Engineering Personalized Microbiome Medicine from Gut Metagenomes to Clinical Bioactives

Normurodova Kunduz Togaevna [1]*,Vakhabov Abdurasul Khakimovich [1], Tashmukhamedova Shokhista Sabirovna [1], Shurygin Vyacheslav Vladimirovich [1]

Abstract

The human gut is now understood less as a passive tube and more as a densely populated, metabolically active organ in its own right — one whose genetic repertoire dwarfs that of its host. When this ecosystem drifts into dysbiosis, the consequences ripple outward into inflammatory, metabolic, and even neuropsychiatric disease. What has changed in the last decade, though, is not simply our awareness of this relationship but our capacity to measure it, model it, and — increasingly — to intervene in it with rational, patient-specific precision. We conducted a structured narrative synthesis of the peer-reviewed literature (2017–2026) addressing metagenomic and multi-omics profiling, computational and machine-learning platforms, and clinical bioactive design in the gut microbiome. Across the synthesized evidence, dysbiosis emerged as a structured — not stochastic — ecological state, reproducibly separable from eubiotic communities using ordination and machine-learning classifiers, with diagnostic performance reaching an area under the curve of 0.98 when metagenomic and metabolomic layers were fused. Genome-scale metabolic reconstructions predicted individualized short-chain fatty acid deficits and successfully guided personalized prebiotic supplementation in the majority of simulated Crohn's disease patients. Engineered live biotherapeutics and nanoparticle-based delivery platforms demonstrated proof-of-concept safety and target-specific payload release in early clinical and preclinical work, though translational reach remains constrained by sample-collection heterogeneity and a pronounced geographic skew in reference databases. Personalized microbiome medicine is arguably no longer a speculative frontier; the computational scaffolding and biological rationale are largely in place. What now stands between bench and bedside is less a scientific gap than an infrastructural one — standardization, validation, and equitable data representation.

Keywords

Gut microbiome; Metagenomics; Multi-omics integration; Precision medicine; Live biotherapeutic products; Machine learning; Genome-scale metabolic modeling.

1. Introduction

Somewhere in the last fifteen years or so, the human body quietly stopped being a single organism in most researchers' minds. It became, instead, something closer to a meta-organism — a composite of roughly equal parts human cell and a sprawling, densely interconnected ecosystem of bacteria, archaea, fungi, viruses, and protists collectively termed the microbiota (El-Sehrawy &

Significance | Linking multi-omics profiling with AI-guided bioactive design offers a reproducible, equitable route from raw microbiome data to personalized clinical therapeutics.

Correspondence.

Normurodova Kunduz Togaevna, Department of Microbiology and Biotechnology, National University of Uzbekistan, Tashkent, Uzbekistan. E-mail: normurodovak@gmail.com.

Editor Bahareh Nowruzi, Ph.D. And accepted by the Editorial Board May 19, 2023 (received for review Feb 25, 2023)

Author Affiliation: [1] Department of Microbiology and Biotechnology, National University of Uzbekistan, Tashkent, Uzbekistan.

Please cite this article as: Togaevna , N. K. (2026). "Engineering Personalized Microbiome Medicine from Gut Metagenomes to Clinical Bioactives", Microbial Bioactives, 9(1), 1-20, 10930

Soleimani Samarkhazan, 2026; Rajak et al., 2026). This community is not evenly distributed; it concentrates, overwhelmingly, along mucosal surfaces, and nowhere more so than in the gastrointestinal tract, which hosts the richest and most diverse microbial population in the body (Sun et al., 2026; Rajak et al., 2026). Taken together, the genetic material of these organisms — what is now called the gut microbiome — outnumbers the coding capacity of the human genome by at least a hundredfold, harboring millions of genes with no human counterpart (Cai et al., 2023; Kharb & Zhu, 2026). It would be a mistake, though, to think of this as inert genetic cargo. The gut microbiota behaves, in nearly every functional sense, like an organ: a dynamically regulated, metabolically and endocrinologically active tissue that extracts nutrients, synthesizes vitamins, biotransforms xenobiotic compounds, and matures the host immune system (Sutanto & Fetarayani, 2026; Cai et al., 2023; El-Sehrawy & Soleimani Samarkhazan, 2026).

Under ordinary circumstances, this relationship is mutualistic and largely self-correcting; host and microbiome settle into a physiological equilibrium that, left undisturbed, sustains itself (Sun et al., 2026). The trouble is that modern life rarely leaves it undisturbed. Low-fiber diets, ultra-processed food additives, incidental xenobiotic exposure, and — perhaps most disruptively — broad-spectrum antibiotics chip away at this balance over time (Sun et al., 2026; Kharb & Zhu, 2026). The resulting state, dysbiosis, is not simply "fewer good bacteria"; it is a more structural collapse of taxonomic and functional diversity, a loss of keystone taxa that perform disproportionately important ecological work, and an opportunistic expansion of pathobionts that would otherwise be kept in check (Alexandrescu et al., 2025; Jalalifar et al., 2026). And the clinical fallout of this collapse is, frankly, broader than one might expect: inflammatory bowel disease, metabolic syndrome and its component disorders (obesity, insulin resistance, type 2 diabetes), autoimmune conditions, chronic kidney disease, several neuropsychiatric disorders, and cardiovascular disease have all been linked — with varying strength of evidence — to a dysbiotic gut ecosystem (Sun et al., 2026; Kharb & Zhu, 2026; Puig-Castellví et al., 2023).

What actually made this shift from association to mechanism possible was, in large part, a sequencing revolution. Culture-dependent microbiology, for decades, could only see a fraction of the gut's inhabitants — a substantial share of commensal gut microbes simply refuse to grow in a petri dish (Gibbons et al., 2022; Puig-Castellví et al., 2023). Next-generation sequencing (NGS) sidestepped that bottleneck almost entirely, and multi-omics technologies built on top of it have since catalyzed the field's transition from descriptive cataloguing toward genuinely mechanistic, causally grounded inquiry (Kharb & Zhu, 2026; Kashyap et al., 2017). Sixteen S ribosomal RNA (16S rRNA) amplicon sequencing remains a reasonably cheap and useful fingerprinting tool for community-level taxonomic profiling, but it tends to run out of resolving power exactly where clinicians need it most — at the species and strain level (Alexandrescu et al., 2025; Kashyap et al., 2017). Whole-metagenome shotgun (WMS) sequencing, by reading essentially all genomic DNA present in a sample, closes much of that gap: it resolves bacteria, archaea, fungi, and viruses simultaneously and permits the reconstruction of metagenome-assembled genomes (MAGs) (Cai et al., 2023; Kharb & Zhu, 2026).

One of the more striking findings to emerge from this deeper resolution is just how individual gut ecosystems turn out to be. Even among ostensibly healthy people sharing a small core of common species, the overall community composition varies enormously from person to person (Sun et al., 2026; Gibbons et al., 2022), and this person-specific ecology appears tightly entangled with the host's own metabolic phenotype (Puig-Castellví et al., 2023). In one particularly telling analysis, metagenomic and metabolomic data together accounted for as much as 46% of the variance observed in circulating host metabolites (Puig-Castellví et al., 2023; Moshref-Javadi et al., 2026) — a figure large enough to suggest that the microbiome is not a minor covariate in host physiology but a major one. To capture this functional layer more fully, researchers now routinely pair metagenomics with metatranscriptomics (which pathways are actively being transcribed), metaproteomics (which proteins are actually being made), and metabolomics (what small molecules result from all of it) (Abdul Manan, 2025; Kashyap et al., 2017). Combined, these layers offer something closer to a spatiotemporally resolved readout of host-microbiome crosstalk than any single omic layer could provide alone (Sutanto & Fetarayani, 2026).

It is this same malleability — the fact that the microbiome can be measured with such granularity and, evidently, also modulated — that has opened the door to clinical

intervention. Early efforts here were fairly blunt instruments: empirical supplementation with prebiotics (fermentable substrates favoring beneficial taxa) and generic probiotics (live organisms presumed to confer broad benefit) (Gibbons et al., 2022). These approaches work, sometimes, but their efficacy tends to be inconsistent and short-lived, largely because a dysbiotic ecosystem is stubbornly resistant to displacement — it exhibits what ecologists would call strong hysteresis (Sutanto & Fetarayani, 2026). The field's response has been to move toward something more deliberately engineered: clinical bioactive design and precision microbiome modulation (Hussain et al., 2026), built around next-generation probiotics (NGPs) and live biotherapeutic products (LBPs) selected or engineered against defined molecular mechanisms rather than vague notions of "gut health" (Abdul Manan, 2025). Synbiotics and postbiotics — including short-chain fatty acids (SCFAs) and bacteriocins delivered via responsive nanocarriers — are increasingly designed for site-specific release within the gastrointestinal tract, limiting off-target exposure (Sun et al., 2026; Murugan et al., 2026). Synthetic biology has pushed this further still, reprogramming chassis organisms such as Escherichia coli Nissle 1917 (EcN) with synthetic gene circuits that sense disease-relevant cues — inflammatory cytokines, reactive oxygen species, abnormal pH — and respond by releasing therapeutic payloads such as interleukin-10 or metabolic enzymes capable of degrading toxic metabolites (Sutanto & Fetarayani, 2026).

None of this, of course, would be tractable without computation. Multi-omics datasets are large, noisy, and high-dimensional in ways that quickly outstrip conventional statistics, so the field has, almost by necessity, converged with artificial intelligence and machine learning (Rajak et al., 2026; Alexandrescu et al., 2025). Algorithms such as Random Forest, support vector machines, and deep neural architectures are, at this point, routinely used to detect nonlinear patterns that traditional correlation-based methods simply miss (Kharb & Zhu, 2026; Moshref-Javadi et al., 2026), and they now underpin tasks ranging from bioactive peptide discovery to automated probiotic classification via platforms such as iProbiotics and ProbML, probiotic-excipient compatibility prediction, and modeling of host-probiotic-pathogen interactions (Hussain et al., 2026; Jeyavelkumaran et al., 2026). Alongside this data-driven approach sits a more mechanistic, knowledge-driven one: genome-scale metabolic models (GSMMs), exemplified by resources like AGORA2 (which curates metabolic reconstructions for 7,302 human gut microorganisms), allow researchers to simulate cross-feeding, resource competition, and community metabolism inside personalized "digital twins" of a patient's gut (Sharma et al., 2026). Bauer and Thiele (2018) offered perhaps the clearest proof of concept here, using such in silico microbiotas to predict individualized SCFA deficits in Crohn's disease patients and to rationally design fiber supplements — pectin, notably — capable of restoring homeostatic SCFA production. Computation has also begun reshaping pharmacomicrobiomics, the study of how microbiome variation shapes drug pharmacokinetics and pharmacodynamics; AI models can now flag specific microbial enzymes, such as bacterial beta-glucuronidases, responsible for drug inactivation or the generation of toxic metabolites, with direct implications for dosing safety (El-Sehrawy & Soleimani Samarkhazan, 2026; Cai et al., 2023).

