

**UCLA**

**UCLA Previously Published Works**

**Title**

Multitissue Multiomics Systems Biology to Dissect Complex Diseases

**Permalink**

<https://escholarship.org/uc/item/0d847415>

**Journal**

Trends in Molecular Medicine, 26(8)

**ISSN**

1471-4914

**Author**

Yang, Xia

**Publication Date**

2020-08-01

**DOI**

10.1016/j.molmed.2020.04.006

Peer reviewed



# HHS Public Access

Author manuscript

*Trends Mol Med.* Author manuscript; available in PMC 2021 August 01.

Published in final edited form as:

*Trends Mol Med.* 2020 August ; 26(8): 718–728. doi:10.1016/j.molmed.2020.04.006.

## Multi-tissue Multi-omics Systems Biology to Dissect Complex Diseases

Xia Yang\*

Department of Integrative Biology and Physiology, University of California, Los Angeles

### Abstract

Most complex diseases involve genetic and environmental risk factors, engage multiple cells and tissues, and follow a polygenic or omnigenic model depicting numerous genes contributing to pathophysiology. These multidimensional complexities pose challenges to traditional approaches that examine individual factors. In turn, multi-tissue multi-omics systems biology has emerged to comprehensively elucidate within- and cross-tissue molecular networks underlying gene-by-environment interactions and contributing to complex diseases. The power of systems biology in retrieving novel insights and formulating new hypotheses has been well-documented. However, the field faces various challenges that call for debate and actions. In this opinion, I discuss the concepts, benefits, current state, and challenges of the field, and point to the next steps towards network-based systems medicine.

### Keywords

systems biology; multi-tissue; multi-omics; networks; gene by environment interactions; complex diseases

### Why multi-tissue multi-omics systems biology?

Eradicating human diseases is the ultimate mission of biomedicine. However, despite decades of extensive research, the majority of common human diseases of high morbidity and mortality, such as cardiovascular disease (CVD) and its associated metabolic risk factor type 2 diabetes (T2D), still face growing prevalence and lack of effective therapeutic strategies to eliminate disease burden.

So, what prevents us from making the next biomedical breakthroughs at a time when biotechnology has seen its peak innovation and when cutting-edge research tools are readily available? One of the key hurdles is the extreme complexity of human diseases (Figure 1). First, besides the revelation of tens to hundreds or even more genetic risk loci identified

---

\*Correspondence: xyang123@ucla.edu (X. Yang).

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest

The author declares no conflict of interest.

through genome-wide association studies (**GWAS**; see Glossary) [1], diverse environmental factors such as diets, physical activity, pollutants, toxins, and pathogens also have profound impact on disease development [2]. Environmental factors also interact with the genome, termed gene-by-environment interactions, to further confer inter-individual differences in disease risks [3]. Secondly, each complex disease is the result of systems level perturbations involving multiple tissues, diverse cell types, and numerous molecular pathways. For example, CVD is promoted by dyslipidemia, inflammation, coagulation and vascular dysfunctions governed by the liver, adipose, immune system, and the vasculature [4]; as a major risk factor for CVD, T2D involves pancreatic beta cells, liver, adipose tissue, intestine, and skeletal muscle [5], as well as various molecular pathways ranging from insulin secretion and insulin signaling to cell cycle and immune pathways [6]. Lastly, the molecular model for complex diseases has been revised from a **polygenic** (multiple genes) to a debatable **omnigenic model** [7, 8]. The omnigenic model states that essentially all genes interact in **molecular networks** (an example for CVD in Figure 1B), and perturbations of any of the interacting genes can propagate into overall network perturbations resulting in disease development. In this model, the scale of contribution of individual genes to disease development differs in accordance with their importance and position in the networks, with central hub genes (e.g., *CAVI* for CVD; Figure 1B) that have more interacting partners playing a more significant role, whereas peripheral genes with few connections (e.g., genes surrounding *CAVI*; Figure 1B) exhibiting subtle to moderate impact. These complexities, ranging from diverse risk factors to a multitude of tissue, cellular, and molecular systems, call for a holistic view of the multidimensional interactions to prioritize key tissues, cell types, and central molecular regulators to guide more effective therapy.

Complex diseases have been mainly investigated using classic **reductionist approaches** [9, 10], which examine one factor at a time and have been successful in elucidating the functions of individual genes, proteins, or other types of molecules. Reductionism requires specific hypotheses, such as the belief that a particular gene is involved in a biological process that affects a particular function or disease. However, such approaches become less efficient in addressing the polygenic/omnigenic disease model and the complex interactions across molecules within and between tissues in different contexts. Specifically, the functions of biomolecules are not only determined by their own properties but by their micro- and macro-environments, such as the specific cell and tissue types where different functional partners are present, as well as the larger genetic, environmental, and physiological contexts of individuals. Examining the functions of a molecule in limited settings (such as in a particular cell line or a specific mouse model) will only offer fragmented biological understanding that does not apply to other conditions. This is particularly significant given the observations that knocking out a gene in different mouse strains has dramatic differential phenotypic consequences between mice with distinct genetic background [11], and the same environmental exposure can lead to a broad spectrum of physiological and disease variations between individuals in both human and rodent populations [12–16]. Yet the majority of existing studies employing reductionist approaches make conclusions based on experiments conducted under a limited number of conditions, if not one. On the other hand, exhaustively testing the functions of individual molecules across all variable conditions is not feasible for reductionist approaches, thereby limiting its power to fully dissect disease complexities.

As a complementary approach, **multi-tissue multi-omics systems biology** has gained momentum in recent years to meet the challenges. The rise of this discipline is the natural result of the recent conceptual and technical breakthroughs that have enabled global examination of the near-complete sets of genes, proteins, metabolites, or bacterial species in health and disease through **multi-omics** technologies.

In this article, I will first introduce the basic concepts and approaches involved in multi-tissue multi-omics systems biology. Subsequently, I will briefly review the current state of the field through recent application studies dissecting the molecular interactions across multiple tissues affected by genetic or environmental risks or by gene-by-environment interactions. Although multi-tissue multi-omics approaches have been applied to numerous diseases, I will focus on select applications investigating *multiple tissues* and *multiple layers of omics* in CVD and its interconnected risk factor T2D as examples to elucidate the main directions and advances. Lastly, I will point out the challenges facing the field and outline the potential strategies to further the path towards comprehensive mechanistic understanding and more effective treatment for complex diseases.

