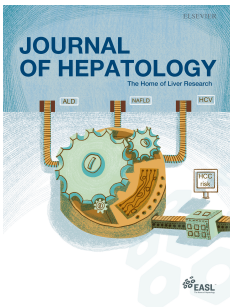
Journal Pre-proof



Review: Opportunities and barriers for omics-based biomarker discovery in steatotic liver diseases

Maja Thiele, Ida Falk Villesen, Lili Niu, Stine Johansen, Karolina Sulek, Suguru Nishijima, Lore Van Espen, Marisa Keller, Mads Israelsen, Tommi Suvitaival, Andressa de Zawadzki, Helene Bæk Juel, Maximilian Joseph Brol, Sara Elisabeth Stinson, Yun Huang, Maria Camilla Alvarez Silva, Michael Kuhn, Ema Anastasiadou, Diana Julie Leeming, Morten Karsdal, Jelle Matthijnssens, Manimozhiyan Arumugam, Louise Torp Dalgaard, Cristina Legido-Quigley, Mathias Mann, Jonel Trebicka, Peer Bork, Lars Juhl Jensen, Torben Hansen, Aleksander Krag, on behalf of the MicrobLiver and GALAXY consortia, GALAXY consortium, MicrobLiver consortium

PII: S0168-8278(24)00219-8

DOI: https://doi.org/10.1016/j.jhep.2024.03.035

Reference: JHEPAT 9566

To appear in: *Journal of Hepatology*

Received Date: 20 September 2023

Revised Date: 16 February 2024

Accepted Date: 19 March 2024

Please cite this article as: Thiele M, Villesen IF, Niu L, Johansen S, Sulek K, Nishijima S, Van Espen L, Keller M, Israelsen M, Suvitaival T, de Zawadzki A, Juel HB, Brol MJ, Stinson SE, Huang Y, Alvarez Silva MC, Kuhn M, Anastasiadou E, Leeming DJ, Karsdal M, Matthijnssens J, Arumugam M, Dalgaard LT, Legido-Quigley C, Mann M, Trebicka J, Bork P, Jensen LJ, Hansen T, Krag A, on behalf of the MicrobLiver and GALAXY consortia, GALAXY consortium, MicrobLiver consortium, Review: Opportunities and barriers for omics-based biomarker discovery in steatotic liver diseases, *Journal of Hepatology* (2024), doi: https://doi.org/10.1016/j.jhep.2024.03.035.

during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

REVIEW: OPPORTUNITIES AND BARRIERS FOR OMICS-BASED BIOMARKER DISCOVERY IN STEATOTIC LIVER DISEASES

Authors: #Maja Thiele[1,2], #Ida Falk Villesen[1,2], Lili Niu[3,4], Stine Johansen[1], Karolina Sulek[5], Suguru Nishijima[6], Lore Van Espen[7], Marisa Keller[6], Mads Israelsen[1,2], Tommi Suvitaival[5], Andressa de Zawadzki[5], Helene Bæk Juel[8], Maximilian Joseph Brol[9], Sara Elisabeth Stinson[8], Yun Huang[8], Maria Camilla Alvarez Silva[8], Michael Kuhn[6], Ema Anastasiadou[10], Diana Julie Leeming[11], Morten Karsdal[11], Jelle Matthijnssens[7], Manimozhiyan Arumugam[8], Louise Torp Dalgaard[12], Cristina Legido-Quigley[5], Mathias Mann[3,4], Jonel Trebicka[9], Peer Bork[6,13,14], Lars Juhl Jensen[3], *Torben Hansen[8] and *Aleksander Krag[1,2] on behalf of the MicrobLiver and GALAXY consortia

Affiliations: 1: Center for Liver Research, Department of Gastroenterology and Hepatology, Odense University Hospital, Odense, Denmark 2: Department for Clinical Research, University of Southern Denmark, Odense, Denmark 3: Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark 4: Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Germany 5: Steno Diabetes Center Copenhagen, Herlev, Denmark 6: Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany 7: KU Leuven, Department of Microbiology, Immunology, and Transplantation, Rega Institute, Laboratory of Viral Metagenomics, Leuven, Belgium 8: Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Science, University of Copenhagen, Copenhagen, Denmark 9: Medizinische Klinik B (Gastroenterologie, Hepatologie, Endokrinologie, Klinische Infektiologie), Universitätsklinikum Münster Westfälische, Wilhelms-Universität Münster, Germany 10: Idryma Iatroviologikon Ereunon Akademias Athinon, Greece 11: Fibrosis, Hepatic and Pulmonary Research, Nordic Bioscience, Herlev, Denmark 12: Department of Science and Environment Roskilde University, Roskilde, Denmark 13: Max Delbrück Centre for Molecular Medicine, Berlin, Germany 14: Department of Bioinformatics, Biocenter, University of Würzburg, Würzburg, Germany.

[#] shared first authorship, * Shared senior authorship

Corresponding author: Aleksander Krag, Center for Liver Research, Department of Gastroenterology and Hepatology, Odense University Hospital, Odense, Denmark. Phone: +45 29647719. E-mail address: Aleksander.Krag@rsyd.dk.

GALAXY consortium: Aleksander Krag (coordinator), Peer Bork, Torben Hansen, Manimozhiyan Arumugam, Jonel Trebicka, Morten Karsdal, Ema Anastasiadou, Hans Israelsen, Hans Olav Melberg, Cristina Legido-Quigley, Maja Thiele.

MicrobLiver consortium: Torben Hansen (coordinator), Matthias Mann, Jelle Matthijnssens, Aleksander Krag.

Keywords: Non-invasive test, genetics, microbiome, metagenomics, metatranscriptomics, viromics, metabolomics, lipidomics, proteomics.

Word count: 4663

Abstract word count: 273

Number of figures/tables: 5/2

Conflicts of interest: MT: Speaker's fee from Echosens, Siemens Healthcare, Norgine, Tillotts Pharma, and advisory fee from GE Healthcare. DJL and MK: Fulltime employee and stockholder of Nordic Bioscience. AK: Speaker for Novo Nordisk, Norgine, Siemens and Nordic Bioscience and participated in advisory boards for Norgine, Siemens, Resalis Therapeutics, Boehringer Ingelheim and Novo Nordisk, all outside the submitted work. Research support; Norgine, Siemens, Nordic Bioscience, Astra, Echosens. Consulting Takeda, Resalis Therapeutics, Zealand Pharma, Novo Nordisk, Boehringer Ingelheim. Board member and co-founder Evido. JT: Speaker

and/or consulting fees from Versantis, Gore, Boehringer-Ingelheim, Falk, Grifols, Genfit and CSL Behring. LJJ: Speaker's fee from Boehringer Ingelheim Fonds.

Financial support: This project has received funding from the GALAXY project from the European Union's Horizon 2020 research and innovation programme under grant agreement number 668031; and from Challenge Grant "MicrobLiver" grant number NNF15OC0016692 from the Novo Nordisk Foundation. LVE is funded by the Fonds Wetenschappelijk Onderzoek (FWO) Vlaanderen (grant number: 1S25720N). MT is funded by a grant from the Novo Nordisk Foundation (NNF20OC0059393). KS is supported by the Novo Nordisk Foundation Excellence Emerging Investigator Grant – Endocrinology and Metabolism 2022 (grant no. NNF 0074491). LTD is funded by grants from the Novo Nordisk Foundation (NNF22OC0078203 and NNF23OC0081177). JT was supported by the German Research Foundation (DFG) project ID 403224013 – SFB 1382 (A09), by the German Federal Ministry of Education and Research (BMBF) for the DEEP-HCC project and by the Hessian Ministry of Higher Education, Research and the Arts (HMWK) for the ENABLE and ACLF-I cluster projects. The MICROB-PREDICT (project ID 825694), DECISION (project ID 847949), GALAXY (project ID 668031), LIVERHOPE (project ID 731875) and IHMCSA (project ID 964590) projects have received funding from the European Union's Horizon 2020 research and innovation programme. LJJ is funded by the Novo Nordisk Foundation (NNF14CC0001). The manuscript reflects only the authors' views, and the European Commission is not responsible for any use that may be made of the information it contains. The funders had no influence on literature used, preparation of the manuscript or where to publish.

Author contribution: Conceptualisation: MT, IFV, SJ, LN, AK, TH. Writing – original draft: MT, IFV, LN, LJJ, SN, SJ, DJL, HBJ, AZ, KS, TS, MJB, MKe, MKu, CLQ, SES,

YH, LVE, MCAS, LTD. Visualisation: MT, IFV, SJ, LN, KS, CLQ, LJJ. Funding acquisition: TH, AK, PB, MT, MK, JM, MA, CLQ, MM, EA, JT. Writing – review and editing: all authors. All authors approved the final version of the manuscript.

Key points

- There is an urgent need for accurate biomarkers in patients with steatotic liver disease, to stage and grade fibrosis and inflammation, for monitoring disease progression and for improving drug development and approval pipelines.
- The rapid development and decreased costs of high-throughput omics technologies in combination with excellent computational power has created a golden opportunity for new types of biomarkers which reflect biological disease processes and can be combined in multiplex systems of molecules. Multi-omics may thereby facilitate an era of accurate, personalised diagnostics.
- Heterogeneity in the development and progression of steatotic liver disease can be disentangled by the interplay between host genetics, transcriptomics, proteomics, metabolomics and lipidomics on the one hand, and gut microbial, viral and fungal metagenomics and meta-transcriptomics on the other hand.
- Hypothesis-free approaches have revealed the potential of omics technologies for the discovery of liver disease biomarkers and have proposed many more candidate biomarkers than the traditional hypothesis-driven studies. However, few of these omics-based biomarker candidates are rigorously tested in independent cohorts, and none have yet been implemented in clinical practice.

Summary

The rising prevalence of liver diseases related to obesity and excessive use of alcohol is fuelling an increasing demand for accurate biomarkers aimed at community screening, diagnosis of steatohepatitis and significant fibrosis, monitoring, prognosis and prediction of treatment efficacy. Breakthroughs in omics methodologies and the power of bioinformatics have created an excellent opportunity to combine clinical needs with technological advancements. Omics technologies allow for advanced investigations into biological processes from the genes to transcription and regulation, to circulating protein, metabolite and lipid levels, as well as the microbiome including bacteria, viruses and fungi. We consequently find ourselves in a period of rapid progress in technology and bioinformatics that may allow for development of precision biomarkers for personalised medicine. However, there are important barriers to consider in omics biomarker discovery and validation, including the use of semi-quantitative measurements from untargeted platforms, which may exhibit high analytical, inter- and intra-individual variance. Standardising methods and the need to validate across diverse populations, presents a challenge, partly due to disease complexity and the dynamic nature of biomarker expression in different disease stages. Lack of validity causes lost opportunities when studies fail to provide the knowledge needed for regulatory approvals, all of which contributes to a delayed translation of these discoveries into clinical practice. While no omics-based biomarkers have matured to clinical implementation, the extent of data generated through omics-technologies holds the power of hypothesis-free discovery of a plethora of candidate biomarkers to be further validated. To explore the many opportunities of omics technologies, hepatologists need detailed knowledge of commonalities and

differences between the various omics layers, and both the barriers to and advantages of these approaches.

Introduction

More than one third of the adult population have steatotic liver disease either metabolic dysfunction associated steatotic liver disease (MASLD), alcohol-related liver disease (ALD) or a combination of these (MetALD).(1-3) Patients with progressive disease experience high liver-related morbidity, extrahepatic complications and premature all-cause mortality.(4, 5) There is consequently an urgent need for accurate risk stratification and effective treatments that modify the natural course of disease.(6, 7) Progression of steatotic liver disease follows a profibrotic path, resulting in pivotal liver-related events that critically affect prognosis. It is consequently important to explore biomarkers that predict precursors of cirrhosis and portal hypertension in the form of significant and advanced fibrosis, these disease stages predict later liver-related events. Relevant biomarker endpoints for how patients function, feel, and survives are decompensation, acute-on-chronic liver failure, hepatocellular carcinoma, and death.(8-10)

The performance of existing and future biomarkers depends on their intended context of use and validation (**Figure 1, Table 1**).(11) General practitioners and hepatologists managing steatotic liver disease from ALD, MetALD and MASLD particularly lack tests for accurate diagnosis of significant fibrosis (≥F2) and steatohepatitis, for prognosis, monitoring and prediction, and for evaluating the efficacy of interventions.(8, 12) Yet, traditionally, the diagnostic accuracy of a biomarker is evaluated by area under the receiver operating characteristic (AUROC), sensitivity specificity and predictive values. However, these performance characteristics depend on disease prevalence in the studied population.(13) Consequently, future biomarkers need to be tailored to the intended population and tested in cohorts which reflect the appropriate disease prevalence.

This review will explore the advantages and limitations of exploring omics technologies for biomarker discovery across the spectrum of steatotic liver disease. We highlight the state of the art of individual omics technologies: genetics, transcriptomics, proteomics, lipidomics, metabolomics, metagenomics, metatranscriptomics, viromics and mycobiomics. These technologies have been selected from a wider list of currently available omics technologies as they represent to most common examples of the promises and obstacles of omics based biomarkers for clinical hepatology.

