



# Systems Medicine as a Transforming Tool for Cardiovascular Genetics

# 11

Melanie Boerries and Tanja Zeller

## Contents

|        |   |     |
|--------|---|-----|
| 11.1   | Introduction .....  | 360 |
| 11.2   | Methods in Systems Medicine .....   | 361 |
| 11.2.1 | Networks and Pathways .....   | 363 |
| 11.2.2 | Cohort Studies and Biobanks .....   | 365 |
| 11.3   | Systems Medicine in Cardiovascular Disease .....                            | 366 |
| 11.3.1 | Systems Medicine Approaches to Reveal Novel Molecular Pathways .....        | 366 |
| 11.3.2 | Systems Medicine Approaches to Reveal Novel Cardiovascular Biomarkers ..... | 367 |
| 11.4   | Challenges in Systems Medicine .....  | 369 |
| 11.4.1 | Big Data: Harmonization of Data/Methods/Formats .....                       | 370 |
| 11.4.2 | Data Exchange, Data Safety, and Ethical Aspects .....                       | 371 |
| 11.4.3 | Interdisciplinary Communication .....                                       | 371 |
| 11.5   | Conclusion and Perspectives .....   | 371 |
| 11.5.1 | The Future of Systems Medicine in Cardiovascular Research .....             | 371 |
|        | References .....  | 374 |

---

M. Boerries

Faculty of Medicine, Institute of Medical Bioinformatics and Systems Medicine, University Medical Center, University of Freiburg, Freiburg, Germany

German Cancer Consortium (DKTK), Partner Site Freiburg, Freiburg, Germany

German Cancer Research Center (DKFZ), Heidelberg, Germany

e-mail: [m.boerries@dkfz-heidelberg.de](mailto:m.boerries@dkfz-heidelberg.de)

T. Zeller (✉)

Clinic for General and Interventional Cardiology, University Heart Center Hamburg, Hamburg, Germany

German Center for Cardiovascular Research (DZHK), Partner site Hamburg/Lübeck/Kiel, Hamburg, Germany

e-mail: [t.zeller@uke.de](mailto:t.zeller@uke.de)

---

**Abstract**

Cardiovascular diseases represent one of the most important causes of morbidity and mortality worldwide. The pathogenesis of cardiovascular disease is complex and remains elusive. Remarkable progress in deciphering the molecular mechanisms of cardiovascular disease has been achieved in different fields of research, ranging from basic-experimental research to molecular epidemiology to clinical research. Each of these isolated fields have successfully improved the pathophysiological understanding in cardiovascular disease. Within the last years, systems medicine has emerged to study the complex genetic, molecular, and physiological interactions leading to cardiovascular diseases by integrating and combining multilevel data sets from different research fields.

This chapter provides an overview of the current understanding of systems medicine and the computational and epidemiological tools applied. First, applications of systems medicine in cardiovascular research are described and challenges and opportunities that arise with systems medicine as a promising tool for cardiovascular genetics are discussed.

---

## 11.1 Introduction

Cardiovascular disease (CVD) comprises numerous complications of the heart, many of which are related to impaired blood flow that might lead to congestive disorders, arrhythmia, or heart valve problems [1]. Causes for organ-wide failure or disorders often have their origin at the subcellular, molecular level, whose effects spread to the level of the whole heart, eventually leading to its spontaneous failure. Thus, to understand the origin and progression of CVD, multiple levels of cellular function—from genomics to transcriptomics and proteomics—to cellular organization—from cell-cell communication, adhesion, and movement—to organ organization including heart physiology and the phenotypic and clinical characteristics of an individual, have to be considered—and must be studied in a dynamic context. Knowledge from each individual level has already been accumulated—although somewhat fragmented—and has been made increasingly accessible with digitalization [2]. To provide a detailed and comprehensive picture, now, we have to learn how to “put the individual pieces together” by learning how to best link and systematically integrate multidimensional datasets.

The concept of systems biology tackles exactly these questions [3, 4]. Systems biology aims at understanding inter- and intracellular dynamic processes through the integration of various high- and low-throughput quantitative data by using mathematical approaches. The concept of a mathematical description of molecular and cellular processes was already applied in simple biological systems more than 50 years ago [5–7]. Driven forward by the development of molecular biology and genetics as research disciplines, the subsequent decades showed large advancements to link molecular events to diseases. In recent years, the generation and evolution of high-dimensional and global molecular data, which are summarized as omics data

(i.e., genomics, transcriptomics, proteomics, metabolomics, lipidomics, glycomics, as well as phenomics or radiomics), have revolutionized approaches to complex biological phenomena. Owing to this development, the last years have witnessed spectacular successes in the identification of new loci involved in the susceptibility to CVD [8–12]. The underlying pathophysiological mechanisms of CVD remain, however, still to be determined. Recently, the idea of a systematic integration of multidimensional datasets in biomedical research and clinical management has percolated through virtually all medical disciplines and has been expanded as the field of systems medicine.

Systems medicine can be described as the implementation of systems biology approaches into medical research (<https://www.casym.eu> [13]). According to this description, a more comprehensive and detailed picture and thus an improvement of disease understanding is taken by the combination of multidimensional omics data from omics-based science, systems biology, network theory, and clinical experience and medical informatics tools for translating the knowledge into improved patient care [14].

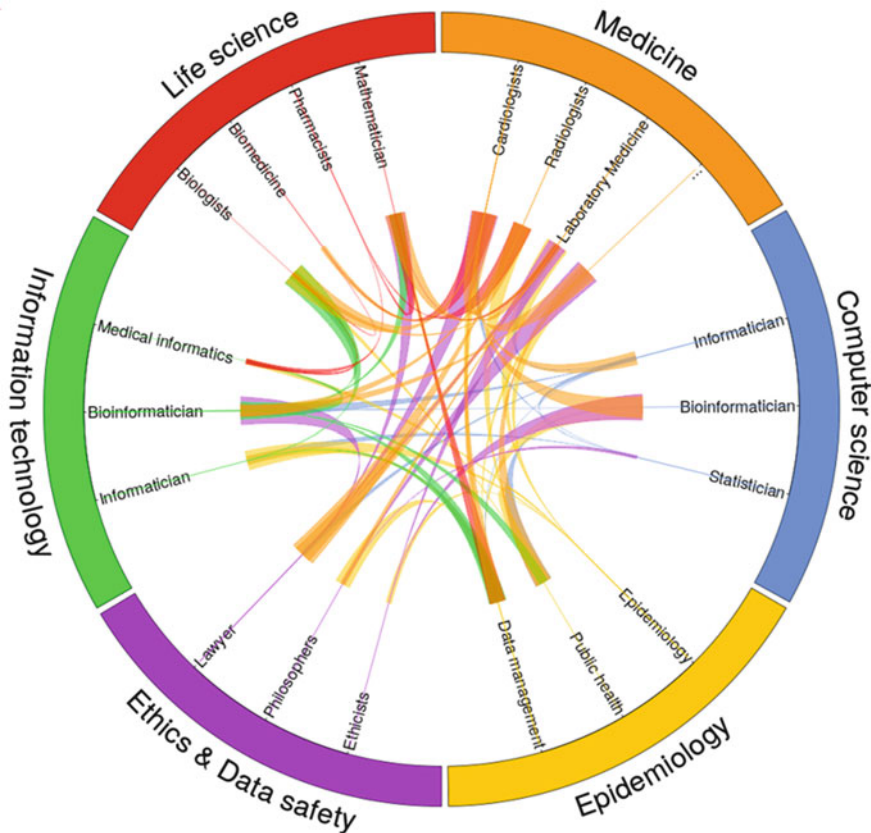
Consequently, systems medicine as a new research field does not only rely on bioinformatics and mathematics, but also on fruitful collaborations across disciplinary boundaries additionally involving computational experts, clinicians, engineers, data manager(s), as well as epidemiologists and researchers in life science (s) (Fig. 11.1).

