

Transancestral TWAS Enrichment of NMN- Informed Pathways in Type 2 Diabetes: CPT2- Linked Fatty-Acid Oxidation and MAP2K1/MAPK3 Drivers as Translational Bridges

Ngo Cheung

info@cheungngomedical.com

Independent Researcher

Article

Keywords:

Posted Date: August 13th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10447512/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

Type 2 diabetes is a heterogeneous disease in which impaired insulin secretion, insulin resistance, altered lipid handling, and tissue-specific stress responses contribute in different proportions across individuals and populations. This heterogeneity complicates the development of broadly effective interventions and may contribute to variable responses to nicotinamide adenine dinucleotide precursors such as nicotinamide mononucleotide and nicotinamide riboside. The present study examined whether pathways linked to NMN-associated transcriptional rescue in aging tissues also carry genetic associations with type 2 diabetes across diverse ancestry groups.

We analyzed S-PrediXcan TWAS results from five type 2 diabetes analysis groups—African American, East Asian, European, Hispanic, and South Asian—using 10 NMN-informed KEGG gene sets. The pipeline combined tissue-level Z-scores using a Stouffer framework, evaluated competitive enrichment by permutation, estimated bootstrap confidence intervals, and performed gene-level leave-one-out influence analysis. Recurrent genes were ranked according to their ancestry breadth, pathway breadth, and influence on the aggregate Stouffer statistic.

All 10 gene sets reached the prespecified absolute Stouffer Z threshold in at least one analysis group. The strongest individual signal was a negative fatty-acid-degradation association in the European group, with Stouffer $Z = -9.70$ and permutation $p = 0.0077$. Autophagy in the African American group was the only other gene-set-by-ancestry result that met the joint absolute-Z and permutation criteria. Stage 3 analysis identified 548 unique leave-one-out-influential genes, including 32 genes observed across at least four ancestry groups. MAP2K1 and MAPK3 were the most recurrent multi-pathway drivers, whereas CPT2 was especially notable because it was directionally concordant across four groups and had also been identified in the NMN-responsive gene analysis. The collapsed top-30 driver module was protective-leaning in the European and Hispanic groups but risk-leaning in the East Asian group.

These findings provide a hypothesis-generating bridge between NMN-responsive metabolic programs and human type 2 diabetes genetics. They support ancestry-aware, pathway-stratified evaluation of NAD⁺ precursors and prioritize CPT2, MAP2K1, MAPK3, and selected trafficking genes for functional validation and biomarker development.

1. Introduction

1.1 Clinical problem and current difficulties

Type 2 diabetes is not a single biological disorder. Individuals diagnosed with the disease can differ substantially in age at onset, body composition, insulin secretion, insulin sensitivity, rate of disease progression, and susceptibility to vascular and microvascular complications. Large multi-ancestry genetic studies have reinforced this view by showing that diabetes risk is distributed across multiple biological processes, including beta-cell dysfunction, obesity-mediated insulin resistance, lipodystrophy, liver and lipid metabolism, and glycaemic regulation. These processes do not contribute equally across populations, and their genetic effects can differ in magnitude between ancestry groups [1–5].

The clinical consequences of this heterogeneity are considerable. A treatment that is effective in a patient whose disease is dominated by insulin resistance may be less effective in a patient whose primary defect is impaired beta-cell compensation. Similarly, an intervention designed to improve mitochondrial substrate utilization may be most relevant in individuals with lipid overload or impaired metabolic flexibility, but less relevant in patients whose dominant pathology is defective insulin secretion. Precision diabetes medicine therefore requires more than an overall disease-risk score. It requires approaches that distinguish the biological pathways through which inherited risk acts.

NAD⁺ precursors have attracted interest as potential metabolic and aging interventions. In a randomized trial of postmenopausal women with prediabetes and overweight or obesity, NMN supplementation increased skeletal-muscle insulin sensitivity and insulin signaling over 10 weeks [6]. By contrast, a randomized trial of nicotinamide riboside in obese, insulin-resistant men did not improve insulin sensitivity over 12 weeks, although the treatment was reported to be well tolerated [7]. These results do not establish that NAD⁺ precursors are ineffective. Rather, they suggest that response may depend on baseline metabolic state, tissue biology, dose, treatment duration, sex, age, and other sources of biological variation.

The lack of predictive biomarkers is a major obstacle to the clinical translation of NAD⁺ precursor research. At present, there is no established method for identifying individuals who are most likely to benefit from NMN or NR, nor is there a validated framework for determining whether an intervention should be expected to work similarly across ancestry groups. This problem is not unique to NAD⁺ biology. Polygenic scores developed primarily in European-ancestry populations can show reduced portability in other populations, and even within a broadly defined ancestry group, prediction accuracy can vary with study design, socioeconomic status, age, sex, and other characteristics of the discovery and target samples [8, 9].

Increasing population diversity improves discovery, fine-mapping, and the potential transferability of genetic findings, but diversity must be interpreted carefully. Genetic ancestry is a continuous property of human demographic history, not a set of discrete biological categories. In addition, ancestry-associated differences in allele frequencies, linkage disequilibrium, environmental exposures, ascertainment, and phenotype definitions can all influence genetic association results [10, 11].

1.2 Mechanistic rationale from NMN biology

The rationale for the present analysis came from an NMN-focused re-analysis of the GSE85718 mouse study originally reported by Mills and colleagues. In the supplied analysis, 35 genes showed age-associated expression rescue in at least two of three metabolic tissues—skeletal muscle, liver, and white adipose tissue—after long-term NMN administration. Twenty-six genes were directionally consistent across the tissues in which they were rescued, while nine showed discordant effects. No gene was rescued in all three tissues. These findings suggested that NMN-associated transcriptional regulation is tissue-dependent rather than a uniform, organism-wide reversal of aging-related expression changes [12, 13].

The NMN-derived gene set highlighted several biological themes. RAB11A was prioritized as a trafficking-related gene, whereas CPT2 was identified as a mitochondrial fatty-acid-oxidation gene. Other candidate genes included MAP2K2, MYO15, GALR1, SLC13A5, TRP53INP2, and APBA2. Because the re-analysis was based on a small mouse experiment and secondary cross-tissue comparisons, the results should be viewed as mechanistic leads rather than definitive evidence of therapeutic targets.

NMN is a precursor in NAD⁺ biosynthesis. In experimental models, increasing NAD⁺ availability can affect NAD⁺-dependent enzymes, including SIRT1 and SIRT3. SIRT1 regulates transcriptional coactivator activity involving PGC-1 α , while SIRT3 regulates mitochondrial protein acetylation and fatty-acid oxidation. Together, these pathways provide a plausible connection between NAD⁺ repletion, mitochondrial substrate use, and metabolic adaptation [14–17]. The important qualification is that these relationships are not equivalent to a direct demonstration that NMN increases CPT2 expression through a single SIRT1–SIRT3–PGC-1 α pathway in humans. They provide a biological framework for testing such a hypothesis.

The carnitine shuttle is central to the oxidation of long-chain fatty acids. CPT1 generates long-chain acyl-carnitines at the outer mitochondrial membrane, whereas CPT2 converts acyl-carnitines back to acyl-CoA inside the mitochondrial

matrix, allowing beta-oxidation to proceed. This pathway is essential for energy homeostasis, particularly in skeletal muscle, liver, heart, and other tissues with substantial fatty-acid utilization [18, 19].

The relationship between fatty-acid oxidation and insulin sensitivity is complex. Studies have linked lipid overload and incomplete fatty-acid oxidation to the accumulation of metabolites that can interfere with insulin signaling, including diacylglycerols and ceramides [20, 21, 24]. However, other work has emphasized that excessive mitochondrial substrate flux, acylcarnitine production, and redox stress may also contribute to insulin resistance, and that increased oxidative capacity is not automatically beneficial under every physiological condition [22, 23]. This complexity is directly relevant to the interpretation of CPT2. A genetic association consistent with higher predicted CPT2 expression and lower diabetes liability should be considered a testable hypothesis, not proof that indiscriminately increasing CPT2 activity will improve metabolic health.

The NMN analysis also suggested a second mechanistic theme involving signal transduction and membrane trafficking. MAPK signaling integrates growth-factor, insulin, stress, and inflammatory inputs. RAB11A-mediated recycling and DNM2-dependent vesicle scission are relevant to the movement of membrane proteins, glucose transporters, and secretory vesicles. DNM2 has been implicated in pancreatic beta-cell secretion, while more recent work has linked dynamin-2 regulation to GLUT4 endocytosis in skeletal muscle [25–27].

1.3 Gap and study aims

The central unanswered question was whether the pathways and genes highlighted by the NMN-responsive mouse analysis also carry genetic associations with human type 2 diabetes. A positive result would not demonstrate that NMN treats diabetes or that any particular gene is causal. It would, however, provide a rational link between an aging-related transcriptional program and human metabolic disease genetics.

To address this question, we analyzed 10 standard KEGG gene sets informed by the NMN-responsive gene analysis. The gene sets included endocytosis, fatty-acid degradation, PPAR signaling, MAPK signaling, RAS signaling, regulation of the actin cytoskeleton, autophagy, synaptic vesicle cycling, neuroactive ligand–receptor interaction, and the citrate cycle. KEGG is a curated framework for representing gene functions in metabolic and regulatory networks [28].