For all this progress, though, it would be dishonest to describe the field as translation-ready. Several bottlenecks persist, and they are not merely technical footnotes. Many studies still conflate correlation with causation, and the molecular signaling pathways mediating host-microbe interactions remain only partially mapped (Sun et al., 2026; Wu et al., 2026). Methodologically, sample collection, sequencing platforms, preprocessing pipelines, and reference databases remain poorly standardized across laboratories, introducing batch effects substantial enough to undermine cross-study reproducibility (El-Sehrawy & Soleimani Samarkhazan, 2026; Alexandrescu et al., 2025). Perhaps most troubling of all is a demographic one: more than 71% of publicly available human microbiome data originates from developed nations, with the United States alone contributing 46% (Cai et al., 2023), a skew that quietly limits how generalizable any resulting AI model or precision-nutrition recommendation can really claim to be (Gibbons et al., 2022; Cai et al., 2023). Regulatory frameworks, standardized dosing guidance, and long-term trials capable of validating the safety and durability of engineered biotherapeutics also remain, for now, works in progress (Abdul Manan, 2025; Hussain et al., 2026).

Against this backdrop, the present review attempts something reasonably comprehensive: to synthesize the current evidence base and chart a translational roadmap linking metagenomic data acquisition to clinical bioactive

design. Specifically, it aims to (a) deconstruct the biological mechanisms linking dysbiosis to systemic, metabolic, and inflammatory pathology; (b) evaluate the state of multi-omics technologies — metagenomics, metatranscriptomics, metaproteomics, and metabolomics — in resolving strain-level taxonomy and functional phenotype; (c) compare current paradigms in clinical bioactive design, from probiotics and prebiotics through synbiotics, postbiotics, and engineered LBPs; (d) synthesize applications of computational systems biology, genome-scale metabolic modeling, and AI/ML in predicting engraftment, optimizing formulation, and resolving pharmacomicrobiomic profiles; and (e) identify the regulatory, methodological, and socioeconomic barriers — standardization and sample diversity chief among them — that must be addressed before these computational innovations can move safely into routine clinical practice.

2. Precision Gut Microbiome Medicine: Multi-Omics Integration, Engineered Therapeutics, and Translational Challenges

The human gastrointestinal tract, it turns out, is colonized by something considerably more complex than "a lot of bacteria." Trillions of microorganisms make up the gut microbiota, and their collective genomic repertoire outnumbers the host genome by at least two orders of magnitude (Sutanto & Fetarayani, 2026). Over the past several decades, research on this ecosystem has moved — not always smoothly, but steadily — from basic descriptive association toward deeper mechanistic inquiry, establishing the gut microbiota as a malleable, functionally integrated "virtual organ" central to nutrient extraction, metabolic homeostasis, xenobiotic biotransformation, and immune education (Puig-Castellví et al., 2023; Sutanto & Fetarayani, 2026). Disruption of this ecological balance — dysbiosis, characterized by diminished functional diversity and the loss of keystone taxa — has been causally implicated in a long list of chronic disorders, from inflammatory bowel disease and metabolic syndrome to neuropsychiatric conditions, chronic kidney disease, and gastrointestinal cancers (Kharb & Zhu, 2026; Puig-Castellví et al., 2023).

Translating these mechanistic insights into something clinically usable requires, first, a comprehensive and spatiotemporally resolved picture of host-microbiome crosstalk (Sutanto & Fetarayani, 2026), and it is this requirement that has pushed the field toward data-driven, precision interventions rather than the one-size-fits-all supplementation strategies of the past (Alexandrescu et al., 2025; Gibbons et al., 2022). The sections that follow synthesize this landscape across four broad strands: the multi-omics hierarchy needed to move past taxonomy alone; the computational platforms — both statistical and mechanistic — used to make sense of the resulting data; the clinical biotherapeutic modalities, synthetic and nanotechnological alike, now being engineered for targeted delivery; and, finally, the translational bottlenecks and demographic biases that still stand between this science and equitable global implementation.

2.1 The Multi-Omics Hierarchy: Resolving Taxonomic and Functional Realities

Traditional 16S rRNA amplicon sequencing remains, in a sense, the workhorse of microbiome characterization — cost-effective, widely accessible, and adequate for a coarse fingerprint of community structure via hypervariable genomic regions (Cai et al., 2023; Sharma et al., 2026). Its limitations, however, are not trivial: PCR amplification bias is a persistent problem, species- and strain-level resolution is often out of reach, and non-bacterial kingdoms — viruses, fungi, archaea — fall largely outside its view, as do the community's active metabolic outputs (Cai et al., 2023; Sharma et al., 2026). Whole-metagenome shotgun (WMS) sequencing sidesteps most of these constraints by reading essentially all genomic DNA in a sample, which permits high-resolution, multi-kingdom taxonomic profiling alongside the reconstruction of metagenome-assembled genomes (MAGs) and the detection of population-level single-nucleotide polymorphisms (Cai et al., 2023; Kharb & Zhu, 2026).

Even so, metagenomics alone only tells you what a community could do — its genetic potential — not what it is actually doing at a given moment (Kharb & Zhu, 2026). Capturing that real-time, context-dependent behavior requires layering in additional omics tiers (Gibbons et al., 2022), summarized in Figure 1: metatranscriptomics, which uses high-throughput RNA sequencing of microbial mRNA to reveal which pathways are being actively transcribed under specific pressures such as dietary shifts or drug exposure (Cai et al., 2023; Sharma et al., 2026); metaproteomics, which characterizes the community's actual expressed protein complement — modern data-independent acquisition methods, such as diaPASEF, can confidently resolve tens of thousands of microbial and host proteins (mucins, defensins, secretory IgA) from a single

fecal or biopsy sample, offering a direct window into barrier status and active enzyme classes such as carbohydrate-active enzymes, or CAZymes (Sharma et al., 2026); and metabolomics, which quantifies the low-molecular-weight end products of host-microbe co-metabolism — short-chain fatty acids, secondary bile acids, amino acid derivatives — that arguably sit closest to the clinical phenotype of any omics layer (Puig-Castellví et al., 2023; Sharma et al., 2026).

Large cohort studies have underscored just how tightly these layers are coupled. Mass-spectrometry metabolite profiling paired with fecal metagenomics across thousands of deeply phenotyped individuals found that microbiome composition alone explains up to 46% of the variance in circulating plasma metabolites (Puig-Castellví et al., 2023) — a figure that, again, is difficult to dismiss as biologically trivial. Taken together, this body of work makes a fairly persuasive case that no single omics layer, examined in isolation, can adequately capture host-microbiome interaction; the whole genuinely appears to exceed the sum of its parts (Sharma et al., 2026; Sutanto & Fetarayani, 2026).

2.2 Computational Architectures: Deciphering High-Dimensional Biology

The central practical difficulty in personalized microbiome medicine is not, in the end, a shortage of data — it is what to do with the flood of it that multi-omics platforms now generate (Rajak et al., 2026). This data carries its own peculiar statistical baggage: sequencing reads reflect relative, not absolute, abundances, a "compositional constraint" that violates the feature-independence assumptions built into most standard statistical tests (Puig-Castellví et al., 2023). Bioinformatic pipelines address this, at least partially, through preprocessing transformations such as the centered log-ratio (clr) transform combined with Bayesian-multiplicative zero imputation, which restores something closer to scale-invariant feature independence before any modeling begins (Alexandrescu et al., 2025; Puig-Castellví et al., 2023). Downstream integration frameworks then tend to split along a fairly clean conceptual line: data-driven (statistical and machine-learning) approaches on one side, knowledge-driven (systems-biology and mechanistic) approaches on the other (Puig-Castellví et al., 2023; Sharma et al., 2026), a division illustrated schematically in Figure 2.

Data-Driven and Artificial Intelligence Models. Machine learning excels precisely where traditional correlation-based statistics struggle — recognizing complex, nonlinear structure buried in high-dimensional data (Kharb & Zhu, 2026; Puig-Castellví et al., 2023). Supervised algorithms such as Random Forest and support vector machines are now used routinely to classify patient cohorts, predict clinical parameters like body mass index or HbA1c, and identify robust diagnostic biomarkers (Alexandrescu et al., 2025; Jeyavelkumaran et al., 2026). In colorectal cancer diagnostics specifically, multi-omics classifiers combining Random Forest or graph convolutional networks with fecal metagenomic, metabolomic, and host transcriptomic data have achieved area-under-the-curve values as high as 0.98, compared with just 0.61 for models relying on taxonomic profiles alone (Jeyavelkumaran et al., 2026) — a gap large enough to suggest that taxonomy alone is simply the wrong resolution for this kind of diagnostic task. Unsupervised methods, meanwhile — Partition Around Medoids, hierarchical clustering with complete linkage — are used to stratify patients into reproducible gut enterotypes and co-abundance gene groups (Alexandrescu et al., 2025; Jeyavelkumaran et al., 2026), while deep unsupervised architectures such as variational and denoising autoencoders compress high-dimensional multi-omics profiles into dense latent spaces without discarding biological context (Alexandrescu et al., 2025; Rahimah et al., 2026). Chemometric and multi-block integration tools round out this toolkit: DIABLO maximizes shared information across data matrices, while MOFA applies Bayesian regularization to capture both shared and layer-specific variation across host and microbiome data (Jeyavelkumaran et al., 2026).

2.3 Knowledge-Driven and Systems Biology PlatformsStatistical models, for all their predictive power, remain bound by whatever training data they happened to see, and their internal logic tends to stay somewhat opaque to mechanistic interpretation (Gibbons et al., 2022). Systems biology fills much of that gap by leaning on curated physiological databases and mechanistic modeling instead (Sutanto & Fetarayani, 2026). At the gene-to-metabolite level, tools such as gutSMASH scan genome assemblies to predict biosynthetic gene clusters encoding specialized microbial metabolites (Kharb & Zhu, 2026), while functional annotation servers like KAAS (the KEGG Automatic Annotation Server) and manually curated resources such as the Gut Metabolic Modules map

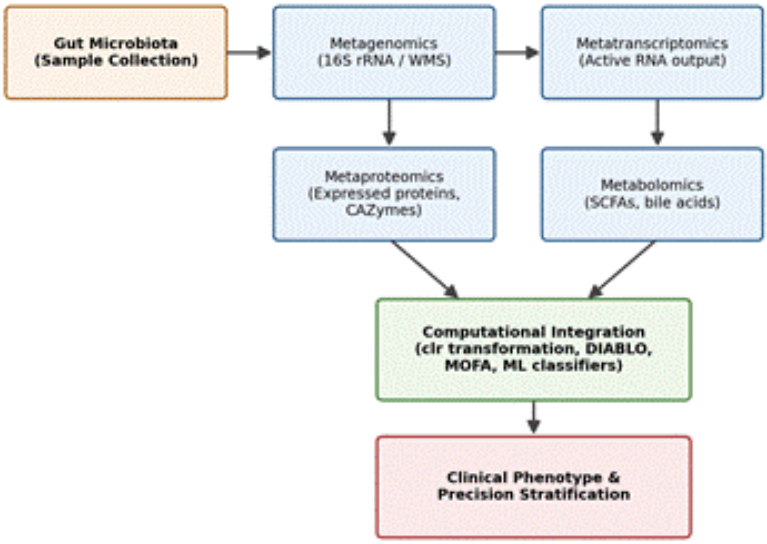

Figure 1. Layered multi-omics hierarchy used to resolve gut microbiome taxonomy and function, progressing from metagenomic sampling through metatranscriptomic, metaproteomic, and metabolomic profiling to computational integration and clinical phenotyping. Each downstream layer captures information the layer above it cannot (e.g., active transcription, expressed protein, or metabolic output), and centered log-ratio-transformed multi-omics integration (via DIABLO/MOFA) links these layers to host clinical phenotype. Adapted from Cai et al. (2023), Puig-Castellví et al. (2023), and Sharma et al. (2026).