This chapter provides an overview of the status of systems medicine in the field of cardiovascular disease, with a particular view on the computational methods and tools currently applied, and discusses the opportunities and challenges that come along with the implementation of systems medicine into cardiology.

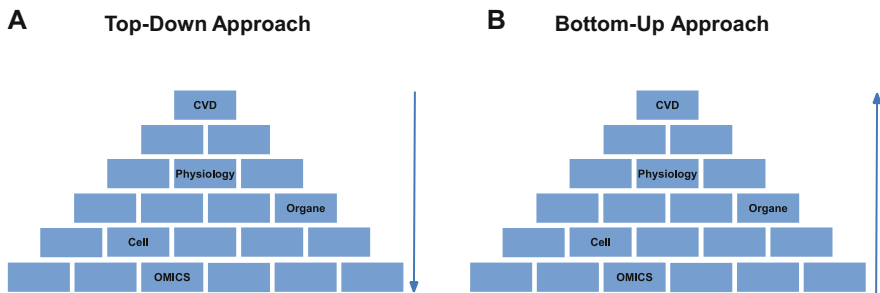
---

## 11.2 Methods in Systems Medicine

Despite technological advances, the main conceptual challenge remains the integration of complex data across entities, space, and time. The difficulty of data integration is deeply rooted in the emergence of novel behavior stemming from molecular and cellular self-organization, which does not allow extrapolation of knowledge across biological processes. For example, current knowledge of protein dynamics cannot explain the emergence of the transcriptional landscape of a cell. To circumvent this problem, the modeling of biological systems generally pursues two major approaches: a top-down and a bottom-up approach. The top-down approach starts from statistical analysis of large-scale omics data with the goal to discern regulatory patterns between different experimental conditions or patient samples. While it allows an unbiased, holistic view of the biology, it usually provides only a limited understanding of molecular mechanisms and thus remains mostly correlative instead of causative. In the case of CVD, a top-down approach can stratify patients with respect to their disease state and pathway activity at a given time point, without, however, considering their pathophysiology in detail or understanding the individual cell, e.g., a cardiomyocyte, all of which might be causing the disease (Fig. 11.2). A



**Fig. 11.1** Interdisciplinarity in systems medicine. Different disciplines jointly work together in a highly connected environment



**Fig. 11.2** Principles of systems biology. (a) The top-down approach allows an unbiased, holistic view of a biological system, without understanding the molecular mechanisms in detail. In contrast, the bottom-up approach attempts to understand and capture molecular mechanisms and interaction with the aim to transfer this knowledge on larger biological network (b). Figure modified from Michael Liebmann, Guest Commentary, Bio IT World, March, 2004

bottom-up approach tries to understand disease starting from molecular interactions and integrating these into larger biological networks. Such an approach requires quantitative dynamic data, which is often not available in sufficient detail, often limiting model up-scaling, and in a limited medical impact.

### 11.2.1 Networks and Pathways

A different approach toward understanding global properties of biological systems is the use of network theory. Cellular behavior is controlled through protein–protein interactions. This so-called interactome can be abstracted into a network, whereby proteins corresponding to nodes and edges represent possible interactions between them. The virtue of this approach is the use of network theory tools to study the importance of individual proteins, to allow dynamic simulation, and to study the effect of network rewiring on the network state, namely, cellular phenotype [15, 16]. The resulting network topology provides insight into organizational principles of the cell. Biological networks are hierarchically organized with few central hubs being connected to many nodes and with many nodes having only few interactors [17]. This makes biological networks robust to random deletions of nodes, i.e., failure of a protein/gene mutation, but fragile to targeted interventions with hub genes or proteins, which often then leads to disease. Likewise, these hub genes/proteins are potential drug targets for a given phenotype [18]. The topology of biological networks has often been exploited for functional enrichment and annotation of genes/proteins of interest [19]. Random [20] and network diffusion methods [21, 22] allow to detect modules and subnetworks among differentially regulated genes or sets of gene mutations, which can subsequently be studied in details. In particular, network diffusion is an interesting approach to find genetic predisposition toward disease and has been successfully applied to cancer [23] or autism [24].

Using this network approach, it is possible to broaden the view from a single protein to larger interaction networks and to explore complex interactions and disease progression in the necessary unbiased way.

It needs to be kept in mind that network analyses crucially depend on the underlying knowledge on network nodes and edges. Various public protein–protein interaction databases exist, as the Biological General Repository for Interaction Datasets [25] (<http://www.thebiogrid.org>), the Biomolecular Interaction Network Database [26] (<https://omictools.com/bind-tool>), the Database of Interacting Proteins [27] (<http://dip.doe-mbi.ucla.edu>), the Molecular INTeraction database [28] (<https://omictools.com/mint-3-tool>), the IntAct molecular interaction database [29] (<http://www.ebi.ac.uk/intact>), or the Human Protein Reference Database [30] (<http://www.hprd.org>). A common protein–protein interaction (PPI) tool is the STRING database (<https://string-db.org/>), which includes indirect (functional) as well as direct interactions (physical) and also predicted protein–protein associations based on co-expression data [31]. From these PPI networks, molecular function and disease association can be predicted based on network topology [32]. However, the identification of relevant network regions (modules), which are linked to diseases

and biological functions, is still a major challenge and goal in the field of systems biology/medicine [33]. The following aspects make it even more difficult: (a) search for a subnetwork (module) depends on the size of the biological network and is in general time consumable, (b) the usage of different omics platforms (sequencing, mass spectrometry, or microarrays) comprise the problems of normalization, batch correction, annotation, and also completeness. Moreover, our biological understanding is incomplete, e.g., only approximately 15% of all protein–protein interactions are known. To gain insight into the interaction properties, network diffusion-based approaches have been developed to calculate a global measure of network proximity by simulation the diffusion of quantity throughout a network. This idea is implemented in the tools HotNet [22] or Response Net [34]. Network-based stratification (NBS) integrates mutations into PPIs to stratify cancer into informative subtypes. This elucidates whether patients share similar network regions based on their mutations. These networks are also predictive on patient survival or tumor response therapy and can be used for RNA expression patterns with the aim to obtain similar information in the absence with DNA sequencing [35]. In this context, weighted gene co-expression network analysis (WGCNA) [36] uses network construction to identify biological functional relevant modules and its respective key drivers. Using this approach, Horvath et al. could identify a molecular target in glioblastoma as a key gene within a gene co-expression module that has a major impact on cell proliferation. The emphasis on pathway modules reduces the internal complexity, e.g., comparison of 20 modules as compared to 20,000 genes. Moreover, intramodular connectivity can be used to identify key drivers, so-called hub nodes. In this context, WGCNA were used to identify the two genes, G6PD and S100A7, that are related to coronary artery disease [37].

Another network approach is transcriptional regulatory associations in pathways (TRAP). In contrast to other gene–gene network approaches, this network inference algorithm is able to identify gene–pathway transcriptional regulatory relationship, in which a single gene is determined as a transcriptional regulator [38]. In this algorithm, regulator–pathway associations are derived from transcriptional co-regulation in more than 3000 individual microarray experiments with transcription factor-coding genes, taken from the Riken Transcription Factor Database and Regulator Pathways.

