REFERENCES

1. Hems TE, Mahmood F. Injuries of the terminal branches of the infraclavicular brachial plexus: patterns of injury, management and outcome. J Bone Joint Surg Br 2012;94:799–804.
2. Atef A, El-Tantawy A, Gad H, Hefeda M. Prevalence of associated injuries after anterior shoulder dislocation: a prospective study. Int Orthop 2016;40:519–524.
3. Mitchell JJ, Chen C, Liechti DJ, Heare A, Chahla J, Bravman JT. Axillary nerve palsy and deltoid muscle atony. JBJS Rev 2017;5(7):e1.
4. Brown SA, Doolittle DA, Bohanon CJ, Jayaraj A, Naidu SG, Huettl EA, *et al.* Quadrilateral space syndrome: the Mayo Clinic experience with a new classification system and case series. Mayo Clin Proc 2015;90(3): 382–394.
5. Maak TG, Osei D, Delos D, Taylor S, Warren RF, Weiland AJ. Peripheral nerve injuries in sports-related surgery: presentation, evaluation, and management: AAOS exhibit selection. J Bone Joint Surg Am 2012; 94(16):e121(1–10).
6. Lee S, Saetia K, Saha S, Kline DG, Kim DH. Axillary nerve injury associated with sports. Neurosurg Focus 2011;31(5):E10.
7. Boon A. Ultrasonography and electrodiagnosis: are they complementary techniques? PM R 2013;5(5 Suppl):S100–S106.
8. Gruber H, Peer S, Gruber L, Loescher W, Bauer T, Loizides A. Ultrasound imaging of the axillary nerve and its role in the diagnosis of traumatic impairment. Ultraschall Med 2014;35(4):332–338.
9. Stecco C, Gagliano G, Lancerotto L, Tiengo C, Macchi V, Porzionato A, *et al.* Surgical anatomy of the axillary nerve and its implication in the transdeltoid approaches to the shoulder. J Shoulder Elbow Surg 2010; 19(8):1166–1174.
10. Leechavengvongs S, Teerawutthichaikit T, Witoonchart K, Uerpairojkit C, Malungpaishrope K, Suppauksorn S, *et al.* Surgical anatomy of the axillary nerve branches to the deltoid muscle. Clin Anat 2015; 28(1):118–122.
11. de Seze MP, Rezzouk J, de Seze M, Uzel M, Lavignolle B, Midy D, *et al.* Does the motor branch of the long head of the triceps brachii arise from the radial nerve? An anatomic and electromyographic study. Surg Radiol Anat 2004;26(6):459–461.
12. Cartwright MS, Passmore LV, Yoon JS, Brown ME, Caress JB, Walker FO. Cross-sectional area reference values for nerve ultrasonography. Muscle Nerve 2008;37(5):566–571.
13. Cartwright MS, Shin HW, Passmore LV, Walker FO. Ultrasonographic reference values for assessing the normal median nerve in adults. J Neuroimaging 2009;19(1):47–51.
14. Zaidman CM, Al-Lozi M, Pestronk A. Peripheral nerve size in normals and patients with polyneuropathy: an ultrasound study. Muscle Nerve 2009;40(6):960–966.
15. Cartwright MS, Mayans DR, Gillson NA, Griffin LP, Walker FO. Nerve cross-sectional area in extremes of age. Muscle Nerve 2013;47(6): 890–893.
16. Won SJ, Kim BJ, Park KS, Kim SH, Yoon JS. Measurement of cross-sectional area of cervical roots and brachial plexus trunks. Muscle Nerve 2012;46(5):711–716.
17. Qrimli M, Ebadi H, Breiner A, Siddiqui H, Alabdali M, Abraham A, *et al.* Reference values for ultrasonography of peripheral nerves. Muscle Nerve 2016;53(4):538–544.
18. Simone JP, Streubel PN, Sanchez-Sotelo J, Steinmann SP, Adams JE. Change in the distance from the axillary nerve to the glenohumeral joint with shoulder external rotation or abduction position. Hand (N Y) 2017;12(4):395–400.
19. James J, Sutton LG, Werner FW, Basu N, Allison MA, Palmer AK. Morphology of the cubital tunnel: an anatomical and biomechanical study with implications for treatment of ulnar nerve compression. J Hand Surg Am 2011;36(12):1988–1995.
20. Thoirs K, Williams MA, Phillips M. Ultrasonographic measurements of the ulnar nerve at the elbow: role of confounders. J Ultrasound Med 2008;27(5):737–743.