Opportunities for omics technologies

A new era of biomarker development has been revealed in recent years thanks to high throughput omics technologies combined with increasing computational power and the ability of running artificial intelligence and machine learning methods with routine hardware and software. This major advancement allows for hypothesis-free testing of thousands or even millions of analytes.(14, 15) Multi-omics is thereby able to disentangle complex molecular interplays between host genes, gene transcription, proteins, metabolites and lipids, in addition to interactions between the host and microbiome consisting of bacteria, viruses and fungi (**Figure 2).** Recent development and promising biomarker targets from omics technology are highlighted in **Table 2**. Omics measurements consequently result in a multitude of candidate biomarkers.(16-19)

To enable the accurate separation of patients with progressive liver disease from those with non-progressive disease, researchers aim at understanding disease heterogeneity and pathophysiology through host-gut-environment interactions.(20)

In the struggle to identify effective anti-fibrotic interventions for MASLD and ALD, omics-based biomarkers that reflect biological fibrotic processes may be used to identify future drug targets, thereby abating the frequent failures of phase III clinical trials.(21) There is a similar search for accurate biomarkers to reduce clinical trial screening failures.(17) Finally, non-invasive biomarkers to replace liver biopsy as the surrogate endpoint would effectively allow for shorter, less costly trials and reduced patient discomfort.(22)

The analysis costs of genetics, transcriptomics, proteomics, lipidomics, metabolomics, metagenomics and metatranscriptomics are decreasing thanks to technological development and an increase in the capacity of high-throughput omics platforms.(23, 24) We therefore expect multi-omics approaches to become increasingly accessible for clinical management of liver disease patients over the next decade.

Barriers to omics technologies

Omics-based biomarkers offer more opportunities for discovery than traditional biomarkers, which quantify a low number of analytes, often only one. However, no omics-based biomarker has penetrated from development to implementation. This shortcoming can be attributed to several barriers across different omics technologies, including 1) technological maturity, 2) cost, 3) analytical validity, 4) untargeted coverage and 5) semi-quantitative measurements, which are usually laboratory or instrument specific.

Except for genetics, omics technologies are in their infancy (**Figure 3**). This immaturity results in several obvious limitations, most notably that the evidence base remains incomplete.

Technological development is moving rapidly from high cost and low throughput to low cost and high throughput.(15, 25) However, finite budgets remain a challenge for the maturation of omics-biomarkers. Current cost pressures create a trade-off between analyte depth and abundance versus sample throughput and sample size.(18) The limited ability to robustly detect low-abundance analytes generates 'technological bias'.(26) Omics studies typically aim for great depth to discover low-abundance biomarkers, but this means that investigators cannot afford as many samples, thus risking spurious findings. The high-dimensional nature of omics data also requires extensive computational protocols and processing power, further increasing time usage and costs.(27) However, increasingly higher demands for omics technologies within the healthcare system will lead to the development of routine protocols and market competition, driving costs downward.

Omics measurements can be divided into two analytical methods: non-targeted and targeted. Non-targeted omics takes a hypothesis-free approach to the semi-quantitative analysis of a very large number of molecules, often aided by machine learning and other advanced bioinformatics. Non-targeted omics is consequently highly suited for discovery of new biomarkers. However, this approach faces three major challenges: 1) semi-quantitative measurements are relative and, as such, study specific. Findings are therefore difficult to replicate in external validation. Candidate biomarkers detected by untargeted approaches must therefore be validated using a targeted platform, such as enzyme-linked immunoassay (ELISA) for absolute concentrations.(28) 2) Non-targeted measurements are more prone to analytical biases such as batch effect and variations related to sample handling and processing.(29) 3) Non-targeted approaches usually require more complex and therefore less standardised- bio-informatics analyses pipelines.

The targeted approach uses quantitative assays to measure concentrations of predefined panels of up to a few hundred molecules.(30, 31) Targeted omics can be done, for example, by using calibration curves and spike-in of internal standards to allow for absolute quantification and is well suited to either searching for high-abundance biomarkers or for hypothesis-driven biomarker evaluation. Discovery of novel targets and pathways is especially useful in drug discovery and searching for disease aetiology; however, its application in routine analysis in the clinic is still being evaluated.

Different omics technologies each have their own set of specific advantages which hold great potential for personalised and precision medicine (**Figure 4; Table 2**). Nevertheless, in order to bring omics-based biomarkers into the clinic, the current process involves transforming them into analytically reproducible assays that can be validated across laboratories and cohorts while also meeting regulatory requirements.(32, 33) These requirements can be insurance against hurried, spurious findings but can also limit the speed of discovery and development to validation.

The subsequent sections delineate the technical complexities and biomarker prospects across diverse omics disciplines.

Genetics

Genetics is the most widely investigated omics technology, linking single nucleotide polymorphisms to cirrhosis, hepatocellular carcinoma and steatosis, particularly for MASLD and ALD.(24, 34, 35) From family and population-based studies, the heritability of MASLD ranges from 20–70% depending on ethnicity and how MASLD is diagnosed.(36) For the heritability of ALD, studies suggest alcohol use disorder heritability ranges from 30–50% and ALD-related cirrhosis ranges from 21–67%.(37)

However, disagreement within the field exists on the proportion of the genetic variance for ALD that is independent of the genetic predisposition to alcohol dependence.(37, 38)

Genotyping of individuals for genome-wide association studies (GWAS) is typically performed using microarrays to measure common variants, due to the higher cost of next-generation sequencing (NGS). NGS methods encompass: 1) whole exome sequencing, which targets coding regions with functional significance and 2) whole genome sequencing, which captures nearly every genotype across the genome, both coding and non-coding, including rare variants. Whole genome sequencing is expected to become the method of choice in the future for untargeted discovery as costs continue to decrease.(39) NGS methods can be effective tools for precision diagnostics in rare monogenic forms of liver disease. Patients who remain undiagnosed despite comprehensive clinical workups may benefit from genomic analysis to improve disease prognostication. Examples include *ABCB4, ABCB11 and ATP8B1* to distinguish idiopathic cholestasis.(40)

Large-scale GWAS and meta-analyses have elucidated the genetic architecture of steatosis, steatohepatitis, and fibrosis from ALD and MASLD, using liver biopsies, imaging, elastography, liver enzymes and electronic health records. These efforts have identified risk loci common to ALD and MASLD, including *PNPLA3*, *TM6SF2*, *GCKR*, *SERPINA1* and *MBOAT7*.(41-45) Novel protective loci include *HSD13B17*, *MTARC1*, *GPAM* and *PSD3*.(35, 45, 46)

Genetic risk scores (GRS) combining multiple genome-wide significant SNPs ($P<5\times10^{-8}$) can be used for risk prediction and stratification. A higher GRS, including *PNPLA3*, *TM6SF2* and *HSD17B13,* confers a 12-fold increased risk of cirrhosis and a

29-fold increased risk of hepatocellular carcinoma in the European population.(47) Likewise, a higher GRS derived from *PNPLA3*, *TM6SF2*, *MBOAT7, GCKR* and *HSD13B17* amplifies the effect of liver steatosis on the risk of subsequent hepatic events.(48) Despite considerable interest, the predictive value of a given GRS over simple biochemical biomarkers has been marginal. Combining *PNPLA3*, *TM6SF2*, *HSD17B13* and *MBOAT7* with metabolic traits slightly increases the area under the curve for diagnosing advanced liver fibrosis, from 0.75 to 0.80 in ALD patients.(49) Prediction of a 10-year cirrhosis risk by adding GRS to the APRI score (age platelet ratio index) increased the prognostic information by less than 5% and improved the C-index from 0.804 to 0.809 in the UK Biobank.(50) This limited impact is likely due to the fact that clinical features from five to ten years before disease explain more variance compared to the few SNPs with small effect sizes identified so far.(51) Yet there is promise: a study based on UK Biobank data demonstrated that a GRS improves risk stratification and diagnostic accuracy, particularly in subgroups of individuals with diabetes, obesity or a fatty liver index above 60. This suggests that integrating a genetic risk GRS with clinical non-invasive markers holds the potential to refine individual risk prediction for severe liver disease, especially in individuals at risk for MASLD.(52)

Polygenic scores have achieved greater predictive power than GRS for complex diseases by including hundreds to thousands of SNPs, rather than being restricted to only those that reach genome-wide significance ($P<5x10^{-8}$).(53) Polygenic scores developed for liver diseases are still under development and require well-powered GWAS studies, validated in independent study populations of varying ancestries to ensure generalisability.

Transcriptomics

The transcriptome is the sum of all RNA transcripts of a tissue or blood sample, commonly used to examine gene expression. Circulating RNA species include several classes of shorter RNAs, with microRNAs (miRNA) being by far the most studied. Quantification of miRNAs can be done by sequencing or reverse transcription quantitative polymerase chain reaction (qPCR), often in targeted or multiplexed panels. These methods are sensitive, often quantitative, and relatively low in cost. In contrast, sequencing all small RNAs is considerably more expensive but allows for measurement of other types, such as PIWI-interacting RNAs, transfer-RNA fragments, ribosomal and nucleolar RNAs, each of which contains tens to thousands of different species.(54, 55) Small RNAs in circulation constitute a novel source of MASLD-related biomarker candidates. For example the hepatocyte enriched miR-122, and other miRNAs (miR-34a, miR-193a).(56-58) Once a promising RNA biomarker has been identified, the RNA can be detected with high sensitivity and accuracy based on targeted RT-qPCR or microfluidics-based nano-sensors.

The extracellular RNAs are an especially interesting subtype of circulating miRNAs.(59) They are enclosed in vesicles or are protein bound, which protects them from degradation and facilitate their transport, in turn allowing for cell-to-cell paracrine communication or long-distance signalling.(60) Liver-derived miRNAs, as extracellular RNA, appear to be important regulators of metabolic disease, particularly MASLD and steatohepatitis.(56) Recent studies show that levels of liver-derived miRNAs are modified by weight-loss or insulin-sensitising treatments.(61, 62)

MiRNAs also show promise as biomarkers for ALD, MASLD and steatohepatitis, prominently miR-34a, which is part of the NIS2+ score.(63, 64) In addition, both miR-193 and miR-122 plasma levels are found to be increased in MASLD patients with steatohepatitis and advanced fibrosis.(65, 66) The liver-specific miR-122 also predicts

type 2 diabetes and decreases following weight loss.(61, 62) Yet low miR-122 is a marker of poor prognosis in patients with cirrhosis.(67) Therefore, it appears that the liver's miR-122 expression is temporary, from upregulation as steatohepatitis progresses, to a decline in cirrhosis patients. A similar non-linear pattern is seen for body weight, and naturally limits the potential use of miR-122 as a diagnostic biomarker, but points toward a possible role in causal pathways. It also illustrates the importance of consecutive recruitment and inclusion across the disease spectrum in biomarker research.

Microbiome

The human body is home to a large number of microbes, on all skin and mucous surfaces.(68) The vast majority reside in the gut, home to ten trillion bacteria.(69) The gut microbiota exerts important effects on host physiology by producing diverse metabolites, modulating the immune system and preventing infection by pathogens.(70) The gut microbiota can profoundly affect the liver, as microbial products can enter the blood circulation and thereby encounter the liver as the very first organ.(23, 71, 72)

Shotgun metagenomic sequencing evaluates both the species-level taxonomic profile and the functional profile of the microbiome but requires resource-heavy sequencing equipment and advanced bioinformatics. The cheaper amplicon sequencing of the bacterial 16S ribosomal RNA genes enables determination of a taxonomic profile without large computational resources, but with lower resolution, at the genus or family level. Metatranscriptomics quantifies microbial RNA to describe how gene transcriptional activity across bacterial species can change according to health or disease.(73)

Several studies have shown alterations in the gut microbiome of patients with cirrhosis or steatohepatitis from ALD or MASLD, compared to healthy individuals.(74-77) The more severe stages of liver disease are associated with dysbiosis, decreased abundance of potentially beneficial families such as *Ruminococcaceae* and *Lachnospiraceae*, and increase in potentially pathogenic families such as *Enterobacteriaceae* and *Bacteroidaceae*.(23, 78) One metagenomic study in decompensated cirrhosis patients found elevated levels of *Veillonella* and *Streptococcus* species, but reduced levels of butyrate-producing commensal bacteria, including *Faecalibacterium prausnitzii* and *Coprococcus comes*.(77) Other studies have demonstrated increased epithelial permeability in liver disease patients, which allows for translocation of bacterial components and metabolites, such as lipopolysaccharides, secondary bile acids and pathogen-associated molecular patterns, fuelling liver inflammation and fibrosis.(79-82) Consequently, microbial derived products can be important biomarkers of treatment effects, as in the RIFSYS trial, where circulating levels of the microbiome-generated metabolite trimethylamine-N-oxide remained stable in cirrhosis patients treated with Rifaximin-α, but increased in placebo treated patients.(83)

While accumulating evidence indicates that microbiota disturbances play a role in the development and progression of liver diseases, the biomarker potential of the gut microbiota is still in its infancy.

Viromics and mycobiomics

The virome and mycobiome, though considered premature omics fields, exhibit promise in light of advancing technologies, making them interesting for future exploration.