In order to obtain a functional insight into gene regulation, Subramanian et al. developed a powerful analytical method called gene set enrichment analysis (GSEA) that is based on focusing on gene sets. The latter is a group of genes that share common biological properties/function. GSEA determines whether the genes are randomly distributed or specifically assigned to a gene set [39]. Such an analysis benefits greatly from the annotation of biological databases, e.g., ConsensusPath DB or Kyoto Encyclopedia of Genes and Genomes (KEGG). ConsensusPathDB is a meta pathway database that provides a comprehensive collection of human (as well as mouse and yeast) molecular interaction data together with a web interface (<http://consensuspathdb.org>) that offers several computational methods and visualization tools to analyze these data. This database comprises network modules interaction, biochemical pathways, and functional information to support data interpretation of

high-throughput data with the aim to prioritize specific biological function and signaling pathways [40, 41]. KEGG is a database resource of large molecular datasets based on genome sequencing and other high-throughput experimental technologies to obtain a high-level functional understanding of biological systems, such as a cell or an organism [42] (<http://www.genome.jp/kegg>). Another database is Reactome that focuses on human pathways, biological processes, and reactions information [43]. It also offers a web interface including tools for pathway browsing and data analysis (<https://reactome.org>). The Gene Ontology Consortium (<http://www.geneontology.org>) has the aim to generate an extensible, ordered, and controlled vocabulary that comprises many eukaryotic processes and is subgrouped into three main ontologies: biological process, molecular function, and cellular component [44].

Further databases for specific diseases and gene expression exist, such as the genotype-tissue expression (GTEx, <https://www.gtexportal.org>) [45] that examines association between genetic variation and gene expression in human tissues, or the Human Protein Atlas (<https://www.proteinatlas.org>) [46] and The Cancer Genome Atlas (TCGA, <https://cancergenome.nih.gov>). The latter links genetic alteration to specific cancer types, whereas the protein Human Protein Atlas maps proteins in cells, tissues, and organs based on both protein and transcriptome data sets. Due to such databases and tools [47], the analysis and understanding of high throughput data, also with regard to CVD, can greatly be improved.

### 11.2.2 Cohort Studies and Biobanks

In order to identify novel risk factors for CVD, large epidemiological cohort studies are particularly suitable. In a cohort study, cardiovascular phenotypes and subsequent fatal and non-fatal events can be studied over time and provide powerful results in identifying relationships between the characteristics of a population and that population's behavior. Two types of cohort studies are prominently used in cardiovascular disease. Prospective studies are carried out from the present time into the future, whereas retrospective studies are carried out at the present time but look to the past to examine disease events and outcomes [48]. The best-known epidemiological cohort study in the cardiovascular field is the Framingham Heart Study (FHS), initiated in 1948, which is currently investigating the third generation of the original participants. The FHS is the longest running prospective cohort study and has contributed enormously to the understanding of various cardiovascular risk factors and to how these factors relate to the overall and cardiovascular-related mortality [49, 50]. The FHS could only be generated based on large populations, generation of a large amount of big data, and by the support of several health institutes. On the basis of the knowledge generated by the FHS, other epidemiological cohort studies were implemented in the last years and have provided additional information on cardiovascular disease risk [51–55].

In the last decades, bio specimens became an important additional resource in epidemiological studies and are collected and stored into various so-called biobanks.

The main bio specimens collected include blood samples, urine, feces, tonsil swaps, saliva, tear fluid, tissue samples and biopsies, as well as genetic material. These materials made modern molecular analyses such as omics analyses feasible on the large scale as these specimens are easily collected and various related phenotypic and linked clinical data are available.

Today, large biobanks are an essential part of a cohort's infrastructure [56, 57] and build the basis for a large part of the biomedical research and consequently are critical for translating advances in molecular biology and new possibilities in the context of systems medicine [56].

---

## 11.3 Systems Medicine in Cardiovascular Disease

In the cardiovascular biomarker field, recent studies that relied on the concept of systems medicine have been proposed and/or applied. A few studies have already successfully identified novel mechanisms of CVD and/or were translated into clinical application.

### 11.3.1 Systems Medicine Approaches to Reveal Novel Molecular Pathways

#### 11.3.1.1 Myocardial Infarction

As a major cause of death, myocardial infarction (MI) is best predicted in middle-aged adults by a positive family history, strongly supporting a genetic basis of MI [58]. Investigating the genetic background of MI with tools applied in systems medicine, a novel pathway and potential drug targets for treatment of cardiovascular disease were identified in elegant studies combining omics and clinical data with bioinformatical and experimental models. In an extended family with several members diagnosed of coronary artery disease (CAD), exome sequencing and linkage analysis identified rare loss-of-function and missense mutations in the genes *GUCY1A3* and the chaperonin containing TCP1 subunit 7 (*CCT7*), respectively [12].

The *GUCY1A3* gene encodes the  $\alpha 1$  subunit of the soluble guanylyl cyclase (sGC). The sGC complex, a heterodimer of the  $\alpha 1$  and a  $\beta 1$  subunit, acts as the receptor for nitric oxide (NO) and catalyzes the formation of the second messenger cGMP [59]. cGMP has several cellular functions, including the inhibition of platelet aggregation, thereby representing an important feature of thrombus formation in MI, and smooth muscle cell relaxation [59]. Further evidence point to a critical involvement of the NO-sGC-cGMP pathway in mediating CAD and MI risk [59].

Transfection of the *GUCY1A3* mutation into human embryonic kidney cells as well as downregulation of *CCT7* by siRNA were preformed to elucidate the functional implications of these variants. These experimental approaches led to a strong reduction of soluble guanylyl cyclase ( $\alpha 1$ -sGC)  $\alpha 1$  levels [12]. Using platelets



extracted from family members carrying either single *GUCY1A3* or *CCT7* mutations, a double mutation, or none of the rare alleles,  $\alpha 1$ -sGC levels and NO-dependent cGMP generation were investigated. Carriers of the digenic mutation exhibited a significant reduction of  $\alpha 1$ -sGC levels and cGMP formation. Subsequently, *Gucy1A3*-deficient mice showed an increased thrombus formation, indicating an increased risk of MI via dysfunctional nitric oxide signaling in rare allele carrying family members.

Further functional evidence that the variants in the risk locus affect *GUCY1A3* gene expression was provided by [59]. Using human samples and cell lines, the SNP rs7692387, located in a DNaseI hypersensitive site, was identified as being involved in the regulation of *GUCY1A3* gene expression. The *GUCY1A3* risk variant rs7692387 seems to act by a modulation of the binding site for a transcription factor, ZEB1, thereby directly affecting the expression of the target gene.

As a consequence of impaired *GUCY1A3* expression, homozygous risk allele carriers exhibited decreased inhibitory effects of sGC stimulation on cell migration, the production of cGMP was inhibited, and platelet aggregation after exposure to an NO donor was impaired. In agreement with these human data, lower *Gucy1a3* expression correlated with more aortic atherosclerosis in mice. By the integration of a series of data derived from experimental settings, human clinical data, and bio specimens, a novel link between impaired soluble-guanylyl-cyclase-dependent nitric oxide signaling and MI risk was identified, possibly providing a new therapeutic target for reducing the risk of MI [12].

### 11.3.2 Systems Medicine Approaches to Reveal Novel Cardiovascular Biomarkers

#### 11.3.2.1 Blood Pressure

Huan et al. [60, 61] conducted a multilevel integration analysis on one of the major cardiovascular risk factors, blood pressure (BP), to identify novel candidate genes involved in BP regulation. Computationally, the authors combined multilevel datasets including genetic, transcriptomic, and phenotype data. Several genes were identified in relation to BP, which jointly explain 5–9% of BP variation. To further seek for molecular key drivers of BP regulation, co-expression sub-networks that were jointly connected by the SH2B adaptor protein 3 (*SH2B3*) were identified [58, 61]. In order to elucidate the molecular role of *SH2B3* and the relation to hypertension, *Sh2b3*<sup>−/−</sup> mice were investigated in response to low-dose angiotensin II treatment [62]. In untreated *Sh2b3*<sup>−/−</sup> mice, kidneys and aortas showed greater levels of inflammation, oxidative stress, and glomerular injury. These effects were accelerated after angiotensin II infusion. A strong indication that the predominant effect of *SH2B3* on BP is mediated by hematopoietic cells was shown by experiments of bone marrow transplantations of *SH2b3*<sup>−/−</sup> into wild-type which reproduced the hypertensive phenotype. Subsequent studies identified the genes of the BP co-expression networks [58, 63] including *CRIP1* (cysteine-rich protein 1). These studies further elucidated that *CRIP1* transcript expression additionally

correlated to measures of cardiac hypertrophy and identified circulating CRIP1 protein levels as a potential biomarker for increased risk for incident stroke, a sequel of high BP [63].

