

Methods Supplement: Fasting and Human Muscle Expression Analysis

This supplement describes the secondary analysis underlying **Fasting and Human Muscle: The Age-Related Gene Signal We Found**. The analysis compares changes in human skeletal-muscle gene expression with a reference associated with chronological age. The alignment score is not calibrated in years and is not a measure of biological age.

1. Study design and datasets

Nine public human skeletal-muscle datasets were assigned distinct analytical roles. The datasets were not pooled as interchangeable fasting replications.

Dataset	Role	Participants / samples	Main comparison
GSE175495	Age-reference cohort	24 men / 24 samples	12 older vs 12 younger
GSE164471	Age-reference cohort	53 people / 53 samples	Continuous age per decade, adjusted for sex
GSE55924	Fasting discovery	12 men / 48 samples	Nominal 4, 10 and 24 h vs nominal 1.5 h
GSE28016	Fasting/refeeding comparison	7 people / 14 paired samples	40 h FASTED minus 6 h REFED
GSE252357	Fed serial-biopsy control	5 people / 24 samples	Nominal 0.5, 3, 8 and 24 h vs PRE
GSE108539	Rhythmicity sensitivity	10 donors / 57 samples	Six scheduled clock times
GSE25941	Third age cohort	36 people / 36 baseline samples	21 older vs 15 younger, adjusted for sex
GSE8479	Age and training challenge	51 people / 65 arrays	Baseline age model; 14 paired POST minus PRE training samples
GSE129843	Eating-window crossover	11 men / 125 samples	TRF and URF/EXF clock-time trajectories and paired differences

The exact external analytical inputs, URLs, file sizes, and SHA-256 hashes are listed in `DATA/SOURCES.tsv`.

Missing samples were retained as missing and were not interpolated or imputed. No participant was excluded on the basis of the alignment score.

Metadata qualifications

For GSE164471, the analysis used GEO sample characteristics giving 33 men and 20 women. These characteristics conflicted with sex counts in the annotated count-file header. Age was parsed from the count-column names; sex was taken from the GEO characteristics. The same GEO characteristics were used for the sex-adjusted age model and bootstrap strata.

For GSE55924, the deposited nominal reference label is 1.5 h after the meal, whereas the publication describes the first biopsy at 105 min. The categorical analysis retains the deposited nominal label.

For GSE252357, nominal PRE/0.5/3/8/24 h labels correspond to approximately 0/1.5/4/9/25 h after PRE. The CTRL4 nominal 8 h sample is absent.

For GSE108539, donor 7 lacks samples at 00:00, 04:00, and 08:00. The publication supplement uses Roman donor labels that could not be linked unambiguously to the GEO numeric donor labels, so donor sex was not inferred from row order.

For GSE129843, URF in the deposited data corresponds to the extended-feeding condition labeled EXF in the publication. Individual crossover period and treatment-order variables were not available and were not inferred.

2. Identifier harmonization

The canonical analysis key is NCBI Gene:<EntrezID>. Mapping used exact, unique identifiers only. The package includes DATA/IDENTIFIERS/HGNC_frozen.tsv, a reduced derived identifier table containing the identifier fields required by the mapping scripts. The full upstream HGNC source file is not included.

GSE175495 was mapped from approved HGNC symbols, followed by unique previous-symbol or alias matches. GSE164471 used unique stable Ensembl identifiers. Duplicate canonical assignments within a cohort were excluded from age-reference harmonization.

For microarray datasets with multiple eligible probes or probe sets per gene, the primary representation selected the unambiguously mapped feature with the highest mean expression across all arrays in that dataset, with lexical feature-ID tie breaking. Selection did not use age or intervention effects. Mean aggregation of eligible log₂ probe values was retained as a sensitivity analysis.

Each downstream score uses the subset of the age signature measurable on that platform. Age-up and age-down means use separate measured denominators; unmeasured genes are not imputed or assigned zero change.

3. Age-reference construction

GSE175495 was filtered at CPM ≥ 1 in at least 12 samples. GSE164471 was filtered at CPM ≥ 1 in at least 10 samples. CPM thresholds of 0.5 and 2, with the same dataset-specific minimum sample counts, were examined as sensitivities.

