Title

Multiomic investigations of Body Mass Index reveal heterogeneous trajectories in response to a lifestyle intervention

Authors

Kengo Watanabe[1], Tomasz Wilmanski[1], Christian Diener[1], Anat Zimmer[1,†], Briana Lincoln[1], Jennifer J. Hadlock[1], Jennifer C. Lovejoy[1,2], Andrew T. Magis[1], Leroy Hood[1], Nathan D. Price[1,3], and Noa Rappaport[1,*]

Affiliations

[1]Institute for Systems Biology, Seattle, WA 98109, USA.
[2]Lifestyle Medicine Institute, Redland, CA 92374, USA.
[3]Thorne HealthTech, New York, NY 10019, USA.
[†]Present address: Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA.
[*]Correspondence to Noa Rappaport (noa.rappaport@isbscience.org)

Abstract

Multiomic profiling is useful in characterizing heterogeneity of both health and disease states. Obesity exerts profound metabolic perturbation in individuals and is a risk factor for multiple chronic diseases. Here, we report a global atlas of cross-sectional and longitudinal changes associated with Body Mass Index (BMI) across 1,100+ blood analytes, as well as their correspondence to host genome and fecal microbiome composition, from a cohort of 1,277 individuals enrolled in a wellness program. Machine learning-based models predicting BMI from blood multiomics captured heterogeneous states of both metabolic and gut microbiome health better than classically measured BMI, suggesting that multiomic data can provide deeper insight into host physiology. Moreover, longitudinal analyses identified variable trajectories of BMI in response to a lifestyle intervention, depending on the analyzed omics platform; metabolomics-based BMI decreased to a greater extent than actual BMI, while proteomics-based BMI exhibited greater resistance. Our analysis further elucidated blood analyte–analyte associations which were significantly modified by obesity and partially reversed in the metabolically obese population through the program. Altogether, our findings provide an atlas of the gradual blood perturbations accompanying obesity and serve as a valuable resource for robustly characterizing metabolic health and identifying actionable targets for obesity.

Introduction

Obesity prevalence has been increasing over the past four decades in adults, adolescents and children around most of the world[1,2]. Many studies have demonstrated that obesity is a major risk factor for multiple chronic diseases such as type 2 diabetes mellitus (T2DM), metabolic syndrome, cardiovascular disease (CVD), and certain types of cancer[3–6]. In obese humans, even 5% loss of body weight can be sufficient for improving metabolic and cardiovascular health[7], and weight loss with lifestyle interventions can reduce the risk for obesity-related chronic diseases[8]. Nevertheless, obesity and its physiological manifestations appear highly heterogeneous, necessitating additional research to better understand this prevalent health condition.

Most commonly, obesity is assessed with an anthropometric Body Mass Index (BMI), defined as the body weight divided by the squared body height [kg m^{-2}]. BMI does not directly measure body composition, but at the population level, BMI correlates well with direct measurements of body fat percentage using computed tomography (CT), magnetic resonance imaging (MRI), or dual-energy X-ray absorptiometry (DXA)[9]. Because BMI is easily calculated and commonly understood among researchers, clinicians, and the general public, it is widely used for the primary diagnosis of obesity and applied as a simple index of efficiency in lifestyle intervention.

At the same time, there are considerable limitations to BMI as a surrogate of health/disease metric; e.g., differences in body composition can lead to misclassification of people with a high muscle-to-fat ratio (e.g., athletes) as obese, and can undervalue metabolic improvement with exercise[10]. In addition, a meta-analysis demonstrated that the common obesity diagnosis using a BMI cutoff has high specificity but low sensitivity in identifying individuals with excess body fat[11], which stems from the difference of appropriate cutoff threshold between ethnic populations as well as the existence of metabolically obese, normal-weight (MONW) individuals[12,13]. Likewise, there are subgroups in the BMI-classified obese group: i.e., metabolically healthy obese (MHO) and metabolically unhealthy obese (MUO). Although most MHO individuals are not necessarily healthy, and are simply healthier than MUO individuals[14], the transition from MHO to MUO may be an important feature for development of obesity-related chronic diseases[15]. Moreover, this transition is potentially preventable through lifestyle interventions[16]. Altogether, BMI is undoubtedly useful but too crude to provide insight into heterogeneous states of obesity.

Omics studies have suggested the usefulness of blood omics data in health-related metrics; e.g., blood proteomics captures many health conditions[17], and blood metabolomics reflects habitual diets and gut microbiome profiles[18,19]. Intriguingly, a recent study showed that a machine learning-based BMI model comprising 49 BMI-associated blood metabolites captures obesity-related health outcomes (e.g., percent visceral fat, blood pressure) better than BMI and genetic risk[20]. Hence, although a targeted metric (e.g., body composition) or a specific biomarker provides useful information, multiomic blood profiling has greater potential to bridge the multifaceted gaps between BMI and the complex physiological states across obese individuals.

Previously, our research group has generated personal, dense, dynamic data (PD3) clouds on 108 healthy participants as part of a scientific wellness pilot study[21]. The PD3 clouds include human genomes and longitudinal measurements of metabolomics, proteomics, clinical lab tests, gut microbiomes, wearable devices, and health/lifestyle questionnaires accompanied by a coaching intervention. The cohort has since been expanded to over 5,000 individuals, contributing to multiple novel scientific findings[19,22–26]. In this study, we leveraged the wealth of the collected dataset to investigate the physiological changes accompanying obesity.

Herein, using BMI as a starting point, we report our investigations into the multiomic perturbations associated with obesity. Blood analytes across all studied omics platforms have a strong capacity to explain a large portion of the variation in BMI. Through machine learning approaches, we also show a heterogeneity in metabolic states accompanying obesity, which is not captured by measured BMI. Furthermore, longitudinal analyses demonstrate variable changes in metabolic health through lifestyle coaching; i.e., the plasma metabolomics exhibits stronger response than measured

BMI, while the plasma proteomics exhibits weaker response. Our findings highlight the power of blood multiomics in investigating the underlying physiology of obesity and weight loss from a clinical standpoint.

Results

Plasma multiomics captures 48–78% of the variance in BMI

To investigate the molecular effects of obesity on metabolic profiles, we defined a study cohort of 1,277 adults who participated in a wellness program (Arivale)[19,21–26] and whose datasets included coupled measurements of metabolomics, proteomics, and clinical labs from the same blood draw (see Methods). This study design allowed us to directly investigate the similarities and differences between omics platforms in regards to how they reflected the metabolic state of each individual. The defined cohort was characteristically female (64.3%), middle aged (mean ± s.d.: 46.6 ± 10.8 years), and white (69.7%) (Supplementary Fig. 1), consistent with our previous studies[19,22–26]. Based on the World Health Organization (WHO) international standards of BMI cutoff (underweight: <18.5 kg m^{-2}, normal: 18.5–25 kg m^{-2}, overweight: 25–30 kg m^{-2}, obese: ≥30 kg m^{-2})[27], the baseline prevalence was similar between normal, overweight, and obese classes, and only 0.8% of participants were underweight class (underweight: 10 participants (0.8%), normal: 426 participants (33.4%), overweight: 391 participants (30.6%), obese: 450 participants (35.2%)).

Leveraging the baseline measurements of plasma molecular analytes (metabolomics: 766 metabolites, proteomics: 274 proteins, clinical labs: 71 clinical lab tests; Supplementary Data 1), we generated omics-based machine-learning models predicting BMI. To address multicollinearity in our dataset (Supplementary Fig. 2a), we applied a tenfold cross-validation (CV) scheme of the least absolute shrinkage and selection operator (LASSO) algorithm for the baseline metabolomics, proteomics, clinical labs, and combined-omics (i.e., using all metabolomics, proteomics, and clinical labs) datasets. This approach generated metabolomics-based, proteomics-based, clinical labs (chemistries)-based, and combined omics-based BMI predictions (MetBMI, ProtBMI, ChemBMI and CombiBMI, correspondingly). The resulting models retained 62 metabolites, 30 proteins, 20 clinical lab tests, and 132 analytes across all ten MetBMI, ProtBMI, ChemBMI, and CombiBMI models, respectively. These selected predictor variables exhibited low collinearity (Supplementary Fig. 2b, c), validating the variable selection during LASSO modeling[28]. The generated models demonstrated remarkably high performance for BMI prediction, ranging from mean out-of-sample $R^2 = 0.48$ (ChemBMI) to 0.70 (ProtBMI) (Fig. 1a, b). The CombiBMI model further improved the performance of BMI prediction ($R^2 = 0.78$; Fig. 1b), suggesting that, although there is a considerable overlap in the signal detected by each omics platform, different omic measurements still contain unique information regarding BMI. The performance ordering between the models (ChemBMI < ProtBMI ~ MetBMI < CombiBMI) were also consistent in sex-stratified models (Supplementary Fig. 3a).

BMI has been reported to be associated with many anthropometric and clinical metrics such as waist circumference, blood pressure, sleep quality, and polygenic risk scores (PRSs)[3,4,14,24,29]. As a first test for the validity of omics-based BMI models, we examined the association of omics-based BMI with each of 51 numeric physiological measures. BMI was significantly associated with 27 features (false discovery rate (FDR) < 0.05) including daily physical activity measures from wearable devices, waist-to-height ratio, blood pressure, and BMI PRS (Fig. 1c). With minor differences in significant features between the models (MetBMI: 25 features, ProtBMI: 25 features, ChemBMI: 25 features, CombiBMI: 25 features), there was concordance among the associations of all omics-based BMI predictions and these BMI-associated features (Fig. 1c), indicating that the omics-based BMIs basically maintain the characteristics of BMI in terms of anthropometric, genetic, lifestyle, and physiological parameters.