### 11.3.2.2 Microbiome

During the last years, the gut microbiome has attracted increasing attention in many medical fields and emerged as a central factor affecting human health and disease [64] and has been linked to cardiovascular disease. In a metagenome-wide association study, a cross-disease cohort integrative analysis was performed, including data from subjects affected by atherosclerotic disease, cardiometabolic diseases, obesity, type 2 diabetes, liver cirrhosis, or the autoimmune disease rheumatoid arthritis [65].

Compared to the healthy controls, the gut microbiome of atherosclerotic disease patients was significantly different in multivariate analyses and showed separation in principle component and distance-based redundancy analyses, which was supported by the relative reduction in the genera *Bacteroides* and *Prevotella*, and an enrichment in *Streptococcus* and *Escherichia* in atherosclerotic disease patients. Co-abundance network of metagenomically (genetically) linked groups confirmed the major compositional differences in individuals with and without atherosclerotic disease.

To explore the diagnostic value of the gut microbiome composition in relation to atherosclerotic disease, random forest classifiers including several different metagenomics linked groups and KEGG (The Kyoto Encyclopedia of Genes and Genomes) functional modules were selected to construct a mathematical model that predicts clinical indices. A discriminatory ability to distinguish between individuals with and without atherosclerotic disease with an area under the receiver operating curve of 0.86 was observed, demonstrating the presence of atherosclerotic disease-associated features in the gut microbiome that may be further developed into non-invasive and inexpensive biomarkers [65]. Insights into the potential possible changes within the gut microbiome in the disease state were provided by KEGG pathway analyses, showing—among others—a higher potential for sugar and amino acid transport, a lower potential for vitamin synthesis altered potential for homocysteine and tetrahydrofolate metabolism, and an enrichment in virulence factors.

As medication can influence the gut microbiome, subsequent analyses were performed and random-forest classifiers for distinguishing between atherosclerotic disease patients treated with and without medication. Overall, the results propose that although drug treatment may affect the composition of the gut microbiota and thus constitute a confounding factor, medication weakened the disease signal, meaning an even more significant difference would be expected in a cohort free of medication [65].

These results were extended to other related diseases such as obesity, type 2 diabetes, liver cirrhosis, and rheumatoid arthritis, and multi-disease microbiome-based classifiers were extracted from these additional microbiome datasets. The results showed that the gut microbiome of each disease exhibited unique features in functional capacity, in species, and gene compositions.

By using human bio specimens, clinical data and bioinformatical approaches, this study implicates a role of the gut microbiome in heart disease. The data represent a

comprehensive resource for further investigations on the role of the gut microbiome in promoting or preventing cardiovascular disease as well as other related diseases.

There is a considerable power of using the microbiome to separate diseased subjects from healthy subjects as well as to predict responses to treatment or the development of diseases in the absence of treatment [66]. To further guide the utility of the microbiome for robust clinical use, data must be validated in larger and more diverse populations, and methodologies must be standardized across different laboratories [66].

### 11.4 Challenges in Systems Medicine

Systems medicine aims to explore medicine beyond linear relationships and single parameters and includes multiple parameters and spatial conditions to achieve a holistic perspective [67].

Over the last years, emerging and interdisciplinary approaches with extensive tools, novel analytical methods and strategies have rapidly been developed.

Promises are made that systems medicine will lead to an innovation in diagnosis and prognosis of cardiovascular diseases and a better understanding of disease mechanisms and pathology and thereby improved therapeutic options.

Nevertheless, at the current state, the adaptation of systems medicine into medical research and, in the long run, into “real-world” clinical practice is a complex endeavor and researchers are facing multiple challenges (Fig. 11.3).

Here, we will discuss the main challenges that arise from working with large amounts of data and from working in an interdisciplinary team.

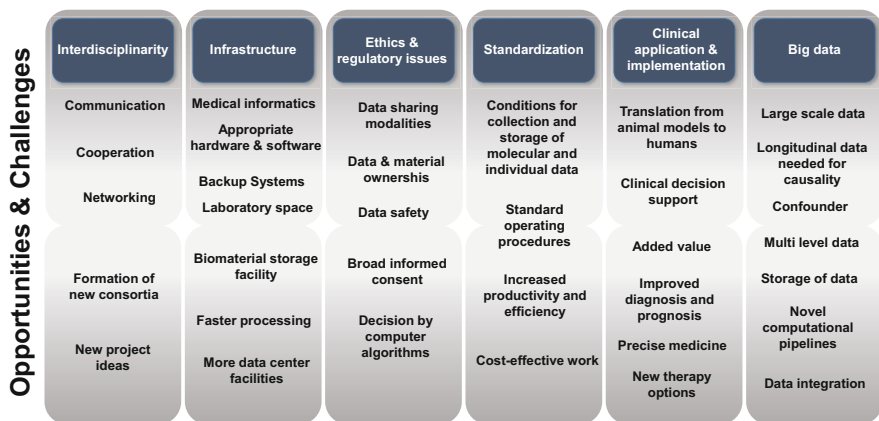


Fig. 11.3 Challenges and opportunities in systems medicine

### 11.4.1 Big Data: Harmonization of Data/Methods/Formats

With the advent of high-throughput technologies, massive data sets are generated that need to be handled, processed, harmonized, and shared [68].

For instance, with the falling sequencing costs, soaring accuracy, and a steadily expanding base of scientific knowledge a vast amount of data, perhaps in the range of petabytes, is generated each year [69]. Thus, some researchers worry that the flood of data could overwhelm the computational pipelines needed for analysis and generate unprecedented demand for storage [69].

Therefore, new infrastructures and information technology systems need to be developed and implemented into the research facilities and clinics. In particular, clinical routine will require different levels of granularity of the information—ranging from a short overview of the patient record in the clinical information system up to an in-depth discussion of potential therapies [70] or to an inclusion for electronic decision support systems [71, 72]. Subsequently, standardization of conditions for collection and storage of molecular and individual data as well as the respective metadata need to be established.

The benefits of harmonizing and pooling research databases will be enormous. Although over the past half-decade both governments and researchers emphasized the importance of harmonizing data as well as collaborative usage of data and samples, the integration and harmonization of different data sets, especially high-throughput data, will continue to be a major challenge [73, 74]. This situation is exacerbated by ethical, legal, and civil rights with regard to the sharing or aggregation of data at an individual level. New database management systems are at the forefront of providing solutions to some of these issues [75]. The idea is that these tools can link different distributed databases with the aim to ensure a secure and effective data analysis. This endeavor can be supported through an interactive and strong collaboration of the different project partners and also through a clear and defined annotation of the diverse data. In this context, BioSHarre (biobank standardization and harmonization for research excellence in the European Union) is a project of the Seventh Framework Program (FP7) and aims to develop tools for data harmonization and standardized IT systems for existing biobanks and cohorts across Europe and apply them to conduct pan-European epidemiological research [76] (<https://www.bioshare.eu/>).

In spite of this, it is important that the different data sets are stored right from the start with the correct annotation and format in the database. This can only be implemented if the consensus on data formats and data annotation and also metadata information is clearly defined. Such data standards have first started with MIAME (minimum information about a microarray experiment) [77] and have found their way into Omics databases like Gene Expression omnibus or Array Express and recently with the Sequence Read Archive. Sequence data on humans from whole exome/whole genome data add an additional layer of controlled access due to ethical reasons. Such databases are the European Genome-Phenome Archive (EGA) [78] the NCBI (NCBI, National Center for Biotechnology Information; <https://www.ncbi.nlm.nih.gov/>) where the information is usually stored.

### **11.4.2 Data Exchange, Data Safety, and Ethical Aspects**

Worldwide, large discussions are currently ongoing regarding the privacy and ownership of data and sharing of data and bio specimens. Traditional models of consent became impractical as more and more large numbers of individuals are included in cohort studies and a much larger amount of data is generated from one individual (e.g., by genome-wide microarray or sequencing approaches). Furthermore, due to the interdisciplinary setting, these large data sets and bio specimens might be shared with internal as well as external partners. Therefore, newer consent strategies include the use of broad or open consent [79] to cover the multilevel assessment of personal, clinical, and biological data as well as the storage and use of data and bio specimens in biobanks [80].

Further issues need to be discussed when medical informatics tools, e.g., electronic decision support and medical informatics systems, might directly be integrated and translated into clinical practice to support patient care [14, 81–83]. For example, is it feasible that clinical decisions will be derived (only) from computer algorithms? What about concerns related to privacy, data protection, and ownership of data?

### **11.4.3 Interdisciplinary Communication**

Besides all the technical difficulties, an important challenge is the multidisciplinary nature that inevitably evolves with systems medicine. Different scientific and clinical disciplines are involved in systems medicine, ranging from epidemiologists and data managers to informaticians, bioinformaticians, statisticians, and life scientists, and most importantly to clinicians. Even within one of these specific groups, several experts (e.g., cardiologist, radiologist, experts in intensive care) are needed for proper interpretation of the data and clinical situation.

A first step in working in an interdisciplinary team is to overcome problems of communication. Each research area has its own language leading to difficulty of understanding and working with each other [13]. However, the closer these disciplines work together, understand, and learn from each other, the more successful will be the progress in system medicine.

---

## **11.5 Conclusion and Perspectives**

### **11.5.1 The Future of Systems Medicine in Cardiovascular Research**

Systems medicine emerged as a powerful tool to study complex diseases by the integration of multidimensional datasets [47]. In the near future, multidimensional, large-scale data are generated, jointly analyzed with phenotypical and clinical

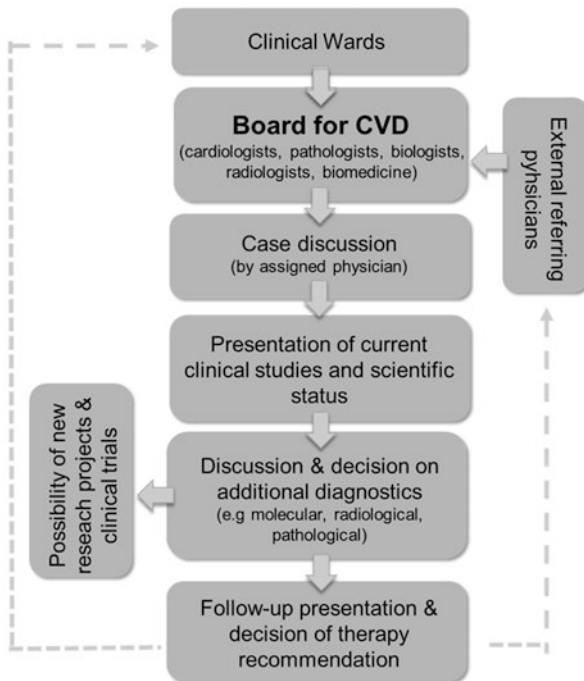
characteristics, hypotheses are functionally examined in basic research models, and findings are translated into epidemiological cohort studies to gain knowledge about the cardiovascular clinical meaning. The expectations are high that the research results will be translated into routine clinical practice.

For an efficient implementation of systems medicine in the cardiovascular field, important cornerstones are: (1) a sustained interplay and communication between experts of multiple disciplines, (2) an optimized usage of resources and infrastructures to efficiently share data, biomaterial, and knowledge, and (3) the extension of the research beyond traditional domains of discovery and disease etiology to accelerate translation into the clinics [84].

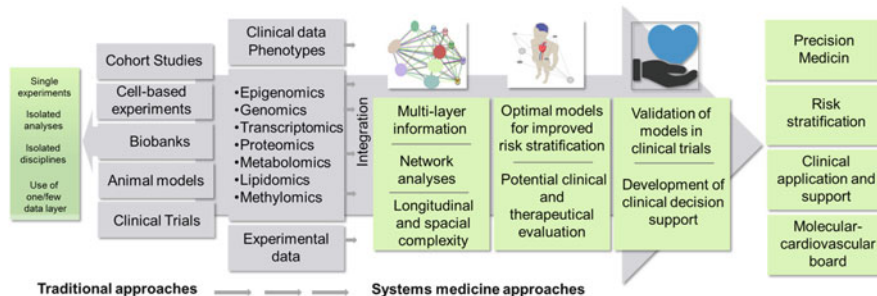
In the cancer field, some of these cornerstones have already been implemented such as the interdisciplinary tumor boards. Therein, the individual patient diagnoses are discussed in order to identify therapeutic opportunities based on extended molecular diagnostics in combination with genome, transcriptome, epigenome sequencing followed by bioinformatics analysis. Such boards are characterized not only by great interdisciplinarity, integrating various clinical research areas, such as oncology, pathology, system biology/system medicine, but also by standardized workflows for patient selection, sample preparation, analysis methods, and reporting. For example, the Molecular Tumor Board of the German Cancer Consortium (DKTK) MASTER program has set itself the task to clarify whether matching treatments with individual molecular profiles lead to improved disease outcome based on the clinical application of whole exome/whole genome and RNA sequencing. This program emphasizes that the main advantage of precision medicine is the collection of characterized molecular data from large patient cohorts and its integration of genetic and clinical information of individual patients. Even if these data are often incomplete, they can be reused for many important future medical questions and research [85].

These kind of “boards” have not been established in the field of cardiovascular disease so far. Certainly, the collection of cardiovascular samples for sequence analysis is much more difficult compared to tumor samples, however, an interdisciplinary board of, e.g., cardiologists, pathologists, nephrologists, systems biologists, bioinformaticians, or further clinical disciplines will tremendously increase the molecular understanding of cardiovascular diseases and therefore will open up new avenues in diagnosis, therapy, and prevention. An example of a proposed cardiovascular board is provided in Fig. 11.4.

Given all the promises, opportunities, but also challenges of systems medicine, the upcoming years will see that researchers of different disciplines are working together to translate results from cardiovascular research into better health strategies and the use of systems medicine as a transforming tool for cardiovascular research (Fig. 11.5).



**Fig. 11.4** Proposed molecular-cardiovascular board. The interdisciplinary board discusses individual patient diagnoses to identify therapeutic opportunities based on extended molecular diagnostics to open new avenues for both research projects and clinical trials



**Fig. 11.5** Systems medicine as a transforming tool in cardiovascular disease. From traditional approaches of single experiments and isolated analyses to large-scale, integrated systems-based approaches to facilitate the translation of discoveries to clinical utility