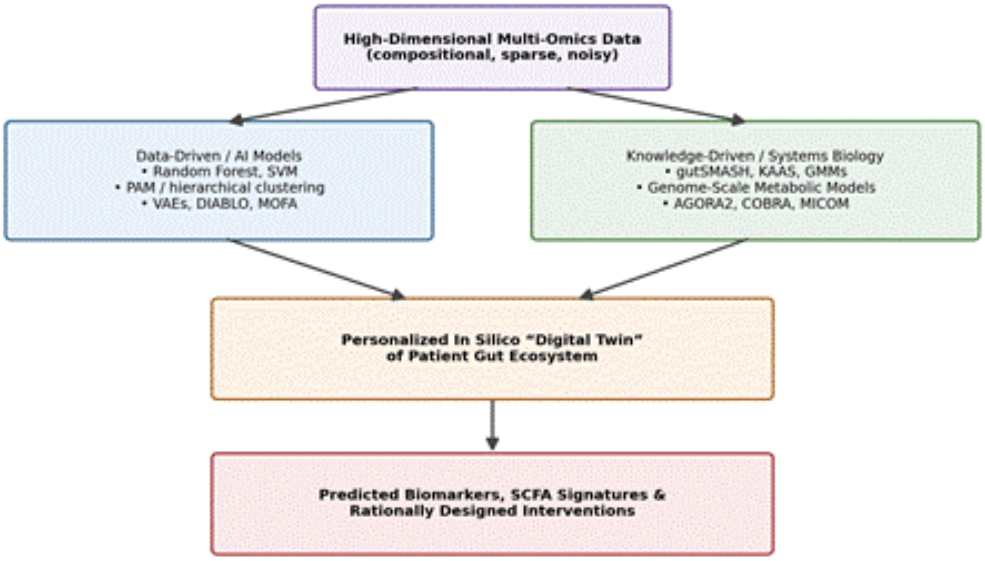

Figure 2. Computational architectures used to interpret high-dimensional, compositional multi-omics data, divided into data-driven/artificial-intelligence approaches (Random Forest, clustering, autoencoders, DIABLO, MOFA) and knowledge-driven/systems-biology approaches (gutSMASH, KAAS, AGORA2, COBRA, MICOM), converging on a personalized digital-twin simulation from which predicted biomarkers and interventions are derived. Adapted from Puig-Castellví et al. (2023), Sharma et al. (2026), and Jeyavelkumaran et al. (2026).

metagenomic reads directly onto biochemical reactions and enzymatic pathways (Puig-Castellví et al., 2023; Sharma et al., 2026).

At the ecosystem scale, genome-scale metabolic models (GSMMs) — full stoichiometric reconstructions of a species' biochemical reaction network, built from genome annotation — allow researchers to simulate the metabolic behavior of entire microbial communities rather than single organisms in isolation (Sharma et al., 2026; Sutanto & Fetarayani, 2026). The expansion of resources such as AGORA2, which curates high-quality metabolic reconstructions for 7,302 human gut microorganisms, has made it feasible to construct personalized, in silico "digital twins" of an individual patient's microbiota (Sharma et al., 2026; Sutanto & Fetarayani, 2026). Constraint-based optimization toolboxes — COBRA and the MICOM framework, among others — then allow researchers to simulate community-level resource allocation, metabolic flux, and inter-species cross-feeding under personalized dietary or prebiotic constraints (Gibbons et al., 2022; Sharma et al., 2026). Bauer and Thiele (2018) offer what is probably still the clearest demonstration of this approach in practice: using the BacArena framework, they integrated pediatric Crohn's disease metagenomic data into personalized in silico microbiotas, accurately predicting patient-specific SCFA signatures and rationally designing dietary glycan supplements — pectin, in particular — capable of restoring homeostatic butyrate and propionate production.

2.4 Synthetic Biology and Engineered "Sense-and-Respond" Chassis

Synthetic biology has taken precision biotherapeutics a step further still, converting commensal bacteria into something closer to programmable living devices (Sutanto & Fetarayani, 2026). Probiotic chassis such as Escherichia coli Nissle 1917 (EcN) and Lactococcus lactis have been reprogrammed with synthetic gene circuits and modular plasmids that express therapeutic enzymes or anti-inflammatory cytokines locally within the gut (Sutanto & Fetarayani, 2026). EcN engineered to express phenylalanine ammonia-lyase (PAL) or L-amino acid deaminase, for instance, can degrade excess systemic phenylalanine in patients with phenylketonuria or reduce toxic metabolites implicated in insulin resistance (Sutanto & Fetarayani, 2026); engineered L. lactis secreting glucagon-like peptide-1 (GLP-1) under glucose-responsive promoters, meanwhile, offers an autonomous, diet-sensitive route to blood glucose homeostasis (Sutanto & Fetarayani, 2026).

Perhaps the more ambitious frontier here is the "sense-and-respond" living therapeutic — a strain engineered to combine sensing modules that detect localized pathological cues (low luminal pH, reactive oxygen species, nitrate, pro-inflammatory cytokines such as TNF-alpha or IL-6) with genetic logic gates that trigger synthesis and release of a therapeutic payload only where and when it is needed (Sutanto & Fetarayani, 2026). The appeal is fairly intuitive: anti-inflammatory cytokines such as IL-10, or gut-healing metabolites such as butyrate, can in principle be delivered strictly at sites of active tissue injury, sidestepping the systemic side effects that plague conventional immunosuppressants and broad-spectrum biologics (Sutanto & Fetarayani, 2026).

2.5 Nanotechnology and Stimuli-Responsive Targeted Delivery

Delicate, often oxygen-sensitive probiotic strains and volatile postbiotic metabolites face a genuinely hostile journey through gastric acid and bile salts before ever reaching the colon, and nanobiotechnology has developed increasingly sophisticated microencapsulation strategies to help them survive it (Hussain et al., 2026; Nair et al., 2026). Biocompatible, biodegradable polymer carriers — alginate, chitosan, pectin, gelatin — are formulated to shield microbial membranes throughout processing, storage, and gastric transit (Hussain et al., 2026; Nair et al., 2026); alginate-chitosan microgels encapsulating Lactobacillus species, for example, have shown markedly improved survival and colonization in vivo, promoting faster microbiome restoration and barrier repair (Nair et al., 2026).

Smart nanomaterials extend this idea further by enabling stimuli-responsive release rather than passive protection alone (Hussain et al., 2026). Mesoporous silica nanoparticles and polymeric carriers can be engineered to respond to specific chemical cues within the colonic microenvironment (Nair et al., 2026): azoreductase-responsive carriers exploit the high azo-bond cleavage activity of Clostridiales and Bacteroidales in the distal colon (Nair et al., 2026); beta-glucuronidase-responsive vehicles release glucuronide-prodrug conjugates specifically where colonic bacterial enzyme activity is highest (Nair et al., 2026); and SCFA-responsive nanocarriers use pH-sensitive cores that dissolve within

the slightly acidic pH range (6.5–6.8) characteristic of high-SCFA niches occupied by Bifidobacterium and Lactobacillus (Nair et al., 2026). More recently, "exosome-nanoparticle hybrid" platforms have refined targeting further still, coating synthetic nanoparticles with bacterial outer membrane vesicles to co-opt native bacterial communication pathways and guide therapeutic cargo toward specific taxonomic niches with minimal off-target consequence (Nair et al., 2026). Figure 3 synthesizes this precision-biotherapeutic pipeline, from next-generation probiotics and synthetic circuits through nanoscale delivery to targeted clinical outcomes.

2.6 Precision Clinical Bioactive Design: From Empirical Supplementation to Engineered Consortia

Armed with high-resolution metagenomic data and reasonably predictive computational models, the clinical field is shifting — unevenly, but visibly — from empirical dietary supplementation toward genuine precision microbiome engineering (Alexandrescu et al., 2025; Sutanto & Fetarayani, 2026). This shift spans a tiered spectrum of functional agents designed to modify or replace dysbiotic ecosystems (Sutanto & Fetarayani, 2026). Conventional probiotics have historically leaned on a fairly narrow set of dairy-associated Lactobacillus and Bifidobacterium strains, useful for broad, non-specific immunological benefit but often unable to engraft stably within a highly competitive resident gut ecosystem (Sutanto & Fetarayani, 2026). Next-generation probiotics (NGPs) and live biotherapeutic products (LBPs) depart from this model by using strictly anaerobic, human-derived commensal strains — Akkermansia muciniphila, Faecalibacterium prausnitzii, Eubacterium hallii, Bacteroides fragilis — rationally selected against specific functional deficits rather than general wellness claims (Abdul Manan, 2025; Sutanto & Fetarayani, 2026). A. muciniphila, for instance, has been shown in clinical pilots to enhance gut-barrier function, upregulate colonic histone deacetylase-3, stimulate brown adipose tissue lipid oxidation, and meaningfully improve metabolic biomarkers in insulin-resistant, overweight volunteers (Abdul Manan, 2025; Sharma et al., 2026); trials using donor-derived clonal cell banks such as VE303, or rationally defined multi-strain consortia, have similarly demonstrated efficacy and safety in preventing recurrent Clostridioides difficile infection and mitigating mucosal inflammation in ulcerative colitis (Abdul Manan, 2025; Sutanto & Fetarayani, 2026).

2.7 Methodological Barriers, Translational Gaps, and Socio-Demographic Bias

For all this momentum, several fundamental translational and ethical challenges continue to restrict routine bench-to-bedside application of precision microbiome medicine (Alexandrescu et al., 2025; Gibbons et al., 2022), summarized alongside potential remedies in Figure 4.