Omics-based BMI captures the variation in BMI better than any single analyte

To confirm the robustness of the variable selection process, we iterated the LASSO modeling while dropping the analyte with the strongest β-coefficient in each iteration step. If a variable is indispensable for a model, the performance should decrease drastically after dropping the variable. In all omics-based BMI models, a steep decay in the mean out-of-sample R^2 across ten models was observed in the first 5–9 iterations (Supplementary Fig. 3b–e), suggesting that the variables that had the largest absolute β-coefficient values in the original LASSO models were the most important in predicting BMI. Interestingly, the overall slope of R^2 decay in the MetBMI model was more gradual than that in the ProtBMI and ChemBMI models (Supplementary Fig. 3b–d), implying that metabolites contain more redundant information to predict BMI. Indeed, the proportion of the variables that were robustly retained across all ten LASSO models to the variables that were retained in at least one of ten LASSO models was lower in the MetBMI model compared to the ProtBMI and ChemBMI models (MetBMI: 62/209 metabolites ≈ 29.7%, ProtBMI: 30/74 proteins ≈ 40.5%, ChemBMI: 20/41 clinical lab tests ≈ 48.9%), confirming the higher level of redundancy in metabolomics. Nevertheless, a large number of metabolites remained in the robust 132 analytes of the CombiBMI model (77 metabolites, 51 proteins, 4 clinical lab tests; Fig. 2a), suggesting that each of the data types possesses unique information about BMI. The strongest positive predictors in the CombiBMI model (mean β-coefficient > 0.02) were leptin (LEP), adrenomedullin (ADM), and fatty acid-binding protein 4 (FABP4), and the strongest negative ones (mean β-coefficient < -0.02) were insulin-like growth factor-binding protein 1 (IGFBP1) and advanced glycosylation end-product specific receptor (AGER; also described as receptor of AGE, RAGE); that is, proteins were the strongest contributors to the model. Furthermore, although it is possible that metabolites which are highly associated with the retained proteins were eliminated from the CombiBMI model due to collinearity, the absolute β-coefficient values of the robustly retained variables were still lower and the differences among them were milder in the MetBMI model compared to the ProtBMI model (Supplementary Fig. 4).

At the same time, the existence of strong and robust predictors in the omics-based BMI model implied that a single analyte may carry sufficient information to predict BMI. To address this possibility, we regressed BMI independently to each analyte that was retained in at least one of ten LASSO models (MetBMI: 209 metabolites, ProtBMI: 74 proteins, ChemBMI: 41 clinical lab tests). Among the analytes that were significantly associated with BMI (180 metabolites, 63 proteins, 30 clinical lab tests), only LEP, FABP4, and interleukin 1 receptor antagonist (IL1RN) univariately explained over 30% of the variance in BMI (Fig. 2b–d), with a maximum of 37.9% (LEP). In contrast, the MetBMI, ProtBMI, and ChemBMI models explained 68.9%, 70.6%, and 48.8% of the variance in BMI, respectively. Moreover, even upon eliminating several strong predictor analytes such as LEP and FABP4 from the omic datasets, the generated models still explained larger variance in BMI than any single analyte (Supplementary Fig. 3b–e). These results indicate that the omics-based BMI models explain a larger portion of the variation in BMI than any single analyte.

Metabolic heterogeneity within standard BMI classes underlies the high rate of misclassification

Although the omics-based BMIs concordantly represented the characteristics of BMI (Fig. 1c), we still observed that the difference between the measured and predicted BMIs was significantly correlated between the omics-based BMI models (Fig. 3a), implying the cases where the omics-based predictions deviated from the measured BMI were in fact a result of a different underlying metabolic state consistently reflected across the omics-based BMIs, rather than an artifact of the model generation. In addition, upon classifying the participants into the WHO international standard BMI classes based on either the measured or the omics-based BMI values, the misclassification rate was approximately 30% across all omics categories and BMI classes (Fig. 3b), consistent with the previously reported misclassification rates derived from the cardiometabolic health classification[30,31].

We then re-examined the BMI-associated features (Fig. 1c) while stratifying by measured BMI class and misclassification status; i.e., each participant was classified using both measured BMI and the predicted BMI (MetBMI or ProtBMI) based on the standard BMI cutoffs, and categorized into "Matched" or "Misclassified" when the BMI-based class matched or mismatched the omics-based

BMI class, respectively. The misclassified group of normal BMI class individuals exhibited significantly higher values of features that are positively associated with BMI, such as waist-to-height ratio, heart rate, and blood pressure, and significantly lower values of features that are negatively associated with BMI, such as daily physical activity measures, compared to the corresponding matched group of normal BMI class individuals (Fig. 3c), suggesting that a participant misclassified into the normal BMI class possesses an unhealthier molecular profile reflected by metabolomics and proteomics, similar to that of overweight and obese individuals, and corresponding to a MONW individual. Conversely, the misclassified group of obese BMI individuals exhibited significantly lower values of features that are positively associated with BMI and significantly higher values of features that are negatively associated with BMI compared to the corresponding matched group of obese BMI class individuals (Fig. 3c), suggesting that a participant misclassified into the obese BMI class has a healthier metabolomic and proteomic signature, similar to that of overweight and normal individuals, and corresponding to a MHO individual. Importantly, there was no difference in BMI PRS between the matched and misclassified groups (Fig. 3c), implying that the discordance between measured and omics-based BMIs may stem from lifestyle or environmental factors, rather than genetic propensity. In this analysis, although a statistical difference in age between the matched and misclassified groups of the normal BMI class was also observed (Fig. 3c), the age difference does not explain the above differences, as the statistical models were adjusted for age. The findings of concordant patterns between matched and misclassified groups of the normal and obese BMI classes were strengthened by consistent trends in multiple other known obesity-related health markers[3,14,32–34], including triglyceride, high-density lipoprotein (HDL) cholesterol, adiponectin, high-sensitivity C-reactive protein (CRP), homeostatic model assessment for insulin resistance (HOMA-IR), glycohemoglobin (HbA1c), and vitamin D (Supplementary Fig. 5a). Taken together, these results suggest that the omics-based BMI models capture the heterogeneous metabolic health states of individuals which are not captured by standard (measured) BMI cutoffs.

To further explore the molecule-level difference between the matched and misclassified groups, we applied unsupervised hierarchical clustering using proteomics and metabolomics data. We observed three clusters for the normal BMI class in a proteomic space defined with the strongest 15 proteins in the ProtBMI models, and Cluster 3 and Cluster 2 were significantly enriched for the matched and misclassified group individuals, respectively (Fig. 3d). In Cluster 3, the expression levels of these proteins prominently corresponded to their contributing directions in the ProtBMI model. In contrast, this correspondence was blurred in Cluster 2, implying that the misclassification may emerge from dysregulation of strong predictor proteins. In addition, we observed three clusters for the obese BMI class in the same proteomic space, and Clusters 2 and 3 and Cluster 1 were significantly enriched for the matched and misclassified group individuals, respectively (Fig. 3e). Likewise, the strong predictor proteins in Cluster 1 (i.e., the misclassified individuals-enriched cluster in obese BMI class) exhibited weaker agreement between their expression levels and their contributing directions in the ProtBMI model. Furthermore, similar patterns were observed in a metabolomic space, clustering the strongest 15 metabolites in the MetBMI models (Supplementary Fig. 5b, c). These results imply that the conventional BMI classification failed to capture the differences in molecular regulatory states of blood metabolites and proteins.

MetBMI reflects gut microbiome profiles better than BMI

Gut microbiome causally affects host obesity phenotypes in a mouse model[35] and obese human individuals generally exhibit lower bacterial diversity and richness[36,37], while some meta-analyses of human studies suggest an inconsistent relationship between gut microbiome and obesity[38,39]. Given our previous finding that the association between blood metabolites and bacterial diversity is dependent on BMI[19] and the current finding that the omics-based BMI models capture heterogeneous metabolic health states of individuals, we hypothesized that the omics-based BMIs represent gut microbiome diversity better than the measured BMI. As expected, while all BMI and omics-based BMIs were significantly associated with different metrics of gut microbiome α-diversity (the number of observed species, Shannon's index, and Chao1 index), the omics-based BMIs explained a larger

portion of the variance in the gut microbiome α-diversity than BMI (Shannon's index: ranging from 7.55% (ChemBMI) to 10.95% (MetBMI) compared to 6.62% (BMI); Fig. 4a, b). In particular, MetBMI explained the largest portion of the variance, consistent with our previous observation that plasma metabolomics showed a much stronger correspondence to gut microbiome structure than either proteomics or clinical labs[19].

We further examined the predictive power of gut microbiome profiles for the omics-based BMI. We generated models classifying normal versus obese individuals using a random forest classifier trained on gut microbiome 16S amplicon sequencing data. The gut microbiome-based classifier for MetBMI categories showed significantly larger area under curve (AUC) in receiver operator characteristic (ROC), sensitivity, and precision compared to these performance parameters of the classifier for measured BMI categories ($P = 0.007$ (AUC), 0.007 (sensitivity), 0.019 (precision); Fig. 4c, d). Therefore, these results suggest that the MetBMI model outperforms BMI even in its capacity to reflect gut microbiome profiles.

Metabolic health of the metabolically obese group was substantially improved following a positive lifestyle intervention

In the Arivale program, personalized lifestyle coaching was provided, resulting in significantly positive clinical outcomes[22]. To investigate the corresponding longitudinal changes in omics-BMI models, we utilized the available longitudinal measurements from a subcohort of 608 participants (see Methods). Given the variability in time between data collection points, we estimated the mean transition of the measured and omics-based BMIs using a linear mixed model (LMM). Consistent with the previous analysis[22], the mean BMI estimate for the overall cohort decreased during the program (Fig. 5a). The decrease of MetBMI was larger than that of BMI while the decrease of ProtBMI was minimal and even smaller than that of BMI (Fig. 5a), suggesting that plasma metabolomics are highly responsive to weight loss, while proteomics (measured from the same blood draw) are resistant to the same lifestyle coaching. Subsequently, we generated LMMs stratified by baseline BMI class, and confirmed that a significant decrease in the mean BMI estimate was observed in the overweight and obese BMI classes, but not in the normal BMI class (Fig. 5b). Concordantly, the mean estimates of ProtBMI and ChemBMI exhibited negative changes over time in the overweight and obese BMI classes, but not in the normal BMI class (Fig. 5b). However, the mean estimate of MetBMI exhibited a significant decrease even in the normal BMI class (Fig. 5b), suggesting that metabolomics data may capture information about the metabolic health response to lifestyle intervention, beyond baseline BMI class or changes in BMI and other omics.

Since we revealed that multiple metabolic health states exist within the standard BMI classes (Fig. 3), we further investigated the difference between misclassification strata based on the baseline MetBMI class. In the normal baseline BMI class, while the mean BMI estimate remained constant in both the matched and misclassified groups, the mean MetBMI estimate exhibited larger reduction in the misclassified group than the matched group (Fig. 5c), suggesting that MONW participants improved their metabolic health to a greater extent than metabolically healthy, normal-weight (MHNW) participants. Likewise, in the obese baseline BMI class, while the decrease in the mean BMI estimate was not significantly different between the matched and misclassified groups, the decrease in the mean MetBMI estimate was larger in the matched group than in the misclassified group (Fig. 5d), suggesting that MUO participants improved their metabolic health to a greater extent than MHO participants. Altogether, these results suggest that metabolic health was substantially improved during the program, in accordance with the baseline metabolomic state rather than the baseline BMI class.