The study had three aims. First, we quantified pathway-level TWAS enrichment across five type 2 diabetes analysis groups. Second, we identified genes that repeatedly influenced the enrichment statistics across pathways and analysis groups using leave-one-out analysis. Third, we evaluated whether the resulting pathway and gene-level patterns could inform ancestry-aware trial design, intermediate biomarker development, and prioritization of candidate metabolic targets.

2. Methods

2.1 Multi-ancestry type 2 diabetes TWAS resource

The analysis (Fig. 1) used S-PrediXcan results corresponding to the type 2 diabetes genetic architecture described by Suzuki and colleagues. The five analysis groups were labeled AFA, EAS, EUR, HIS, and SAS, corresponding to African American, East Asian, European, Hispanic, and South Asian study groups in the supplied pipeline [29, 30]. These labels represent genetically and statistically grouped study collections and should not be interpreted as discrete biological races.

The upstream genetic study aggregated genome-wide association evidence across multiple populations and emphasized the heterogeneity of type 2 diabetes pathophysiology. It identified numerous association signals and

grouped them into mechanistic clusters involving beta-cell dysfunction, obesity, lipodystrophy, liver and lipid metabolism, body fat, metabolic syndrome, and residual glycaemic effects [3]. The present work did not reanalyze individual-level genotype data or reconstruct the original GWAS. It consumed the S-PrediXcan output files supplied for the five analysis groups. (nature.com)

2.2 Gene-set construction

The gene sets were constructed from standard KEGG pathway definitions and selected from a collection of NMN-informed gene sets derived from the supplied NMN/GSE85718 analysis. The 10 sets were KEGG endocytosis, fatty-acid degradation, PPAR signaling, MAPK signaling, RAS signaling, regulation of the actin cytoskeleton, autophagy in animals, synaptic vesicle cycle, neuroactive ligand–receptor interaction, and the citrate cycle.

The relevant KEGG pathway identifiers were hsa04144 for endocytosis, hsa00071 for fatty-acid degradation, hsa03320 for PPAR signaling, hsa04010 for MAPK signaling, hsa04014 for RAS signaling, hsa04810 for regulation of the actin cytoskeleton, hsa04140 for autophagy, hsa04721 for the synaptic vesicle cycle, hsa04080 for neuroactive ligand–receptor interaction, and hsa00020 for the citrate cycle. These pathway labels were used as biological annotations for the supplied gene sets. The current analysis tested whether the genes within those sets showed aggregate TWAS evidence for type 2 diabetes; it did not constitute a separate de novo KEGG over-representation analysis using an independently defined background.

The upstream NMN analysis was based on a small mouse experiment and was not designed as a human genetic study. It therefore served as a biological prior for pathway selection. The use of the NMN-informed sets was intended to test whether these pathways were also represented in human type 2 diabetes genetics, not to assume that a mouse transcriptional response must correspond directly to a human disease mechanism.

2.3 Statistical pipeline

The supplied pipeline first loaded the gene-level S-PrediXcan results from each tissue file. For each gene, it collected the available tissue-level Z-scores and calculated a cross-tissue meta-Z as the sum of the tissue-level Z-scores divided by the square root of the number of contributing tissues. This is a Stouffer-type combination. Gene names were harmonized using gene symbols and available Ensembl identifiers, and duplicate rows within files were reduced by retaining the first occurrence.

For each gene set and analysis group, the pipeline calculated the number of genes found, the percentage of genes covered, the Stouffer Z, the mean Z, the mean absolute Z, the median Z, and a trimmed mean. It also calculated a Wilcoxon signed-rank p-value against zero when enough genes were available. Competitive permutation p-values were generated by sampling genes from the background distribution of available meta-Z scores without replacement. The run used 10,000 permutations for the primary enrichment analyses and 2,000 bootstrap resamples for percentile confidence intervals. The random seed was 42.

The robust option was enabled. For statistics affected by extreme values, the pipeline winsorized Z-scores at the first and ninety-ninth percentiles. The code retained non-winsorized values for several nonparametric tests and calculated both ordinary and winsorized Stouffer statistics. The analysis also generated QQ plots, profile correlations, Euclidean and Manhattan distances, and hierarchical clustering based on one minus the Pearson correlation.

Gene-level influence was assessed by leave-one-out analysis. For each gene present in a gene set and analysis group, the pipeline recalculated the Stouffer statistic after removing that gene. A gene was flagged as influential if its removal changed the absolute Stouffer statistic by more than 20 percent or caused a sign change. In the downstream

analysis, a high-impact gene was defined more stringently as having an absolute change in Stouffer Z greater than 0.30 or a leave-one-out sign change.

Pairwise comparisons between analysis groups were conducted using Mann–Whitney U, Kolmogorov–Smirnov, Welch t, Levene, Brunner–Munzel, paired t, paired Wilcoxon, sign-flip permutation, sign concordance, Lin's concordance correlation coefficient, and Kendall's tau. Profile proximity was evaluated with Pearson and Spearman correlations and cosine similarity. Family-wise FDR was calculated by pooling p-values produced within the run, including enrichment, differential, proximity, and pairwise tests. The multi-gene-set run was configured with robust statistics, bootstrap resampling, 10,000 primary permutations, and family-wise FDR.

The downstream priority score combined the number of analysis groups exceeding the absolute Stouffer threshold, the number with permutation p below 0.05, the mean absolute Z, whether the direction was consistent, and the magnitude of the cross-ancestry meta-Z. Stage 3 then aggregated influential genes according to ancestry breadth, pathway breadth, and mean absolute leave-one-out influence. The Stage 3 transancestral score was calculated as twice the number of analysis groups, plus the number of gene sets, plus mean absolute leave-one-out influence.

A core driver was defined as a gene observed in at least three analysis groups and at least two gene sets. A transancestral driver was defined as a gene observed in at least four analysis groups. Directional concordance was calculated from the mean gene-level Z-score within each analysis group. The top-30 driver module was formed from the highest-ranked Stage 3 genes and then collapsed by analysis group by averaging one gene-level Z-score per gene across gene sets before calculating a module-level approximate Stouffer statistic.

2.4 Translational framing criteria

Signals were prioritized when they combined statistical strength, recurrence across analysis groups or pathways, directional consistency, and biological relevance to actionable processes. Particular attention was given to CPT2 and fatty-acid oxidation, MAP2K1 and MAPK3 within MAPK signaling, and DNM2 and related genes within endocytosis and vesicle dynamics.

The translational interpretation was deliberately conservative. A negative TWAS direction was treated as a protective-leaning association with type 2 diabetes liability, not as proof that increasing expression would be therapeutic. Likewise, a leave-one-out driver was interpreted as a gene that contributes strongly to an aggregate statistic, not as a confirmed causal gene.

3. Results

3.1 Pathway-level multi-ancestry enrichment

All 10 gene sets crossed the prespecified absolute Stouffer threshold of 1.96 in at least one analysis group. The downstream priority analysis therefore classified all 10 pathways as having at least one ancestry-level signal. This result should be understood as an uncorrected aggregate-Z criterion; it does not mean that every pathway had a permutation-supported or family-wise FDR-supported association (Table 1).

Table 1
Gene-Set Priority and Cross-Ancestry Enrichment

Rank	Gene set	Evidence score	Significant ancestries	Permutation-significant ancestries	Mean Z	Mean absolute Z	Cross-meta Z	Direction consistent
1	KEGG ENDOCYTOSIS	11.67	3/5	0	2.20	2.20	4.93	True
2	KEGG SYNAPTIC VESICLE CYCLE	11.13	4/5	0	0.44	2.64	0.98	False
3	KEGG AUTOPHAGY ANIMAL	10.71	3/5	1	-0.95	2.14	-2.13	False
4	KEGG REGULATION OF ACTIN CYTOSKELETON	10.67	3/5	0	-2.09	2.33	-4.67	False
5	KEGG NEUROACTIVE LIGAND RECEPTOR INTERACTION	9.99	3/5	0	-1.62	2.18	-3.61	False
6	KEGG PPAR SIGNALING PATHWAY	9.14	2/5	0	-2.41	2.44	-5.40	False
7	KEGG FATTY ACID DEGRADATION	8.60	1/5	1	-1.88	3.00	-4.20	False
8	KEGG RAS SIGNALING PATHWAY	7.02	2/5	0	-1.41	1.44	-3.14	False
9	KEGG CITRATE CYCLE TCA CYCLE	6.50	2/5	0	0.78	1.63	1.75	False
10	KEGG MAPK SIGNALING PATHWAY	3.23	1/5	0	-0.04	1.19	-0.09	False

Endocytosis received the highest composite evidence score, 11.67. It crossed the absolute Stouffer threshold in three of the five analysis groups and showed a consistent positive direction in the significant ancestry-level results. However, no endocytosis result had a permutation p below 0.05. Its high priority score was therefore driven by repeated aggregate-Z evidence and directional consistency rather than by the strict combined pathway criterion.

