proteins, H2A, H2B, H3, and H4, which undergo myriad post-translational modification, including acetylation, methylation, phosphorylation, and sumoylation, that affect how certain genes are turned on or off. Early technologies to analyze histone modifications used immunoprecipitation with antibodies against specific histone modification sites on DNA but were limited by expense and throughput. The development of ChIP-chip, where genomic DNA sites enriched in specific modifications were identified by DNA hybridization to a microarray (ChIP-chip) (19–23), improved epigenomics studies. This approach, however, was noisy and expensive to apply genome-wide. The advent of NGS further enabled high-resolution genome-wide mapping (ChIP-seq) of chromatin modifications and the location of the bound regions (24–26). Epigenomics, being an integrator of genome and environment, has also found wider application in disease diagnosis, prognosis, and therapy (27).

Transcriptomics

Transcriptomics measures the complete set of RNA transcripts and their quantity in a cell or a population of cells as a read-out of cell state (28). Progress in transcriptomics has paralleled rapid developments in NGS and analysis technologies (28). Initial transcriptomics approaches used both hybridization-based (Microarray) and sequencing-based approaches for the quantification of transcripts. However, microarray technologies could only detect genes and exons previously incorporated into the array and could not detect novel transcripts. Additionally, low-expressed genes were not detectable (sensitivity), and microarrays failed to differentiate between genes with sequence homology (specificity) (29, 30). In contrast, sequence-based approaches, such as serial analysis of gene expression or massively parallel signature

sequencing, were limited in their ability to detect all transcript isoforms and were expensive (31, 32). Several studies in 2008 reported high-throughput sequencing of the whole transcriptome, known as RNA sequencing (RNA-Seq), which revealed new parts of the genome that are transcribed while also enabling more accurate RNA quantitation, detection of transcripts with low expression, and identification of new genes, exons, and transcript isoforms at the same time (33, 34). Continued progress in transcriptomics has also revolutionized modern medicine with applications in disease diagnosis and prognosis, enabling the definition of how different genes interact in unique cell types over time (35).

Proteomics

Proteomics, the quantification of all protein identity and abundance in a sample, has similarly seen major advances in technologies and instrumentation, enabling faster, more efficient, sensitive, and accurate detection of proteins (36). Mass spectrometry (MS)-based proteomics began with the development of soft ionization techniques such as electrospray ionization (ESI) (37) and matrix-assisted laser desorption ionization (MALDI) (38) in volatilizing and ionizing proteins and peptides in the 1990s but was limited in how many proteins could be identified. The first high-throughput analysis of proteins was achieved using protein array methodologies based on prefabricated chips with specified protein detection. Despite being sensitive, this approach could not capture the entire proteome (39). From this followed multidimensional protein identification technology, which used two-dimensional liquid chromatography to separate proteins before tandem-MS analysis (40). After this followed shotgun proteomics with better sensitivity, dynamic range, molecular weight, and hydrophobicity (41). The early 21st century has witnessed significant improvements in both liquid chromatography (LC) and MS parameters, especially the higher scanning frequency and mass accuracy, enabling an era of LC-MS/MS-based "next generation proteomics". More recent labeling strategies, such as tandem mass tag (42) and isobaric tagging for relative and absolute quantitation (43), offer improved multiplexing and sensitivity to significantly reduce LC-MS analysis times and increase throughput. With rapid, robust, and high-throughput analysis, including options for multiplexing large numbers of samples, the costs of proteome analysis per sample have dropped from $3250 in 2006 to just $375 in 2021 (44). MS-based proteomics has already shown promise by successfully revealing complex and predictive biomarker signatures leading to improved clinical decision-making as well as enabling the prediction of patient trajectories *via* machine learning (45). Furthermore, recent advances in data-independent acquisition methods such as the sequential windowed acquisition of all theoretical fragment ion spectra-MS are expected to transform clinical diagnostics and prognosis through scalable and affordable proteomics (46, 47).

Metabolomics

Metabolomics refers to the study of small molecules in the body <1500 Da in mass and has similarly seen a dramatic improvement in technologies and instrumentation in the past several decades. Major metabolomics approaches include targeted metabolomics, untargeted metabolomics, fluxomics, and metabolite imaging. Targeted metabolomics aims to identify and quantify a small subset of metabolites (50–500) and is ideal for biomarker detection. Untargeted metabolomics attempts to characterize all possible number of metabolites (>10,000). Fluxomics is a branch of targeted metabolomics that monitors the movement of isotopic labels through metabolic intermediates and measures metabolite reaction rates. Metabolite imaging is an emerging field of metabolomics that involves the detection and visualization of metabolites in tissues (48). Being the substrate on which genetics, environment, microbiota, and exposome interact, metabolomics studies have propelled biomedical research with applications in biomarker discovery, disease diagnosis, and prognosis (4). Lipidomics in particular has seen significant progress with MS-based technological advances. Analysis of intact cellular lipids was greatly accelerated with the advances in ionization technologies, which were in large part fueled by the development of ESI and MALDI in the late 1980s (37, 38). Advances in MS methods have improved both resolution and mass accuracy while multiplexing has enabled exponential growth in lipidomics throughput and utility since the 1990s, offering exciting new possibilities to understand health and disease (49).

Wearables

Wearable devices (wearables) refer to any miniaturized electronic device with sensors that can be donned on the body or integrated into clothing or other body-worn accessories (50). Wearable technologies are revolutionizing biomedicine through mobile and digital health by enabling continuous, longitudinal monitoring of vital physiological parameters including heart rate, sleep, pulse oximetry, blood pressure, steps, and temperature (51). Along with multi-omics, wearable data can track transitions from health and disease at an exquisite resolution and is considered an important tool for precision health (50, 51). Recent studies have demonstrated the potential of wearables in detecting inflammation, predicting cardiometabolic health, and passively predicting atrial fibrillation (52–56). In fact, measurements from consumer smart watches could reliably predict clinical measurements of inflammation, infection, and even insulin sensitivity status. Continuous glucose monitoring could longitudinally track glucose dynamics and uncover highly personal glucotypes to provide nutritional guidance (57). More recently, smart watch-based physiological monitoring was shown to successfully detect symptomatic and pre-symptomatic COVID-19 infections (58–60). This shows promise in expanding the use of wearables in clinical applications for detecting both acute health events and for monitoring and managing chronic diseases.

APPLICATIONS IN PRECISION HEALTH

Broadly, multi-omics integrative approaches have been critical in (1) predicting disease risk, (2) disease subtyping (e.g., glucotypes, ageotypes) and classification, (3) biomarker discovery, (4) deriving biological insights, and (5) stratifying patients for therapy (e.g., mild, moderate, and severe COVID-19) among others (Fig. 2). Multi-omics integrative approaches have enabled deep phenotyping of individuals in health and disease, leading to many clinically actionable discoveries (61–63). A prime example of this is a longitudinal integrative personal omics profiling (iPOP) study performed on a 54-year-old individual at 20 time points over a 14-month period. This study characterized the transition from a healthy to an insulin-resistant state following a viral infection and uncovered extensive, dynamic changes in diverse molecular components and biological pathways during this transition, prompting lifestyle changes in the individual and quantifying their impact (64). Further biological pathway expression analysis integrating metabolomics and proteomics data in more patients was found to predict and monitor disease (65). Since then, the iPOP study has been expanded to >116 individuals for health discoveries and molecular understanding of response to perturbations including weight gain/loss, exercise, and vaccination.

In a similar vein, the Pioneer 100 Wellness Project (P100) initiated by the Institute of Systems Biology studied 108 individuals over the course of 9 months to create a personal, dense, and dynamic data cloud for each individual. Further analysis of inter-omics correlations led to the identification of putative biomarkers for cardiometabolic disease (55). The Pioneer 100 study has been expanded to 100,000 participants in 100K Wellness Project examining blood, saliva, and stool as well as other physiological and psychological measurements to capture the initiation and progression of many common diseases. A longitudinal integrated multi-omics, physiological and behavioral analysis performed on a pair of monozygotic (identical) twin astronauts (One twin on board International space station and the other on Earth) for the first time revealed the impact of long-duration space flight on human body. The molecular insights revealed pathways and mechanisms that are vulnerable to spaceflight and could serve as a guide for targeted countermeasures/monitoring during future missions (66). Longitudinal saliva multi-omics was used to monitor immune response to vaccination is emerging as a non-invasive diagnostic approach (67). Together these initial studies showcase the immense potential of multi-omics approaches for assessing health status, discovering clinically actionable insights into illness, and guiding personalized medical treatment to ultimately provide better health management.
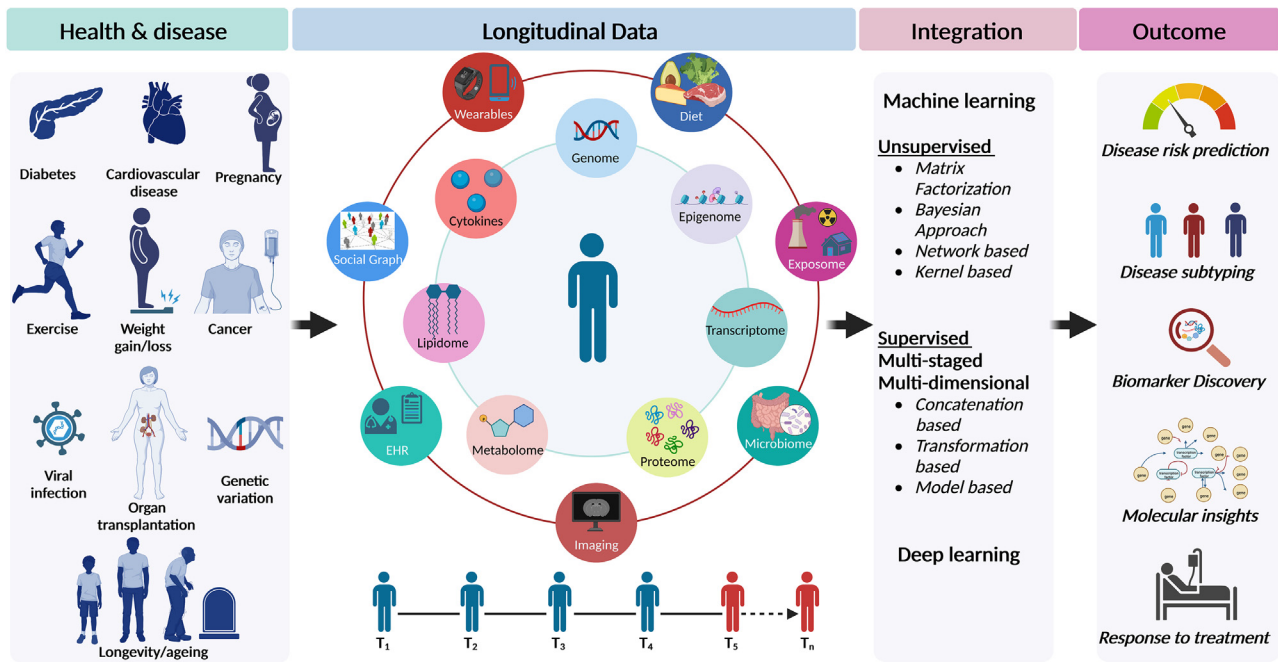
FIG. 2. **Longitudinal multi-omics and wearable data enabled deep phenotyping for precision health.** Omics and non-omics data across times (T_1-T_n) could be integrated using machine learning and deep learning approaches to predict disease risk, subtyping, biomarker discovery, molecular insights, and response to treatment among others. Figure was created using Biorender (https://biorender.com/).