Plasma analyte correlation network in the metabolically obese group reverted back to normal state following lifestyle intervention

Finally, we explored longitudinal changes in plasma analytes correlation network, focusing on the metabolically obese group. Based on the importance of baseline metabolomic state (Fig. 5c, d), we

first assessed relationships between each plasma analyte–analyte correlation and baseline MetBMI, using their interaction term of a generalized linear model (GLM) for each analyte–analyte pair in 608 participants (See Methods). In this type of model, the statistical test assesses whether the relationship between any two analytes is dependent on a third variable (in this case, baseline MetBMI). Among 608,856 pairwise relationships of plasma analytes, 91 analyte–analyte correlation pairs derived from 75 metabolites, 26 proteins, and 13 clinical lab tests were significantly modified by baseline MetBMI (FDR < 0.05). Subsequently, we assessed longitudinal changes of the significant 91 pairs, using their interaction term (days in the program) of a generalized estimating equation (GEE) for each pair in 184 metabolically obese participants (See Methods). Among the significant 91 pairs from the GLM models, 14 analyte–analyte correlation pairs were significantly modified by days in the program (Fig. 6a). The significant 14 pairs were mainly derived from metabolites (16 metabolites, 3 clinical lab tests). For instance, homoarginine is a recently discovered biomarker of CVD[40] and was a robust positive predictor in the MetBMI and CombiBMI models (Fig. 2a and Supplementary Fig. 4a), while phenyllactate (PLA) is a gut microbiome-derived phenylalanine derivative having antimicrobial activity and natural antioxidant activity[41,42]; and the positive correlation between homoarginine and PLA was observed in metabolically obese group at baseline (Fig. 6b). However, this correlation in the metabolically obese group was attenuated during the program (Fig. 6c), implying that some types of plasma metabolic regulation specific to the metabolically obese group was improved during the program.

Discussion

Obesity is a significant risk factor for many chronic diseases[3–6]. The heterogeneous nature of human health conditions, with variable manifestation ranging from metabolic abnormalities to cardiovascular symptoms, calls for a deeper molecular characterization in order to optimize wellness and reduce the current global epidemic of chronic diseases. In this study, we demonstrated that obesity profoundly perturbs human physiology, as reflected across all the omics modalities studied. The key findings of this study are: (1) machine learning-based multiomic BMI was a more reliable measure of metabolic health than traditional BMI, while maintaining a high level of interpretability and intuitiveness attributed to the original metric (Fig. 1–3); (2) metabolomic reflection of obesity exhibited the strongest correspondence to gut microbiome profiles in all omics studied (Fig. 4); (3) the plasma metabolomics exhibited stronger (and/or earlier) response to lifestyle coaching than measured BMI, while the plasma proteomics exhibited a weaker (and/or more delayed) response (Fig. 5a, b); (4) MONW (i.e., normal-weight by BMI but overweight or obese by MetBMI) participants exhibited a greater improvement in metabolic state (but not in weight loss itself) in response to lifestyle coaching compared to MHNW (i.e., normal-weight by both BMI and MetBMI) participants, and vice versa for MHO and MUO participants (Fig. 5c, d); (5) dozens of analyte–analyte associations were modified by the metabolic state, stressing the functional context dependence of many blood-measured analytes (Fig. 6).

Multiple observational studies have explored obesity biomarkers. With regard to obesity, the involvements of insulin/insulin-like growth factor (IGF) axis and chronic low-grade inflammation have been discussed in the context of obesity-related disease risks[5,6], backed by robust findings about the association of obesity with IGFBP1/2 (−), adipokines such as LEP (+), adiponectin (−), FABP4 (+), and ADM (+) and proinflammatory cytokines such as interleukin 6 (IL6, +)[32,43]. Consistent with these well-known changes, we observed the positive significant association of BMI with LEP, FABP4, IL1RN, IL6, ADM, and insulin and the negative association with IGFBP1/2 and adiponectin (Fig. 2c, d). Importantly, all these known biomarkers were incorporated into the omics-based BMI models, and most of them emerged as the robustly retained strong components of the models (Fig. 2a and Supplementary Fig. 4b, c). At the same time, we observed that RAGE explains a relatively small proportion of the variance in BMI (Fig. 2c), while being a strong negative predictive variable in all ten models of both ProtBMI and CombiBMI (Fig. 2a and Supplementary Fig. 4b). The importance of soluble RAGE (sRAGE) has been gradually highlighted in T2DM and CVD[44], with several reports on

the negative association of sRAGE with BMI[45]. Therefore, the omics-based BMI may not only reflect obesity, but also reflect early transition to clinical manifestation of obesity-related chronic diseases.

Likewise, many epidemiological studies have revealed metabolomic biomarkers for obesity[46,47]. In line with the previous knowledge, we confirmed the positive significant association of BMI with mannose, uric acid (urate), and glutamate and the negative correlation with asparagine and glycine (Fig. 2b), and all of them were robustly incorporated into all ten models of MetBMI and CombiBMI (Fig. 2a and Supplementary Fig. 4a). Furthermore, we observed that many lipids emerged as the strong components in the MetBMI and CombiBMI models; in particular, glycerophosphocholines (GPCs) are negative variable components while sphingomyelins (SMs) are positive ones (Fig. 2a and Supplementary Fig. 4a), even though both have a phosphocholine group in common. Although lipids were traditionally regarded as positively associated factors with obesity, recent metabolomics studies have uncovered variable trends of fatty acid species; e.g., plasma lysophosphatidylcholines (LPCs) are decreased in obese mice (high-fat diet model)[48], which was in accordance with our analysis (e.g., LPC(18:1), described as 1-oleoyl-GPC(18:1), in Fig. 2b and Supplementary Fig. 4b). However, because there are many combinations of acyl residues in lipids and many potential confounding factors with obesity, a meta-analysis to systematically understand the species-level lipid biomarkers for obesity is still challenging[46,47]. Our approach, applying machine learning to metabolomics data, addresses this challenge by automatically and systematically providing a molecular signature of obesity, while reflecting the versatile and complex metabolite species. Altogether, the omics-based BMI can be regarded as a multidimensional metric of obesity, possessing molecular mechanistic information. Although targeted measurement of a specific metrics (e.g., body composition) or biomarker provides a detailed, albeit narrow, view into the anthropometric parameters of obesity, the blood multiomic measurements can therefore provide a broader view into its heterogeneous molecular dimensions.

Recently, Cirulli and colleagues have reported a machine learning-based model of BMI computed from blood metabolomics captured obesity-related phenotypes[20]. Their main model explained 39.1% of the variance in BMI, while our MetBMI model exhibited higher explained variance (68.9%; Fig. 2b). Other than the difference in studied cohorts, the performance gap may also result from the difference in modeling strategies. Cirulli and colleagues stringently selected 49 metabolites out of their metabolomics panel (1,007 metabolites) based on a pre-screening for significant association with BMI, and subsequently applied a tenfold CV implementation of the Ridge or LASSO method. In contrast, we used the LASSO method for feature selection, applying it to the full metabolomics panel (766 metabolites). In addition to the increased number of modeled metabolites, our higher performance may stem from the existence of metabolites which were critical for BMI prediction but not significantly associated with BMI. In fact, our MetBMI model contained many metabolites which were weakly associated with BMI but robustly retained across all ten models (Fig. 2b and Supplementary Fig. 4a). At the same time, many of their 49 metabolites (14–20 metabolites among the 31–41 corresponding metabolites in our metabolomics panel) were also retained in at least one of ten MetBMI models. Therefore, our strategy of feature selection through machine learning may be preferable for predicting BMI from metabolomics.

A recent study that investigated multiomic changes in response to weight perturbation demonstrated that some weight gain-associated blood signatures reverted during a subsequent weight loss, while others persisted[49]. Interestingly, we revealed that MetBMI is more responsive to lifestyle intervention than BMI and ChemBMI, while ProtBMI is more resistant (Fig. 5a, b). Our analyses of components in the omics-based BMI models (Fig. 2 and Supplementary Fig. 3b–e, 4) suggested that various metabolites share a wider spread of the feature importance, while a small subset of proteins (~5 proteins) predominantly reflects obesity profiles. Therefore, the effect of lifestyle coaching may consist of small additive contributions in the short range rather than affecting the root cause. However, longitudinal analysis with longer periods must be investigated to infer the physiological meaning of the metabolomic responsiveness and the proteomic resistance. For instance, it is possible that ProtBMI shows a delayed response to the same weight loss (over a span greater than a year measured presently; Fig. 5a), indicating blood metabolites and proteins may be early and late responders to a lifestyle

intervention, respectively. If the difference between BMI and the omics-based BMIs remains constant even after one year, it is reinforced that blood metabolites and proteins are more and less sensitive factors of weight loss, respectively. In either scenario, the advantage of monitoring blood multiomics during weight loss programs would be supported as a tool to keep participants motivated, since lifestyle changes generally take longer to manifest as weight loss. At the same time, long-term maintenance of the improvement is an important challenge for lifestyle interventions. Although there is variability between studies, a previous study indicated that only ≈20% of overweight individuals successfully maintained weight loss[50], raising the possibility that BMI and the other omics-based BMIs revert back to the baseline after one year of the Arivale program. Nevertheless, previous studies revealed that lifestyle interventions had benefits in preventing diabetes incidence as long as 20 years post-intervention even if body weight was regained[51,52], implying that the observed larger improvement of MetBMI compared to BMI may persist in the longer term. Hence, deeper investigations on our findings are required, especially the long-term dynamics of the MetBMI and ProtBMI responses, which may provide a foothold to develop a scientific strategy to maintain metabolic health in the long term. In addition to the studied time period, there are additional limitations to be noted in this study. This study was not designed as a randomized control trial, and we cannot strictly evaluate the effectiveness of the lifestyle intervention (e.g., improvement in obese group due to the regression-toward-the-mean effect). As an observational study derived from a consumer facing cohort, generalizability of the findings may be limited. Our measurements did not cover all biomolecules in blood; in particular, proteomics was based on targeted Olink panels. Hence, our findings on metabolomic and proteomic states are restricted to the studied space. Nevertheless, this study will serve as a valuable resource for robustly characterizing metabolic health from the blood and identifying actionable targets for health interventions.