The synaptic vesicle cycle also scored highly, with an evidence score of 11.13 and four of five analysis groups exceeding the absolute-Z threshold. Its mean Z was only 0.44 and the direction was inconsistent, indicating that the high score reflected the magnitude and recurrence of the signals rather than a shared biological direction. Autophagy scored 10.71, with three of five analysis groups crossing the absolute-Z threshold and one permutation-supported result. Regulation of the actin cytoskeleton scored 10.67 and was predominantly negative in the replicated results. Neuroactive ligand–receptor interaction scored 9.99 and was also predominantly negative. PPAR signaling scored 9.14, with a protective-leaning direction in the analysis groups that crossed the threshold.

Fatty-acid degradation ranked seventh by the composite cross-ancestry score, despite producing the most extreme individual enrichment result. This apparent discrepancy reflects the fact that the strongest signal was concentrated in the European analysis group and was not broadly replicated at the same threshold in the other groups. The pathway nevertheless remains central to the biological interpretation because its European result was the strongest single pathway-by-ancestry signal in the entire analysis.

The European fatty-acid-degradation result had Stouffer $Z = -9.704$, a 95% bootstrap confidence interval from -17.25 to -2.62 , permutation $p = 0.007699$, and a Wilcoxon $p = 0.02071$. The negative direction was interpreted as protective-leaning at the aggregate TWAS level. The top leave-one-out drivers reported for this particular publication-grade result were GCDH, ACADS, and EHHADH. CPT2 was not the leading leave-one-out contributor to this specific European aggregate result, although it emerged as a recurrent transancestral driver across the broader analysis and directly overlapped the NMN-responsive gene analysis.

Autophagy in the African American analysis group was the only other gene-set-by-ancestry signal to meet the joint criterion of an absolute Stouffer Z greater than 1.96 and permutation p below 0.05. Its Stouffer Z was -5.065 , with a 95% bootstrap interval from -8.62 to -1.68 and permutation $p = 0.0121$. The leading drivers listed for this signal were ATG5, RRAS2, and MAPK3. The publication-ready signals are summarized in Table 2, and the per-disease and tissue-specific enrichment results are summarized in Table 3.

Table 2
Publication-Ready Enrichment Signals

Gene set	Analysis group	n	Stouffer Z	95% bootstrap CI	Permutation p	Wilcoxon p	Leading drivers
KEGG FATTY ACID DEGRADATION	dm cvs EUR	38	-9.70	[- 17.25, - 2.62]	0.007699	0.02071	GCDH (- 1.61); ACADS (- 1.53); EHHADH (- 1.19)
KEGG AUTOPHAGY ANIMAL	dm cvs AFA	132	-5.06	[- 8.62, - 1.68]	0.0121	0.06213	ATG5 (- 0.60); RRAS2 (- 0.56); MAPK3 (- 0.43)

Table 3
Per-Disease and Tissue-Specific Enrichment Results

Gene set	Analysis group	Tissue	n	Stouffer Z	Permutation p	Wilcoxon p
KEGG FATTY ACID DEGRADATION	dm cvs EUR	Whole Blood spreadixcan	23	-7.058	NA	0.00166
KEGG CITRATE CYCLE TCA CYCLE	dm cvs AFA	Whole Blood spreadixcan	20	2.977	NA	0.002325
KEGG FATTY ACID DEGRADATION	dm cvs EUR	Not specified	38	-9.704	0.007699	0.02071
KEGG PPAR SIGNALING PATHWAY	dm cvs SAS	Muscle Skeletal spreadixcan	37	-3.292	NA	0.01056
KEGG AUTOPHAGY ANIMAL	dm cvs AFA	Adipose Subcutaneous spreadixcan	86	-3.284	NA	0.01141
KEGG AUTOPHAGY ANIMAL	dm cvs AFA	Not specified	132	-5.065	0.0121	0.06213
KEGG REGULATION OF ACTIN CYTOSKELETON	dm cvs EAS	Pancreas spreadixcan	98	-4.735	NA	0.0128
KEGG AUTOPHAGY ANIMAL	dm cvs EUR	Liver spreadixcan	73	5.854	NA	0.01545
KEGG NEUROACTIVE LIGAND RECEPTOR INTERACTION	dm cvs EAS	Muscle Skeletal spreadixcan	97	-3.727	NA	0.01551
KEGG CITRATE CYCLE TCA CYCLE	dm cvs AFA	Muscle Skeletal spreadixcan	19	2.024	NA	0.01808
KEGG SYNAPTIC VESICLE CYCLE	dm cvs EUR	Adipose Visceral Omentum spreadixcan	34	5.805	NA	0.01868
KEGG AUTOPHAGY ANIMAL	dm cvs AFA	Whole Blood spreadixcan	77	-3.370	NA	0.02101
KEGG AUTOPHAGY ANIMAL	dm cvs EUR	Muscle Skeletal spreadixcan	86	5.997	NA	0.02418
KEGG SYNAPTIC VESICLE CYCLE	dm cvs HIS	Muscle Skeletal spreadixcan	33	-4.160	NA	0.02533
KEGG FATTY ACID DEGRADATION	dm cvs EUR	Adipose Subcutaneous spreadixcan	20	-5.766	NA	0.02664
KEGG PPAR SIGNALING PATHWAY	dm cvs HIS	Muscle Skeletal spreadixcan	37	-2.163	NA	0.02784
KEGG PPAR SIGNALING PATHWAY	dm cvs EUR	Muscle Skeletal spreadixcan	37	-5.195	NA	0.0298
KEGG MAPK SIGNALING PATHWAY	dm cvs HIS	Liver spreadixcan	92	-3.477	NA	0.03191

Gene set	Analysis group	Tissue	n	Stouffer Z	Permutation p	Wilcoxon p
KEGG ENDOCYTOSIS	dm cvs HIS	Adipose Visceral Omentum spredixcan	126	3.501	NA	0.03385
KEGG PPAR SIGNALING PATHWAY	dm cvs EUR	Adipose Subcutaneous spredixcan	39	-4.210	NA	0.03996
KEGG REGULATION OF ACTIN CYTOSKELETON	dm cvs AFA	Muscle Skeletal spredixcan	120	3.488	NA	0.04627
KEGG ENDOCYTOSIS	dm cvs SAS	Whole Blood spredixcan	126	2.205	NA	0.04763

Six gene sets showed significance in at least two analysis groups with a consistent direction among the significant results. Regulation of the actin cytoskeleton, neuroactive ligand–receptor interaction, PPAR signaling, and RAS signaling were protective-leaning. Endocytosis and the citrate cycle were risk-leaning. Fatty-acid degradation did not meet this replicated same-direction criterion, because its strongest result was concentrated in the European group. The downstream summary also identified 20 significant profile-proximity pairs, indicating that several analysis groups had correlated gene-level profiles within particular gene sets. These correlations were strongest for RAS signaling between the European and Hispanic groups, for citrate-cycle profiles between the Hispanic and South Asian groups, and for endocytosis and MAPK-related profiles involving European and East Asian groups.

3.2 Recurrent and transancestral driver genes

The primary leave-one-out analysis identified 1,076 influential gene-by-analysis-group entries. After collapsing across gene sets and analysis groups, 548 unique genes were flagged as influential. These genes were not equivalent to genome-wide significant causal genes. They were genes whose removal changed the aggregate gene-set statistic beyond the prespecified threshold. The leading recurrent influential genes are summarized in Table 4.

Table 4
Top Recurrent Influential Genes

Gene	Entries	Ancestries	Gene sets	Mean absolute Δ Stouffer	Maximum absolute Δ Stouffer	Direction	Sign flip
MAP2K1	11	5	4	0.74	1.82	Negative / protective	No
MAPK3	14	5	4	0.56	1.32	Negative / protective	7% of entries
VEGFB	9	5	2	0.84	1.67	Negative / protective	No
CS	5	5	1	1.00	1.60	Positive / risk	40% of entries
RRAS2	10	4	4	0.41	0.61	Negative / protective	10% of entries
FGFR1	7	4	3	0.34	0.49	Positive / risk	14% of entries
RAC3	6	4	3	0.27	0.43	Positive / risk	17% of entries
AKT2	5	4	3	0.17	0.22	Positive / risk	No
MAP2K2	6	4	3	0.13	0.21	Negative / protective	No
MAPK9	5	4	3	0.09	0.17	Negative / protective	No
DNM2	6	4	2	1.01	2.21	Positive / risk	No
CPT2	6	4	2	0.86	1.81	Negative / protective	33% of entries
ACOX1	8	4	2	0.86	1.72	Positive / risk	No
PLA2G4B	7	4	2	0.58	1.26	Positive / risk	29% of entries
NFKB1	6	4	2	0.35	0.74	Negative / protective	17% of entries
EFNA5	5	4	2	0.22	0.28	Negative / protective	20% of entries
PIK3CA	5	4	2	0.20	0.26	Negative / protective	No
GNG12	5	4	2	0.19	0.37	Positive / risk	20% of entries
MAPK8	5	4	2	0.15	0.33	Positive / risk	No
PDGFA	5	4	2	0.13	0.21	Negative / protective	No
FGF17	5	4	2	0.12	0.17	Positive / risk	20% of entries