GENETIC VARIATION TO DISEASE

The completion of the Human Genome Project marked the beginning of a new era in biomedical research (68). This was followed by large-scale GWAS for identifying thousands of genetic variations associated with diseases or complex traits. However, the functional relationship of these variations to patient phenotype or the translation of GWAS results to clinical applications has been lacking. Since most GWAS loci fall within non-coding regions, assigning functions to these variants has been challenging (69–71). In this regard, multi-omics integration of whole-genome sequencing (WGS) or whole-exome sequencing (WES) data with transcriptome information has been critical in identifying genes and pathways that may have a role in a particular disease. Importantly, proteomics has revealed the effect of genetic variants in conditions otherwise undetectable by RNA analysis (72). For example, through integrative analysis of ribosome profiling, RNA sequencing, and MS of lymphoblastoid cell lines from 95 ethnically diverse individuals, we discovered distinct mechanisms of gene expression variation among humans and found that genetic variants can cause changes in protein levels through effects on translation (73).

While WGS and WES can characterize ~10,000 variants per genome, computational algorithms for accurately predicting and prioritizing functional pathogenic variants have been challenging. Towards this end, Mohammadi *et al.* (74) recently developed a new method, ANEVA-DOT test, to compare the expression activity of maternal and paternal alleles to identify heterozygous DNA variants with a strong effect on gene expression in rare genetic diseases and other complex conditions. Multi-omics approaches can be utilized for the construction of gene-regulatory networks to prioritize disease-risk genes and for the prediction of drug efficacy (75). Visscher and Yang have developed a method called omics-data-based complex trait analysis to identify associations between omics data, such as DNA methylation, and complex traits while also accounting for confounding factors (76). Marioni *et al.* carried out both GWAS and epigenome-wide association studies on 92 plasma proteins with known neurological links from 750 healthy older adults and identified both genetic and epigenetic factors associated with the protein biomarkers (77). Similarly, a multi-omics approach integrating functional genomics with GWAS summary statistics identified 650 amyotrophic lateral sclerosis-associated genes that represent a fivefold increase in recovered heritability, extensive conservation, and transcriptome network changes associated with disease development. Rare variant analyses have demonstrated the functional significance of candidate genes in healthy and diseased motor neurons and brain tissues. These studies demonstrate the power of multi-omics in dissecting the genetic basis of complex diseases that were not possible with single omics approaches, opening new avenues for precision interventions (78).

CARDIO-METABOLIC DISEASE

Complex diseases including cardiovascular, metabolic disorders, and cancer evolve over time and show incredible heterogeneity among individuals. Thus, longitudinal

multi-omics integrative analysis can identify temporal molecular shifts indicative of physiological transitions between healthy to disease states (Fig. 2). A prime example of this is a longitudinal multi-omics study of weight perturbation that demonstrated activation of strong inflammatory and hypertrophic signatures in the blood associated with weight gain. Although weight loss reversed some changes, several signatures persisted, indicative of long-term physiological changes due to weight gain. Additionally, omics signatures revealed an association with insulin resistance that could serve as a novel diagnostic. Interestingly, specific biomolecules were highly individualized and stable in response to perturbations, potentially representing "personalized biomarkers" (79). In a similar vein, the prediabetes to clinical type 2 diabetes mellitus (T2DM) transition was captured by performing multi-omics on prediabetic individuals for over 4 years. This rich longitudinal data set revealed many insights into precision health: first, healthy profiles were distinct among individuals while displaying diverse patterns of intra-and/or inter-personal variability. Second, extensive host and microbial changes were found during respiratory viral infections while immunization was found to trigger potentially protective responses that are distinct from responses to respiratory viral infections. Moreover, during respiratory viral infections, insulin-resistant (IR) participants' immune signatures responded differently than insulin-sensitive (IS) participants. Third, global co-association analyses among the thousands of profiled molecules revealed specific host-microbe interactions that differed between IR and IS individuals. Lastly, this study identified early personal molecular signatures in one individual that preceded the onset of T2DM, including the inflammation markers interleukin-1 receptor agonist and high-sensitivity C-reactive protein paired with dysregulated xenobiotic-induced immune signaling. Overall, this study revealed insights into myriad pathways and responses that differ between glucose-dysregulated and healthy individuals during health and disease (80).

A recent study used deep longitudinal multi-omics profiling including emerging technologies like immunome, microbiome, and wearable monitoring of a cohort enriched for risk factors for T2DM for up to 8 years and discovered more than 67 clinically actionable health discoveries while also identifying multiple molecular pathways associated with metabolic, cardiovascular, and oncologic pathophysiology. Additionally, omics measurements could reliably predict insulin resistance, illustrating their potential to replace burdensome clinical tests (81). Similarly, multi-omics has also been applied to understand the molecular basis of hypertrophic cardiomyopathy (HCM). Comprehensive molecular analysis using transcriptome, metabolome, and lipidome profiling of myocardial samples from HCM and normal controls revealed perturbed metabolic signaling and mitochondrial dysfunction as common pathogenic mechanisms underlying HCM, highlighting potential new drug targets for attenuation of HCM (82).

CANCER

Cancer etiology is multifactorial and highly heterogenous in nature, requiring multi-modal approaches to dissect its underlying mechanisms and develop new therapies. In this direction, Liu *et al.* (83) utilized multi-omics analysis of genomic CNVs, DNA methylation, and gene expression in 256 hepatocellular carcinoma samples and identified five subgroups with distinct molecular signatures and a distinct survival rate. Kamoun *et al.* (84) performed multi-omics integrative analysis on oligodendroglial tumors to identify three subgroups of 1p/19q co-deleted gliomas. Single omics measurements have been used in melanoma prognosis prediction. Unfortunately, these approaches cannot comprehensively describe the biological processes underlying prognosis and the prognostic models developed were less accurate for clinical implementation. To this end, Jiang *et al.* (85) performed an integrative analysis of clinical variables, genomic CNVs, DNA methylation, and gene expression data from The Cancer Genome Atlas and found that integrated analysis led to models with improved prediction, with a mean C-statistic of 0.724. Although chromatin alterations are reported in several cancers, their relevance for cancer gene expression phenotypes remains unclear. Recently, multi-omics profiling of chromatin accessibility, RNA, and protein abundance of human thyroid cancer primary tumors, metastases, and patient-match normal tissue identified gene body enhancers predictive of correlated RNA and protein expression. This study demonstrates the utility of multi-omics in identifying potential targets and better understanding cancer treatments (86). In a similar vein, Zhang *et al.* (87) performed genome sequencing and proteomics analysis on high-grade ovarian serous carcinomas to unravel the influence of different gene copy-number variations on the proteome, post-translational modification levels, and clinical outcomes, providing a mechanistic link between copy-number variations and potential progression events in ovarian cancer. These studies demonstrate the power of multi-omics in understanding mechanisms of cancer and subtyping for precision therapies. Recent reviews provide an excellent and in-depth overview of multi-omics applications in cancer (88–90).