Methods

Study cohort

This study relied on a cohort consisting of over 5,000 individuals who participated in a wellness program offered by a currently closed commercial company (Arivale Inc., Washington, USA). We collected personal, dense, dynamic data (PD3) clouds on individuals in this program between 2015–2019. An individual was eligible for enrollment if the individual was over 18 years old, not pregnant, and a resident of any US state except New York; participants were primarily recruited from Washington, California, and Oregon. During the Arivale program, each participant was provided personalized lifestyle coaching via telephone by registered dietitians, certified nutritionists, or registered nurses (see the previous paper[22] for details). Participants also had access to their clinical data via an online data dashboard. This study was conducted with deidentified data of the participants who had consented to the use of their anonymized data in research. All procedures were approved by the Western Institutional Review Board (WIRB) with Institutional Review Board (IRB) (Study Number: 20170658 at Institute for Systems Biology and 1178906 at Arivale).

In this study, to confidently compare the association between Body Mass Index (BMI) and host phenotypes across multiomics, we limited our study cohort to only the participants whose datasets contained (1) all main omic measurements (metabolomics, proteomics, clinical lab tests) from the same first blood draw, (2) genetic information, and (3) a baseline BMI measurement within ±1.5 month from the first blood draw. We also eliminated "outlier" participants whose baseline BMI was beyond ±3 standard deviations from mean in the cohort distribution. The final cohort consists of 1,277 participants, whose demographics (Supplementary Fig. 1) were confirmed to be consistent with the overall Arivale demographics previously reported in our studies[19,22–26].

Data collection and PD3 clouds

The PD3 clouds consist of human genomes, longitudinal measurements of metabolomics, proteomics, clinical lab tests, gut microbiomes, and wearable devices, and health/lifestyle questionnaires. Peripheral venous blood draws for all measurements were performed by trained phlebotomists at LabCorp (Laboratory Corporation of America Holdings, North Carolina, USA) or Quest (Quest Diagnostics, New Jersey, USA) service centers. For some participants, saliva was also sampled at home to measure analytes such as diurnal cortisol and dehydroepiandrosterone (DHEA) using a standardized kit (ZRT Laboratory, Oregon, USA). Likewise, stool samples for gut microbiome measurements were obtained by participants at home using a standardized kit (DNA Genotek, Inc., Ottawa, Canada).

– Genome

DNA extracted from whole blood underwent whole genome sequencing or single-nucleotide polymorphisms (SNP) microarray genotyping. Genetic ancestry was calculated with principal components (PCs) using a set of ~100,000 ancestry-informative SNP markers as described previously[22]. Polygenic risk scores were constructed using publicly available summary statistics from published genome-wide association studies (GWAS) as described previously[24].

– Blood-measured omics

Metabolomic data was generated by Metabolon, Inc. (North Carolina, USA), using ultra-high-performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS) for plasma derived from whole blood samples. Proteomic data was generated using proximity extension assay (PEA) for plasma derived from whole blood samples with several Olink Target panels (Olink Proteomics, Uppsala, Sweden), and measurements with the Cardiovascular II, Cardiovascular III and Inflammation panels were used in this study since the other panels were not necessarily applied to all samples. All clinical laboratory tests were performed by LabCorp or Quest in a Clinical Laboratory Improvement Amendments (CLIA)-certified lab,

and measurements by LabCorp were selected in this study to eliminate potential differences between vendors. In this study, analytes missing in more than 10% of the baseline samples were removed from each omic dataset, and observations missing in more than 10% of the remaining analytes were further removed. The final filtered metabolomics, proteomics, and clinical labs consist of 766 metabolites, 274 proteins, 71 clinical lab tests, respectively (Supplementary Data 1).

– Gut microbiome

Gut microbiome data was generated based on 16S amplicon sequencing of the V4 region using a MiSeq sequencer (Illumina, Inc., California, USA) for DNA extracted from stool samples, as previously described[25]. All samples were first rarefied to an even sampling depth of 38,813, the minimum number of reads per sample in the cohort. Gut microbiome α-diversity was calculated at the amplicon sequence variant (ASV) level using Shannon's index calculated by $H = -\sum_{i=1}^{S} p_i \ln p_i$, where p_i is the proportion of species i to the total number of species S in a community represented by ASVs or using Chao1 diversity score calculated by $S_{\text{Chao1}} = S_{\text{obs}} + \frac{n_1^2}{2n_2}$, where S_{obs} is the number of observed ASVs, n_1 is the number of singletons (ASVs captured once), and n_2 is the number of doubletons (ASVs captured twice).

– Anthropometrics, saliva-measured analytes, daily physical activity measures

Anthropometrics including weight, height, and waist circumference and blood pressure were measured at the time of blood draw and also reported by participants, which generated diverse timing and number of observations depending on each participant. The measured weight and height simultaneously generated BMI measurements with dividing the weight by the squared height. Measurements of saliva samples were performed in the testing laboratory of ZRT Laboratory. Daily physical activity measures such as heart rate, moving distance, number of steps, burned calories, floors climbed, and sleep quality were tracked using the Fitbit wearable device (Fitbit, Inc., California, USA). To manage variations between days, monthly averaged data was used for these daily measures. In this study, the baseline measurement for these longitudinal measures was defined with the closest observation to the first blood draw per participant and data type, and each dataset was eliminated from analyses when its baseline measurement was beyond ±1.5 month from the first blood draw.

Generation of omics-based BMI models

For each omic dataset, missing values were first imputed with a random forest (RF) algorithm using Python missingpy (version 0.2.0) library corresponding to MissForrest in R[53]. For sex-stratified models, the datasets after imputation were divided into sex-dependent datasets. Each value was subsequently standardized with the *Z*-score using mean and standard deviation per analyte. Then, ten iterations of least absolute shrinkage and selection operator (LASSO) modeling with tenfold cross-validation (CV) were performed for the log-transformed BMI and the processed omic datasets using Python scikit-learn (version 0.22.1) library. Training and testing datasets were generated by splitting participants into ten sets with one set as a "testing dataset" and the remaining nine sets as a "training dataset", and iterating all combinations over those ten sets; i.e., overfitting was controlled using an internal tenfold CV in each "training dataset". Consequently, this procedure generated a "testing dataset"-derived BMI prediction value for each participant and ten fitted models for each omics. Model performance was conservatively estimated by the R^2 score from out-of-sample predictions. Pearson's *r* was calculated using measured and predicted BMI values for the entire cohort.

For the LASSO-modeling iterations while dropping the strongest variable, the generation of ten LASSO models was repeated as the same as the above except for eliminating the strongest

variable analyte from the dataset at the end of each iteration. The strongest variable was defined as the variable that was retained across ten models and had the highest absolute value of mean β-coefficient.

For longitudinal models, the standardization distribution and training dataset were restricted to the baseline measurements from all 1,277 participants and only one LASSO model with tenfold CV was generated per sex-stratified cohort, because those measurements were minimally affected by lifestyle coaching and each participant had a different number of observations.

Generation of obesity-classifying models

To compare the ability of the gut microbiome to accurately distinguish between obese and normal weight individuals across different omics measurements of BMI (Fig. 4c, d), ASVs were collapsed into species, genus, family, order, class, and phylum, respectively, and merged into a single dataframe using the SILVA database (version 132). This dataframe served as input for a RF classifier implemented with Python scikit-learn library. Briefly, two classifiers were trained on taxon abundances, one predicting whether an individual is obese based on observed BMI, and one predicting whether an individual is obese based on metabolomics-based BMI. Both models were constructed using RF with a fivefold CV scheme. In this approach, 80% of the data is used for training while the remaining 20% is used as a testing set. This process is repeated fivefold where each participant serves as part of the testing set once. Performance for each of the classifiers was then assessed by averaging the performance across the five testing sets.

Longitudinal analysis for BMI and omics-based BMIs

For longitudinal analyses, 608 participants, whose datasets contained more than two time-series datasets of both BMI and omics during 18 months after enrollment, were further extracted from the study cohort of 1,277 participants. To estimate the mean transition of the measured and omics-based BMIs, a linear mixed model (LMM) for the rate of change in each measured BMI or sex-stratified LASSO models-predicted BMI was generated with random intercepts for participants and random slopes for days in the program, following the previous approach[22]. As the fixed effects regarding time, linear regression splines with knots at 0, 6, 12, and 18 months were applied to days in program to fit time as a continuous variable rather than a categorical variable because the timing of data collection was different between the participants, allowing for differences in the trajectory of changes throughout the program. In addition to the linear regression splines for days in the program, each LMM included sex, baseline age, ancestry PCs, and meteorological seasons as fixed effects to adjust potential confounding effects. For the baseline BMI class-stratified LMMs, the interaction terms between the categorical baseline BMI-based class and the linear regression splines for time were further added. All LMMs were modeled using Python statsmodels (version 0.11.1) library. Of note, the underweight participants were eliminated in the LMMs stratified with baseline BMI class because the sample size was too small for convergence.

Plasma analyte correlation network analysis

In advance, outlier values which were beyond ±3 standard deviations from mean in the longitudinal cohort distribution of 608 participants were eliminated from the dataset per analyte, and seven clinical lab tests which became almost invariant across the participants were eliminated from analyses, allowing convergence in the following modeling. Against each analyte, values were converted with a transformation method producing the lowest skewness (e.g., no transformation, the logarithm transformation for right skewed distribution, the square root transformation with mirroring for left skewed distribution) and standardized with the *Z*-score using mean and standard deviation.

Against 608,856 pairwise combinations of the analytes (766 metabolites, 274 proteomics, 64 clinical lab tests), generalized linear models (GLMs) for the baseline measurements of 608 participants were independently generated with the Gaussian distribution and identity link function

using Python statsmodels library. Each GLM constitutes of an analyte as dependent variable, another analyte and the baseline MetBMI as independent variables with their interaction term, and sex, baseline age, and ancestry PCs as covariates. The significant analyte–analyte correlation pairs modified by the baseline MetBMI were obtained based on the β-coefficient (two-sided t-test) of the interaction term between independent variables in GLM, while correcting the multiple-hypothesis testing with the Benjamini–Hochberg method (false discovery rate (FDR) < 0.05).

Against the significant 91 pairs (75 metabolites, 26 proteomics, 13 clinical lab tests) from the GLM analysis, generalized estimating equations (GEEs) for the longitudinal measurements of 184 metabolically obese participants were independently generated with the exchangeable covariance structure using Python statsmodels library. Each GEE constitutes of an analyte as dependent variable, another analyte and days in the program as independent variables with their interaction term, and sex, baseline age, ancestry PCs, and meteorological seasons as covariates. The significant analyte–analyte correlation pairs modified by days in the program were obtained based on the β-coefficient (two-sided t-test) of the interaction term between independent variables in GEE, while correcting the multiple-hypothesis testing with the Benjamini–Hochberg method (FDR < 0.05).