## Basic concepts and approaches for multi-tissue multi-omics systems biology

Systems biology aims to understand the holistic complex interactions in biological systems [17]. Multi-tissue multi-omics systems biology is a subdiscipline that relies on diverse types of high throughput omics data (genome, epigenome, transcriptome, metabolome, proteome, and microbiome; Figure 1A) from disease-relevant tissues, to derive the molecular interactions in the form of **molecular networks** across organ systems using mathematical, statistical, and computational analyses [6, 18–21]. As reviewed previously [22, 23], there are many types of molecular networks, such as protein-protein interaction networks, gene regulatory networks, metabolic networks, and hybrid networks (Figure 2), which can be derived based on correlation, regression, ordinary differential equation, mutual information, Gaussian graphical models, and Bayesian approaches. Despite some recent debate [24, 25], the organization of biological networks has long been viewed to follow a “scale-free” pattern [26, 27] where a small number of nodes have many more connections than average (“hub” in Figure 1) whereas the majority of the nodes have few connections (“peripheral node” in Figure 1).

Genetic and environmental risk factors influence multiple parts of the networks (“subnetworks”; Figure 1A), disruptions of which perturb specific biological pathways or functions which in turn promote disease development. As many genes and their associated subnetworks could contribute to a complex disease, it is critical to pinpoint the hubs that are central in the disease network as well as their cell and tissue of origin (e.g., adipose *CAVI* in CVD network; Figure 1B) to prioritize target cells/tissues, pathways, and genes.

In contrast to hypothesis-driven reductionist approaches that focus on individual molecules and pathways in a given cell or tissue, systems biology is data-driven and attend to interactions across biomolecules, biological pathways, and networks within and between cell types and tissues to detect key features. These two approaches are complementary, rather

than conflicting, and both are important tools to tackle complex biological questions. A logical order is to use systems biology to provide the global maps and point to important features, followed by reductionist approaches to investigate the detailed events.

## Where does the field stand?

Over the past decade, major progress has been made to dissect cross-tissue mechanisms underlying genetic risks, environmental risks, and gene by environment interactions by simultaneously investigating multiple tissues and diverse omics domains (select examples for diverse diseases in Table 1). In the main discussion I use examples in the cardiometabolic field, CVD and its associated risk factor T2D, as a case-study.

### Multi-tissue multi-omics systems biology to understand genetically perturbed gene networks.

Despite the enormous success of human GWAS in unraveling tens to hundreds of common genetic variants that are associated with human diseases [1], the target genes and pathways as well as the tissue/cell context of the disease-associated variants are largely unclear. In one of the early applications of multi-tissue multi-omics systems biology [28], 16 human GWAS of CVD were integrated with independent tissue-specific transcriptome data and the corresponding **expression quantitative trait loci** (eQTLs) from liver, adipose, blood, and vascular endothelial cells, to link genetic variants with potential downstream target genes and pathways in individual tissues based on tissue-specific gene regulation. In doing so, this single study not only recapitulated essentially all previously known pathways involved in CVD pathogenesis such as cholesterol and lipid metabolism, vascular dysfunction, and inflammation, but revealed novel pathways such as erythropoietin-mediated neuroprotection through NF- $\kappa$ B, cell cycle, cell stress, and spliceosome. Additionally, through the use of tissue-specific gene regulation information, the study determined the tissue context of the genes and pathways, highlighting the adipose tissue as the most informative among the four relevant tissues examined. Moreover, the study used tissue-specific gene regulatory networks to pinpoint hub genes that regulate individual disease processes and their interactions, such as *SQLE* and *PLG* for lipid pathways and *PTPRC* for immune processes. The integration of multi-omics data across tissues helped piece together the genetic puzzle of CVD and highlighted key perturbation points governing interacting pathways that could serve as novel therapeutic targets. As a step forward, a recent study generated genetic and transcriptomic datasets from seven vascular and metabolic tissues across 600 CVD patients, offering the scientific community a much richer resource to understand CVD and revealing additional tissue-specific CVD targets [29] (details in Box 1).

As T2D is a significant risk factor for CVD, multi-tissue multi-omics systems biology has also been applied to investigate the genetically perturbed networks connecting these two prevalent cardiometabolic diseases [30]. Here, integration of five multi-ethnic GWAS with transcriptome and eQTL data from ~20 tissues identified converging pathways in numerous tissues between CVD and T2D across African American, Caucasian, and Hispanic populations. The converging pathways included lipid, glucose, and branched-chain amino acid metabolism, oxidative phosphorylation, extracellular matrix, immune response, and

neuronal processes. Network analysis further prioritized tissue-specific key regulators of these shared processes such as *CAVI* (adipose) for inflammation, lipid metabolism, and vascular functions, and *PCOLCE* (hypothalamus) for extracellular matrix and energy homeostasis (Figure 1B). This unique study comparing two interconnected cardiometabolic diseases across multi-ethnic populations highlights shared pathways and regulators that could be targeted to treat both diseases in the general population.

The above studies demonstrate the power of utilizing genetic associations in conjunction with tissue-specific gene regulation to obtain global views of the genetically perturbed networks and pathways in a tissue-specific manner.

### **Multi-tissue multi-omics systems biology to predict response to environmental exposure and to understand the environmentally perturbed gene networks.**

Environmental exposures are particularly difficult to dissect in human populations, owing to the challenges in accurately quantifying and controlling exposure levels and difficulties in obtaining internal human tissues (see Clinician's Corner). Nevertheless, peripheral samples such as blood, urine, and stool samples are readily available and multi-omics studies can be pursued to identify peripheral biomarkers predictive of environmental exposures for prognostic and diagnostic purposes. Artificial intelligence and machine learning approaches are uniquely positioned for efficient biomarker identification. For example, in a landmark precision nutrition study, physiological measurements, blood parameters, and gut microbiota profiling were used to build a machine learning model of biomarkers that can predict glycemic responses (indicative of T2D risks) to various diets more accurately than experienced dietitians can achieve [15], thus paving the path for omics-driven personal nutrition. Similar efforts from the Integrative Personal Omics Profiles (iPOP) and the Pioneer 100 Wellness Project (P100) have also yielded highly predictive dynamic peripheral multi-omics biomarkers of CVD, prediabetes, and diabetes, as well as biomarkers reflective of environmental influences of viral infection, immunization, and weight gain/loss [31–34]. These laudable studies revealed not only concordant dynamic molecular and phenotypic responses across study participants, but, more importantly, great inter-individual variability, pointing to the need for personalized biomarkers for translational precision medicine.