INFECTIOUS DISEASES

Multi-omics have enabled deep characterization of antibody responses to infections and vaccines. Bulk and single-cell multi-omics were instrumental in dissecting cellular responses to SARS-CoV-2 viral infection and subsequent vaccine responses during the COVID pandemic. While most investigations involved focused on analyzing one information layer at a time (*e.g.*, fluorescence-activated cell sorting) to understand the dynamics of circulating immune cells in COVID-19, Bernardes *et al.* performed longitudinal multi-omics on peripheral blood mononuclear cells (PBMCs) from patients with COVID-19 infection throughout the disease

course for a comprehensive understanding of longitudinal cellular features. Interferon-activated circulating megakaryocytes and increased erythropoiesis coincided with critical illness while megakaryocytes- and erythroid-cell-derived co-expression modules were predictive of fatal disease outcomes. This multi-omics approach demonstrated the broad cellular effects of SARS-CoV-2 infection at the epigenetic and transcriptional level beyond just phenotypic analysis of immune cells providing insights to develop biomarkers and precision treatments for patients with COVID-19 (91). A recent study using an integrated single-cell multi-omics profiling of human lungs discovered and validated over 1000 risk genes underlying severe COVID-19 across 19 cell types. Genetic risk for severe COVID-19, covering both common and rare variants, was particularly enriched in natural killer cells. Further, RefMap, a machine learning algorithm, enabled sensitive prediction of severe disease in non-elderly patients based on GWAS and single-cell omics. Individualized predictions were accurate independent of age and sex and were consistent across multiple populations and cohorts. When combined with machine learning, this single-cell multi-omics approach provided novel insights into the molecular mechanisms of severe disease, leading to new therapeutic targets and sensitive detection of at-risk individuals (92). Similarly, Sacco *et al.* applied longitudinal multi-omics (analysis of soluble biomarkers, proteomics, single-cell gene expression, and immune repertoire analysis) to identify immunopathological signatures between pediatric COVID-19 and multisystem inflammatory syndrome in children (MIS-C). Pediatric COVID-19 was characterized by robust type I interferon (IFN) responses, whereas increased levels of circulating spike protein, matrisome activation, and prominent type II IFN-dependent or NF-kB-dependent signatures were detected in MIS-C. This approach thus better defines the pathophysiology of these disorders and helps design precision therapies (93).

Multi-omics has also been used to characterize the molecular shifts between mild and moderate COVID-19. By characterizing circulating immune cell classes and plasma multi-omics profiles form two longitudinal blood draws, Su *et al.* demonstrated elevated inflammatory signaling accompanied by loss of specific classes of metabolites and metabolic processes during a shift from mild to moderate COVID-19 disease. This integrated approach revealed that moderate disease may provide the most effective setting for therapeutic intervention (94). Multi-omics approaches are also accelerating vaccine development by helping construct global maps of the complex immune responses that occur during vaccination to identify cellular and molecular correlates of vaccine efficacy (For a comprehensive overview of multi-omics approaches for precision medicine in infectious diseases, See Refs (95, 96)). MS-based proteomics, with its fast turn-around and high throughput, has been instrumental in revealing classifiers of COVID-19 infection. Recently, Messner *et al.* developed a low-cost platform (less than 10€ for consumables per sample) for ultra-high-throughput serum and plasma proteomics. In a cohort-based epidemiological study, the platform could identify 27 potential markers that revealed the severity grade of COVID-19. The platform demonstrates the power of MS-based large-scale proteomics in clinical decision support in situations needing rapid responses such as the COVID-19 pandemic (97).

ORGAN TRANSPLANTATION

Organ transplantation remains the ultimate treatment option for patients with end-stage disease with organ failure, yet mortality rates are high due to frequent rejection. This is due to a limited understanding of complex post-transplant immune adaptation mechanisms. A better understanding of donor–recipient matching and longitudinal multi-omics tracking after transplant is needed to improve rejection rates and design personalized therapies. Towards this end, Watzenboeck *et al.* combined profiling of the alveolar microbiome, cellular composition, metabolome, and lipidome in bronchoalveolar lavage samples from organ recipients and donors to identify recipient-specific and environmental factors that shape the long-term lung microbiome. The abundance of certain bacterial strains correlated with underlying lung diseases even after transplantation. By applying machine learning models to this data, they could accurately predict changes in forced expiratory volume during the first second (FEV_1, a major characteristic of lung allograft dysfunction) from multi-omics data, whereby lung microbiome composition showed a high predictive power (98).

Wigger *et al.* conducted a comprehensive multi-omics analysis of pancreatic islets obtained from metabolically profiled pancreatectomized living human donors stratified along the glycemic continuum (from normoglycemia to T2DM) and found remarkable heterogeneity in the transcriptomic and proteomic profiles in patients with diabetes compared to non-diabetic controls. Differential regulation of islet gene expression is already observed in prediabetic individuals with impaired glucose tolerance, suggesting a progressive, but disharmonic, remodeling of mature beta cells and thus challenging the current model of a linear trajectory toward precursor or transdifferentiation stages in T2DM development. Furthermore, through the integration of islet transcriptomics with preoperative plasma lipidomics, this study also defined the relative importance of gene coexpression modules and lipids that are positively or negatively associated with HbA1c levels, pointing to potential prognostic biomarkers. This approach helps define subtypes of T2DM, and biomarkers thereof, thus enabling precision approaches for the treatment of T2DM (99).

PREGNANCY

Multi-omics approaches have also been instrumental in unraveling biological signatures and transitions predictive of

pregnancy-related complications, including pre-term birth (PTB) and preeclampsia. Multi-omics modeling integrating transcriptomic, immunological, microbiomic, metabolomic, and proteomic measurements during the course of full-term pregnancy was used to measure the ability of each dataset to predict gestational age. Among the individual dataset, plasma proteomics had the strongest predictive power. Additionally, combining all datasets increased the predictive power and revealed novel interactions among different biological modalities (100). Longitudinal multi-omics (metabolome, proteome, and immunome) profiling captured a distinct molecular shift from pregnancy maintenance to pre-labor biology occurring 2 to 4 weeks before delivery. A surge in steroid hormone metabolites and interleukin-1 receptor type 4 preceded labor onset and coincided with a switch from immune activation to the regulation of inflammatory responses. This approach could help in developing blood-based methods predicting the day of labor, anchored in mechanisms shared in pre-term and full-term pregnancies (101). Multi-omics analysis combined with machine learning modeling was also used to identify early biological measurements associated with pre-term birth in five biorepository cohorts in low- and middle-income countries (102). These studies reveal the power of combining multi-omics and machine learning for developing valuable predictive tests and intervention candidates for preventing PTB.

LONGEVITY/AGING

Longitudinal multi-omics profiling was also used to reveal myriad molecular changes during aging, identifying both known and new markers, as well as distinct molecular patterns of aging in insulin-resistant as compared to insulin-sensitive individuals. Molecular pathways that changed over time in each individual suggested different aging patterns (ageotype) that may ultimately be useful in monitoring and intervening in the aging process (45). In a similar vein, Nie *et al*. utilized multi-omics data, including clinical tests, immune repertoire, targeted metabolomics, gut microbiome, physical fitness tests, and facial skin examination to estimate the biological ages of different organs to identify diversity in aging. This study revealed different aging patterns across the study population, suggesting precision interventions may be necessary to decrease the impact of aging (103). In another study, multi-omics profiling was used to understand variability in reprogramming old or young fibroblasts to induced pluripotent stem cells (iPSC) akin to ageotypes. This approach revealed that fibroblast cultures from older mice contained "activated fibroblasts" that secrete inflammatory cytokines and that the proportion of activated fibroblasts in each cell culture correlated with the reprogramming efficiency. This could help in developing personalized strategies to improve iPSC cell generation and wound healing in elderly individuals (104). These studies highlight the promise of multi-omics integrative approaches in developing personalized aging interventions.

EMERGING FRONTIERS IN PRECISION HEALTH
Host–Microbiome Interactions

The microbiome, often considered our second genome, shapes health and plays a crucial role in a plethora of diseases. Characterization of diverse microbes in healthy individuals found extensive variation in both body site habitat and between different individuals, giving rise to the concept of a "personal microbiome" (105). Microbial interactions with their human hosts change across health and disease and thus serve as a modifiable factor to manage health (106). A recent host-microbial multi-omics study demonstrated taxonomic and functional differences between insulin-resistant and insulin-sensitive individuals in various measurements, both at baseline and in response to stresses such as weight loss and respiratory viral infections (80). Along similar lines, Heintz-Buschart *et al*. (107) performed microbial multi-omics on four families with type 1 diabetes and observed intra- and inter-individual variation demonstrating a pronounced effect of family membership on the structural and functional composition of the gut microbiome. Lloyd-Price *et al*. performed an integrated host-microbial multi-omics longitudinal profiling study that provided a comprehensive view of functional dysbiosis in the gut microbiome during inflammatory bowel syndrome activity. They demonstrated that a characteristic increase in facultative anaerobes at the expense of obligate anaerobes, as well as molecular disruptions in microbial transcription, metabolite pools, and levels of antibodies in host serum, was correlated with the development of bowel inflammation. Periods of disease activity were also marked by increases in temporal variability in the microbiome with characteristic taxonomic, functional, and biochemical shifts. Finally, the integrative analysis identified microbial, biochemical, and host factors central to this dysregulation (108). Similarly, through longitudinal sampling and integrative host–microbial multi-omics, a recent study identified inflammatory bowel syndrome subtype-specific and symptom-related variations in microbial composition and function. Furthermore, purine metabolism was identified as a key host–microbial metabolic pathway as a therapeutic target for inflammatory bowel syndrome (109). Thasis *et al*. utilized host methylome, transcriptome, metabolome, and gut microbial metagenome and imaging data to quantify the global reprogramming of host biology by microbiota. They showed a tight link between the host and microbial circadian activities and further found that disruption of microbial rhythmicity abrogates normal host oscillations in the intestine and liver, influencing host diurnal fluctuations (110). Apart from host-microbial omics analysis, extra-cellular vesicles (EVs) secreted by host and microbial cells are emerging as critical players in cell-to-cell communication under various conditions. EVs are lipid bilayerd

structures containing transmembrane proteins, cytosolic proteins membrane-associated proteins, and nucleic acids. Upon release by cells, EVs can interact with adjacent or distant cells and modulate their function through signaling *via* surface contact or by transferring cargo (111). A comprehensive omics profiling of EVs thus could serve as biomarkers for a range of clinical conditions including COVID-19 (112), cancers (113), and so on. These studies demonstrate the power of integrated longitudinal host–microbial multi-omics analyses in revealing the complex interactions that shape host physiology and may be amenable to precision interventions for preventing diseases.