Gene	Entries	Ancestries	Gene sets	Mean absolute Δ Stouffer	Maximum absolute Δ Stouffer	Direction	Sign flip
MYL7	4	4	1	1.32	2.54	Positive / risk	No
CYP2U1	4	4	1	1.31	2.47	Positive / risk	25% of entries
ACLY	4	4	1	1.29	1.60	Negative / protective	75% of entries
BCAR1	4	4	1	0.87	2.06	Negative / protective	No
OGDH	4	4	1	0.68	1.22	Negative / protective	50% of entries
PDHB	4	4	1	0.62	0.72	Negative / protective	25% of entries
SDHD	4	4	1	0.47	0.85	Positive / risk	No
TAOK2	4	4	1	0.44	0.73	Positive / risk	No
DUSP6	4	4	1	0.43	0.73	Positive / risk	25% of entries

Stage 3 identified 32 genes in the strict transancestral tier, defined as influence observed in at least four analysis groups. Seventy-six genes were observed in exactly three analysis groups, 177 were bi-ancestral, and 263 were analysis-group-private. Sixty genes met the core-driver criterion of appearing in at least three analysis groups and at least two gene sets. This tier structure indicates that many genes can influence a particular pathway in a particular analysis group, but only a minority exerted broad leverage across multiple populations and biological modules.

MAP2K1 and MAPK3 were the strongest recurrent multi-pathway drivers. Each appeared across five analysis groups and four gene sets. MAP2K1 had 11 influential entries, a mean absolute leave-one-out change of 0.74, and a maximum change of 1.82. Its gene-level direction was negative in four of five groups. MAPK3 had 14 influential entries, a mean absolute change of 0.56, and a maximum change of 1.32. It was fully direction-concordant in the Stage 3 analysis, with negative or protective-leaning gene-level Z-scores in all five groups.

VEGFB was also observed in all five analysis groups and two gene sets, with a mean absolute leave-one-out change of 0.84 and a maximum of 1.67. Its dominant direction was negative. RRAS2 appeared in four analysis groups and four gene sets, while FGFR1 and RAC3 appeared in four groups and three gene sets. Other recurrent genes included AKT2, MAP2K2, MAPK9, DNM2, CPT2, ACOX1, PLA2G4B, NFKB1, EFNA5, PIK3CA, GNG12, MAPK8, PDGFA, and FGF17.

CPT2 was particularly notable for three reasons. First, it appeared in four analysis groups and two gene sets. Second, its direction was fully concordant and protective-leaning across those four groups in the Stage 3 analysis. Third, it was the only mitochondrial gene among the 35 NMN-responsive genes identified in the supplied GSE85718 re-analysis. CPT2 therefore represents a direct molecular overlap between the NMN-informed gene set and the type 2 diabetes TWAS analysis. This overlap is important, but it should not be confused with proof that CPT2 was the sole or dominant driver of the European fatty-acid-degradation signal.

DNM2 appeared in four analysis groups and two gene sets, with a mean absolute leave-one-out change of 1.01 and a maximum of 2.21. It was especially influential in endocytosis and the synaptic vesicle cycle. ACOX1 appeared in four analysis groups and two gene sets and had a mean absolute change of 0.86. Its recurrence provided additional

support for a fatty-acid-handling and peroxisomal oxidation theme. MAP2K2, MAPK9, and PIK3CA linked the autophagy, RAS, and MAPK-related modules.

Nineteen core drivers were fully direction-concordant across the analysis groups in which they appeared. Protective-leaning examples included MAPK3, CPT2, AKT1, INSR, NTRK1, PDGFD, and RASGRF2. Risk-leaning examples included FGFR1, PLA2G4B, HRAS, FGFR3, PIK3R3, CSF1, ACSL3, PCK1, VEGFA, and RRAS. MAP2K1 showed strong but incomplete concordance, with four negative and one positive analysis-group direction. This mixture pathway-level effects arise from coordinated networks rather than from a single universal direction of gene regulation.

3.3 Driver-module ancestry heterogeneity

The top-30 driver module comprised MAP2K1, MAPK3, VEGFB, CS, RRAS2, FGFR1, RAC3, AKT2, MAP2K2, MAPK9, DNM2, CPT2, ACOX1, PLA2G4B, NFKB1, EFNA5, PIK3CA, GNG12, MAPK8, PDGFA, FGF17, MYL7, CYP2U1, ACLY, BCAR1, OGDH, PDHB, SDHD, TAOK2, and DUSP6.

When these genes were collapsed into one module, the European group had a mean Z of - 1.96 and an approximate Stouffer statistic of - 10.75, with 60% of genes negative. The Hispanic group had a mean Z of - 0.86 and an approximate Stouffer statistic of - 4.72, with 57% of genes negative. In contrast, the East Asian group had a mean Z of 1.09 and an approximate Stouffer statistic of 5.99, with 47% of genes negative. The African American group was weaker and risk-leaning, with mean Z = 0.25 and approximate Stouffer Z = 1.37. The South Asian group was also modestly risk-leaning, with mean Z = 0.44 and approximate Stouffer Z = 2.40. The collapsed driver-module results are shown in Table 5.

Table 5
Top-30 Transancestral Driver Module by Analysis Group

Analysis group	Module size	Mean Z	Approximate Stouffer Z	Percentage negative	Interpretation
dm cvs AFA	30	0.25	1.37	40%	Risk-leaning
dm cvs EAS	30	1.09	5.99	47%	Risk-leaning
dm cvs EUR	30	-1.96	-10.75	60%	Protective-leaning
dm cvs HIS	30	-0.86	-4.72	57%	Protective-leaning
dm cvs SAS	30	0.44	2.40	53%	Risk-leaning

This pattern indicates that the top driver module was not uniformly protective across analysis groups. The result could reflect genuine ancestry-associated differences in effect-size distributions, differences in linkage disequilibrium and allele frequencies, differences in gene-expression prediction models, or study-level differences in phenotype and covariate composition. It should not be interpreted as evidence that the module has opposite causal effects in different human populations without additional fine-mapping and functional validation.

Variance heterogeneity was more prominent than mean separation in the pairwise analysis. Levene tests implicated RAS, endocytosis, autophagy, MAPK, neuroactive ligand-receptor interaction, actin-cytoskeleton regulation, synaptic vesicle cycling, fatty-acid degradation, PPAR signaling, and the citrate cycle. European comparisons appeared frequently among these signals. This pattern may partly reflect the greater sample size and power of European TWAS models, but it may also reflect real differences in the distribution of gene-level association effects.

3.4 Pathway crosstalk and biological coherence

The crosstalk analysis identified 14 gene-set pairs that shared at least one high-impact influential gene. The largest overlap was between MAPK and RAS signaling, with 20 shared genes and a Jaccard index of 0.22. This overlap is expected because RAS signaling is a principal upstream component of the canonical RAF–MEK–ERK cascade.

The second major overlap was between fatty-acid degradation and PPAR signaling, with 11 shared genes. These included CPT2, ACOX1, ACOX3, ACSL1, ACSL3, ACSL5, ACSL6, CPT1A, CPT1B, and EHHADH. This was the clearest metabolic crosstalk module and linked the NMN-responsive CPT2 signal to a broader fatty-acid-utilization program.

RAS and regulation of the actin cytoskeleton shared 11 high-impact genes, including FGFR1, FGFR3, HRAS, MAP2K1, MAPK3, PIK3R3, RAC3, RHOA, and RRAS2. MAPK and actin-cytoskeleton regulation shared eight genes. Autophagy also shared genes with RAS and MAPK signaling, including HRAS, MAP2K1, MAPK3, MAPK8, PIK3R3, and RRAS2. Endocytosis and the synaptic vesicle cycle shared AP2A1, AP2S1, CLTB, and DNMT2.

These overlaps support a connected network model involving substrate oxidation, growth-factor signaling, cytoskeletal regulation, autophagy, and vesicle trafficking. They also provide a mechanistic context for why the same genes repeatedly influenced several gene sets. The crosstalk analysis does not prove physical interaction among all highlighted genes, but it shows that the statistical drivers are concentrated in biologically related network components rather than being randomly distributed among unrelated pathways (Table 6).

Table 6
Pathway Crosstalk Among High-Impact Influential Genes

Gene-set pair	Shared genes	Jaccard index
KEGG MAPK SIGNALING PATHWAY ↔ KEGG RAS SIGNALING PATHWAY	20	0.22
KEGG FATTY ACID DEGRADATION ↔ KEGG PPAR SIGNALING PATHWAY	11	0.20
KEGG RAS SIGNALING PATHWAY ↔ KEGG REGULATION OF ACTIN CYTOSKELETON	11	0.12
KEGG MAPK SIGNALING PATHWAY ↔ KEGG REGULATION OF ACTIN CYTOSKELETON	8	0.10
KEGG AUTOPHAGY ANIMAL ↔ KEGG RAS SIGNALING PATHWAY	7	0.06
KEGG ENDOCYTOSIS ↔ KEGG REGULATION OF ACTIN CYTOSKELETON	6	0.07
KEGG AUTOPHAGY ANIMAL ↔ KEGG REGULATION OF ACTIN CYTOSKELETON	5	0.06
KEGG AUTOPHAGY ANIMAL ↔ KEGG MAPK SIGNALING PATHWAY	5	0.05
KEGG ENDOCYTOSIS ↔ KEGG SYNAPTIC VESICLE CYCLE	4	0.05
KEGG ENDOCYTOSIS ↔ KEGG RAS SIGNALING PATHWAY	3	0.03
KEGG NEUROACTIVE LIGAND RECEPTOR INTERACTION ↔ KEGG REGULATION OF ACTIN CYTOSKELETON	2	0.03
KEGG ENDOCYTOSIS ↔ KEGG MAPK SIGNALING PATHWAY	2	0.02
KEGG CITRATE CYCLE TCA CYCLE ↔ KEGG PPAR SIGNALING PATHWAY	1	0.02
KEGG AUTOPHAGY ANIMAL ↔ KEGG ENDOCYTOSIS	1	0.01

4. Discussion

4.1 Principal findings in clinical context

This study identified a relationship between NMN-informed metabolic pathways and human type 2 diabetes TWAS architecture across five ancestry-analysis groups. The most robust individual result was the negative fatty-acid-degradation signal in the European group. The most reproducible cross-ancestry driver signals involved MAP2K1 and MAPK3, while CPT2 provided the clearest direct overlap between the NMN-responsive mouse gene analysis and the human type 2 diabetes TWAS results.

The findings do not establish that NMN, NR, CPT2, MAP2K1, MAPK3, or DNM2 is an effective treatment target. They instead identify a group of pathways and genes that are plausible candidates for testing in genetically stratified studies. The central result is therefore not a single therapeutic claim. It is a framework in which NAD⁺ precursor response, mitochondrial substrate utilization, signaling-state regulation, and vesicle trafficking can be evaluated together and in an ancestry-aware manner.

4.2 Mechanistic synthesis

CPT2 and the fatty-acid-oxidation axis

CPT2 completes the mitochondrial carnitine shuttle by converting long-chain acyl-carnitines back to acyl-CoA within the mitochondrial matrix (Fig. 2). This reaction permits the subsequent beta-oxidation of long-chain fatty acids. In tissues such as skeletal muscle and liver, the pathway contributes to energy production and metabolic flexibility, particularly during fasting or increased energy demand [18, 19].

The European fatty-acid-degradation result was strongly negative at the aggregate TWAS level. A negative direction is compatible with lower type 2 diabetes liability under genetically predicted higher expression of one or more genes in the set, but it cannot by itself establish that higher CPT2 expression is protective. The leading drivers of the European publication-grade result were GCDH, ACADS, and EHHADH. CPT2 was a recurrent cross-ancestry driver and a direct NMN overlap, but it was not the top leave-one-out contributor to that specific European pathway-level statistic.

The biological interpretation remains plausible. Accumulation of ectopic lipid metabolites, including diacylglycerols and ceramides, can interfere with insulin signaling in liver and skeletal muscle. Incomplete fatty-acid oxidation and mitochondrial overload have also been linked to insulin resistance in experimental models [20, 21, 23].

At the same time, the relationship between fatty-acid oxidation and insulin action is not linear. Muoio and Neufer emphasized that excessive mitochondrial substrate load, acylcarnitine production, and redox stress may contribute to insulin resistance, challenging the assumption that more fatty-acid oxidation is always beneficial [22]. In addition, a muscle-specific CPT2 deletion model unexpectedly protected mice from diet-induced obesity and insulin resistance despite long-chain acylcarnitine accumulation, demonstrating that CPT2 biology is context-dependent [31].

For these reasons, the appropriate interpretation is that CPT2 identifies a candidate metabolic axis rather than a validated therapeutic target. The most defensible hypothesis is that the balance among fatty-acid uptake, mitochondrial entry, beta-oxidation, TCA-cycle capacity, lipid storage, and redox control determines whether changes in the CPT2-centered program are beneficial or harmful.

The NMN connection is biologically coherent but remains indirect. Experimental work has shown that NAD⁺-dependent sirtuin pathways can regulate mitochondrial metabolism. SIRT1 interacts with PGC-1 α -related transcriptional programs, and SIRT3 regulates fatty-acid oxidation through reversible deacetylation of mitochondrial enzymes [14, 15]. (pmc.ncbi.nlm.nih.gov) The present data therefore support a testable model in which NAD⁺

repletion could influence metabolic flexibility through a CPT2-linked program. They do not demonstrate that NMN acts through CPT2 in humans or that genetically predicted CPT2 expression modifies NMN response.

MAP2K1 and MAPK3

MAP2K1 and MAPK3 occupy central positions in the RAF–MEK–ERK signaling cascade. MAP2K1 encodes MEK1, and MAPK3 encodes ERK1. The pathway integrates signals from growth-factor receptors, insulin-related inputs, inflammatory mediators, and cellular stress. The consequences of ERK signaling depend on signal strength, duration, subcellular localization, scaffold proteins, and crosstalk with other pathways [32].

The Stage 3 results placed MAP2K1 and MAPK3 at the top of the recurrent driver hierarchy. Both were observed across five analysis groups and four gene sets, principally involving MAPK, RAS, actin-cytoskeleton, and autophagy-related signals. MAPK3 showed fully concordant negative direction across all five groups, while MAP2K1 was negative in four and positive in one. This pattern is compatible with a protective-leaning association within the range of common genetically predicted expression effects, but it does not mean that sustained activation of the pathway is necessarily beneficial.

Experimental evidence supports a context-dependent interpretation. MEK/ERK signaling participates in beta-cell mass regulation and glucose-stimulated insulin secretion, and its activity may help maintain glucose homeostasis under some conditions [33]. The same signaling system can also contribute to adipocyte differentiation, inflammatory responses, and insulin resistance when activation is excessive or chronically mislocalized. Thus, the direction observed in the TWAS analysis may reflect a change in signaling balance or cellular resilience rather than a simple “more is better” relationship.

The translational implication is not to activate MAP2K1 or MAPK3 indiscriminately. Broad MEK inhibition may be metabolically undesirable if it disrupts beta-cell compensation, insulin signaling, or tissue repair. A more plausible strategy would be to characterize pathway dynamics, including phosphorylation state, duration of activation, and tissue-specific localization. MAP2K1 and MAPK3 may be most useful as components of a signaling-flexibility biomarker rather than as isolated drug targets.

Trafficking and endocytosis

DNM2 was a recurrent Stage 3 driver in four analysis groups and two gene sets. It was particularly influential in endocytosis and the synaptic vesicle cycle. DNM2 is a large GTPase involved in membrane scission and endocytic vesicle formation. Its function is relevant to secretory-cell physiology and to the recycling and internalization of membrane proteins.

In pancreatic beta-cells, DNM2 deficiency has been shown to impair biphasic insulin secretion and disturb glucose homeostasis [25]. In skeletal muscle, insulin signaling regulates DNM2-dependent GLUT4 endocytosis through GSK3 α and Bin1, linking dynamin activity to glucose transporter residence at the cell surface [26, 27]. These findings make DNM2 a plausible bridge between the endocytosis pathway identified in the TWAS analysis and the trafficking theme highlighted by RAB11A in the NMN-responsive mouse analysis.

The relationship to RAB11A should nevertheless be described as a network-level hypothesis. RAB11A regulates recycling endosomes, whereas DNM2 is directly involved in vesicle scission. Their functions can converge on membrane-protein recycling, but the present analysis did not demonstrate that NMN changes DNM2 or RAB11A activity in human metabolic tissues. Similarly, the supplied analysis did not directly measure insulin-receptor recycling, GLUT4 trafficking, endosomal pH, or mitochondrial-derived vesicles.

The most informative follow-up experiments would therefore measure trafficking competence directly. These could include insulin-receptor surface recovery, GLUT4 internalization and recycling, endosomal cargo flow, vesicle acidification, and beta-cell secretory dynamics. Such assays would provide a more direct test of whether the statistical convergence between endocytosis, synaptic vesicle cycling, and the NMN-responsive trafficking program corresponds to altered cellular function.

4.3 Translational opportunities and how they address current difficulties

Addressing heterogeneous response to NMN and NR

The first translational difficulty is the variable response to NAD⁺ precursors. NMN improved insulin sensitivity in one small randomized study of prediabetic women, whereas NR did not improve insulin sensitivity in a separate study of obese, insulin-resistant men [6, 7]. The studies differed in precursor, dose, duration, sex, age, metabolic phenotype, and design, so they cannot be compared as a direct test of one universal treatment effect. Nevertheless, they illustrate why baseline biology and treatment context are likely to matter.

The present results suggest one testable approach: pre-stratify participants according to a fatty-acid-degradation or CPT2-linked pathway score. Individuals with low genetically predicted fatty-acid-utilization capacity might be hypothesized to benefit more from an intervention that improves NAD⁺ availability and metabolic flexibility. This hypothesis should be tested prospectively rather than assumed. An adaptive enrichment design could initially enroll a broad population, estimate treatment effects across prespecified genetic strata, and then expand recruitment in groups showing evidence of differential response.