Compared to human studies, animal models offer unique advantages, such as better control of environmental exposure and genetic variables as well as easy access to internal tissues, which can go beyond peripheral biomarkers to tackle mechanistic causal insights into the tissue-specific network perturbations across organ systems induced by specific environmental exposures (Table 1). For instance, a recent multi-tissue multi-omics study was carried out to understand whether and how prenatal exposure to Bisphenol A (BPA), a prevalent environmental chemical, poses risks to cardiometabolic disorders [35]. Through comprehensive examination of both the transcriptome and epigenome of the hypothalamus, liver, and adipose tissues in a mouse model, this systems biology investigation revealed extensive molecular perturbations in metabolic pathways across tissues as well as tissue-specific processes such as extracellular matrix in the hypothalamus and histone subunits in the adipose tissue. Compared to previous hypothesis-driven research which focused on the effects of BPA on specific pathways such as estrogen and PPAR signaling, this multi-tissue

multi-omics study uncovered numerous less-studied targets of BPA, such as *Cyp51* and long noncoding RNAs across tissues, *Fasn* in both liver and adipose, *Hnf4a* in liver, *Fa2h* in hypothalamus, and *Nfya* in adipose tissue. Further integration of the tissue-specific molecular alterations with GWAS of >60 human diseases revealed strong enrichment of the BPA target genes for association with CVD and T2D. These findings significantly expand our knowledge about the selective tissue and molecular sensitivity to BPA and offer molecular insights into the mechanisms connecting BPA exposure to cardiometabolic diseases. Such animal model investigations can complement human biomarker studies to illuminate predictive models that inform on mechanisms, which offer both therapeutic targets and prognostic/diagnostic tools to counteract environmental risks of human diseases.

### **Use of multi-tissue multi-omics systems biology to explore gene-by-environment interactions.**

Rodent populations with diverse genetic background serve as a powerful model to examine how specific genetic makeup interacts with a given environmental exposure to determine inter-individual differences in disease risks. As an example, the interactions between a high fat high sucrose diet and the host genome have been investigated using >100 inbred and recombinant mouse strains, revealing vast differences in cardiometabolic phenotypes across mouse strains fed the same diet [12, 13]. By integrating the genomic sequence variants with the liver and adipose transcriptome, plasma metabolome, and gut microbiome, these studies uncovered numerous genes such as *Npc1*, *Gpl2r*, and *Klf14* as well as microbial species such as *Akkermansia* and *Lactococcus* as potential determinants of differential susceptibility to cardiometabolic dysfunctions in response to high fat high sucrose diet among genetically distinct individuals.

The studies outlined above are only a few examples among numerous multi-tissue multi-omics studies to reveal completely new insights that were not possible by purely hypothesis-driven approaches (see additional studies in Table 1 and detailed examples in Box 1 and Box 2). Importantly, subsequent molecular perturbation experiments were carried out in many of the studies and novel predictions were substantiated by experimental evidence. The prioritized regulators, biomarkers, networks, and tissues that are affected by genetic and environmental factors serve as guidance for future development of preventative and therapeutic strategies that can be tailored to the specific types of causal risks and pathways. In particular, agents that have the capacity to target the central subnetworks and regulators are likely to be more effective in counteracting disease pathogenesis in the general population, whereas individual-specific molecular alterations may inform on personalized strategies for prevention and treatment. The diverse pathways connected in the networks can also guide the selection of combinatorial treatment strategies that engage drugs targeting different pathways and subnetworks. Once a holistic understanding of disease networks is achieved, this type network-based medicine is the critical next step for multi-tissue multi-omics systems biology [36].

## Current debate and challenges over multi-tissue multi-omics systems biology

Despite the accumulating evidence supporting the value and enormous discovery potential of multi-tissue multi-omics systems biology and the growing interest in such approaches by the scientific community, the field has been under intense debate. Proponents supporting the field recognize the importance of global comprehensive views of interactions among biomolecules, cells, tissues, and organs, and value its objective and data-driven nature as well as the great potential for novel discoveries that do not rely on prior hypotheses. Such recognition and acceptance can be reflected in the rising numbers of multi-tissue multi-omics systems biology publications in high impact journals. The same strengths, however, are perceived by critics as “fishing expedition”, exploratory, lack of hypothesis, descriptive, overwhelming, and open-ended. Some view the field as a technology-driven hypothesis-generating tool, rather than an independent research discipline. In turn, data scientists are often perceived as supporting analysts or even “research parasites” [37, 38] but not independent investigators driving biological discoveries. Other common critiques of multi-omics systems biology include the perceived lack of mechanistic insights, correlative nature of the findings, and an unclear translational path because of the large numbers of new targets and hypotheses it usually generates. The harshest criticism among all, is perhaps that “this approach is bound to fail” [39].

Given the strong skepticism and criticism, it is not surprising that omics-driven systems biology has been facing numerous challenges. To address the common critiques and to convince skeptics, robust demonstrations of the values added by systems biology research in terms of mechanistic follow-ups and translational successes are required. This is a high bar for a relatively new field, however, since it is costly and time-consuming to generate multidimensional data, develop tools for data integration, carry out extensive data modeling or “the \$1000 genome, \$100000 analysis” [40], and then conduct validation and functional studies. This is in stark contrast with traditional research fields where one starts with a prior hypothesis focusing on a gene or protein or a molecular pathway and directly tests the hypothesis experimentally, which constitutes only one of the many steps in systems biology studies. Additionally, high throughput omics data can also be viewed as less accurate than traditional low throughput measures such as qPCR and Western blotting, and hence technical validations are often demanded, despite the fact that technical comparison studies support that high throughput methods are not any less accurate than traditional methods [41, 42]. Overall, peers, collaborators, and reviewers from non-systems biology disciplines tend to under-value of the multidimensional datasets and findings that can form watersheds for future research, the innovation involved in the design and implementation of multidimensional studies and analytical strategies, the higher cost, longer time and effort commitment, and the higher standards for systems biologists to complete and publish a study. From the translational perspective, multi-tissue multi-omics systems biology studies can be difficult to implement in human studies and clinical settings due to the explicit need for diverse types of biospecimen (see Clinician’s Corner).