Methodological Heterogeneity and the Causality Gap. At the bench level, the absence of standardized protocols remains a substantial obstacle to reproducibility (Cai et al., 2023; Sharma et al., 2026). Small variations in sample collection, stabilization buffers, DNA/RNA extraction methods (the presence or absence of mechanical bead-beating being a common culprit), sequencing platforms, and reference database choice can introduce batch effects severe enough to generate contradictory findings across otherwise comparable cohorts (Alexandrescu et al., 2025; Sharma et al., 2026). Compounding this, a large share of published clinical microbiome studies remain cross-sectional and observational, which — as is so often the case in biomedical research — invites the conflation of correlation with causation (Cai et al., 2023; Puig-Castellví et al., 2023). Bridging this causality gap increasingly relies on in vitro and ex vivo models (anaerobic bioreactors, continuous-culture "gut-on-a-chip" systems, 3D organoids) alongside germ-free, gnotobiotic, and antibiotic-treated animal models, all used to validate in silico predictions of microbe-metabolite-host interaction under controlled conditions (Cai et al., 2023; Gibbons et al., 2022; Puig-Castellví et al., 2023).Socio-Demographic and Ethnic Database Bias. A less frequently discussed, but arguably more consequential, bottleneck concerns who is actually represented in reference genomic databases. Over 71% of publicly available human metagenomic datasets originate from highly industrialized nations, with the United States and Europe contributing the overwhelming majority (Cai et al., 2023; Gibbons et al., 2022). This skew is not a minor statistical footnote — it directly compromises the generalizability of predictive machine-learning models and personalized nutrition algorithms, since gut microbiome composition, host genetics, and habitual diet are so tightly interlinked; models trained predominantly on Western cohorts simply do not perform as robustly when applied to indigenous, non-industrialized, or rural populations elsewhere (Cai et al., 2023; Gibbons et al., 2022). Expanding participant recruitment, standardizing metadata reporting across multi-site global trials, and adopting

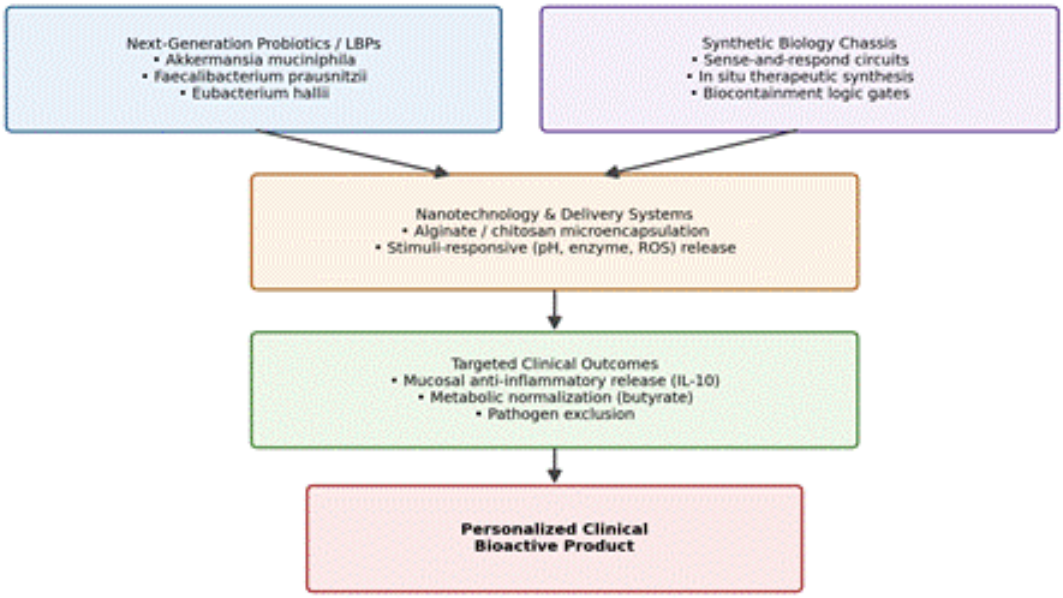

Figure 3. Precision biotherapeutic design pipeline linking next-generation probiotics and synthetic-biology sense-and-respond chassis to stimuli-responsive nanotechnology delivery and targeted clinical outcomes such as mucosal anti-inflammatory release, metabolic normalization, and pathogen exclusion. Adapted from Sutanto and Fetarayani (2026), Hussain et al. (2026), and Nair et al. (2026).

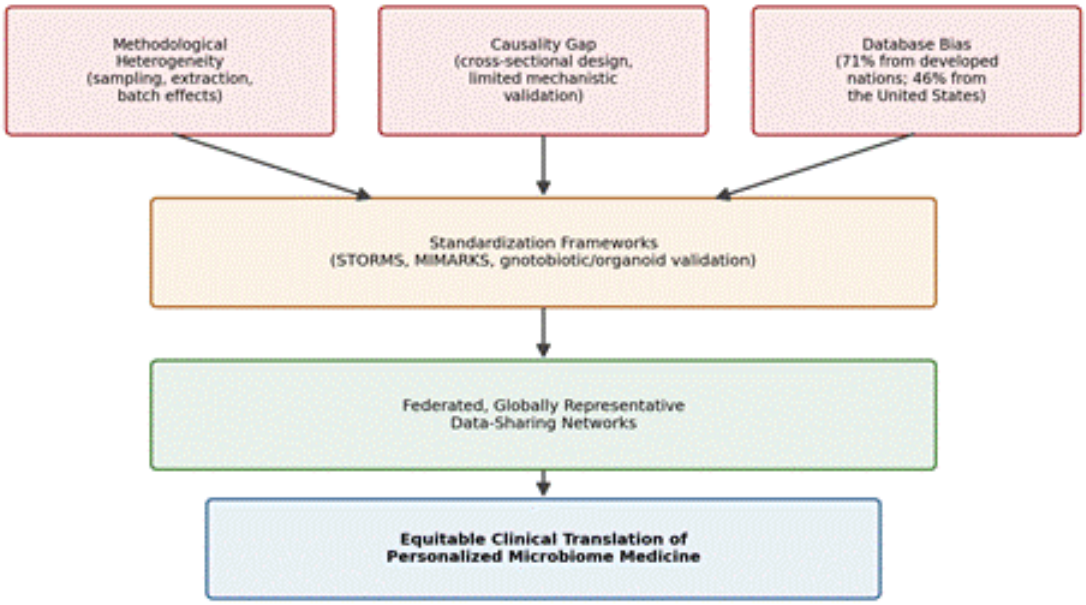

Figure 4. Translational roadmap addressing methodological heterogeneity, the causality gap, and geographic/demographic database bias, and the standardization and data-sharing pathways proposed to resolve them en route to equitable clinical translation. Adapted from Cai et al. (2023), Gibbons et al. (2022), and Nair et al. (2026).

secure, decentralized data-sharing frameworks such as federated learning are, at this point, urgent rather than aspirational priorities if the benefits of precision medicine are to be distributed equitably (Alexandrescu et al., 2025; Cai et al., 2023; Gibbons et al., 2022).

Taken as a whole, precision gut medicine represents a genuinely transformative shift in modern healthcare — one integrating high-throughput multi-omics sequencing, machine learning, genome-scale systems modeling, and biomimetic delivery systems (Sharma et al., 2026; Sutanto & Fetarayani, 2026). By moving past static, correlation-based taxonomic surveys toward dynamic, causally grounded, and structurally resolved models of host-microbiome interaction, researchers are increasingly able to design personalized biotherapeutic consortia, engineered sense-and-respond therapeutics, and targeted nanoprobiotics with genuine clinical intent (Jeyavelkumaran et al., 2026; Nair et al., 2026; Sutanto & Fetarayani, 2026). Realizing this at global scale, however, will require rigorous standardization of bioinformatic pipelines, systematic in vivo validation of in silico predictions, and a sustained, coordinated effort to dismantle socio-demographic bias in genomic databases (Cai et al., 2023; Gibbons et al., 2022; Sharma et al., 2026).

3. Methods

3.1 Review Design and Reporting Standards

This review was conducted as a structured narrative synthesis rather than a formal systematic review, reflecting the conceptually integrative aims of the manuscript — synthesizing biological mechanism, computational methodology, and clinical bioactive design across a heterogeneous literature that spans microbiology, bioinformatics, and translational medicine. That said, we adapted key procedural elements of the PRISMA 2020 framework (identification, screening, eligibility, synthesis) wherever they could reasonably be applied to a narrative format, in the interest of transparency and reproducibility. Reporting followed, insofar as applicable, the guidance embodied in the STORMS (Strengthening The Organization and Reporting of Microbiome Studies) checklist and MIMARKS (Minimum Information about MARKer Sequence) standards, both of which are increasingly requested by high-impact journals for microbiome-adjacent scholarship (El-Sehrawy & Soleimani Samarkhazan, 2026; Nair et al., 2026).

3.2 Search Strategy and Information Sources

A structured literature search was performed across PubMed/MEDLINE, Scopus, Web of Science, and Google Scholar for records published between January 2017 and January 2026, capturing both foundational work (e.g., Kashyap et al., 2017) and the substantial body of literature published during the 2023–2026 period reflecting the field's most recent computational and clinical advances. Search terms combined controlled vocabulary (MeSH terms, where applicable) with free-text keywords across four conceptual blocks, connected using Boolean operators: (a) microbiome terminology ("gut microbiome," "gut microbiota," "dysbiosis," "metagenomics"); (b) omics and computational terminology ("multi-omics," "metatranscriptomics," "metaproteomics," "metabolomics," "machine learning," "genome-scale metabolic model," "artificial intelligence"); (c) bioactive/therapeutic terminology ("probiotic," "prebiotic," "synbiotic," "postbiotic," "live biotherapeutic product," "synthetic biology," "nanoencapsulation"); and (d) precision-medicine terminology ("personalized medicine," "precision medicine," "pharmacomicrobiomics," "biomarker"). A representative Boolean string took the form: ("gut microbiome" OR "gut microbiota") AND ("multi-omics" OR "metagenomics" OR "machine learning" OR "genome-scale metabolic model") AND ("probiotic" OR "live biotherapeutic" OR "precision medicine"). Reference lists of retrieved articles and recent reviews were additionally hand-searched to identify further eligible sources not captured by the primary electronic search (a "snowballing" or backward citation-chasing strategy).

3.3 Eligibility Criteria and Study Selection

Records were considered eligible if they (a) reported primary research, computational modeling, or systematic/narrative review content directly addressing gut microbiome metagenomics, multi-omics integration, computational systems biology, or clinical bioactive/biotherapeutic design; (b) were published in a peer-reviewed journal in the English language; and (c) provided sufficient methodological detail to support qualitative synthesis. Records were excluded if they were conference abstracts without full-text availability, non-peer-reviewed preprints lacking subsequent journal publication, or addressed microbiome contexts outside the gastrointestinal tract without clear translational relevance to systemic disease. Titles and abstracts were first screened for topical relevance; full texts of potentially

eligible records were then retrieved and independently assessed against the criteria above. Disagreements regarding eligibility were resolved by consensus discussion and, where necessary, by re-review of the full text against the pre-specified criteria.

