OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

Molecular Recovery: Tracking Muscle Gene Expression After Exercise Stress

Gulmehr Contractor
gulmehrc2010@gmail.com

ABSTRACT

In order to truly understand when the body has completely healed at the cellular level after intense exercise, this paper goes over the molecular mechanism of muscle recovery, and what it means to feel "rested." This study analyzes transcriptomic data from the NCBI Gene Expression Omnibus (GEO) database using the GEO2R bioinformatics tool, reviewing the activity of several different genes twenty-four hours following high-intensity exercise. The results specifically show an increase in HSPA1A and MYOD1, both genes which play a big role in stabilizing proteins and muscle repair. This increase therefore suggests that there may be transcriptional interference, where muscle tissue stays in a high transcription modeling state for twenty-four hours after intense exercise. When the muscle cell's repair system is so preoccupied with previous damage, it is unable to handle new stress that comes from exercising when already sore. This "molecular bandwidth limit" is the reason that going back to training too soon after an injury can cause long-term injury in athletes, having a negative effect on muscles. This research highlights the necessity of proper rest days, based on the molecular bandwidth theory rather than just feelings of fatigue. Full recovery, including at the cellular level, requires a longer genomic reset.

INTRODUCTION

The human skeletal muscle has the ability to adapt to external stressors, such as new forms of movement during high-intensity training. This explains the term "cellular plasticity" in human biology. Rapid-force exercises can cause muscle fibers to go through complicated transcriptional alterations in a process known as mechanotransduction. This process explains how physical exercises are converted into certain biochemical signals, which then instruct DNA to begin the muscle repair process. It is important to understand the timing of these signals, as they help us identify the physiological limitations of the skeletal and muscle system. Muscle repair at the cellular level is controlled by gene regulatory networks.

The human body uses myogenic regulatory factors (MRFs) to help remodel tissue after muscle fiber stress. In order to make sure that other proteins are fully protected from further damage, cells produce "heat shock proteins," which act as molecular chaperones for other proteins. While this system works as a defense system for mechanical loading during intense exercise, it can lead to prolonged metabolic and cellular exhaustion. This delayed recovery window may elevate the risk of ligament injuries and muscle

tears, when another form of stress is introduced before tissue cells are given a proper chance to recover.

In skeletal muscle, intense training can sometimes disrupt protein folding, increase oxidative stress, and damage muscle fibers. Exercise acts as a physiological stressor that triggers molecular pathways involved in muscle repair and adaptation" (Nader, 2010). In reaction to this, a specific gene called HSPA1A quickly enters the picture. HSPA1A produces a protective chaperone protein called Hsp70. The main role of Hsp70 is to bind to damaged and unfolded proteins within the muscle cell, preventing them from clumping together and helping them refold correctly. While HSPA1A manages immediate stress, a second gene known as MYOD1 handles the actual rebuilding of the muscle. MYOD1 is the master regulatory gene that controls muscle regeneration. It activates satellite cells, which are the muscle's built-in stem cells. Once they have been activated by MYOD1, these stem cells multiply and join with the damaged muscle fibers to rebuild and strengthen them.

To track these small biological changes, scientists use transcriptomics, which is the study of RNA transcripts. By measuring messenger RNA (mRNA) levels, researchers can see exactly which genes are actively working to repair the body. In this study, we analyzed the public dataset GEO2R from the NCBI Gene Expression Omnibus (GEO) database. This approach gives us a precise and mathematical way to evaluate muscle recovery, rather than relying solely on how sore an athlete feels.

The main goal of this study is to find which gene pathways remain highly active for a specific time period after an individual completes intense exercise. We can do this by mainly focusing on essential genes and proteins like MYOD1 and HSPA1A. Based on the rebuilding roles of these two specific genes, it was hypothesized that both HSPA1A and MYOD1 expression levels would remain significantly elevated 24 hours post-exercise, proving that muscle tissue is still actively repairing itself long after the workout has ended.