CLINICAL UTILITY OF RNA SEQUENCING TO RESOLVE UNUSUAL GNE MYOPATHY WITH A NOVEL PROMOTER DELETION

SAMYA CHAKRAVORTY, PHD,[1] KIERA BERGER, BS,[2] DALIA ARAFAT, MS,[2] BABI RAMESH REDDY NALLAMILLI, PHD,[1] HARI PRASANNA SUBRAMANIAN, MS,[2] SOUMYA JOSEPH, PHD,[3] MARY E. ANDERSON, MS,[3] KEVIN P. CAMPBELL, PHD,[3] JONATHAN GLASS, MD,[4] GREG GIBSON, PHD,[2] and MADHURI HEGDE, PHD[1,5]

[1] Department of Human Genetics, Emory University School of Medicine, Whitehead Building Suite 301, 615 Michael Street NE, Georgia, USA

[2] Center for Integrative Genomics, School of Biology, Georgia Institute of Technology, Atlanta, Georgia, USA

[3] Howard Hughes Medical Institute, Department of Molecular Physiology and Biophysics, Department of Neurology, University of Iowa Roy J. and Lucille A. Carver College of Medicine, Iowa City, Iowa, 52242, USA

[4] Department of Neurology and Pathology, Emory University School of Medicine, Atlanta, Georgia, USA

[5] Global Laboratory Services/Diagnostics, Perkin Elmer, Waltham, Massachusetts, USA

Accepted 12 April 2019

ABSTRACT: *Introduction:* UDP N-acetylglucosamine2-epimerase/N-acetylmannosamine-kinase (*GNE*) gene mutations can cause mostly autosomal-recessive myopathy with juvenile-onset known as hereditary inclusion-body myopathy (HIBM). *Methods:* We describe a family of a patient showing an unusual HIBM with both vacuolar myopathy and myositis without quadriceps-sparing, hindering diagnosis. We show how genetic testing with functional assays, clinical transcriptome sequencing (RNA-seq) in particular, helped facilitate both the diagnosis and a better understanding of the genotype-phenotype relationship. *Results:* We identified a novel 7.08 kb pathogenic deletion upstream of *GNE* using array comparative genomic hybridization (aCGH) and a common Val727Met variant. Using RNA-seq, we found only monoallelic (Val727Met-allele) expression, leading to ~50% *GNE* reduction in muscle. Importantly, α-dystroglycan is hypoglycosylated in the patient muscle, suggesting HIBM could be a "dystroglycanopathy." *Conclusions:* Our study shows the importance of considering aCGH for GNE-myopathies, and the potential of RNA-seq for faster, definitive molecular diagnosis of unusual myopathies.

Muscle Nerve **60**:98–103, 2019

Additional supporting information may be found in the online version of this article.

Key words: aCGH, GNE myopathy (HIBM), molecular diagnostics, myositis, next generation sequencing, transcriptome sequencing (RNA-seq)

This work was also supported by Paul D. Wellstone Muscular Dystrophy Cooperative Research Center Grant (1U54NS053672)

Funding: This work was supported by the Muscular Dystrophy Association grant MDA578400 to S.C. and grant MDA418496 to M.H.

Conflicts of Interest: None of the authors has any conflict of interest to disclose.

Correspondence to: S. Chakravorty; e-mail: samya.chakravorty@emory.edu. and M. Hegde; e-mail: madhuri.hegde@emory.edu

Published online 16 April 2019 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/mus.26486

To enhance molecular diagnostic yield in neuromuscular disorders (NMDs), functional assays downstream of genomic DNA[1–4] is recommended by the American College of Medical Genetics and Genomics (ACMG).[5] GNE-myopathy (OMIM #605820) or hereditary inclusion-body myopathy (HIBM) is a vacuolar myopathy generally sparing the quadriceps. It

is a rare, recessive, inherited degenerative skeletal muscle disorder caused by *GNE* gene (MIM# 603824; NCBI Gene ID: 10020; NC_000009.12) variants with - early-adult onset.[6–8] The GNE enzyme catalyzes the first 2 rate-limiting steps in the biosynthesis of 5-N-acetylneuraminic acid (Neu5Ac)[9,10] found as the terminal glycans on various glycoproteins/glycolipids, such as the sarcoglycans and dystroglycan (DG), functioning in variety of cellular pathways.[11]