## References

1. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, De Ferranti SD, Floyd J, Fornage M, Gillespie C, Isasi CR, Jimenez MC, Jordan LC, Judd SE, Lackland D, Lichtman JH, Lisabeth L, Liu S, Longenecker CT, Mackey RH, Matsushita K, Mozaffarian D, Mussolino ME, Nasir K, Neumar RW, Palaniappan L, Pandey DK, Thiagarajan RR, Reeves MJ, Ritchey M, Rodriguez CJ, Roth GA, Rosamond WD, Sasson C, Towfighi A, Tsao CW, Turner MB, Virani SS, Voeks JH, Willey JZ, Wilkins JT, Wu JH, Alger HM, Wong SS, Muntner P. Heart disease and stroke statistics-2017 update: A report from the American Heart Association. *Circulation*. 2017;135:e146–603.
2. Rebhan M. Towards a systems approach for chronic diseases, based on health state modeling. *F1000Res*. 2017;6:309.
3. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet*. 2004;5:101–13.
4. Kitano H. Computational systems biology. *Nature*. 2002;420:206–10.
5. Adam G, Delbruck M. Reduction of dimensionality in biological diffusion processes. San Francisco: W.H. Freeman & Company; 1968.
6. Hopfield JJ. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci U S A*. 1974;71:4135–9.
7. Luria SE, Delbruck M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*. 1943;28:491–511.
8. Auer PL, Stitzel NO. Genetic association studies in cardiovascular diseases: Do we have enough power? *Trends Cardiovasc Med*. 2017;27:397–404.
9. Howson JMM, Zhao W, Barnes DR, Ho WK, Young R, Paul DS, Waite LL, Freitag DF, Fauman EB, Salfati EL, Sun BB, Eicher JD, Johnson AD, Sheu WHH, Nielsen SF, Lin WY, Surendran P, Malarstig A, Wilk JB, Tybjaerg-Hansen A, Rasmussen KL, Kamstrup PR, Deloukas P, Erdmann J, Kathiresan S, Samani NJ, Schunkert H, Watkins H, Do R, Rader DJ, Johnson JA, Hazen SL, Quyyumi AA, Spertus JA, Pepine CJ, Franceschini N, Justice A, Reiner AP, Buyske S, Hindorff LA, Carty CL, North KE, Kooperberg C, Boerwinkle E, Young K, Graff M, Peters U, Absher D, Hsiung CA, Lee WJ, Taylor KD, Chen YH, Lee IT, Guo X, Chung RH, Hung YJ, Rotter JI, Juang JJ, Quertermous T, Wang TD, Rasheed A, Frossard P, Alam DS, Majumder AAS, Di Angelantonio E, Chowdhury R, Chen YI, Nordestgaard BG, Assimes TL, Danesh J, Butterworth AS, Saleheen D. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. *Nat Genet*. 2017;49:1113–9.
10. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, Zeng L, Ntalla I, Lai FY, Hopewell JC, Giannakopoulou O, Jiang T, Hamby SE, Di Angelantonio E, Assimes TL, Bottinger EP, Chambers JC, Clarke R, Palmer CNA, Cubbon RM, Ellinor P, Ermel R, Evangelou E, Franks PW, Grace C, Gu D, Hingorani AD, Howson JMM, Ingelsson E, Kastrati A, Kessler T, Kyriakou T, Lehtimaki T, Lu X, Lu Y, Marz W, McPherson R, Metspalu A, Pujades-Rodriguez M, Ruusalepp A, Schadt EE, Schmidt AF, Sweeting MJ, Zalloua PA, Alghalayini K, Keavney BD, Kooner JS, Loos RJF, Patel RS, Rutter MK, Tomaszewski M, Tzoulaki I, Zeggini E, Erdmann J, Dedoussis G, Bjorkegren JLM, Schunkert H, Farrall M, Danesh J, Samani NJ, Watkins H, Deloukas P. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet*. 2017;49:1385–91.
11. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjornes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lytykainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, De Andrade M, De Vries PS, Van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L, Lieb W,



- Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissono D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015;47:1121–30.
12. Erdmann J, Stark K, Esslinger UB, Rumpf PM, Koesling D, De Wit C, Kaiser FJ, Braunholz D, Medack A, Fischer M, Zimmermann ME, Tennstedt S, Graf E, Eck S, Aherrahrou Z, Nahrstaedt J, Willenborg C, Bruse P, Braenne I, Nothen MM, Hofmann P, Braund PS, Mergia E, Reinhard W, Burgdorf C, Schreiber S, Balmforth AJ, Hall AS, Bertram L, Steinhagen-Thiessen E, Li SC, Marz W, Reilly M, Kathiresan S, McPherson R, Walter U, Ott J, Samani NJ, Strom TM, Meitinger T, Hengstenberg C, Schunkert H. Dysfunctional nitric oxide signalling increases risk of myocardial infarction. *Nature.* 2013;504:432–6.
  13. Wolkenhauer O, Auffray R, Jaster R, Steinhoff G, Dammann O. The road from systems biology to systems medicine. *Pediatr Res.* 2013;73:502–7.
  14. Fischer T, Brothers KB, Erdmann P, Langanke M. Clinical decision-making and secondary findings in systems medicine. *BMC Med Ethics.* 2016;17:32.
  15. Barabasi AL, Albert R. Emergence of scaling in random networks. *Science.* 1999;286:509–12.
  16. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. *Nature.* 1998;393:440–2.
  17. Strogatz SH. Exploring complex networks. *Nature.* 2001;410:268–76.
  18. Rajalingam K, Schreck R, Rapp UR, Albert S. Ras oncogenes and their downstream targets. *Biochim Biophys Acta.* 2007;1773:1177–95.
  19. Hu JX, Thomas CE, Brunak S. Network biology concepts in complex disease comorbidities. *Nat Rev Genet.* 2016;17:615–29.
  20. Chipman KC, Singh AK. Predicting genetic interactions with random walks on biological networks. *BMC Bioinform.* 2009;10:17.
  21. Bersanelli M, Mosca E, Remondini D, Castellani G, Milanese L. Network diffusion-based analysis of high-throughput data for the detection of differentially enriched modules. *Sci Rep.* 2016;6:34841.
  22. Vandin F, Upfal E, Raphael BJ. Algorithms for detecting significantly mutated pathways in cancer. *J Comput Biol.* 2011;18:507–22.
  23. Network CGAR. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature.* 2013;499:43–9.
  24. Mosca E, Bersanelli M, Gnocchi M, Moscatelli M, Castellani G, Milanese L, Mezzelani A. Network diffusion-based prioritization of autism risk genes identifies significantly connected gene modules. *Front Genet.* 2017;8:129.
  25. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 2006;34:D535–9.
  26. Bader GD, Donaldson I, Wolting C, Ouellette BF, Pawson T, Hogue CW. Bind – The biomolecular interaction network database. *Nucleic Acids Res.* 2001;29:242–5.
  27. Xenarios I, Rice DW, Salwinski L, Baron MK, Marcotte EM, Eisenberg D. DIP: the database of interacting proteins. *Nucleic Acids Res.* 2000;28:289–91.
  28. Zanzoni A, Montecchi-Palazzi L, Quondam M, Ausiello G, Helmer-Citterich M, Cesareni G. MINT: a Molecular INTERaction database. *FEBS Lett.* 2002;513:135–40.
  29. Hermjakob H, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, Vingron M, Roechert B, Roepstorff P, Valencia A, Margalit H, Armstrong J, Bairoch A, Cesareni G, Sherman D, Apweiler R. IntAct: an open source molecular interaction database. *Nucleic Acids Res.* 2004;32:D452–5.
  30. Peri S, Navarro JD, Amanchy R, Kristiansen TZ, Jonnalagadda CK, Surendranath V, Niranjan V, Muthusamy B, Gandhi TK, Gronborg M, Ibarrola N, Deshpande N, Shanker K, Shivashankar HN, Rashmi BP, Ramya MA, Zhao Z, Chandrika KN, Padma N, Harsha HC,