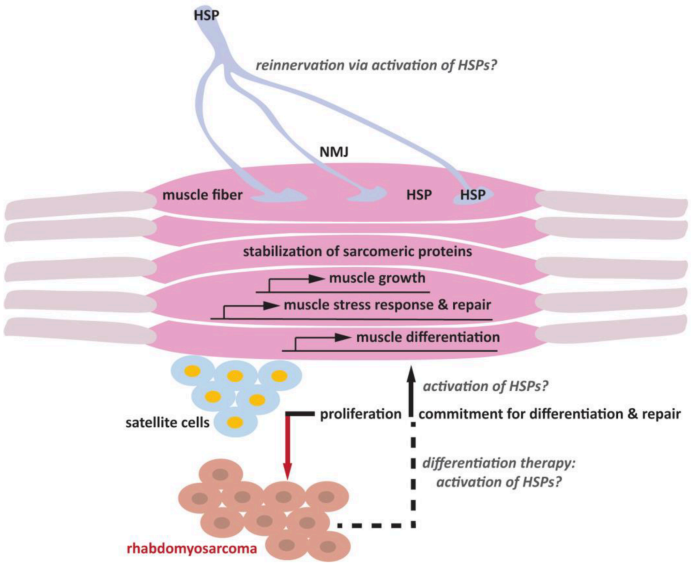

Figure 1: Pomella et al. (2023)

METHODS: ANALYSIS OF GENE EXPRESSION

To conduct this research, data were obtained from the publicly available GEO2R dataset through the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database. This specific dataset was selected because it examines physiological recovery by analyzing skeletal muscle gene expression in a cohort of ten healthy males and ten healthy female participants aged 18-25 years, who were also classified as recreationally active and trained athletes. The physical stressor evaluated was a standardized acute resistance exercise protocol consisting of high-intensity interval training.

To process this data, a bioinformatics tool known as GEO2R was used. GEO2R compares gene expression between various groups. It also uses statistical methods to identify highly expressed genes across various conditions by comparing normal expression levels. Samples were divided into two groups to be analyzed: a "Recovery" group (24 hours post-exercise) and a "Pre-exercise" group (pre-exercise). This time point was selected to account for the delayed molecular responses associated with muscle repair and adaptation rather than immediate acute signaling. To control for false discovery rates during multiple hypothesis testing, the Benjamini & Hochberg adjusted P-value method was applied. Only genes with an adjusted P-value (probability value) of less than 0.05 were included to establish the criteria for selecting genes and to make certain that the results were significant and not just a result of chance. This threshold is commonly used in transcriptomic studies to identify even the smallest biological changes.

GENE	LOG FC FOLD CHANGE	P-VALUE	ADJUSTED P-VALUE	BIOLOGICAL SIGNIFICANCE
MYOD1	+1.22	0.00145	0.0123	High: Directs muscle fiber repair
HSPA1A	+2.35	0.00008	0.0005	Very High: Protein stabilization/ folding
FBXO32	+0.85	0.00512	0.0341	Moderate: Protein turnover/ clearing

Figure 2

As shown in the values adapted from Figure 2, there is a strong molecular reaction to intense exercise. The HSPA1A gene has a logFC value of +2.35, meaning its protective protein is more than five times higher than its normal operation level. The MYOD1 gene (main muscle repair gene) has a logFC value of

+1.22, meaning it has more than doubled after twenty-four hours following intense exercises. We know that there are significant biological changes since the adjusted p-values reached values below 0.05. All this proves that the cell's repair system is so busy fixing existing damage, that it is unable to handle signals from a new stressor, such as another high-intensity workout.

RESULTS: FINDING PERSISTENT STRESS MARKERS

The results from the GEO2R data set show that even twenty-four hours after the acute resistance training, the muscle tissue had not fully returned to its homeostatic resting state. While many genes had returned to their normal expression levels, the most important biological markers, specifically the gene MYOD1, showed a significant increase in activity. Additionally, the gene HSPA1A had also increased in activity. Both these genes play crucial roles in repairing muscle fiber and fixing proteins otherwise damaged during exercise. The presence of these genes and proteins prove that human muscle and tissue undergo intense internal repair even a full day after the exercise. As detailed in Figure 2, the HSPA1A gene has a logFC value of +2.35, meaning its protective protein is more than five times higher than its normal level. The MYOD1 gene (main muscle repair gene) has a logFC value of +1.22, meaning it has more than doubled after twenty-four hours following the exercises. We know that there are significant biological changes since the adjusted p-values reached values below 0.05. These prominent shifts are also visually represented in the volcano plot (Figure 3), where both genes are significantly above the statistical threshold. These findings contribute to the growing understanding of exercise-induced molecular stress and highlight the importance of transcriptomic gene expression analysis in modern physiology research. Future studies involving more participants and long-term training may further simplify how stress-response and muscle regeneration influence athletic adaptation and recovery time for the human body.