Library-size normalization used TMM. Differential expression used edgeR robust dispersion estimation and quasi-likelihood modeling. GSE175495 estimated OLD minus YOUNG. GSE164471 modeled age per decade with adjustment for sex. [12,13]

For each cohort, the signed statistic used for ranking was

```
sign(logFC) * sqrt(QL F).
```

The shared uniquely mapped reference universe contained 10,603 genes. A full-universe age-ranking component was defined as

```
sign(statistic) * rank(abs(statistic), average ties) / N,
```

and the two cohort components were averaged.

The targeted age signature required a concordant nonzero raw-effect direction in both cohorts and BH FDR < 0.05 in at least one cohort. This produced 1,074 genes: 504 age-up and 570 age-down. Only the two age-reference cohorts contributed to signature selection.

The reference is cross-sectional and describes expression differences associated with age across individuals. It does not estimate an individual's age or longitudinal aging rate.

4. Dataset-specific preprocessing and models

GSE55924 fasting discovery

Positive deposited intensities were log2-transformed and jointly quantile-normalized across all 48 arrays. [16] The primary detection rule required the reported detection flag to equal zero in at least 12 arrays. The highest-mean unambiguous probe was selected after normalization. Detection minima of 6 and 24 arrays, mean-probe aggregation, and the authors' normalized representation were treated as separate sensitivities.

The limma model included participant fixed effects and categorical time, with `trend=TRUE` and `robust=TRUE`. [14] Primary contrasts were nominal 4, 10, and 24 h minus nominal 1.5 h. A separate temporal family compared later time points with one another.

GSE28016 fasting/refeeding

All 14 CEL arrays were processed jointly by oligo RMA at the core level using `pd.huex.1.0.st.v2`. [15] GPL5175 gene assignments were retained only when they resolved to one unique numeric Entrez identifier represented in the derived identifier table. No additional expression filter was imposed.

The paired limma model included participant and state. The contrast was FASTED minus REFED, comparing the state after 40 h without food with the state 6 h after refeeding. Primary highest-mean feature selection, mean-feature aggregation, and the deposited processed representation were kept distinct.

GSE252357 fed serial-biopsy control

Only the CONTROL/biopsy-only arm entered this analysis. Deposited DESeq2-normalized noninteger values were analyzed as $\log_2(\text{value} + 1)$, not as raw counts. The filter required a deposited value ≥ 10 in at least four samples. The model included participant and categorical nominal time with limma trend/robust moderation. A $\log_2(\text{value} + 0.5)$ transformation was examined separately.

GSE108539 rhythmicity

The deposited expression matrix is already log2 RPKM despite its count-like filename. The primary analysis used the exonic mRNA representation without another logarithmic transformation and retained features with mean log2 RPKM > -2 .

The model included participant fixed effects and 24 h sine and cosine terms. Their joint moderated F test defined rhythmicity, with BH FDR < 0.05 applied over expression-filtered source Ensembl features before restriction to uniquely mapped genes. Two author-derived rhythmic classifications were retained as additional predefined exclusions.

GSE25941 third age cohort

The 36 CEL files were processed jointly using affy RMA with `hgu133plus2cdf`. [15] GPL570 ENTREZ_GENE_ID entries were retained only when they specified one unique recognized identifier. The limma age model included sex and age group, with OLD minus YOUNG as the contrast. No additional expression filter was imposed.

GSE8479 age and training challenge

The 65 deposited Illumina HumanRef-8 signal vectors were log2-transformed and jointly quantile-normalized. Features nonpositive or nonfinite in any sample were excluded without adding a pseudocount. Annotation used GPL2700 and the derived Entrez reference.

The cross-sectional age model used the 51 baseline samples with sex, chip, and age group. The training comparison paired 14 older participants before and after training by documented subject code and modeled participant plus time, with POST minus PRE as the contrast.

All post-training observations were measured on chips absent from baseline, so training time and chip batch are completely confounded. The training-associated contrast is therefore not interpretable as a training-specific causal effect.