Our understanding of the molecular basis of GNE myopathy is unclear.[12,13] Here, we describe the lessons learned by using functional genomic approaches to characterize an unusual GNE myopathy in a family with a novel deletion variant in which relative quadriceps sparing in association with both vacuolar myopathy and myositis made diagnosis challenging.

MATERIALS AND METHODS

All protocols were approved by institutional IRB with written consents and are presented here in chronological order. First, trio (patient and parents) exome sequencing and simultaneous patient muscle biopsy immunohistochemistry, and then RNA sequencing of the biopsy, were performed, which prompted us to do array comparative genomic hybridization (aCGH) to identify any deletions/duplications. Immunoblotting was then done using biopsy to identify any glycosylation defect related to the HIBM pathophysiology. Gene ontology-pathway analysis[14] led to further understanding of patient muscle glycosylation defects. Molecular-dynamics (MD) simulation of the V727M variant on the GNE kinase-domain crystal-structure was performed. For methods details, see Supplementary Methods, which are available online.

RESULTS

Clinical Phenotype. The patient was a 21-year-old man, ethnically Indian-Guyanese, with progressive muscle-weakness without any cardiac or respiratory comorbidity. Symptoms began at age 20 years with asymmetric leg pain and weakness, initially with bilateral-foot-drop and mild (Medical Research Council 4/5) weakness in the quadriceps, followed by rapidly progressive and severe bilateral lower extremity distal weakness, additional quadriceps weakness, and upper extremity weakness in the deltoids and the long-finger-flexors. There was no facial or bulbar weakness and tendon reflexes were reduced throughout. Needle electromyography of the quadriceps, tibialis anterior, iliopsoas, and medial gastrocnemius muscles demonstrated abnormal spontaneous activity (fibrillation potentials and positive sharp waves) and early motor unit potential recruitment, compatible with multiple types of myopathies. There was no family history of neuromuscular disease.

Quadriceps Showed Unusual HIBM: Both Vacuolar and Inflammatory Myopathy. Quadriceps biopsy (Fig. 1) showed increased fiber-size variability with both muscle-fiber necrosis and perivascular inflammation. Prominent rimmed vacuoles were seen on modified Gomori trichrome that were positive by immunohistochemistry