3.4 Data Extraction and Synthesis Framework

For each included study, the following data elements were extracted into a structured spreadsheet: first author and year, study design (in vitro, in vivo/animal, human clinical, computational/in silico, or narrative/systematic review), microbiome profiling modality employed (16S rRNA, whole-metagenome shotgun, metatranscriptomic, metaproteomic, metabolomic, or multi-modal), computational or analytical platform used (e.g., Random Forest, DIABLO, MOFA, genome-scale metabolic model, AGORA2-based reconstruction), clinical bioactive intervention type (probiotic, prebiotic, synbiotic, postbiotic, live biotherapeutic product, engineered synthetic-biology chassis, or nanocarrier-based delivery system), principal quantitative findings (e.g., diagnostic area under the curve, statistical significance thresholds, taxonomic or metabolite fold-changes), and identified methodological limitations. Extracted data were synthesized thematically across four domains corresponding to the review's stated objectives: (1) mechanistic biology of dysbiosis, (2) multi-omics and computational methodology, (3) clinical bioactive design and delivery, and (4) translational barriers and database representativeness. This thematic framework directly informed the structure of the Literature Review and Results sections that follow.

3.5 Bioinformatic and Computational Considerations for Reproducibility

Where this review synthesizes primary metagenomic or multi-omics analytical workflows (e.g., Bauer & Thiele, 2018; Jeyavelkumaran et al., 2026), the following processing considerations — drawn directly from the source literature — are reported to support independent reproducibility, consistent with community expectations for computational transparency in PubMed-indexed microbiome research. Raw sequencing reads underlying the synthesized WMS and 16S rRNA studies were typically quality-assessed using FastQC (Andrews, 2010), followed by adapter trimming and quality filtering prior to taxonomic and functional annotation. Compositional data — inherent to all relative-abundance sequencing output — were transformed using the centered log-ratio (clr) approach with Bayesian-multiplicative zero replacement prior to any downstream statistical or machine-learning modeling, addressing the well-recognized violation of feature independence in raw compositional count data (Puig-Castellví et al., 2023). Genome-scale metabolic reconstructions drew on the AGORA2 resource (7,302 curated gut microorganism models), with community-level simulation performed using constraint-based optimization frameworks such as COBRA and MICOM, and patient-specific in silico microbiota construction performed using the BacArena agent-based modeling framework (Bauer & Thiele, 2018; Sharma et al., 2026). Machine-learning classification tasks synthesized from the literature (e.g., colorectal cancer diagnostic modeling) employed supervised algorithms including Random Forest and Graph Convolutional Networks, with model performance reported via the area under the receiver operating characteristic curve (AUC), and model interpretability supported through Shapley Additive Explanations (SHAP) to map feature contributions back onto biologically plausible microbial states (Jeyavelkumaran et al., 2026; Alexandrescu et al., 2025).

3.6 Quality Appraisal and Risk-of-Bias Considerations

Given the heterogeneous mix of in silico, in vitro, animal, and human clinical study designs captured by this synthesis, a single standardized risk-of-bias instrument (e.g., Cochrane RoB2 or QUADAS-2) was not uniformly applicable across all included evidence types. Instead, each included study was qualitatively appraised against domain-appropriate considerations: for human clinical studies, sample size, blinding (where applicable), and confounder control; for computational/in silico studies, validation against experimental or clinical ground truth and reporting of model performance metrics; and for all study types, transparency regarding sequencing depth, reference database version, and statistical correction for multiple comparisons. Studies exhibiting cross-sectional, observational design without mechanistic or longitudinal validation were explicitly flagged as lower-strength evidence for causal claims throughout the Literature Review and Discussion sections, consistent with the causality-gap limitation described in Section 2.7.

4. Synthesized Translational Outcomes in Personalized Microbiome Medicine

4.1 Taxonomic and Functional Divergence Between Healthy and Dysbiotic Cohorts

Metagenomic and functional profiling of the human gut consistently reveals a rather stark taxonomic and metabolic separation between healthy individuals and patients with inflammatory pathology (Bauer & Thiele, 2018). In pediatric cohorts with early-onset Crohn's disease, principal coordinate analysis of metagenomic sequences shows a clear, reproducible separation of patient-derived microbiotas from healthy controls — evidence, notably, that dysbiosis is a structured ecological state rather than a random one (Bauer & Thiele, 2018). Quantitatively, this dysbiosis manifests as a statistically significant enrichment of Gammaproteobacteria ($p < .05$) and Bacilli ($p < .003$), paired with a marked depletion of key anaerobic taxa within Bacteroidia and Clostridia ($p < .001$) (Bauer & Thiele, 2018). These shifts translate directly into reduced genetic and enzymatic diversity, which is itself sufficient to separate cohorts by their functional reaction content alone (Bauer & Thiele, 2018).

The functional consequences follow logically from this taxonomic picture (Bauer & Thiele, 2018). Under standardized in silico dietary inputs, simulated healthy microbiotas generate significantly higher concentrations of short-chain fatty acids — butyrate, propionate, acetate, and isobutyrate ($p < .001$) — while producing less L-lactate than Crohn's disease-derived communities (Bauer & Thiele, 2018). Flux analysis further attributes butyrate production almost entirely to Clostridia and propionate production predominantly to Bacteroidia (Bauer & Thiele, 2018), which means the selective depletion of precisely these two taxa constitutes a direct mechanistic bottleneck impairing mucosal SCFA synthesis in inflammatory bowel disease.

4.2 Computational Optimization of Personalized Dietary Interventions

Building on these functional deficits, constraint-based metabolic modeling has been used to screen and optimize personalized dietary supplementation strategies (Bauer & Thiele, 2018). Simulations using the BacArena framework showed that increasing specific fermentable glycan concentrations successfully restored SCFA levels in 24 of 28 simulated Crohn's disease patients (Bauer & Thiele, 2018). For the remaining four, however, no dietary intervention could be identified computationally — their dysbiosis had already eliminated the baseline microbial taxa needed to ferment any prebiotic input at all (Bauer & Thiele, 2018), a finding that underscores just how far dysbiosis can progress before dietary strategies alone stop being viable.

The predicted supplementation regimens were, notably, highly individualized — requiring anywhere from 1 to 55 distinct metabolites per patient (median: 19) (Bauer & Thiele, 2018). Pectin emerged as a particularly frequent therapeutic candidate, identified for 17 of the 24 responsive patients, alongside other plant- and host-derived glycans such as lavanbiose, amylose, larch arabinogalactan, and human milk-type mucus glycans (Bauer & Thiele, 2018). Over a simulated 24-hour period, these predicted glycans produced statistically significant increases in acetate, propionate, and butyrate, whereas isobutyrate and L-lactate proved comparatively resistant to dietary rescue (Bauer & Thiele, 2018). Interestingly, these interventions did not substantially shift the relative abundance of resident taxa — likely because the depleted baseline species were simply absent — suggesting that prebiotic supplementation works primarily by optimizing the metabolic output of an existing, if diminished, commensal network (Bauer & Thiele, 2018). This, in turn, points toward a fairly clear clinical implication: prebiotics alone may not suffice in severely depleted patients, and combining them with next-generation probiotics or live biotherapeutics is likely necessary for durable ecosystem restoration (Bauer & Thiele, 2018).

4.3 Machine Learning Classifiers and Multi-Omics Diagnostic Integration

At the diagnostic level, machine-learning architectures have shown strong performance extracting predictive signal from noisy, high-dimensional multi-omics data (Alexandrescu et al., 2025; Jeyavelkumaran et al., 2026). In colorectal cancer classification specifically, supervised models — Random Forest algorithms in particular — have consistently outperformed conventional clinical tumor markers such as CEA, CA19-9, and CA242 (Jeyavelkumaran et al., 2026). Large-scale meta-analyses of metagenomic datasets report area-under-the-curve values ranging from 0.80 (discriminating adenoma from healthy controls) up to 0.98 when fecal metagenomics and serum metabolomics are integrated (Jeyavelkumaran et al., 2026), reinforcing the broader pattern — already visible in Section 2.2 — that multi-omics fusion meaningfully outperforms single-layer taxonomic profiling.

At the taxonomic level, these predictive models consistently flag a recognizable panel of colorectal cancer-associated pathogens: enrichment of Fusobacterium

(specifically F. nucleatum), Peptostreptococcus, and Parvimonas, alongside depletion of beneficial commensals such as Faecalibacterium and Anaerostipes (Jeyavelkumaran et al., 2026). Quantitative PCR assays targeting Parvimonas micra have validated these associations, yielding a diagnostic sensitivity of 60.5% and specificity of 87.3% (Jeyavelkumaran et al., 2026). Multi-omics fusion algorithms such as DIABLO and MOFA have further linked fecal metagenomic profiles to circulating metabolites, identifying urea-cycle metabolites and microbial lipopolysaccharide and menaquinone biosynthesis pathways as robust diagnostic and prognostic biomarkers in precision oncology (Jeyavelkumaran et al., 2026). To bridge the gap between raw statistical prediction and clinical trust — a gap that matters more than it might initially seem — explainable AI frameworks using Shapley Additive Explanations (SHAP) have confirmed the biological plausibility of these models, successfully mapping transitional microbial states along the adenoma-carcinoma sequence (Alexandrescu et al., 2025; Jeyavelkumaran et al., 2026).

4.4 Precision Biotherapeutic Engineering and Delivery Architectures

Where dietary and statistical approaches offer non-invasive management, clinical translation increasingly relies on programmable live biotherapeutic products (LBPs) engineered for targeted physiological correction (Abdul Manan, 2025; Sutanto & Fetarayani, 2026). Traditional broad-spectrum probiotics — standard Lactobacillus and Bifidobacterium strains — frequently show inconsistent or short-lived clinical benefit, largely because they fail to engraft stably within a resilient, already-dysbiotic gut ecosystem (Sutanto & Fetarayani, 2026; Kashyap et al., 2017). To work around this, researchers have deployed human-derived, strictly anaerobic commensals — Akkermansia muciniphila and Faecalibacterium prausnitzii among them — as next-generation probiotics (Abdul Manan, 2025; El-Sehrawy & Soleimani Samarkhazan, 2026). A pilot study of overweight, insulin-resistant volunteers found that prophylactic A. muciniphila supplementation significantly reduced circulating liver dysfunction and inflammatory markers, offering fairly direct clinical support for its therapeutic potential (Puig-Castellví et al., 2023).