To further the challenge, funding mechanisms and balanced peer-review systems are lagging behind for the type of discovery studies promoted by multi-omics systems biology. Most of the regular NIH study sections are in favor of hypothesis-driven research that is considered more “mechanistic” and hence more valuable than “exploratory” discovery studies. Although systems biology in fact has specific hypotheses, these do not fit the traditional definition. Even the very definition of being “mechanistic” is debatable, since different fields may have different standards for what qualifies as a mechanistic study. For a biochemist it could mean where the exact binding site is for a protein; for a molecular biologist, it could mean the linear signaling cascade from biomolecule A to B to C; for systems biologists it means which molecules and pathways are interacting in which cells or tissues to perform a function and, when perturbed, lead to disease. These are different levels of mechanistic insights that are all biologically meaningful and should be broadly recognized and appreciated. In my opinion, penalizing holistic data-driven research for not fitting the narrow-sense mechanistic definition is harmful to the very mission of science to make discoveries and to explore uncharted domains.

## Concluding remarks

Despite the numerous challenges outlined above, I see a bright future for multi-tissue multi-omics systems biology to help accelerate our understanding of complex diseases. We have entered a golden age to conduct such research, with numerous cutting-edge high-throughput omics technologies and maturing analytical methodologies under our belt. Most encouragingly, accumulating evidence substantiates the enormous discovery potential as well as the validity and accuracy of the findings from a systems approach. The growing acceptance and adaptation of the discipline in basic and clinical research further facilitate the maturation of the field, although broader appreciation and support are needed to ensure the healthy growth of the field to maximize its impact. In terms of future directions, in addition to broadening the applications of multi-tissue multi-omics systems biology to diverse types of complex diseases, the field needs to embrace new opportunities to make conceptual and technical leaps. There are numerous exciting new directions that the field is uniquely positioned to address, particularly with regards to how systems biology can help formulate new data-driven hypotheses on novel targets and biomarkers that are tailored to individualized disease risks to fill in the gaps towards network-based medicine (see Outstanding Questions). With a mission towards unbiased and objective research and with an utter respect for data and discovery, I see no boundaries for this field that is in my view not bound to fail.

## Acknowledgements

I thank Montgomery Blencowe for helping prepare the figure. XY is funded by NIH R01 DK104363, R01 DK117850, R21 NS103088, R01 HL145708, R01 ES027237, and P30 ES07048.

## Glossary

### Expression quantitative trait loci (eQTL)

Genetic loci that are associated with the expression levels of a gene transcript. eQTLs are identified by examining the correlation between the genotype at each genetic locus with the

copy numbers of an expressed transcript. eQTLs are usually tissue-specific because gene expression is regulated in a tissue-specific manner

### **Genome-wide association studies (GWAS)**

A popular and common type of genetic study in human populations to identify genetic variants that are associated with increased or decreased risk of a disease

### **Molecular networks**

Comprised of nodes which are indicative of the molecular entities such as genes, proteins, or metabolites, and edges that illustrate the connections or links between nodes. Edges can be regulatory relations, correlations, physical interactions, or enzymatic and biochemical reactions (Figure 2)

### **Multi-omics**

Concerted analyses of various “omics”, or all biological entities participating in the functions of a cell, tissue, or organism. Examples of omics domains include the genome (the DNA sequence and its variations across individuals), epigenome (chemical modifications and structural conformations of the DNA sequence; noncoding RNA molecules), transcriptome (the expressed mRNA transcripts from a genome), proteome (all proteins translated from gene transcripts as well as various chemical modifications of proteins), metabolome (complete set of small molecules and chemicals), and microbiome (microbial species inhabiting human body along with their genome information)

### **Multi-tissue multi-omics systems biology**

A research discipline that utilizes modern high throughput multi-omic technologies to examine diverse types of biomolecules in multiple tissues simultaneously to generate comprehensive global views of molecular interactions within and between tissues

### **Omnigenic disease model**

The model hypothesizes that all genes likely play a role in disease development because of their interconnections with other genes. The further away a gene is from an important disease gene, the weaker effect it likely has on a disease. However, even weak genes can accumulate disease risks over the life span

### **Polygenic disease model**

The model states that more than one gene contribute to disease development

### **Reductionist approach**

The type of research that investigate the effect of individual biological factors, such as genes or proteins, one at a time. Typically it involves activating or inhibiting the molecule to observe downstream molecular or functional changes

## **References**

1. Buniello A, et al., The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*, 2019 47(D1): p. D1005–D1012. [PubMed: 30445434]

2. Rappaport SM and Smith MT, Epidemiology. Environment and disease risks. Science, 2010 330(6003): p. 460–1. [PubMed: 20966241]
3. Hunter DJ, Gene–environment interactions in human diseases. Nat Rev Genet, 2005 6(4): p. 287–98. [PubMed: 15803198]
4. Libby P, et al., Atherosclerosis. Nat Rev Dis Primers, 2019 5(1): p. 56. [PubMed: 31420554]
5. O’Rahilly S, Human genetics illuminates the paths to metabolic disease. Nature, 2009 462(7271): p. 307–14. [PubMed: 19924209]
6. Meng Q, et al., Systems Biology Approaches and Applications in Obesity, Diabetes, and Cardiovascular Diseases. Curr Cardiovasc Risk Rep, 2013 7(1): p. 73–83. [PubMed: 23326608]
7. Boyle EA, Li YI, and Pritchard JK, An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell, 2017 169(7): p. 1177–1186. [PubMed: 28622505]
8. Wray NR, et al., Common Disease Is More Complex Than Implied by the Core Gene Omnigenic Model. Cell, 2018 173(7): p. 1573–1580. [PubMed: 29906445]
9. Fang FC and Casadevall A, Reductionistic and holistic science. Infect Immun, 2011 79(4): p. 1401–4. [PubMed: 21321076]
10. Mazzocchi F, Complexity and the reductionism-holism debate in systems biology. Wiley Interdiscip Rev Syst Biol Med, 2012 4(5): p. 413–27. [PubMed: 22761024]
11. Sittig LJ, et al., Genetic Background Limits Generalizability of Genotype-Phenotype Relationships. Neuron, 2016 91(6): p. 1253–1259. [PubMed: 27618673]
12. Parks BW, et al., Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. Cell Metab, 2013 17(1): p. 141–52. [PubMed: 23312289]
13. Parks BW, et al., Genetic architecture of insulin resistance in the mouse. Cell Metab, 2015 21(2): p. 334–46. [PubMed: 25651185]
14. Schoenrock SA, et al., Perinatal nutrition interacts with genetic background to alter behavior in a parent-of-origin-dependent manner in adult Collaborative Cross mice. Genes Brain Behav, 2018 17(7): p. e12438. [PubMed: 29125223]
15. Zeevi D, et al., Personalized Nutrition by Prediction of Glycemic Responses. Cell, 2015 163(5): p. 1079–1094. [PubMed: 26590418]
16. Zhang G, et al., Differential metabolic and multi-tissue transcriptomic responses to fructose consumption among genetically diverse mice. Biochim Biophys Acta Mol Basis Dis, 2019: p. 165569. [PubMed: 31669422]
17. Chuang HY, Hofree M, and Ideker T, A decade of systems biology. Annu Rev Cell Dev Biol, 2010 26: p. 721–44. [PubMed: 20604711]
18. Yan J, et al., Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. Brief Bioinform, 2018 19(6): p. 1370–1381. [PubMed: 28679163]
19. Hasin Y, Seldin M, and Lusis A, Multi-omics approaches to disease. Genome Biol, 2017 18(1): p. 83. [PubMed: 28476144]
20. Arneson D, et al., Multidimensional Integrative Genomics Approaches to Dissecting Cardiovascular Disease. Front Cardiovasc Med, 2017 4: p. 8. [PubMed: 28289683]
21. Chen Y, et al., Variations in DNA elucidate molecular networks that cause disease. Nature, 2008 452(7186): p. 429–35. [PubMed: 18344982]
22. Blencowe M, et al., Network Modeling Approaches and Applications to Unravelling Non-Alcoholic Fatty Liver Disease. Genes (Basel), 2019 10(12).
23. De Smet R and Marchal K, Advantages and limitations of current network inference methods. Nat Rev Microbiol, 2010 8(10): p. 717–29. [PubMed: 20805835]
24. Holme P, Rare and everywhere: Perspectives on scale-free networks. Nat Commun, 2019 10(1): p. 1016. [PubMed: 30833568]
25. Broido AD and Clauset A, Scale-free networks are rare. Nat Commun, 2019 10(1): p. 1017. [PubMed: 30833554]
26. Albert R, Jeong H, and Barabasi AL, Diameter of the World-Wide Web. Nature, 1999 401: p. 130–131.
27. Barabasi A, Network Science. 2016, Cambridge, UK: Cambridge University Press.