A second possibility is to use the top-30 driver module as a broader pathway score. Because the module includes fatty-acid oxidation, MAPK/RAS signaling, actin regulation, vesicle trafficking, and TCA-cycle genes, it may capture a more integrated metabolic state than a CPT2-only score. However, the ancestry-dependent direction of this module means that it cannot be used without ancestry-specific calibration and external validation.

Addressing ancestry-related disparities and trial generalizability

The second difficulty is that genetic prediction and clinical trial generalizability can vary across ancestry groups. European-centric polygenic scores may not transfer reliably to other populations, and prediction accuracy can vary even within a broadly defined ancestry group [8, 9]. The present analysis showed a related problem at the pathway level: the collapsed top-30 driver module was strongly protective-leaning in the European and Hispanic groups but risk-leaning in the East Asian group.

This pattern does not prove that NMN would be beneficial in one population and harmful in another. It does show that a pathway score trained or interpreted in one group should not be assumed to have the same meaning elsewhere. Any future genetic stratification strategy should therefore estimate score distributions, effect sizes, calibration, and treatment interactions separately across the relevant analysis groups.

Clinical trials should also be powered to detect ancestry-specific efficacy rather than treating diversity as a descriptive subgroup characteristic. That would require prespecified interaction analyses, adequate recruitment of underrepresented populations, ancestry-calibrated scoring procedures, and careful reporting of model performance. The aim should not be to assign individuals to rigid biological categories, but to test whether the relationship between a molecular score and treatment response differs across populations and contexts.

Developing intermediate biomarkers

The third difficulty is the lack of practical intermediate biomarkers of NMN target engagement. The CPT2-centered interpretation suggests several possibilities. Plasma acyl-carnitines could provide a minimally invasive measure of altered fatty-acid transport and oxidation, although their interpretation must account for the fact that acylcarnitines can be both markers and regulators of mitochondrial metabolic stress. Additional measures could include lipid intermediates, ketone-body handling, indirect calorimetry, and muscle metabolic flux.

CPT2-related biomarkers could also include muscle expression or protein abundance, although tissue sampling would limit scalability. Imaging-based assessments of intramyocellular lipid, mitochondrial function, or substrate utilization may offer more practical alternatives. Any such biomarker should be evaluated longitudinally and linked to changes in insulin sensitivity rather than treated as a surrogate endpoint without validation.

For the MAPK pathway, potential biomarkers include phospho-MEK and phospho-ERK states, pathway-responsive transcriptional signatures, and measures of the temporal response to insulin or other physiological stimuli. For trafficking, assays could measure insulin-receptor surface recovery, GLUT4 endocytosis and recycling, endosomal cargo movement, and beta-cell granule exocytosis. These read-outs would allow early-phase trials to distinguish pharmacodynamic engagement from clinical efficacy.

Candidate adjunctive interventions

The analysis also raises the possibility of evaluating CPT2- or PPAR-related interventions in genetically defined subgroups. Such a strategy should be approached cautiously. PPAR agonism affects lipid metabolism and insulin sensitivity but can have substantial tissue-specific and systemic effects. Similarly, the unexpected metabolic phenotype of muscle-specific CPT2 loss indicates that direct CPT2 manipulation may not produce a simple dose-dependent benefit [31]. The immediate priority should therefore be pathway characterization and biomarker development, not direct clinical targeting of CPT2.

The MAPK results argue against broad MEK inhibition as a general metabolic strategy. Instead, the pathway should be studied as a dynamic signaling system. Context-specific modulation, tissue-selective delivery, or interventions that restore normal signaling responses may be more plausible than continuous global inhibition or activation.

4.4 Limitations

Several limitations constrain the conclusions. First, the analysis is based on TWAS associations. A TWAS signal reflects the relationship between genetically predicted expression and a phenotype, but it does not by itself establish that the measured gene is causal. The association may reflect linkage disequilibrium, horizontal pleiotropy, correlated expression prediction, or another gene regulated by the same genetic region. Colocalization and fine-mapping would be necessary to determine whether the relevant GWAS and expression signals share a causal variant [34, 35].

Second, the cross-tissue meta-Z calculation in the supplied code uses a Stouffer-type combination of tissue-level Z-scores without an explicit covariance matrix for cross-tissue gene-level statistics. Tissue-level TWAS associations can be correlated because of shared genetic predictors, linkage disequilibrium, and shared expression prediction models. The resulting aggregate statistics should therefore be viewed as prioritization statistics rather than perfectly calibrated independent tests.

Third, leave-one-out influence is not equivalent to causal importance. A gene can be influential because it has a large Z-score, because the gene set is relatively small, because it is correlated with other genes, or because it is an outlier in a particular analysis group. The Stage 3 labels “core,” “transancestral,” and “direction-concordant” describe recurrence and statistical leverage within this pipeline; they do not establish that the genes are therapeutic targets.

Fourth, the gene sets were informed by an NMN-responsive mouse analysis. Cross-species gene mapping and pathway conservation are useful for hypothesis generation, but they do not guarantee that the same pathway is regulated in the same tissue, cell type, or physiological state in humans. The NMN re-analysis itself was based on a small experimental design, with four animals per group according to the supplied description, and did not produce individual genes that reached genome-wide significance after multiple-testing correction. The 35-gene set should therefore be considered a biological prior, not a definitive molecular signature.

Fifth, only two gene-set-by-analysis-group signals satisfied the joint threshold of absolute Stouffer Z greater than 1.96 and permutation p below 0.05. The large number of additional nominal, pathway-level, pairwise, proximity, and influence findings should not be presented as equivalent evidence. The 78 family-wise FDR survivors were concentrated largely in pairwise variance and profile-similarity tests, and they do not replace independent replication of the primary enrichment signals.

Finally, ancestry-associated differences may reflect several sources at once, including linkage disequilibrium, allele frequency, expression-model accuracy, environmental exposures, phenotype definitions, ascertainment, and sample size. The European and East Asian differences in the driver module are therefore important for trial design but should not be described as fixed biological differences between populations.

4.5 Future directions

The next step is independent multi-ancestry replication using external S-PrediXcan or TWAS resources, preferably with tissue-specific models and covariance-aware gene-set tests. Replication should distinguish the European fatty-acid-degradation signal from the broader CPT2-linked hypothesis and should test whether the pathway-level results persist after conditioning on correlated genes and known type 2 diabetes loci.

Colocalization and fine-mapping should be performed for CPT2, MAP2K1, MAPK3, DNM2, ACOX1, and the principal genes driving the two publication-grade signals. These analyses should be conducted separately across ancestry groups where sample size allows. Functional experiments should use allelic series, CRISPR perturbation, metabolic flux measurements, and insulin-response assays in skeletal muscle, liver, adipocyte, and beta-cell models.

The most direct translational test would be a prospective NMN or NR study that prespecifies pathway scores and intermediate biomarkers. Such a trial should measure insulin sensitivity, acyl-carnitines, lipid intermediates, MAPK activity, and trafficking-related read-outs, while ensuring adequate representation of multiple ancestry groups. The objective would be to determine whether genetic pathway scores modify treatment response, not merely whether they correlate with baseline diabetes risk.

5. Conclusion

This analysis establishes a hypothesis-generating bridge between NMN-responsive transcriptional programs and the multi-ancestry genetic architecture of type 2 diabetes. Ten NMN-informed KEGG gene sets showed evidence of enrichment in at least one analysis group, while six showed same-direction significance in at least two groups. The strongest individual results were the European fatty-acid-degradation signal and the African American autophagy signal, the only two gene-set-by-analysis-group combinations that satisfied the joint absolute-Z and permutation threshold.

The most important transancestral drivers were MAP2K1, MAPK3, VEGFB, RRAS2, CPT2, DNM2, ACOX1, and related genes within the MAPK/RAS, fatty-acid-oxidation, actin, autophagy, and trafficking networks. MAPK3 and CPT2 showed the clearest directional concordance, while MAP2K1 was highly recurrent but not completely directionally

uniform. CPT2 is especially important because it links the human TWAS results to the NMN-responsive mouse gene analysis, although it was not the leading leave-one-out driver of the specific European fatty-acid-degradation result.

The direction of the aggregate top-30 driver module differed across analysis groups, with strong protective-leaning profiles in European and Hispanic groups and a risk-leaning profile in the East Asian group. This finding argues against universal interpretation of pathway scores and supports ancestry-calibrated genetic prediction and trial design.

The translational value of the work lies in prioritization. CPT2-centred fatty-acid metabolism, MAP2K1/MAPK3 signaling, and DNMT2-related trafficking are reasonable candidates for functional studies, pharmacodynamic biomarker development, and genetically stratified NAD⁺ precursor trials. The current results do not justify treating these genes as proven causal therapeutic targets. They do, however, provide a structured and testable framework for connecting aging biology, mitochondrial substrate utilization, adaptive signaling, vesicular trafficking, and the heterogeneous genetic pathophysiology of type 2 diabetes.