Statistical analysis

All data preprocessing and ordinary least squares (OLS) regression analyses were performed using Python NumPy (version 1.18.1), pandas (version 1.0.3), SciPy (version 1.4.1) and statsmodels libraries. Only the baseline datasets were utilized in regression analyses, and each numeric variable was scaled and centered in advance. When assessing available numeric physiological features and obesity-related health markers in the PD3 clouds, the baseline dataset of each metric variable was also preprocessed with the elimination steps for outliers and invariant variables and the conversion step for skewness reduction, as same as those described in the above subsection except for the basis of whole study cohort distribution. Relationships of the preprocessed numeric physiological features with the measured or omics-based BMIs (Fig. 1c) were independently assessed using OLS linear regression with the log-transformed measured or omics-based BMI as dependent variable and sex, age, and ancestry PCs as covariates, while correcting the multiple-hypothesis testing with the Benjamini–Hochberg method (FDR < 0.05). Relationships between BMI and analytes which were retained in at least one of ten LASSO models (210 metabolites, 75 proteins, 42 clinical lab tests) (Fig. 2b–d) were independently assessed using OLS linear regression with the log-transformed BMI as dependent variable and sex, age, and ancestry PCs as covariates, while correcting the multiple-hypothesis testing with the Benjamini–Hochberg method (FDR < 0.05). Differences in the BMI-associated features and obesity-related health markers between the matched and misclassified groups in the normal or obese BMI class (Fig. 3c and Supplementary Fig. 5a) were independently assessed using OLS linear regression with sex, age, and ancestry PCs as covariates (sex and ancestry PCs as covariates for the regression of age). Misclassification distribution in hierarchical clustering (Figs. 3d, e and Supplementary Fig. 5b, c) was assessed using Fisher's exact tests with the Bonferroni correction (family-wise error rate (FWER) < 0.05). Relationships between measured or omics-based BMI and α-diversity metrics (Fig. 4a, b) were independently assessed using OLS linear regression with α-diversity as dependent variable and sex, age, and ancestry PCs as covariates. Difference in classifier performance parameters (Fig. 4d) was assessed using Student's t-test. All statistical tests were performed using a two-sided hypothesis.

Data visualization

Almost all results were visualized using Python matplotlib (version 3.2.1) and seaborn (version 0.10.1) libraries. Data were summarized as the mean ± standard error of the mean (s.e.m.), the mean with 95% confidence interval (CI), or the boxplot (median: center line, 95% CI around median: notch, [Q_1, Q_3]: box limits, [max(minimum value, $Q_1 - 1.5 \times$ IQR), min(maximum value, $Q_3 + 1.5 \times$ IQR)]: whiskers, where Q_1, Q_3, and IQR are the 1st quartile, the 3rd quartile, and the interquartile range, respectively), as indicated in each figure legend. For presentation purpose, s.e.m. and CI were

simultaneously calculated during visualization using seaborn barplot or boxplot (utilizing matplotlib) application programming interface (API) with default setting (1,000 times bootstrapping or a Gaussian-based asymptotic approximation, respectively). The OLS linear regression line with 95% CI was simultaneously generated during visualization using seaborn lmplot API with default setting (1,000 times bootstrapping). Hierarchical clustering was simultaneously performed during visualization using seaborn clustermap API (utilizing SciPy library) with the Ward's linkage method for Euclidean distance. The plasma analyte correlation network was visualized with a circos plot using R circlize (version 0.4.11) package[54].

References

1. NCD Risk Factor Collaboration (NCD-RisC). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants. *Lancet (London, England)* **387**, 1377–1396 (2016).
2. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet (London, England)* **390**, 2627–2642 (2017).
3. Kopelman, P. G. Obesity as a medical problem. *Nature* **404**, 635–43 (2000).
4. Haslam, D. W. & James, W. P. T. Obesity. *Lancet (London, England)* **366**, 1197–209 (2005).
5. Kahn, S. E., Hull, R. L. & Utzschneider, K. M. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* **444**, 840–6 (2006).
6. Van Gaal, L. F., Mertens, I. L. & De Block, C. E. Mechanisms linking obesity with cardiovascular disease. *Nature* **444**, 875–80 (2006).
7. Magkos, F. *et al.* Effects of Moderate and Subsequent Progressive Weight Loss on Metabolic Function and Adipose Tissue Biology in Humans with Obesity. *Cell Metab.* **23**, 591–601 (2016).
8. Hamman, R. F. *et al.* Effect of weight loss with lifestyle intervention on risk of diabetes. *Diabetes Care* **29**, 2102–7 (2006).
9. Sun, Q. *et al.* Comparison of dual-energy x-ray absorptiometric and anthropometric measures of adiposity in relation to adiposity-related biologic factors. *Am. J. Epidemiol.* **172**, 1442–54 (2010).
10. Prentice, A. M. & Jebb, S. A. Beyond body mass index. *Obes. Rev.* **2**, 141–7 (2001).
11. Okorodudu, D. O. *et al.* Diagnostic performance of body mass index to identify obesity as defined by body adiposity: A systematic review and meta-analysis. *Int. J. Obes.* **34**, 791–799 (2010).
12. Ruderman, N., Chisholm, D., Pi-Sunyer, X. & Schneider, S. The metabolically obese, normal-weight individual revisited. *Diabetes* **47**, 699–713 (1998).
13. Ding, C., Chan, Z. & Magkos, F. Lean, but not healthy: the 'metabolically obese, normal-weight' phenotype. *Curr. Opin. Clin. Nutr. Metab. Care* **19**, 408–417 (2016).
14. Smith, G. I., Mittendorfer, B. & Klein, S. Metabolically healthy obesity: facts and fantasies. *J. Clin. Invest.* **129**, 3978–3989 (2019).
15. Appleton, S. L. *et al.* Diabetes and cardiovascular disease outcomes in the metabolically healthy obese phenotype: a cohort study. *Diabetes Care* **36**, 2388–94 (2013).
16. Schröder, H. *et al.* Determinants of the transition from a cardiometabolic normal to abnormal overweight/obese phenotype in a Spanish population. *Eur. J. Nutr.* **53**, 1345–53 (2014).
17. Williams, S. A. *et al.* Plasma protein patterns as comprehensive indicators of health. *Nat. Med.* **25**, 1851–1857 (2019).
18. Bar, N. *et al.* A reference map of potential determinants for the human serum metabolome. *Nature* **588**, 135–140 (2020).
19. Wilmanski, T. *et al.* Blood metabolome predicts gut microbiome α-diversity in humans. *Nat. Biotechnol.* **37**, 1217–1228 (2019).
20. Cirulli, E. T. *et al.* Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk. *Cell Metab.* **29**, 488-500.e2 (2019).
21. Price, N. D. *et al.* A wellness study of 108 individuals using personal, dense, dynamic data clouds. *Nat. Biotechnol.* **35**, 747–756 (2017).
22. Zubair, N. *et al.* Genetic Predisposition Impacts Clinical Changes in a Lifestyle Coaching Program. *Sci. Rep.* **9**, 6805 (2019).
23. Earls, J. C. *et al.* Multi-Omic Biological Age Estimation and Its Correlation With Wellness and Disease Phenotypes: A Longitudinal Study of 3,558 Individuals. *J. Gerontol. A. Biol. Sci. Med. Sci.* **74**, S52–S60 (2019).
24. Wainberg, M. *et al.* Multiomic blood correlates of genetic risk identify presymptomatic disease alterations. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 21813–21820 (2020).
25. Wilmanski, T. *et al.* Gut microbiome pattern reflects healthy ageing and predicts survival in humans. *Nat. Metab.* **3**, 274–286 (2021).
26. Zimmer, A. *et al.* The geometry of clinical labs and wellness states from deeply phenotyped humans. *Nat. Commun.* **12**, 3578 (2021).
27. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet (London, England)* **363**, 157–63 (2004).