28. Makinen VP, et al., Integrative genomics reveals novel molecular pathways and gene networks for coronary artery disease. *PLoS Genet*, 2014 10(7): p. e1004502. [PubMed: 25033284]
29. Franzen O, et al., Cardiometabolic risk loci share downstream cis- and transgene regulation across tissues and diseases. *Science*, 2016 353(6301): p. 827–30. [PubMed: 27540175]
30. Shu L, et al., Shared genetic regulatory networks for cardiovascular disease and type 2 diabetes in multiple populations of diverse ethnicities in the United States. *PLoS Genet*, 2017 13(9): p. e1007040. [PubMed: 28957322]
31. Price ND, et al., A wellness study of 108 individuals using personal, dense, dynamic data clouds. *Nat Biotechnol*, 2017 35(8): p. 747–756. [PubMed: 28714965]
32. Schussler-Fiorenza Rose SM, et al., A longitudinal big data approach for precision health. *Nat Med*, 2019 25(5): p. 792–804. [PubMed: 31068711]
33. Zhou W, et al., Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature*, 2019 569(7758): p. 663–671. [PubMed: 31142858]
34. Piening BD, et al., Integrative Personal Omics Profiles during Periods of Weight Gain and Loss. *Cell Syst*, 2018 6(2): p. 157–170 e8. [PubMed: 29361466]
35. Shu L, et al., Prenatal Bisphenol A Exposure in Mice Induces Multitissue Multiomics Disruptions Linking to Cardiometabolic Disorders. *Endocrinology*, 2019 160(2): p. 409–429. [PubMed: 30566610]
36. Barabasi AL, Gulbahce N, and Loscalzo J, Network medicine: a network-based approach to human disease. *Nat Rev Genet*, 2011 12(1): p. 56–68. [PubMed: 21164525]
37. Longo DL and Drazen JM, More on Data Sharing. *N Engl J Med*, 2016 374(19): p. 1896–7.
38. Longo DL and Drazen JM, Data Sharing. *N Engl J Med*, 2016 374(3): p. 276–7. [PubMed: 26789876]
39. Brenner S, Sequences and consequences. *Philos Trans R Soc Lond B Biol Sci*, 2010 365(1537): p. 207–12. [PubMed: 20008397]
40. Mardis ER, The \$1,000 genome, the \$100,000 analysis? *Genome Med*, 2010 2(11): p. 84. [PubMed: 21114804]
41. SEQC/maQC-III Consortium, A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat Biotechnol*, 2014 32(9): p. 903–14. [PubMed: 25150838]
42. Bock C, et al., Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat Biotechnol*, 2010 28(10): p. 1106–14. [PubMed: 20852634]
43. Meng Q, et al., Systems Nutrigenomics Reveals Brain Gene Networks Linking Metabolic and Brain Disorders. *EBioMedicine*, 2016 7: p. 157–66. [PubMed: 27322469]
44. Williams EG, et al., Systems proteomics of liver mitochondria function. *Science*, 2016 352(6291): p. aad0189. [PubMed: 27284200]
45. Lempiainen H, et al., Network analysis of coronary artery disease risk genes elucidates disease mechanisms and druggable targets. *Sci Rep*, 2018 8(1): p. 3434. [PubMed: 29467471]
46. Chella Krishnan K, et al., Integration of Multi-omics Data from Mouse Diversity Panel Highlights Mitochondrial Dysfunction in Non-alcoholic Fatty Liver Disease. *Cell Syst*, 2018 6(1): p. 103–115 e7. [PubMed: 29361464]
47. Kurt Z, et al., Tissue-specific pathways and networks underlying sexual dimorphism in non-alcoholic fatty liver disease. *Biol Sex Differ*, 2018 9(1): p. 46. [PubMed: 30343673]
48. Zhao Y, et al., Multi-omics integration reveals molecular networks and regulators of psoriasis. *BMC Syst Biol*, 2019 13(1): p. 8. [PubMed: 30642337]
49. Zhang B, et al., Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell*, 2013 153(3): p. 707–20. [PubMed: 23622250]
50. Parker BL, et al., An integrative systems genetic analysis of mammalian lipid metabolism. *Nature*, 2019 567(7747): p. 187–193. [PubMed: 30814737]
51. Zhao Y, et al., Integrative Genomics Analysis Unravels Tissue-Specific Pathways, Networks, and Key Regulators of Blood Pressure Regulation. *Front Cardiovasc Med*, 2019 6: p. 21. [PubMed: 30931314]

