













High-throughput mediation analysis of human proteome and metabolome identifies mediators of post-bariatric surgical diabetes control

Jonathan M. Dreyfuss ^{1,2,3,18}, Yixing Yuchi ^{3,4,11,18}, Xuehong Dong^{3,4,12}, Vissarion Efthymiou ^{3,4}, Hui Pan ^{1,2}, Donald C. Simonson ^{3,5}, Ashley Vernon^{3,6}, Florencia Halperin^{3,5,13}, Pratik Aryal ^{3,7}, Anish Konkar^{8,14}, Yinong Sebastian⁸, Brandon W. Higgs^{8,15}, Joseph Grimsby^{8,16}, Cristina M. Rondinone⁸, Simon Kasif ², Barbara B. Kahn ^{3,7}, Kathleen Foster⁴, Randy Seeley ⁹, Allison Goldfine ^{3,4,5,17}, Vera Djordjilović ¹⁰ & Mary Elizabeth Patti ^{3,4}✉

To improve the power of mediation in high-throughput studies, here we introduce High-throughput mediation analysis (Hitman), which accounts for direction of mediation and applies empirical Bayesian linear modeling. We apply Hitman in a retrospective, exploratory analysis of the SLIMM-T2D clinical trial in which participants with type 2 diabetes were randomized to Roux-en-Y gastric bypass (RYGB) or nonsurgical diabetes/weight management, and fasting plasma proteome and metabolome were assayed up to 3 years. RYGB caused greater improvement in HbA1c, which was mediated by growth hormone receptor (GHR). GHR's mediation is more significant than clinical mediators, including BMI. GHR decreases at 3 months postoperatively alongside increased insulin-like growth factor binding proteins IGFBP1/BP2; plasma GH increased at 1 year. Experimental validation indicates (1) hepatic GHR expression decreases in post-bariatric rats; (2) GHR knockdown in primary hepatocytes decreases gluconeogenic gene expression and glucose production. Thus, RYGB may induce resistance to diabetogenic effects of GH signaling.

Trial Registration: Clinicaltrials.gov NCT01073020.

¹Bioinformatics and Biostatistics Core, Research Division, Joslin Diabetes Center, Boston, MA, USA. ²Biomedical Engineering, Boston University, Boston, MA, USA. ³Harvard Medical School, Boston, MA, USA. ⁴Integrative Physiology and Metabolism, Research Division, Joslin Diabetes Center, Boston, MA, USA. ⁵Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA. ⁶Department of Surgery, Brigham and Women's Hospital, Boston, MA, USA. ⁷Division of Endocrinology, Diabetes & Metabolism, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA. ⁸MedImmune, Gaithersburg, MD, USA. ⁹Department of Surgery, University of Michigan, Ann Arbor, MI, USA. ¹⁰Department of Economics, Ca' Foscari University of Venice, Venice, Italy. ¹¹Present address: Vertex Pharmaceuticals, Boston, MA, USA. ¹²Present address: Department of Endocrinology, Diabetes & Metabolism, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China. ¹³Present address: Form Health, Boston, MA, USA. ¹⁴Present address: Eli Lilly and Company, Indianapolis, IN, USA. ¹⁵Present address: Genmab, Plainsboro, NJ, USA. ¹⁶Present address: AstraZeneca, Gaithersburg, MD, USA. ¹⁷Present address: Novartis Institute for Biomedical Research, Cambridge, MA, USA. ¹⁸These authors contributed equally: Jonathan M. Dreyfuss, Yixing Yuchi.

✉email: MaryElizabeth.Patti@joslin.harvard.edu

Given the major public health and personal burden of T2D, the identification of new approaches to therapy for both T2D and obesity is critical. Increasing evidence supports that surgical approaches to T2D are effective for both long-term weight loss and improved glucose metabolism, resulting in remission of T2D in $\approx 90\%$ of patients at 1 year and 45% at 5 years^{1–5}, and that these effects are superior to traditional nonsurgical diabetes management. Surgery's beneficial effects have been attributed to weight loss and improved insulin sensitivity, but improved glycemic control occurs within days after RYGB, before substantial weight loss, supporting an important role for weight-independent mechanisms⁶. Additional metabolic effects of surgery which may contribute to improved glycemic outcomes include increased postprandial insulin and incretin hormone secretion^{7,8}, reduction in circulating amino acids^{9,10}, alterations in bile acids, FXR signaling and FGF19^{11–15}, and changes in microbiome metabolism^{16–18}, but preclinical studies indicate these cannot fully explain clinical improvement^{19–21}. Thus, the primary molecular factors mediating improved metabolism and remission of T2D in response to surgical procedures remain uncertain.

To identify mediators of metabolic improvement after surgical vs. medical therapy for T2D, we perform a retrospective exploratory analysis of the SLIMM-T2D trial, in which participants with type 2 diabetes were randomized to Roux-en-Y gastric bypass (RYGB) or nonsurgical diabetes weight management (DWM), and fasting plasma proteome and metabolome were assayed up to 3 years. Greater clinical improvement was seen in RYGB than in DWM for multiple outcomes, including reductions in body weight (assessed by BMI), glycemia (assessed by HbA1c), triglycerides, HDL cholesterol, blood pressure, and reductions in the number of antidiabetic, antihypertensive, and lipid-lowering medications⁵. Moreover, all RYGB participants achieved 10% weight loss before 3 months, whereas only 37% of DWM participants achieved 10% weight loss before 3 months. Given that participants in SLIMM-T2D were randomized to either RYGB or DWM, we can infer causal effects of RYGB (relative to the DWM control) on the outcomes at, for example, one year, and then ask if any clinical outcome or analyte measured at 3 months mediates this causal effect. Identifying mediators among approximately 2000 measured analytes would be a key step to develop new nonsurgical approaches to weight loss and glycemic control in T2D. Many mediation methods are available, but they have been found to lack power in high-throughput studies²². A previous comprehensive comparison of mediation methods concluded that the best balance between false positive rate and power was offered by the joint significance method²³. The joint significance method²³ in our example assesses if two tests are jointly significant for each analyte: (I) is RYGB associated with analyte abundance at 3 months? and (II) is analyte abundance at 3 months associated with HbA1c at one year given RYGB? The joint significance method was also recommended by another comparative study²², where a bootstrap-based method that assesses the product of effects I and II²⁴ performed only slightly worse than the joint significance test. However, the bootstrap-based method could not be used in their genome-wide application, since it was too computationally expensive to perform sufficient numbers of bootstraps to reach genome-wide significance²². The joint significance method was also shown to control its false-positive rate in theory and to be more powerful than the commonly-used product significance method for mediation²⁵. Another popular mediation test not included in these comparative simulations is based on the potential outcomes framework²⁶. This framework supports very general causal models, but also assesses significance with the bootstrap or other resampling methods, and so is also computationally burdensome for high-throughput data.

Here we report the development of a mediation method termed High-throughput mediation analysis (Hitman) and apply this method to proteomic and metabolomic data from the SLIMM-T2D randomized trial. We identify and validate the growth hormone receptor (GHR) as a mediator of the metabolic impact of RYGB (Fig. 1).

Results

Clinical characteristics of participants. Of the 38 SLIMM-T2D participants, proteomic and metabolomic data at baseline, the 3 month time point, and at 1 year were available from 35 participants. Metabolic characteristics of these 35 participants did not differ between groups at baseline (Supplementary Data 1), and HbA1c and BMI of those with proteomics or metabolomics measurements did not differ from the SLIMM-T2D participants within their group at any time point (Supplementary Fig. 1). After 3 years, no DWM participants achieved study-defined glycemic goals (HbA1c < 6.5% and fasting glucose < 126 mg/dL), whereas eight RYGB participants did, and seven of these eight participants were not receiving any anti-diabetes medications⁵. The metabolic characteristics per participant per time point are tabulated in our Zenodo repository, which also includes participants' diabetes medication use per medication class per time point.

Longitudinal analysis of the plasma proteome. Fasting proteomics were profiled at baseline, the 3 month time point, and years 1, 2, and 3 in RYGB and DWM. At baseline (pre-randomization), there were no statistically significant proteins (FDR < 0.15). For post-randomization time points, reduction in BMI was greater in RYGB, so groups were compared both with and without adjustment for each person's BMI change. Differences in the baseline-corrected proteome in RYGB vs. DWM (i.e., differences between groups in changes from baseline) emerged at the 3 month time point, with 14 significant proteins in the unadjusted analysis and 8 in the BMI-adjusted analysis. Differences of baseline-corrected protein abundance in RYGB vs. DWM persisted at years 1, 2, and 3. The baseline-corrected log₂ fold change values for the 19 proteins differentially abundant with fold-change absolute value >1.5 at any time point without BMI adjustment are depicted in the heatmap in Fig. 2a. Statistics for all comparisons in all analytes are presented in Supplementary Data 2.

We observed several proteins with important roles in systemic metabolism to be significant in unadjusted analyses at several time points, but never significant with BMI adjustment. We considered such analytes as potentially regulated by weight-dependent mechanisms. For example, proteins upregulated post-baseline included sex hormone binding globulin (SHBG), upregulated at all post-baseline time points, and adiponectin, upregulated at 12, 24, and 36 months. Conversely, GHR was downregulated at all post-baseline time points; leptin was downregulated at 12 and 24 months, while C-reactive protein and insulin were downregulated at 12 and 36 months. Multiple complement factor proteins were downregulated at 12, 24, and 36 months, including complement factor H (CFH), C1s, and C5.

Several proteins remained significant after BMI adjustment, indicating potential weight-independence. These include increases in (1) bone structural proteins integrin-binding sialoprotein (IBSP), which was increased at all post-baseline time points in the unadjusted analysis and at 3 and 12 months in the BMI-adjusted analysis, and osteomodulin (OMD), which was increased with and without BMI adjustment at 12, 24, and 36 months, (2) fibroblast growth factor 19 (FGF19), the intestinally-derived target of bile acid and FXR signaling¹⁵, which was increased with and without BMI adjustment at 36 months, and (3) insulin-like growth factor binding protein IGFBP2,

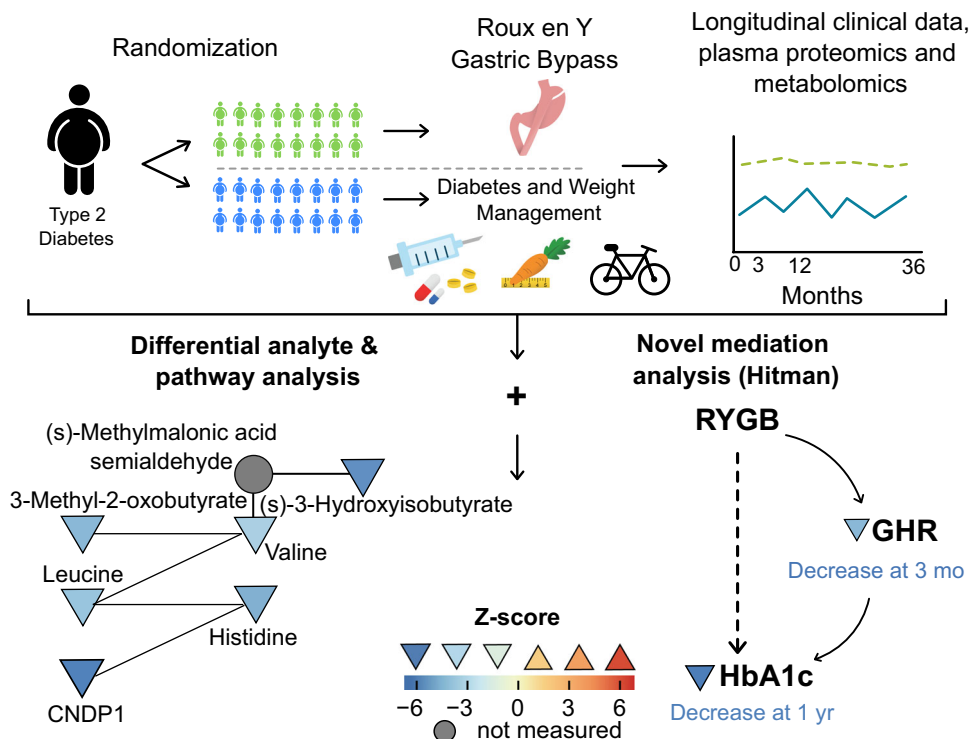


Fig. 1 Graphical overview shows study flow including clinical study schematic (top) and analysis (bottom). Differential analyte and pathway analysis (bottom left) identified a network composed of top analytes from Valine, Leucine, and Isoleucine Degradation and Beta-Alanine Metabolism. Mediation analysis (bottom right) identified GHR as mediator of decrease in HbA1c. Triangles denote measured proteins or metabolites, while circles denote proteins or metabolites in the pathway not measured; symbols are colored according to between-group Z scores, using heat map indicated, with the orientation of triangle also indicating up- or downregulation, respectively. Source data are provided as a Source Data file.

recently shown to be increased after RYGB^{27–29}, potentially via improved leptin sensitivity³⁰, was increased at all post-baseline time points in the unadjusted analysis and at the 3 month time point in the BMI-adjusted analysis. Of the complement factors, only complement factor B (CFB) remained significant after BMI adjustment, at 12, 24, and 36 months.