Host–Environmental Interactions

Genetic loci identified from GWAS have been able to explain only a small proportion of complex disease heritability/etiology, leading to "missing heritability" concept (114). Nongenetic factors including lifestyle, diet, and environmental exposures (exposome) are suggested to explain the "missing heritability" of complex diseases (115). Lifestyle factors, especially physical exercise, favorably impact overall health and protect against complex diseases including obesity, diabetes, and other cardiometabolic diseases (116–119). However, the molecular mechanisms underpinning exercise-induced benefits have not been clearly defined. Along this direction, a longitudinal multi-omics profiling study of plasma and PBMCs from 36 well-characterized volunteers, before and after a controlled bout of symptom-limited exercise, detected distinct molecular changes and an orchestrated choreography of biological processes involving energy metabolism, oxidative stress, inflammation, tissue repair, and growth factor response as well as regulatory pathways governing magnitude and duration of those responses. Interestingly, these processes were dampened, and some were even reversed, in insulin-resistant participants. Machine learning models based on multi-omics data from this study were able to predict potential blood-based biomarkers of peak oxygen consumption during exercise (120). Recently, Li *et al.* (121) discovered an exercise-induced metabolite, *N*-lactoyl-phenylalanine (Lac-Phe), as a suppressor of feeding and obesity in mouse, humans, and racehorse models of exercise and provided new insights into molecular responses to physical activity. To tap the immense potential of multi-omics, datasets with larger samples and more tissue- and disease-specific repositories are essential. In this direction, a larger consortium involving pre-clinical and clinical studies is also examining systemic response to acute and chronic exercise using multi-tissue multi-omics. This will serve as a public database to enhance our understanding of the health benefits of exercise and could provide insights into how exercise mitigates disease (122).

Diet, another lifestyle factor, has a profound impact on host physiology and exerts a "personalized effect" (123). Integrative multi-omics, wearable data, and machine learning approaches are revealing molecular insights into individuals' responses to diet, enabling improved precision nutrition approaches. For example, Zeevi *et al.* (124) created a machine learning algorithm that included blood parameters, dietary habits, anthropometrics, physical activity, and gut microbiome data that could accurately predict personalized postprandial glycemic response to real-life meals. Recently, Berry *et al.* assessed postprandial metabolic responses in ~1002 twins and unrelated healthy adults and found a large interindividual variability in blood triglycerides, glucose, and insulin. Machine learning models implementing meal composition, habitual diet, meal context, anthropometry, genetics, microbiome, clinical, and biochemical parameters could accurately predict postprandial triglyceride and glucose responses (125). A recent study used continuous glucose monitoring (CGM) to longitudinally track glucose dynamics in response to standardized meals and uncovered highly personal glucotypes unique to participants (57). In a similar vein, personalized responses to dietary fiber (arabinoxylan & inulin) supplementation were also discovered using host multi-omics, microbiome, and clinical parameters (126). These studies show the utility of integrative multi-omics, microbiome, and machine learning approaches to dissect individual responses to diet.

The exposome is another non-genetic modifier of health that includes both biological (*e.g.*, pollen, viral particles) and chemical components (*e.g.*, pollutants, disinfectants, and insecticides). This diverse repertoire of components can exert distinct biological responses through methylation, gene expression changes, microbial shifts, and inflammatory cytokine secretion (127). The external exposome can influence internal omics responses, including metabolomics, linking functional environmental changes to chronic disease (128). However, the impact of diverse environmental exposures on individuals' health is not clearly understood and thus needs large-scale efforts comparable to human genome sequencing (129). To this end, a recent study longitudinally profiled the personal exposome of 15 adult individuals for up to 890 days using a portable exposometer. Combined with deep sequencing and mass spectrometry profiling, over 2500 microbial species and 2796 putative chemical features were identified in these collected personal airborne exposures and showed highly dynamic changes in exposome composition in response to varying environments and lifestyles (130). Similarly, a recent study from Human Early Life Exposure (HELIX) project investigated the biological effects of early life exposure in a multicenter cohort of 1301 mother–child pairs and associated individual exposomes consisting of 100 chemical, physical, and lifestyle exposures assessed in pregnancy and childhood followed by multi-omics profiling in childhood. They identified 1170 associations, 249 in pregnancy and 921 in childhood, which revealed potential biological responses and sources of exposure. The methylome best captured the persistent influence of pregnancy exposures, including maternal smoking, while childhood exposures were associated with features from all omics layers, revealing novel

signatures for indoor air quality, essential trace elements, endocrine disruptors, and weather conditions (131). To better understand how the exposome shapes an individual's phenotype, a recent study used deep longitudinal personal exposome and internal multi-omics profiling and annotated thousands of chemical and biological components in the personal exposome cloud, finding a significant correlation with thousands of internal biomolecules which were cross-validated using corresponding clinical data. These results showed that agrochemicals and fungi dominated the highly diverse and dynamic personal exposome, while the bio-molecules and pathways related to the individual's immune system, kidney, and liver were most highly associated with their personal external exposome. This data-driven longitudinal monitoring study showed the depth of dynamic interactions between the personal exposomes and internal multi-omics, underlining the need for further study and tool development (132).

WEARABLE AND ELECTRONIC HEALTH RECORD DATA INTEGRATION

Electronic health records (EHR) can complement multi-omics and wearables with longitudinal clinical data, including diagnostic codes, procedure codes, lab results, physical measurements, clinical notes, and medical images. In this direction, a recent study leveraged a machine learning framework to integrate genomes, EHR data, and lifestyle factors to accurately predict the occurrence of abdominal aortic aneurysm (133). Despite its clinical utility, EHR data are usually sparse with records from discrete clinical visits. Thus, wearable-based continuous physiological monitoring and integration with other multi-omics data within EHR will be critical to speed up clinical decisions and potentially reduce medical costs. For instance, initiatives by the NIH such as the "All of US" project are building health databases collecting EHR, questionnaires, physical measurements, digital health technologies, and the collection and analysis of biospecimens of a million diverse individuals that will characterize the intersection of biology, lifestyle, and environment in health (134).

LARGE-SCALE EFFORTS ADVANCING MULTI-OMICS ENABLED PRECISION HEALTH

Multi-omics–enabled precision health is propelled by continued technical advances in human genomics, proteomics, lipidomics, and metabolomics. For instance, long-read sequencing has enabled the completion of the human genome with telomere-to-telomere sequencing (T2T Consortium) and will provide a gold standard reference for mapping genetic variation to the genome and detecting pathogenic variants (135). Similarly, the human proteome project launched by The Human Proteome Organization (HUPO) in 2010 has made enormous progress in enhancing accurate annotation of genome-encoded proteins and reached a 90.4% complete high-stringency human proteome blueprint in 2021 (136). As a part of HUPO, we quantified the relative protein levels from over 12,000 genes across 32 normal human tissues, identified tissue-specific and tissue-enriched proteins, and compared them to transcriptome data. Discordance of RNA and protein levels revealed potential sites of protein synthesis and action of secreted proteins. Most importantly, our study demonstrated that protein tissue-enrichment information can explain phenotypes of genetic diseases that cannot be obtained by transcript information alone. Furthermore, we demonstrated how understanding protein level patterns can provide insights into gene regulation, the secretome, metabolism, and human diseases (137). Similarly, the human metabolome database (HMDB) created by the Human Metabolome Project has curated detailed information on small molecule metabolites found in the human body, serving as an up-to-date reference database for metabolomics studies (138).

Multi-omics repositories are providing a rich data resource to understand health and disease at the population level (Table 1). For example, two large-scale epigenetics studies, the Encyclopedia of DNA Elements (ENCODE) project and Roadmap Epigenomics, have mapped regions of transcription, transcription factor association, histone modification, DNA methylation, and chromatin structure to delineate all functional elements encoded in the human genome (139), and develop (140) critical reference epigenomic maps of human tissues, respectively. Similarly, to understand the functional consequences of genetic variation and its impact on complex human diseases, the Genotype-Tissue Expression (GTEx) project was initiated in 2010 (141). The Enhancing GTEx (eGTEx) project was later introduced to complement gene expression phenotypes determined in the GTEx project by extending data depth and introducing new methods (142). The Cancer Genome Atlas, which includes genomic, epigenomic, transcriptomic, proteomic, and clinical data for 32 cancers, is another landmark multi-omics study that has revolutionized precision oncology (143).

ADVANCES AND OUTLOOK OF COMPUTATIONAL METHODS IN MULTI-OMICS DATA INTEGRATION

Heterogeneous and high dimensional nature of multi-omics data requires robust integrative approaches to avoid information burden from an individual data type. Several machine learning methods, including unsupervised (matrix factorization, Bayesian, network-based, and kernel-based) and supervised approaches (multi-staged, multidimensional) have been successfully applied for fast and efficient integrative analysis of multi-omics data and are commonly used in research settings as described in the previous sections (Fig. 2) (144–146). Supervised approaches relay on labeled data (train data) to learn the underlying patterns and discern similar patterns in the independent data set (test data). Supervised approaches include random forests, hidden Markov models, decision trees, support vector machines, elastic nets, and neural networks among others. Supervised approaches are

TABLE 1
Multi-omics repositories for precision medicine research

Consortium	Year of launch	Status	Sample size	Omics assays	Reference
The *FANTOM* Consortium	2000	Healthy	Variable	CAGE, RNA-Seq, RADICL-Seq	https://fantom.gsc.riken.jp/
ENCODE	2003	Healthy, Cancer	Variable	RNA-Seq, Chip-Seq, DNase-Seq, eCLIP-Seq, ChIA-PET, Hi-C, CAGE, ScRNA-Seq, ATAC-Seq	https://www.encodeproject.org/
Roadmap Epigenomics	2007	Healthy	Variable	RNA-Seq, ChIP-Seq, DNase-Seq, methylation	http://www.roadmapepigenomics.org/
1000 Genomes Project	2007	Healthy	1000	WGS, Targeted Exome sequencing	https://www.internationalgenome.org/
UK Biobank	2007	Various	500,000	Genotyping, WES, WGS	https://www.ukbiobank.ac.uk/
GTEx and eGTEx	2010	Healthy	948	WGS, WES, RNA-Seq,	https://www.gtexportal.org/home/
eQTLGen consortium	2018	Various	31,684	Whole genome, Transcriptomics	https://www.eqtlgen.org/
MoTrPAC	2019	Healthy	Variable	RNA-seq, ATAC-seq, Methyl-cap, RRBS, WGS, Proteomics, Lipidomics and Metabolomics	https://www.motrpac.org/
All of Us	2020	Healthy	1 million	Surveys, wearables, physical measurements, EHR	https://www.researchallofus.org/
COSMIC	2004	Cancer	Variable	Genomics, Epigenomics, Transcriptomics	https://cancer.sanger.ac.uk/cosmic
TCGA	2006	Cancer	20,000	Genomics, Epigenomics, Transcriptomics	https://portal.gdc.cancer.gov/
CPTAC	2011	Cancer	Variable	Copy number variation, whole genome and whole exome sequencing, DNA methylation, RNA-seq, miRNAs, global proteome, phosphoproteome, acetylome and ubiquitinome, and immune subtyping	https://cptac-data-portal.georgetown.edu/cptacPublic/
TARGET	2016	Pediatric cancers	Variable	Clinical, genomic, transcriptomic, and epigenomic data	https://ocg.cancer.gov/programs/target
ADNI	2004	Alzheimer's disease patients, mild cognitive impairment subjects, and elderly controls	Variable	Clinical, genetic, magnetic resonance imaging, and positron emission tomography imaging	https://adni.loni.usc.edu/
CommonMind	2012	Schizophrenia, bipolar disorder, and unaffected controls	1000	RNA and DNA sequencing, genotyping, epigenetics	https://www.nimhgenetics.org/resources/commonmind
PsychENCODE	2015	Neuropsychiatric disease	Variable	WGS, Transcriptomics	https://psychencode.synapse.org/
AMP-PD	2018	Alzheimer's disease, type 2 diabetes, rheumatoid arthritis, systemic lupus erythematosus and Parkinson's disease	Variable	Transcriptomics, epigenomics, whole genome sequencing, metabolomics, and proteomics	https://amp-pd.org/about

ATAC, Assay for Transposase-Accessible Chromatin; CAGE, Cap Analysis of Gene Expression; ChIA-PET, Chromatin Interaction Analysis by Paired-End Tag; Chip, Chromatin immunoprecipitation; eCLIP, enhanced Crosslinking and Immunoprecipitation; EHR, Electronic Health Record; Hi-C, High-throughput chromosome conformation capture; RADICL, RNA and DNA Interacting Complexes Ligated; RRBS, Reduced-Representation Bisulfite sequencing; ScRNA, Single-cell RNA; WGS, Whole Genome Sequencing.

ideal for the prediction of continuous tasks such as survival or pain scores and the classification of discrete outcomes such as disease/healthy status. Unsupervised approaches discern patterns in the data without the need for labeled data and in an unbiased manner. Unsupervised approaches include principal component analysis (PCA), hierarchical clustering, self-organizing maps, and k-means clustering among others. Unsupervised approaches are well suited for the discovery of disease subtypes, biomarkers, and early diagnosis of disease. A comprehensive list of tools for multi-omics data integration can be found at https://github.com/mikelove/awesome-multi-omics. Moving forward, multi-omics data will be increasingly utilized for precision medicine framework, where incorporation of deep learning, artificial intelligence, and cloud-computing systems will play a crucial role in integrative analysis, interpretation, and visualization of multi-omics, imaging, clinical, wearable, and epidemiological data. Additionally, this has the potential to provide clinicians with automated, real-time, and interpretable platforms in assisting disease diagnosis, treatment strategy, and prognosis.

CHALLENGES AND FUTURE PROSPECTS

Individual omics like WGS and WES have already entered clinics for routine genetic screening, understanding response to treatments, and discovering disease biomarkers (2–5). However, the clinical implementation of multi-omics for precision health has been challenging for several practical reasons. Firstly, omics data acquisition and analysis require specialized equipment, trained personnel, and large financial commitments. Secondly, while cost and turnaround times are rapidly declining, there are parallel challenges in data storage and analysis. Multi-omics data is often heterogenous and poses myriad challenges associated with "Big data," that is, volume, variety, velocity, and veracity. Datasets with thousands of variables come with the "curse of dimensionality" where the variance between samples becomes large and sparse, rendering clustering analysis uninformative and posing further challenges in interpreting integrated data (147). In addition, missing values, lack of samples, data complexity, class imbalance, dataset shifts, batch effects, and unavailability of some data types can pose significant challenges. In this, there is a lack of standardization for sample collection and omics data analysis. Third, the heterogenous nature of datasets requires rigorous statistical tools to integrate and interpret the results. As data complexity grows with the inclusion of wearables and EHR data, approaches for deep learning, data mining, and artificial intelligence will be necessary to integrate and interpret them. Fourth, multi-modal data complexity and scale (multi-omics data sets can easily exceed tera byte (TB) scale) require robust data management systems to ensure adequate data handling capacity, privacy, and security. Health management platforms like Personal Health Dashboard (PHD), which utilizes state-of-the-art security and

scalable technologies to provide an end-to-end solution for big biomedical data analytics both at an individual and cohort level, were developed to meet these challenges. PHD can also be used for collecting and visualizing diverse data types (wearable, clinical, omics) as demonstrated recently in the investigation of insulin resistance and the detection of pre-symptomatic COVID-19 (59, 60, 148). Similar data management infrastructure tools have been proposed for the integration of imaging data and omics data (149, 150). Fifth, so far multi-omics analysis has been typically restricted to few hundreds of participants rising questions on the scalability. To this end, recent studies have demonstrated the promise of expanding it to large populations (>4000) (151). In addition to these limitations, there is also a general resistance to change among health-care systems and policy makers, who need robust evidence for adopting multi-omics widely. In addition, training among clinicians for interpreting multi-omics results is currently lacking, as is training for scientists to flexibly work across different 'omes.

In addition to the practical considerations listed earlier, there are equally important ethical considerations surrounding discrimination, consent for testing, data privacy, security, data aggregation, and data re-use. As individualized medical big data becomes commonly utilized in health care settings for personalized medicine, clarity on data ownership, management, distribution, and access needs carefully considered. For example, while data leading to diagnoses are of interest to providers and payers, patients have a right to their privacy. Access to personal medical big data revealing debilitating or expensive disease conditions might prompt employers and insurers to discriminate a person. Thus, legal frameworks are needed to protect privacy as well as providing minimum information necessary for other stake holders (152). Additionally, cloud-based data management platforms hosting personal data must have rigorous regulatory compliance (For *e.g.*, HIPAA, GDPR) to prevent inadvertent/malicious access to data. Some of the commercially available cloud-based platforms such as AWS, Google cloud and MS genomics (www.microsoft.com/en-us/genomics/) have approved and necessary tools for multi-omics data management (153).

Despite these hurdles, with an ongoing decrease in omics analysis costs, availability of robust computational tools for data analysis and management, and integration of data informatics in health-care systems, and adequate training of clinicians, it is projected that by 2030, multi-omics–based precision medicine will increasingly transform clinical medicine with the routine use of multi-omics, microbiome analysis, real-time monitoring of environmental exposures, wearable based continuous monitoring of physical activity, sleep, and metabolic parameters for better management of health (154).

Author contributions—M. B. and M. S. conceptualization; M. B. visualization; M. B. writing – original draft; M. S. writing - reviewing and editing; M. S. supervision; M. S. funding acquisition.

Conflict of interest—M. S. is a cofounder and scientific advisor of Personalis, SensOmics, Qbio, January AI, Fodsel, Filtricine, Protos, RTHM, Iollo, Marble Therapeutics and Mirvie. He is a scientific advisor of Genapsys, Jupiter, Neuvivo, Swaza, Mitrix.

Abbreviations—The abbreviations used are: EHR, Electronic Health Record; ESI, Electrospray Ionization; EV, extra-cellular vesicles; GTEx, Genome-Tissue Expression; GWAS, genome-wide association studies; HCM, Hypertrophic Cardiomyopathy; IFN, interferon; iPOP, Integrative Personal Omics Profiling; iPSC, induced Pluripotent Stem Cells; IR, insulin-resistant; IS, insulin-sensitive; LC, liquid chromatography; MALDI, Matrix-Assisted Laser Desorption/Ionization; NGS, next-generation sequencing; PBMCs, peripheral blood mononuclear cells; PHD, Personal Health Dashboard; PTB, Pre-Term Birth; T2DM, Type 2 diabetes; WES, whole-exome sequencing; WGS, Whole Genome Sequencing.