52. Wang D, et al., Comprehensive functional genomic resource and integrative model for the human brain. *Science*, 2018 362(6420).
53. Meng Q, et al., Traumatic Brain Injury Induces Genome-Wide Transcriptomic, Methylomic, and Network Perturbations in Brain and Blood Predicting Neurological Disorders. *EBioMedicine*, 2017 16: p. 184–194. [PubMed: 28174132]
54. Ahn IS, et al., Host Genetic Background and Gut Microbiota Contribute to Differential Metabolic Responses to Fructose Consumption in Mice. *bioRxiv*, 2018.
55. Lee S, et al., Integrated Network Analysis Reveals an Association between Plasma Mannose Levels and Insulin Resistance. *Cell Metab*, 2016 24(1): p. 172–84. [PubMed: 27345421]
56. Lee S, et al., Network analyses identify liver-specific targets for treating liver diseases. *Mol Syst Biol*, 2017 13(8): p. 938. [PubMed: 28827398]

**Box 1.****An additional example of studying genetically perturbed CVD networks.**

One of the main challenges in dissecting genetically perturbed disease networks in human populations is the difficulty to obtain genetic information (accessible from blood or buccal cells) along with molecular profiles of internal tissues (requiring biopsies) relevant to a given disease in the same study population. In a breakthrough study, genetic and transcriptomic datasets from seven vascular and metabolic tissues (liver, subcutaneous adipose, visceral abdominal adipose, skeletal muscle, blood, atherosclerotic-lesion-free internal mammary artery, atherosclerotic aortic artery) were collected from 600 CVD patients in the Stockholm-Tartu Atherosclerosis Reverse Networks Engineering Task (STARNET) study [29]. Compared to other studies that utilized omics data from tissue samples lacking specific disease information [28, 30], the STARNET study pointed to the important of examining tissues from disease populations, as many of the gene regulatory relations were unique in these disease tissues and proven more informative for inferring candidate genes for CVD GWAS loci. For instance, *PCSK9*, a known cholesterol regulator important in the *liver tissue under physiological conditions*, was found to be a key network regulator in the *adipose tissue in CVD patients*, thus supporting context-specific gene functions and disease mechanisms. Before this study, the role of *PCSK9* in CVD was considered to be primarily through the liver tissue, and the results from STARNET added adipose tissue as an important action site for *PCSK9* in CVD pathogenesis. The importance of the adipose tissue as a central regulator of CVD risks in STARNET is in agreement with the conclusions from other independent studies utilizing non-disease tissues discussed in the main text [28, 30]. However, the unique insights revealed through tissues from a disease population that cannot be retrieved from non-disease-specific individuals emphasize the need for tissue biopsies in clinical studies.

**Box 2.****Additional examples of multi-tissue multi-omics systems biology studies of environmental factors.**

High fructose consumption has emerged as a significant risk for cardiometabolic diseases, while DHA, an omega-3 fatty acid, has been associated with beneficial effects. To understand the molecular mechanisms underlying these contrasting dietary effects, a recent study interrogated the global transcriptome and epigenome of two brain regions of a rat model – hypothalamus which is the control center of appetite and metabolism, and hippocampus which is important for cognitive and feeding behavior [43]. This multi-tissue multi-omics study uncovered broad impact of long-term high fructose consumption on numerous tissue-specific and cross-tissue processes ranging from metabolic and immune pathways to neuronal processes and extracellular matrix organization, and predicted extracellular matrix genes such as *Bgn* and *Fmod* as key mediators of the fructose effects on brain gene networks. Further integration of the rat multi-omics data with human GWAS studies of a broad spectrum of diseases supports that the genes and pathways affected by fructose are significantly enriched for genes involved in human cardiometabolic disorders. Interestingly, dietary supplement with DHA, an omega-3 fatty acid, was found to reverse fructose-perturbed pathways and networks in both brain tissues. This study examining two diets, two brain regions, and both the transcriptome and epigenome provides unique insights into the converging CNS networks involved in cardiometabolic diseases that are modulated by two opposing diets, suggesting the promise of using diets to modify disease networks.

Although not a multi-tissue study, another carefully designed multi-omics systems study serves as an excellent example to uncover novel insights of gene-by-diet interactions for cardiometabolic disorders that could not be retrieved by any individual layer of omics data [44]. This study examined the genome as well as the liver transcriptome, metabolome, and proteome from 80 mouse strains with different genetic composition under a chow diet or a high fat diet condition. Their key finding regarding the gene-by-environmental impact on mitochondria function in cardiometabolic diseases was revealed by signals across different omics domains, such as the link between protein D2HGDH and metabolite D-2-hydroxyglutarate, the BCKDHA protein connecting to the gene *Bckdhb*, and the association between protein COX7A2L and the mitochondrial supercomplex assembly. Therefore, broadening the coverage of multi-omics domains can better piece together the cascades of molecular events that execute key metabolic functions that determine gene-by-environment interactions.

### Clinician's Corner

- Complex diseases manifest as subtypes determined by unique genetic makeup and environmental exposures, making the one-size-fit-all approaches to diagnosis and therapy less optimal for individual patients.
- Comprehensive maps of tissues and molecular entities influenced by different risk factors offer insights into the determinants or markers of health or disease to guide more accurate and individualized diagnosis and treatment strategies.
- Multi-tissue multi-omics systems biology has emerged as a powerful research discipline that fully leverages modern technologies and big data analytics to deconvolute the molecular interactions across tissues to dissect disease complexity.
- Implementation of multi-tissue multi-omics systems biology in clinical settings remains challenging due to the need for collection of diverse types of biospecimen. However, it is essential to push for such efforts to gradually close the gap in translational medicine.