GSE129843 eating-window crossover

The author CPM matrix was analyzed as $\log_2(\text{CPM} + 0.5)$. Canonical mapping used exact unique Ensembl stable identifiers. The primary gene model included participant plus condition-by-categorical-clock-time terms. `duplicateCorrelation` estimated a common correlation within participant-condition blocks; `lmFit` used the corresponding block and correlation. Empirical Bayes moderation used trend and robust settings.

All 125 observed samples were modeled without imputation. The six clock times were 07:00, 11:00, 15:00, 19:00, 23:00, and 03:00 the following day.

5. Alignment score and inference

For participant i and a paired contrast, the alignment score was

$$A_i = 1/2 * [\text{mean}(\text{delta } \log_2 \text{ expression in measured age-up genes}) - \text{mean}(\text{delta } \log_2 \text{ expression in measured age-down genes})].$$

Positive A indicates net alignment with the separately constructed age-reference directions. Negative A indicates the opposite direction. A is not calibrated in years.

Group summaries use participants as the unit of inference. Two-sided exact sign-flip tests enumerate all 2^n participant sign configurations and count absolute null means at least as large as the observed absolute mean, using a floating-point tolerance of $1e-12$. Displayed 95% confidence intervals are pointwise t intervals and are not inversions of the sign-flip tests. [18]

BH correction families were defined separately for the three discovery baseline contrasts, the three discovery temporal contrasts, the four fed-control contrasts, and the five later times within each TRF, URF, and paired-difference trajectory. Validation and training each had one primary contrast. Sensitivity analyses retained their own specified families. [17]

GSE129843 joint score test

For each later clock time, the paired difference was the participant's anchored TRF A minus anchored URF A . One sign was applied to each eligible participant across all available later times, preserving the missing-data pattern. All 512 sign configurations of the nine eligible participants were enumerated. The joint statistic summed squared time-specific mean differences scaled by the corresponding mean squared individual difference divided by the time-specific n . This is distinct from the gene-level condition-by-time interaction and is not an equivalence test.

6. Rank and pathway analyses

Full-reference Spearman correlations used all measurable genes from the shared age-reference universe. Participant-sign-flip nulls applied the same sign to all gene changes from one participant before recalculating ranks of mean gene changes. Fitted moderated-statistic rank correlations were descriptive. Centered ranks use $(\text{average rank} - 0.5) / N - 0.5$; the difference of

their age-up and age-down means is a directional enrichment statistic. Wilcoxon/Mann-Whitney and rank-biserial comparisons were descriptive gene-level summaries, not independent-participant replication.

All 50 Hallmark gene sets from MSigDB v2025.1.Hs were evaluated. CAMERA used a minimum of 10 eligible mapped genes per set and inter-gene correlation 0.01, with BH correction within each contrast. [10,11] The five crossover interaction CAMERA P values were additionally combined by $\min(1, 5 * \min(P))$, then BH-adjusted across sets; this is a conservative union, not a competitive omnibus F test. Age RNA-seq pathway analyses used voom. The fed-control analysis used `trend.var=TRUE`. The TRF pathway analysis transformed both expression and design to respect the modeled covariance.

The complete Hallmark result comprises 50 sets across 34 reported contrasts from eight datasets; the circadian joint analysis is separate. Enrichment direction and FDR are expression-level results and are not direct measurements of pathway activity, metabolic flux, or a pathway's contribution to A.

7. Sensitivity analyses

Prespecified sensitivity analyses included:

- discovery leave-one-participant-out estimates;
- median-based A;
- common-feature analyses using the exact robust-signature intersection measurable in GSE55924, GSE28016, and GSE252357;
- contributor trimming at 1%, 5%, 10%, and 20% of the largest absolute signed gene contributions;
- baseline-expression matching with 10 quantile bins and 1,000 accepted gene-subset draws for each of discovery, validation, and fed control;
- exclusion of three predefined rhythmic-gene sets;
- exclusion of the full union of prespecified fiber, immune, stromal, and vascular marker panels;
- alternative probe collapse and source-processing representations;
- age-reference filter sensitivities at CPM thresholds 0.5 and 2;
- demographic and metadata sensitivities where numerically estimable.