REFERENCES

1. Hasin, Y., Seldin, M., and Lusis, A. (2017) Multi-omics approaches to disease. *Genome Biol.* **18**, 83
2. Marshall, C. R., Chowdhury, S., Taft, R. J., Lebo, M. S., Buchan, J. G., Harrison, S. M., *et al*. (2020) Best practices for the analytical validation of clinical whole-genome sequencing intended for the diagnosis of germline disease. *NPJ Genom. Med.* **5**, 47–49
3. Geyer, P. E., Holdt, L. M., Teupser, D., and Mann, M. (2017) Revisiting biomarker discovery by plasma proteomics. *Mol. Syst. Biol.* **13**, 942
4. Wishart, D. S. (2016) Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* **15**, 473–484
5. Meikle, T. G., Huynh, K., Giles, C., and Meikle, P. J. (2021) Clinical lipidomics: realizing the potential of lipid profiling. *J. Lipid Res.* **62**, 100127
6. Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., *et al*. (2017) 10 Years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet.* **101**, 5–22
7. Ritchie, M. D., Holzinger, E. R., Li, R., Pendergrass, S. A., and Kim, D. (2015) Methods of integrating data to uncover genotype-phenotype interactions. *Nat. Rev. Genet.* **16**, 85–97
8. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
9. Gorzynski, J. E., Goenka, S. D., Shafin, K., Jensen, T. D., Fisk, D. G., Grove, M. E., *et al*. (2022) Ultrarapid nanopore genome sequencing in a critical care setting. *N. Engl. J. Med.* **386**, 700–702
10. Korbel, J. O., Urban, A. E., Affourtit, J. P., Godwin, B., Grubert, F., Simons, J. F., *et al*. (2007) Paired-end mapping reveals extensive structural variation in the human genome. *Science* **318**, 420–426
11. Shendure, J., Findlay, G. M., and Snyder, M. W. (2019) Genomic medicine-progress, pitfalls, and promise. *Cell* **177**, 45–57
12. Hotchkiss, R. D. (1948) The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J. Biol. Chem.* **175**, 315–332
13. Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **51**, 786–794
14. Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., *et al*. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 1827–1831
15. Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S., and Jaenisch, R. (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877
16. Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H., *et al*. (2008) Highly integrated single-base resolution maps of the epigenome in arabidopsis. *Cell* **133**, 523–536
17. Taiwo, O., Wilson, G. A., Morris, T., Seisenberger, S., Reik, W., Pearce, D., *et al*. (2012) Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat. Protoc.* **7**, 617–636
18. Vaisvila, R., Ponnaluri, V. K. C., Sun, Z., Langhorst, B. W., Saleh, L., Guan, S., *et al*. (2021) Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. *Genome Res.* **31**, 1280–1289
19. Ren, B., Robert, F., Wyrick, J. J., Aparicio, O., Jennings, E. G., Simon, I., *et al*. (2000) Genome-wide location and function of DNA binding proteins. *Science* **290**, 2306–2309
20. Iyer, V. R., Horak, C. E., Scafe, C. S., Botstein, D., Snyder, M., and Brown, P. O. (2001) Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**, 533–538
21. Lieb, J. D., Liu, X., Botstein, D., and Brown, P. O. (2001) Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat. Genet.* **28**, 327–334
22. Horak, C. E., and Snyder, M. (2002) ChIP-chip: a genomic approach for identifying transcription factor binding sites. *Methods Enzymol.* **350**, 469–483
23. Weinmann, A. S., Yan, P. S., Oberley, M. J., Huang, T. H., and Farnham, P. J. (2002) Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* **16**, 235–244
24. Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., *et al*. (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* **4**, 651–657
25. Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007) Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* **316**, 1497–1502
26. Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., *et al*. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837
27. Lauschke, V. M., Ivanov, M., and Ingelman-Sundberg, M. (2017) Pitfalls and opportunities for epigenomic analyses focused on disease diagnosis, prognosis, and therapy. *Trends Pharmacol. Sci.* **38**, 765–770
28. Wang, Z., Gerstein, M., and Snyder, M. (2009) RNA-seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63
29. Fodor, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T., and Solas, D. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science* **251**, 767–773
30. Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470
31. Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D. H., Johnson, D., *et al*. (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **18**, 630–634
32. Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995) Serial analysis of gene expression. *Science* **270**, 484–487
33. Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., *et al*. (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344–1349

34. Wilhelm, B. T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., et al. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**, 1239–1243

35. Byron, S. A., Van Keuren-Jensen, K. R., Engelthaler, D. M., Carpten, J. D., and Craig, D. W. (2016) Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat. Rev. Genet.* **17**, 257–271

36. Muntel, J., Gandhi, T., Verbeke, L., Bernhardt, O. M., Treiber, T., Bruderer, R., et al. (2019) Surpassing 10 000 identified and quantified proteins in a single run by optimizing current LC-MS instrumentation and data analysis strategy. *Mol. Omics* **15**, 348–360

37. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64–71

38. Karas, M., Bachmann, D., and Hillenkamp, F. (1985) Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal. Chem.* **57**, 2935–2939

39. Zhu, H., Klemic, J. F., Chang, S., Bertone, P., Casamayor, A., Klemic, K. G., et al. (2000) Analysis of yeast protein kinases using protein chips. *Nat. Genet.* **26**, 283–289

40. Washburn, M. P., Wolters, D., and Yates, J. R., 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19**, 242–247

41. Kubota, K., Kosaka, T., and Ichikawa, K. (2009) Shotgun protein analysis by liquid chromatography-tandem mass spectrometry. *Methods Mol. Biol.* **519**, 483–494

42. Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., et al. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **75**, 1895–1904

43. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., et al. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3**, 1154–1169

44. Xiao, Q., Zhang, F., Xu, L., Yue, L., Kon, O. L., Zhu, Y., et al. (2021) High-throughput proteomics and AI for cancer biomarker discovery. *Adv. Drug Deliv. Rev.* **176**, 113844

45. Ahadi, S., Zhou, W., Schüssler-Fiorenza Rose, S. M., Sailani, M. R., Contrepois, K., Avina, M., et al. (2020) Personal aging markers and ageotypes revealed by deep longitudinal profiling. *Nat. Med.* **26**, 83–90

46. Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., et al. (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **11**, O111.016717

47. Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B. C., and Aebersold, R. (2018) Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **14**, e8126

48. Wishart, D. S. (2019) Metabolomics for investigating physiological and pathophysiological processes. *Physiol. Rev.* **99**, 1819–1875

49. Yang, K., and Han, X. (2016) Lipidomics: techniques, applications, and outcomes related to biomedical sciences. *Trends Biochem. Sci.* **41**, 954–969

50. Gambhir, S. S., Ge, T. J., Vermesh, O., Spitler, R., and Gold, G. E. (2021) Continuous health monitoring: an opportunity for precision health. *Sci. Transl. Med.* **13**, eabe5383

51. Dunn, J., Runge, R., and Snyder, M. (2018) Wearables and the medical revolution. *Per. Med.* **15**, 429–448

52. Li, X., Dunn, J., Salins, D., Zhou, G., Zhou, W., Schüssler-Fiorenza Rose, S. M., et al. (2017) Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. *PLoS Biol.* **15**, e2001402

53. Lim, W. K., Davila, S., Teo, J. X., Yang, C., Pua, C. J., Blöcker, C., et al. (2018) Beyond fitness tracking: the use of consumer-grade wearable data from normal volunteers in cardiovascular and lipidomics research. *PLoS Biol.* **16**, e2004285

54. Tison, G. H., Sanchez, J. M., Ballinger, B., Singh, A., Olgin, J. E., Pletcher, M. J., et al. (2018) Passive detection of atrial fibrillation using a commercially available smartwatch. *JAMA Cardiol.* **3**, 409–416

55. Price, N. D., Magis, A. T., Earls, J. C., Glusman, G., Levy, R., Lausted, C., et al. (2017) A wellness study of 108 individuals using personal, dense, dynamic data clouds. *Nat. Biotechnol.* **35**, 747–756

56. Ballinger, B., Hsieh, J., Singh, A., Sohoni, N., Wang, J., Tison, G., et al. (2018) DeepHeart: semi-supervised sequence learning for cardiovascular risk prediction. *Proc. AAAI Conf. Artif. Intell.* **32**, 2079–2086

57. Hall, H., Perelman, D., Breschi, A., Limcaoco, P., Kellogg, R., McLaughlin, T., et al. (2018) Glucotypes reveal new patterns of glucose dysregulation. *PLoS Biol.* **16**, e2005143

58. Quer, G., Radin, J. M., Gadaleta, M., Baca-Motes, K., Ariniello, L., Ramos, E., et al. (2021) Wearable sensor data and self-reported symptoms for COVID-19 detection. *Nat. Med.* **27**, 73–77

59. Mishra, T., Wang, M., Metwally, A. A., Bogu, G. K., Brooks, A. W., Bahmani, A., et al. (2020) Pre-symptomatic detection of COVID-19 from smartwatch data. *Nat. Biomed. Eng.* **4**, 1208–1220

60. Alavi, A., Bogu, G. K., Wang, M., Rangan, E. S., Brooks, A. W., Wang, Q., et al. (2022) Real-time alerting system for COVID-19 and other stress events using wearable data. *Nat. Med.* **28**, 175–184

61. Van Eyk, J. E., and Snyder, M. P. (2019) Precision medicine: role of proteomics in changing clinical management and care. *J. Proteome Res.* **18**, 1–6

62. Karczewski, K. J., and Snyder, M. P. (2018) Integrative omics for health and disease. *Nat. Rev. Genet.* **19**, 299–310

63. Kellogg, R. A., Dunn, J., and Snyder, M. P. (2018) Personal omics for precision health. *Circ. Res.* **122**, 1169–1171

64. Chen, R., Mias, G. I., Li-Pook-Than, J., Jiang, L., Lam, H. Y., Chen, R., et al. (2012) Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell* **148**, 1293–1307

65. Stanberry, L., Mias, G. I., Haynes, W., Higdon, R., Snyder, M., and Kolker, E. (2013) Integrative analysis of longitudinal metabolomics data from a personal multi-omics profile. *Metabolites* **3**, 741–760

66. Garrett-Bakelman, F. E., Darshi, M., Green, S. J., Gur, R. C., Lin, L., Macias, B. R., et al. (2019) The NASA twins study: a multidimensional analysis of a year-long human spaceflight. *Science* **364**, eaau8650