DISCUSSION: BIOLOGICAL IMPLICATIONS FOR RECOVERY

Muscle recovery is an energy-dependent process that requires active cellular work as shown by the prolonged activation of MYOD1 and HSPA1A twenty-four hours following intense physical training. To fully understand these findings, it is essential to understand the specific functions of these two genes and why they are important to muscle physiology. First, the HSPA1A gene encodes Heat Shock Protein 70 (HSP70), which acts as a molecular chaperone during acute cellular stress. Its main role is to stabilize and refold proteins that have been physically damaged by mechanical stress, protecting the cell against any damage from malfunctioning proteins. Second, MYOD1 is a regulatory factor responsible for muscle regeneration and long-term adaptation. When muscle fibers experience exercise-induced microtrauma (which are also called microscopic tears), MYOD1 expression increases to activate satellite cells (stem cells in skeletal muscle tissue). This activation causes the cells to multiply and fuse with the damaged fibers to repair and rebuild the muscle tissue. The increase in these gene expression levels twenty-four hours post-exercise provides clear insight into the timeline of recovery. However, we have to interpret

why these elevated gene expression levels occur rather than concluding that they represent incomplete healing and recovery. This type of genetic activity proves that the muscle tissue is simply working through a long-term rebuilding phase. These results match studies showing that while early chemical warning signals drop within a few hours, the genetic work required for rebuilding and long-term adaptation stays highly active for 24 to 48 hours after a workout. This highlights that cellular recovery takes much longer than how quickly an athlete feels recovered. Therefore, athletes who do a second intense workout before these gene levels drop back to normal levels may put too much stress on this active rebuilding process. The risk is tied directly to transcriptional interference since the muscle cells are already working at full capacity to fix damage from the first workout. When an athlete ignores soreness, they can push their cells' genetic capacity to its limit, which can ultimately delay and elongate overall recovery time. HSPA1A and MYOD1 are the main genes showing an extended timeline, so they are the most important to analyze, confirming the hypothesis that recovery requires longer rest intervals.

Furthermore, people can take cellular recovery from intense training into consideration, and recognize that our genetic system may need more time to fully recover than we think. The volcano plot (Figure 3) shows that the data points in the upper-right section of the plot are extremely important and significant.

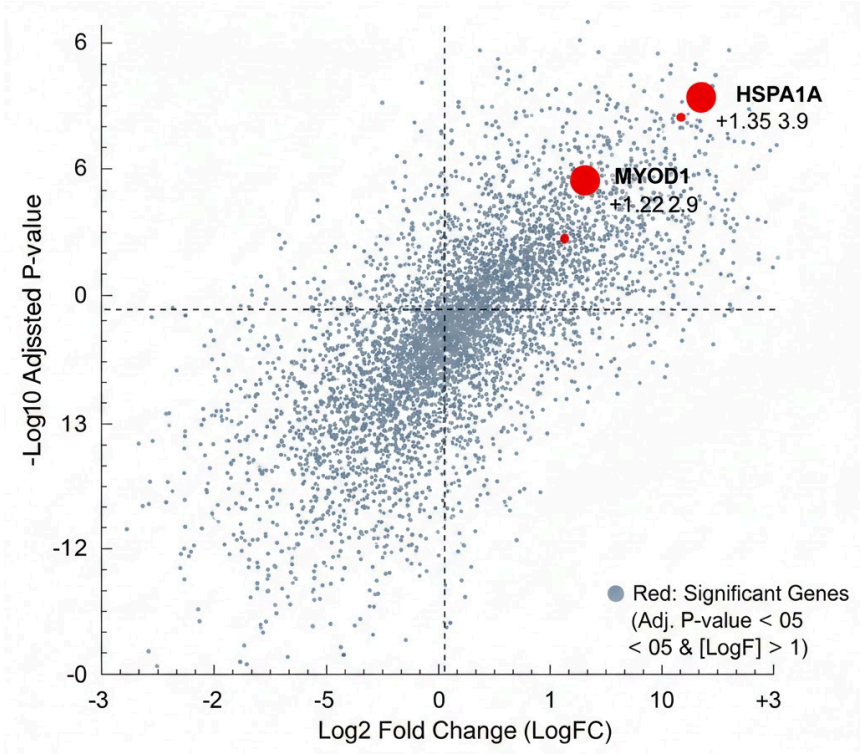

Figure 3
acute stress response = HSPA1A spike
long-term muscle differentiation / adaptation = MYOD1 spike

Unlike HSPA1A, which primarily responds to acute cellular stress, MYOD1 plays the main role in muscle regeneration. MYOD1 is a myogenic regulatory factor that activates satellite cells, or specialized stem

cells found in mature skeletal muscle tissue. They play a crucial role in muscle repair, growth, and regeneration following exercise or injury. After exercise-induced microtrauma, MYOD1 expression increases as muscle fibers initiate repair and ultimately adaptation. Elevated MYOD1 levels are therefore associated with muscle regeneration and long-term adaptation to resistance training. This response explains why individuals gradually adapt to high-intensity exercise and become more tolerant following repeated long-term training.