Contributor trimming ranked genes by $\text{abs}(\text{direction} * \text{mean_change} / (2 * n_{\text{direction}}))$, with canonical-ID lexical tie breaking. The removed count is the ceiling of the specified fraction times the measured signature size; A is then recalculated with the remaining direction-specific denominators. Baseline-expression matching ordered genes by mean baseline log2 expression and canonical ID, then assigned bin floor($10 * \text{zero-based rank} / \text{gene count}$). Baselines were the nominal 1.5 h discovery sample, REFED in GSE28016, and PRE in GSE252357. Equal age-up and age-down counts were sampled within each bin without replacement, using the smaller group size; bins containing only one direction contributed no genes. Draws required both directions and finite scores. Duplicate gene subsets were valid repeated Monte Carlo draws. NumPy default_rng generated the draws. The first 1,000 accepted draws were retained. Seeds were 20261559 for GSE55924, 20261280 for GSE28016, and 20262523 for GSE252357.

The marker union comprised MYH7, TNNT1, TNNI1, ATP2A2 (slow fibers); MYH1, MYH2, TNNT3, TNNI2, ATP2A1 (fast fibers); PTPRC, CD3D, CD3E, LST1, TYROBP (immune); COL1A1, COL1A2, COL3A1, DCN, PDGFRA (stromal); and PECAM1, VWF, KDR, EMCN (vascular). This exclusion is a diagnostic, not a cellular deconvolution.

The three rhythmic exclusions used the independently fitted source-feature BH FDR < 0.05 set, the publication exonic q-value < 0.05 set, and the publication Model classifications R.E or R.I-R.E. Signed discovery 24 h contributions were also compared between rhythmic and remaining measured signature genes by a Mann-Whitney test, with a higher-contribution directional alternative and a two-sided descriptive result.

The race-adjusted age-reference sensitivity produced materially invalid negative quasi-likelihood F statistics and was not interpreted as a successful robustness analysis. It did not define the primary age reference. F values below $-1e-7$ were classified as invalid and their signed statistics left missing. Tiny negatives in $[-1e-7, 0)$ were set to zero for the reported signed sensitivity statistic. An explicitly labeled legacy-QL sensitivity used the same counts, design and normalization. Finite-only downstream scores are restricted diagnostics, not successful complete demographic adjustment.

8. Resampling procedures

GSE25941 participant resampling

Participants were resampled within sex-by-age-group strata, with the frozen age-reference membership held fixed. The first 1,000 valid bootstrap draws used seed 20262594. A separate 1,000-valid-draw permutation procedure shuffled age group within sex using seed 20262595. Empty, rank-deficient, or nonfinite fits were rejected. NumPy default_rng generated each sequence. Bootstrap intervals use the 2.5th and 97.5th percentiles. The permutation P value is $(1 + \text{the number of } | \text{permuted } A | \text{ at least } | \text{observed } A |, \text{ allowing } 1e-12 \text{ numerical tolerance}) / 1001$.

Age-reference reconstruction

GSE175495 participants were resampled within age group and GSE164471 participants within the 33-male/20-female strata. Each attempt repeated filtering, TMM normalization, dispersion estimation, quasi-likelihood fitting, construction of the shared universe, and the concordant-sign/ either-cohort-significance rule. The observed fasting datasets were held fixed.

Attempt k used seed $2026092201 + k$ with R Mersenne-Twister, Inversion normal generation, and Rejection sampling. Attempts with materially negative QL F statistics below $-1e-7$, non-estimable designs, nonfinite test statistics, or no selected genes in one direction were rejected. The first 1,000 accepted attempts were retained. Each accepted draw supplies shared-gene directions and membership for subsequent alignment. Direction-stability fractions use 1,000 as the denominator, including draws in which a gene is absent from the shared universe. Full per-fit F/P/ FDR diagnostics were not retained for the reported bootstrap.

These distributions quantify uncertainty in the estimated age reference conditional on the observed fasting datasets. They do not represent new fasting participants or joint reference-plus-participant confidence intervals.