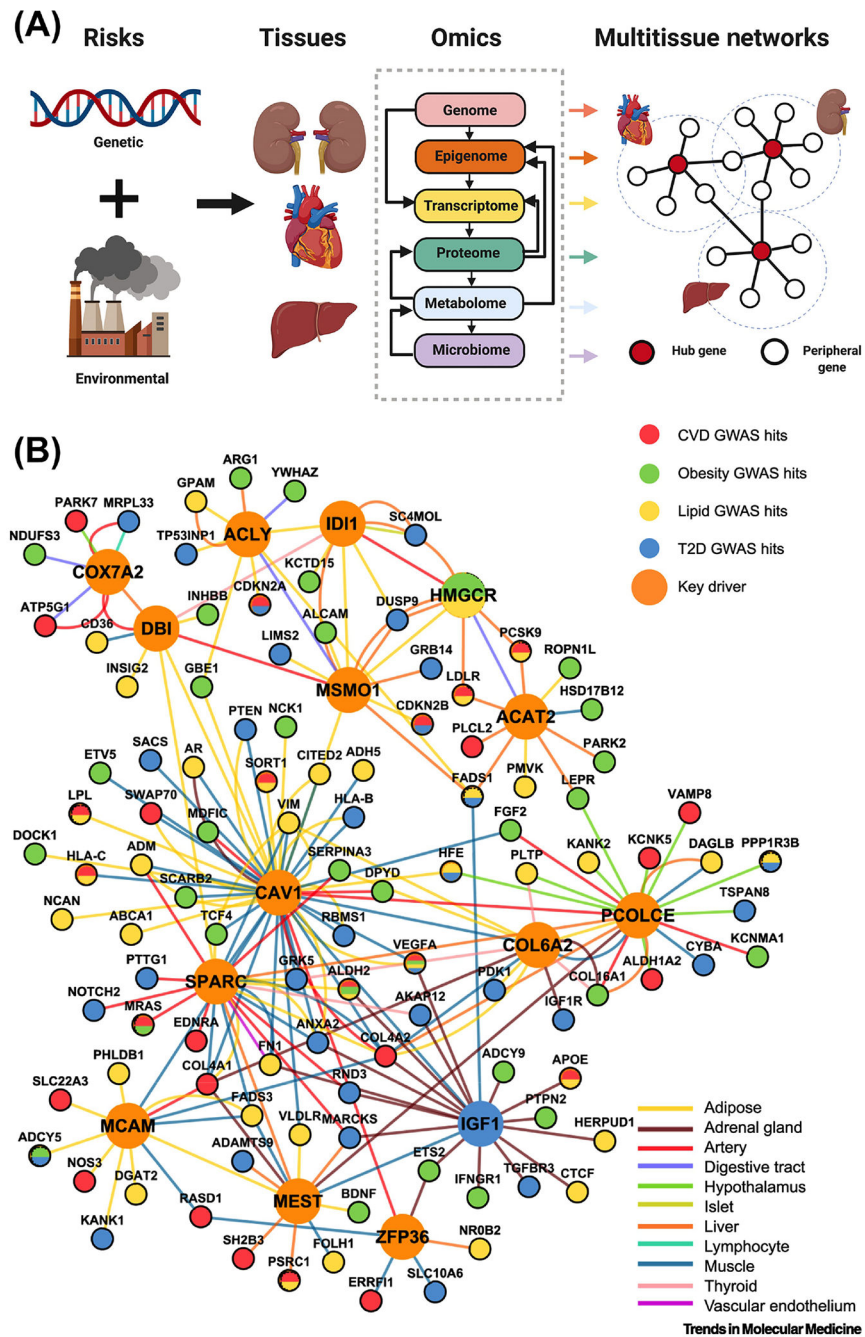


### Outstanding Questions

- What is the systems impact of the exposome? Given the limited advances in our understanding of environmental exposures and gene-by-environment interactions in disease pathogenesis, coordinated efforts to construct multi-tissue multi-omics data repositories and knowledgebases for nutrients/diets, environmental pollutants, toxins, drugs, pathogens, and physical activities are warranted to better control these modifiable health factors.
- Do genetically and environmentally perturbed networks converge or diverge? Detailed partitioning and comparison of networks affected by genetic risks and environmental risks are required to guide disease subtyping and personalized medicine.
- Do complex diseases interconnect through multi-tissue networks? Growing epidemiological evidence points to comorbidity of multiple disease conditions (e.g., metabolic and immune influences on brain functions). Multi-tissue multi-omics systems biology has the unique power to offer deeper insights into inter-disease connections to guide strategies that normalize networks linking to multiple diseases.
- What are the vulnerable cell types and cell-cell interactions in complex diseases? With the advent and maturation of single cell multi-omics technologies, a shift from tissue-based studies to higher resolution cellular studies is needed and attainable.
- How can we use molecular networks to guide drug discovery and network medicine? The field should fully utilize the comprehensive cross-tissue molecular network models to prioritize targets and formulate novel hypotheses to guide preclinical and clinical studies.
- What are the similarities and differences in disease networks between human and model organisms? The lack of cross-species comparative studies has been a limiting factor for translational successes in drug discovery. Multi-species studies are warranted to improve our understanding of between-species disparities to guide translational studies.

### Highlights

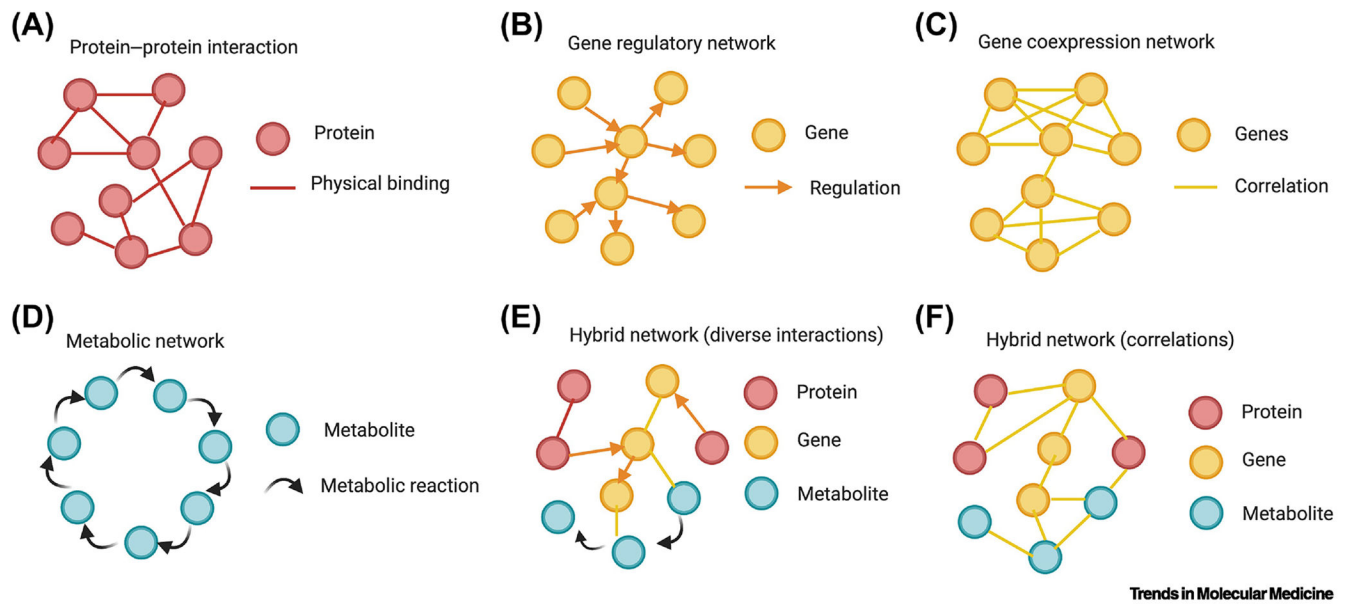
- Recent advances in omics technologies have empowered multi-tissue multi-omics systems biology, a discipline that aims to dissect the multidimensional complexities of human diseases.
- Studies integrating large-scale genetic associations with other omics have resolved tissue-specific molecular networks and pathways perturbed by genetic risks of diseases.
- Systematic investigation of multi-omic domains across tissues in response to environmental exposure such as diet, chemicals, and pathogens has revealed fresh insights into the target tissues, genes, and molecular networks underlying environmental risks of diseases.
- The holistic within- and between-tissue networks of diseases unraveled by systems biology offer comprehensive mechanistic insights and help formulate data-driven hypotheses to guide network-based medicine targeting specific tissues, genes, and pathways tailored to specific risks.