Upregulation of IGFBP2 (>50%, FDR < 5% at all post-baseline time points in SOMAscan, Supplementary Fig. 1a) was confirmed by ELISA, which showed significant differences in baseline-corrected values ($p < 0.05$) at 12, 24, and 36 months (Supplementary Fig. 2b), and ELISA changes per person over time were significantly correlated to corresponding SOMAscan changes ($r = 0.78$, $p < 10^{-7}$).

Longitudinal analysis of the plasma metabolome. At baseline (pre-randomization), there were no significant differences in the fasting metabolome between groups. Significant differences in the baseline-corrected metabolome emerged at the 3 month time point, with 96 metabolites differentially abundant in the unadjusted analysis and 74 in the BMI-adjusted analysis; 45 metabolites were common to both analyses.

Differences between arms of baseline-corrected metabolite abundance persisted at years 1, 2, and 3. The baseline-corrected \log_2 fold change values in RYGB vs. DWM for the 85 metabolites differentially abundant with fold-change absolute value >1.5 at any time point without BMI adjustment are depicted in the heatmap in Fig. 2b. Although the color scale is the same as that in Fig. 2a to permit comparison, fold changes were much greater for the metabolome. For example, the BCAA-related metabolite 3-hydroxyisobutyrate was 88% lower in RYGB (i.e., down by >9-fold) with significance in both the unadjusted and BMI-adjusted analysis at the 3 month time point.

Prolylhydroxyproline, a marker of bone collagen degradation, was increased in RYGB at the 3, 12, and 36 month time points, and was also increased with BMI adjustment at the 3 and 12 month time points, again suggesting that postoperative bone metabolic changes are independent of weight loss. The abundance of metformin was reduced with or without BMI adjustment at 12 months, in agreement with reduced medication use in the RYGB group. Interestingly, increases in the conjugated bile acid glycochenodeoxycholate sulfate emerged with time, with significance at the 12 month time point in unadjusted analysis and both the 12 and 36 month time point in the BMI-adjusted analysis.

Having both metabolites and proteins on the same subjects is an advantage of our study, so we tested the correlation of top metabolites' vs. top proteins' baseline-corrected changes at 3 months (Supplementary Data 2). Multiple analytes were highly correlated. For example, the metabolite most correlated to GHR was prolylhydroxyproline ($r = -0.68$, $p < 10^{-5}$, FDR < 0.002); prolylhydroxyproline was also negatively correlated with CNDP1 (beta-ala-his dipeptidase or carnosine dipeptidase 1; $r = -0.65$, $p < 10^{-4}$, FDR < 0.004) and positively correlated with IGFBP2 ($r = 0.69$, $p < 10^{-5}$, FDR < 0.002). GHR and CNDP1 were positively correlated to the BCAA-related metabolites beta-hydroxyisovalerate and isovalerylcarnitine, and these were negatively correlated with IGFBP2 (FDR < 0.08 for all these correlations).

Proteomic and metabolomic integrative pathway analysis. We integrated proteomics and metabolomics for pathway analysis by creating a single, integrated dataset, and testing pathways from the Small Molecule Pathway Database³¹, which has pathways composed of both proteins and metabolites. We identified 50

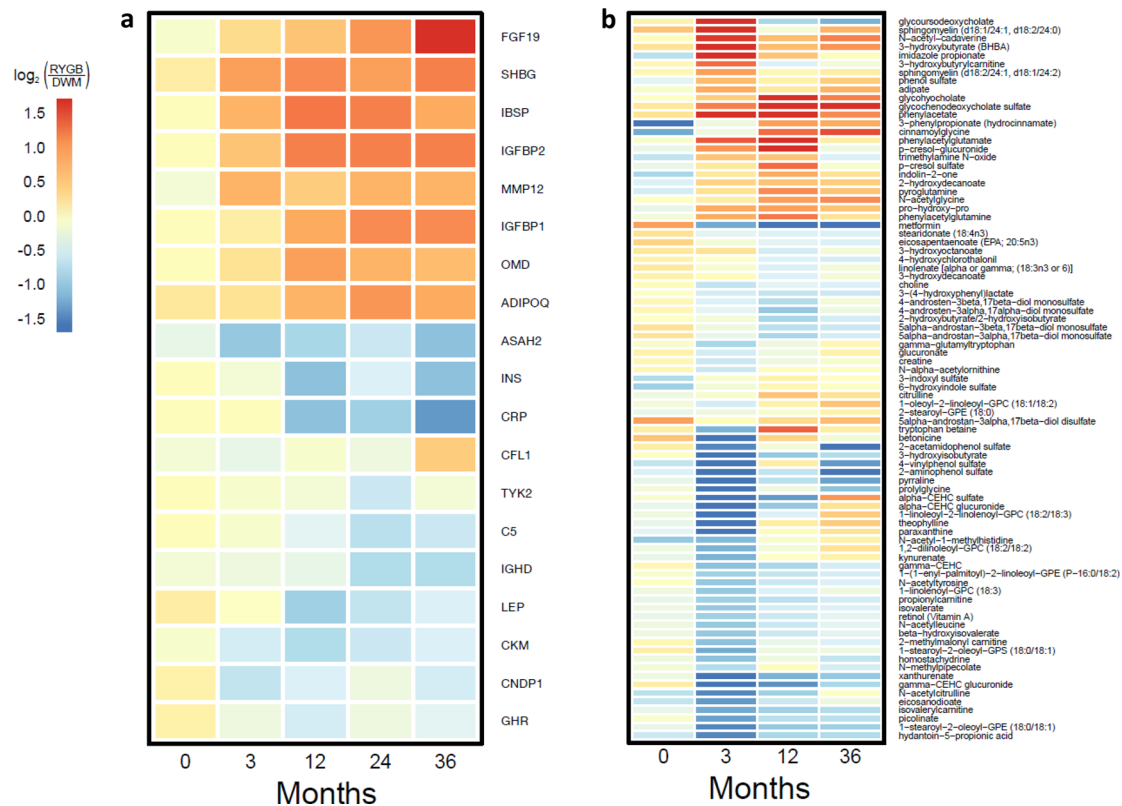


Fig. 2 Differentially abundant analytes with time. Heatmaps of **a** proteins and **b** metabolites. Heatmap of $\log_2(\text{RYGB}/\text{DWM})$ at all time points (post-baseline \log_2 abundance values are baseline-corrected) for analytes that are differentially abundant at any time point (FDR < 0.15 and fold-change absolute value > 1.5). The range of colors show $\log_2(\text{RYGB}/\text{DWM})$ from -1.6 to 1.6, which corresponds to fold changes from -3 to 3. Analytes with a larger absolute value of fold-change are shown as only having fold change absolute value of 3, so that weaker fold changes are easily observed. Source data are provided as a Source Data file.

differentially abundant pathways (FDR < 15%) between groups without BMI adjustment at the 3 month time point.

The top-ranking pathways are presented in Fig. 3a. The top-ranking pathway is Phospholipid Biosynthesis (Fig. 3b), whose top analytes were choline (Supplementary Fig. 2c) and choline phosphate (Supplementary Fig. 2d), both reduced in RYGB relative to DWM at the 3 month time point. Choline was also reduced at 12 and 36 months, and the 3 month time point with BMI adjustment, but choline phosphate was not significant in other comparisons.

The second pathway was Valine, Leucine, and Isoleucine Degradation (schematic in Fig. 4a). Its top analytes (Fig. 4b), reduced at the 3 month time point in both the unadjusted and BMI-adjusted analysis, were the BCAAs valine and leucine, the ketoacid 3-methyl-2-oxobutyrate (ketoisovalerate), and the valine catabolic intermediate 3-hydroxyisobutyrate. All but 3-hydroxyisobutyrate were also reduced at 12 months; with BMI adjustment, none were significant at 12 or 36 months, suggesting an important contribution of weight loss during the early postoperative period.

Similar patterns were observed for other ketoacids and downstream acylcarnitines. For example, propionylcarnitine, a C3 acylcarnitine product of valine and isoleucine metabolism, was reduced in RYGB at the 3 month time point in both the unadjusted and BMI-adjusted analyses (but not at later time points with or without BMI adjustment) and was also the top analyte of the related pathway Oxidation of Branched Chain Fatty Acids.

Beta-alanine Metabolism's (Fig. 5a) and Histidine Metabolism's top analytes were CNDP1 and its enzymatic product histidine. Histidine was reduced after RYGB at the 3 month time point with

or without BMI adjustment but not later (Fig. 5b), and its reduction was correlated with that of CNDP1 ($r = 0.40$, $p = 0.02$). CNDP1 was reduced by 43% after RYGB at the 3 month time point and was reduced at 12 months, and at the 3 month time point with BMI adjustment (Fig. 5c). A previous study was unable to validate SOMAscan CNDP1 measurements³². We measured CNDP1 with ELISA and our ELISA-determined CNDP1 levels showed little correlation with SOMAscan levels per time point, but changes within person over time showed stronger correlation ($r = 0.21$, $p = 0.055$). We confirmed significant reductions in CNDP1 after RYGB, demonstrating a 68% decrease in baseline-corrected change for RYGB vs. DWM ($p = 0.004$ at the 3 month time point, Fig. 5d).

Retinol Metabolism's top analyte was retinol (vitamin A), whose baseline-corrected abundance decreased by 33% in RYGB at the 3 month time point, when it was also decreased with BMI adjustment, but over time retinol reverted towards baseline in both groups (Supplementary Fig. 3a). Changes in retinol were significantly correlated to changes in retinol-binding protein 4 (RBP4), shown in Supplementary Fig. 3b, over all time points ($r = 0.33$, $p = 0.001$). Given the role of RBP4 as a mediator of insulin resistance³³, we measured RBP4 and its partner transthyretin (TTR) in a random subset of 12 subjects by quantitative western blot. From these western blots, RBP4 and TTR were not different between-arms at the 3 month time point (Supplementary Fig. 3c, d), but RBP4 changes (i.e., differences per person over time) correlated significantly to RBP4 changes by SOMAscan ($r = 0.28$, $p = 0.03$), and changes in both RBP4 and TTR significantly correlated to each other ($r = 0.63$, $p < 10^{-5}$) and to changes in retinol (both $r > 0.6$, $p < 10^{-4}$).

