Does resistance exercise training significantly alter skeletal muscle gene expression in humans?

Group 4
Maria Lara Moran, James Liu, Matthew MacFarlane, Yu-Hsin Yeh
11/17/2025

1. General Research Question

Resistance exercise training is known to improve muscular strength and metabolic health, but individuals vary widely in their biological responses. A central question in exercise physiology is whether resistance training induces consistent and measurable changes in skeletal muscle gene expression, and whether such molecular responses differ by age.

2. Dataset Description :

We analyze the GSE47881 dataset, also known as the DERBY BBSRC RET Study: Impact of resistance exercise on human skeletal muscle gene expression, which is a human skeletal muscle microarray experiment measuring gene expression changes before and after a 20-week supervised resistance exercise training program.

The dataset originally contains 45 participants, but one individual did not complete the program. After excluding this incomplete case, the final analytical sample includes 44 participants. For each participant, the microarray platform measured expression levels for 54,675 genes at two time points (pre-training and post-training).

In the raw GSE47881 matrix:

- Rows represent genes,
- Columns represent individual samples.

To prepare the data for statistical modeling, we performed the following preprocessing steps:

a. Removed the participant who did not complete the training.
b. Computed the expression difference (Post – Pre) for every gene in every participant, representing the individual's molecular response to training.
c. Transposed the expression matrix so that each row corresponds to one participant and each column represents one gene.

After processing, we obtained a dataset with:

- 44 rows (participants),
- 54,677 columns (gene difference variables + subject_id + age).

Except for the two metadata columns (subject_id and age), every column represents the change in gene expression induced by resistance training.

3. Statistical Plan

Our goal is to identify which genes show statistically significant expression changes following resistance exercise and to assess whether age influences the magnitude of these changes. To achieve this, we will conduct the following analyses:

a. Test for Exercise Induced Gene Changes
   For each gene, we will fit a simple linear model:

   $GeneDiff_i = \beta_0 + \varepsilon_i$

   and test whether the mean Post–Pre change for that gene differs significantly from zero.

b. Test for Age Effects
   For each gene, we will fit:

   $GeneDiff_i = \beta_0 + \beta_1 (Age_i) + \varepsilon_i$

   and test whether age is significantly associated with the magnitude or direction of training induced gene expression change.

4. Possible Reference Academic Papers:

a. Phillips BE, Williams JP, Gustafsson T, Bouchard C, Rankinen T, Knudsen S, et al. (2013) Molecular Networks of Human Muscle Adaptation to Exercise and Age. PLoS Genet 9(3): e1003389. https://doi.org/10.1371/journal.pgen.1003389

b. Kumar, V., Selby, A., Rankin, D., Patel, R., Atherton, P., Hildebrandt, W., Williams, J., Smith, K., Seynnes, O. R., Hiscock, N., & Rennie, M. J. (2009). Age-related differences in the dose–response relationship of muscle protein synthesis to resistance exercise in young and old men. *Journal of Applied Physiology*, 107(5), 1614–1622. https://physoc.onlinelibrary.wiley.com/doi/10.1113/jphysiol.2008.164483