The gut virome mainly consists of bacteriophages (viruses infecting bacteria) and viruses infecting eukaryotic cells. Viruses are the most diverse genetic elements on earth, which poses several technical challenges for virome research.(86)

Due to the small genome size of viruses compared to prokaryotes and eukaryotes, the enrichment of faecal samples for viruses before DNA and RNA extraction is recommended. A reverse transcription step is necessary to also capture RNA viruses. As bacteriophages are highly diverse and highly individual specific, they are not sufficiently represented in databases. Hence, a *de novo* genome assembly approach and a viral identification method that is, at least partially, independent of databases is crucial to also identify novel viruses from sequencing data.(87)

Recent developments in bioinformatics tools have allowed for improved identification (geNomad), taxonomic classification (vConTACT2), host prediction (iPHoP) and functional annotation (Cenote-Taker2) of viral sequences, advancing the field to help identify associations between the virome and human health and disease.(88-92) Viruses can directly affect the human host by killing target cells such as hepatocytes or by modulating the immune system. The human host can also be indirectly affected by the gut virome through the effect of the gut phages on the composition and function of the gut bacterial community.(93)

Changes in the gut virome have been linked to the presence and severity of liver diseases such as MASLD, ALD, alcohol-related hepatitis and cirrhosis.(94-97) The high inter-individual variability of the human gut virome, however, limits the identification of robust viral biomarkers.(98) Overall, viral diversity might be a better biomarker than a set of individual viruses, but viral diversity lacks disease specificity, similar to dysbiosis.(94, 96) Other approaches which overcomes the low prevalence

of individual viral genomes are to look for virome biomarkers of higher taxonomical orders (e.g. families) or grouping bacteriophages by their bacterial host, but these more diverse groups of viruses will be more difficult to detect using qPCR tests.(99) Finally, viral-encoded genes might be less individual specific, for example, toxins or auxiliary metabolic genes, and hence better suited as biomarkers. These genes could be horizontally transferred to their bacterial hosts, thus altering the functional capacities of the targeted bacteria and thereby indirectly affecting the human host.

The fungal fraction of the microbiome, the mycobiome, are important in maintaining intestinal homeostasis and immunity. But although there has been advancement in the field of mycobiome research, this omics technology is still in its infancy. Early studies have shown that *Candida* overgrowth can be linked to ALD and cirrhosis, and that elevated levels of anti-S. *cerevisiae* antibodies, cross-react with *Candida albicans* found to associate with increased mortality in ALD.(100-102)

Proteomics

Proteins are the most prominent source of biomarkers and drug targets in human diseases. Routine laboratory testing is dominated by proteins (42% of all analytes) and as of 2017, 75% of drugs approved by the US Food and Drug Administration (FDA) targets human proteins.(28) Aminotransferases, albumin, bilirubin and coagulation factors are examples of routinely measured proteins for assessing liver function.

Proteomics seeks to map all proteins in a biological sample, with existing platforms quantifying hundreds to tens of thousands of proteins, depending on the sample type. Several cell type-resolved human liver proteome maps have been published, establishing a robust reference for the abundance of over ten thousand proteins in human liver cells.(103) Mass spectrometry (MS)-based proteomics and affinity-based

proteomics are commonly used technologies for the large-scale study of proteins. MS-based proteomics is the most comprehensive approach and the gold standard for the quantitative profiling of proteins, post-translational modifications and protein-protein interactions.(104) MS-based proteomics is an ideal approach for unbiased protein profiling across all organisms and sample types (**Table 2**). The untargeted approach, also known as discovery proteomics, offers a global view of the proteome and is often used for uncovering novel biomarkers. However, the lack of standardisation as well as its semi-quantitative nature is a significant hurdle for discovery proteomics – values obtained in a specific study can typically only be compared horizontally to other samples acquired within the same study. In contrast, targeted MS-based proteomics focuses on specific proteins of interest, providing precise quantification, validation and clinical applications.

Recent technological advancements in MS-based proteomics, including the automation of sample preparation, improvements in liquid chromatography, as well as the development of novel MS acquisition methods and sophisticated informatics solutions, have made it feasible to generate thousands of proteome profiles in a single clinical study. (105) This further translates into reproducible and robust results. At the same time, researchers have started to apply machine learning-based classification algorithms to demonstrate the predictive or discriminative power of proposed biomarkers in liver disease.

Affinity-based proteomics platforms such as Olink and SomaScan have been widely applied in human plasma and serum studies.(45, 106, 107) These platforms offer measurements for dozens and up to thousands of proteins, with standardised workflows allowing for value comparison across studies. However, studies comparing the two platforms have highlighted inconsistencies in quantification for a significant

number of proteins.(108) Consequently, findings from these platforms often require validation by an orthogonal method, ideally mass spectrometry, which excels in its specificity of identification and quantification.(109) Other methods include ELISA and similar techniques, which measure the concentration of a single protein, making them better suited for biomarker validation and implementation.

The FDA-approved OVA1 test for ovarian cancer serves as an example of a biomarker identified by MS-based proteomics but which was ultimately developed using immunoassays. The test consists of a panel of five proteins, four of which were first published in 2004. Five years later the test received FDA clearance.(110)

More than 200 candidate protein biomarkers for MASLD and 22 for ALD have been reported, although none have matured into clinical practice.(15, 111-113) The two most recent proteomics biomarker studies were selected from 2,201 candidate proteins for MASLD fibrosis and 1,235 candidates for ALD fibrosis, resulting in 8- and 9-protein biomarker panels.(15, 113) Complement component C7 was part of both panels, while the other proteins differed. Consequently, much work remains to be done in terms of evaluation of disease specificity and external validation of these signatures.

Metabolomics and lipidomics

The metabolome comprises all small molecules in the human body, originating from both endogenous and environmental sources, and encompasses a biochemically diverse array of metabolites such as sugars, lipids, amino acids, fatty acids, alkaloids, and polyphenols.(114) One example of a lipid metabolite biomarker is phosphatidylethanol, used to detect alcohol consumption, derived from the trans-phosphatidylation of phosphatidylcholine in the presence of ethanol.(115)

Humans are thought to contain around 3,000 endogenous or common metabolites while the plant kingdom harbours around 200,000 metabolites, of which 90% are still unquantified or unidentified.(114) Metabolomics platforms are usually a combination of different chemical analyses using mass spectrometry. The platforms detect anything between 100 and 1000 metabolites, and their quality is based on prior work identifying the metabolites with pure standards in in-house identification libraries. Public libraries are available, to characterise molecular features, but they only provide putative identifications as the certainty is insufficient to derive meaningful conclusions. In addition, machine learning approaches are used to identify the large number of new metabolites.(116) MS- and affinity-based metabolomics can detect several thousand human metabolites, although, as mentioned, the diverse nature of the metabolome necessitates the use of multiple analytical chemistry techniques (**Table 2**).(114)

Lipidomics is an especially promising metabolomics technique for biomarker discovery in steatotic liver disease. In a study of early ALD, the lipidomic signature of ALD patients began to differ from matched healthy controls as early as minimal fibrosis.(117) The bioactive lipid classes sphingomyelins and phosphocholines were downregulated in both liver tissue and plasma with increasing fibrosis stages and were both diagnostic of significant fibrosis and predicted liver-related outcomes. This finding was validated in an independent cohort of advanced ALD cirrhosis.(118) Other studies suggest that lipid panels can predict advanced forms of MASLD: molecular lipids in blood have shown good diagnostic performance for MASLD and MASH in well-powered studies, with elevated triglycerides and reduced lysophosphatidylcholines and phospholipids.(119, 120) Interestingly, unsaturated triglycerides are increased with the presence of the *PNPLA3* risk variant.(121) A 10-metabolite panel including eight eicosanoid molecules predicted advanced fibrosis with an area under the ROC

curve of 0.94.(122) Finally, recent data suggest that the liver lipidome of patients with ALD respond differently to acute alcohol intoxication than that of MASLD patients.(123) This finding indicates that there are likely distinct molecular differences between the two diseases, which may explain the marked difference in disease progression and risk of liver-related complications.

The use of metabolomics and lipidomics in hepatology is challenged by specificity, as most known metabolites have common disease pathways.(124) Furthermore, while some metabolites are found to be stable, others, such as glucose and cholesterol, have been shown to have a daily flux or to be affected by diet.(125) Hence, the establishment of a baseline level is important, especially when measured longitudinally throughout liver disease progression or regression.

Multi-omics

Clinical studies are increasingly generating multiple omics layers, allowing for integrated multi-omics investigations of liver disease.(126, 127) Machine-learning based feature selection from several omics layers can help determine the diagnostic and prognostic weight of each omics layer, but more importantly, multi-omics integration can capture disease complexity by addressing biologically relevant interactions between genes, their expression and their products. Unfortunately, integrating multiple types of omics remains a computational barrier. Consequently, current multi-omics studies rarely integrate more than two omics layers, and often instead interpret the outputs in parallel.(73, 128)

One study of multi-omics integration, performed GWAS in 9,491 MASLD patients and detected 20 gene variants predictive of steatosis and/or cirrhosis.(45) From this, the researchers combined GWAS with transcriptomics and proteomics to derive

expression quantitative trait loci and protein quantitative trait loci in the European populations. This multi-omics integration resulted in 16 putative genes associated with 273 circulating proteins, enriched in order to enable multiple metabolic and catabolic processes, including the metabolism of hormones, lipids, alcohol, vitamins, steroids and monocarboxylic acid. This represents an integrative step forward in understanding disease mechanisms.

The regulatory landscape from an omics perspective

The regulatory qualification of a biomarker requires thorough planning and patience.(11) For example, the Enhanced Liver Fibrosis test (Siemens Healthcare) obtained FDA approval in 2021, with the first core clinical study published in 2004 (**Figure 5**).(129, 130) The *nordic*PRO-C3™ biomarker (pro-peptide of type III collagen, Nordic Bioscience and Roche Diagnostics) took five years to complete assay development, minimising pre-analytical measurement uncertainty, followed by six years to create clinical evidence before having a Letter of Intent accepted by the FDA (**Figure 5**).

Every year, thousands of papers on biomarkers are published, yet very few enter clinical practice.(131) This so-called *valley of death* often happens because the transition from academic studies to implementation and commercialisation fails.

There are many reasons for the transition to fail. First, understanding the biological, pre-analytical and analytical factors that contribute to measurement uncertainty is important.(132) Second, when validating a biomarker, the FDA mandates the establishment of a predefined hypothesis and statistical analysis plan. Hence, the distribution of the cohort needs to allow for sufficient statistical power to address the potential context of use, whether it is diagnostic, prognostic or predictive. The 2016

BEST (Biomarkers, EndpointS and other Tools) resource from the FDA and National Institutes of Health Biomarker Work Group provides a notable glossary of biomarker definitions.(133) These considerations are important in moving from discovery to the internal and external validation of a biomarker. Third, for a study to adhere to Good Clinical Practice, regulatory standards, protocols and documents needs to be in place, describing procedures for sample collection and handling, measurement techniques and quality assurance systems. Fourth, biomarker measurements need to be conducted within certified laboratories and the informed consent process should encompass the explicit acceptance of sample utilisation for research, as well as for registration and commercialisation. To make a real difference, a biomarker needs to be implemented on a worldwide platform, and while many biomarkers may be interesting in a research setting, very few qualify according to the Clinical and Standards Institute guidelines.

The current failure of omics to transition from academic research to implementation and commercialisation may be partly due to the untargeted nature of most omics analyses, rendering them best suited for discovery. But the field also remains hampered by study designs dominated by retrospective studies without adherence to regulatory issues.(134) However, the burden is not only on biomarker research and development units, but also on regulatory agencies such as the FDA and the European Medicines Agency, which have been slow to adapt their approval procedures to the large data generated by omics on novel measurement platforms, associated by advanced biostatistical methods. Only in 2019 was the report from the Head of Medicines Agencies on Big Data issued, with a subgroup report on Bioanalytical Omics.(135, 136)

Conclusion

Omics technologies offer several advantages. They can identify associations between biomolecules and diseases, uncover underlying mechanisms and identify new biomarkers with untargeted hypothesis-free or targeted hypothesis-driven approaches. Despite the growing enthusiasm, we currently find ourselves in an exploratory phase where there is a lack of sufficient high-quality studies to provide the conclusive evidence of analytical validity, discovery, development and validation that would meet the requirements of regulatory authorities. The next five to ten years will inevitably provide crucial improvements in the evidence base and maturity of multi-omics, allowing for the first omics-based biomarkers to enter into clinical practice as precision tools for personalised medicine.

Abbreviations: ALD, Alcohol-related Liver Disease; cACLD, compensated Advanced Chronic Liver Disease; DC, Decompensated Cirrhosis; ELISA, enzyme-linked immunoassay; FDA, U.S. Food and Drug Administration; GCKR, glucokinase regulator; GPAM, glycerol-3-phosphate acyltransferase, mitochondrial; GRS, Genetic Risk Scores; GWAS, Genome-Wide Association Studies; HSD17B13, hydroxysteroid 17-beta dehydrogenase13; MASLD, Metabolic dysfunction associated Steatotic Liver Disease; MAF, minor allele frequency; MetALD, MASLD with increased alcohol intake; MBOAT7, membrane bound O-acyltransferase domain-containing 7; mRNA, messenger RNA; miRNA, microRNA; MTARC1, mitochondrial amidoxime reducing component 1; NMR, Nuclear Magnetic Resonance; NPV, Negative Predictive Value; PNPLA3, patatin-like phospholipase domain-containing protein 3; PPV, Positive Predictive Value; PSD3, pleckstrin and Sec7 domain-containing 3; qPCR, quantitative polymerase chain reaction; ROC, Receiver Operating Characteristics curve; rRNA, ribosomal RNA; SERPINA1, serpin family A member 1; SNP, Single Nucleotide Polymorphisms; TM6SF2, transmembrane 6 superfamily member 2.

Acknowledgements:

We acknowledge Louise Skovborg Just and Lise Ryborg for their invaluable work in the GALAXY and MicrobLiver consortia. We also wish to acknowledge the researchers and managerial staff who have contributed to the GALAXY and MicrobLiver at various times during the project. For the creation of Figure 5, we would like to acknowledge the invaluable input received from Siemens Healthcare, William Rosenberg and Echosens.

References

Author names in bold designate shared co-first authorship

1. Rehm J, Shield KD. Global Burden of Alcohol Use Disorders and Alcohol Liver Disease. Biomedicines 2019;7 |DOI: 10.3390/biomedicines7040099.

2. Riazi K, Azhari H, Charette JH, et al. The prevalence and incidence of NAFLD worldwide: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2022;7:851-861 |DOI: 10.1016/s2468-1253(22)00165-0.

3. Rinella ME, Lazarus JV, Ratziu V, et al. A multi-society Delphi consensus statement on new fatty liver disease nomenclature. J Hepatol 2023 |DOI: 10.1016/j.jhep.2023.06.003.

4. Hagström H, Thiele M, Roelstraete B, Söderling J, Ludvigsson JF. Mortality in biopsy-proven alcohol-related liver disease: a population-based nationwide cohort study of 3453 patients. Gut 2021;70:170-179 |DOI: 10.1136/gutjnl-2019-320446.

5. Sanyal AJ, Van Natta ML, Clark J, et al. Prospective Study of Outcomes in Adults with Nonalcoholic Fatty Liver Disease. New England Journal of Medicine 2021;385:1559-1569 |DOI: 10.1056/NEJMoa2029349.

6. Crabb DW, Im GY, Szabo G, Mellinger JL, Lucey MR. Diagnosis and Treatment of Alcohol-Associated Liver Diseases: 2019 Practice Guidance From the American Association for the Study of Liver Diseases. Hepatology 2020;71:306-333 |DOI: 10.1002/hep.30866.

7. Harrison SA, Allen AM, Dubourg J, Noureddin M, Alkhouri N. Challenges and opportunities in NASH drug development. Nature Medicine 2023;29:562-573 |DOI: 10.1038/s41591-023-02242-6.

8. Anstee QM, Castera L, Loomba R. Impact of non-invasive biomarkers on hepatology practice: Past, present and future. Journal of Hepatology 2022;76:1362-1378 |DOI: 10.1016/j.jhep.2022.03026.

9. Israelsen M, Torp N, Johansen S, et al. Validation of the new nomenclature of steatotic liver disease in patients with a history of excessive alcohol intake: an analysis of data from a prospective cohort study. Lancet Gastroenterol Hepatol 2024;9:218-228 |DOI: 10.1016/s2468-1253(23)00443-0.

10. Thiele M, Johansen S, Israelsen M, et al. Noninvasive assessment of hepatic decompensation. Hepatology 2023 |DOI: 10.1097/hep.0000000000000618.

11. Rasmussen DGK, Anstee QM, Torstenson R, et al. NAFLD and NASH biomarker qualification in the LITMUS consortium - Lessons learned. J Hepatol 2023;78:852-865 |DOI: 10.1016/j.jhep.2022.11.028.

12. Berzigotti A, Tsochatzis E, Boursier J, et al. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis – 2021 update. Journal of Hepatology 2021;75:659-689 |DOI: 10.1016/j.jeh.2021.05.025.

13. Usher-Smith JA, Sharp, Stephen J, Griffin SJ. The spectrum effect in tests for risk prediction, screening, and diagnosis. BMJ 2016;353 |DOI: 10.1136/bmj.i3139.

14. Ferkingstad E, Sulem P, Atlason BA, et al. Large-scale integration of the plasma proteome with genetics and disease. Nature Genetics 2021;53:1712-1721 |DOI: 10.1038/s41588-021-00978-w.

15. Niu L, Thiele M, Geyer PE, et al. Noninvasive proteomic biomarkers for alcohol-related liver disease. Nat Med 2022;28:1277-1287 |DOI: 10.1038/s41591-022-01850-y.

16. Masoodi M, Gastaldelli A, Hyötyläinen T, et al. Metabolomics and lipidomics in NAFLD: biomarkers and non-invasive diagnostic tests. Nature Reviews Gastroenterology & Hepatology 2021;18:835-856 |DOI: 10.1038/s41575-021-00502-9.

17. Angelini G, Panunzi S, Castagneto-Gissey L, et al. Accurate liquid biopsy for the diagnosis of non-alcoholic steatohepatitis and liver fibrosis. Gut 2022:gutjnl-2022-327498 |DOI: 10.1136/gutjnl-2022-327498.

18. Niu L, Sulek K, Vasilopoulou CG, et al. Defining NASH from a Multi-Omics Systems Biology Perspective. J Clin Med 2021;10:4673 |DOI: 10.3390/jcm10204673.

19. Niu L, Geyer PE, Gupta R, et al. Dynamic human liver proteome atlas reveals functional insights into disease pathways. Mol Syst Biol 2022;18:e10947 |DOI: 10.15252/msb.202210947.

20. **Qian T**, **Fujiwara N**, Koneru B, et al. Molecular Signature Predictive of Long-Term Liver Fibrosis Progression to Inform Antifibrotic Drug Development. Gastroenterology 2022;162:1210-1225 |DOI: 10.1053/j.gastro.2021.12.250.

21. Wouters OJ, McKee M, Luyten J. Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. JAMA 2020;323:844-853 |DOI: 10.1001/JAMA.2020.1166.

22. Wong VW, Adams LA, de Lédinghen V, Wong GL, Sookoian S. Noninvasive biomarkers in NAFLD and NASH - current progress and future promise. Nat Rev Gastroenterol Hepatol 2018;15:461-478 |DOI: 10.1038/s41575-018-0014-9.

23. Trebicka J, Bork P, Krag A, Arumugam M. Utilizing the gut microbiome in decompensated cirrhosis and acute-on-chronic liver failure. Nature Reviews Gastroenterology & Hepatology 2021;18:167-180 |DOI: 10.1038/s41575-020-00376-3.

24. Whitfield JB, Schwantes-An T-H, Darlay R, et al. A genetic risk score and diabetes predict development of alcohol-related cirrhosis in drinkers. Journal of Hepatology 2021 |DOI: 10.1016/j.jhep.2021.10.005.

25. Wetterstrand KA. *"The Cost of Sequencing a Human Genome"*. National Human Genome Research Institute: https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost. Published November 1, 2021. Accessed August 1, 2023

26. Schölz C, Lyon D, Refsgaard JC, et al. Avoiding abundance bias in the functional annotation of post-translationally modified proteins. Nat Methods 2015;12:1003-1004 |DOI: 10.1038/nmeth.3621.

27. Mato JM, Martínez-Chantar ML, Lu SC. Systems biology for hepatologists. Hepatology 2014;60:736-743 |DOI: 10.1002/hep.27023.

28. Geyer PE, Holdt LM, Teupser D, Mann M. Revisiting biomarker discovery by plasma proteomics. Mol Syst Biol 2017;13:942 |DOI: 10.15252/msb.20156297.

29. Geyer PE, Voytik E, Treit PV, et al. Plasma Proteome Profiling to detect and avoid sample-related biases in biomarker studies. EMBO Molecular Medicine 2019;11:e10427 |DOI: 10.15252/emmm.201910427.

30. **Feldman A**, **Eder SK**, Felder TK, et al. Clinical and Metabolic Characterization of Lean Caucasian Subjects With Non-alcoholic Fatty Liver. Am J Gastroenterol 2017;112:102-110 |DOI: 10.1038/ajg.2016.318.

31. López-Vicario C, Checa A, Urdangarin A, et al. Targeted lipidomics reveals extensive changes in circulating lipid mediators in patients with acutely decompensated cirrhosis. J Hepatol 2020 |DOI: 10.1016/j.jhep.2020.03.046.

32. Marchand CR, Farshidfar F, Rattner J, Bathe OF. A Framework for Development of Useful Metabolomic Biomarkers and Their Effective Knowledge Translation. Metabolites 2018;8:59 |DOI: 10.3390/metabo8040059.

33. de Gonzalo-Calvo D, Marchese M, Hellemans J, et al. Consensus guidelines for the validation of qRT-PCR assays in clinical research by the CardioRNA consortium. Mol Ther Methods Clin Dev 2022;24:171-180 |DOI: 10.1016/j.omtm.2021.12.007.

34. Buch S, Stickel F, Trepo E, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nature Genetics 2015;47:1443-1448 |DOI: 10.1038/ng.3417.

35. Anstee QM, Darlay R, Cockell S, et al. Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically-characterised cohort. J Hepatol 2020;73:505-515 |DOI: 10.1016/j.jhep.2020.04.003.

36. Sookoian S, Pirola CJ. Genetic predisposition in nonalcoholic fatty liver disease. Clin Mol Hepatol 2017;23:1-12 |DOI: 10.3350/cmh.2016.0109.

37. Stickel F, Moreno C, Hampe J, Morgan MY. The genetics of alcohol dependence and alcohol-related liver disease. Journal of Hepatology 2017;66:195-211 |DOI: 10.1016/j.jhep.2016.08.011.

38. **Kranzler HR**, **Zhou H**, **Kember RL**, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. Nat Commun 2019;10:1499 |DOI: 10.1038/s41467-019-09480-8.

39. Uffelmann E, Huang QQ, Munung NS, et al. Genome-wide association studies. Nature Reviews Methods Primers 2021;1:59 |DOI: 10.1038/s43586-021-00056-9.

40. Zheng M, Hakim A, Konkwo C, et al. Advancing diagnosis and management of liver disease in adults through exome sequencing. EBioMedicine 2023;95:104747 |DOI: 10.1016/j.ebiom.2023.104747.

41. Emdin CA, Haas M, Ajmera V, et al. Association of genetic variation with cirrhosis: a multi-trait genome-wide association and gene-environment interaction study. Gastroenterology 2021;160:1620-1633 |DOI: 10.1053/j.gastro.2020.12.011.

42. **Haas ME**, **Pirruccello JP**, **Friedman SN**, et al. Machine learning enables new insights into genetic contributions to liver fat accumulation. Cell Genom 2021;1 |DOI: 10.1016/j.xgen.2021.100066.

43. **Parisinos CA**, **Wilman HR**, Thomas EL, et al. Genome-wide and Mendelian randomisation studies of liver MRI yield insights into the pathogenesis of steatohepatitis. J Hepatol 2020;73:241-251 |DOI: 10.1016/j.jhep.2020.03.032.

44. **Vujkovic M**, **Ramdas S**, Lorenz KM, et al. A multiancestry genome-wide association study of unexplained chronic ALT elevation as a proxy for nonalcoholic fatty liver disease with histological and radiological validation. Nat Genet 2022;54:761-771 |DOI: 10.1038/s41588-022-01078-z.

45. **Sveinbjornsson G**, **Ulfarsson MO**, Thorolfsdottir RB, et al. Multiomics study of nonalcoholic fatty liver disease. Nat Genet 2022;54:1652-1663 |DOI: 10.1038/s41588-022-01199-5.

46. **Mancina RM**, **Sasidharan K**, Lindblom A, et al. PSD3 downregulation confers protection against fatty liver disease. Nat Metab 2022;4:60-75 |DOI: 10.1038/s42255-021-00518-0.

47. Gellert-Kristensen H, Richardson TG, Davey Smith G, et al. Combined Effect of PNPLA3, TM6SF2, and HSD17B13 Variants on Risk of Cirrhosis and Hepatocellular Carcinoma in the General Population. Hepatology 2020;72:845-856 |DOI: 10.1002/hep.31238.

48. Liu Z, Suo C, Shi O, et al. The Health Impact of MAFLD, a Novel Disease Cluster of NAFLD, Is Amplified by the Integrated Effect of Fatty Liver Disease-Related Genetic Variants. Clin Gastroenterol Hepatol 2022;20:e855-e875 |DOI: 10.1016/j.cgh.2020.12.033.

49. Israelsen M, Juel HB, Detlefsen S, et al. Metabolic and Genetic Risk Factors Are the Strongest Predictors of Severity of Alcohol-Related Liver Fibrosis. Clin Gastroenterol Hepatol 2022;20:1784-1794.e1789 |DOI: 10.1016/j.cgh.2020.11.038.

50. Innes H, Morling JR, Buch S, et al. Performance of routine risk scores for predicting cirrhosis-related morbidity in the community. Journal of Hepatology 2022;77:365-376 |DOI: 10.1016/j.jhep.2022.02.022.

51. Johansen S, Thiele M, Juel HB, Hansen T, Krag A. External validation of a genetic risk score that predicts development of alcohol-related cirrhosis. J Hepatol 2022;77:1720-1721 |DOI: 10.1016/j.jhep.2022.06.006.

52. **De Vincentis A**, **Tavaglione F**, Jamialahmadi O, et al. A Polygenic Risk Score to Refine Risk Stratification and Prediction for Severe Liver Disease by Clinical Fibrosis Scores. Clin Gastroenterol Hepatol 2022;20:658-673 |DOI: 10.1016/j.cgh.2021.05.056.

53. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nature Protocols 2020;15:2759-2772 |DOI: 10.1038/s41596-020-0353-1.

54. **Ozata DM**, **Gainetdinov I**, Zoch A, O'Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. Nature Reviews Genetics 2019;20:89-108 |DOI: 10.1038/s41576-018-0073-3.

55. Pandey KK, Madhry D, Ravi Kumar YS, et al. Regulatory roles of tRNA-derived RNA fragments in human pathophysiology. Mol Ther Nucleic Acids 2021;26:161-173 |DOI: 10.1016/j.omtn.2021.06.023.

56. Atic AI, Thiele M, Munk A, Dalgaard LT. Circulating miRNAs associated with nonalcoholic fatty liver disease. American Journal of Physiology-Cell Physiology 2023;324:C588-C602 |DOI: 10.1152/ajpcell.00253.2022.

57. Sletten AC, Davidson JW, Yagabasan B, et al. Loss of SNORA73 reprograms cellular metabolism and protects against steatohepatitis. Nat Commun 2021;12:5214 |DOI: 10.1038/s41467-021-25457-y.

58. Huang P, Tu B, Liao H-J, et al. Elevation of plasma tRNA fragments as a promising biomarker for liver fibrosis in nonalcoholic fatty liver disease. Scientific reports 2021;11:5886 |DOI: 10.1038/s41598-021-85421-0.

59. Srinivasan S, Yeri A, Cheah PS, et al. Small RNA Sequencing across Diverse Biofluids Identifies Optimal Methods for exRNA Isolation. Cell 2019;177:446-462.e416 |DOI: 10.1016/j.cell.2019.03.024.

60. **Murillo OD**, **Thistlethwaite W**, Rozowsky J, et al. exRNA Atlas Analysis Reveals Distinct Extracellular RNA Cargo Types and Their Carriers Present across Human Biofluids. Cell 2019;177:463-477.e415 |DOI: 10.1016/j.cell.2019.02.018.

61. Hess AL, Larsen LH, Udesen PB, et al. Levels of Circulating miR-122 are Associated with Weight Loss and Metabolic Syndrome. Obesity (Silver Spring) 2020;28:493-501 |DOI: 10.1002/oby.22704.

62. Willeit P, Skroblin P, Moschen AR, et al. Circulating MicroRNA-122 Is Associated With the Risk of New-Onset Metabolic Syndrome and Type 2 Diabetes. Diabetes 2017;66:347-357 |DOI: 10.2337/db16-0731.

63. Hendy OM, Rabie H, El Fouly A, et al. The Circulating Micro-RNAs (-122, -34a and -99a) as Predictive Biomarkers for Non-Alcoholic Fatty Liver Diseases. Diabetes Metab Syndr Obes 2019;12:2715-2723 |DOI: 10.2147/dmso.S231321.

64. Harrison SA, Ratziu V, Magnanensi J, et al. NIS2+™, an optimisation of the blood-based biomarker NIS4® technology for the detection of at-risk NASH: A prospective derivation and validation study. J Hepatol 2023;79:758-767 |DOI: 10.1016/j.jhep.2023.04.031.

65. Zhang X, Mens MMJ, Abozaid YJ, et al. Circulatory microRNAs as potential biomarkers for fatty liver disease: the Rotterdam study. Aliment Pharmacol Ther 2021;53:432-442 |DOI: 10.1111/apt.16177.

66. Johnson K, Leary PJ, Govaere O, et al. Increased serum miR-193a-5p during non-alcoholic fatty liver disease progression: Diagnostic and mechanistic relevance. JHEP Rep 2022;4:100409 |DOI: 10.1016/j.jhepr.2021.100409.

67. Waidmann O, Köberle V, Brunner F, et al. Serum microRNA-122 predicts survival in patients with liver cirrhosis. PLoS One 2012;7:e45652 |DOI: 10.1371/journal.pone.0045652.

68. **Huttenhower C**, **Gevers D**, Knight R, et al. Structure, function and diversity of the healthy human microbiome. Nature 2012;486:207-214 |DOI: 10.1038/nature11234.

69. Velasquez-Manoff M. Gut Microbiome: The Peacekeepers. Nature 2015;518:S3-S11 |DOI: 10.1038/518S3a.

70. Lynch SV, Pedersen O. The Human Intestinal Microbiome in Health and Disease. New England Journal of Medicine 2016;375:2369-2379 |DOI: doi:10.1056/NEJMra1600266.

71. Lang S, Schnabl B. Microbiota and Fatty Liver Disease-the Known, the Unknown, and the Future. Cell Host Microbe 2020;28:233-244 |DOI: 10.1016/j.chom.2020.07.007.

72. Tripathi A, Debelius J, Brenner DA, et al. The gut-liver axis and the intersection with the microbiome. Nat Rev Gastroenterol Hepatol 2018;15:397-411 |DOI: 10.1038/s41575-018-0011-z.

73. Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature 2019;569:655-662 |DOI: 10.1038/s41586-019-1237-9.

74. Loomba R, Seguritan V, Li W, et al. Gut Microbiome-Based Metagenomic Signature for Non-invasive Detection of Advanced Fibrosis in Human Nonalcoholic Fatty Liver Disease. Cell Metab 2017;25:1054-1062.e1055 |DOI: 10.1016/j.cmet.2017.04.001.

75. Zhu L, Baker SS, Gill C, et al. Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: a connection between endogenous alcohol and NASH. Hepatology 2013;57:601-609 |DOI: 10.1002/hep.26093.

76. Boursier J, Mueller O, Barret M, et al. The severity of nonalcoholic fatty liver disease is associated with gut dysbiosis and shift in the metabolic function of the gut microbiota. Hepatology 2016;63:764-775 |DOI: 10.1002/hep.28356.

77. **Qin N**, **Yang F**, **Li A**, et al. Alterations of the human gut microbiome in liver cirrhosis. Nature 2014;513:59-64 |DOI: 10.1038/nature13568.

78. Trebicka J, Macnaughtan J, Schnabl B, Shawcross DL, Bajaj JS. The microbiota in cirrhosis and its role in hepatic decompensation. Journal of Hepatology 2021;75:S67-S81 |DOI: 10.1016/j.jhep.2020.11.013.

79. Parlesak A, Schäfer C, Schütz T, Bode JC, Bode C. Increased intestinal permeability to macromolecules and endotoxemia in patients with chronic alcohol abuse in different stages of alcohol-induced liver disease. J Hepatol 2000;32:742-747 |DOI: 10.1016/s0168-8278(00)80242-1.

80. Miele L, Valenza V, La Torre G, et al. Increased intestinal permeability and tight junction alterations in nonalcoholic fatty liver disease. Hepatology 2009;49:1877-1887 |DOI: 10.1002/hep.22848.

81. Bajaj JS, Reddy KR, O'Leary JG, et al. Serum Levels of Metabolites Produced by Intestinal Microbes and Lipid Moieties Independently Associated With Acute-on-Chronic Liver Failure and Death in Patients With Cirrhosis. Gastroenterology 2020;159:1715-1730.e1712 |DOI: 10.1053/j.gastro.2020.07.019.

82. Seki E, De Minicis S, Osterreicher CH, et al. TLR4 enhances TGF-beta signaling and hepatic fibrosis. Nat Med 2007;13:1324-1332 |DOI: 10.1038/nm1663.

83. Patel VC, Lee S, McPhail MJW, et al. Rifaximin-α reduces gut-derived inflammation and mucin degradation in cirrhosis and encephalopathy: RIFSYS randomised controlled trial. J Hepatol 2022;76:332-342 |DOI: 10.1016/j.jhep.2021.09.010.

84. Wirbel J, Pyl PT, Kartal E, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nat Med 2019;25:679-689 |DOI: 10.1038/s41591-019-0406-6.

85. **Kartal E**, **Schmidt TSB**, Molina-Montes E, et al. A faecal microbiota signature with high specificity for pancreatic cancer. Gut 2022;71:1359-1372 |DOI: 10.1136/gutjnl-2021-324755.

86. **Garmaeva S**, **Sinha T**, Kurilshikov A, et al. Studying the gut virome in the metagenomic era: challenges and perspectives. BMC Biol 2019;17:84 |DOI: 10.1186/s12915-019-0704-y.

87. **Espen LV**, **Bak E**, Beller L, et al. A previously undescribed highly prevalent phage identified in a Danish enteric virome catalog. mSystems 2021;6:e0038221 |DOI: 10.1128/mSystems.00382-21.

88. Townsend EM, Kelly L, Muscatt G, et al. The Human Gut Phageome: Origins and Roles in the Human Gut Microbiome. Front Cell Infect Microbiol 2021;11:643214 |DOI: 10.3389/fcimb.2021.643214.

89. **Bin Jang H**, **Bolduc B**, Zablocki O, et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nature Biotechnology 2019;37:632-639 |DOI: 10.1038/s41587-019-0100-8.

90. Roux S, Camargo AP, Coutinho FH, et al. iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. PLoS Biol 2023;21:e3002083 |DOI: 10.1371/journal.pbio.3002083.

91. Tisza MJ, Belford AK, Domínguez-Huerta G, Bolduc B, Buck CB. Cenote-Taker 2 democratizes virus discovery and sequence annotation. Virus Evol 2021;7:veaa100 |DOI: 10.1093/ve/veaa100.

92. Camargo AP, Roux S, Schulz F, et al. Identification of mobile genetic elements with geNomad. Nature Biotechnology 2023 |DOI: 10.1038/s41587-023-01953-y.

93. Virgin HW. The virome in mammalian physiology and disease. Cell 2014;157:142-150 |DOI: 10.1016/j.cell.2014.02.032.

94. **Lang S**, **Demir M**, Martin A, et al. Intestinal Virome Signature Associated With Severity of Nonalcoholic Fatty Liver Disease. Gastroenterology 2020;159:1839-1852 |DOI: 10.1053/j.gastro.2020.07.005.

95. Hsu CL, Zhang X, Jiang L, et al. Intestinal virome in patients with alcohol use disorder and after abstinence. Hepatol Commun 2022 |DOI: 10.1002/hep4.1947.

96. **Jiang L**, **Lang S**, Duan Y, et al. Intestinal Virome in Patients With Alcoholic Hepatitis. Hepatology 2020;72:2182-2196 |DOI: 10.1002/hep.31459.

97. Bajaj JS, Sikaroodi M, Shamsaddini A, et al. Interaction of bacterial metagenome and virome in patients with cirrhosis and hepatic encephalopathy. Gut 2021;70:1162-1173 |DOI: 10.1136/gutjnl-2020-322470.

98. Shkoporov AN, Clooney AG, Sutton TDS, et al. The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. Cell Host Microbe 2019;26:527-541.e525 |DOI: 10.1016/j.chom.2019.09.009.

99. **de Jonge PA**, **Wortelboer K**, Scheithauer TPM, et al. Gut virome profiling identifies a widespread bacteriophage family associated with metabolic syndrome. Nat Commun 2022;13:3594 |DOI: 10.1038/s41467-022-31390-5.

100. Lang S, Duan Y, Liu J, et al. Intestinal Fungal Dysbiosis and Systemic Immune Response to Fungi in Patients With Alcoholic Hepatitis. Hepatology 2020;71:522-538 |DOI: 10.1002/hep.30832.

101. **Yang AM**, **Inamine T**, Hochrath K, et al. Intestinal fungi contribute to development of alcoholic liver disease. J Clin Invest 2017;127:2829-2841 |DOI: 10.1172/jci90562.

102. Bajaj JS, Liu EJ, Kheradman R, et al. Fungal dysbiosis in cirrhosis. Gut 2018;67:1146-1154 |DOI: 10.1136/gutjnl-2016-313170.

103. Sun A, Jiang Y, Wang X, et al. Liverbase: a comprehensive view of human liver biology. J Proteome Res 2010;9:50-58 |DOI: 10.1021/pr900191p.

104. Sinha A, Mann M. A beginner's guide to mass spectrometry–based proteomics. The Biochemist 2020;42:64-69 |DOI: 10.1042/bio20200057.

105. Geyer PE, Kulak NA, Pichler G, et al. Plasma Proteome Profiling to Assess Human Health and Disease. Cell Syst 2016;2:185-195 |DOI: 10.1016/j.cels.2016.02.015.

106. Luo Y, Wadhawan S, Greenfield A, et al. SOMAscan Proteomics Identifies Serum Biomarkers Associated With Liver Fibrosis in Patients With NASH. Hepatology Communications 2021;5:760-773 |DOI: 10.1002/hep4.1670.

107. Fourman LT, Stanley TL, Ockene MW, et al. Proteomic Analysis of Hepatic Fibrosis in Human Immunodeficiency Virus–Associated Nonalcoholic Fatty Liver Disease Demonstrates Up-regulation of Immune Response and Tissue Repair Pathways. The Journal of Infectious Diseases 2022;227:565-576 |DOI: 10.1093/infdis/jiac475.

108. Katz DH, Robbins JM, Deng S, et al. Proteomic profiling platforms head to head: Leveraging genetics and clinical traits to compare aptamer- and antibody-based methods. Science Advances 2022;8:eabm5164 |DOI: doi:10.1126/sciadv.abm5164.

109. Bader JM, Albrecht V, Mann M. MS-Based Proteomics of Body Fluids: The End of the Beginning. Mol Cell Proteomics 2023;22:100577 |DOI: 10.1016/j.mcpro.2023.100577.

110. Fung ET. A recipe for proteomics diagnostic test development: the OVA1 test, from biomarker discovery to FDA clearance. Clin Chem 2010;56:327-329 |DOI: 10.1373/clinchem.2009.140855.

111. Niu L, Geyer PE, Wewer Albrechtsen NJ, et al. Plasma proteome profiling discovers novel proteins associated with non-alcoholic fatty liver disease. Molecular Systems Biology 2019;15:e8793 |DOI: 10.15252/msb.20188793.