676 28. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B* **58**, 267–288 (1996).
677 29. Xu, X. *et al.* Habitual sleep duration and sleep duration variation are independently associated with body
678 mass index. *Int. J. Obes. (Lond).* **42**, 794–800 (2018).
679 30. Shah, N. R. & Braverman, E. R. Measuring adiposity in patients: the utility of body mass index (BMI),
680 percent body fat, and leptin. *PLoS One* **7**, e33308 (2012).
681 31. Tomiyama, A. J., Hunger, J. M., Nguyen-Cuu, J. & Wells, C. Misclassification of cardiometabolic health
682 when using body mass index categories in NHANES 2005-2012. *Int. J. Obes.* **40**, 883–886 (2016).
683 32. Nimptsch, K., Konigorski, S. & Pischon, T. Diagnosis of obesity and use of obesity biomarkers in science
684 and clinical medicine. *Metabolism.* **92**, 61–70 (2019).
685 33. Bennett, C. M., Guo, M. & Dharmage, S. C. HbA(1c) as a screening tool for detection of Type 2 diabetes: a
686 systematic review. *Diabet. Med.* **24**, 333–43 (2007).
687 34. Pereira-Santos, M., Costa, P. R. F., Assis, A. M. O., Santos, C. A. S. T. & Santos, D. B. Obesity and vitamin
688 D deficiency: a systematic review and meta-analysis. *Obes. Rev.* **16**, 341–9 (2015).
689 35. Ridaura, V. K. *et al.* Gut microbiota from twins discordant for obesity modulate metabolism in mice.
690 *Science (80-.).* **341**, (2013).
691 36. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484 (2009).
692 37. Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**,
693 541–546 (2013).
694 38. Walters, W. A., Xu, Z. & Knight, R. Meta-analyses of human gut microbes associated with obesity and IBD.
695 *FEBS Lett.* **588**, 4223–4233 (2014).
696 39. Duvallet, C., Gibbons, S. M., Gurry, T., Irizarry, R. A. & Alm, E. J. Meta-analysis of gut microbiome
697 studies identifies disease-specific and shared responses. *Nat. Commun.* **8**, 1784 (2017).
698 40. Karetnikova, E. S. *et al.* Is Homoarginine a Protective Cardiovascular Risk Factor? *Arterioscler. Thromb.*
699 *Vasc. Biol.* **39**, 869–875 (2019).
700 41. Dieuleveux, V., Lemarinier, S. & Guéguen, M. Antimicrobial spectrum and target site of D-3-phenyllactic
701 acid. *Int. J. Food Microbiol.* **40**, 177–83 (1998).
702 42. Beloborodova, N. *et al.* Effect of phenolic acids of microbial origin on production of reactive oxygen
703 species in mitochondria and neutrophils. *J. Biomed. Sci.* **19**, 89 (2012).
704 43. Li, Y. *et al.* Adrenomedullin is a novel adipokine: adrenomedullin in adipocytes and adipose tissues.
705 *Peptides* **28**, 1129–43 (2007).
706 44. Egaña-Gorroño, L. *et al.* Receptor for Advanced Glycation End Products (RAGE) and Mechanisms and
707 Therapeutic Opportunities in Diabetes and Cardiovascular Disease: Insights From Human Subjects and
708 Animal Models. *Front. Cardiovasc. Med.* **7**, 37 (2020).
709 45. Norata, G. D. *et al.* Circulating soluble receptor for advanced glycation end products is inversely associated
710 with body mass index and waist/hip ratio in the general population. *Nutr. Metab. Cardiovasc. Dis.* **19**, 129–
711 34 (2009).
712 46. Rauschert, S., Uhl, O., Koletzko, B. & Hellmuth, C. Metabolomic biomarkers for obesity in humans: A
713 short review. *Ann. Nutr. Metab.* **64**, 314–324 (2014).
714 47. Rangel-Huerta, O. D., Pastor-Villaescusa, B. & Gil, A. Are we close to defining a metabolomic signature of
715 human obesity? A systematic review of metabolomics studies. *Metabolomics* **15**, 93 (2019).
716 48. Barber, M. N. *et al.* Plasma lysophosphatidylcholine levels are reduced in obesity and type 2 diabetes. *PLoS*
717 *One* **7**, e41456 (2012).
718 49. Piening, B. D. *et al.* Integrative Personal Omics Profiles during Periods of Weight Gain and Loss. *Cell Syst.*
719 **6**, 157-170.e8 (2018).
720 50. Wing, R. R. & Phelan, S. Long-term weight loss maintenance. *Am. J. Clin. Nutr.* **82**, 222S-225S (2005).
721 51. Li, G. *et al.* The long-term effect of lifestyle interventions to prevent diabetes in the China Da Qing Diabetes
722 Prevention Study: a 20-year follow-up study. *Lancet* **371**, 1783–1789 (2008).
723 52. Diabetes Prevention Program Research Group. 10-year follow-up of diabetes incidence and weight loss in
724 the Diabetes Prevention Program Outcomes Study. *Lancet* **374**, 1677–1686 (2009).
725 53. Stekhoven, D. J. & Bühlmann, P. Missforest-Non-parametric missing value imputation for mixed-type data.
726 *Bioinformatics* **28**, 112–118 (2012).
727 54. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. Circlize implements and enhances circular visualization
728 in R. *Bioinformatics* **30**, 2811–2812 (2014).
729

Acknowledgements

We thank Sergey A. Kornilov, Gustavo Glusman, and Max Robinson (Institute for Systems Biology; ISB) for providing comments to the manuscript. We are grateful to all Arivale participants who consented to using their deidentified data for research purposes. This work was supported by the M.J. Murdock Charitable Trust (Reference No. 2014096:MNL:11/20/2014, awarded to ISB), Arivale, and a generous gift from C. Ellison (K.W., T.W., and A.Z.). K.W. was supported by The Uehara Memorial Foundation (Overseas Postdoctoral Fellowships). C.D. was supported by the Washington Research Foundation Distinguished Investigator Award and startup funds from ISB.

Author Contribution

K.W., T.W., L.H., N.D.P., and N.R. conceptualized the study. K.W., T.W., A.Z., N.D.P., and N.R. participated in the study design. K.W., T.W., C.D., B.L., and N.R. performed data analysis and figure generation. J.J.H., J.C.L., C.D., A.T.M., and L.H. assisted in results interpretation. J.C.L. and A.T.M. managed the logistics of data collection and integration. K.W., T.W., and N.R. were the primary authors of the paper, with contributions from all other authors. All authors read and approved the final manuscript.

Competing Interests

J.J.H. has received grants from Pfizer and Novartis for research unrelated to this study. All other authors declare no competing interests.

Figures

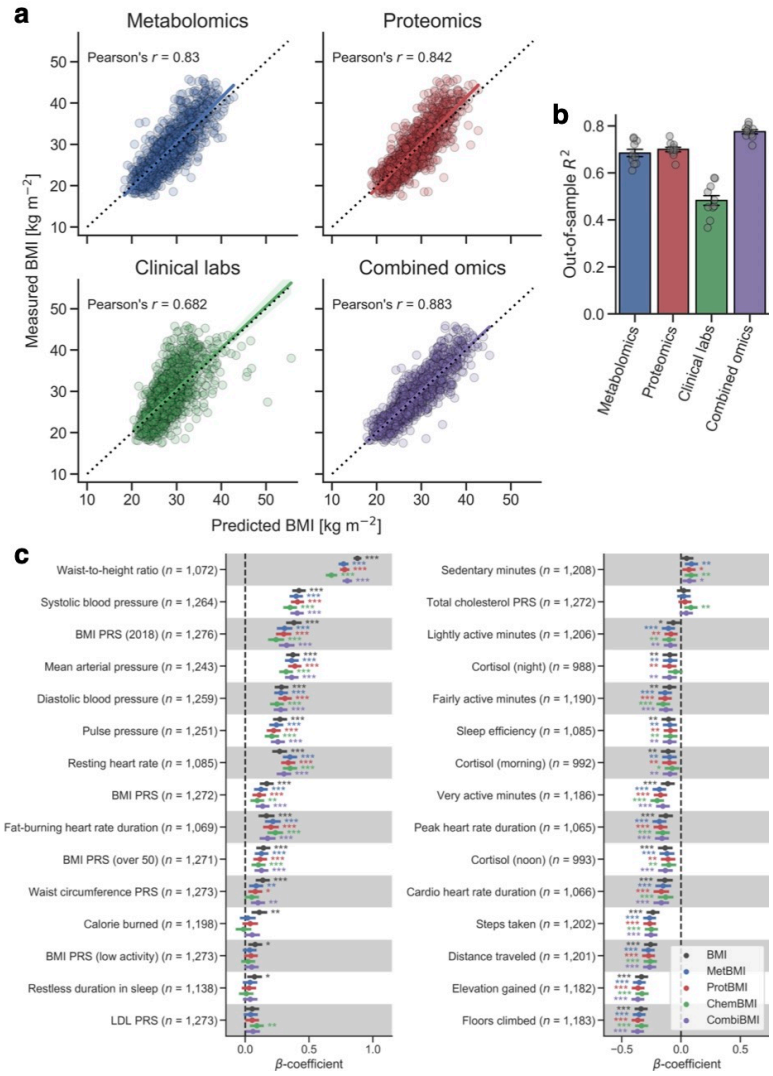


Figure 1. Plasma multiomics captures 48–78% of the variance in BMI.

a Scatterplot of measured Body Mass Index (BMI) versus predicted BMI using least absolute shrinkage and selection operator (LASSO) with tenfold cross-validation (CV). The solid line in each panel is the ordinary least squares (OLS) linear regression line with 95% confidence interval (CI), and the dotted line is measured BMI = predicted BMI. n = 1,277 participants. **b** Mean out-of-sample R^2 across the tenfold CV for each omics. Data: mean ± s.e.m., n = 10 LASSO models. **c** β-coefficients for numeric physiological feature in each OLS linear regression model with BMI or omics-based BMI as dependent variable and sex, age, and ancestry principal components (PCs) as covariates. All presented 30 features are significantly associated with at least one of BMI or omics-based BMIs in the Benjamini–Hochberg method (false discovery rate (FDR) < 0.05; *P < 0.05, **P < 0.01, ***P < 0.001). BMI: measured BMI, MetBMI: metabolomics-based BMI, ProtBMI: proteomics-based BMI, ChemBMI: clinical chemistries-based BMI, CombiBMI: combined omics-based BMI, PRS: polygenic risk score, LDL: low-density lipoprotein, n: the number of participants. Data: β-coefficient with 95% CI.