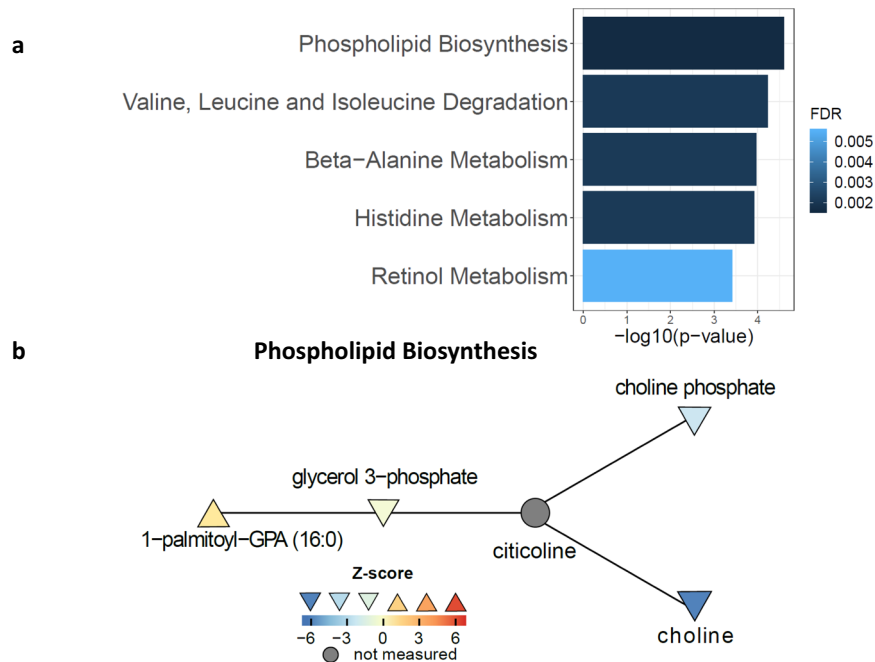


Fig. 3 Phospholipid biosynthesis and other top-ranking differential pathways. **a** Top-ranking differential pathways. Graph illustrates $-\log_{10}(p\text{-values})$ and false discovery rates (FDRs), with bars colored by FDR, from Limma Roast's Mixed statistic⁸⁷ (which is one-sided) adjusted with FDRs. **b** Phospholipid biosynthesis network. Nodes are colored by between-group z-score, whereas unmeasured nodes are colored dark gray. Orientation of triangle also indicates the directionality of regulation. Connections are from the Pathway Commons network. Source data are provided as a Source Data file.

Development of high-throughput mediation analysis (Hitman). We sought to identify analytes whose change in abundance at the postoperative 3 month time point mediated RYGB's improvement in diabetes control at one year. We know that RYGB improves diabetes outcomes from prior studies, so we can be confident that RYGB actually improves diabetes, as we observe, rather than the opposite direction of effect (in which case RYGB would exacerbate diabetes). A reasonable mediation method to use would be the joint significance method. However, one drawback of the joint significance test and other mediation methods when the direction of the causal effect of the exposure on the outcome is known is that they do not account for the direction of effect of the mediator, so they could call a significant an "inconsistent" mediator, which suppresses or inhibits the causal effect (Fig. 6)³⁴. For example, relative to DWM, RYGB decreases the baseline-corrected abundance of retinol at the 3 month time point, and RYGB decreases baseline-corrected 1 year HbA1c (Fig. 6, left side). However, a decrease in retinol is associated with an increase in HbA1c—a direction of mediation inconsistent with RYGB decreasing HbA1c, so retinol does not appear to mediate RYGB's beneficial effect.

A second drawback of the joint significance and other mediation methods is that they are intended for testing only one or a few mediators. Thus they do not employ powerful statistical methods for high-throughput data, such as linear regression modeling with empirical Bayesian modeling of an analyte's variance with the R package Limma³⁵, which has been validated in multiple omics platforms, and is particularly powerful for small sample sizes.

To address these limitations, we developed high-throughput mediation analysis (Hitman). Hitman is designed for studies where the exposure causes a significant change in the outcome in the same direction as known a priori and high-throughput mediators are assayed. Hitman improves power by accounting for the direction of effect and by employing empirical Bayesian linear modeling³⁵. If a mediator's direction of effect is consistent with

the prespecified direction of the causal effect of the exposure on the outcome, its p -value from the joint significance test is halved. Otherwise, its p -value becomes one. In the case of the inconsistent mediator retinol, Hitman assigns a p -value of one. The description of the algorithm, implemented in our R package *Hitman*, mathematical proof that Hitman provides valid p -values when there is a causal effect whose direction is known, and details of our simulations and their results are provided in Supplementary Note 1.

Using simulations similar to those of Barfield et al.³⁶ and MacKinnon et al.²³, we compared Hitman against the joint significance test and the potential outcome mediation test²⁶, which had not been included in previous simulations. Consistent with our mathematical proof, our simulation using 50 samples demonstrates that Hitman properly controls its false positive rate when the causal effect's direction is known and specified prior to the analysis. Our simulations also test the ability of the methods to detect consistent mediators, and confirm that Hitman has more power. For example, when the true causal mediation effects are of a size termed "small"²³, the joint significance and the potential outcome mediation test identify mediation in <2.5% of simulations, whereas Hitman identifies mediators in >6%, which is significantly greater than both (odds ratio (OR) = 2.89, 95% CI = [2.48, 3.39], $p < 10^{-15}$). When the true effect is of "medium" size²³, the proportion of simulations that identified mediation by the joint significance test and the potential outcome mediation test were 53.9 and 57.4%, respectively, whereas Hitman identifies 68.5% (OR = 1.62, 95% CI = [1.52, 1.71], $p < 10^{-15}$).

The directional nature of Hitman accounts for much but not all of Hitman's power, since the empirical Bayesian variance estimation adds power, particularly with smaller sample sizes. Smaller sample sizes are common in basic research with model organisms, where variance between units in a group is expected to be small. Using Limma³⁵ also allows for testing high-throughput data types that are not normally distributed, such as RNA-seq³⁷. We demonstrate the benefit of using Limma by simulating as

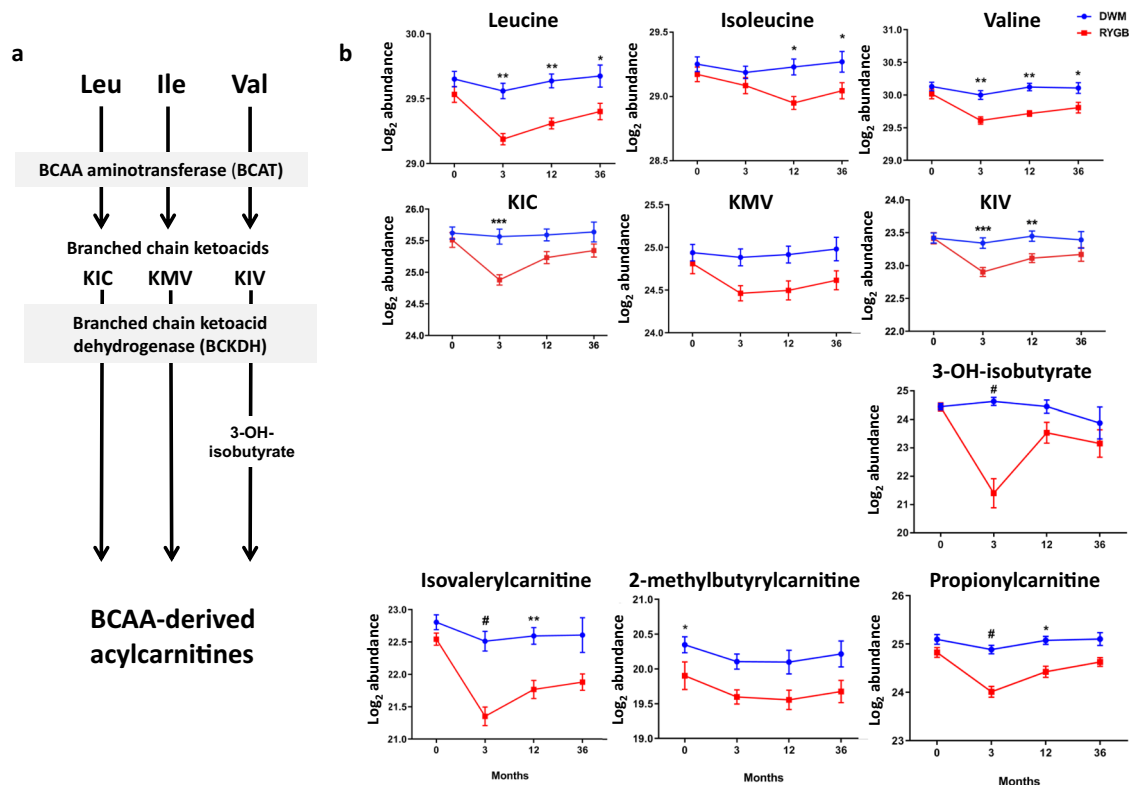


Fig. 4 Branched chain amino acids (BCAA) and downstream metabolites. **a** Schematic of BCAA metabolic pathway showing measured analytes. KIC, KMV, and KIV indicate the branched chain ketoacids ketoisocaproate, ketomethylvalerate, and ketoisovalerate, respectively. **b** Relative abundance of measured analytes, aligned to position in pathway in (a). Data were analyzed by two-sided moderated t-tests; post-baseline time points were analyzed using change from baseline. Data are reported as mean \pm SEM on the log₂ scale. Analysis was derived from samples from independent human participants. RYGB at 0, 3, 12, and 36 months: $n = 19, 18, 19, 13$; DWM: $n = 19, 18, 16, 9$. Nominal p -values are as follows: Leucine: 3 mo ** = 0.001, 12 mo ** = 0.0041, 36 mo * = 0.03; Isoleucine: 12 mo * = 0.017, 36 mo * = 0.036; Valine: 3 mo ** = 0.0019, 12 mo ** = 0.0013, 36 mo * = 0.016; KIC: 3 mo *** = 0.00016; KIV: 3 mo: *** = 0.00026, 12 mo: ** = 0.0045. 3-OH-isobutyrate: 3 mo: # < 0.0001; isovalerylcarnitine: 3 mo: # < 0.0001, 12 mo: ** = 0.002; 2-methylbutyrylcarnitine: 0 mo: * = 0.025; propionylcarnitine: 3 mo: # < 0.0001, 12 mo: * = 0.016. isobutyrylcarnitine: 0 mo: * = 0.024. Source data and FDRs are available in Supplementary Data 2.

above but with only 15 samples. We also compare these simulations against a simplified version of Hitman that does not use Limma, which is applicable to scenarios other than high-throughput data. We find with medium effect size that the joint significance and the potential outcome method have power <7%, the simplified version of Hitman that does not use Limma has power 11.8%, and Hitman has the power of 16.1%, which is significantly greater than the others (OR = 1.43, 95% CI = [1.32, 1.55], $p < 10^{-15}$).

We next conducted omics simulations using datasets of 500 analytes where p -values were adjusted with the false discovery rate (FDR). We simulated sample sizes of 15 and 50, either 1, 5, or 25 consistent mediators, and 200 analytes associated with the exposure but not the outcome in three scenarios: (1) analytes are independent and mediators are consistent; (2) analytes are dependent on each other and mediators are consistent; and (3) analytes are dependent on each other and there are an equal number of consistent and inconsistent mediators of equal effect. We compared the joint significance method, Hitman, and the version of Hitman that does not use Limma. We did not include the potential outcome mediation test in these simulations because of its computational cost. We modeled dependence between analytes based on the dependence structure from the GTEx expression dataset³⁸, which was adjusted for measured and inferred³⁹ covariates, like the processing by Oliva et al.⁴⁰.

We found that power increased with more samples and fewer mediators, whether these mediators were consistent or inconsistent, and that including dependence from GTEx had little effect

on the results, most likely because its covariates were properly accounted for. We also found that Hitman controls its FDR in all scenarios and has the most power, with the greatest advantage when $N = 15$. For example, in scenario 1 when $N = 15$ and there was one true mediator, the joint significance method had power of 31.5%, whereas Hitman had power of 42.4% (OR = 1.6, 95% CI = [1.33, 1.93], $p = 5 * 10^{-7}$).