- Yatish AJ, Kavitha MP, Menezes M, Choudhury DR, Suresh S, Ghosh N, Saravana R, Chandran S, Krishna S, Joy M, Anand SK, Madavan V, Joseph A, Wong GW, Schiemann WP, Constantinescu SN, Huang L, Khosravi-Far R, Steen H, Tewari M, Ghaffari S, Blobe GC, Dang CV, Garcia JG, Pevsner J, Jensen ON, Roepstorff P, Deshpande KS, Chinnaiyan AM, Hamosh A, Chakravarti A, Pandey A. Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res.* 2003;13:2363–71.
31. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, Von Mering C. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015;43:D447–52.
  32. Wang X, Gulbahce N, Yu H. Network-based methods for human disease gene prediction. *Brief Funct Genomics.* 2011;10:280–93.
  33. Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011;12:56–68.
  34. Lan A, Smoly IY, Rapaport G, Lindquist S, Fraenkel E, Yeger-Lotem E. ResponseNet: revealing signaling and regulatory networks linking genetic and transcriptomic screening data. *Nucleic Acids Res.* 2011;39:W424–9.
  35. Hofree M, Shen JP, Carter H, Gross A, Ideker T. Network-based stratification of tumor mutations. *Nat Methods.* 2013;10:1108–15.
  36. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol.* 2005;4:Article 17.
  37. Liu J, Jing L, Tu X. Weighted gene co-expression network analysis identifies specific modules and hub genes related to coronary artery disease. *BMC Cardiovasc Disord.* 2016;16:54.
  38. Kwong LN, Costello JC, Liu H, Jiang S, Helms TL, Langsdorf AE, Jakubosky D, Genovese G, Muller FL, Jeong JH, Bender RP, Chu GC, Flaherty KT, Wargo JA, Collins JJ, Chin L. Oncogenic NRAS signaling differentially regulates survival and proliferation in melanoma. *Nat Med.* 2012;18:1503–10.
  39. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102:15545–50.
  40. Herwig R, Hardt C, Lienhard M, Kamburov A. Analyzing and interpreting genome data at the network level with ConsensusPathDB. *Nat Protoc.* 2016;11:1889–907.
  41. Kamburov A, Stelzl U, Lehrach H, Herwig R. The ConsensusPathDB interaction database: 2013 update. *Nucleic Acids Res.* 2013;41:D793–800.
  42. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28:27–30.
  43. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, Caudy M, Garapati P, Gopinath G, Jassal B, Jupe S, Kalatskaya I, Mahajan S, May B, Ndegwa N, Schmidt E, Shamovsky V, Yung C, Birney E, Hermjakob H, D'eustachio P, Stein L. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 2011;39:D691–7.
  44. Ashburner M, Ball CA, Blake JA, BOTSTEIN D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25:25–9.
  45. Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobank.* 2015;13:307–8.
  46. Thul PJ, Lindskog C. The human protein atlas: A spatial map of the human proteome. *Protein Sci.* 2018;27(1):233–44.
  47. Haase T, Bornigen D, Muller C, Zeller T. Systems medicine as an emerging tool for cardiovascular genetics. *Front Cardiovasc Med.* 2016;3:27.
  48. Song JW, Chung KC. Observational studies: cohort and case-control studies. *Plast Reconstr Surg.* 2010;126:2234–42.

49. Chen G, Levy D. Contributions of the Framingham Heart Study to the Epidemiology of Coronary Heart Disease. *JAMA Cardiol.* 2016;1:825–30.
50. Long MT, Fox CS. The Framingham Heart Study--67 years of discovery in metabolic disease. *Nat Rev Endocrinol.* 2016;12:177–83.
51. Investigators TA. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol.* 1989;129:687–702.
52. Koch B, Schaper C, Ittermann T, Volzke H, Felix SB, Ewert R, Glaser S. Reference values for lung function testing in adults--results from the study of health in Pomerania (SHIP). *Dtsch Med Wochenschr.* 2009;134:2327–32.
53. Lowel H, Lewis M, Hormann A, Keil U. Case finding, data quality aspects and comparability of myocardial infarction registers: results of a south German register study. *J Clin Epidemiol.* 1991;44:249–60.
54. Wild PS, Zeller T, Beutel M, Blettner M, Dugi KA, Lackner KJ, Pfeiffer N, Munzel T, Blankenberg S. The Gutenberg Health Study. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2012;55:824–9.
55. Zeller T, Hughes M, Tuovinen T, Schillert A, Conrads-Frank A, Ruijter H, Schnabel RB, Kee F, Salomaa V, Siebert U, Thorand B, Ziegler A, Breek H, Pasterkamp G, Kuulasmaa K, Koenig W, Blankenberg S. BiomarCaRE: rationale and design of the European BiomarCaRE project including 300,000 participants from 13 European countries. *Eur J Epidemiol.* 2014;29:777–90.
56. Mayrhofer MT, Holub P, Wutte A, Litton JE. BBMRI-ERIC: the novel gateway to biobanks. From humans to humans. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2016;59:379–84.
57. Simeon-Dubach D, Zeisberger SM, Hoerstrup SP. Quality assurance in biobanking for pre-clinical research. *Transfus Med Hemother.* 2016;43:353–7.
58. Rotival M, Zeller T, Wild PS, Maouche S, Szymczak S, Schillert A, Castagne R, Deiseroth A, Proust C, Brocheton J, Godefroy T, Perret C, Germain M, Eleftheriadis M, Sinning CR, Schnabel RB, Lubos E, Lackner KJ, Rossmann H, Munzel T, Rendon A, Erdmann J, Deloukas P, Hengstenberg C, Diemert P, Montalescot G, Ouwehand WH, Samani NJ, Schunkert H, Tregouet DA, Ziegler A, Goodall AH, Cambien F, Tiret L, Blankenberg S. Integrating genome-wide genetic variations and monocyte expression data reveals trans-regulated gene modules in humans. *PLoS Genet.* 2011;7:e1002367.
59. Kessler T, Wobst J, Wolf B, Eckhold J, Vilne B, Hollstein R, Von Ameln S, Dang TA, Sager HB, Moritz Rumpf P, Aherrahrou R, Kastrati A, Bjorkegren JLM, Erdmann J, Lusic AJ, Civelek M, Kaiser FJ, Schunkert H. Functional characterization of the GUCY1A3 coronary artery disease risk locus. *Circulation.* 2017;136:476–89.
60. Huan T, Esko T, Peters MJ, Pilling LC, Schramm K, Schurmann C, Chen BH, Liu C, Joehanes R, Johnson AD, Yao C, Ying SX, Courchesne P, Milani L, Raghavachari N, Wang R, Liu P, Reinmaa E, Dehghan A, Hofman A, Uitterlinden AG, Hernandez DG, Bandinelli S, Singleton A, Melzer D, Metspalu A, Carstensen M, Grallert H, Herder C, Meitinger T, Peters A, Roden M, Waldenberger M, Dorr M, Felix SB, Zeller T, Vasani R, O'donnell CJ, Munson PJ, Yang X, Prokisch H, Volker U, Van Meurs JB, Ferrucci L, Levy D. A meta-analysis of gene expression signatures of blood pressure and hypertension. *PLoS Genet.* 2015a;11:e1005035.
61. Huan T, Meng Q, Saleh MA, Norlander AE, Joehanes R, Zhu J, Chen BH, Zhang B, Johnson AD, Ying S, Courchesne P, Raghavachari N, Wang R, Liu P, O'donnell CJ, Vasani R, Munson PJ, Madhur MS, Harrison DG, Yang X, Levy D. Integrative network analysis reveals molecular mechanisms of blood pressure regulation. *Mol Syst Biol.* 2015b;11:799.
62. Saleh MA, McMaster WG, Wu J, Norlander AE, Funt SA, Thabet SR, Kirabo A, Xiao L, Chen W, Itani HA, Michell D, Huan T, Zhang Y, Takaki S, Titze J, Levy D, Harrison DG, Madhur MS. Lymphocyte adaptor protein Lnk deficiency exacerbates hypertension and end-organ inflammation. *J Clin Invest.* 2015;125:1189–202.