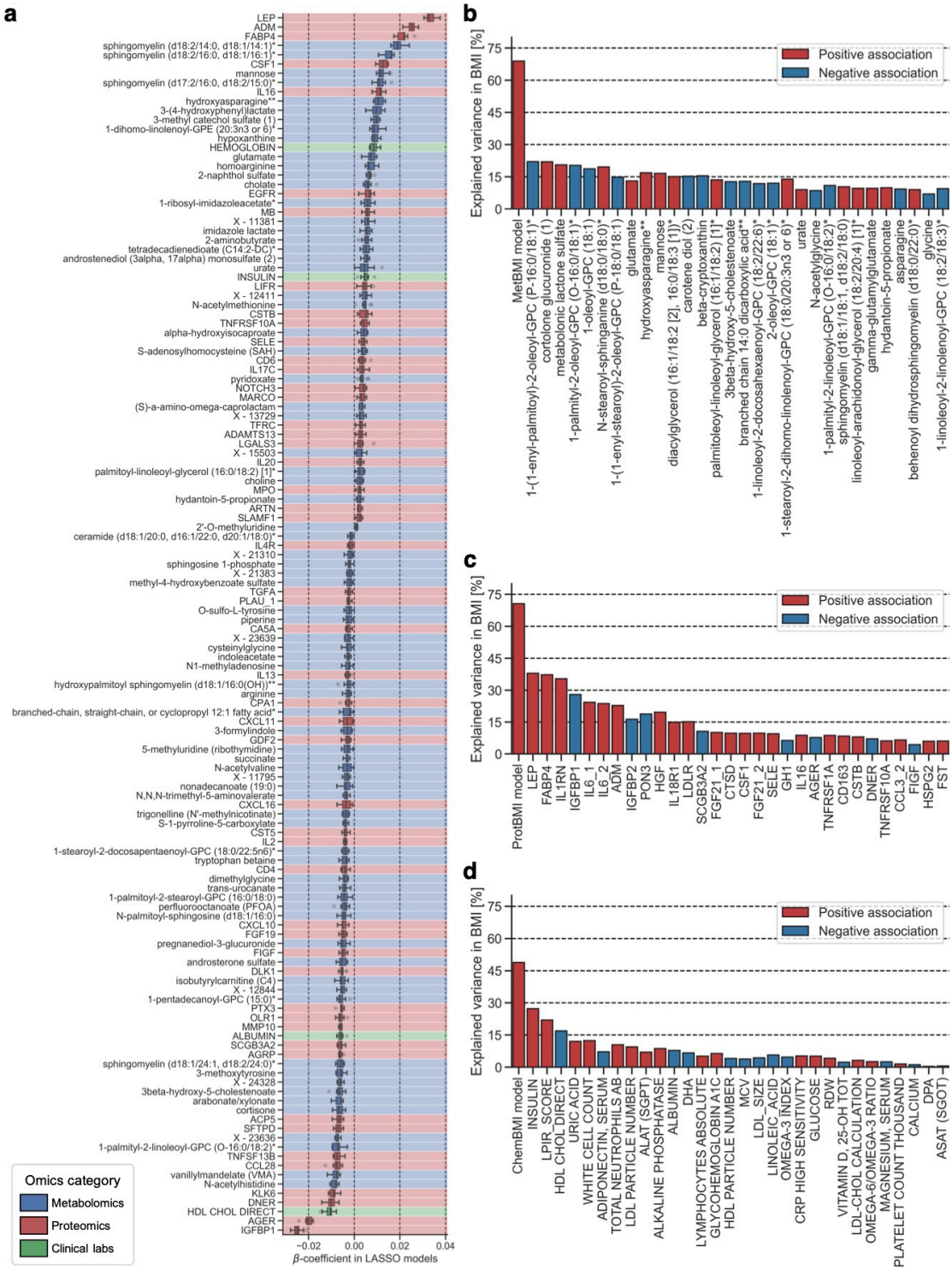
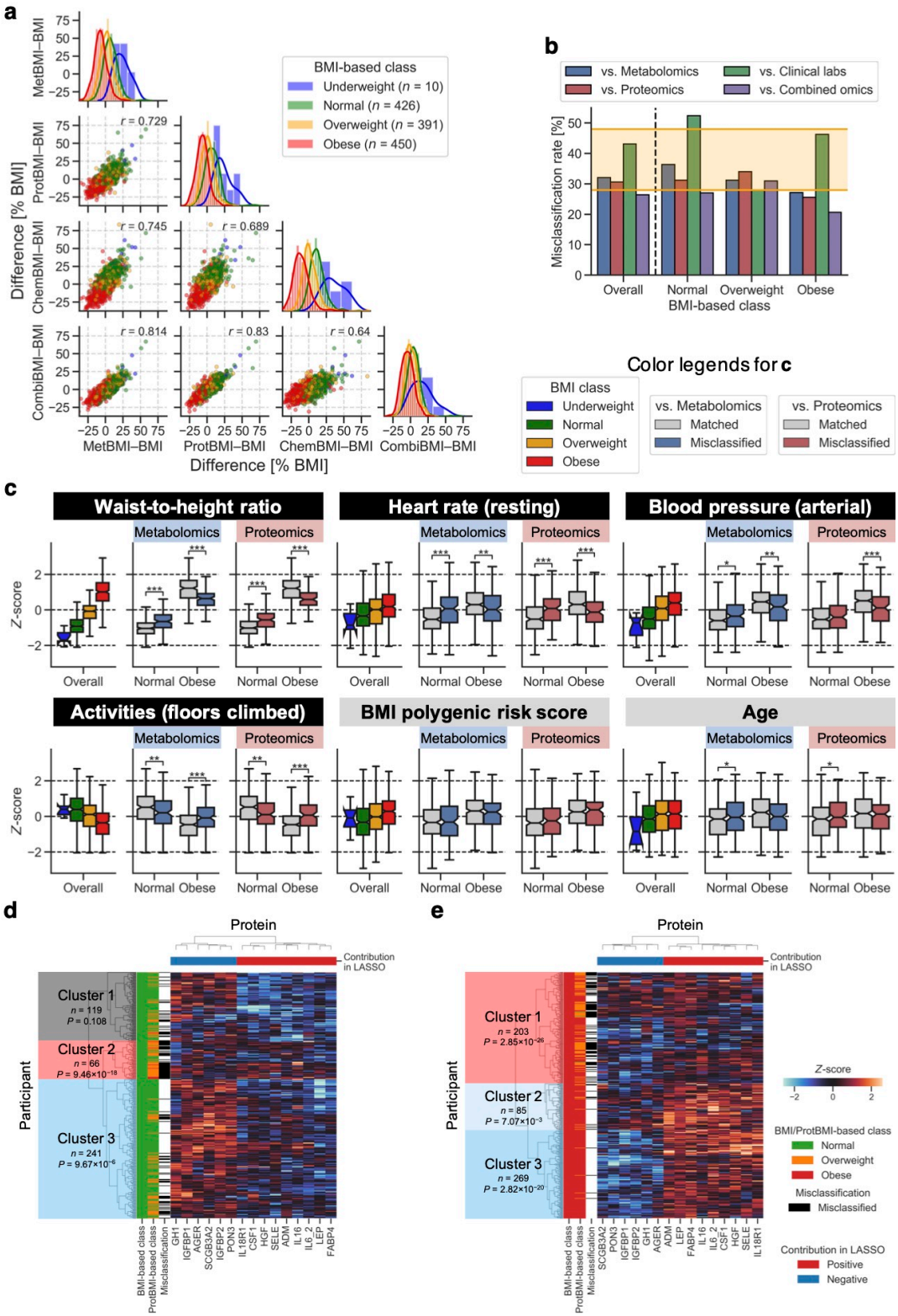


Figure 2. Omics-based BMI captures the variance in BMI better than any single analyte.

a β-coefficient estimates for the variables that were retained across all ten combined omics-based Body Mass Index (BMI) models (132 analytes). Color of each row corresponds to the analyte category (blue: 77 metabolites, red: 51 proteins, green: 4 clinical lab tests). Data: median (center line), [Q_1, Q_3] (box limits), [max(minimum value, $Q_1 - 1.5 \times \text{IQR}$), min(maximum value, $Q_3 + 1.5 \times \text{IQR}$)] (whiskers), where Q_1, Q_3, and IQR are the 1st quartile, the 3rd quartile, and the interquartile range,

respectively; n = 10 least absolute shrinkage and selection operator (LASSO) models. **b**–**d** Percentage of variance in BMI explained by each metabolite (**b**), protein (**c**), or clinical lab test (**d**). BMI was independently regressed to each analyte which was retained in at least one of ten LASSO models (210 metabolites, 75 proteins, 42 clinical lab tests). The strongest 30 analytes among the analytes significantly associated with BMI (180 metabolites, 63 proteins, 30 clinical lab tests) are presented. Significance was assessed using ordinary least squares (OLS) linear regression with sex, age, and ancestry principal components (PCs) as covariates, while correcting for multiple-hypothesis testing with the Benjamini–Hochberg method (false discovery rate (FDR) < 0.05). Each omics-based BMI is included for comparison (MetBMI: metabolomics-based BMI, ProtBMI: proteomics-based BMI, ChemBMI: clinical chemistries-based BMI).



786

787 **Figure 3. Metabolic heterogeneity within standard BMI classes underlies the high rate of**
788 **misclassification.**

a Scatterplot and distribution of difference between Body Mass Index (BMI) and omics-based BMI. BMI: measured BMI, MetBMI: metabolomics-based BMI, ProtBMI: proteomics-based BMI, ChemBMI: clinical chemistries-based BMI, CombiBMI: combined omics-based BMI, r: Pearson's correlation coefficient, n: the number of participants. The line in histogram panels indicates the kernel density estimate. **b** Misclassification rate of BMI-based class. Range of previously reported misclassification rate[30,31] is highlighted with orange-colored lines. Note that the underweight BMI class is not presented due to small sample size, and its misclassification rate is 80% against combined omics and 100% against the others. **c** Comparison of BMI-associated feature between the matched and misclassified groups in the normal or obese BMI class. Data: median (center line), 95% confidence interval (CI) around median (notch), [Q_1, Q_3] (box limits), [max(minimum value, $Q_1 - 1.5 \times \mathrm{IQR}$), min(maximum value, $Q_3 + 1.5 \times \mathrm{IQR}$)] (whiskers), where Q_1, Q_3, and IQR are the 1st quartile, the 3rd quartile, and the interquartile range, respectively. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ according to ordinary least squares (OLS) linear regression with sex, age, and ancestry principal components (PCs) as covariates (sex and ancestry PCs as covariates for the regression of age). **d**, **e** Heatmap with hierarchical clustering of the normal (**d**) and obese (**e**) BMI class using proteomics data. The strongest 15 proteins among the analytes retained across more than eight ProtBMI models and significantly associated with BMI were used as variables. Z-score was calculated from the overall population. n: the number of participants, P: the adjusted P-value in multiple-hypothesis testing for the misclassification distribution in each cluster using two-sided Fisher's exact tests with the Bonferroni correction.

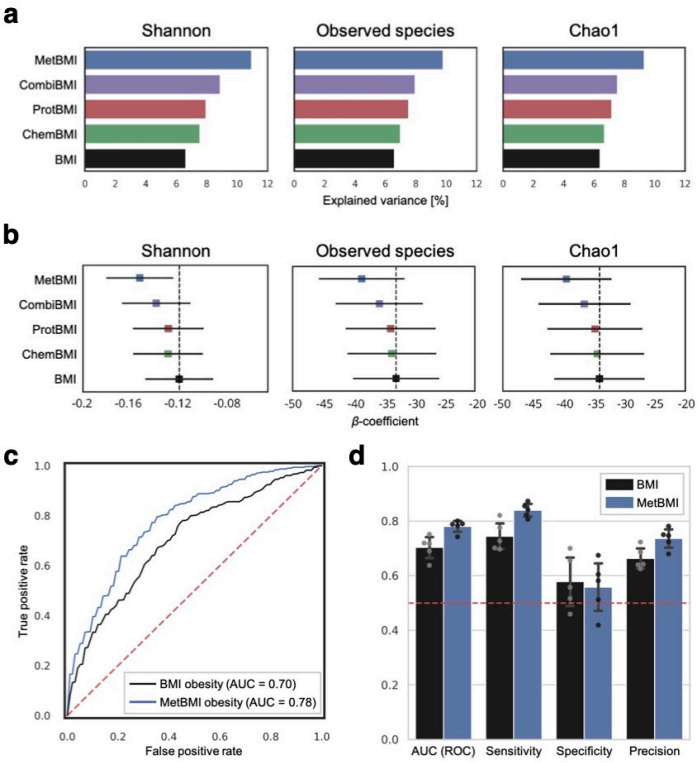


Figure 4. Metabolomics-based BMI reflects gut microbiome profiles better than BMI.

a Percentage of variance in gut microbiome α-diversity that is explained by each Body Mass Index (BMI) or omics-based BMI. BMI: measured BMI, MetBMI: metabolomics-based BMI, ProtBMI: proteomics-based BMI, ChemBMI: clinical chemistries-based BMI, CombiBMI: combined omics-based BMI. **b** β-coefficient for BMI or omics-based BMI in each ordinary least squares (OLS) linear regression model with α-diversity as dependent variable and sex, age, and ancestry principal components (PCs) as covariates. The dashed line indicates the β-coefficient estimate for BMI. Data: β-coefficient with 95% confidence interval (CI). **c** Receiver operator characteristic (ROC) curve of a gut microbiome-based model classifying participants to the normal vs. obese class. Gut microbiome 16S ribosomal RNA datasets were used for generating the random forest (RF) classifier with fivefold cross-validation (CV). Each ROC curve indicates the average curve across five RF models. The red dashed line indicates a random classification line. AUC: area under curve. **d** Comparison of AUC of ROC curve, sensitivity, specificity, and precision between the classifying models of BMI and MetBMI. Each performance parameter was calculated as the mean out-of-sample value across the fivefold CV. Data: mean with 95% CI, n = 5 RF models.

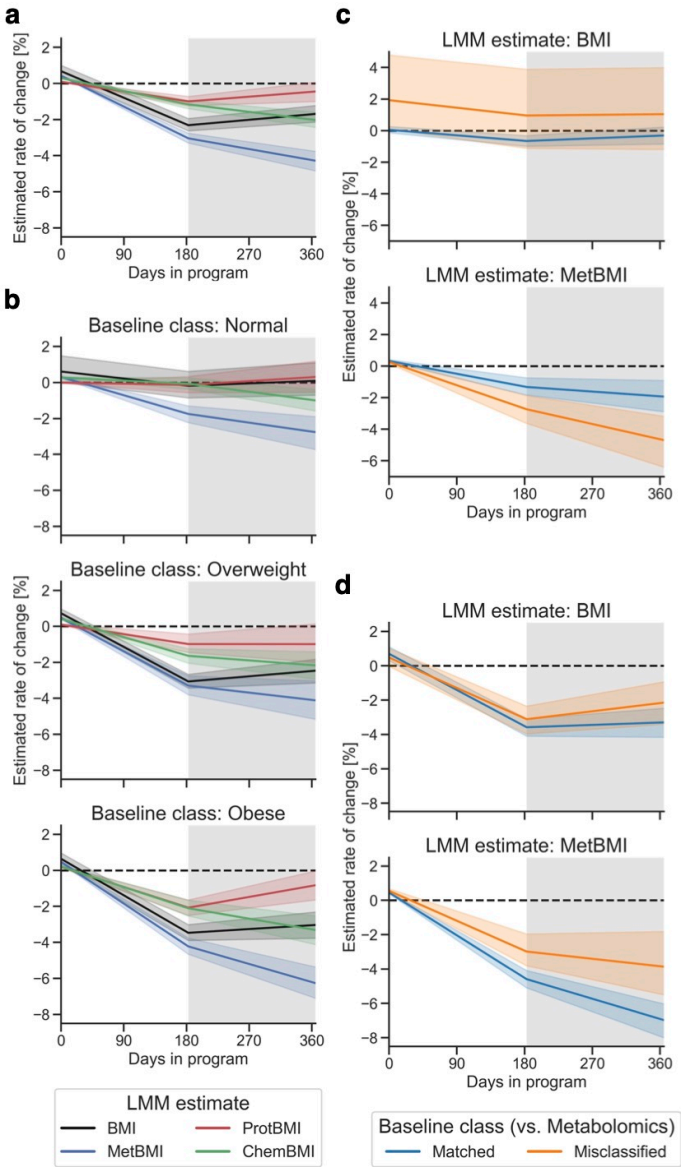


Figure 5. Metabolic health of the metabolically obese group was substantially improved following a positive lifestyle intervention.

a Longitudinal change in Body Mass Index (BMI) or omics-based BMI for overall cohort. Rate of change in BMI and omics-based BMIs was estimated using each linear mixed model (LMM) with random intercepts for participants and random slopes for days in the program (see Methods). n = 608 participants. **b** Longitudinal change in BMI or omics-based BMI for each baseline BMI-based class. Rate of change in BMI and omics-based BMIs was estimated using each baseline BMI-based class-stratified LMM with random intercepts for participants and random slopes for days in the program. n = 222 (Normal), 185 (Overweight), 196 (Obese) participants. **c**, **d** Longitudinal change in BMI or metabolomics-based BMI of the participants misclassified with the normal (**c**) or obese (**d**) BMI class. n = 156 (Normal, Matched), 66 (Normal, Misclassified), 151 (Obese, Matched), 45 (Obese, Misclassified) participants. **a**–**d** The dashed line and gray shading correspond to the baseline value of each estimate and the 2nd period of linear regression spline for time, respectively. BMI: measured BMI, MetBMI: metabolomics-based BMI, ProtBMI: proteomics-based BMI, ChemBMI: clinical chemistries-based BMI. Data: mean with 95% confidence interval (CI).

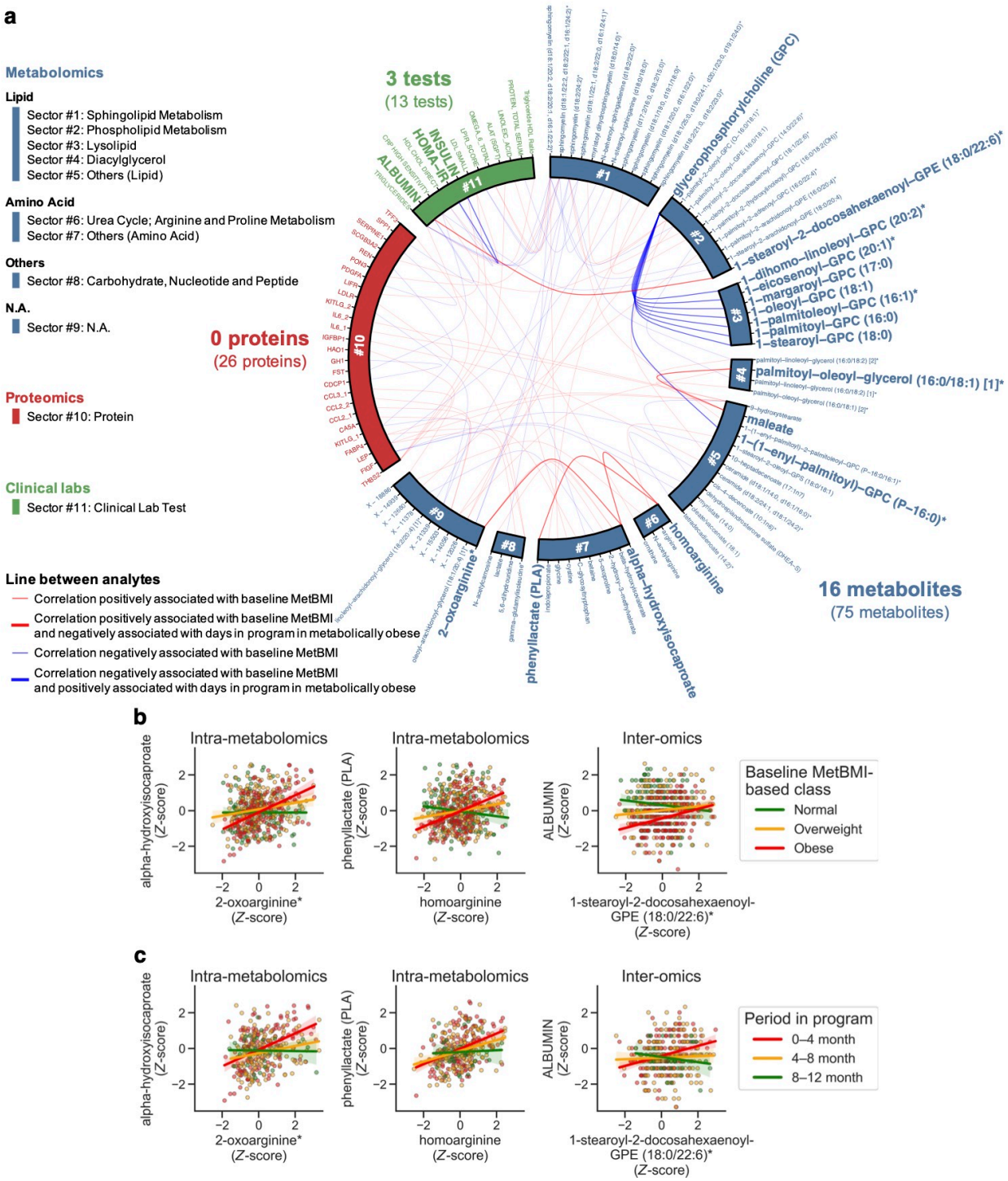


Figure 6. Plasma analyte correlation network in the metabolically obese group reverted back to normal state following lifestyle intervention.

a Circos plot of cross-omic interactions modified by metabolomics-based Body Mass Index (MetBMI) and days in the program. Among 608,856 pairwise relationships of plasma analytes (766 metabolites, 274 proteomics, 64 clinical lab tests) from 608 participants, 91 analyte–analyte pairs significantly modified by the baseline MetBMI are presented (75 metabolites, 26 proteomics, 13 clinical lab tests), whose significance was assessed using their interaction term in each generalized linear model (GLM;

see Methods) while correcting the multiple-hypothesis testing with the Benjamini–Hochberg method (false discovery rate (FDR) < 0.05). Among the significant 91 pairs from 184 metabolically obese participants, 14 analyte–analyte pairs significantly modified by days in the program are highlighted by line width and label font size (16 metabolites, 3 clinical lab tests), whose significance was assessed using their interaction term in each generalized estimating equation (GEE; see Methods) while correcting the multiple-hypothesis testing with the Benjamini–Hochberg method (FDR < 0.05). **b**, **c** Representative examples of the analyte–analyte pair significantly modified by the baseline MetBMI (**b**) or days in the program (**c**) in **a**. The solid line in each panel is the ordinary least squares (OLS) linear regression line with 95% confidence interval (CI). n = 530 (**b**, Intra-metabolomics (left)), 553 (**b**, Intra-metabolomics (right)), 566 (**b**, Inter-omics) participants; n = 329 (**c**, Intra-metabolomics (left)), 344 (**c**, Intra-metabolomics (right)), 353 (**c**, Inter-omics) measurements from 184 metabolically obese participants. Of note, data points outside of plot range are trimmed in these presentations.