Mediation analysis: analytes, clinical data, and pathways. We applied Hitman to identify analytes whose early change (baseline to 3 month time point) mediates HbA1c improvement at one year (Supplementary Data 2). The only significant analyte was growth hormone receptor (GHR; $p = 10^{-4}$, FDR = 0.12; Fig. 6, right), which was reduced by 24% in RYGB at the 3 month time point. We also tested the change in proteins and metabolites at the 3 month time point vs HbA1c change at the 18, 24, and 36 month time points. However, we found no significant protein or metabolite mediators of HbA1c at these later time points, potentially because after 12 months the number of participants with HbA1c measurements declines substantially, or there are additional secondary changes, such as progressive weight loss or inter-individual variation in diet or other factors (Supplementary Data 1).

As a comparison to Hitman, we applied the joint significance method to test mediation of proteins and metabolites whose change at the 3 month time point mediates HbA1c improvement at one year. We found similar top analytes, but with weaker

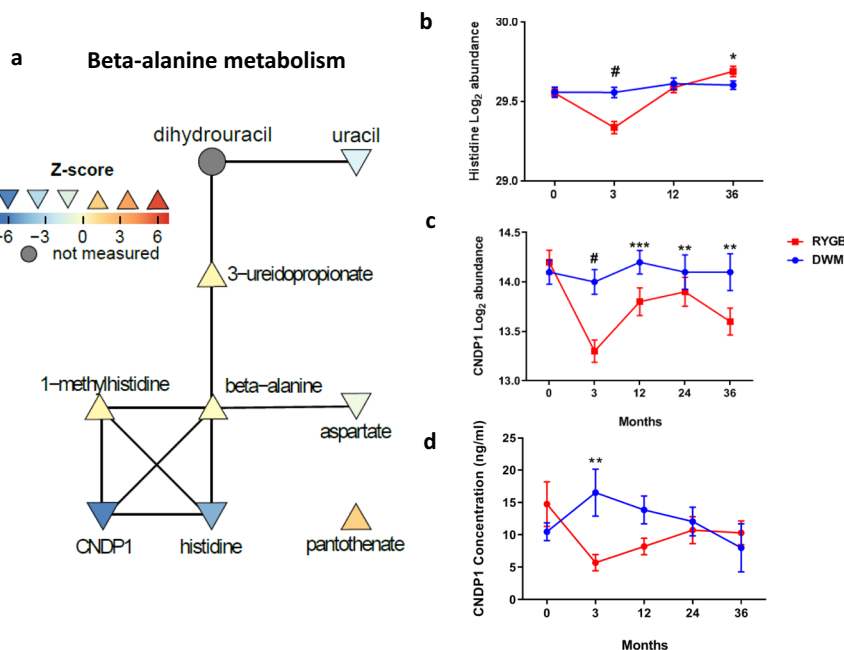


Fig. 5 Beta-alanine metabolism. **a** Network nodes are colored by between-group z-score, whereas unmeasured nodes are colored gray. Orientation of triangle also indicates the directionality of regulation. Connections are from the Pathway Commons network. **b** Log₂ abundance of histidine measured by metabolomics in samples from independent human participants at baseline, 3, 12, and 36 months from 19, 18, 16, and 9 DWM and 19, 18, 19, and 13 RYGB participants. **c** Log₂ abundance of CNDP1 measured by SOMAscan, measured in samples from independent human participants at baseline, 3, 12, 24, and 36 months from 19, 19, 16, 10, and 9 DWM and 19, 19, 19, 15, and 14 RYGB participants. **d** CNDP1 plasma levels measured by ELISA at baseline, 3, 12, 24, and 36 months from 10, 10, 10, 6, and 4 DWM participants and 10, 10, 9, 5, and 6 RYGB participants, respectively. Data were analyzed by two-sided moderated t-tests; post-baseline time points were analyzed using change from baseline. Data in **b–d** are reported as mean ± SEM. Nominal *p*-values are as follows: **b** histidine: 3 mo: #<0.0001, 36 mo: * = 0.019; **c** CNDP1 (SOMAscan): 3 mo: #<0.0001, 12 mo: *** = 0.00014; 24 mo: ** = 0.0054; 36 mo: ** = 0.0067; **d** CNDP1 ELISA: 3 mo: ** = 0.0036. Supplementary Data 2 includes FDRs and source data for (**b/c**); source data for **d** is provided as a Source Data file.

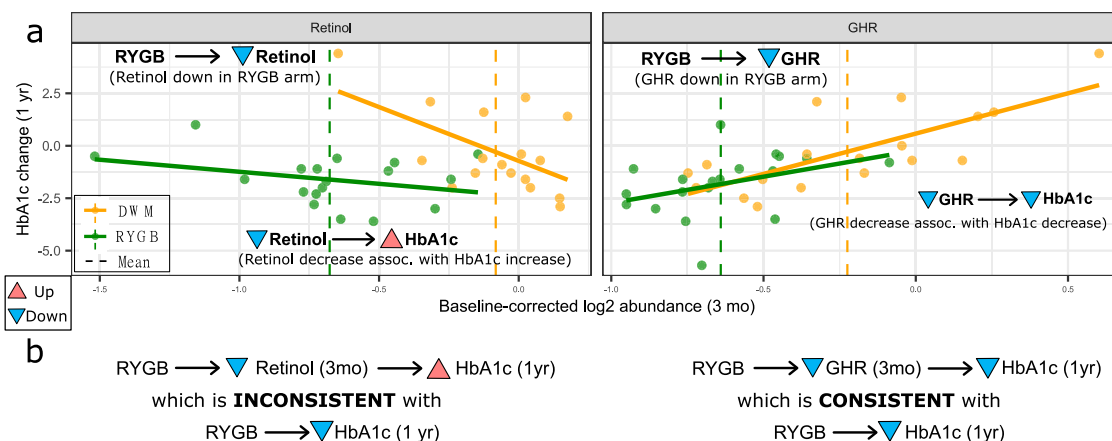


Fig. 6 GHR is a consistent mediator of change in HbA1c. Retinol is an inconsistent mediator (i.e., it suppresses or inhibits the causal effect; left side) whereas GHR is a consistent mediator of HbA1c change at 1 year (right side). **a** X-axis represents log₂ abundance change from baseline to the 3 month time point, and Y-axis represents HbA1c change from baseline at 1 year. **b** Shows putatively causal links with arrows based on (A) and direction of change with triangles pointed up (increase) or down (decrease). Source data are provided as a Source Data file.

significance, consistent with our simulations demonstrating that Hitman offers significantly more power (Supplementary Note 1). The top-ranking analytes identified by the joint significance method were GHR ($p = 2.3 \times 10^{-4}$; FDR = 0.26) and prolylhydroxyproline ($p = 10^{-3}$; FDR = 0.73). Prolylhydroxyproline was also the second-ranking analyte in Hitman ($p = 5 \times 10^{-4}$, FDR = 0.32; Supplementary Data 2).

We next asked whether the early postoperative change (baseline to the 3 month time point) in 40 clinical markers mediated HbA1c improvement at one year. The top-ranking early

postoperative mediators identified by Hitman were 6 min walk test distance and BMI ($p < 0.03$, FDR = 0.4; Supplementary Data 2). Strikingly, GHR was more significant (by raw and FDR-adjusted *p*-value) than any clinical markers, suggesting its utility as a potential clinical biomarker.

We next sought to test the mediation of our integrated pathways. Several approaches test pathway mediation^{41,42}. We utilized the Camera procedure to test pathway enrichment as it can utilize Hitman’s scores and accounts for the correlation between genes⁴³. We first tested for pathways whose change at

the 3 month time point mediate HbA1c improvement at 1 year, but none were significant.

Given that improvements in glycemic control after RYGB are known to be related to changes in insulin sensitivity and/or insulin secretion⁴⁴, we applied Hitman to identify analytes whose change from baseline at the 3 month time point mediated insulin secretion and insulin sensitivity change from baseline at 1 year (Supplementary Data 2) and followed this with integrative pathway mediation analysis. Insulin secretion, defined as the change in insulin from 0 to 30 min during a mixed meal tolerance test, was improved in RYGB vs. DWM, as expected (means of changes: RYGB = 37.4, DWM = 0.808; $p = 0.001$). No single analyte was identified as a significant mediator, but one pathway, Caffeine Metabolism, significantly mediated improved insulin secretion ($FDR < 10^{-4}$). Similarly, insulin sensitivity, defined by reduction in HOMA-IR, also improved in RYGB vs. DWM at 1 year (means of changes: RYGB = -2.12, DWM = -0.102; $p = 0.02$). No individual proteins or metabolites were significant mediators of insulin sensitivity, but many top-ranking analytes were BCAA-related metabolites. Consequently, pathway analysis identified Valine, Leucine, and Isoleucine Degradation as a significant pathway mediator of insulin sensitivity ($FDR < 10^{-7}$), together with 12 other pathways, primarily involved in lipid and amino acid metabolism (Supplementary Data 2).

Validation of growth hormone receptor as candidate mediator.

GHR was decreased by 24% at the 3 month time point after RYGB, relative to DWM, with similar reductions of 35, 28, and 30% at 12, 24, and 36 months, respectively. Robust GHR mediation led us to hypothesize that reductions in plasma GHR reflected reduced tissue content or altered receptor shedding, and thus could be associated with reduced growth hormone (GH) signaling. In the liver, GH is best recognized as a determinant of insulin-like growth factor 1 (IGF1) secretion during periods of nutritional sufficiency, but GHR also exerts complex pro-diabetogenic effects on glucose metabolism via IGF1-independent pathways⁴⁵. Indeed, IGF1 was not identified as a mediator by Hitman, and neither SOMAscan nor ELISA measures of IGF1 differed between groups at the 3 month time point (Supplementary Fig. 4). Levels of IGFBP1 and IGFBP2, which are repressed by GH signaling, were increased in RYGB vs. DWM at multiple time points (Fig. 2a and Supplementary Fig. 2). Plasma GH levels did not differ at the 3 month time point, but were >6-fold higher in baseline-corrected RYGB vs. DWM at 12 months (Supplementary Fig. 4). Collectively, these data support the hypothesis that pathway-selective growth hormone resistance in post-RYGB participants could contribute to sustained improvements in glucose metabolism⁴⁶.

Given that reduction in plasma GHR was a top-ranking mediator of improvements in glycemia, we asked whether expression of *Ghr* in the liver (where *Ghr* is most highly expressed) is similarly reduced after experimental bariatric surgery in rodents. We analyzed *Ghr* expression in liver of HFD-fed rats 8 weeks after vertical sleeve gastrectomy. Expression of *Ghr* was reduced by 26% ($p < 0.05$) in VSG-treated as compared with sham rats (Fig. 7b). Thus, reductions in hepatic *Ghr* follow both VSG and RYGB⁴⁷ forms of bariatric surgery in rodents, paralleling the reduction in plasma GHR we observed in humans following RYGB.

Given that changes in hepatic transcriptional patterns could reflect both the direct impact of surgery and secondary effects of altered metabolism, we next investigated whether experimental reduction of *Ghr* could affect glucose metabolism in a cell-autonomous fashion in mouse primary hepatocytes. siRNA-mediated *Ghr* knockdown led to a >90% reduction in expression of *Ghr* and 50% reduction in expression of *Igf-1* (Fig. 7c, d) as

well as increased expression of the GH signaling inhibitors *Socs1* and *Socs2* (Fig. 7e). Furthermore, *Ghr* knockdown enhanced insulin action to repress gluconeogenesis, as revealed by (1) increased insulin-mediated suppression of cAMP-stimulated glucose production (Fig. 7f), and (2) downregulation of the gluconeogenic enzymes *G6pc* and *Pepck* (Fig. 7g). Therefore, inhibition of hepatocyte GH signaling reduces transcriptional activation of gluconeogenesis and reduces glucose production. Since these key pathways contribute to hyperglycemia in T2D, reductions in GH signaling could contribute to improved glucose metabolism.