Comparable proof-of-concept results appear outside metabolic disease as well. Ex vivo human gut organoid models derived from pre-celiac individuals showed that cell-free supernatants from Bacteroides vulgatus strain 20220303-A2 attenuated gliadin-induced epithelial damage, preserved tight-junction integrity, and reduced pro-inflammatory cytokine secretion via epigenetic and microRNA-mediated host cellular reprogramming (Fousekis et al., 2026). Synthetic biology has pushed this further still: Escherichia coli Nissle 1917 has been engineered to express phenylalanine ammonia-lyase, providing an autonomous, gut-based route to degrading excess systemic phenylalanine in phenylketonuria (Sutanto & Fetarayani, 2026; Wu et al., 2026), and a first-in-human trial of an EcN strain engineered to consume systemic ammonia demonstrated both safety and proof-of-concept efficacy in patients with hepatic encephalopathy — an early but genuinely encouraging step toward similarly engineered "sense-and-respond" cancer immunotherapies and metabolic interventions (El-Sehrawy & Soleimani Samarkhazan, 2026; Wu et al., 2026). To keep these sensitive engineered strains viable long enough to matter clinically, nanobiotechnology delivery platforms — alginate-chitosan microencapsulation being a representative example — have been shown to dramatically improve probiotic survival through simulated gastric transit, supporting more robust gut mucosal re-establishment and anti-inflammatory activity (Murugan et al., 2026).

4.5 Translational Gaps, Database Biases, and Standardization Pathways

Despite these genuine advances, several technical and ethical barriers still stand between the current evidence base and widespread clinical adoption (Cai et al., 2023). Chief among them is the demographic and geographic skew embedded in existing reference genomic databases (Cai et al., 2023; Gibbons et al., 2022): more than 71% of publicly available human gut microbiome datasets originate from developed nations, with the United States alone contributing 46% of all samples (Cai et al., 2023; Gibbons et al., 2022). This imbalance directly compromises the generalizability of predictive AI models and precision-nutrition algorithms, since models trained largely on Western cohorts do not perform as robustly when applied to more genetically, dietarily, and environmentally diverse populations (Cai et al., 2023; Gibbons et al., 2022).

The field also continues to grapple with substantial methodological heterogeneity (Amiri et al., 2026). Variability in DNA extraction protocols, sequencing

Table 1. Comparative analysis of gut microbiome multi-omics profiling technologies. For each omics dimension (16S rRNA amplicon profiling through host-omics), the table summarizes the biological targets measured, the core diagnostic benefits, the principal technical bottlenecks, representative precision-medicine applications, and the corresponding in-text citation(s). Rows are ordered from lowest to highest biological resolution, illustrating why single-omics profiling is progressively supplemented by higher layers (Section 2.1).

Omics Dimension	Biological Targets & Measured Features	Core Benefits & Diagnostic Resolution	Technical Bottlenecks & Limitations	Precision Medicine & Clinical Applications	In-Text Citations
16S rRNA Amplicon Profiling	Hypervariable regions (e.g., V3–V4) of the bacterial 16S ribosomal RNA gene; taxonomic markers.	Highly cost-effective; establishes baseline bacterial taxonomic “fingerprints”; well-suited for high-throughput population screening.	High susceptibility to PCR amplification biases; fails to resolve at the species- or strain-level; cannot profile non-bacterial kingdoms (fungi, viruses) or active functional outputs.	Large-scale epidemiological cohort profiling; early screening of general dysbiosis patterns in inflammatory and metabolic disorders.	(Cai et al., 2023; Sharma et al., 2026)
Whole-Metagenome Shotgun (WMS) Sequencing	All genomic DNA extracted from the fecal or mucosal specimen; covers bacteria, archaea, fungi, and viruses.	Achieves species- and strain-level taxonomic resolution; enables reconstruction of Metagenome-Assembled Genomes (MAGs); profiles biosynthetic gene clusters and SNPs.	Extremely expensive and computationally intensive; reveals only functional potential rather than active transcription; does not account for native microenvironment.	Multi-cohort biomarker discovery; risk-stratification models in oncology; profiling the fecal resistome and tracking probiotic engraftment.	(Cai et al., 2023; Kharb & Zhu, 2026)
Metatranscriptomics (RNA-Seq)	Total transcribed microbial messenger RNA (mRNA) in the community.	Captures real-time functional gene expression; profiles active metabolic shifts and transcriptional responses to external stressors (e.g., xenobiotics, diet).	Unstable target molecule (short RNA half-life); requires complex extraction and rRNA depletion; low abundance of active transcripts.	Assessing inter-individual variability in active drug metabolism (pharmacomicrobiomics); monitoring acute gut-barrier responses to dietary interventions.	(Cai et al., 2023; Sharma et al., 2026)
Metaproteomics	The translated, physically operational protein landscape (e.g., CAZymes, active enzymes, host mucosal proteins).	Directly confirms presence of functional, active enzymes and host-barrier proteins; links genomic/transcriptomic potential to physical biology.	Highly restricted proteome coverage (2–20% of theoretical proteome); high dynamic range of protein abundances; complex peptide-spectrum matching.	Identifying direct functional markers of host-microbiome crosstalk; discovering active mucosal defense proteins and therapeutic targets.	(Sharma et al., 2026)
Metabolomics (MS and NMR)	Low-molecular-weight chemical endpoints (SCFAs, secondary bile acids, amino acid derivatives, indoles, TMAO precursors).	Represents the functional endpoint of host-microbe co-metabolism; provides a direct readout of the host's actual metabolic phenotype and systemic signaling.	Matrix effects and ionization suppression in MS; poor sensitivity of NMR; demands extensive derivatization for volatiles (GC-MS); large library annotation gaps.	Precision nutrition titration; pharmacomicrobiomic tracking of gut microbial drug metabolism; diagnostic classifier development in oncology and cardiovascular prevention.	(Puig-Castellví et al., 2023; Rajak et al., 2026)
Host-Omics (Genomics, Epigenomics, Transcriptomics)	Host germline genetic variants (SNPs), DNA methylation states, histone acetylation, and tissue-specific mRNA profiles.	Contextualizes host-barrier susceptibility; maps recipient receptor configurations (e.g., SCFA-sensing GPCRs) and mucosal immune cell activation.	High dimensionality and complex data-fusion requirements; cross-cohort generalizability issues; requires temporal sampling to capture dynamic changes.	Stratifying patients for customized dietary and prebiotic interventions; predicting individual metabolic benefits based on host GPCR genotypes.	(Rahimah et al., 2026)

Table 2. Key computational tools, databases, and systems-biology platforms supporting microbiome research. Each row lists a named platform, its resource category, its principal analytical function, its main AI/ML or computational strength, its recognized curation pitfalls or limitations, and the corresponding in-text citation(s), providing a practical reference for selecting a platform appropriate to a given multi-omics integration task (Section 2.2–2.3).

Platform / Database / Framework	Resource Category	Primary Analytical Function & Core Features	AI/ML & Computational Strengths	Curation Pitfalls & Technical Limitations	In-Text Citations
AGORA2	Systems Biology Knowledge Base	High-quality, manually curated genome-scale metabolic reconstructions representing 7,302 human gut microorganisms.	Enables stoichiometric constraint-based modeling (COBRA/MICOM) to predict community-level resource allocation and metabolic fluxes.	Demands intensive computational resources; limited by current knowledge of non-metabolic and immune interactions; lacks kinetic parameters.	(Kharb & Zhu, 2026; Sharma et al., 2026; Sutanto & Fetarayani, 2026)
gutSMASH	Genome Mining Tool	Systematically profiles genome assemblies to predict specialized primary metabolic gene clusters (BGCs) in the gut.	Identified 19,890 primary metabolic clusters across 4,240 high-quality microbial genomes, clarifying specific taxonomic pathways.	Restricted to genome-centric potential; cannot predict actual in vivo transcription or translation levels of predicted pathways.	(Kharb & Zhu, 2026)
gutMGene	Curated Reference Database	Systematically compiles experimentally validated relationships among gut microbes, microbial metabolites, host target genes, and diseases.	Facilitates generation of data-driven, mechanistic hypotheses for target-gene interactions and potential microbiome-based drugs.	Subject to manual curation latency; reference biased toward pathogens and highly abundant commensal species; lacks quantitative context.	(Kharb & Zhu, 2026)
MicrobioLink	Integrated Pipeline	Predicts physical host-microbe protein-protein interactions (PPIs) using domain-domain and domain-motif patterns.	Employs network diffusion theories and systems methodologies to map downstream host cellular signaling and immune cascades.	High false-positive rate of predicted interactions; relies heavily on eukaryotic structural database completeness; requires experimental validation.	(Sharma et al., 2026)
metaGEM	Metabolic Modeling Pipeline	Reconstructs genome-scale, flux-balance analysis-ready metabolic models directly from metagenomic datasets.	Supports community-level flux modeling; allows in silico simulation of niche competition and metabolic cross-feeding in clinical samples.	Strongly dependent on quality of Metagenome-Assembled Genome (MAG) binning; vulnerable to errors in fragmented genomes.	(Kharb & Zhu, 2026)
MintTea	Intermediate-Integration Framework	Combines metabolomics and molecular omics data to identify robust, disease-associated host-microbe-metabolite modules.	Leverages network-based inference to extract robust signatures across multi-omics layers, bypassing single-variable statistical noise.	Vulnerable to dataset-specific batch effects; requires matched multi-omics cohorts with substantial sample sizes to avoid overfitting.	(Kharb & Zhu, 2026)
MicrobeRX & GutBug	In Silico Predictive Frameworks	Uses machine learning and chemical similarity algorithms to forecast enzyme commission numbers and gut bacterial biotransformations.	Successfully mapped over 5,878 metabolites generated by microbial enzymes from 1,083 orally administered therapeutics.	Rely heavily on structural chemical inference; lack quantitative metabolic flux context; performance constrained by limited training data.	(Kharb & Zhu, 2026)
Paired Omics Data Platform (PoDP)	Data-Sharing Repository	Community-curated platform that systematically links untargeted metabolomic datasets with paired metagenomic/genomic assemblies.	Facilitates mining of biosynthetic origins of natural products and structural metabolite annotation using empirical mass spectra.	Limited database scale; heavily reliant on voluntary, high-quality user depositions; metadata reporting lacks universal standardization.	(Puig-Castellví et al., 2023)

platforms, reference database selection, and bioinformatic pipelines introduces batch effects large enough to generate genuinely contradictory, non-reproducible taxonomic associations across studies (Cai et al., 2023; El-Sehrawy & Soleimani Samarkhazan, 2026). In response, multidisciplinary consortia have established standardized reporting frameworks — the STORMS checklist, MIMARKS guidelines, and the Human Microbiome Action Project among them — that, if consistently enforced across clinical trial design, could meaningfully improve data interoperability and cross-study reproducibility (El-Sehrawy & Soleimani Samarkhazan, 2026; Nair et al., 2026).