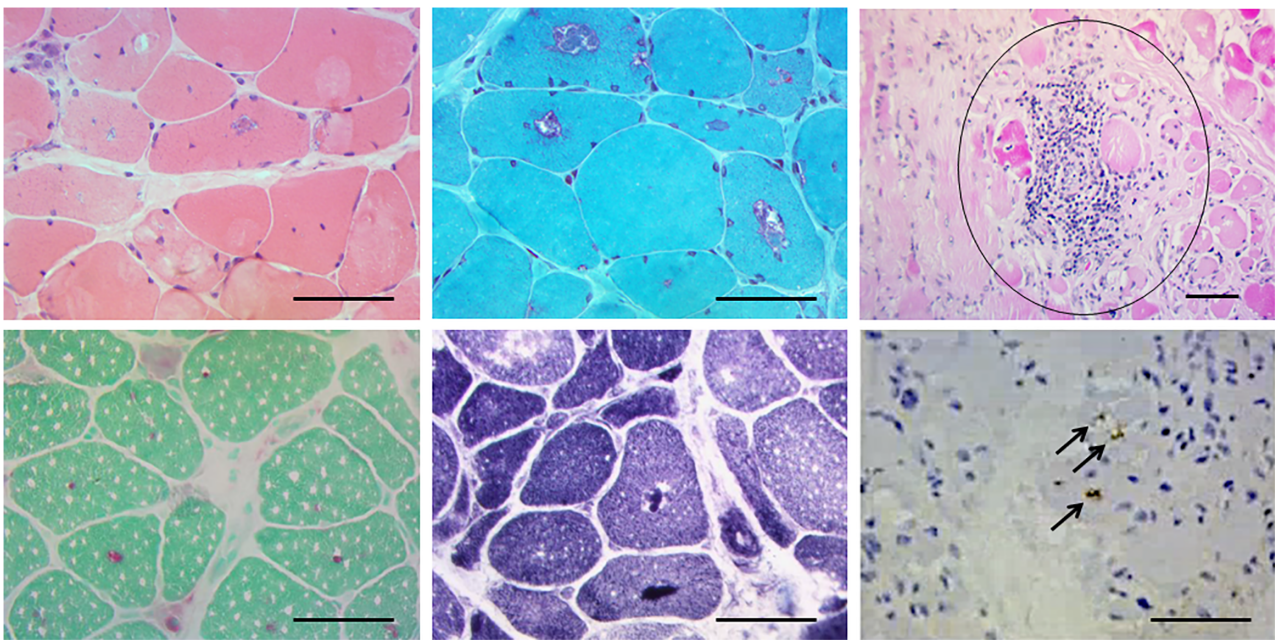


FIGURE 1. Muscle biopsy showing central and sub-sarcolemmal vacuoles in hemotoxylin and eosin (H&E) (top panel left), which demonstrates red "rimming" with modified Gomori trichrome (top center panel). An example of inflammation is seen in the top right panel (circled). Bottom panel: Acid phosphatase (left) and NADH (center) showing positive material within the vacuoles and vacuoles are stained positive for ubiquitin (black arrows) (right). Scale bar = 50 μm. Additionally, the connective tissue was mildly increased. Atrophic fibers were round and pyknotic nuclear clumps were not seen, and the biopsy showed a moderate number of fibers with internalized nuclei. Regenerating fibers were not seen. The following stains were normal: cytochrome oxidase (COX), myosin ATPase (normal distribution of fiber types), Oil red O, periodic acid–Schiff (PAS), phosphorylase, Congo red. Neither muscle fiber-type grouping, nor type specific atrophy was seen. [Color figure can be viewed at wileyonlinelibrary.com]