Discussion

RYGB yielded weight loss and improvement in multiple clinical parameters. Our proteomic and metabolomics analysis in the fasting state revealed both weight-dependent and -independent protein and metabolite changes. Weight-dependent changes included decreases in insulin, leptin, proinflammatory proteins, amino acids, and lipid-related metabolites; these changes are consistent with improved insulin sensitivity, loss of adipose mass, and reductions in inflammation observed in prior studies of RYGB^{9,10,48,49}. In contrast, the differential abundance of several analytes linked to bone or bile acid metabolism remained after BMI adjustment, suggesting weight-independent mechanisms underlying these changes. For example, the highly abundant bile acid and FXR ligand glycochenodeoxycholate sulfate and the intestinal FXR-dependent hormone FGF-19 were upregulated at later time points even with BMI adjustment in participants randomized to RYGB, consistent with prior reports in patients post-RYGB^{11,12,14,15} and in those with hypoglycemia after gastric bypass⁵⁰.

Valine, Leucine, and Isoleucine Degradation was the most significantly changed pathway at the 3 month time point between groups and a significant HOMA-IR pathway mediator. There were robust post-RYGB decreases in BCAA and multiple downstream catabolic intermediates, including C3 and C5 acylcarnitines. Our results are in agreement with reduced BCAA in response to RYGB in nonrandomized studies^{9,10,51,52} and prior findings of increased BCAA in insulin resistance^{53,54}. Lower BCAA levels post-RYGB may be a consequence of weight loss-related improvements in insulin sensitivity or altered microbial metabolism¹⁸, but may also contribute directly to improved insulin sensitivity⁵⁵. Our data reveal a marked 88% reduction in the valine catabolic intermediate 3-hydroxyisobutyrate at the 3 month time point, which is particularly interesting since 3-hydroxyisobutyrate can exit mitochondria and serve as a signaling molecule, promoting muscle lipid uptake and insulin resistance^{56,57}.

The significant protein putative mediator of improved glycemic control was GHR, which was reduced after RYGB. GHR at the 3 month time point is a more significant mediator of glycemia at 1 year than any clinical markers, including BMI, indicating that GHR may have clinical significance as a biomarker of bariatric surgery or as a therapeutic target. In agreement, liver *Ghr* expression was significantly reduced after VSG in rats with dietary obesity. These data are consistent with the >50% reduction in *Ghr* mRNA expression and protein (assessed by PCR and immunohistochemistry) in liver and multiple intestinal segments recently reported at both 9 days and 9 weeks post-RYGB in mice⁴⁷. These reductions in *Ghr* were also observed in weight-matched mice, indicating weight-independence. In our human cohort, reduction in GHR was closely correlated with reductions in BCAA-related metabolites, which have also been linked to improved post-surgical metabolism.

Beyond reductions in GHR content, we observed changes in plasma levels of downstream targets of GH signaling, with a robust increase in IGFBP1 at the 3 month time point and both IGFBP1 and

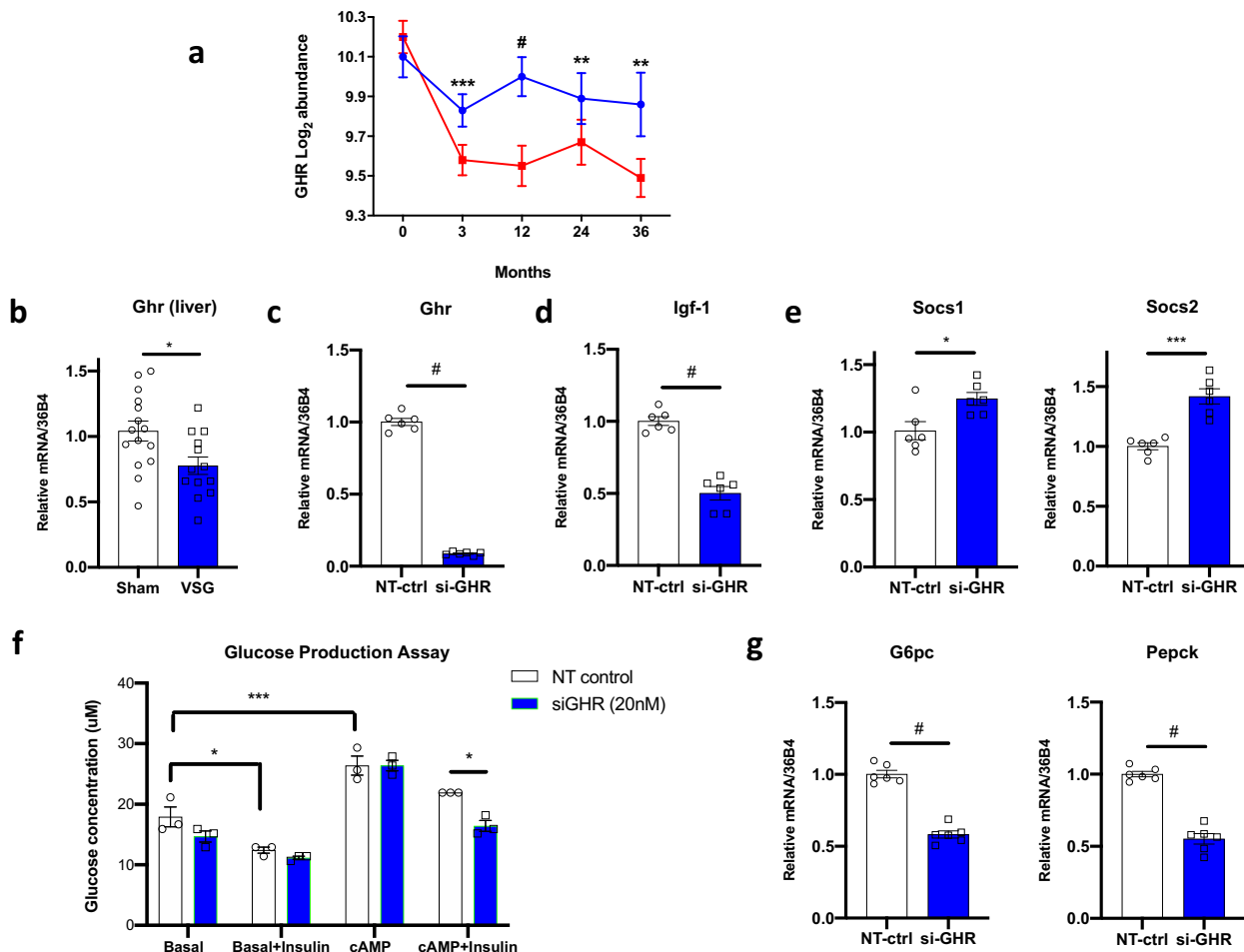


Fig. 7 GHR is reduced after bariatric surgery in humans and rodents and regulates hepatic glucose metabolism. **a** GHR in plasma is reduced after RYGB in humans, measured by SOMAscan in samples from independent human participants at baseline, 3, 12, 24, and 36 months from 19, 19, 16, 10, and 9 DWM and 19, 19, 19, 15, and 14 RYGB participants. **b** *Ghr* expression in liver, assessed by qRT-PCR, from diet-induced obese rats, 8 weeks post-VSG bariatric surgery (VSG) or sham controls (Sham) ($n = 13$ VSG, $n = 15$ sham). **c–g** Expression of GH signaling/axis genes, assessed by qRT-PCR, in cells treated with siRNA targeting *Ghr* (si-Ghr) or non-targeting scrambled control siRNA (NT-ctrl) ($n = 6$ per group). **c** *Ghr*, **d** *Igf-1*, and **e** *Socs1/2* in mouse primary hepatocytes treated with siRNA-*Ghr* or NT-control. **f** Glucose production in response to si-Ghr or NT-ctrl. Data are reported as mean \pm SEM. For **a** the p -values are nominal, using a two-sided T-test; FDRs are reported in Supplementary Data 2. For **b–e** and **g** two-sided t-tests were applied; for **f** two-way ANOVA with Bonferroni’s multiple comparison test was applied in GraphPad Prism. P values are as follows: **a**: GHR: 3 mo: *** = 0.00099; 12 mo: # < 0.0001, 24 mo: ** = 0.0024, 36 mo: ** = 0.0013. **b**: * = 0.0155, **c**: # < 0.0001, **d**: # < 0.0001; **e**: *Socs1*: * = 0.017, *Socs2*: *** = 0.0002; **f**: basal vs. insulin in NT control: * = 0.0322; basal vs. cAMP in NT control: *** = 0.0004; NT control vs. siGHR in cAMP + insulin: * = 0.0328. **g**: *G6pc* and *Pepck* both # < 0.0001. Source data for **a** is provided in Supplementary Data 2. Source data for panels **b–g** are provided in the Source Data file. For expression analysis, two independent experiments were performed, while glucose production data represent a single experiment.

IGFBP2 at 12, 24, and 36 months; these changes are concordant with reduced GH effect within IGF1-independent pathways regulating hepatic and systemic metabolism. The time course of these effects provides potential clues. Reduction in GHR at the 3 month time point was accompanied by an increase in the GHR-dependent target IGFBP1⁴⁷. By contrast, increases in GH measured by ELISA were not observed at the 3 month time point, but only emerged at 12 months. Taken together, these data support the hypothesis that late post-operative increases in GH may reflect both acquired GH resistance following RYGB, together with the reversal of obesity- and inflammation-associated reductions in GH secretion⁵⁸. The emergence of increases in GH at 12 months in the surgical group could also contribute to sustained lipolysis in adipose tissue⁵⁹. Additional experiments will be required to identify the tissue-specific changes in GH action associated with improved metabolic control after RYGB in patients with T2D.

Given that GH excess (due to acromegaly or pharmacologic use of GH in adults) is associated with increased glucose levels and diabetes in humans⁶⁰, a net reduction of GH signaling would be expected to improve systemic metabolism. Indeed, rodent models with experimental ablation of either *Gh* or *Ghr* have increased healthspan, delayed aging, and improved insulin sensitivity^{45,61}. Mice with liver-specific ablation of *Ghr* have impaired gluconeogenesis, with overt hypoglycemia⁶². Moreover, inhibition of GH action by the GHR antagonist pegvisomant improves systemic insulin action in humans⁴⁶.

Our data in isolated primary hepatocytes with experimental knockdown of *Ghr* provide further support for an important role for reduction in GHR-dependent signals as a mediator of improved glucose metabolism. Efficient *Ghr* knockdown reduced expression of gluconeogenic enzymes *Pepck* and *G6pc*, and enhanced the action of insulin to repress glucose production.

Taken together, our study adds to the growing evidence that RYGB decreases GHR-dependent signaling in the liver, and supports that this reduction may contribute to lowering of glucose and robust improvements in diabetes control^{63–66}.

We have developed and applied a mediation method that can be applied to studies where samples are randomized to a treatment (or can be modeled as randomized given covariates) whose causal effect on the outcome is consistent with prior knowledge and these samples' potential mediators are assayed. As others have noted⁶⁷, mediation analyses can be misled by confounding variables, such as an unmeasured factor that both affects GHR at the 3 month time point and (independently of GHR) affects HbA1c at one year. To avoid confounding among measured mediators, some high-throughput mediation approaches decompose measured analytes to identify latent mediators that are independent of measured confounders^{42,68}. However, these latent mediators are composed of multiple analytes and can be difficult to interpret and follow-up experimentally. Furthermore, in complex biological systems, many potential confounders are unmeasured, such as microbiome changes in this study, so confounders' impact often cannot be prevented by decomposing measured analytes. To avoid confounding, it is helpful to control for relevant covariates. For example, in this study Hitman analyzed changes within individuals, which obviates the need to control for some demographic variables. For our omics simulations, we measured inter-gene dependence in GTEx after controlling for previously identified covariates and unknown factors using surrogate variable analysis (SVA)³⁹, which decreased dependence. However, Hitman tests mediators individually and does not attempt to control for confounders or inter-analyte dependence, so its results should be considered exploratory, like other causal mediation methods that test analytes individually in omics datasets⁴¹. Future work should focus on combining Hitman with methods that consider multiple mediators simultaneously, such as pathway analyses⁴¹, and account for some level of confounding⁶⁹.