5. Bridging Computational Promise and Clinical Reality in Precision Microbiome Medicine

5.1 From Correlation to Mechanism: What the Evidence Actually Supports

Read together, the findings synthesized above make a fairly compelling case that dysbiosis is not simply noise around a healthy baseline but a structurally distinct, reproducibly detectable ecological state (Bauer & Thiele, 2018) (Table 1; Figure 1). The taxonomic shifts documented in pediatric Crohn's disease — enrichment of Gammaproteobacteria and Bacilli, depletion of Bacteroidia and Clostridia — translate mechanistically into measurable functional deficits in SCFA production (Bauer & Thiele, 2018), which is precisely the kind of taxon-to-function linkage that correlation-only studies have historically struggled to establish. It is worth being honest, though, about the limits of this evidence: much of it derives from computational simulation rather than direct in vivo measurement, and while genome-scale metabolic models are increasingly well-validated (Sharma et al., 2026), they remain, at bottom, models — approximations that require continued anchoring in wet-lab and clinical data (Section 2.7). This is not a minor caveat; it goes some way toward explaining why the field, for all its computational sophistication, has not yet achieved a comparable degree of clinical standardization.

5.2 Computation as a Bridge, Not a Replacement, for Biological Validation

The comparative performance figures reported across Section 4.3 — an AUC of 0.98 for integrated multi-omics classifiers versus 0.61 for taxonomy-only models (Jeyavelkumaran et al., 2026) (Table 4) — offer a fairly unambiguous argument for multi-omics integration over single-layer profiling. Yet this same result also carries an implicit warning about interpretability: a model's numerical performance says relatively little about whether its underlying logic maps onto real biology, which is exactly the gap that explainable AI approaches such as SHAP were built to address (Alexandrescu et al., 2025; Jeyavelkumaran et al., 2026). The fact that SHAP-derived feature importances have successfully traced transitional microbial states along the adenoma-carcinoma sequence (Jeyavelkumaran et al., 2026) is genuinely reassuring, but it also underscores how much interpretability work remains necessary before regulators, let alone clinicians, are likely to trust these models unsupervised. Similarly, the genome-scale metabolic modeling work reviewed in Sections 2.3 and 4.2 (Bauer & Thiele, 2018) demonstrates that computation can generate individualized, testable hypotheses — which patients will and will not respond to pectin supplementation, for instance — but it cannot substitute for the clinical trials needed to confirm those hypotheses at scale.

5.3 Engineered Biotherapeutics: Encouraging Early Signals, Still-Narrow Evidence BaseThe clinical proof-of-concept data reviewed in Section 4.4 — A. muciniphila supplementation improving metabolic biomarkers (Puig-Castellví et al., 2023), engineered EcN safely reducing systemic ammonia in hepatic encephalopathy (El-Sehrawy & Soleimani Samarkhazan, 2026; Wu et al., 2026), and organoid-validated B. vulgatus supernatants protecting against gliadin-induced epithelial damage (Fousekis et al., 2026) (Table 3) — collectively suggest that engineered and next-generation biotherapeutics are moving past pure concept and into demonstrable early efficacy. That said, these remain, almost without exception, pilot-scale or first-in-human studies; the sample sizes are small, the follow-up durations are short, and — as Section 2.7 makes clear — regulatory frameworks and standardized dosing guidance for living biotherapeutics are still very much under construction (Abdul Manan, 2025; Hussain et al., 2026). The broader oncology literature on multi-omics-guided precision therapeutics, including recent work applying integrative multi-omics and AI approaches to breast cancer management (Sabit et al., 2026), reinforces this same pattern across disease areas: computational and biotherapeutic tools are maturing quickly, but the clinical evidence supporting any single intervention typically lags

Table 3. Comparative analysis of precision biotherapeutic modalities. The table contrasts conventional probiotics, next-generation probiotics/live biotherapeutic products, prebiotics, synbiotics, postbiotics, and engineered living biotherapeutics across their key formulations, biochemical mechanism of action, delivery systems and nanotechnologies employed, and current clinical development or translational status, with corresponding in-text citations (Section 2.4–2.6).

Therapeutic Modality	**Key Formulations & Bioactive Agents**	**Biochemical Mechanism & Mode of Action**	**Key Delivery Systems & Nanotechnologies**	**Clinical Development & Translational Status**	**In-Text Citations**
Conventional Probiotics	Live *Lactobacillus* and *Bifidobacterium* species; *Saccharomyces boulardii*.	Non-specific pathogen exclusion, competitive colonization of epithelial binding sites, and broad host mucosal immunomodulation.	Standard gelatin or cellulose capsules; vulnerable to rapid inactivation in severe gastric acid and bile environments.	Historically mature; widely sold as commercial dietary supplements; modest and highly variable clinical engraftment.	(El-Sehrawy & Soleimani Samarkhazan, 2026; Sutanto & Fetarayani, 2026)
Next-Generation Probiotics (NGPs) / LBPs	Strict anaerobic commensals: *Akkermansia muciniphila*, *Faecalibacterium prausnitzii*, *Eubacterium hallii*.	Targeted metabolic correction, restoration of homeostatic SCFA (butyrate) pathways, and selective upregulation of host tight junctions.	Protective polymeric hydrogel encapsulation (alginate, chitosan) to buffer anaerobic cells during oral passage.	Emerging biotherapeutic products; Phase II/III clinical trials completed for recurrent *C. difficile* prevention and IBD.	(Abdul Manan, 2025; Sutanto & Fetarayani, 2026)
Prebiotics	Inulin, fructooligosaccharides (FOS), pectin, resistant starch, mucin glycans.	Substrates selectively fermented by indigenous beneficial taxa, promoting expansion of specific SCFA-producing guilds.	Polymeric nanoparticles formulated to transport fermentable glycans directly to specialized distal colonic niches.	Highly accessible dietary adjuncts; early-stage precision nutrition trials utilizing genome-scale modeling for personalized fiber dosing.	(Gibbons et al., 2022; Kashyap et al., 2017)
Synbiotics	Synergistically acting combinations of tailored prebiotics and functional probiotic strains.	Co-administration ensures immediate metabolic fuel for the engrafting strain, improving ecological survival and colonization.	Multi-layered responsive nanocarriers that sequentially release prebiotics and probiotics under target physiological triggers.	Emerging therapeutic platform; pre-clinical evaluation and clinical trials demonstrating superior gut restoration over probiotics alone.	(Alexandrescu et al., 2025; Nair et al., 2026)
Postbiotics & Metabolites	Inanimate microbial preparations, SCFAs (butyrate), cell lysates, bacteriocins, and outer membrane vesicles.	Direct anti-inflammatory and epithelial-barrier reinforcement without the viability risks of live biotherapeutic products.	Stimuli-responsive mesoporous silica nanoparticles (MSNs) designed for colonic enzyme- or pH-responsive release.	Highly stable pharmaceutical alternatives; clinical use of topical/rectal SCFA preparations in distal ulcerative colitis.	(Abdul Manan, 2025; Nair et al., 2026; Sutanto & Fetarayani, 2026)
Engineered Live Biotherapeutics	Reprogrammed commensals (e.g., *E. coli* Nissle 1917, *L. lactis*) carrying synthetic gene circuits.	“Sense-and-respond” logic: detects inflammation markers (nitrate, ROS) and executes site-specific synthesis and release of therapeutic payloads (IL-10).	Advanced microencapsulation paired with synthetic biological containment networks (auxotrophic kill-switches).	Preclinical and early-phase clinical development; first-in-human trial of ammonia-consuming EcN for hepatic encephalopathy completed.	(El-Sehrawy & Soleimani Samarkhazan, 2026; Sutanto & Fetarayani, 2026)

Table 4. Key artificial intelligence and machine learning methodologies used for disease biomarker discovery and patient stratification in microbiome research. The table summarizes each computational method family, the algorithmic implementations used, the features and input dimensions resolved, reported diagnostic or predictive performance metrics, and the validation or causal-verification steps still required, with corresponding in-text citations (Section 2.2 and 4.3).

Computational Method	Key Algorithmic Families & Implementations	Features Resolved & Input Dimensions	Reported Diagnostic / Predictive Metrics	Required Validation & Causal Verification	In-Text Citations
Supervised Learning	Random Forest (RF), Support Vector Machines (SVM), Extreme Gradient Boosting (XGBoost).	High-dimensional taxonomic tables (OTU/ASV), metabolic pathway abundance profiles, and target serum metabolites.	RF models achieve diagnostic AUCs of ~0.83 (species-level), ~0.78 (enzyme-level), and ~0.90 (integrated microbial-clinical panels).	Demands rigorous feature ablation analysis (leave-top-1-out evaluation) and blind validation across geographically diverse cohorts.	(El-Sehrawy & Soleimani Samarkhazan, 2026; Jeyavelkumaran et al., 2026)
Unsupervised Learning	Partition Around Medoids (PAM), Hierarchical Clustering with Complete Linkage, Autoencoders.	Beta-diversity matrices (Bray–Curtis dissimilarity, unweighted UniFrac) and compressed multi-omics latent feature spaces.	Identifies three highly stable, reproducible gut enterotypes and co-abundance gene groups (CAGs) across global populations.	Verification of cluster stability across independent validation cohorts; evaluation of sensitivity to choice of distance metrics.	(Alexandrescu et al., 2025; Jeyavelkumaran et al., 2026)
Deep Learning	Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Transformers.	Multi-View Convolutional Variational Information Bottlenecks (MV-CVIB) and Phylogenetic Multi-path CNNs (PM-CNN).	Achieved diagnostic AUCs exceeding 0.90 in predicting metastatic colorectal cancer (mCRC) by modeling disease–disease states.	Requires confirmation in human-relevant ex vivo systems (gut-on-a-chip, organoids) and germ-free (gnotobiotic) animal models.	(Jeyavelkumaran et al., 2026)
Reinforcement Learning	Q-Learning, Deep Q-Networks (DQN), ABIOME System + MARS.	In silico simulated host–microbiome growth trajectories, species interaction matrices, and metabolite flux constraints.	Predicts and optimizes synergistic multi-strain probiotic combinations to maximize localized, targeted metabolite synthesis.	Empirical validation of predicted metabolic fluxes utilizing in vitro continuous-culture bioreactors (e.g., SHIME, TIM-1).	(Jeyavelkumaran et al., 2026; Sutanto & Fetarayani, 2026)
Explainable AI (XAI)	SHapley Additive exPlanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME).	Post-hoc feature attribution of discriminative taxa (e.g., *Fusobacterium*, *Peptostreptococcus*, *Parvimonas* vs. *Eubacterium eligens*).	Random Forest integrated with SHAP achieved predictive precision of 0.729 ± 0.038 and AU-PRC of 0.668 ± 0.016.	Essential cross-model consistency validation; benchmarking of attribution values against mechanistic, in vitro co-culture assays.	(Alexandrescu et al., 2025; Jeyavelkumaran et al., 2026)

well behind the underlying technology.