**Figure 1. The concept, framework, and an example of multi-tissue multi-omics systems biology of complex diseases.**

A) Complex diseases are the result of causal risks that affect tissue functions by perturbing multi-omics molecular entities which interact in tissue-specific gene networks. Multi-tissue multi-omics systems biology aims to identify the gene subnetworks in multiple tissues (depicted in dotted circles) relevant to a particular complex disease that are influenced by genetic and/or environmental risks. A molecular network is comprised of nodes representing molecular entities such as genes or proteins and edges that connect the nodes. Network hubs (red nodes) have more connections than peripheral nodes (white nodes) and likely play more

important roles in disease etiology. B) Cross-tissue gene networks shared by CVD and T2D. GWAS candidate genes of various cardiometabolic diseases are network nodes indicated with different colors. Key driver or hub genes are large nodes. Edge color denotes tissue origin of the network connections between nodes. Figure is obtained from Shu et al. [30] under open access agreement and author permission.



**Figure 2. Main types of molecular networks.**

A) Protein-protein interaction networks. B) Gene regulatory networks. C) Gene coexpression networks. D) Metabolic networks. E) Hybrid networks based on various interaction types. F) Hybrid networks based on correlations.

**Table 1.**

Examples of multi-tissue multiomics studies of molecular networks affected by genetic, environment, and gene-by-environment interactions for complex diseases.

Category	Complex disease/trait	Species	Tissue/cell types	Omics layers	References
Genetic centric	CVD	Human	Vascular endothelial cells, liver, adipose, blood	GWAS, transcriptome, eQTLs	[28]
	CVD, cardiometabolic disease	Human	6 vascular and metabolic tissues, blood	GWAS, transcriptome, eQTLs	[29]
	CVD	Human	6 vascular and metabolic	GWAS, transcriptome, eQTLs	[45]
	CVD, T2D	Human	~20 tissues	GWAS, transcriptome, eQTLs	[30]
	Nonalcoholic fatty liver disease (NAFLD)	Mouse	Liver, adipose	GWAS, transcriptome, eQTLs	[46]
	NAFLD (sexual dimorphism)	Mouse	Liver, adipose	GWAS, transcriptome, eQTLs	[47]
	Psoriasis	Human	Blood, skin	GWAS, EWAS, TWAS, eQTLs	[48]
	Alzheimer's disease	Human	Prefrontal cortex, visual cortex, cerebellum	GWAS, transcriptome, eQTL	[49]
	Lipid metabolism	Mouse	Plasma, liver	Genetics, proteomics, lipidomics	[50]
	Hypertension	Human	>40 tissues	GWAS, transcriptome, eQTL	[51]
Psychiatric diseases	Human	Prefrontal cortex, temporal cortex, cerebellum	Bulk and single cell transcriptome Epigenome (Hi-C, ATAC-seq) Transcription factor binding (ChIP-seq)	[52]	
Environment-centric	Cardiometabolic (Bisphenol A)	Mouse	Hypothalamus, liver, adipose	Transcriptome, DNA methylome, human GWAS	[35]
	Cardiometabolic (Fructose)	Rat	Hypothalamus, hippocampus	Transcriptome, DNA methylome, human GWAS	[43]
	Prediabetes (viral infection, immunization, antibiotic)	Human	Plasma, serum, PBMC, stool, nares	Transcriptome, proteome, cytokines, metabolome, microbiome	[33]
	Cardiometabolic (weight gain, weight loss)	Human	Plasma, serum, PBMC, stool	Transcriptome, proteome, cytokines, metabolome, microbiome	[34]
	Traumatic brain injury	Mouse	Hippocampus, blood	Transcriptome, DNA methylome, human GWAS	[53]
Gene by environment interactions	Obesity (high fat high sucrose diet)	Mouse	Adipose, plasma	GWAS, transcriptome, metabolome, microbiome	[12]
	Insulin resistance (high fat high sucrose diet)	Mouse	Liver, adipose	GWAS, transcriptome	[13]
	Cardiometabolic (high fructose diet)	Mouse	Liver, adipose, hypothalamus, fecal/cecal samples	Transcriptome, gut microbiome	[16, 54]
Other	Multiple	Human	Saliva, stool, urine	Genome, metabolomes, proteomes, microbiomes	[31]
	Diabetes, CVD	Human	Plasma, PBMC, stool	Genome, transcriptome, proteome, immunome, metabolome, microbiome	[32]
	Insulin resistance	Human	Hepatocytes, myocytes, adipocytes, liver, adipose	DNAase-seq, ChIP-seq, Transcriptome	[55]

Category	Complex disease/trait	Species	Tissue/cell types	Omics layers	References
	NAFLD, hepatocellular carcinoma	Human	46 tissues, with a focus on liver, adipose, and muscle	Transcriptome, Transcription factor binding, protein-protein interactions, metabolic reactions	[56]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript