Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Longitudinal Transcriptomic Remodeling of Adipose Tissue After Bariatric Surgery Revealed by Differential Expression and Explainable Machine Learning

Soumaya Allouch
Qatar University
Md. Shaheenur Islam Sumon
Qatar University
Aisha Naeem
Qatar University
Søren Risom Kristensen
Aalborg University
Muhammad E. H. Chowdhury
Qatar University
Shona Pedersen
spedersen@qu.edu.qa

Qatar University

Article

Keywords: Bariatric surgery, White adipose tissue, Longitudinal transcriptomics, Machine learning, Explainable artificial intelligence, Metabolic remodeling

Posted Date: June 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9232975/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Bariatric surgery leads to major improvements in metabolic health, but the long-term molecular changes in white adipose tissue (WAT) that support these benefits are not fully understood. We aimed to characterize how WAT gene expression remodels over time after Roux-en-Y gastric bypass (RYGB) and to identify stable markers of long-term adaptation. We reanalyzed a public longitudinal transcriptomic dataset from women undergoing RYGB, with WAT biopsies collected before surgery and at 2 and 5 years after surgery. We combined traditional differential expression analysis with machine learning and SHAP-based interpretation to identify biologically meaningful and reproducible gene signatures. Both approaches showed a clear shift in WAT gene-expression profiles after surgery. We observed strong and persistent downregulation of immune and inflammatory pathways and significant remodeling of extracellular matrix components, especially at 5 years. Machine-learning models accurately separated pre- and post-surgery samples, and SHAP analysis confirmed that the most predictive genes aligned with biologically relevant pathways. Together, the analyses converged on a stable set of 50 genes that reflect long-term reprogramming of adipose tissue. RYGB induces durable molecular changes in WAT characterized by reduced inflammation and structural remodeling that likely contribute to long-term metabolic improvement. Integrating transcriptomics with explainable machine learning provides a powerful framework for identifying robust biomarkers and for gaining deeper insight into tissue adaptation after bariatric surgery.

1. Introduction

Obesity is a chronic, multifactorial disease associated with substantial disturbances in metabolic homeostasis, including insulin resistance, dyslipidemia, type 2 diabetes, and increased cardiovascular risk. Central to the development of these complications is white adipose tissue (WAT), which functions not only as an energy storage site but also as a dynamic endocrine and immune organ. In obesity, WAT undergoes extensive structural and functional remodeling characterized by adipocyte hypertrophy, altered adipokine secretion, immune cell infiltration, and chronic low-grade inflammation, all of which contribute to systemic metabolic dysfunction [1, 2]

Roux-en-Y gastric bypass (RYGB) surgery is one of the most effective interventions for severe obesity, producing substantial and sustained weight loss together with marked improvements in insulin sensitivity, lipid metabolism, and glycemic control. Importantly, the metabolic benefits of RYGB often persist long after the period of maximal weight loss and may remain evident even in the presence of partial weight regain [3, 4]. These observations indicate that the long-term efficacy of bariatric surgery cannot be explained solely by changes in body weight or fat mass, but instead reflects durable biological adaptations at the tissue and molecular levels.

Growing evidence suggests that adipose tissue plays a central role in mediating the sustained metabolic benefits of bariatric surgery. Transcriptomic profiling of human subcutaneous WAT has revealed extensive gene-expression remodeling following RYGB, involving pathways related to inflammation, immune signaling, lipid metabolism, mitochondrial function, and adipocyte differentiation [5, 6]. Notably, although many transcriptional changes occur within the first one to two years after surgery, a subset of genes; particularly those associated with inflammatory and immune processes, continue to show progressive regulation up to five years postoperatively, independent of weight regain [5]. These persistent molecular adaptations may contribute to improved adipose tissue morphology and endocrine function, thereby supporting long-term metabolic health.

Despite these insights, comprehensive longitudinal analyses capturing both intermediate and long-term transcriptomic changes in human adipose tissue remain limited. Many existing studies focus on short-term follow-up or cross-sectional comparisons, which restricts the ability to distinguish transient from sustained molecular responses to surgery. Moreover, given known sex-specific differences in adipose tissue biology and metabolic regulation, there is a particular need for well-characterized longitudinal studies in female cohorts.

High-throughput transcriptomic datasets present substantial analytical challenges due to their high dimensionality, complex correlation structure, and potential redundancy among features. Traditional statistical approaches, such as differential expression analysis, are essential for identifying genes that change significantly between conditions but may not fully capture multivariate expression patterns that optimally distinguish physiological states over time [7]. Machine learning (ML) approaches provide a complementary framework by enabling feature prioritization, classification, and pattern recognition in high-dimensional data, thereby improving the identification of biologically informative gene signatures [7, 8].

In recent years, ML methods have been increasingly applied to biomedical and metabolic research, including studies of bariatric surgery outcomes, to enhance predictive performance and uncover complex molecular relationships [9, 10]. However, the interpretability of ML models remains a critical concern, particularly in translational research. Explainable artificial intelligence approaches, such as SHapley Additive exPlanations (SHAP), address this limitation by quantifying the contribution of individual features to model predictions, thereby improving transparency and biological interpretability [11, 12].

In the present study, we performed a comprehensive longitudinal transcriptomic analysis of subcutaneous white adipose tissue from obese women before and up to five years after Roux-en-Y gastric bypass surgery, with comparison to age-matched non-operated controls. By integrating conventional differential expression analysis with machine learning based feature selection, classification, and SHAP-based model interpretation, we aimed to identify robust and sustained gene-expression signatures associated with long-term metabolic improvement following bariatric surgery. This integrative analytical strategy provides new insights into the durable molecular remodeling of adipose tissue after RYGB and highlights transcriptomic pathways that may underlie the long-term clinical benefits of surgical weight loss.

2. Materials and Methods

Study Population

This study presents a novel *in silico* re-analysis of a publicly available dataset. The dataset used in this study consists of longitudinal adipose tissue transcriptomic and clinical data collected from obese women undergoing Roux-en-Y gastric bypass (RYGB) surgery [13]. Abdominal subcutaneous white adipose tissue (WAT) biopsies were obtained before surgery (n = 50), 2 years post-surgery (n = 49), and 5 years post-surgery (n = 38), with 37 women providing samples at all three time points. For long-term comparison, an additional control group of 28 age-matched, nonoperated women with stable body weight was included. Global gene expression profiling was performed using Affymetrix Clariom™ D microarrays. Alongside transcriptomic data, extensive clinical, metabolic, and adipose phenotyping was collected, including BMI, fat mass (DXA-derived ESAT and EVAT), insulin resistance (HOMA-IR), lipid profiles, adipokines, adipocyte size and number, and lipolysis markers [13]. All participants in this study were women, and the research comprised three human cohorts. Ethical approval was obtained from the local ethics committee in Stockholm (Regional etikprövningsnämnden i Stockholm, Karolinska Institutet, Tomtebodavägen 18 A 17165 Solna, Sweden), and written informed consent was provided by all participants; no financial compensation was offered. All methods were performed in accordance with the relevant guidelines and regulations. All examinations were performed at 8:00 AM after an overnight fast. Anthropometric measurements and clinical chemistry analyses were conducted according to previously published protocols, and total body fat mass was estimated using bioelectrical impedance analysis [13].

Sample collection and preparation

The gene expression analyses were based on white adipose tissue (WAT) obtained by needle biopsy from abdominal subcutaneous fat, collected lateral to the umbilicus at each sampling time point. Immediately after collection, biopsies were rinsed in a 9 mg/mL sodium chloride solution and divided into two portions: approximately 300 mg of tissue was snap-frozen in liquid nitrogen and stored for subsequent gene expression profiling, while the remaining fraction was used for adipocyte isolation to assess fat cell size as part of the phenotypic characterization [13].

Gene expression profiling platform

Total RNA was isolated from frozen white adipose tissue using the RNeasy kit (Qiagen). Global transcriptome profiling was carried out with Affymetrix Clariom™ D microarrays according to the manufacturer's protocols. Sense-strand, biotin-labeled target cDNA was generated using the WT Plus Kit (Thermo Fisher), after which 5.5 µg of cDNA was fragmented and hybridized to the arrays in a GeneChip Hybridization Oven 645 at 45°C for 16–18 hours. Following hybridization, arrays were washed and stained using a GeneChip Fluidics Station 450 and subsequently scanned with an Affymetrix GeneChip Scanner 7 [13]. Microarray CEL files were processed using the SST-RMA workflow and analyzed in Qlucore Omics Explorer following the original analytical framework, with lncRNAs extracted and longitudinal comparisons performed using paired designs. Selected findings were validated by qRT-PCR with SYBR Green chemistry and ΔΔCT/ΔCT normalization using 18S rRNA or B2M as reference genes;

additional methodological details are described in [13]. All probe-set identifiers were mapped to HGNC-approved gene symbols, and suffixes indicating transcript variants or platform-specific probe labels were removed for consistency.

Statistical analysis

Gene expression differences between time points were evaluated using paired t-tests for within-subject comparisons, specifically comparing post-surgery samples with pre-surgery (baseline) samples at 2- and 5-year follow-ups. Results were reported as p-values with the direction of change. To account for multiple testing and reduce the risk of false positives, p-values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) method [14], and genes with FDR ≤ 0.05 were considered statistically significant [15]. Exploratory and presentation analyses included principal component analysis (PCA) [16] to visualize global transcriptomic separation across groups/time points [17], volcano plots [18] to jointly display effect size and statistical significance, and box plots to illustrate group-wise distributions for representative genes.

Machine Learning analysis

Machine learning (ML) models were employed to classify and analyze the data following a structured preprocessing and modeling pipeline. To enhance training efficiency and ensure that all features contributed comparably, the input data were normalized using the Standard Scaler [19], which also supported robust training and improved generalization. Given the high dimensionality of the dataset, feature ranking was applied as a key strategy to reduce overfitting. Feature importance was computed using three tree-based feature selection models, Random Forest [20], XGBoost [21], and ExtraTrees [22]—and the final importance score for each feature was obtained by averaging the scores from these models. Feature importance scores were averaged across three tree-based models to reduce model-specific bias and capture features consistently prioritized across diverse ensemble architectures. The dataset was divided into training (80%) and testing (20%) subsets, and model performance was assessed using 5-fold cross-validation. Using the selected features, several ML classifiers were trained and evaluated, including Linear Discriminant Analysis, XGBoost Classifier, Random Forest Classifier, Logistic Regression, and ExtraTrees Classifier. All ML-selected genes, together with their corresponding ML importance score, $\log_2$FC, FDR are provided in Supplementary Table S1. Evaluation metrics, including accuracy, precision, recall, and F1-score, were computed, and the corresponding equations are provided in Supplementary Section S1.

Gene Set Enrichment Analysis

To identify the overarching biological processes remodeled after surgery, pathway enrichment analysis was performed. The list of the top 500 most significant differentially expressed genes, ranked by p-value, was used as the input gene set. This list was uploaded to Metascape (www.metascape.org) for functional annotation against the Gene Ontology (GO) Biological Processes database. Pathways were considered significantly enriched if the false discovery rate (FDR) adjusted p-value was less than 0.05.

Figure 1 illustrates the overall workflow of this longitudinal transcriptomics study. After defining the study cohorts and timepoints, the analysis proceeds through two complementary branches: machine learning and statistical testing. In the machine learning branch, the data are preprocessed, informative features are selected, and models are trained using a train/test split, followed by test-set evaluation to identify the best-performing model. Model interpretability is then performed using SHapley Additive explanation (SHAP) [23] to determine the most influential genes driving predictions. In the statistical analysis branch, group-wise testing is carried out, followed by multiple-testing correction (FDR) to control false discoveries and identify significant genes. Key visual outputs (e.g., PCA, volcano, ROC, and box plots) summarize expression patterns and discrimination performance. Finally, results from both branches are compared to highlight genes that are consistently supported by both statistical significance and ML-based importance.

3. Results

Characteristics of study populations

Table 1 summarizes key demographic, anthropometric, and metabolic parameters measured at baseline and at 2 and 5 years after RYGB surgery. Values are reported as mean ± standard deviation for each follow-up timepoint. Changes from baseline to 2 years and from baseline to 5 years are evaluated using the corresponding P-values shown in the last two columns. The results

indicate marked improvements in body weight, BMI, total fat mass, and insulin resistance indices after surgery. Glycemic control metrics (glucose, insulin, HOMA-IR) show significant reductions, reflecting improved metabolic status. Lipid-related measures also shift favorably, including lower triglycerides and higher HDL cholesterol over time. Overall, the table highlights sustained post-surgical benefits, with some attenuation between 2 and 5 years for select measures.

Differential Expression Analysis and Machine Learning Results

In the statistical analysis workflow, group-wise differential expression analysis was performed, followed by false discovery rate (FDR) correction to control for multiple testing. After quality control and filtering, including the exclusion of unannotated or unidentified probe sets, the final analysis was carried out on approximately 25,000 annotated genes selected from the original 40,590 transcript features. Genes with an FDR ≤ 0.05 were considered statistically significant and were subsequently used for downstream visualization and candidate gene prioritization. Following FDR correction, 6,825 significant genes were identified in the comparison between before surgery and 2 years post-surgery, while 6,479 significant genes were identified in the comparison between before surgery and 5 years post-surgery. In Supplementary Figure S1, a Venn diagram summarizes shared and unique significant DE genes across the three longitudinal comparisons. A total of 3030 genes are common to all three comparisons, while 510 are shared only between Before Surgery vs Control and 2-Years vs Before Surgery, and 337 are shared only between Before Surgery vs Control and 5-Years vs Before Surgery. The post-surgery comparisons share 1528 genes (2-Years vs Before Surgery ∩ 5-Years vs Before Surgery), with unique genes of 644 (Before Surgery vs Control), 1757 (2-Years vs Before Surgery), and 1584 (5-Years vs Before Surgery).

Figure 2 (a–b) presents PCA score plots computed using the statistically significant genes (FDR ≤ 0.05). In both panels, samples from Before Surgery form a distinct cluster that separates from 2 Years Post-Surgery and 5 Years Post-Surgery, indicating a clear shift in global gene-expression patterns after surgery. The separation is primarily driven by PC1, suggesting that the major variance reflects differences between the pre-surgery and post-surgery states. Overall, the PCA results confirm that the selected significant genes effectively distinguish Before Surgery from long-term follow-up timepoints. Figure 2 (c–d) shows the volcano plots for differential expression comparing Before Surgery vs 2 Years Post-Surgery and Before Surgery vs 5 Years Post-Surgery. The x-axis represents log2 fold change, while the y-axis shows −log10(P), so genes farther from the center and higher on the plot indicate stronger and more significant changes. In the 2 Years Post-Surgery comparison, 418 genes are upregulated and 126 are downregulated relative to Before Surgery. In the 5 Years Post-Surgery comparison, the number of changes is slightly larger, with 504 upregulated and 138 downregulated genes. Overall, both panels indicate a predominance of upregulated genes after surgery and demonstrate sustained transcriptomic remodeling at both follow-up timepoints. In Supplementary Figure S2, the volcano plot comparing 2-year and 5-year post-surgery samples shows that 4 genes are significantly upregulated and 23 genes are significantly downregulated between the two timepoints.

Figure 2 (e–f) shows the PLS-DA loading plots highlighting the genes that contribute most to separating the groups in each comparison. For Before Surgery vs 2 Years Post-Surgery (e) and Before Surgery vs 5 Years Post-Surgery (f), genes with larger absolute loadings have a stronger influence on the discrimination between timepoints. In the PLS-DA Before Surgery vs 2 Years Post-surgery, the strongest positive loadings are observed for ACSS3.1 and ACVR1C, followed by GLUL (and GLULP3), indicating that these genes contribute most strongly to the separation toward the post-surgery group at 2 years and are likely associated with metabolic adaptation after surgery. In contrast, for Before Surgery vs 5 Years Post-surgery, RNU11 stands out as the dominant positively loaded gene, suggesting a distinct long-term molecular signature compared with the earlier post-surgical phase. The common gene between both analyses is C1orf162, which appears in both.

In the machine-learning analysis, the top 20 features (genes) were selected using a combined importance score derived from three tree-based models: XGBoost, ExtraTrees, and Random Forest. Figure 2 (g–h) illustrates the resulting mean normalized importance rankings for the before surgery comparisons with 2 Years Post-Surgery and 5 Years Post-Surgery, respectively. These panels highlight the genes that most consistently contributed to classification performance across the ensemble feature-importance approach. In the Before Surgery vs 2 Years Post-surgery model, BMP2K emerges as the most important feature, followed by IL1RN, ACSS3, and HMOX1, highlighting the dominance of metabolic regulation and inflammation-related pathways shortly after surgery. In contrast, the Before Surgery vs 5 Years Post-surgery model is strongly driven by immune signaling genes,

with FCGR2A showing the highest importance, followed by NCKAP1L, BTK, and RNU11, indicating a shift toward immune and leukocyte-associated processes in long-term outcomes.

Using the top features from Fig. 2 (g–h), we trained ML models and summarized the top 5 in Table 2. Random Forest performed best for both comparisons, achieving 97.97% (Before vs 2 years) and 97.72% (Before vs 5 years) accuracy/F1. For 2 years, LGBM and AdaBoost also showed high performance (~ 97% and 96%). For 5 years, LDA was the second-best model (96.59%). Overall, the selected features enable strong separation of post-surgery samples from baseline.

Table 2
Performance metrics of the top five machine-learning models.

Before Surgery Vs After 2 Years						**Before Surgery Vs After 5 Years**					
Models	A	P	R	S	F1	Models	A	P	R	S	F1
Random Forest	97.97	98.05	97.97	97.93	97.97	Random Forest	97.72	97.72	97.72	97.64	97.72
LGBM	96.96	97.14	96.96	97.03	96.96	LDA	96.59	96.62	96.59	96.77	96.59
AdaBoost	95.95	96.03	95.95	95.91	95.95	AdaBoost	93.18	93.31	93.18	93.55	93.19
Gradient Boosting	92.92	92.94	92.92	92.94	92.92	Gradient Boosting	90.90	91.83	90.9	92.45	90.95
XGB	89.89	90.16	89.89	89.97	89.88	LGBM	90.9	90.9	90.9	90.56	90.9

The XGBoost model demonstrated the highest predictive performance (AUC = 0.9X). From this model, we identified the 50 most influential genes that discriminate between the pre- and 5-year post-surgery states. These top discriminative genes, which form the core of our predictive signature, are presented in Table 3.

To visualize the expression patterns and individual predictive performance of the most significant genes from this signature, we plotted the top upregulated and downregulated genes in Fig. 3 (A) and (B), respectively. These figures highlight key genes and confirm their robust changes in expression following surgery. To provide a unified view of the predictive and statistical relevance of each ML-selected gene, we compiled an integrated summary of ML importance scores, $\log_2$ fold changes, and FDR values (Supplementary Table S1).

Table 3
Top discriminative genes ranked by adjusted p-value (FDR) and AUC.

Before Surgery Vs 2 Years Post-Surgery					Before Surgery Vs 5 Years Post-Surgery				
Gene Name	**Full Form**	**log2FC**	**Adj p value**	**AUC**	**Gene Name**	**Full Form**	**log2FC**	**Adj p value**	**AUC**
ACSS3	Acyl-CoA synthetase short chain family member 3	2.22	6.71E-23	0.98	RNU11	RNA, U11 small nuclear	2.24	1.35E-19	0.97
ACVR1C	Activin A receptor type 1C	2.09	4.66E-22	0.98	TLR8	Toll like receptor 8	-2.25	1.35E-19	0.97
BMP2K	BMP2 inducible kinase	-1.39	1.07E-20	0.97	CHST11	Carbohydrate sulfotransferase 11	-1.88	1.35E-19	0.98
ALCAM	Activated leukocyte cell adhesion molecule	-3.15	1.58E-20	0.99	CLEC7A	C-type lectin domain containing 7A	-2.82	6.24E-19	0.97
NLN	Neurolysin	-1.31	1.58E-20	0.96	PPT1	Palmitoyl-protein thioesterase 1	-1.51	7.37E-19	0.97
HMOX1	Heme oxygenase 1	-1.60	3.11E-20	0.98	CTSS	Cathepsin S	-3.64	7.37E-19	0.98
CD248	CD248 molecule (endosialin)	-2.69	3.58E-20	0.98	C1orf162	Chromosome 1 open reading frame 162	-3.01	1.06E-18	0.97
NCKAP1L	NCK associated protein 1 like	-2.32	3.58E-20	0.97	IL1RN	Interleukin 1 receptor antagonist	-4.30	1.06E-18	0.97
MSR1	Macrophage scavenger receptor 1	-3.63	3.63E-20	0.97	PLA2G7	Phospholipase A2 group VII	-4.48	1.06E-18	0.97
PSAP	Prosaposin	-0.93	8.75E-20	0.97	TIMP1	TIMP metallopeptidase inhibitor 1	-2.38	1.09E-18	0.99
GLUL	Glutamate-ammonia ligase (glutamine synthetase)	2.13	8.87E-20	0.96	C5AR1.1	Complement C5a receptor 1	-2.70207	1.30E-18	0.97
FPR3	Formyl peptide receptor 3	-3.17	1.56E-19	0.97	ALCAM	Activated leukocyte cell adhesion molecule	-3.29	2.72E-18	0.97
C1orf162	Chromosome 1 open reading frame 162	-2.55677	1.74E-19	0.97	MSR1	Macrophage scavenger receptor 1	-4.00	2.72E-18	0.96
TFEC	Transcription factor EC	-2.04106	1.74E-19	0.96	LCP1	Lymphocyte cytosolic protein 1	-4.16	4.79E-18	0.97
GAS2L3	Growth arrest specific 2 like 3	-1.68174	1.74E-19	0.97	NPL	N-acetylneuraminate pyruvate lyase	-2.62	5.20E-18	0.98
PLA2G7	Phospholipase A2 group VII	-3.98	1.78E-19	0.98	VAMP8	Vesicle associated membrane protein 8	-1.95256	1.54E-17	0.97

Before Surgery Vs 2 Years Post-Surgery					Before Surgery Vs 5 Years Post-Surgery				
Gene Name	**Full Form**	**log2FC**	**Adj p value**	**AUC**	**Gene Name**	**Full Form**	**log2FC**	**Adj p value**	**AUC**
CD68	CD68 molecule	-2.31	2.68E-19	0.97	LAPTM5	Lysosomal protein transmembrane 5	-2.59584	2.08E-17	0.97
CXCL16	C-X-C motif chemokine ligand 16	-2.45	3.78E-19	0.97	ARPC5	Actin related protein 2/3 complex subunit 5	-1.10349	2.08E-17	0.96
PTPN13	Protein tyrosine phosphatase non-receptor type 13	0.92	3.97E-19	0.96	CD86	CD86 molecule	-1.99893	2.08E-17	0.97
NPL	N-acetylneuraminate pyruvate lyase	-2.21	4.32E-19	0.97	IL18	Interleukin 18	-1.82347	3.23E-17	0.95

Pathway Enrichment Analysis

Having identified and visualized these key predictive genes, we next sought to validate the broader biological themes they represent. We performed a comparative pathway enrichment analysis to test whether our machine learning approach and the standard statistical analysis converged on similar biological functions. The results, shown in Fig. 4, reveal a high degree of concordance between the two analytical techniques.

Specifically, pathways central to tissue remodeling were identified as highly significant by both methods, including "Inflammatory response," "Extracellular matrix organization," and "Wound healing." This strong overlap demonstrates that the biological findings are robust and not an artifact of a single analytical method. Furthermore, both approaches independently identified "Response to lipid," reinforcing the conclusion that a sustained immune-metabolic adaptation is a key outcome five years after surgery.

Finally, to directly link the predictive features of our model to their biological impact, we applied SHapley Additive exPlanations (SHAP). This analysis quantifies how much each gene contributes to the model's prediction for a sample being pre- or post-surgery. As illustrated in Fig. 5, higher gene expression values (red) with positive SHAP scores increased the probability of classifying samples as post-surgery, while lower expression values (blue) with negative SHAP scores contributed toward baseline classification. This analysis confirms that the genes with the highest statistical significance, such as RNU11 (upregulated) and CLEC7A (downregulated), were also the most impactful drivers of the model's predictions, thereby strengthening the connection between our predictive model and its underlying biology.

4. Discussion

Bariatric surgery remains one of the most effective interventions for achieving sustained weight loss and improving metabolic health in individuals with severe obesity. Yet, the biological mechanisms that underpin these long-term benefits, particularly within white adipose tissue (WAT), are yet to be unraveled [4]. As a metabolically active and immunologically responsive organ, WAT undergoes profound remodeling in response to weight loss, including changes in adipocyte function, immune cell composition, and gene expression [1, 2]. Recent longitudinal studies have highlighted that these molecular adaptations can persist for years after surgery, suggesting a form of metabolic memory that extends beyond the initial weight loss phase [5, 24, 25]. In this study, we sought to better understand the temporal dynamics of WAT remodeling by analyzing gene expression profiles in subcutaneous adipose tissue collected before, and at two and five years after, Roux-en-Y gastric bypass (RYGB). By combining traditional differential expression analysis with machine learning and explainable artificial intelligence (XAI) techniques, we identified distinct early and late gene signatures, as well as a core set of stable markers that may reflect durable biological reprogramming. These findings offer new insights into the long-term molecular trajectory of adipose tissue following bariatric surgery and highlight the value of interpretable computational tools in uncovering clinically meaningful biomarkers.

Although this dataset was originally generated by Kerr et al. (2022), the present work provides several layers of novelty that extend well beyond the scope of the original publication. First, we performed a complete re-analysis of all longitudinal time points,

integrating baseline, 2-year, and 5-year samples within a unified analytical framework rather than examining individual comparisons in isolation. Second, we applied a new integrative computational pipeline that combines traditional differential expression with supervised machine-learning models and SHAP-based explainability, enabling the identification of genes that are not only statistically significant but also predictive and mechanistically interpretable [11, 26]. Third, by merging statistical significance, ML feature-importance, and SHAP attribution, we identified a stable, long-term gene signature that was not reported in Kerr et al., including classifiers such as RNU11 that only emerge when considering multivariate interactions. Together, these advances provide a systems-level view of adipose remodeling that complements and substantially expands upon the original findings.

A major finding of our study is the profound and sustained suppression of immune and inflammatory pathways within adipose tissue five years after surgery. The downregulation of key genes such as TLR8, a sensor for viral single-stranded RNA, CLEC7A (Dectin-1), a pattern recognition receptor crucial for fungal recognition, and CTSS (Cathepsin S), an enzyme essential for antigen presentation by MHC class II molecules, collectively points to a deep and lasting latency of the resident immune cell populations [27]. In the obese state, adipose tissue is characterized by a chronic, low-grade inflammation driven by stressed adipocytes and pro-inflammatory macrophages, a state directly linked to the development of insulin resistance [28, 29]. The coordinated downregulation of these specific genes suggests a multi-faceted resolution of this state, involving a reduction in both innate immune sensing and the capacity for antigen presentation, which would limit T-cell activation. Our pathway analysis, demonstrating "Inflammatory response" as a top theme in both our statistical and machine learning models (Fig. 4), provides robust, systems-level confirmation of this conclusion. More importantly, the durability of this anti-inflammatory signature, persisting even with partial weight regain as noted in other longitudinal studies [24, 30], suggests that bariatric surgery may induce a fundamental "reset" of the adipose tissue immune environment, rather than simply responding to reduced adipocyte size. This lasting immunomodulation is likely a primary mechanism underpinning the long-term improvements in systemic insulin sensitivity and overall metabolic health.

Complementing the profound suppression of inflammation, our analysis revealed a dynamic and significant remodeling of the extracellular matrix (ECM) and a fundamental shift in cellular metabolism. In obesity, adipose tissue ECM often becomes excessively cross-linked and fibrotic, impairing healthy adipocyte expansion and promoting a pro-inflammatory environment [31]. The persistent alteration of genes like ACVR1C and ALCAM at both timepoints (Table 3) suggests a durable change in cell-cell and cell-matrix interactions that counteracts this pathological state. Our comparative pathway analysis strongly supports this, identifying "Extracellular matrix organization" as a core point of concordance between our methods (Fig. 4). This indicates a stable restructuring of the tissue architecture away from a fibrotic state towards a more flexible and healthier scaffold, which is critical for restoring proper adipocyte function and insulin sensitivity [32]. Furthermore, the emergence of a distinct 5-year signature, marked by the upregulation of transcription factors like MLXIPL (carbohydrate-responsive element-binding protein), a master regulator of de novo lipogenesis [33] points towards a sophisticated, long-term adaptation in metabolic programming. This suggests that after an initial phase of rapid catabolism, the tissue establishes a new homeostatic set point, actively regulating lipid metabolism in a manner that likely contributes to the sustained maintenance of metabolic improvements [26].

A key strength and novel contribution of our study is the integration of explainable AI (XAI) to move beyond prediction and into biological interpretation. While machine learning models are powerful, their "black box" nature can limit their translational value [11]. By applying SHAP analysis, we were able to deconstruct our model's decisions on a per-gene, per-sample basis, confirming that the model's predictions were driven by biologically plausible features. The SHAP summary plots (Fig. 6) provide a striking visualization of this, showing, for instance, that higher expression of the upregulated gene RNU11 consistently pushed the model's prediction toward a "post-surgery" classification. Conversely, higher expression of the pro-inflammatory gene CLEC7A drove the prediction toward "baseline." This demonstrates a tight coupling between a gene's statistical significance (Figs. 3(A) & (B)), its predictive importance for classification (Table 3), and its mechanistic contribution to the model's output (Fig. 6). This multi-layered validation provides exceptional confidence in our identified gene signature. More broadly, our integrative framework serves as a powerful template for future longitudinal studies, illustrating how explainable AI can be used to transform complex transcriptomic data into testable hypotheses with clear clinical and biological relevance [12, 34].

RNU11, a core component of the minor (U12-dependent) spliceosome, emerged as a dominant long-term classifier in our analysis. Although small nuclear RNAs (snRNAs) are not typically highlighted in adipose-tissue transcriptomics [35], their

expression can reflect broader changes in cellular activity, including altered splicing demands in immune cells or shifts in adipocyte turnover following weight-loss surgery. At present, the biological significance of RNU11 upregulation remains unclear, and we cannot exclude the possibility that its strong signal partly reflects platform-specific characteristics of microarrays, which often detect snRNAs with high sensitivity and dynamic range. Nonetheless, its consistent importance across statistical and machine-learning approaches suggests that changes in the minor spliceosome may accompany long-term adipose remodeling. Further studies using RNA-seq or targeted snRNA quantification will be necessary to determine whether RNU11 represents a true biological marker or a technical artifact.

While this study provides a comprehensive longitudinal and interpretable analysis of adipose tissue transcriptomic remodeling after RYGB, several considerations should be noted when interpreting the findings. First, the study cohort consisted exclusively of women, which reduces biological heterogeneity but may limit the generalizability of the results to male populations. Furthermore, long-term transcriptomic changes at the 5-year timepoint could be influenced by age-related processes and hormonal transitions such as menopause, factors known to independently alter adipose tissue biology, especially in female populations ([36, 37]. Second, analyses were confined to subcutaneous white adipose tissue and therefore may not fully capture molecular remodeling processes occurring in visceral adipose depots, which are also critically involved in metabolic risk and inflammation [2, 38]. Third, gene expression profiling was performed using microarray technology, which offers robust longitudinal comparability but may have reduced sensitivity for low-abundance transcripts or novel isoforms compared with RNA-sequencing approaches [39]. Finally, although the paired longitudinal design strengthens inference and reduces inter-individual variability, the observational nature of the study precludes causal conclusions regarding the functional roles of individual genes. Future studies integrating multi-omic profiling, functional validation, and more diverse cohorts will be important to further elucidate the mechanisms underlying long-term adipose tissue remodeling after bariatric surgery.

5. Conclusion

In summary, this study provides a comprehensive and interpretable analysis of long-term white adipose tissue transcriptomic remodeling following Roux-en-Y gastric bypass surgery. By integrating longitudinal differential expression analysis with machine learning based feature prioritization and SHAP-driven model interpretability, we identified distinct early and late molecular signatures that reflect dynamic metabolic adaptation and sustained suppression of adipose immune and inflammatory pathways. Importantly, the strong concordance between statistical significance, classification performance, and explainable-AI attribution highlights a subset of stable genes that may serve as robust markers of durable adipose tissue reprogramming. Our findings support the concept that the long-term metabolic benefits of bariatric surgery are underpinned by persistent molecular changes in adipose tissue that extend beyond weight loss alone. More broadly, our integrative analytical framework demonstrates how combining transcriptomics with interpretable machine learning can enhance biological insight in longitudinal metabolic studies and may inform future efforts to identify biomarkers and therapeutic targets for sustained metabolic health following obesity interventions.

Declarations

Additional Information

The authors declare that they have no competing interests.

Funding statement:

This research was funded by the Qatar Research, Development and Innovation (QRDI) council through the National Research Program, grant number PPM 07-0421-240054.

Data Availability Statement:

The dataset analyzed during the current study was originally published by Kerr et al. (2022) and is publicly available in the NCBI Gene Expression Omnibus (GEO) repository under the accession number GSE199076 and GSE199063.

Author Contributions

S.A.: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization.

M.S.I.S.: Data Curation, Formal Analysis, Validation, Writing – Review & Editing.

A.N.: Data Curation, Investigation, Writing – Review & Editing.

S.R.K.: Writing – Review & Editing.

M.E.H.C.: Conceptualization, Methodology, Supervision, Project Administration, Writing – Review & Editing.

S.P.: Conceptualization, Supervision, Project Administration, Funding Acquisition, Writing – Review & Editing.

References

1. Hotamisligil, G. S. Inflammation, metaflammation and immunometabolic disorders. *Nature* **542**, 177–185. https://doi.org:10.1038/nature21363 (2017).
2. Crewe, C., An, Y. A. & Scherer, P. E. The ominous triad of adipose tissue dysfunction: inflammation, fibrosis, and impaired angiogenesis. *J. Clin. Invest.* **127**, 74–82. https://doi.org:10.1172/jci88883 (2017).
3. Sjöström, 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. https://doi.org:10.1001/jama.2014.5988 (2014).
4. Arterburn, D. E., Telem, D. A., Kushner, R. F. & Courcoulas, A. P. Benefits and Risks of Bariatric Surgery in Adults: A Review. *Jama* **324**, 879–887. https://doi.org:10.1001/jama.2020.12567 (2020).
5. Kerr, A. G., Andersson, D. P., Rydén, M., Arner, P. & Dahlman, I. Long-term changes in adipose tissue gene expression following bariatric surgery. *J. Intern. Med.* **288**, 219–233. https://doi.org:10.1111/joim.13066 (2020).
6. van der Kolk, B. W. et al. Differential Mitochondrial Gene Expression in Adipose Tissue Following Weight Loss Induced by Diet or Bariatric Surgery. *J. Clin. Endocrinol. Metab.* **106**, 1312–1324. https://doi.org:10.1210/clinem/dgab072 (2021).
7. Libbrecht, M. W. & Noble, W. S. Machine learning applications in genetics and genomics. *Nat. Rev. Genet.* **16**, 321–332. https://doi.org:10.1038/nrg3920 (2015).
8. Camacho, D. M., Collins, K. M., Powers, R. K., Costello, J. C. & Collins, J. J. Next-Generation Machine Learning for Biological Networks. *Cell* **173**, 1581–1592. https://doi.org:10.1016/j.cell.2018.05.015 (2018).
9. Saux, P. et al. Development and validation of an interpretable machine learning-based calculator for predicting 5-year weight trajectories after bariatric surgery: a multinational retrospective cohort SOPHIA study. *Lancet Digit. Health*. **5**, e692–e702. https://doi.org:10.1016/s2589-7500(23)00135-8 (2023).
10. Pantelis, A. G., Epiphaniou, P. & Lapatsanis, D. P. Machine learning and artificial intelligence for predicting short and long-term complications following metabolic bariatric surgery - a systematic review. *Artif. Intell. Surg.* **5**, 322–344. https://doi.org:10.20517/ais.2024.104 (2025).
11. Lundberg, S. M. & Lee, S. I. in *Neural Information Processing Systems*.
12. Salih, A. M. et al. A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. *Adv. Intell. Syst.* **7**, 2400304. https://doi.org/10.1002/aisy.202400304 (2025). https://doi.org:.
13. Kerr, A. G. et al. The long noncoding RNA ADIPINT regulates human adipocyte metabolism via pyruvate carboxylase. *Nat. Commun.* **13**, 2958 (2022).
14. Kwong, K. S., Holland, B. & Cheung, S. H. A modified Benjamini–Hochberg multiple comparisons procedure for controlling the false discovery rate. *J. Stat. Plann. Inference*. **104**, 351–362 (2002).
15. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.: Ser. B (Methodol.)*. **57**, 289–300 (1995).
16. Maćkiewicz, A. & Ratajczak, W. Principal components analysis (PCA). *Comput. Geosci.* **19**, 303–342 (1993).

17. Abdi, H. & Williams, L. J. Principal component analysis. *Wiley interdisciplinary reviews: Comput. Stat.* **2**, 433–459 (2010).
18. Li, W. Volcano plots in analyzing differential expressions with mRNA microarrays. *J. Bioinform. Comput. Biol.* **10**, 1231003 (2012).
19. Singh, D. & Singh, B. Investigating the impact of data normalization on classification performance. *Appl. Soft Comput.* **97**, 105524 (2020).
20. Fawagreh, K., Gaber, M. M. & Elyan, E. Random forests: from early developments to recent advancements. *Syst. Sci. Control Engineering: Open. Access. J.* **2**, 602–609 (2014).
21. Chen, T. & XGBoost A Scalable Tree Boosting System. *Cornell University* (2016).
22. Geurts, P., Ernst, D. & Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **63**, 3–42 (2006).
23. Mangalathu, S., Hwang, S. H. & Jeon, J. S. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Eng. Struct.* **219**, 110927 (2020).
24. Bouchard-Mercier, A. et al. Molecular remodeling of adipose tissue is associated with metabolic recovery after weight loss surgery. *J. Transl Med.* **20**, 283. https://doi.org:10.1186/s12967-022-03485-6 (2022).
25. Loft, A. et al. Single cell-resolved transcriptional dynamics of human subcutaneous adipose tissue during lifestyle- and bariatric surgery-induced weight loss. *bioRxiv*, 2025.2001.2030.634294 (2025). https://doi.org:10.1101/2025.01.30.634294
26. Andersson, D. P., Tseng, B. T. P., Arner, P. & Dahlman, I. Long-term improvement of adipocyte insulin action during body weight relapse after bariatric surgery: a longitudinal cohort study. *Surg. Obes. Relat. Dis.* **18**, 683–692. https://doi.org:10.1016/j.soard.2022.02.013 (2022).
27. Aruwa, C. E. & Sabiu, S. Adipose tissue inflammation linked to obesity: A review of current understanding, therapies and relevance of phyto-therapeutics. *Heliyon* **10**, e23114. https://doi.org:10.1016/j.heliyon.2023.e23114 (2024).
28. Hotamisligil, G. S. Inflammation and metabolic disorders. *Nature* **444**, 860–867. https://doi.org:10.1038/nature05485 (2006).
29. Lumeng, C. N. & Saltiel, A. R. Inflammatory links between obesity and metabolic disease. *J. Clin. Invest.* **121**, 2111–2117. https://doi.org:10.1172/jci57132 (2011).
30. Shao, Y. et al. Immune and metabolic remodeling following bariatric surgery: Implications for targeted immunotherapy (Review). *Int. J. Mol. Med.* **57** https://doi.org:10.3892/ijmm.2025.5676 (2026).
31. Sun, K., Kusminski, C. M. & Scherer, P. E. Adipose tissue remodeling and obesity. *J. Clin. Invest.* **121**, 2094–2101. https://doi.org:10.1172/jci45887 (2011).
32. Sun, K., Li, X. & Scherer, P. E. Extracellular Matrix (ECM) and Fibrosis in Adipose Tissue: Overview and Perspectives. *Compr. Physiol.* **13**, 4387–4407. https://doi.org:10.1002/cphy.c220020 (2023).
33. Hua, Z. G. et al. Glucose and Insulin Stimulate Lipogenesis in Porcine Adipocytes: Dissimilar and Identical Regulation Pathway for Key Transcription Factors. *Mol. Cells*. **39**, 797–806. https://doi.org:https://doi.org/10.14348/molcells.2016.0144 (2016).
34. Hayakawa, J., Seki, T., Kawazoe, Y. & Ohe, K. Pathway importance by graph convolutional network and Shapley additive explanations in gene expression phenotype of diffuse large B-cell lymphoma. *PLOS ONE*. **17**, e0269570. https://doi.org:10.1371/journal.pone.0269570 (2022).
35. Ghaben, A. L. & Scherer, P. E. Adipogenesis and metabolic health. *Nat. Rev. Mol. Cell. Biol.* **20**, 242–258. https://doi.org:10.1038/s41580-018-0093-z (2019).
36. Karastergiou, K., Smith, S. R., Greenberg, A. S. & Fried, S. K. Sex differences in human adipose tissues – the biology of pear shape. *Biology Sex. Differences*. **3**, 13. https://doi.org:10.1186/2042-6410-3-13 (2012).
37. Palmer, B. F. & Clegg, D. J. The sexual dimorphism of obesity. *Mol. Cell. Endocrinol.* **402**, 113–119. https://doi.org:10.1016/j.mce.2014.11.029 (2015).
38. Ibrahim, M. M. Subcutaneous and visceral adipose tissue: structural and functional differences. *Obes. Rev.* **11**, 11–18. https://doi.org:10.1111/j.1467-789X.2009.00623.x (2010).
39. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63. https://doi.org:10.1038/nrg2484 (2009).

Tables

Table 1 is not available with this version.

Table 1. Baseline Clinical and Anthropometric Characteristics of the Study Cohort.

Figures

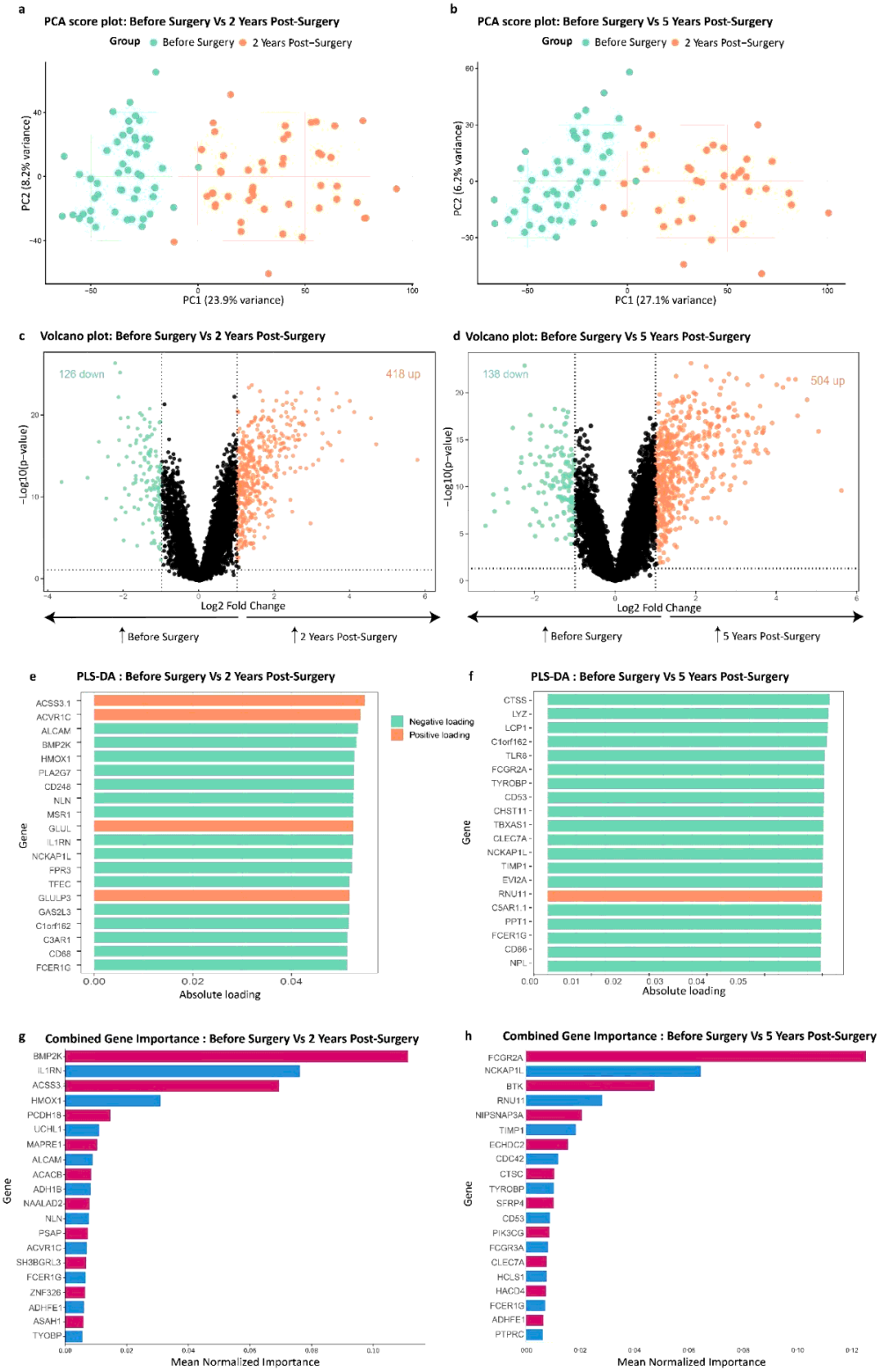

Figure 1

Overview of the statistical and machine-learning workflow for gene expression analysis following bariatric surgery.

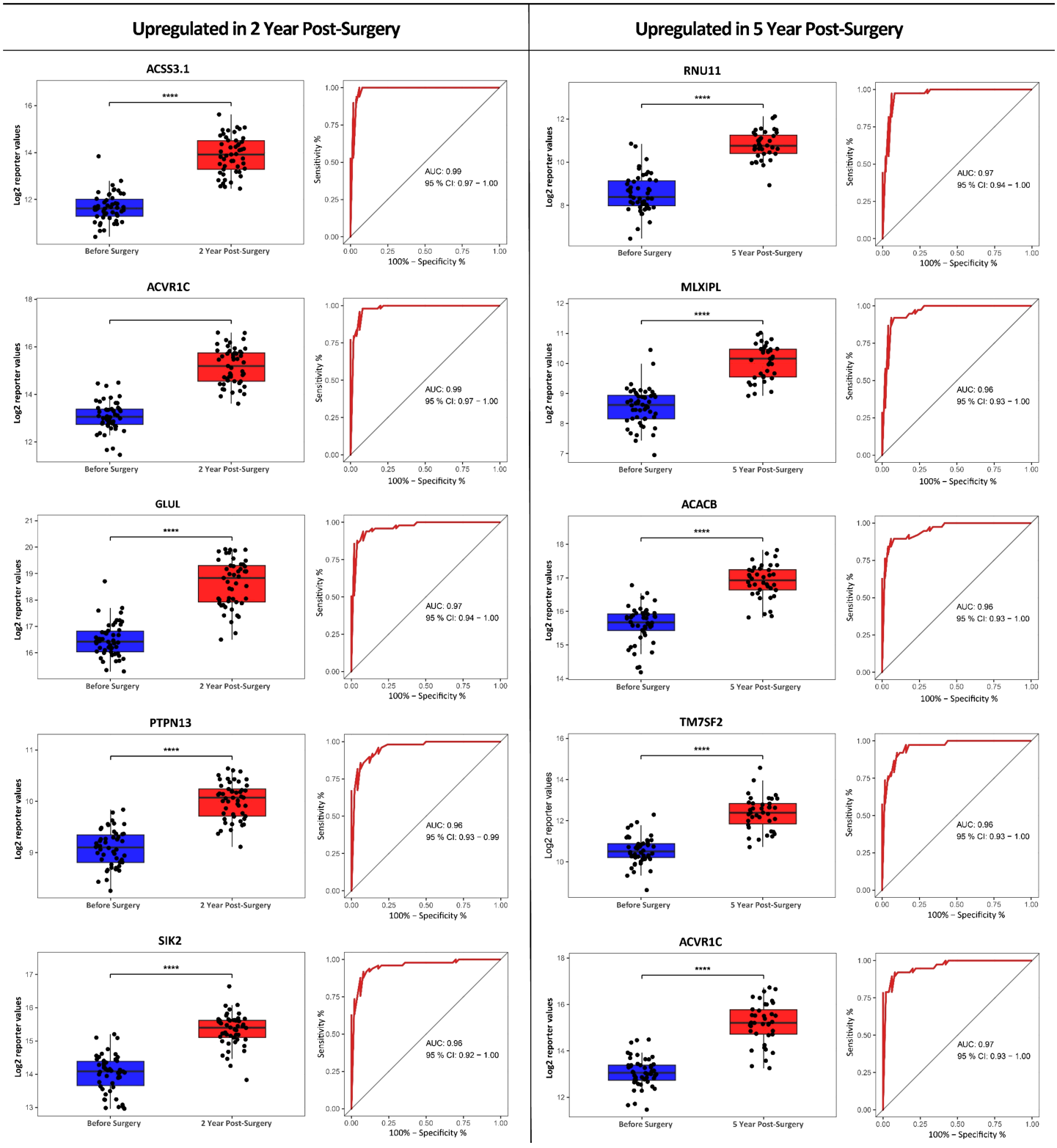

Figure 2

Global Transcriptomic Remodeling Following RYGB Surgery.
(A) Principal Component Analysis (PCA) of all samples, showing clear separation of the post-surgery groups (2-year and 5-year) from the baseline (Pre) group along PC1. (B) Volcano plot comparing the 5-year vs. baseline transcriptomes, highlighting significantly upregulated (red) and downregulated (blue) genes. (C) Partial Least Squares-Discriminant Analysis (PLS-DA)

confirming robust separation between timepoints. (D) PLS-DA loadings plot, indicating top genes contributing to the separation in panel C. (E) Upset plot illustrating the overlap of differentially expressed genes (DEGs) between the 2-year and 5-year timepoints.

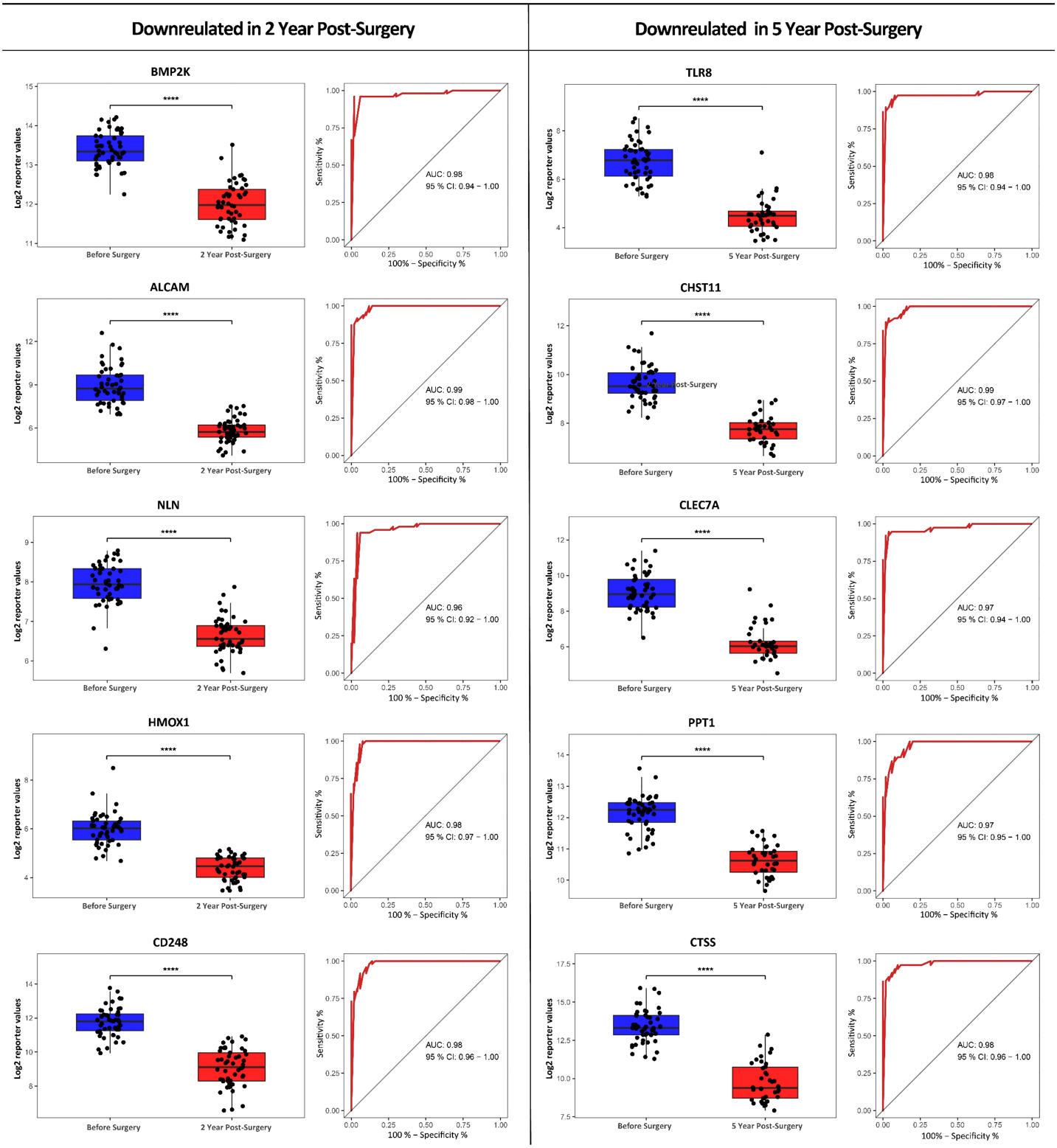

Figure 3

The figure displays data for a selection of the most significantly (A) upregulated and (B) downregulated genes. For each gene, boxplots (left) show normalized expression for baseline (Pre, n=50), 2-year (2Y, n=49), and 5-year (5Y, n=38) groups. Boxes denote the interquartile range (IQR) and whiskers 1.5×IQR. Statistical significance was determined by a paired t-test with FDR correction

(****, p < 0.0001). Corresponding ROC curves (right) illustrate each gene's individual classification performance, with the Area Under the Curve (AUC) noted.

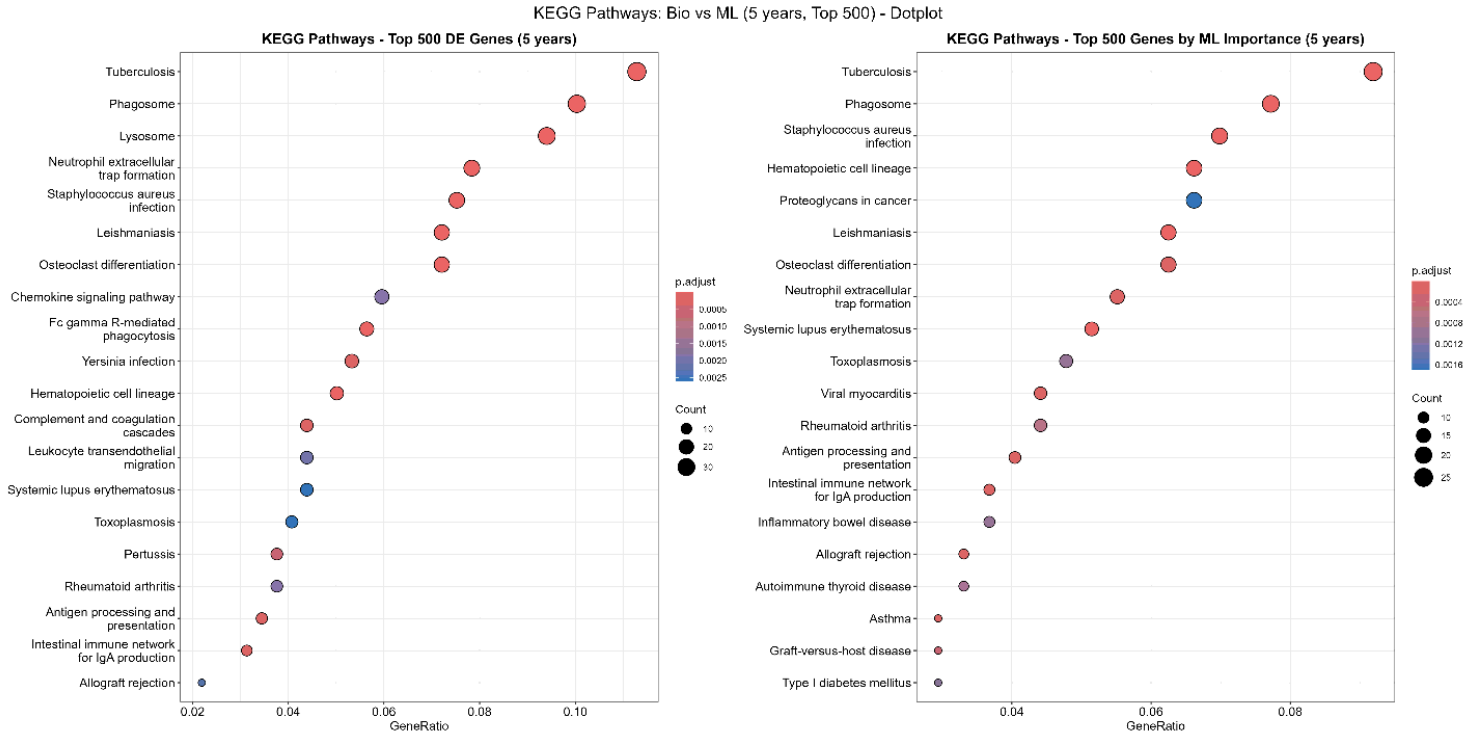

Figure 4

Concordance of GO Pathway Enrichment between Analytical Methods.
The dot plot contrasts the top Gene Ontology (GO) terms from the 50 most significant genes at 5 years post-surgery. Pathways were identified via a standard statistical analysis ('Bio', blue) versus the machine learning model's top features ('ML', red). The x-axis represents statistical significance (-log10(p-value)), showing strong concordance between the two methods.

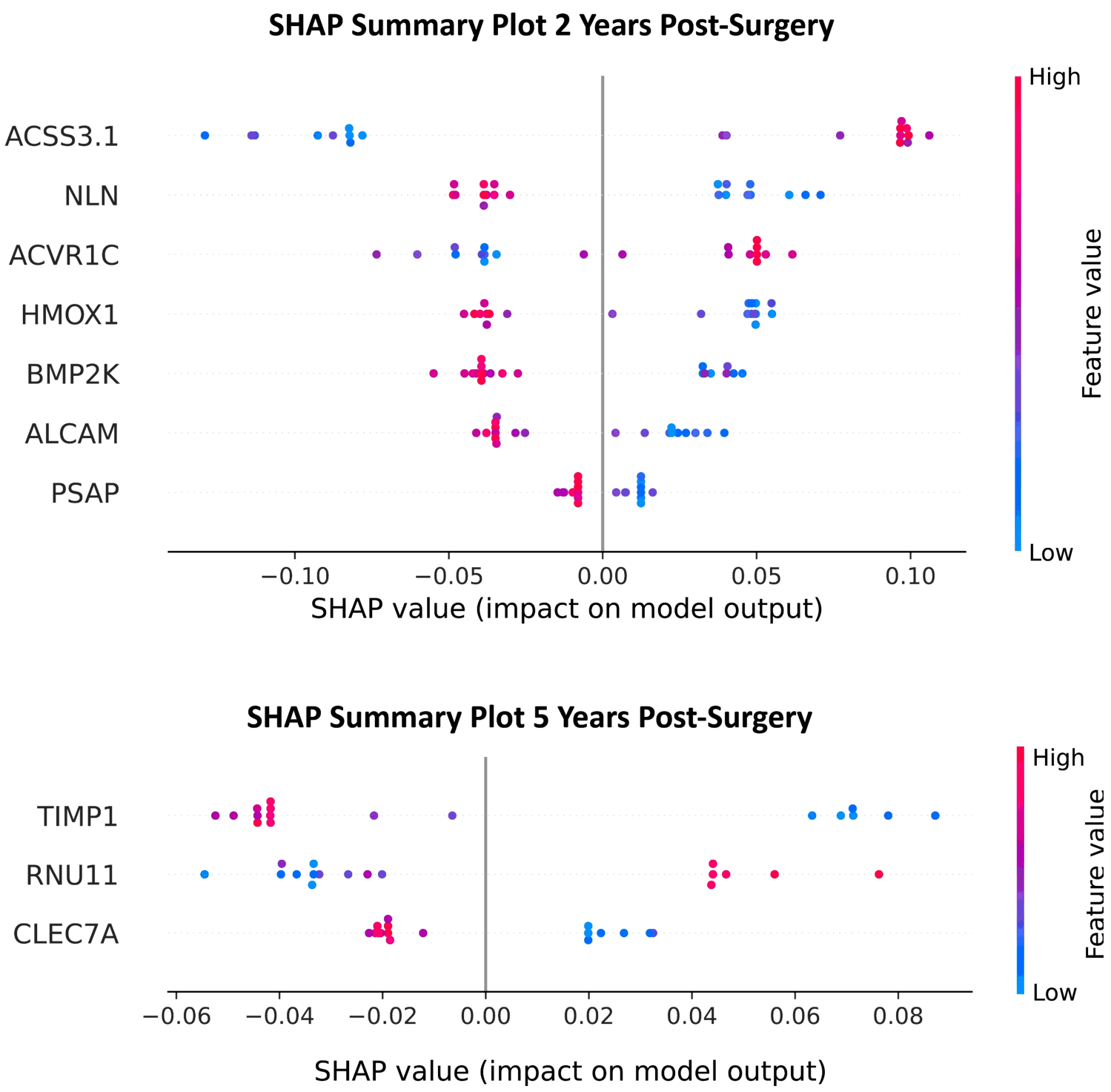

Figure 5

SHAP Summary Plots Revealing the Top Drivers of Model Predictions.
The plots illustrate the feature importance and impact of the top genes for the 2-year (left) and 5-year (right) predictive models. Each point represents a single gene for a single sample. The position on the x-axis indicates the SHAP value—the impact of that gene on the model's output, where positive values drive the prediction toward a "post-surgery" classification. The color represents the gene's expression level (red indicates high expression; blue indicates low).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryNaturereports.docx