67. Mias, G. I., Singh, V. V., Rogers, L. R. K., Xue, S., Zheng, M., Domanskyi, S., et al. (2021) Longitudinal saliva omics responses to immune perturbation: a case study. *Sci. Rep.* **11**, 710–716

68. International Human Genome Sequencing Consortium. (2004) Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945

69. Schaub, M. A., Boyle, A. P., Kundaje, A., Batzoglou, S., and Snyder, M. (2012) Linking disease associations with regulatory information in the human genome. *Genome Res.* **22**, 1748–1759

70. Corradin, O., and Scacheri, P. C. (2014) Enhancer variants: evaluating functions in common disease. *Genome Med.* **6**, 85

71. Boix, C. A., James, B. T., Park, Y. P., Meuleman, W., and Kellis, M. (2021) Regulatory genomic circuitry of human disease loci by integrative epigenomics. *Nature* **590**, 300–307

72. Battle, A., Khan, Z., Wang, S. H., Mitrano, A., Ford, M. J., Pritchard, J. K., et al. (2015) Genomic variation. Impact of regulatory variation from RNA to protein. *Science* **347**, 664–667

73. Cenik, C., Cenik, E. S., Byeon, G. W., Grubert, F., Candille, S. I., Spacek, D., et al. (2015) Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory variation across humans. *Genome Res.* **25**, 1610–1621

74. Mohammadi, P., Castel, S. E., Cummings, B. B., Einson, J., Sousa, C., Hoffman, P., et al. (2019) Genetic regulatory variation in populations informs transcriptome analysis in rare disease. *Science* **366**, 351–356

75. Võsa, U., Claringbould, A., Westra, H. J., Bonder, M. J., Deelen, P., Zeng, B., et al. (2021) Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310

76. Zhang, F., Chen, W., Zhu, Z., Zhang, Q., Nabais, M. F., Qi, T., et al. (2019) OSCA: a tool for omic-data-based complex trait analysis. *Genome Biol.* **20**, 107

77. Hillary, R. F., McCartney, D. L., Harris, S. E., Stevenson, A. J., Seeboth, A., Zhang, Q., et al. (2019) Genome and epigenome wide studies of neurological protein biomarkers in the Lothian Birth Cohort 1936. *Nat. Commun.* **10**, 3160

78. Zhang, S., Cooper-Knock, J., Weimer, A. K., Shi, M., Moll, T., Marshall, J. N. G., et al. (2022) Genome-wide identification of the genetic basis of amyotrophic lateral sclerosis. *Neuron* **110**, 992–1008.e11

79. Piening, B. D., Zhou, W., Contrepois, K., Röst, H., Gu Urban, G. J., Mishra, T., et al. (2018) Integrative personal omics profiles during periods of weight gain and loss. *Cell Syst.* **6**, 157–170.e8

80. Zhou, W., Sailani, M. R., Contrepois, K., Zhou, Y., Ahadi, S., Leopold, S. R., *et al*. (2019) Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature* **569**, 663–671

81. Schussler-Fiorenza Rose, S. M., Contrepois, K., Moneghetti, K. J., Zhou, W., Mishra, T., Mataraso, S., *et al*. (2019) A longitudinal big data approach for precision health. *Nat. Med.* **25**, 792–804

82. Ranjbarvaziri, S., Kooiker, K. B., Ellenberger, M., Fajardo, G., Zhao, M., Vander Roest, A. S., *et al*. (2021) Altered cardiac energetics and mitochondrial dysfunction in hypertrophic cardiomyopathy. *Circulation* **144**, 1714–1731

83. Liu, G., Dong, C., and Liu, L. (2016) Integrated multiple "-omics" data reveal subtypes of hepatocellular carcinoma. *PLoS One* **11**, e0165457

84. Kamoun, A., Idbaih, A., Dehais, C., Elarouci, N., Carpentier, C., Letouzé, E., *et al*. (2016) Integrated multi-omics analysis of oligodendroglial tumours identifies three subgroups of 1p/19q co-deleted gliomas. *Nat. Commun.* **7**, 11263

85. Jiang, Y., Shi, X., Zhao, Q., Krauthammer, M., Rothberg, B. E., and Ma, S. (2016) Integrated analysis of multidimensional omics data on cutaneous melanoma prognosis. *Genomics* **107**, 223–230

86. Sanghi, A., Gruber, J. J., Metwally, A., Jiang, L., Reynolds, W., Sunwoo, J., *et al*. (2021) Chromatin accessibility associates with protein-RNA correlation in human cancer. *Nat. Commun.* **12**, 5732

87. Zhang, H., Liu, T., Zhang, Z., Payne, S. H., Zhang, B., McDermott, J. E., *et al*. (2016) Integrated proteogenomic characterization of human high-grade serous ovarian cancer. *Cell* **166**, 755–765

88. Mani, D. R., Krug, K., Zhang, B., Satpathy, S., Clauser, K. R., Ding, L., *et al*. (2022) Cancer proteogenomics: current impact and future prospects. *Nat. Rev. Cancer* **22**, 298–313

89. Rodriguez, H., Zenklusen, J. C., Staudt, L. M., Doroshow, J. H., and Lowy, D. R. (2021) The next horizon in precision oncology: proteogenomics to inform cancer diagnosis and treatment. *Cell* **184**, 1661–1670

90. Zhang, B., Whiteaker, J. R., Hoofnagle, A. N., Baird, G. S., Rodland, K. D., and Paulovich, A. G. (2019) Clinical potential of mass spectrometry-based proteogenomics. *Nat. Rev. Clin. Oncol.* **16**, 256–268

91. Bernardes, J. P., Mishra, N., Tran, F., Bahmer, T., Best, L., Blase, J. I., *et al*. (2020) Longitudinal multi-omics analyses identify responses of megakaryocytes, erythroid cells, and plasmablasts as hallmarks of severe COVID-19. *Immunity* **53**, 1296–1314.e9

92. Zhang, S., Cooper-Knock, J., Weimer, A. K., Shi, M., Kozhaya, L., Unutmaz, D., *et al*. (2022) Multiomic analysis reveals cell-type-specific molecular determinants of COVID-19 severity. *Cell Syst.* **13**, 598–614.e6

93. Sacco, K., Castagnoli, R., Vakkilainen, S., Liu, C., Delmonte, O. M., Oguz, C., *et al*. (2022) Immunopathological signatures in multisystem inflammatory syndrome in children and pediatric COVID-19. *Nat. Med.* **28**, 1050–1062

94. Su, Y., Chen, D., Yuan, D., Lausted, C., Choi, J., Dai, C. L., *et al*. (2020) Multi-omics resolves a sharp disease-state shift between mild and moderate COVID-19. *Cell* **183**, 1479–1495.e20

95. Wimmers, F., and Pulendran, B. (2020) Emerging technologies for systems vaccinology - multi-omics integration and single-cell (epi)genomic profiling. *Curr. Opin. Immunol.* **65**, 57–64

96. Ward, R. A., Aghaeepour, N., Bhattacharyya, R. P., Clish, C. B., Gaudillière, B., Hacohen, N., *et al*. (2021) Harnessing the potential of multiomics studies for precision medicine in infectious disease. *Open Forum Infect. Dis.* **8**, ofab483

97. Messner, C. B., Demichev, V., Wendisch, D., Michalick, L., White, M., Freiwald, A., *et al*. (2020) Ultra-high-throughput clinical proteomics reveals classifiers of COVID-19 infection. *Cell. Syst.* **11**, 11–24.e4

98. Watzenboeck, M. L., Gorki, A. D., Quattrone, F., Gawish, R., Schwarz, S., Lambers, C., *et al*. (2022) Multi-omics profiling predicts allograft function after lung transplantation. *Eur. Respir. J.* **59**, 2003292

99. Wigger, L., Barovic, M., Brunner, A. D., Marzetta, F., Schöniger, E., Mehl, F., *et al*. (2021) Multi-omics profiling of living human pancreatic islet donors reveals heterogeneous beta cell trajectories towards type 2 diabetes. *Nat. Metab.* **3**, 1017–1031

100. Ghaemi, M. S., DiGiulio, D. B., Contrepois, K., Callahan, B., Ngo, T. T. M., Lee-McMullen, B., *et al*. (2019) Multiomics modeling of the immunome, transcriptome, microbiome, proteome and metabolome adaptations during human pregnancy. *Bioinformatics* **35**, 95–103

101. Stelzer, I. A., Ghaemi, M. S., Han, X., Ando, K., Hédou, J. J., Feyaerts, D., *et al*. (2021) Integrated trajectories of the maternal metabolome, proteome, and immunome predict labor onset. *Sci. Transl. Med.* **13**, eabd9898

102. Jehan, F., Sazawal, S., Baqui, A. H., Nisar, M. I., Dhingra, U., Khanam, R., *et al*. (2020) Multiomics characterization of preterm birth in low- and middle-income countries. *JAMA Netw. Open* **3**, e2029655

103. Nie, C., Li, Y., Li, R., Yan, Y., Zhang, D., Li, T., *et al*. (2022) Distinct biological ages of organs and systems identified from a multi-omics study. *Cell Rep.* **38**, 110459

104. Mahmoudi, S., Mancini, E., Xu, L., Moore, A., Jahanbani, F., Hebestreit, K., *et al*. (2019) Heterogeneity in old fibroblasts is linked to variability in reprogramming and wound healing. *Nature* **574**, 553–558

105. Integrative HMP (iHMP) Research Network Consortium. (2014) The integrative human microbiome project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. *Cell. Host Microbe* **16**, 276–289

106. Sommer, F., and Bäckhed, F. (2013) The gut microbiota–masters of host development and physiology. *Nat. Rev. Microbiol.* **11**, 227–238