63. Zeller T, Schurmann C, Schramm K, Muller C, Kwon S, Wild PS, Teumer A, Herrington D, Schillert A, Iacoviello L, Kratzer A, Jagodzinski A, Karakas M, Ding J, Neumann JT, Kuulasmaa K, Gieger C, Kacprowski T, Schnabel RB, Roden M, Wahl S, Rotter JI, Ojeda F, Carstensen-Kirberg M, Tregouet DA, Dorr M, Meitinger T, Lackner KJ, Wolf P, Felix SB, Landmesser U, Costanzo S, Ziegler A, Liu Y, Volker U, Palmas W, Prokisch H, Guo X, Herder C, Blankenberg S, Homuth G. Transcriptome-wide analysis identifies novel associations with blood pressure. *Hypertension*. 2017;70:743–50.
64. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. *Cell*. 2012;148:1258–70.
65. Jie Z, Xia H, Zhong SL, Feng Q, Li S, Liang S, Zhong H, Liu Z, Gao Y, Zhao H, Zhang D, Su Z, Fang Z, Lan Z, Li J, Xiao L, Li R, Li X, Li F, Ren H, Huang Y, Peng Y, Li G, Wen B, Dong B, Chen JY, Geng QS, Zhang ZW, Yang H, Wang J, Zhang X, Madsen L, Brix S, Ning G, Xu X, Liu X, Hou Y, Jia H, He K, Kristiansen K. The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun*. 2017;8:845.
66. Gilbert JA, Quinn RA, Debelius J, Xu ZZ, Morton J, Garg N, Jansson JK, Dorrestein PC, Knight R. Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature*. 2016;535:94–103.
67. Ahn AC, Tewari M, Poon CS, Phillips RS. The limits of reductionism in medicine: could systems biology offer an alternative? *PLoS Med*. 2006;3:e208.
68. Gijzen H. Development: Big data for a sustainable future. *Nature*. 2013;502:38.
69. Eisenstein M. Big data: The power of petabytes. *Nature*. 2015;527:S2–4.
70. Lowes M, Kleiss M, Lueck R, Detken S, Koenig A, Nietert M, Beissbarth T, Stanek K, Langer C, Ghadimi M, Conradi LC, Homayounfar K. The utilization of multidisciplinary tumor boards (MDT) in clinical routine: results of a health care research study focusing on patients with metastasized colorectal cancer. *Int J Color Dis*. 2017;32(10):1463–9.
71. Ginsburg GS, Willard HF. Genomic and personalized medicine: foundations and applications. *Transl Res*. 2009;154:277–87.
72. Mirnezami R, Nicholson J, Darzi A. Preparing for precision medicine. *N Engl J Med*. 2012;366:489–91.
73. Harris JR, Burton P, Knoppers BM, Lindpaintner K, Bledsoe M, Brookes AJ, Budin-Ljosne I, Chisholm R, Cox D, Deschenes M, Fortier I, Hainaut P, Hewitt R, Kaye J, Litton JE, Metspalu A, Ollier B, Palmer LJ, Palotie A, Pasterk M, Perola M, Riegman PH, Van Ommen GJ, Yuille M, Zatloukal K. Toward a roadmap in global biobanking for health. *Eur J Hum Genet*. 2012;20:1105–11.
74. Walport M, Brest P. Sharing research data to improve public health. *Lancet*. 2011;377:537–9.
75. Muilu J, Peltonen L, Litton JE. The federated database--a basis for biobank-based post-genome studies, integrating phenome and genome data from 600,000 twin pairs in Europe. *Eur J Hum Genet*. 2007;15:718–23.
76. Doiron D, Burton P, Marcon Y, Gaye A, Wolffenbuttel BHR, Perola M, Stolk RP, Foco L, Minelli C, Waldenberger M, Holle R, Kvaloy K, Hillege HL, Tasse AM, Ferretti V, Fortier I. Data harmonization and federated analysis of population-based studies: the BioSHaRE project. *Emerg Themes Epidemiol*. 2013;10:12.
77. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansong W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet*. 2001;29:365–71.
78. Lappalainen I, Almeida-King J, Kumanduri V, Senf A, Spalding JD, Ur-Rehman S, Saunders G, Kandasamy J, Caccamo M, Leinonen R, Vaughan B, Laurent T, Rowland F, Marin-Garcia P, Barker J, Jokinen P, Torres AC, De Argila JR, Llobet OM, Medina I, Puy MS, Alberich M, De La Torre S, Navarro A, Paschall J, Flicek P. The European Genome-phenome Archive of human data consented for biomedical research. *Nat Genet*. 2015;47:692–5.

79. Caulfield T, Murdoch B. Genes, cells, and biobanks: Yes, there's still a consent problem. *PLoS Biol.* 2017;15:e2002654.
80. Grabe HJ, Assel H, Bahls T, Dorr M, Endlich K, Endlich N, Erdmann P, Ewert R, Felix SB, Fiene B, Fischer T, Flessa S, Friedrich N, Gadebusch-Bondio M, Salazar MG, Hammer E, Haring R, Havemann C, Hecker M, Hoffmann W, Holtfreter B, Kacprowski T, Klein K, Kocher T, Kock H, Krafczyk J, Kuhn J, Langanke M, Lendeckel U, Lerch MM, Lieb W, Lorbeer R, Mayerle J, Meissner K, Zu Schwabedissen HM, Nauck M, Ott K, Rathmann W, Rettig R, Richardt C, Salje K, Schminke U, Schulz A, Schwab M, Siegmund W, Stracke S, Suhre K, Ueffing M, Ungerer S, Volker U, Volzke H, Wallaschofski H, Werner V, Zygumt MT, Kroemer HK. Cohort profile: Greifswald approach to individualized medicine (GANI\_MED). *J Transl Med.* 2014;12:144.
81. Auffray C, Balling R, Barroso I, Bencze L, Benson M, Bergeron J, Bernal-Delgado E, Blomberg N, Bock C, Conesa A, Del Signore S, Delogne C, Devilee P, Di Meglio A, Eijkemans M, Flicek P, Graf N, Grimm V, Guchelaar HJ, Guo YK, Gut IG, Hanbury A, Hanif S, Hilgers RD, Honrado A, Hose DR, Houwing-Duistermaat J, Hubbard T, Janacek SH, Karanikas H, Kievits T, Kohler M, Kremer A, Lanfear J, Lengauer T, Maes E, Meert T, Muller W, Nickel D, Oledzki P, Pedersen B, Petkovic M, Pliakos K, Rattray M, I Mäs JR, Schneider R, Sengstag T, Serra-Picamal X, Spek W, Vaas LA, Van Batenburg O, Vandelaer M, Varnai P, Villoslada P, Vizcaino JA, Wubbe JP, Zanetti G. Making sense of big data in health research: Towards an EU action plan. *Genome Med.* 2016;8:71.
82. Hazin R, Brothers KB, Malin BA, Koenig BA, Sanderson SC, Rothstein MA, Williams MS, Clayton EW, Kullo IJ. Ethical, legal, and social implications of incorporating genomic information into electronic health records. *Genet Med.* 2013;15:810–6.
83. Lamas E, Barh A, Brown D, Jaulent MC. Ethical, Legal and Social Issues related to the health data-warehouses: re-using health data in the research and public health research. *Stud Health Technol Inform.* 2015;210:719–23.
84. Vasan RS, Benjamin EJ. The future of cardiovascular epidemiology. *Circulation.* 2016;133:2626–33.
85. Horak P, Klink B, Heining C, Groschel S, Hutter B, Frohlich M, Uhrig S, Hubschmann D, Schlesner M, Eils R, Richter D, Pfitze K, Georg C, Meissburger B, Wolf S, Schulz A, Penzel R, Herpel E, Kirchner M, Lier A, Endris V, Singer S, Schirmacher P, Weichert W, Stenzinger A, Schlenk RF, Schrock E, Brors B, Von Kalle C, Glimm H, Frohling S. Precision oncology based on omics data: The NCT Heidelberg experience. *Int J Cancer.* 2017;141:877–86.