5.4 The Equity Problem: Why Database Bias Is Not Just a Technical Footnote

If there is one finding in this synthesis that deserves more attention than it typically receives, it is probably the demographic skew in reference microbiome databases — over 71% of public data from developed nations, 46% from the United States alone (Cai et al., 2023; Gibbons et al., 2022) (Figure 4). This is not merely an academic completeness issue; because gut microbiome composition is so tightly linked to host genetics, diet, and environment (Cai et al., 2023), a model trained almost exclusively on Western cohorts will, almost by construction, underperform when applied elsewhere. Precision nutrition frameworks that integrate host genetics, epigenetics, and microbiome data (Rahimah et al., 2026) are only as generalizable as the populations used to build them, and right now, that generalizability is genuinely limited. Addressing this will likely require more than good intentions — federated learning architectures that allow model training across geographically distributed, non-centralized datasets (Alexandrescu et al., 2025; Cai et al., 2023; Gibbons et al., 2022) offer one plausible technical route, but they will need to be paired with deliberate, well-funded recruitment of underrepresented populations if the resulting therapeutics are to serve more than a narrow demographic slice of the world.

5.5 Toward Reproducible, Standardized Precision Microbiome Science

Finally, the methodological heterogeneity documented in Section 2.7 and reinforced in Section 4.5 — variable extraction protocols, inconsistent sequencing platforms, incompatible bioinformatic pipelines (Amiri et al., 2026; Cai et al., 2023) — is, in a sense, the connective tissue linking every other limitation discussed above. Batch effects introduced by inconsistent methodology can just as easily masquerade as biological signal, undermining both the mechanistic claims discussed in Section 5.1 and the diagnostic performance metrics discussed in Section 5.2. The emergence of standardized reporting frameworks such as STORMS and MIMARKS (El-Sehrawy & Soleimani Samarkhazan, 2026; Nair et al., 2026) (Table 2) represents a genuinely constructive response, but standards only work if journals, funders, and regulatory bodies actually enforce them — a step that, admittedly, remains inconsistent across the current publication landscape. Taken together, then, the path from metagenomic data to a validated clinical bioactive is scientifically plausible and, in several documented instances, already partially realized; what remains is less a question of further discovery than of coordinated infrastructure — standardized methods, mechanistic validation, and globally representative data — sufficient to carry these computational innovations reliably into everyday clinical practice.

6. Conclusion

This review traces a single throughline—from raw metagenomic sequence to a bioactive product a clinician could actually prescribe—and finds each link matured unevenly. Multi-omics profiling is now genuinely comprehensive, resolving not just who inhabits the gut but what they are doing and producing. Computational platforms, from supervised machine learning to genome-scale metabolic reconstruction, translate that complexity into predictions that are, at least in silico, patient-specific and mechanistically interpretable. Clinical bioactive design has kept pace too, shifting from generic probiotic supplementation toward engineered live biotherapeutics and stimuli-responsive nanocarriers that release their payload precisely where needed. Yet none of this progress travels far without methodological standardization, mechanistic validation beyond correlation, and reference databases reflecting genuine global diversity. Until those gaps close, predictive models trained overwhelmingly on Western cohorts will keep underperforming elsewhere. The path forward is therefore not simply more data, but the right data—collected and reported consistently enough to replicate across laboratories, and distributed equitably enough to serve more than a narrow demographic slice of the world's population.

References

Abdul Manan, M. (2025). The role of probiotics in personalized therapeutics: Advances in gut microbe-driven interventions. The Microbe, 8, Article 100497. https://doi.org/10.1016/j.microb.2025.100497

Alexandrescu, L., Tofolean, I. T., Condur, L. M., Tofolean, D. E., Nicoara, A. D., Serbanescu, L., Rusu, E., Twakor, A. N., Dumitru, E., Dumitru, A., Tocia, C., Herlo, L. F., Alexandrescu, D. M., & Stanigut, A. M. (2025). Smart microbiomes: How AI is revolutionizing personalized medicine. Bioengineering, 12(9), Article 944. https://doi.org/10.3390/bioengineering12090944

Amiri, D., Briziarelli, L., Shah, S. J., Karlsson, B., Amiri, S., Vakili, A., & Mustakova, A. (2026).

Methodological heterogeneity and reproducibility in psychiatric microbiomics. Journal of Affective Disorders Reports, 25, Article 101133. https://doi.org/10.1016/j.jadr.2026.101133

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics.

Bauer, E., & Thiele, I. (2018). From metagenomic data to personalized in silico microbiotas: Predicting dietary supplements for Crohn's disease. npj Systems Biology and Applications, 4, Article 27. https://doi.org/10.1038/s41540-018-0063-2

Cai, J., Auster, S., Cho, S., Leone, V., Hubert, N., Chang, E., & Dumas, M. E. (2023). Dissecting the human gut microbiome to better decipher drug liability: A once-forgotten organ takes center stage. Journal of Advanced Research, 52, 171–201. https://doi.org/10.1016/j.jare.2023.07.002

El-Sehrawy, A. A. M. A., & Soleimani Samarkhazan, H. (2026). Translational oncology: The gut microbiota in cancer immunity and immunotherapy. Translational Oncology, 70, Article 102833. https://doi.org/10.1016/j.tranon.2026.102833

Fousekis, F., Lianos, G. D., Stavropoulou, E., Patrikiou, E., Vradelis, S., Cassimos, D., & Tsigalou, C. (2026). Decoding gut microbiome alterations in celiac disease: Implications for pathogenesis and treatment. Autoimmunity Reviews, 104127.. https://doi.org/10.1016/j.autrev.2026.104127

Gibbons, S. M., Gurry, T., Lampe, J. W., Chakrabarti, A., Dam, V., Everard, A., Goas, A., Gross, G., Kleerebezem, M., Lane, J., Maukonen, J., Penna, A. L. B., Pot, B., Valdes, A. M., Walton, G., Weiss, A., Zanzer, Y. C., Venlet, N. V., & Miani, M. (2022). Perspective: Leveraging the gut microbiota to predict personalized responses to dietary, prebiotic, and probiotic interventions. Advances in Nutrition, 13(5), 1450–1461. https://doi.org/10.1016/j.advnut.2022.08.002

Hussain, A., Ilyas, R., Tasleem, S., Mojgani, N., Shah, H., & Ali, S. A. (2026). Integrated probiotic, nanotechnology, and artificial intelligence-based strategies for the management of multidrug-resistant pathogens: A review. Next Nanotechnology, 10, Article 100685. https://doi.org/10.1016/j.nexnan.2026.100685

Jalalifar, S., Bajelan, B., Mohammadi, R., Ghafoury, R., Kalhori, Z., Pooshang-Bagheri, K., ... & Faranoush, M. (2026). The impact of gut microbiota on leukemia and prospects for novel therapies. Infectious Medicine, 100239.

Jeyavelkumaran, R., Saravanan, M. D., Keerthivasan, N., Jeya, S. H., & Kumaran, J. (2026). Machine learning driven multi omics and probiotic integration for precision colorectal cancer management. Biomedical Analysis, 3, 24–37. https://doi.org/10.1016/j.bioana.2026.05.002

Kashyap, P. C., Chia, N., Nelson, H., Segal, E., & Elinav, E. (2017). Microbiome at the frontier of personalized medicine. Mayo Clinic Proceedings, 92(12), 1855–1864. https://doi.org/10.1016/j.mayocp.2017.10.004

Kharb, A., & Zhu, X. (2026). Unlocking therapeutic impacts of the gut microbiota with computational tools. Current Opinion in Biotechnology, 97, Article 103431. https://doi.org/10.1016/j.copbio.2025.103431

Moshref-Javadi, M., Ahmadbeigi, G., Ataollahi, H., Boroomand, S., Kheradmand Kia, S., Soleimani, N., Zia-Jahromi, N., & Hedayati, M. (2026). Bidirectional microbiota-cancer crosstalk: Emerging platforms for advanced diagnostics and precision therapeutics in oncology. Biomedicine & Pharmacotherapy, 201, Article 119745. https://doi.org/10.1016/j.biopha.2026.119745

Murugan, S., Kannan, S., Balakrishnan, J., & Bhosale, N. K. (2026). Therapeutic applications of nano-microbiome interactions. Medicine in Microecology, 30, Article 100184. https://doi.org/10.1016/j.medmic.2026.100184

Nair, V., Vijayakumar, K., Balakrishnan, J., & Suganya, K. (2026). Gut-targeted nutraceutical delivery: Engineering microbiome-responsive nutraceutical interfaces. Next Materials, 13, Article 103125. https://doi.org/10.1016/j.nxmate.2026.103125

Puig-Castellví, F., Pacheco-Tapia, R., Deslande, M., Jia, M., Andrikopoulos, P., Chechi, K., Bonnefond, A., Froguel, P., & Dumas, M. E. (2023). Advances in the integration of metabolomics and metagenomics for human gut microbiome and their clinical applications. Trends in Analytical Chemistry, 167, Article 117248. https://doi.org/10.1016/j.trac.2023.117248

Rahimah, S., Tallei, T. E., Savitri, M., Yamada, C., Kim, H. J., Choi, M., Park, M. N., Ophinni, Y., & Kim, B. (2026). Modular model of precision nutrition: Integrating host genetics, epigenetics, and the gut microbiota. Food Chemistry: Molecular Sciences, 12, Article 100411. https://doi.org/10.1016/j.fochms.2026.100411

Rajak, M., Dash, D., & Koiri, R. K. (2026). Microbiota-derived metabolites in cancer: Dual roles in pathogenesis and opportunities for therapy. Medicine in Drug Discovery, 29, Article 100245. https://doi.org/10.1016/j.medidd.2025.100245

Sabit, H., Yadav, A. K., Salimy, S., Sakr, A., Abdel-Ghany, S., Wadan, A. S., Alqosaibi, A. I., Rashwan, R., AlGosaibi, Y. S., Alnamshan, M. M., Almulhim, J., Alaqeel, N. K., & Arneth, B. (2026). Integrative multi-omics and AI in precision breast cancer management. Cancer Letters, 649, Article 218468. https://doi.org/10.1016/j.canlet.2026.218468

Sharma, I., Bharti, S., Pandey, S., Gupta, S., Bisen, P. S., Kango, N., & Sharma, K. K. (2026). Leveraging computational methodologies for probiotic research: Unraveling host-microbiota interactions and therapeutic potential. Food Nutrition, 2, Article 100067. https://doi.org/10.1016/j.fnutr.2026.100067

Sun, C., Zhu, J., Sun, X., Zhang, Z., Sun, Y., Jin, Y., & Wu, T. (2026). Targeting the human gut microbiome: A comparative review of probiotics, prebiotics, synbiotics, and postbiotics. Journal of Advanced Research, 87, 727–752. https://doi.org/10.1016/j.jare.2025.12.025

Sutanto, H., & Fetarayani, D. (2026). Engineering the gut microbiome: Synthetic biology approaches for human health and disease. Gut Microbiology, 2, Article 100005. https://doi.org/10.1016/j.gutmic.2026.100005

Wu, Y., Zhang, K., Xu, B., Zengin, G., & Li, M. Y. (2026). Challenges and future perspectives in microbiome-metabolism-based therapeutics. TransMed, 1, Article 100032. https://doi.org/10.1016/j.tmed.2026.100032