As required for a directional (“one-sided”) test, the direction to be tested in Hitman is defined using prior knowledge before Hitman is applied⁷⁰; this must be the same as the observed direction of the exposure's effect on the outcome. The question of when directional tests are appropriate is often debated, but the primary criteria are that the direction of the test should align with prior evidence⁷¹ and the scientific hypothesis^{72–75}. These two criteria are related since scientific hypotheses are based on prior evidence. For example, multiple studies (including ours) have found that RYGB causes a robust decrease in HbA1c. This provides confidence that RYGB truly decreases HbA1c, which supports the hypothesis that there exist mediators of RYGB's decrease in HbA1c. If the observed causal effect is not in the pre-specified direction or no information about the direction of a potential causal effect is available, then Hitman should not be applied, as the approach of using a directional test is invalid, and the false positive rate could be inflated by two-fold.

Here we are interested in consistent mediators, which can explain the causal effect of the exposure on the mediator and the mechanism of action⁷⁶. The mediation hypothesis in the inconsistent direction is that there exist mediators that suppress or inhibit RYGB's reduction of HbA1c. Detecting inconsistent mediators could be of value, as these mediators could be blocked to enhance RYGB's reduction of HbA1c. If these exist, though, their effect size should be small relative to the magnitude of RYGB-related reduction of HbA1c. If only inconsistent mediators are of interest, though, Hitman can be slightly modified so that consistent mediators are assigned a *p*-value of one and the *p*-value of inconsistent mediators are halved. Without this modification, Hitman has no power to detect inconsistent mediators.

We acknowledge that profiling semi-quantitative plasma metabolomics and proteomics, with emphasis on the secreted proteome, cannot fully define the pleiotropic effects of bariatric surgery, including both weight-dependent and weight-independent changes in complex inter-organ communication, gut microbiome effects, tissue-specific protein abundance, or flux in metabolic pathways. Moreover, medication usage, activity, and diet composition could impact both clinical and omics differences between groups in the human cohort. Although studies in rodents cannot fully recapitulate human systemic metabolism, we were able to identify and validate changes in *Ghr* and related pathway genes under controlled experimental conditions. The analytes we have identified and validated can be modulated in future studies to determine whether they can be utilized for non-surgical control of glucose metabolism in T2D, and our mediation method can be applied in translational omics studies.

Methods

Clinical study. All relevant ethical regulations for work with human participants were followed, and informed consent was obtained from all participants. The protocol was approved by the Partners Healthcare Institutional Review Board, and an independent data monitoring committee reviewed patient safety.

We performed a retrospective, exploratory analysis of fasting plasma samples from the Surgery or Lifestyle with Intensive Medical Management in the Treatment of Type 2 Diabetes (SLIMM-T2D) clinical trial ([clinicaltrials.gov:NCT01073020](https://clinicaltrials.gov/NCT01073020)), in which 38 individuals with T2D and obesity were randomized to RYGB using standard operative protocols (*n* = 19) or nonsurgical intensive diabetes weight management (DWM; *n* = 19) and followed longitudinally for 3 years; pre-specified primary study outcomes were previously reported⁵. Metabolic assessments were performed at baseline (pre-randomization) and repeated at the 3 month time point (defined as achieving 10% of initial body weight loss or at 3 months, permitting assessment at similar weight loss in both groups) and at 12, 18, 24, and 36 months.

Blood samples were obtained after an overnight fast and analyzed for HbA1c, plasma glucose, lipids (Quest Diagnostics), and insulin (Mercodia ELISA); additional aliquots were stored at –80 °C and used for the current analysis. Samples with HbA1c were available from 19, 19, 19, 17, 16, and 15 participants from the RYGB arm, and 19, 19, 18, 11, 10, and 10 participants from the DWM arm at the baseline, 3, 12, 18, 24, and 36 month time points. Groups were metabolically similar at baseline⁵.

Meal tolerance and 6 min walk testing were performed at baseline and at 3 and 12 months as described⁷⁷. In brief, Ensure (9 g protein, 40 g carbohydrate, 6 g fat) was consumed over 5 min, and blood was sampled at 30, 60, and 120 min. Insulin sensitivity was calculated by HOMA-IR, while insulin secretion was calculated as the change in insulin from 0 to 30 min during the meal test.

Proteomic profiling and validation. Plasma proteome profiling was performed using the high-throughput DNA aptamer-based SOMAscan assay platform (SomaLogic, Inc.)⁷⁸. The abundance of 1129 proteins (enriched for extracellular proteins) was quantified as relative fluorescent units (RFU), normalized, calibrated, and log₂-transformed. Samples were available at baseline, 3, 12, 24, and 36 month time points for 19, 19, 19, 15, and 14 participants in the RYGB arm, and for 19, 19, 16, 10, and 9 participants in the DWM arm (Supplementary Data 1).

Selected proteomic data were validated by ELISA in a subset of fasting plasma samples, including IGFBP2 (22-BP2HU-E01, ALPCO, NH), CNDP1 (F34010, LifeSpan Biosciences, WA), growth hormone (DGH00, R&D Systems, MN), and total IGF-1 (DG100, R&D Systems, MN). RBP4 and TTR were assayed by quantitative western blotting using polyclonal anti-human RBP4 (Dako) and anti-human TTR (Dako) with standard curves of purified human RBP4 or TTR (Sigma) on each blot⁷⁹. Changes per individual over time were tested for positive correlation to corresponding SOMAscan changes with a one-sided test of Pearson correlation.

Metabolomic profiling. Plasma metabolomics were profiled using a commercial semi-quantitative mass spectrometry-based platform (Metabolon, Inc.) as described^{80,81}. In brief, samples were prepared using the automated MicroLab STAR[®] system from Hamilton Company. Several recovery standards were added prior to extraction for QC purposes. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The organic solvent was removed using the TurboVap[®] (Zymark); after overnight storage under nitrogen, samples were reconstituted in solvents prior to separation using ultrahigh performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) with scan range 70–1000 *m/z*. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection

and chromatographic consistency. Raw data were extracted, and peaks were quantified by area-under-the-curve.

The majority of metabolites were identified by confirming a structure with a minimum of two orthogonal properties (here, an accurate mass m/z , accurate mass fragmentation, and a retention time) from a pure reference standard acquired under identical analytical conditions. These metabolites have confidence level 1, as per the Metabolomics Standard Initiative⁸². Those metabolites not officially confirmed based on a standard have confidence level 3 and their biochemical name is marked with an asterisk (*) in the metabolite annotation (Supplementary Data 2 and 3); these are primarily lipids.

Quality control was ensured by assessment of instrument and process variability, determined by calculating the median relative standard deviation (RSD) for the (a) internal standards added to each sample prior to injection into the mass spectrometers, and (b) endogenous metabolites present in 100% of technical replicates of a large pool of extensively characterized human plasma. Values for median RSD for instrument and process variability were 5 and 9%, respectively.

This dataset had 17% of values missing, and the metabolite average abundance from non-missing values was strongly negatively correlated with the metabolite's proportion of missing values (Spearman $\rho = -0.59$, $p < 10^{-15}$). This indicated that missingness is abundance-dependent, i.e., data were missing not at random (MNAR)⁸³. For MNAR data, potential imputation approaches include imputing half of the minimum abundance per metabolite (HM) and quantile regression imputation of left-censored data (QRILC), which imputes the left-censored data by randomly drawing values from a truncated normal distribution⁸⁴.

Metabolites that had missing values in more than 85% of samples were filtered out, leaving the dataset with 10% missing values. Missing values were imputed with HM, and abundance values were \log_2 -transformed. Samples were available at baseline, 3, 12, and 36 months from 19, 18, 19, and 13 participants in the RYGB arm and from 19, 18, 16, and 9 participants in the DWM arm. Metabolomics were not profiled at 24 months due to cost. Metabolomics data are provided in Supplementary Data 3.

Differential abundance of proteomics, metabolomics, and pathways. To test differential abundance of \log_2 normalized analytes between groups at baseline, we applied two-sided moderated t -tests with the R package *Limma*^{35,85}, and accounted for multiple hypothesis testing by controlling the FDR with the Benjamini–Hochberg method⁸⁶. We analyzed proteins and metabolites separately. *Limma* applies linear regression modeling with empirical Bayesian methods to improve each analyte's variance estimation using analytes' shared systematic variance. At post-baseline time points, we calculated change in analyte abundance from baseline for each individual, and then applied moderated t -tests to test if these changes varied by group. We repeated these post-baseline analyses accounting for each person's BMI change as a covariate.

Significance for this and other analyses of high-throughput data (either from SOMAscan or Metabolon) was defined as FDR < 0.15 . However, significance for analytes measured without high-throughput platforms, such as clinical data (e.g., BMI) or proteins analyzed by ELISA, was defined as $p < 0.05$.

To test the correlation of top metabolites to top proteins, we tested the correlation of baseline-corrected changes, and chose top analytes as those shown in the heatmaps in Fig. 2a, b.

To identify differentially abundant pathways, we constructed an integrated dataset of proteins and metabolites for samples that had both data types, where we averaged SomaLogic probes that had identical gene symbols. We applied *Limma*'s Roast method to this dataset against Small Molecule Pathway Database³¹ pathways. We were interested in pathways that change, even if, for example, proteins tend to increase and metabolites tend to decrease, so we reported statistics from Roast's "Mixed" test, which tests if a pathway's analytes tend to change without regard to direction (i.e., if analyte changes have large absolute values), so a pathway can be significant even if some analytes go up and others go equally down⁸⁷.

To plot top nodes in a network, we plot the most significant analytes as nodes with edges between them from the Pathway Commons⁸⁸ network version 9. When disconnected nodes can be easily connected, we add less significant or unmeasured intermediates. Nodes are colored by their t -statistics, which had sufficiently high degrees of freedom to approach z -scores, and were labeled as such.

In plots of individual analytes with standard error of the mean (SEM), ordinary (i.e., unmoderated) SEM are displayed.

Mediation analysis. We tested mediation of each clinical variable with the causal chain: group \rightarrow clinical variable change \rightarrow clinical outcome change. We tested each clinical variable's mediation by defining *group* as a binary variable representing RYGB or DWM per individual; *clinical variable change* as each clinical variable's change between baseline and the 3 month time point per individual; and *clinical outcome change* as the change in clinical outcome between baseline and 12 months per individual.

We tested mediation of analytes with the causal chain: group \rightarrow analyte change \rightarrow clinical outcome change, with *analyte change* as each analyte's change (on the \log_2 scale) between baseline and the 3 month time point per individual and the *clinical outcome change* as the change between baseline and 12 months (or in some cases later time points). As a comparison to Hitman, we similarly tested the mediation of analytes using the joint significance method and using the counterfactual approach with the mediation package²⁶ in the R software. In our simulations to evaluate

methods, the number of significant simulations was compared statistically with a two-sided Fisher exact test in the R software. We tested these Hitman results against pathways from the Small Molecule Pathway Database (SMPDB)³¹ with the CAMERA pre-ranked pathway analysis method⁴³ from the Limma package.

Validation studies in rodent model of bariatric surgery and in primary mouse hepatocytes. The study complied with all relevant ethical regulations for animal testing and research. All protocols were reviewed and approved by the University of Michigan (Ann Arbor, MI) and Joslin Diabetes Center Animal Care and Use Committees. Rat bariatric surgery was performed in high fat diet (HFD) fed male mice as described⁸⁹. Liver tissue was dissected 8 weeks postoperatively and flash-frozen for subsequent RNA extraction.

Primary hepatocyte isolation and siRNA-mediated gene knockdown. Primary hepatocytes were isolated from wild-type C57Bl6 male mice at age 8 weeks. After anesthesia (pentobarbital, 50 mg/kg), the vena cava was cannulated (30G) and the portal vein was cut for drainage. The liver was perfused for 5 min with 20 mL of perfusion solution (NaCl 140 mM, Tricine 25 mM, KCl 5.5 mM, EGTA 0.5 mM, KH_2PO_4 0.5 mM, Na_2HPO_4 0.3 mM, pH 7.4) and then for 5 min with 25 mL of digestion buffer (Eagle's balanced salts (NaCl 115 mM, NaHCO_3 25 mM, D-Glucose 5.5 mM, KCl 5.5 mM, CaCl_2 0.3 mM, Na_2HPO_4 0.8 mM, MgSO_4 0.8 mM), 1% penicillin/streptomycin and collagenase type II) for 5 min. The liver was harvested and minced; the cell suspension was filtered through a 100 μm cell strainer; cells were collected via centrifugation ($200 \times g$, 3 min), and the pellet was resuspended in 15 ml of complete DMEM (glucose 450 mg/dL) containing 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. The cell suspension was mixed with 10 ml of Percoll and centrifuged at $340 \times g$, 10 min; the pellet was resuspended in 25 mL complete DMEM, centrifuged again ($200 \times g$, 5 min), and resuspended in 10 mL complete DMEM. Cells were counted and seeded at 1×10^5 cells/mL onto collagen-coated plates.

Pools of four siRNAs targeting GHR or non-targeting controls (Horizon-Dharmacon) (final concentration 20 nM) were added to a mixture of OptiMEM (Gibco) and transfection reagent (DharmaFECT, 15 μl per 1 ml OptiMEM); the mixture was added to collagen-coated wells (200 μL /well) and incubated for 20 min at room temperature (RT). Freshly isolated primary hepatocytes ($n = 6$ in si-GHR; $n = 6$ in non-targeting controls) were added to each well (800 μL /well), and medium containing siRNA and transfection reagents was replaced after 24 h.

Hepatocyte glucose production. At 48 h post-seeding, cells were incubated in serum-free medium for 4 h, washed twice with PBS, and incubated for 2 h with DMEM containing glutamine, sodium lactate, and sodium pyruvate, in the presence or absence of pCPT-cAMP (100 μM , Sigma) and/or 10 nM insulin. After 2 h, the supernatant was collected, centrifuged twice ($200 \times g$, 3 min, $340 \times g$, 3 min); 50 μL of the supernatant were used for glucose quantification (Thermo). The attached hepatocytes were washed once with PBS and frozen in -80°C .

RNA extraction and gene expression. RNA from primary hepatocytes and liver tissue was extracted using TRIzol reagent using the manufacturer's protocol. cDNA was synthesized using random hexamers (High-Capacity cDNA Reverse Transcription Kit, ThermoFisher-Applied Biosystems). Real-time quantitative PCR was performed using SYBR green qPCR Supermix (Biorad) and QuantStudio 6 (Applied Biosystems). Primers are provided in Supplementary Table 1. Gene expression was quantified using the ddCt method.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The SOMAscan proteomics and clinical data generated in this study have been deposited in the Gene Expression Omnibus database under accession code [GSE122279](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122279). The metabolomics data are provided in Supplementary Data 3. SOMAscan proteomics, metabolomics, clinical data, and the R/Bioconductor⁸⁵ code to reproduce this article's main results are publicly available on GitHub at <https://github.com/jdreyf/slimm-t2d-omics> and at Zenodo at <https://doi.org/10.5281/zenodo.5485746>⁹⁰. A Source Data file is included. All remaining data generated or analyzed during this study are included in this published article (and its supplementary information files).

Code availability

The Hitman R package is available at <https://github.com/jdreyf/Hitman>, which depends on the ezlimma package (<https://github.com/jdreyf/ezlimma>). We generated plots from its companion package, ezlimmaplot (<https://github.com/jdreyf/ezlimmaplot>). SOMAscan proteomics, metabolomics, clinical metadata, and the R/Bioconductor⁸⁵ code to reproduce this article's main results are available on GitHub at <https://github.com/jdreyf/slimm-t2d-omics> and at Zenodo at <https://doi.org/10.5281/zenodo.5485746>⁹⁰. To reproduce the headline results with one click even when software versions of dependencies change, data and code have also been deposited in a Code Ocean capsule at <https://doi.org/10.24433/CO.9548519.v2>, which is open, exportable, reproducible, and interoperable.

Received: 10 February 2020; Accepted: 11 November 2021;

Published online: 29 November 2021

References

- Schauer, P. R. et al. Bariatric surgery versus intensive medical therapy for diabetes—5-year outcomes. *N. Engl. J. Med.* **376**, 641–651 (2017).
- Mingrone, G. et al. Bariatric-metabolic surgery versus conventional medical treatment in obese patients with type 2 diabetes: 5 year follow-up of an open-label, single-centre, randomised controlled trial. *Lancet* **386**, 964–973 (2015).
- Courcoulas, A. P. et al. Three-year outcomes of bariatric surgery vs lifestyle intervention for type 2 diabetes mellitus treatment: a randomized clinical trial. *JAMA Surg.* **150**, 931–940 (2015).
- Cummings, D. E. et al. Gastric bypass surgery vs intensive lifestyle and medical intervention for type 2 diabetes: the CROSSROADS randomised controlled trial. *Diabetologia* **59**, 945–953 (2016).
- Simonson, D., Halperin, F., Foster, K., Vernon, A. & Goldfine, A. Clinical and patient-centered outcomes in obese patients with type 2 diabetes 3 years after randomization to Roux-en-Y gastric bypass surgery versus intensive lifestyle management: the SLIMM-T2D study. *Diabetes Care* **41**, 670–679 (2018).
- Jorgensen, N. B. et al. Acute and long-term effects of Roux-en-Y gastric bypass on glucose metabolism in subjects with type 2 diabetes and normal glucose tolerance. *Am. J. Physiol. Endocrinol. Metab.* **303**, E122–E131 (2012).
- Ramracheya, R. D. et al. PYY-dependent restoration of impaired insulin and glucagon secretion in type 2 diabetes following Roux-En-Y gastric bypass surgery. *Cell Rep.* **15**, 944–950 (2016).
- Salehi, M., Prigeon, R. L. & D'Alessio, D. A. Gastric bypass surgery enhances glucagon-like Peptide 1-stimulated postprandial insulin secretion in humans. *Diabetes* **60**, 2308–2314 (2011).
- Laferrere, B. et al. Differential metabolic impact of gastric bypass surgery versus dietary intervention in obese diabetic subjects despite identical weight loss. *Sci. Transl. Med.* **3**, 80re82 (2011).
- Arora, T. et al. Roux-en-Y gastric bypass surgery induces early plasma metabolomic and lipidomic alterations in humans associated with diabetes remission. *PLoS One* **10**, e0126401 (2015).
- Dutia, R. et al. Temporal changes in bile acid levels and 12 α -hydroxylation after Roux-en-Y gastric bypass surgery in type 2 diabetes. *Int. J. Obes.* **39**, 806–813 (2015).
- Patti, M. E. et al. Serum bile acids are higher in humans with prior gastric bypass: potential contribution to improved glucose and lipid metabolism. *Obesity* **17**, 1671–1677 (2009).
- Ryan, K. K. et al. FXR is a molecular target for the effects of vertical sleeve gastrectomy. *Nature* **509**, 183–188 (2014).
- Liu, H., Hu, C., Zhang, X. & Jia, W. Role of gut microbiota, bile acids and their cross-talk in the effects of bariatric surgery on obesity and type 2 diabetes. *J. Diabetes Investig.* **9**, 13–20 (2018).
- DePaoli, A. M. et al. FGF19 analog as a surgical factor mimetic that contributes to metabolic effects beyond glucose homeostasis. *Diabetes* **68**, 1315–1328 (2019).
- Saeidi, N. et al. Reprogramming of intestinal glucose metabolism and glycemic control in rats after gastric bypass. *Science* **341**, 406–410 (2013).
- Magouliotis, D., Tasiopoulou, V., Sioka, E., Chatedaki, C. & Zacharoulis, D. Impact of bariatric surgery on metabolic and gut microbiota profile: a systematic review and meta-analysis. *Obes. Surg.* **27**, 1345–1357 (2017).
- Ilhan, Z. E. et al. Distinctive microbiomes and metabolites linked with weight loss after gastric bypass, but not gastric banding. *ISME J.* **11**, 2047–2058 (2017).
- Aron-Wisniewsky, J. et al. Major microbiota dysbiosis in severe obesity: fate after bariatric surgery. *Gut* **68**, 70–82 (2018).
- Mokadem, M., Zechner, J. F., Margolskee, R. F., Drucker, D. J. & Aguirre, V. Effects of Roux-en-Y gastric bypass on energy and glucose homeostasis are preserved in two mouse models of functional glucagon-like peptide-1 deficiency. *Mol. Metab.* **3**, 191–201 (2014).
- Ye, J. et al. GLP-1 receptor signaling is not required for reduced body weight after RYGB in rodents. *Am. J. Physiol. Regulatory, Integr. Comp. Physiol.* **306**, R352–R362 (2014).
- Barfield, R. et al. Testing for the indirect effect under the null for genome-wide mediation analyses. *Genet. Epidemiol.* **41**, 824–833 (2017).
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G. & Sheets, V. A comparison of methods to test mediation and other intervening variable effects. *Psychol. Methods* **7**, 83 (2002).
- Fritz, M. S. & Mackinnon, D. P. Required sample size to detect the mediated effect. *Psychol. Sci.* **18**, 233–239 (2007).
- Huang, Y.-T. Joint significance tests for mediation effects of socioeconomic adversity on adiposity via epigenetics. *Ann. Appl. Stat.* **12**, 1535–1557 (2018).
- Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. mediation: r package for causal mediation analysis. *J. Stat. Softw.* <https://doi.org/10.18637/jss.v059.i05> (2014).
- Al-Regaiey, K. et al. Effects of gastric sleeve surgery on the serum levels of GH, IGF-1 and IGF-binding protein 2 in healthy obese patients. *BMC Gastroenterol.* **20**, 199 (2020).
- Ceccarini, G. et al. Serum IGF-binding protein 2 (IGFBP-2) concentrations change early after gastric bypass bariatric surgery revealing a possible marker of leptin sensitivity in obese subjects. *Endocrine* **65**, 86–93 (2019).
- Shah, R. V. et al. Proteins altered by surgical weight loss highlight biomarkers of insulin resistance in the community. *Arterioscler. Thromb. Vasc. Biol.* **39**, 107–115 (2019).
- Hedbacker, K. et al. Antidiabetic effects of IGFBP2, a leptin-regulated gene. *Cell Metab.* **11**, 11–22 (2010).
- Jewison, T. et al. SMPDB 2.0: big improvements to the small molecule pathway database. *Nucleic Acids Res.* **42**, D478–D484 (2014).
- Rhodes, C. J. et al. Plasma proteome analysis in patients with pulmonary arterial hypertension: an observational cohort study. *Lancet Respir. Med.* **5**, 717–726 (2017).
- Yang, Q. et al. Serum retinol binding protein 4 contributes to insulin resistance in obesity and type 2 diabetes. *Nature* **436**, 356–362 (2005).
- Mackinnon, D. P., Lockwood, C. M. & Williams, J. Confidence limits for the indirect effect: distribution of the product and resampling methods. *Multivar. Behav. Res.* **39**, 99 (2004).
- Ritchie, M. et al. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* **43**, gkv007–e047 (2015).
- Barfield, R. et al. Testing for the indirect effect under the null for genome-wide mediation analyses. *Genet. Epidemiol.* **41**, 824–833 (2017).
- Law, C., Chen, Y., Shi, W. & Smyth, G. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
- Consortium, G. T. The genotype-tissue expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
- Leek, J. T. & Storey, J. D. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet.* **3**, 1724–1735 (2007).
- Oliva, M. et al. The impact of sex on gene expression across human tissues. *Science* <https://doi.org/10.1126/science.aba3066> (2020).
- Djordjilović, V. et al. Global test for high-dimensional mediation: testing groups of potential mediators. *Stat. Med.* **38**, 3346–3360 (2019).
- Huang, Y. T. & Pan, W. C. Hypothesis test of mediation effect in causal mediation model with high-dimensional continuous mediators. *Biometrics* **72**, 402–413 (2015).
- Wu, D. & Smyth, G. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res.* **40**, e133–e133 (2012).
- Yoshino, M. et al. Effects of diet versus gastric bypass on metabolic function in diabetes. *N. Engl. J. Med.* **383**, 721–732 (2020).
- Aguiar-Oliveira, M. H. & Bartke, A. Growth hormone deficiency: health and longevity. *Endocr. Rev.* **40**, 575–601 (2019).
- Pedersen, M. H. et al. Substrate metabolism and insulin sensitivity during fasting in obese human subjects: impact of GH blockade. *J. Clin. Endocrinol. Metab.* **102**, 1340–1349 (2017).
- Ben-Zvi, D. et al. Time-dependent molecular responses differ between gastric bypass and dieting but are conserved across species. *Cell Metabolism* **28**, 310–323 (2018).
- Illan-Gomez, F. et al. Obesity and inflammation: change in adiponectin, C-reactive protein, tumour necrosis factor- α and interleukin-6 after bariatric surgery. *Obes. Surg.* **22**, 950–955 (2012).
- Wolf, R. M. et al. Cytokine, chemokine, and cytokine receptor changes are associated with metabolic improvements after bariatric surgery. *J. Clin. Endocrinol. Metab.* **104**, 947–956 (2019).
- Mulla, C. M. et al. Plasma FGF-19 levels are increased in patients with post-bariatric hypoglycemia. *Obes. Surg.* **29**, 2092–2099 (2019).
- Lips, M. et al. Roux-en-Y gastric bypass surgery, but not calorie restriction, reduces plasma branched-chain amino acids in obese women independent of weight loss or the presence of type 2 diabetes. *Diabetes Care* **37**, 3150–3156 (2014).
- Wijayatunga, N. et al. Roux-en-Y gastric bypass surgery alters serum metabolites and fatty acids in patients with morbid obesity. *Diabetes/Metab. Res. Rev.* <https://doi.org/10.1002/dmrr.3045> (2018).
- Wang, T. et al. Metabolite profiles and the risk of developing diabetes. *Nat. Med.* **17**, 448–453 (2011).
- Newgard, C. B. et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* **9**, 311–326 (2009).
- White, P. J. et al. Branched-chain amino acid restriction in Zucker-fatty rats improves muscle insulin sensitivity by enhancing efficiency of fatty acid oxidation and acyl-glycine export. *Mol. Metab.* **5**, 538–551 (2016).
- Jang, C. et al. A branched-chain amino acid metabolite drives vascular fatty acid transport and causes insulin resistance. *Nat. Med.* **22**, 421–426 (2016).
- Viegas, C. M. et al. Evidence that 3-hydroxyisobutyric acid inhibits key enzymes of energy metabolism in cerebral cortex of young rats. *Int. J. Dev. Neurosci.: Off. J. Int. Soc. Dev. Neurosci.* **26**, 293–299 (2008).

58. Juiz-Valina, P. et al. Altered GH-IGF-1 axis in severe obese subjects is reversed after bariatric surgery-induced weight loss and related with low-grade chronic inflammation. *J. Clin. Med.* <https://doi.org/10.3390/jcm9082614> (2020).
59. Kopchick, J. J., Berryman, D. E., Puri, V., Lee, K. Y. & Jorgensen, J. O. L. The effects of growth hormone on adipose tissue: old observations, new mechanisms. *Nat. Rev. Endocrinol.* **16**, 135–146 (2020).
60. Melmed, S. Pathogenesis and diagnosis of growth hormone deficiency in adults. *N. Engl. J. Med.* **380**, 2551–2562 (2019).
61. Dominici, F. P., Arostegui Diaz, G., Bartke, A., Kopchick, J. J. & Turyn, D. Compensatory alterations of insulin signal transduction in liver of growth hormone receptor knockout mice. *J. Endocrinol.* **166**, 579–590 (2000).
62. Fang, F., Shi, X., Brown, M. S., Goldstein, J. L. & Liang, G. Growth hormone acts on liver to stimulate autophagy, support glucose production, and preserve blood glucose in chronically starved mice. *Proc. Natl Acad. Sci. USA* **116**, 7449–7454 (2019).
63. Sjostrom, L. et al. Association of bariatric surgery with long-term remission of type 2 diabetes and with microvascular and macrovascular complications. *Jama* **311**, 2297–2304 (2014).
64. Arterburn, D. Microvascular outcomes in patients with diabetes after bariatric surgery. *Ann. Intern. Med.* **170**, 507–508 (2019).
65. Fisher, D. P. et al. Association between bariatric surgery and macrovascular disease outcomes in patients with type 2 diabetes and severe obesity. *Jama* **320**, 1570–1582 (2018).
66. Aminian, A. et al. Association of metabolic surgery with major adverse cardiovascular outcomes in patients with type 2 diabetes and obesity. *Jama* <https://doi.org/10.1001/jama.2019.14231> (2019).
67. Zhang, H. et al. Estimating and testing high-dimensional mediation effects in epigenetic studies. *Bioinformatics* **32**, 3150–3154 (2016).
68. Chén, O. Y. et al. High-dimensional multivariate mediation with application to neuroimaging data. *Biostatistics* **19**, 121–136 (2018).
69. VanderWeele, T. J. & Vansteelandt, S. Mediation analysis with multiple mediators. *Epidemiol. Method* **2**, 95–115 (2014).
70. Bland, J. M. & Altman, D. G. One and two sided tests of significance. *BMJ* **309**, 248 (1994).
71. Freedman, L. S. An analysis of the controversy over classical one-sided tests. *Clin. Trials* **5**, 635–640 (2008).
72. Ludbrook, J. Should we use one-sided or two-sided P values in tests of significance? *Clin. Exp. Pharm. Physiol.* **40**, 357–361 (2013).
73. Wolterbeek, R., Enkin, M. W., Bland, J. M. & Altman, D. G. One and two sided tests of significance: statistical hypothesis should be brought into line with clinical hypothesis. *BMJ* **309**, 873–874 (1994).
74. Knottnerus, J. A. & Bouter, L. M. The ethics of sample size: two-sided testing and one-sided thinking. *J. Clin. Epidemiol.* **54**, 109–110 (2001).
75. Cho, H.-C. & Abe, S. Is two-tailed testing for directional research hypotheses tests legitimate? *J. Bus. Res.* **66**, 1261–1266 (2013).
76. VanderWeele, T. J. Mediation and mechanism. *Eur. J. Epidemiol.* **24**, 217–224 (2009).
77. Panosian, J. et al. Physical activity in obese type 2 diabetes after gastric bypass or medical management. *Am. J. Med.* **130**, 83–92 (2017).
78. Gold, L., Walker, J., Wilcox, S. & Williams, S. Advances in human proteomics at high scale with the SOMAScan proteomics platform. *N. Biotechnol.* **29**, 543–549 (2012).
79. Graham, T. E. et al. Retinol-binding protein 4 and insulin resistance in lean, obese, and diabetic subjects. *N. Engl. J. Med.* **354**, 2552–2563 (2006).
80. Gall, W. E. et al. alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS One* **5**, e10883 (2010).
81. Ford, L. et al. Precision of a clinical metabolomics profiling platform for use in the identification of inborn errors of metabolism. *J. Appl. Lab. Med.* **5**, 342–356 (2020).
82. Sumner, L. W. et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **3**, 211–221 (2007).
83. Little, R. J. & Rubin, D. B. *Statistical Analysis with Missing Data* 793 (John Wiley & Sons, 2019).
84. Wei, R. et al. Missing value imputation approach for mass spectrometry-based metabolomics data. *Sci. Rep.* <https://doi.org/10.1038/s41598-017-19120-0> (2018).
85. Gentleman, R. et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, 1–16 (2004).
86. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
87. Wu, D. et al. ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176–2182 (2010).
88. Cerami, E. et al. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res.* **39**, D685–D690 (2011).
89. Stefater, M. A. et al. Sleeve gastrectomy induces loss of weight and fat mass in obese rats, but does not affect leptin sensitivity. *Gastroenterology* **138**, 2426–2436 (2010).
90. Jonathan M. Dreyfuss Analysis of omics from the SLIMM-T2D trial (v1.0.4). *Zenodo* <https://doi.org/10.5281/zenodo.5485746>. (2021).

Acknowledgements

We thank Danyel Cavazos for help with figures; Grace Daher, G. H. Nguyen, Ronaldo Da Silva Francisco Junior, and Lado Gholijashvili for assistance with R package development; Wilson Pacheco, and Oleksii S. Rukhlenko for helpful discussion; Jack Magrisso, Stace Kernodle, Alfor Lewis, Andriy Myronovych, and Mouhamadou Toure for experiments; and Research Computing at Harvard Medical School and at Joslin Diabetes Center for computational resources. This work was supported by unrestricted grant support from Medimmune (to M.E.P.) and the Charles King Trust (fellowship to Y.Y.). The parent trial was supported by the National Institute of Health grants RC1DK086918, R56DK095451, and P30DK03836. J.M.D., H.P., and S.K. were also supported by P30DK03836. B.K. acknowledges support from R01DK43051 and P30DK57521. V.E. was supported by the Swiss National Science Foundation (SNSF). RS acknowledges support from P01DK117821. We also thank the Herbert Graetz Fund at Joslin Diabetes Center and Patient-Centered Outcomes Research Institute (PCORI) grant CE-1304-6756. For the parent trial, Covidien provided funds for the surgical costs of participants with BMI less than 35 who were randomized to undergo surgery; Lifescan, a Division of Johnson & Johnson, provided home glucose monitoring supplies; Nestle Nutrition Inc. provided Boost; and NovoNordisk provided drug supplies.

Author contributions

J.M.D. and M.E.P. wrote the manuscript. All authors (J.M.D., Y.Y., X.D., V.E., H.P., D.S.C., A.V., F.H., P.A., A.K., Y.S., B.W.H., J.G., C.M.R., S.K., B.B.K., K.F., R.S., A.G., V.D., and M.E.P.) contributed to data generation or analysis and reviewed/edited the manuscript. J.M.D. and M.E.P. take responsibility for the content of the manuscript.

Competing interests

For the parent study (SLIMM-T2D) from which samples were derived for analysis in the current study, D.S., A.V., F.H., and A.B.G. received unrestricted funding from the Herbert Graetz Fund and from Covidien, to support surgical procedures for research participants with BMI 30–35 (not covered by insurance). Some supplies for the SLIMM-T2D clinical study were received from Lifescan, a division of Johnson and Johnson, Nestle Inc, and Novo Nordisk. M.E.P. received unrestricted investigator-initiated research grant funding from Medimmune to support the present work, and funding to support assay costs from SomaLogic. ABG served on advisory boards for Baranova, and Kowa. DCS received funds from PCORI and is a stock/shareholder of GI Windows. A.K., B.H., C.M.R., J.G., and Y.S. were employees of Medimmune when the work was initiated. Y.Y. is now employed at Vertex, F.H. at Form Health, A.K. at Sanofi, B.H. at Immunocore, J.G. at AstraZeneca, and A.B.G. at Novartis Institutes of Biomedical Research. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-27289-2>.

Correspondence and requests for materials should be addressed to Mary Elizabeth Patti.

Peer review information *Nature Communications* thanks Adil Mardinoglu, Geltrude Mingrone and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021