Declarations

Ethics approval and consent to participate

Not applicable. This study is a secondary analysis of publicly available transcriptome-wide association study (TWAS) summary statistics and does not involve human participants, individual-level data, or new biological samples.

Consent for publication

Not applicable.

Availability of data and materials

The multi-ancestry type 2 diabetes S-PrediXcan results analyzed in this study were derived from the genetic architecture described by Suzuki et al. (2024). The NMN-responsive gene set was obtained from the re-analysis of the publicly available GSE85718 dataset as reported in Cheung (2026). All gene-set definitions correspond to standard KEGG pathway annotations (hsa04144, hsa00071, hsa03320, hsa04010, hsa04014, hsa04810, hsa04140, hsa04721, hsa04080, and hsa00020). Processed enrichment statistics, leave-one-out influence rankings, and Stage 3 driver lists generated in this work are available from the corresponding author upon reasonable request.

Competing interests

The author declares that there are no competing interests.

Funding

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

N.C. conceived and designed the study, performed all analyses, interpreted the results, and wrote the manuscript. The author read and approved the final manuscript.

Acknowledgements

The author thanks the original investigators of the multi-ancestry type 2 diabetes GWAS and the GSE85718 study for making their data publicly available.

References

1. Mahajan A, Spracklen CN, Zhang W, et al. Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. *Nat Genet.* 2022;54(5):560-572. doi:10.1038/s41588-022-01058-3
2. Smith K, Deutsch AJ, McGrail C, et al. Multi-ancestry polygenic mechanisms of type 2 diabetes. *Nat Med.* 2024;30(4):1065-1074. doi:10.1038/s41591-024-02865-3
3. Suzuki K, Hatzikotoulas K, Southam L, et al. Genetic drivers of heterogeneity in type 2 diabetes pathophysiology. *Nature.* 2024;627(8003):347-357. doi:10.1038/s41586-024-07019-6
4. Vujkovic M, Keaton JM, Lynch JA, et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat Genet.* 2020;52(7):680-691. doi:10.1038/s41588-020-0637-y
5. Møller JB, Pedersen M, Tanaka H, et al. Body composition is the main determinant for the difference in type 2 diabetes pathophysiology between Japanese and Caucasians. *Diabetes Care.* 2014;37(3):796-804. doi:10.2337/dc13-0598
6. Yoshino M, Yoshino J, Kayser BD, et al. Nicotinamide mononucleotide increases muscle insulin sensitivity in prediabetic women. *Science.* 2021;372(6547):1224-1229. doi:10.1126/science.abe9985
7. Døllerup OL, Christensen B, Svart M, et al. A randomized placebo-controlled clinical trial of nicotinamide riboside in obese men: safety, insulin-sensitivity, and lipid-mobilizing effects. *Am J Clin Nutr.* 2018;108(2):343-353. doi:10.1093/ajcn/nqy132
8. Martin AR, Kanai M, Kamatani Y, et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet.* 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x
9. Mostafavi H, Harpak A, Agarwal I, et al. Variable prediction accuracy of polygenic scores within an ancestry group. *Elife.* 2020;9:e48376. doi:10.7554/eLife.48376
10. Lewis ACF, Molina SJ, Appelbaum PS, et al. Getting genetic ancestry right for science and society. *Science.* 2022;376(6590):250-252. doi:10.1126/science.abm7530
11. Wojcik GL, Graff M, Nishimura KK, et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature.* 2019;570(7762):514-518. doi:10.1038/s41586-019-1310-4
12. Cheung N. NMN prevents age-associated transcriptional drift in a tissue-dependent manner: mechanistic leads from RAB11A-mediated trafficking and CPT2-linked fatty acid oxidation [preprint]. Figshare. Published 2026. doi:10.6084/m9.figshare.32800638.v1
13. Mills KF, Yoshida S, Stein LR, et al. Long-term administration of nicotinamide mononucleotide mitigates age-associated physiological decline in mice. *Cell Metab.* 2016;24(6):795-806. doi:10.1016/j.cmet.2016.09.013
14. Gerhart-Hines Z, Rodgers JT, Bare O, et al. Metabolic control of muscle mitochondrial function and fatty acid oxidation through SIRT1/PGC-1 α . *EMBO J.* 2007;26(7):1913-1923. doi:10.1038/sj.emboj.7601633
15. Hirschey MD, Shimazu T, Goetzman E, et al. SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation. *Nature.* 2010;464(7285):121-125. doi:10.1038/nature08778
16. Yoshino J, Baur JA, Imai SI. NAD⁺ intermediates: the biology and therapeutic potential of NMN and NR. *Cell Metab.* 2018;27(3):513-528. doi:10.1016/j.cmet.2017.11.002

17. Rodgers JT, Lerin C, Haas W, et al. Nutrient control of glucose homeostasis through a complex of PGC-1 α and SIRT1. *Nature*. 2005;434(7029):113-118. doi:10.1038/nature03354
18. Houten SM, Violante S, Ventura FV, Wanders RJA. The biochemistry and physiology of mitochondrial fatty acid β -oxidation and its genetic disorders. *Annu Rev Physiol*. 2016;78:23-44. doi:10.1146/annurev-physiol-021115-105045
19. McGarry JD, Brown NF. The mitochondrial carnitine palmitoyltransferase system: from concept to molecular analysis. *Eur J Biochem*. 1997;244(1):1-14. doi:10.1111/j.1432-1033.1997.00001.x
20. Koves TR, Ussher JR, Noland RC, et al. Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance. *Cell Metab*. 2008;7(1):45-56. doi:10.1016/j.cmet.2007.10.013
21. Samuel VT, Shulman GI. Mechanisms for insulin resistance: common threads and missing links. *Cell*. 2012;148(5):852-871. doi:10.1016/j.cell.2012.02.017
22. Muoio DM, Neufer PD. Lipid-induced mitochondrial stress and insulin action in muscle. *Cell Metab*. 2012;15(5):595-605. doi:10.1016/j.cmet.2012.04.010
23. Petersen MC, Shulman GI. Mechanisms of insulin action and insulin resistance. *Physiol Rev*. 2018;98(4):2133-2223. doi:10.1152/physrev.00063.2017
24. Yu C, Kim YB, Line DM, et al. Mechanism by which fatty acids inhibit insulin activation of insulin receptor substrate-1-associated phosphatidylinositol 3-kinase in muscle. *J Biol Chem*. 2002;277(52):50230-50236. doi:10.1074/jbc.M200958200
25. Fan F, Ji C, Wu Y, et al. Dynamin 2 regulates biphasic insulin secretion and plasma glucose homeostasis. *J Clin Invest*. 2015;125(11):4026-4041. doi:10.1172/JCI80652
26. Laiman J, Hsu YJ, Loh J, et al. GSK3 α phosphorylates dynamin-2 to promote GLUT4 endocytosis in muscle cells. *J Cell Biol*. 2023;222(2):e202102119. doi:10.1083/jcb.202102119
27. Zeigerer A, McBrayer MK, McGraw TE. Insulin stimulation of GLUT4 exocytosis, but not its inhibition of endocytosis, is dependent on RabGAP AS160. *Mol Biol Cell*. 2004;15(10):4406-4415. doi:10.1091/mbc.e04-04-0333
28. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000;28(1):27-30. doi:10.1093/nar/28.1.27
29. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48(3):245-252. doi:10.1038/ng.3506
30. Wainberg M, Sinnott-Armstrong N, Mancuso N, et al. Opportunities and challenges for transcriptome-wide association studies. *Nat Genet*. 2019;51(4):592-599. doi:10.1038/s41588-019-0385-z
31. Pereyra AS, Rajan A, Ferreira CR, Ellis JM. Loss of muscle carnitine palmitoyltransferase 2 prevents diet-induced obesity and insulin resistance despite long-chain acylcarnitine accumulation. *Cell Rep*. 2020;33(6):108374. doi:10.1016/j.celrep.2020.108374
32. Shaul YD, Seger R. The MEK/ERK cascade: from signaling specificity to diverse functions. *BBA Mol Cell Res*. 2007;1773(8):1213-1226. doi:10.1016/j.bbamcr.2006.10.005
33. Matsumoto Ikushima Y, Awazawa M, Kobayashi N, et al. MEK/ERK signaling in β -cells bifunctionally regulates β -cell mass and glucose-stimulated insulin secretion response to maintain glucose homeostasis. *Diabetes*. 2021;70(7):1519-1535. doi:10.2337/db20-1295
34. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383. doi:10.1371/journal.pgen.1004383

35. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet.* 2016;48(5):481-487. doi:10.1038/ng.3538