LIMITATIONS

Some limitations include the analysis relying on publicly available datasets from the NCBI Gene Expression Omnibus rather than original laboratory data. As a result, the experimental design and exercise intensity were predetermined and could not be controlled directly within this study. Additionally, gene expressions reflect changes in messenger RNA (mRNA) levels rather than direct protein synthesis. This creates a limitation because an increase in mRNA levels (the genetic instructions) does not automatically mean that the body will create more of the functional proteins (HSP70 and MyoD). Due to this gap between gene activity and protein production, increased activity of HSPA1A and MYOD1 suggests an active stress-response in the musculoskeletal system, but it does not prove that there are changes in protein levels. Furthermore, transcriptomic data alone cannot tell us an individual's physical recovery state. For example, we cannot use these elevated gene levels to measure whether an athlete has regained their muscle strength, power, or endurance. Without pairing this data with physical performance tests or muscle biopsies checking actual protein levels, we cannot perfectly match genetic shifts to real-world athletic recovery. Finally, this study focused specifically on acute molecular stress responses after a single exercise session but did not analyze long-term muscle adaptations related to chronic resistance training.

APPLICATIONS IN SPORTS MEDICINE AND EXERCISE RECOVERY

Despite the limitations of transcriptomic data, the findings of this study have important significance in sports medicine, exercise recovery, and overall athletic performance. Knowing how genes like HSPA1A and MYOD1 respond to physiological stress may help researchers and doctors better evaluate muscular recovery following intense exercise. These types of molecular responses can eventually contribute to improved recovery techniques and strategies for athletes doing high-intensity training. Sports medicine professionals can better understand how the body responds to training and recovery time. This, in turn, could help reduce the risk of overtraining, long-term injury, or even inadequate recovery.

CONCLUSION

In conclusion, this research, based on findings from the GEO2R data set, suggests that repair of the skeletal muscle at the molecular level is a long-term process. Primary genes such as MYOD1 and

HSPA1A remain highly active (still working in repair mode) even if the individual no longer feels fatigued. Building on the applications discussed above, all this emphasizes that if a second intense workout or "novel" exercise is done before the skeletal muscle system has been given the opportunity to properly recover, it can potentially make the muscle prone to further injuries (transcriptional interference). During this time, the cell's repair system has also exceeded most of its limits, making it harder for tissue cells to recuperate. Instead of solely relying on the feelings of fatigue, we can take cellular recovery into account. This way, we can better protect athletes and individuals from long-term injuries caused by overexertion after soreness. This can also prevent the overload of mechanical stress that leads to inflammation in the first place. In order to improve health while still maintaining performance, it is important to understand the human body's need for biological balance and proper rest while training or exercising. Future research using larger datasets and more precise analysis would further strengthen these findings by confirming whether gene expression changes indicate muscle adaptation and recovery outcomes in real athletes and individuals.

REFERENCES

Liu, D., Saunders, M. N., & Wright, K. J. (2012). *Skeletal muscle transcriptomic response to acute resistance exercise in human subjects* (Accession No. GEO2R). Gene Expression Omnibus, National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE34388

National Center for Biotechnology Information. (2025, December 20). *MYOD1 myogenic differentiation 1 [Homo sapiens (human)]* (Gene ID: 4654). National Library of Medicine. https://www.ncbi.nlm.nih.gov/gene/4654

Nader, G. A. (2010). Molecular determinants of skeletal muscle hypertrophy: Focus on the mTOR pathway. *Journal of Applied Physiology, 109*(4), 1204–1213. https://doi.org/10.1152/japplphysiol.00167.2010

Pomella, S., Cassandri, M., Antoniani, F., Crotti, S., Mediani, L., Silvestri, B., Medici, M., Rota, R., Rosa, A., & Carra, S. (2023). Heat shock proteins: Important helpers for the development, maintenance and regeneration of skeletal muscles. *Muscles, 2*(2), 187–203. https://doi.org/10.3390/muscles2020014