112. Govaere O, Hasoon M, Alexander L, et al. A proteo-transcriptomic map of non-alcoholic fatty liver disease signatures. Nature Metabolism 2023 |DOI: 10.1038/s42255-023-00775-1.

113. Sanyal AJ, Williams SA, Lavine JE, et al. Defining the serum proteomic signature of hepatic steatosis, inflammation, ballooning and fibrosis in non-alcoholic fatty liver disease. Journal of Hepatology 2023;78:693-703 |DOI: 10.1016/j.jhep.2022.11.029.

114. Wishart DS, Guo A, Oler E, et al. HMDB 5.0: the Human Metabolome Database for 2022. Nucleic Acids Res 2022;50:D622-d631 |DOI: 10.1093/nar/gkab1062.

115. Cabezas J, Lucey MR, Bataller R. Biomarkers for monitoring alcohol use. Clin Liver Dis (Hoboken) 2016;8:59-63 |DOI: 10.1002/cld.571.

116. Guijas C, Montenegro-Burke JR, Domingo-Almenara X, et al. METLIN: A Technology Platform for Identifying Knowns and Unknowns. Anal Chem 2018;90:3156-3164 |DOI: 10.1021/acs.analchem.7b04424.

117. **Thiele M**, **Suvitaival T**, Trošt K, et al. Sphingolipids Are Depleted in Alcohol-Related Liver Fibrosis. Gastroenterology 2023;164:1248-1260 |DOI: 10.1053/j.gastro.2023.02.023.

118. Kronborg TM, Gao Q, Trošt K, et al. Low sphingolipid levels predict poor survival in patients with alcohol-related liver disease. JHEP Rep 2024;6:100953 |DOI: 10.1016/j.jhepr.2023.100953.

119. **Luukkonen PK**, **Zhou Y**, Sädevirta S, et al. Hepatic ceramides dissociate steatosis and insulin resistance in patients with non-alcoholic fatty liver disease. J Hepatol 2016;64:1167-1175 |DOI: 10.1016/j.jhep.2016.01.002.

120. **Mayo R**, **Crespo J**, **Martínez-Arranz I**, et al. Metabolomic-based noninvasive serum test to diagnose nonalcoholic steatohepatitis: Results from discovery and validation cohorts. Hepatol Commun 2018;2:807-820 |DOI: 10.1002/hep4.1188.

121. Orešič M, Hyötyläinen T, Kotronen A, et al. Prediction of non-alcoholic fatty-liver disease and liver fat content by serum molecular lipids. Diabetologia 2013;56:2266-2274 |DOI: 10.1007/s00125-013-2981-2.

122. Caussy C, Ajmera VH, Puri P, et al. Serum metabolites detect the presence of advanced fibrosis in derivation and validation cohorts of patients with non-alcoholic fatty liver disease. Gut 2019;68:1884-1892 |DOI: 10.1136/gutjnl-2018-317584.

123. **Israelsen M**, **Kim M**, Suvitaival T, et al. Comprehensive lipidomics reveals phenotypic differences in hepatic lipid turnover in ALD and NAFLD during alcohol intoxication. JHEP Rep 2021;3:100325 |DOI: 10.1016/j.jhepr.2021.100325.

124. Pietzner M, Stewart ID, Raffler J, et al. Plasma metabolites to profile pathways in noncommunicable disease multimorbidity. Nat Med 2021;27:471-479 |DOI: 10.1038/s41591-021-01266-0.

125. Kervezee L, Cermakian N, Boivin DB. Individual metabolomic signatures of circadian misalignment during simulated night shifts in humans. PLoS Biol 2019;17:e3000303 |DOI: 10.1371/journal.pbio.3000303.

126. Atabaki-Pasdar N, Ohlsson M, Viñuela A, et al. Predicting and elucidating the etiology of fatty liver disease: A machine learning modeling and validation study in the IMI DIRECT cohorts. PLoS Med 2020;17:e1003149 |DOI: 10.1371/journal.pmed.1003149.

127. Wood GC, Chu X, Argyropoulos G, et al. A multi-component classifier for nonalcoholic fatty liver disease (NAFLD) based on genomic, proteomic, and phenomic data domains. Sci Rep 2017;7:43238 |DOI: 10.1038/srep43238.

128. Tarazona S, Arzalluz-Luque A, Conesa A. Undisclosed, unmet and neglected challenges in multi-omics studies. Nature Computational Science 2021;1:395-402 |DOI: 10.1038/s43588-021-00086-z.

129. Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. Gastroenterology 2004;127:1704-1713 |DOI: 10.1053/j.gastro.2004.08.052.

130. Gee Matthew SHDI. Prognostic test for assessment of liver related disease progression. In: Kelm KB, editor. www.accessdata.fda.gov; 2021.

131. Ioannidis JPA, Bossuyt PMM. Waste, Leaks, and Failures in the Biomarker Pipeline. Clin Chem 2017;63:963-972 |DOI: 10.1373/clinchem.2016.254649.

132. Selby PJ, Banks RE, Gregory W, et al. Methods for the evaluation of biomarkers in patients with kidney and liver diseases: multicentre research programme including ELUCIDATE RCT. Programme Grants Appl Res 2018;6 |DOI: 10.3310/pgfar06030.

133. Group F-NBW. 2016. BEST (Biomarkers, EndpointS, and other Tools) National Institutes of Health (US), Bethesda (MD). https://www.ncbi.nlm.nih.gov/books/NBK326791/

134. McShane LM, Cavenagh MM, Lively TG, et al. Criteria for the use of omics-based predictors in clinical trials: explanation and elaboration. BMC Medicine 2013;11:220 |DOI: 10.1186/1741-7015-11-220.

135. König R, Cave A, Goldammer M, Meulendijks D. 2019. Bioanalytical Omics Subgroup report www.ema.europa.eu.

136. Agency EM. 2019. HMA-EMA Joint Big Data Taskforce. Heads of Medicines Agencies; www.hma.eu/about-hma/working-groups/hma/ema-joint-big-data-steering-group/hma/ema-joint-big-data-steering-group.html.

137. **Lehmann R**, **Franken H**, Dammeier S, et al. Circulating lysophosphatidylcholines are markers of a metabolically benign nonalcoholic fatty liver. Diabetes Care 2013;36:2331-2338 |DOI: 10.2337/dc12-1760.

138. Borisevich D, Schnurr TM, Engelbrechtsen L, et al. Non-linear interaction between physical activity and polygenic risk score of body mass index in Danish and Russian populations. PLOS ONE 2021;16:e0258748 |DOI: 10.1371/journal.pone.0258748.

139. **Duan Y**, **Llorente C**, Lang S, et al. Bacteriophage targeting of gut bacterium attenuates alcoholic liver disease. Nature 2019;575:505-511 |DOI: 10.1038/s41586-019-1742-x.

140. **Harrison SA**, **Ratziu V**, Boursier J, et al. A blood-based biomarker panel (NIS4) for non-invasive diagnosis of non-alcoholic steatohepatitis and liver fibrosis: a prospective derivation and global validation study. Lancet Gastroenterol Hepatol 2020;5:970-985 |DOI: 10.1016/s2468-1253(20)30252-1.

141. Ramachandran P, Dobie R, Wilson-Kanamori JR, et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. Nature 2019;575:512-518 |DOI: 10.1038/s41586-019-1631-3.

142. Indira Chandran V, Wernberg CW, Lauridsen MM, et al. Circulating TREM2 as a noninvasive diagnostic biomarker for NASH in patients with elevated liver stiffness. Hepatology 2023;77:558-572 |DOI: 10.1002/hep.32620.

143. Kothari V, Savard C, Tang J, et al. sTREM2 is a plasma biomarker for human NASH and promotes hepatocyte lipid accumulation. Hepatol Commun 2023;7 |DOI: 10.1097/hc9.0000000000000265.

144. Gaggini M, Carli F, Rosso C, et al. Altered amino acid concentrations in NAFLD: Impact of obesity and insulin resistance. Hepatology 2018;67:145-158 |DOI: 10.1002/hep.29465.

145. **Zhou Y**, **Llauradó G**, Orešič M, et al. Circulating triacylglycerol signatures and insulin sensitivity in NAFLD associated with the E167K variant in TM6SF2. J Hepatol 2015;62:657-663 |DOI: 10.1016/j.jhep.2014.10.010.

146. Muta K, Saito K, Kemmochi Y, et al. Phosphatidylcholine (18:0/20:4), a potential biomarker to predict ethionamide-induced hepatic steatosis in rats. J Appl Toxicol 2022;42:1533-1547 |DOI: 10.1002/jat.4324.

147. Slieker RC, Donnelly LA, Fitipaldi H, et al. Distinct Molecular Signatures of Clinical Clusters in People With Type 2 Diabetes: An IMI-RHAPSODY Study. Diabetes 2021;70:2683-2693 |DOI: 10.2337/db20-1281.

148. Wittenbecher C, Cuadrat R, Johnston L, et al. Dihydroceramide- and ceramide-profiling provides insights into human cardiometabolic disease etiology. Nature Communications 2022;13:936 |DOI: 10.1038/s41467-022-28496-1.

149. Collins JM, Isaacs C. Management of breast cancer risk in BRCA1/2 mutation carriers who are unaffected with cancer. Breast J 2020;26:1520-1527 |DOI: 10.1111/tbj.13970.

150. Nashed AL, Rao KW, Gulley ML. Clinical applications of BCR-ABL molecular testing in acute leukemia. J Mol Diagn 2003;5:63-72 |DOI: 10.1016/s1525-1578(10)60454-0.

1018 **Table 1**

	Diagnostic	Prognostic	Monitoring	Prediction*	Surrogate endpoint
Outcome of interest	Disease present or not; disease staging	Development of clinical events, mortality	Change in disease severity	Effect of treatment	Substitute for one or more clinical outcomes
Subclasses of biomarkers	Screening	Susceptibility/ risk stratification	Efficacy of intervention; pharmacodynamic response	Safety (adverse events)	Reasonably likely surrogate endpoint
Measurement timing	Baseline	Baseline	Longitudinal	Baseline, before intervention	Start and end of intervention study
Clinical characteristic	Reflects true disease state	Reflects patient or disease characteristics	Biomarker changes correlate with changes in extent or status of disease	Reflects patient or disease characteristics	Effect on the surrogate endpoint predicts a clinical benefit
Statistics used	Discriminative accuracy, sensitivity, specificity, NPV, PPV, calibration curves, goodness of fit, information criterium, odds ratio	C-statistics, hazard ratio, time-dependent receiver operating characteristics curve, Aalen-Johansen or Kaplan-Meier estimator	Correlation coefficients: diagnostic and prognostic accuracy of Δbiomarker**	Treatment effect in biomarker positive vs. biomarker negative patients if patient groups have the same prognosis	Correlation coefficients: diagnostic accuracy of Δbiomarker to detect change; prognostic accuracy of Δbiomarker
Examples of omics-based biomarkers	Proteomics for diagnosis of ALD fibrosis, inflammation and steatosis(15)	Genetic risk polymorphisms for development of hepatocellular carcinoma in the population(47)	Changes in Lyso-phosphocholines by lipidomics in MASLD during dietary intervention(137)	A polygenic score to predict weight loss in response to physical activity(138)	No omics markers approved as surrogate endpoints, but single molecules may arise from omics discovery

*A prognostic biomarker is used to identify the likelihood of a clinical event in a patient, while predictive biomarkers identify patients who are more likely to experience beneficial or adverse effects of an intervention. **Δ means change from baseline. **Abbreviations:** ALD, alcohol-related liver disease; MASLD, metabolic dysfunction-associated steatotic liver disease; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristics curve.

1019 **Table 1.** Biomarker indications and clinical use.