Figures

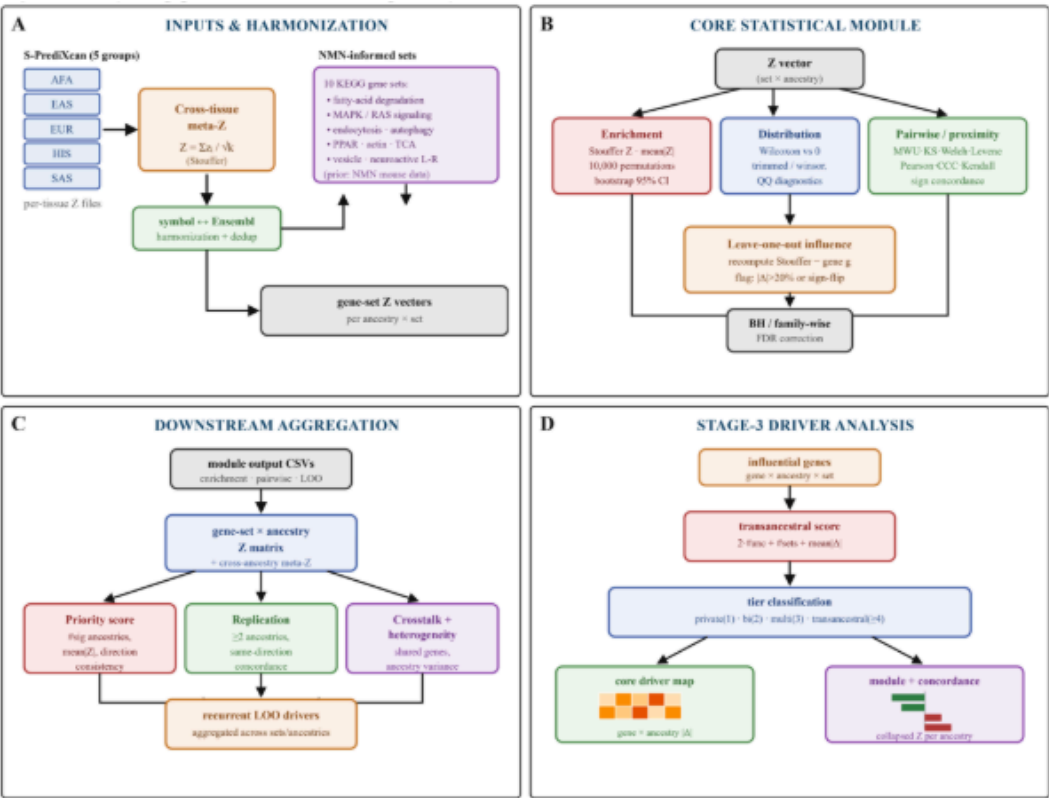


Figure 1

Analytical pipeline for transancestral, pathway-stratified TWAS enrichment. Schematic of the four analytical stages; panels depict the workflow only and no results are shown. (A) Inputs and harmonization. S-PrediXcan association statistics from five ancestry groups (AFA, EAS, EUR, HIS, SAS) are combined across tissues into gene-level meta-Z scores using Stouffer's method, and gene identifiers are harmonized between symbol and Ensembl namespaces. Ten NMN-informed KEGG gene sets are intersected with the harmonized genes to yield per-ancestry gene-set Z vectors. **(B) Core statistical module.** Applied independently to each gene-set × ancestry combination: permutation- and bootstrap-based enrichment (Stouffer Z, mean |Z|), distribution testing against zero with robust (winsorized/trimmed) estimators and QQ diagnostics, pairwise and proximity tests (Mann–Whitney U, Kolmogorov–Smirnov, Welch, Levene, Pearson, Lin's CCC, Kendall τ, sign concordance), and leave-one-out (LOO) influence analysis flagging genes that shift the Stouffer Z by >20% or invert its sign. Results are corrected by Benjamini–Hochberg and family-wise FDR. **(C) Downstream aggregation.** Per-run tables are pooled into a gene-set × ancestry Z matrix (with a cross-ancestry meta-Z), from which composite priority scoring, cross-ancestry replication, pathway crosstalk and heterogeneity analyses, and recurrent-driver aggregation are derived. **(D) Stage-3 transancestral driver analysis.** Influential genes are scored by ancestry and pathway breadth, assigned to tiers (ancestry-private, bi-ancestral, multi-ancestry, transancestral), and summarized as a core driver × ancestry influence map and a collapsed driver module with direction-concordance metrics.

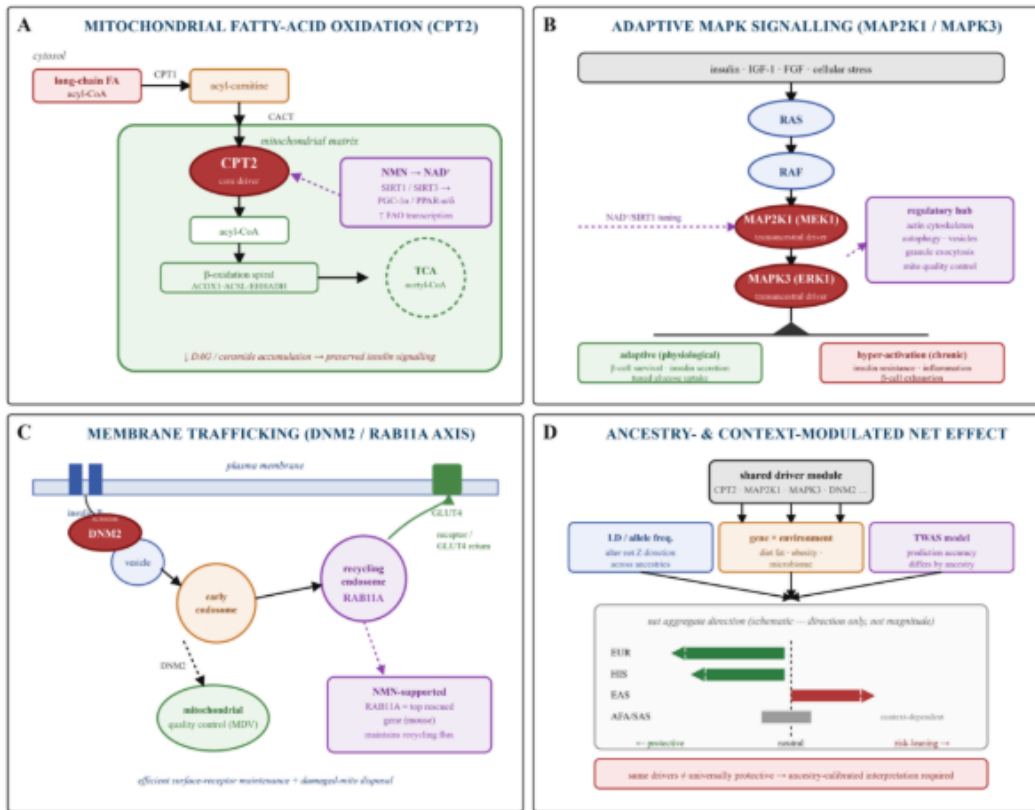


Figure 2

Proposed mechanistic bridges between NMN-responsive programmes and T2D-associated genetic pathways. (A) Mitochondrial fatty-acid oxidation centred on CPT2. Long-chain acyl-CoA is converted to acyl-carnitine by CPT1, transported across the inner membrane (CACT), and reconverted to acyl-CoA by CPT2, the core driver, feeding the β -oxidation spiral (ACOX1, ACSL, EHHADH) and delivering acetyl-CoA to the TCA cycle. The NMN $\rightarrow$ NAD⁺ $\rightarrow$ SIRT1/SIRT3 $\rightarrow$ PGC-1 α /PPAR- α/δ axis transcriptionally up-regulates these FAO enzymes, and greater oxidative clearance of fatty acids is expected to limit accumulation of lipotoxic diacylglycerols and ceramides that impair insulin signalling. **(B) Adaptive RAF–MEK–ERK signalling.** Metabolic and stress inputs converge on RAS $\rightarrow$ RAF $\rightarrow$ MAP2K1 (MEK1) $\rightarrow$ MAPK3 (ERK1), the two strongest transancestral drivers. A balance beam depicts the dose-dependent physiology: appropriately tuned activity supports β -cell survival, insulin secretion, and glucose uptake, whereas chronic hyper-activation promotes insulin resistance and β -cell exhaustion. Both nodes also act as a regulatory hub linking cytoskeletal dynamics, autophagy, vesicle trafficking, and mitochondrial quality control, with possible NAD⁺/SIRT1-dependent modulation. **(C) Membrane trafficking via the DNM2/RAB11A axis.** DNM2 mediates scission of clathrin-coated vesicles carrying internalized insulin receptor and GLUT4; cargo passes through early endosomes to RAB11A-positive recycling endosomes that return receptors and transporters to the surface, and DNM2 additionally supports mitochondrial-derived-vesicle formation for quality control. RAB11A was the top NMN-rescued gene in mouse data, positioning this recycling network as a shared node maintaining surface-receptor availability and disposal of damaged mitochondria. **(D) Ancestry- and context-modulated net effect.** The same shared driver module is filtered through differences in linkage disequilibrium and allele frequency, gene-by-environment interactions (dietary fat, obesity, microbiome), and ancestry-specific TWAS model accuracy, which can change the net directional interpretation. This emphasizes that recurrent drivers are not assumed to be universally protective and that any polygenic or pathway score built from them must be ancestry-calibrated.

Supplementary Files

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

- [Combined088CvDp2Appendix.docx](#)
- [GraphicalAbstract.docx](#)