9. Software

Principal versions were R 4.5.3; edgeR 4.8.2; limma 3.66.0; statmod 1.5.2; affy 1.88.0; oligo 1.74.0; Biobase 2.70.0; jsonlite 2.0.0; preprocessCore 1.72.0; hgu133plus2cdf 2.18.0; and pd.huex.1.0.st.v2 3.14.1. Python analyses used Python 3.12.14 with NumPy 2.3.5, pandas 2.2.3, SciPy 1.17.0, Matplotlib 3.10.8, openpyxl 3.1.5, and lxml 6.1.1.

The distributed scripts are grouped by acquisition, preprocessing, statistical modeling, alignment scoring, sensitivity/resampling analyses, and figure generation. The scientific order follows the sections above.

10. Data and code availability

The analyzed data originate from GEO series GSE175495, GSE164471, GSE55924, GSE28016, GSE252357, GSE108539, GSE25941, GSE8479, and GSE129843. Exact analytical files are listed in DATA/SOURCES.tsv.

The study datasets and full upstream HGNC source file are not redistributed. The package includes the reduced derived identifier table used by the mapping code. The MSigDB Hallmark file is not redistributed and must be obtained through MSigDB under its applicable terms.

The package includes the analysis scripts, analysis parameters encoded in the scripts, principal software versions, selected scientific outputs, nine figures, and 18 figure source tables.

References

1. Trim et al. (2022 [online 2021]). Divergent immunometabolic changes in adipose tissue and skeletal muscle with ageing in healthy humans. doi:10.1113/JP280977.
2. Tumasian et al. (2021). Skeletal muscle transcriptome in healthy aging. doi:10.1038/s41467-021-22168-2.
3. Wijngaarden, Bakker et al. (2014). Regulation of skeletal muscle energy/nutrient-sensing pathways during metabolic adaptation to fasting in healthy humans. doi:10.1152/ajpendo.00215.2014.
4. Kunkel et al. (2011). mRNA expression signatures of human skeletal muscle atrophy identify a natural compound that increases muscle mass. doi:10.1016/j.cmet.2011.03.020.
5. Edman et al. (2024). The 24-hour molecular landscape after exercise in humans reveals MYC is sufficient for muscle growth. doi:10.1038/s44319-024-00299-z.
6. Perrin et al. (2018). Transcriptomic analyses reveal rhythmic and CLOCK-driven pathways in human skeletal muscle. doi:10.7554/eLife.34114.
7. Raue et al. (2012). Transcriptome signature of resistance exercise adaptations: mixed muscle and fiber type specific profiles in young and old adults. doi:10.1152/japplphysiol.00435.2011.
8. Melov, Tarnopolsky, Beckman, Felkey and Hubbard (2007). Resistance Exercise Reverses Aging in Human Skeletal Muscle. doi:10.1371/journal.pone.0000465.
9. Lundell, Parr, Devlin et al. (2020). Time-restricted feeding alters lipid and amino acid metabolite rhythmicity without perturbing clock gene expression. doi:10.1038/s41467-020-18412-w.
10. Liberzon et al. (2015). The Molecular Signatures Database hallmark gene set collection. doi:10.1016/j.cels.2015.12.004.
11. Wu and Smyth (2012). Camera: a competitive gene set test accounting for inter-gene correlation. doi:10.1093/nar/gks461.
12. Chen, Chen, Lun, Baldoni and Smyth (2025). edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. doi:10.1093/nar/gkaf018.
13. Robinson and Oshlack (2010). A scaling normalization method for differential expression analysis of RNA-seq data. doi:10.1186/gb-2010-11-3-r25.

14. Ritchie et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. doi:10.1093/nar/gkv007.
15. Irizarry et al. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. doi:10.1093/biostatistics/4.2.249.
16. Bolstad et al. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. doi:10.1093/bioinformatics/19.2.185.
17. Benjamini and Hochberg (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. doi:10.1111/j.2517-6161.1995.tb02031.x.
18. Hemerik and Goeman (2018 [online 2017]). Exact testing with random permutations. doi:10.1007/s11749-017-0571-1.
19. Malhan et al. (2023). Skeletal muscle gene expression dysregulation in long-term spaceflights and aging is clock-dependent. doi:10.1038/s41526-023-00273-4.