1020

Table 2

	Specimen	Outcomes of interest	Technology (untargeted)	Technology (targeted)	Number of analytes (targeted tech.)	Examples of biomarker candidates
Genetics	Whole blood, buffy coat	SNPs, candidate genes, GRS, polygenic scores	Whole genome sequencing	Microarray-based genotyping or whole exome sequencing	>6*10^6 common SNPs (MAF > 0.01)	*PNPLA3*, *TM6SF2*, *GCKR*, *MBOAT7*, *HSD17B13, SERPINA1*(14, 41, 45, 139)
Transcriptomics	All tissue types, plasma, serum, whole blood	RNA sequences: non-coding RNA (miRNA, long noncoding RNA), coding mRNA, steady state RNA levels	Reverse transcription-quantitative polymerase chain reaction. Small RNA-sequencing	Reverse transcription-quantitative polymerase chain reaction Targeted sequencing panels	10^5	miR-34a(140), miR-122, miR-21
Proteomics	All tissue types and body fluids	Protein abundance	Mass spectrometry, Proximity Extension Assay (commercialized by Olink Explore) and SomaScan Assay (commercialized by SomaLogic)	Mass spectrometry (parallel or multiple reaction monitoring). Proximity Extension Assay (used by Olink Target), ELISA	1 -10^4	TREM-2 was discovered by single-cell sequencing, subsequently developed into an ELISA assay.(141-143) Compliment component C7 identified as a fibrosis marker in two independent biomarker studies.(15, 113)
Metabolomics and lipidomics	Plasma, urine, stool, liver, adipose tissue	Metabolite abundance. Lipid abundance w.r.t. lipid class, lipid saturation / unsaturation, lipid size	Gas or liquid chromatography coupled to mass-spectrometry	Triple-quadrupole mass-spectrometry, NMR spectroscopy	10^2 -10^3	Glutamate and glutamine (144) Triglycerides, such as TG(48:0)(145) and TG(50:2)(117); Phosphatidylcholines, such as PC(36:4)(146); Sphingomyelins, such as SM(41:1)(117, 118) The Metabolomics-Advanced steatohepatitis fibrosis score developed to detect at-risk MASH
Viromics	Stool, saliva, plasma, skin	Viral genomes (DNA or RNA) and their encoded genes	Shotgun metagenomic sequencing	Quantitative polymerase chain reaction	Variable (depending on sequencing depth and sample diversity) with limited overlap between samples	None, but bacteriophages which target cytolytic *Enterococcus faecalis* could potentially be markers of resistance against alcohol-induced liver injury.(139)
Microbiomics	Stool, saliva, skin, mucosa	Bacteriomics	Shotgun metagenomics or amplicon sequencing	Quantitative polymerase chain reaction or Antibody test	100-1,000 species per sample, with 10^5 - 10^6 genes	Cytolytic *Enterococcus faecalis* (139)

Abbreviations: ELISA, enzyme-linked immunoassay; GRS, Genetic Risk Scores; MAF, minor allele frequency; SNP, Single NucleotidePolymorphisms; miRNA, microRNA; NMR, nuclear magnetic resonance

1021 **Table 2.** Omics based biomarkers in hepatology

Legends

Figure 1. **Intended use of biomarkers and the spectrum bias.**

Due to the spectrum effect, diagnostic accuracy for the same biomarker will change when tested in populations with different prevalence of disease. Discrete types of omics allow biomarkers to be tailored to the different contexts of use and different disease spectrums. Plots illustrate variability in sensitivity and specificity, as well as, PPV and NPV with disease prevalence in the studied cohort, derived from Usher-Smith et al.(13). cACLD, compensated advanced chronic liver disease; DC, decompensated cirrhosis; F0 – F2, denotes liver fibrosis stage; NPV, negative predictive value; PPV, positive predictive value.

Figure 2. The potential of omics-based biomarkers

Illustrated by layers of biological signals and the complexity of biological molecules within the human body. The environmental signals introduce another layer of complexity as individual risk factors of disease. MAF, minor allele frequency; SNP, Single-Nucleotide Polymorphism.

Figure 3. Omics timeline with major scientific and technological breakthroughs, using genetics as reference.

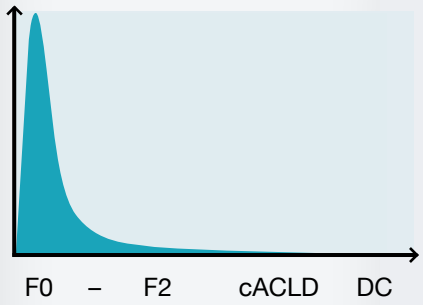
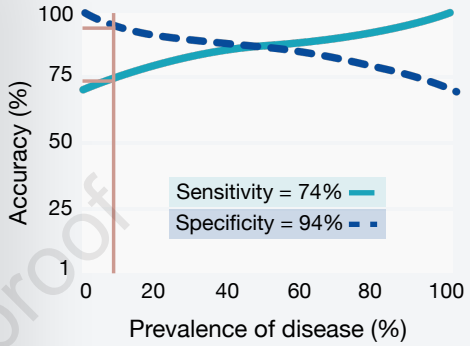
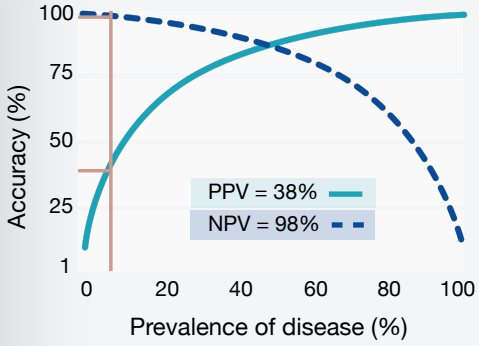
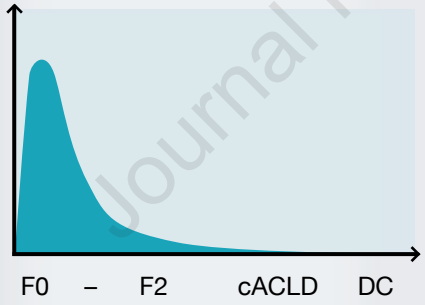
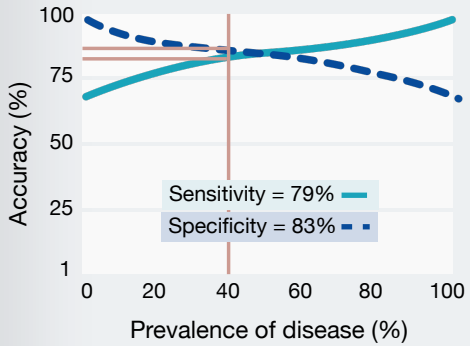
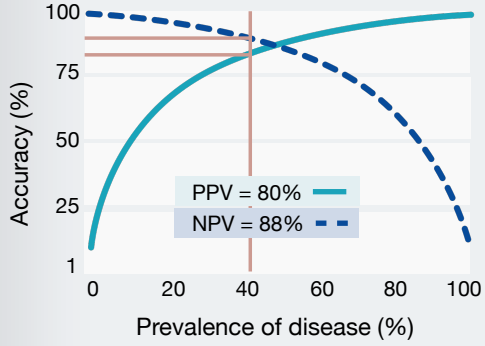
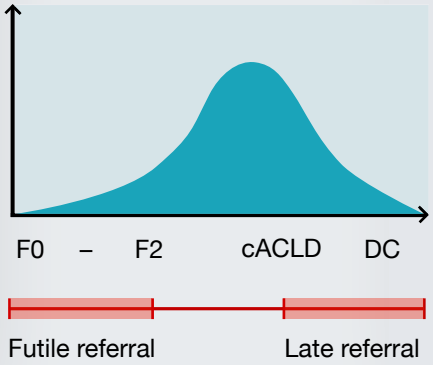
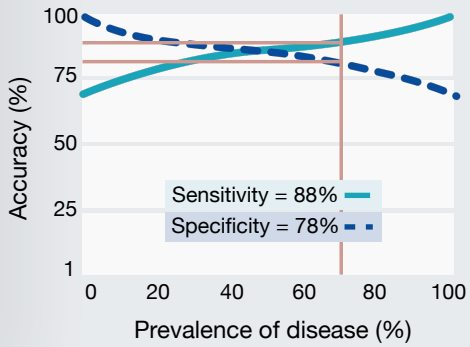
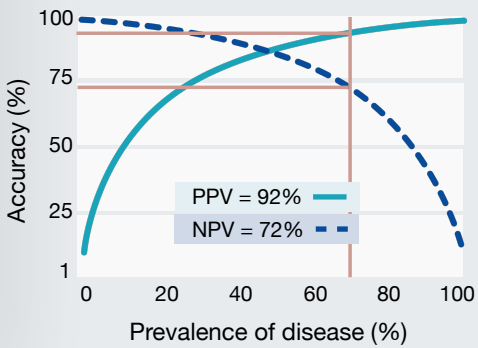
The immaturity of most omics technologies result in a shortage of (a) high-quality diagnostic studies, (b) independent validation of novel biomarkers, (c) established cut-offs for clinical decision making, (d) analytical standardisation. GWAS, Genome-Wide Association Studies; PCR, Polymerase Chain Reaction; mRNA, messenger RNA; MS, Mass-Spectrometry; FDA, U.S. Food and Drug Administration; NMR, Nuclear Magnetic Resonance; miRNA, microRNA; exRNA, extracellular RNA; NASH, Non-Alcoholic Steatohepatitis.

Figure 4. Population-based versus personalized omics biomarkers, promises and challenges.

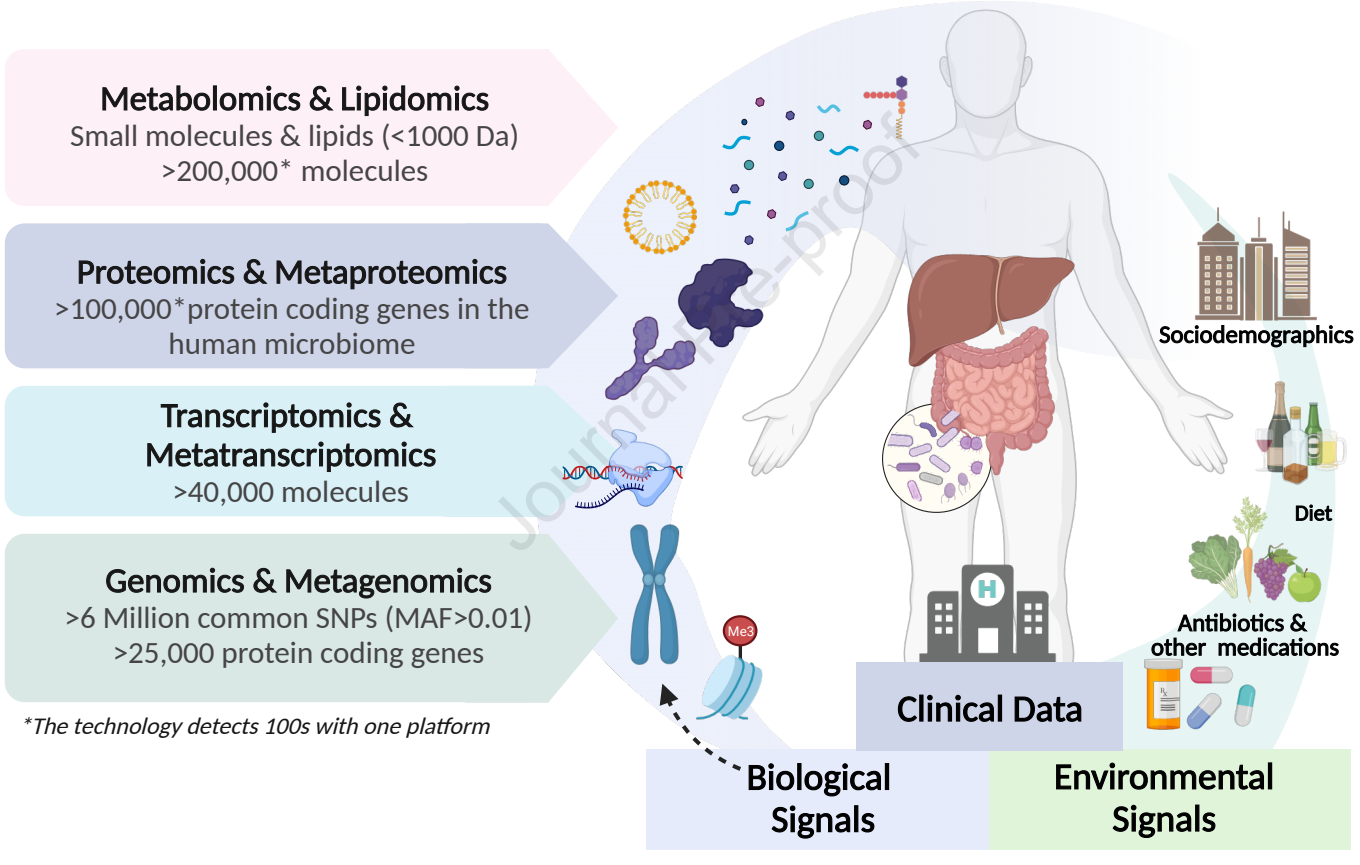
miR, microRNA. Examples are based on references (64, 117, 147-150)