107. Heintz-Buschart, A., May, P., Laczny, C. C., Lebrun, L. A., Bellora, C., Krishna, A., *et al*. (2016) Integrated multi-omics of the human gut microbiome in a case study of familial type 1 diabetes. *Nat. Microbiol.* **2**, 16180

108. Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., *et al*. (2019) Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662

109. Mars, R. A. T., Yang, Y., Ward, T., Houtti, M., Priya, S., Lekatz, H. R., *et al*. (2020) Longitudinal multi-omics reveals subset-specific mechanisms underlying irritable bowel syndrome. *Cell* **182**, 1460–1473.e17

110. Thaiss, C. A., Levy, M., Korem, T., Dohnalová, L., Shapiro, H., Jaitin, D. A., *et al*. (2016) Microbiota diurnal rhythmicity programs host transcriptome oscillations. *Cell* **167**, 1495–1510.e12

111. van Niel, G., D'Angelo, G., and Raposo, G. (2018) Shedding light on the cell biology of extracellular vesicles. *Nat. Rev. Mol. Cell Biol.* **19**, 213–228

112. Yim, K. H. W., Borgoni, S., and Chahwan, R. (2022) Serum extracellular vesicles profiling is associated with COVID-19 progression and immune responses. *J. Extracell Biol.* **1**, e37

113. Chronopoulos, A., and Kalluri, R. (2020) Emerging role of bacterial extracellular vesicles in cancer. *Oncogene* **39**, 6951–6960

114. Génin, E. (2020) Missing heritability of complex diseases: case solved? *Hum. Genet.* **139**, 103–113

115. Li, J., Li, X., Zhang, S., and Snyder, M. (2019) Gene-environment interaction in the era of precision medicine. *Cell* **177**, 38–44

116. Eriksson, K. F., and Lindgärde, F. (1991) Prevention of type 2 (non-insulin-dependent) diabetes mellitus by diet and physical exercise. The 6-year Malmö feasibility study. *Diabetologia* **34**, 891–898

117. Rejeski, W. J., Ip, E. H., Bertoni, A. G., Bray, G. A., Evans, G., Gregg, E. W., *et al*. (2012) Lifestyle change and mobility in obese adults with type 2 diabetes. *N. Engl. J. Med.* **366**, 1209–1217

118. Helmrich, S. P., Ragland, D. R., Leung, R. W., and Paffenbarger, R. S., Jr. (1991) Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. *N. Engl. J. Med.* **325**, 147–152

119. Rawshani, A., Rawshani, A., Franzén, S., Sattar, N., Eliasson, B., Svensson, A. M., *et al*. (2018) Risk factors, mortality, and cardiovascular outcomes in patients with type 2 diabetes. *N. Engl. J. Med.* **379**, 633–644

120. Contrepois, K., Wu, S., Moneghetti, K. J., Hornburg, D., Ahadi, S., Tsai, M. S., *et al*. (2020) Molecular choreography of acute exercise. *Cell* **181**, 1112–1130.e16

121. Li, V. L., He, Y., Contrepois, K., Liu, H., Kim, J. T., Wiggenhorn, A. L., *et al*. (2022) An exercise-inducible metabolite that suppresses feeding and obesity. *Nature* **606**, 785–790

122. Sanford, J. A., Nogiec, C. D., Lindholm, M. E., Adkins, J. N., Amar, D., Dasari, S., *et al*. (2020) Molecular Transducers of Physical Activity Consortium (MoTrPAC): mapping the dynamic responses to exercise. *Cell* **181**, 1464–1474

123. Barabási, A., Menichetti, G., and Loscalzo, J. (2020) The unmapped chemical complexity of our diet. *Nat. Food* **1**, 33–37

124. Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., *et al*. (2015) Personalized nutrition by prediction of glycemic responses. *Cell* **163**, 1079–1094

125. Berry, S. E., Valdes, A. M., Drew, D. A., Asnicar, F., Mazidi, M., Wolf, J., *et al.* (2020) Human postprandial responses to food and potential for precision nutrition. *Nat. Med.* **26**, 964–973

126. Lancaster, S. M., Lee-McMullen, B., Abbott, C. W., Quijada, J. V., Hornburg, D., Park, H., *et al.* (2022) Global, distinctive, and personal changes in molecular and microbial profiles by specific fibers in humans. *Cell. Host Microbe* **30**, 848–862.e7

127. Renz, H., Holt, P. G., Inouye, M., Logan, A. C., Prescott, S. L., and Sly, P. D. (2017) An exposome perspective: early-life events and immune development in a changing world. *J. Allergy Clin. Immunol.* **140**, 24–40

128. Smith, M. T., de la Rosa, R., and Daniels, S. I. (2015) Using exposomics to assess cumulative risks and promote health. *Environ. Mol. Mutagen.* **56**, 715–723

129. Vermeulen, R., Schymanski, E. L., Barabási, A. L., and Miller, G. W. (2020) The exposome and health: where chemistry meets biology. *Science* **367**, 392–396

130. Jiang, C., Wang, X., Li, X., Inlora, J., Wang, T., Liu, Q., *et al.* (2018) Dynamic human environmental exposome revealed by longitudinal personal monitoring. *Cell* **175**, 277–291.e31

131. [preprint] Maitre, L.É., Bustamante, M., Hernández-Ferrer, C., Thiel, D., Lau, C., Siskos, A., *et al.* (2021) Multi-omics signatures of the human early life exposome. *medRxiv*. https://doi.org/10.1101/2021.05.04.21256605

132. Gao, P., Shen, X., Zhang, X., Jiang, C., Zhang, S., Zhou, X., *et al.* (2022) Precision environmental health monitoring by longitudinal exposome and multi-omics profiling. *Genome Res.* **32**, 1199–1214

133. Li, J., Pan, C., Zhang, S., Spin, J. M., Deng, A., Leung, L. L. K., *et al.* (2018) Decoding the genomics of abdominal aortic aneurysm. *Cell* **174**, 1361–1372.e10

134. All of Us Research Program Investigators, Denny, J. C., Rutter, J. L., Goldstein, D. B., Philippakis, A., Smoller, J. W., *et al.* (2019) The "all of us" research program. *N. Engl. J. Med.* **381**, 668–676

135. Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., *et al.* (2022) The complete sequence of a human genome. *Science* **376**, 44–53

136. Adhikari, S., Nice, E. C., Deutsch, E. W., Lane, L., Omenn, G. S., Pennington, S. R., *et al.* (2020) A high-stringency blueprint of the human proteome. *Nat. Commun.* **11**, 5301–5309

137. Jiang, L., Wang, M., Lin, S., Jian, R., Li, X., Chan, J., *et al.* (2020) A quantitative proteome map of the human body. *Cell* **183**, 269–283.e19

138. Wishart, D. S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., *et al.* (2022) HMDB 5.0: the human metabolome database for 2022. *Nucleic Acids Res.* **50**, D622–D631

139. ENCODE Project Consortium. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74

140. Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330

141. GTEx Consortium, Laboratory, Data Analysis & Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, *et al.* (2017) Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213

142. eGTEx Project. (2017) Enhancing GTEx by bridging the gaps between genotype, gene expression, and disease. *Nat. Genet.* **49**, 1664–1670

143. Cancer Genome Atlas Research Network, Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R., Ozenberger, B. A., *et al.* (2013) The cancer genome atlas pan-cancer analysis project. *Nat. Genet.* **45**, 1113–1120

144. Misra, B. B., Langefeld, C. D., Olivier, M., and Cox, L. A. (2018) Integrated omics: tools, advances, and future approaches. *J. Mol. Endocrinol.* https://doi.org/10.1530/JME-18-0055

145. Subramanian, I., Verma, S., Kumar, S., Jere, A., and Anamika, K. (2020) Multi-omics data integration, interpretation, and its application. *Bioinform. Biol. Insights* **14**, 1177932219899051

146. Krassowski, M., Das, V., Sahu, S. K., and Misra, B. B. (2020) State of the field in multi-omics research: from computational needs to data mining and sharing. *Front. Genet.* **11**, 610798

147. Ronan, T., Qi, Z., and Naegle, K. M. (2016) Avoiding common pitfalls when clustering biological data. *Sci. Signal.* **9**, re6

148. Bahmani, A., Alavi, A., Buergel, T., Upadhyayula, S., Wang, Q., Ananthakrishnan, S. K., *et al.* (2021) A scalable, secure, and interoperable platform for deep data-driven health management. *Nat. Commun.* **12**, 5757

149. Kuhn Cuellar, L., Friedrich, A., Gabernet, G., de la Garza, L., Fillinger, S., Seyboldt, A., *et al.* (2022) A data management infrastructure for the integration of imaging and omics data in life sciences. *BMC Bioinformatics* **23**, 61–63

150. Zanfardino, M., Castaldo, R., Pane, K., Affinito, O., Aiello, M., Salvatore, M., *et al.* (2021) MuSA: a graphical user interface for multi-OMICs data integration in radiogenomic studies. *Sci. Rep.* **11**, 1550

151. Zhang, W., Wan, Z., Li, X., Li, R., Luo, L., Song, Z., *et al.* (2022) A population-based study of precision health assessments using multi-omics network-derived biological functional modules. *Cell Rep. Med.* **3**, 100847

152. Price, W. N., 2nd, and Cohen, I. G. (2019) Privacy in the age of medical big data. *Nat. Med.* **25**, 37–43

153. Koppad, S., B, A., Gkoutos, G. V., and Acharjee, A. (2021) Cloud computing enabled big multi-omics data analytics. *Bioinform. Biol. Insights* **15**, 11779322211035921

154. Denny, J. C., and Collins, F. S. (2021) Precision medicine in 2030-seven ways to transform healthcare. *Cell* **184**, 1415–1419