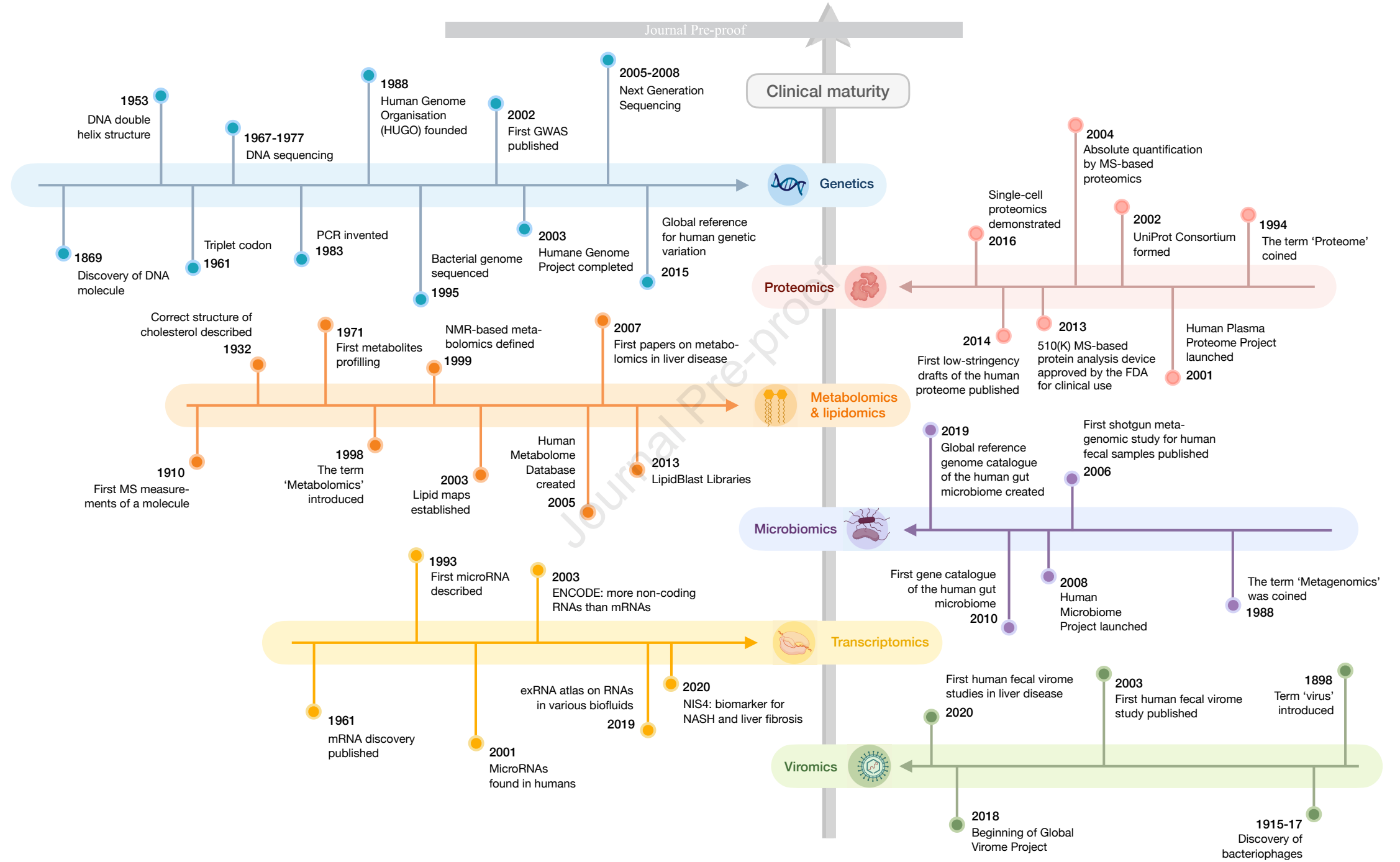
Figure 5. Regulatory pathways of three commercial biomarkers

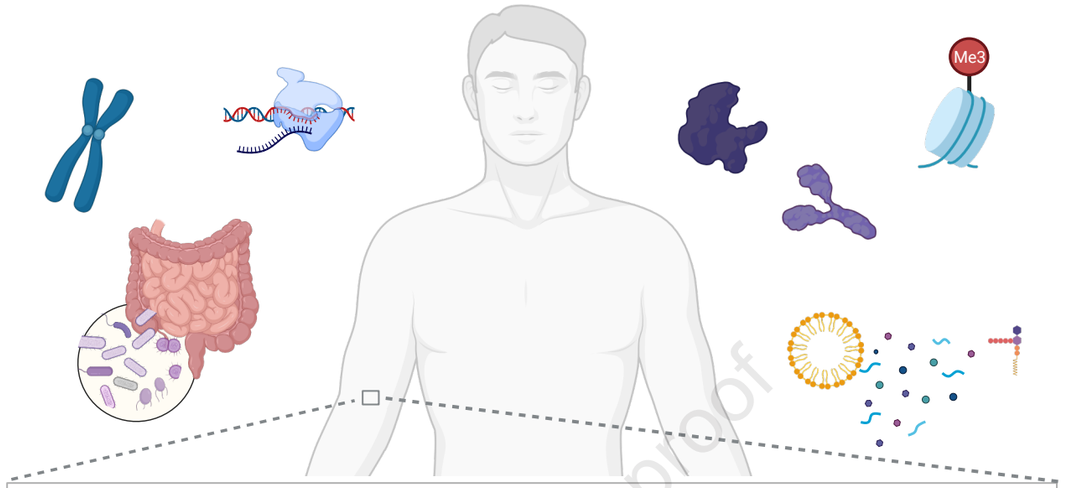
Illustrating the regulatory timeline of *nordic*PRO-C3^{TM}, Enhanced Liver Fibrosis test (ELF) and FibroScan. Eash timeline shows significant publications and regulatory milestones. Please refer to the supplementary materials for specific publications and milestones.

	Intended use of biomarkers	Patient characteristics	Variability in sensitivity and specificity	Spectrum effect: Variability in positive and negative predictive value
Citizens at risk of liver disease	Diagnostic: Go to general practitioner? Monitoring: Liver health tracker			
General practice	Diagnostic: Refer to secondary care? Prognostic: Intervene in general practice to improve prognosis? Monitoring: Monitor effect of lifestyle intervention			
Secondary care	Diagnostic: Diagnose steatohepatitis, Diagnose cACLD Prognostic: Predict development of liver-related events Monitoring: Monitor change in disease severity, Measure efficacy of intervention Prediction: Predict expected effect of treatment			

THE OMICS POTENTIAL IN BIOMARKER DEVELOPMENT







Population

Promise: Population health through the perspective of environmental exposure, nutrition and lifestyle. **Challenge**: Lacks specificity and a complete understanding of a healthy (liver) status from the omics perspective.

Example 1. miR-34a is part of the NIS2+ score, used to diagnose steatohepatitis in at-risk patients. **Example 2.** Glucose is commonly used to diagnose diabetes and determine treatment. **Example 3.** Plasma alanine aminotransferease and aspartate aminotransferase is used individually and as a ratio, in the general practice to indicate presence of liver damage.

Precision

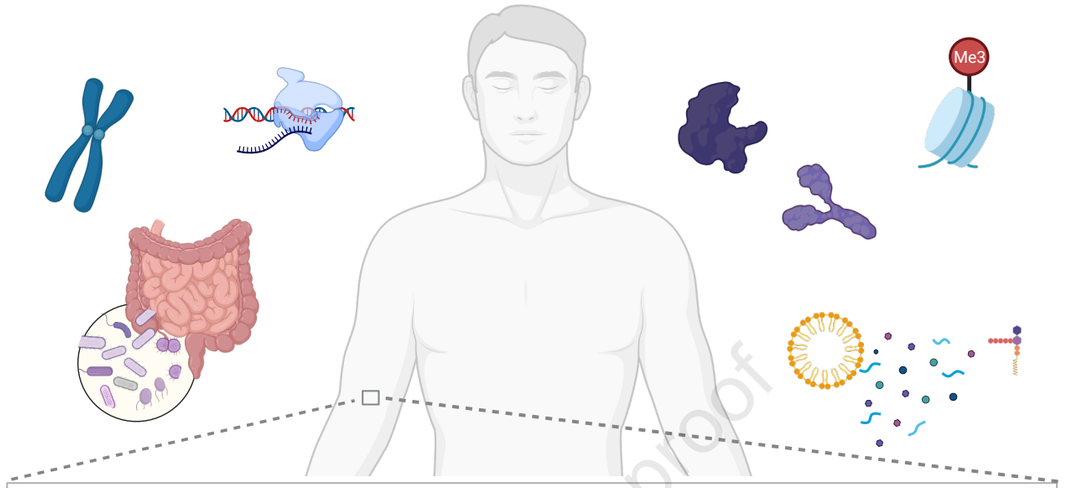
Promise: Stratification of patients to improve outcome of treatment and reduce side effects. Monitoring of disease development. **Challenge:** Translational omics research in the clinics is still in its infancy; lack of bench-to-bedside investigations.

Example 1. Adding the genetic risk polymorphisms: PNPLA3, TM6SF6, GCKR, and MBOAT7 to known metabolic traits aids prediction of outcome. **Example 2.** Branched-Chain Amino Acids, diacylglycerol, triglyceride, phosphatidylcholines, phosphatidylethanolamine, sphingomyelin levels differentiates clinical clusters of people with type 2 diabetes **Example 3**. Distinct patterns of lipid depletion can be measured in circulation and are found to associate with progressive alcohol-related liver fibrosis.

Personalized

Promise: Response-guided therapy and medication on the individual basis. **Challenge:** Individual and daily variations may lead to significant noise level.

Example 1. Levels of ceramides are found to link genetic predisposition and dietary habits to cardiometabolic disease risk. **Example 2.** Genetic polymorphy of HSD17β13 as a predictive biomarker for the effect of future treatments in liver disease. **Example 3.** Interleukin 28B gene on chromosome 19 as predictive for the extend of treatment needed in patients with hepatitis C infection.



Population

Promise: Population health through the perspective of environmental exposure, nutrition and lifestyle. **Challenge**: Lacks specificity and a complete understanding of a healthy (liver) status from the omics perspective.

Example 1. miR-34a is part of the NIS2+ score, used to diagnose steatohepatitis in at-risk patients. **Example 2.** Glucose is commonly used to diagnose diabetes and determine treatment. **Example 3.** Plasma alanine aminotransferease and aspartate aminotransferase is used individually and as a ratio, in the general practice to indicate presence of liver damage.

Precision

Promise: Stratification of patients to improve outcome of treatment and reduce side effects. Monitoring of disease development. **Challenge:** Translational omics research in the clinics is still in its infancy; lack of bench-to-bedside investigations.

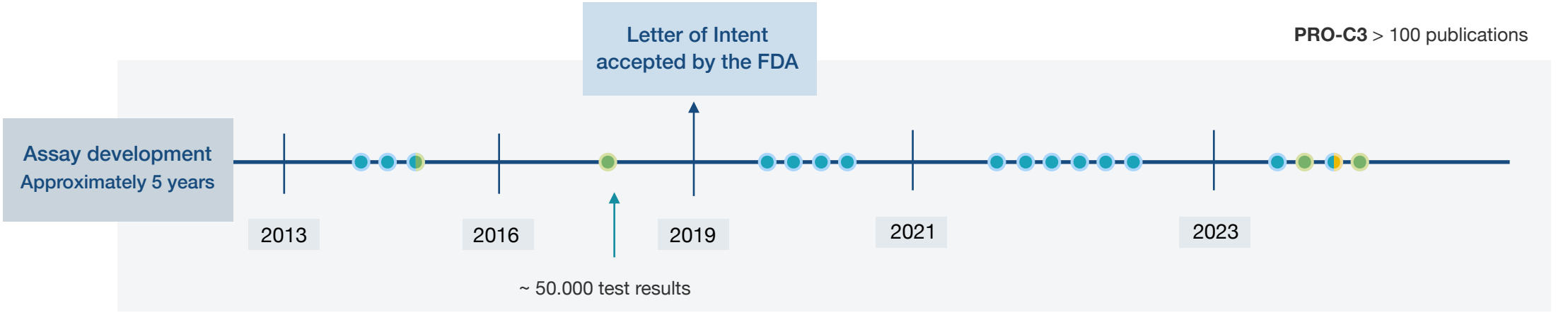
Example 1. Adding the genetic risk polymorphisms: PNPLA3, TM6SF6, GCKR, and MBOAT7 to known metabolic traits aids prediction of outcome. **Example 2.** Branched-Chain Amino Acids, diacylglycerol, triglyceride, phosphatidylcholines, phosphatidylethanolamine, sphingomyelin levels differentiates clinical clusters of people with type 2 diabetes **Example 3**. Distinct patterns of lipid depletion can be measured in circulation and are found to associate with progressive alcohol-related liver fibrosis.

Personalized

Promise: Response-guided therapy and medication on the individual basis. **Challenge:** Individual and daily variations may lead to significant noise level.

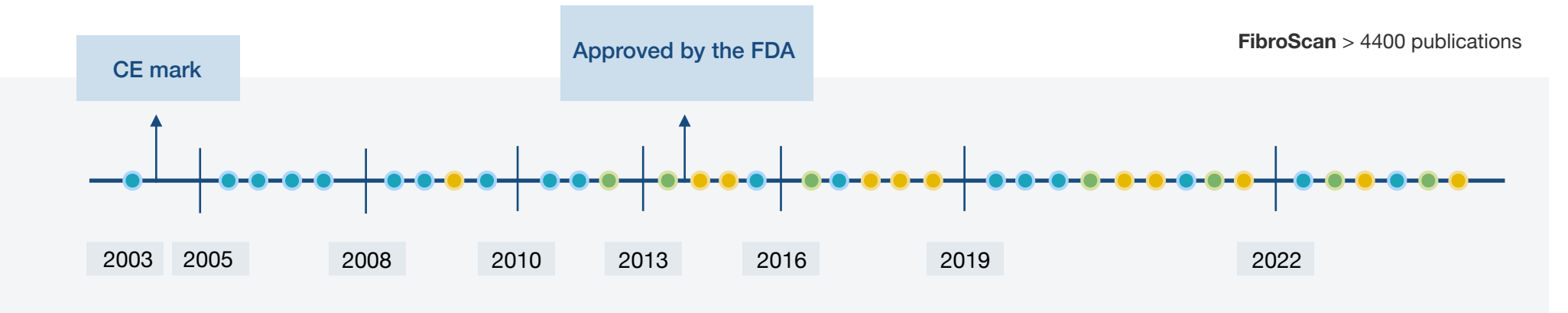
Example 1. Levels of ceramides are found to link genetic predisposition and dietary habits to cardiometabolic disease risk. **Example 2 (from oncology).** BRCA1 gene mutations are used in ovarian and breast cancers to determine treatment. **Example 3 (from oncology).** BCR-ABL fusion gene is used in leukemia to determine treatment and predict response to targeted therapy.

A) **PRO-C3**



B) **Enhanced Liver Fibrosis (ELF) test**

C) **FibroScan**



Diagnosis
Prognosis
Monitoring