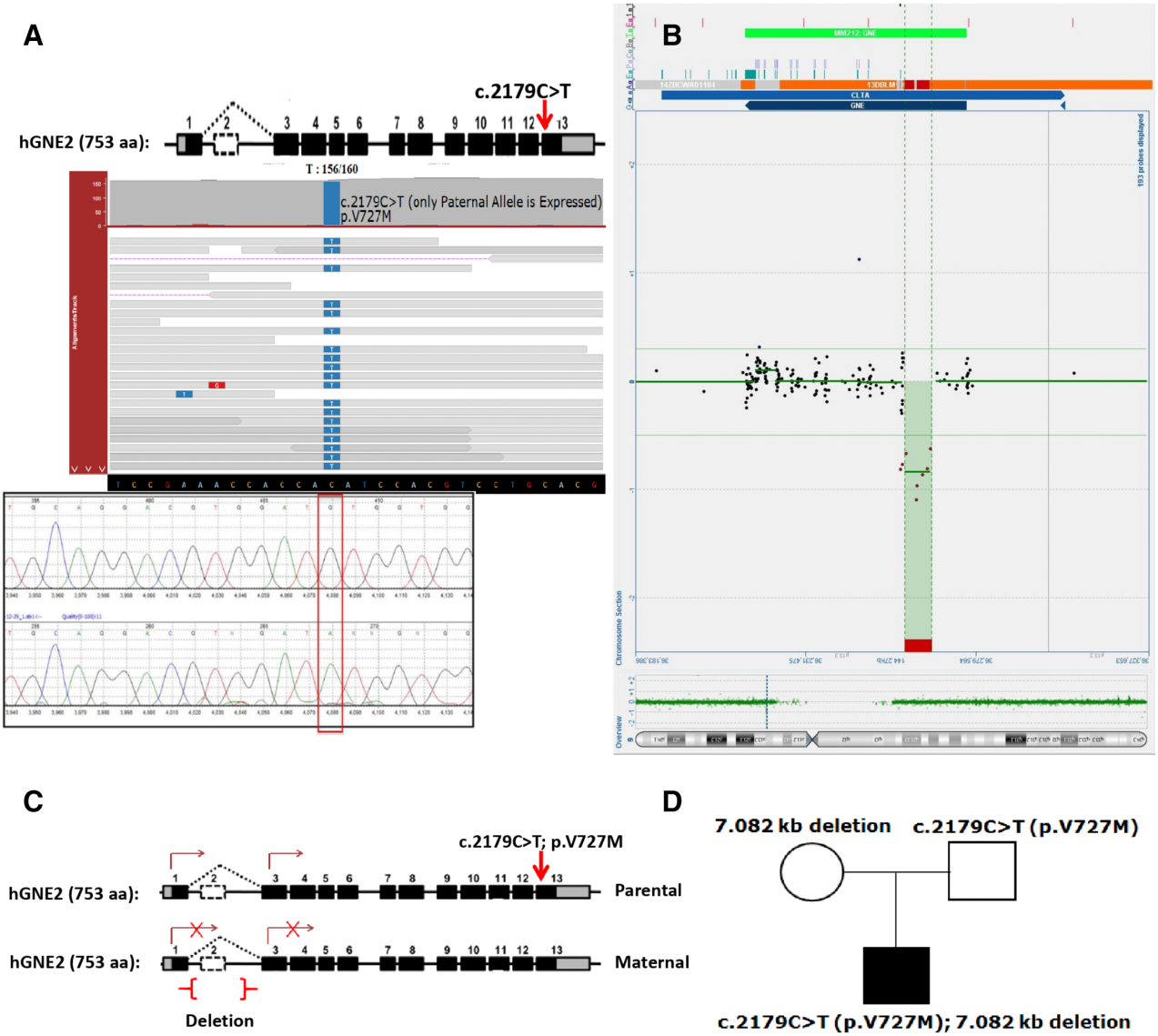


FIGURE 2. (A) Integrated Genomics Viewer (IGV) pile up of RNA-sequencing showing monoallelic expression of *GNE* gene with only the allele harboring c.2179C>T:G>A (p.V727M) missense "likely pathogenic" variant expressed. The red arrow indicates the position of the V727M variant in exon 13 of the *GNE* gene. Sanger sequencing confirmation was performed on cDNA showing monoallelic expression as shown below. **(B)** aCGH signal showing a deletion upstream of the *GNE* gene with genomic breakpoints at nucleotide positions g.36,259,402 and 36,266,483 was detected in this individual (SCV000599234). This deletion is 7.08 kb in size and encompasses the untranslated exon 2 of the *hGNE2* transcript but upstream of the *hGNE1* transcript of the *GNE* gene. **(C,D)** Exome sequencing and later aCGH of trios reveal that monoallelic expression was due to expression of only the paternal allele of *GNE* in the proband. [Color figure can be viewed at wileyonlinelibrary.com]

for both ubiquitin and TDP-43. There was no lipid or glycogen accumulation. Acid phosphatase stain showed increased lysosomal activity.

Exome and RNA Sequencing Revealed Monoallelic Expression of V727M Allele. A known "likely pathogenic" heterozygous missense variant (c.2179C>T (p.V727M)) (Supplementary Fig. S1; Supplementary Table S1) was identified in *GNE* (rs121908627; allele frequency of 0.0141) prevalent in South East Asian populations.[6,15–21] Exome analysis did not identify a second variant resulting in no molecular diagnosis. RNA sequencing using target muscle biopsy revealed the presence of only the *GNE* V727M allele (Fig. 2A) suggesting mono-allelic expression. Absence of transcription from the alternate chromosome could be due to a deletion/duplication not detected by ES. Thus, we performed aCGH using patient genomic DNA.

aCGH Identified Novel Deletion Encompassing Exon2 of *hGNE2*, Upstream of *hGNE1*. Recently, a ~11.3-kb deletion encompassing exon 2 was found in a patient along with a single V727M variant.[22] In our study, aCGH revealed a novel 7.08 kb deletion (g.36,259,402 to g.36,266,483) (SCV000599234) upstream of the *GNE* gene (different from deletions identified in Zhu et al.[23]) (Fig. 2B). This reports a novel compound-heterozygous variant combination (Supplementary Table S1) of a

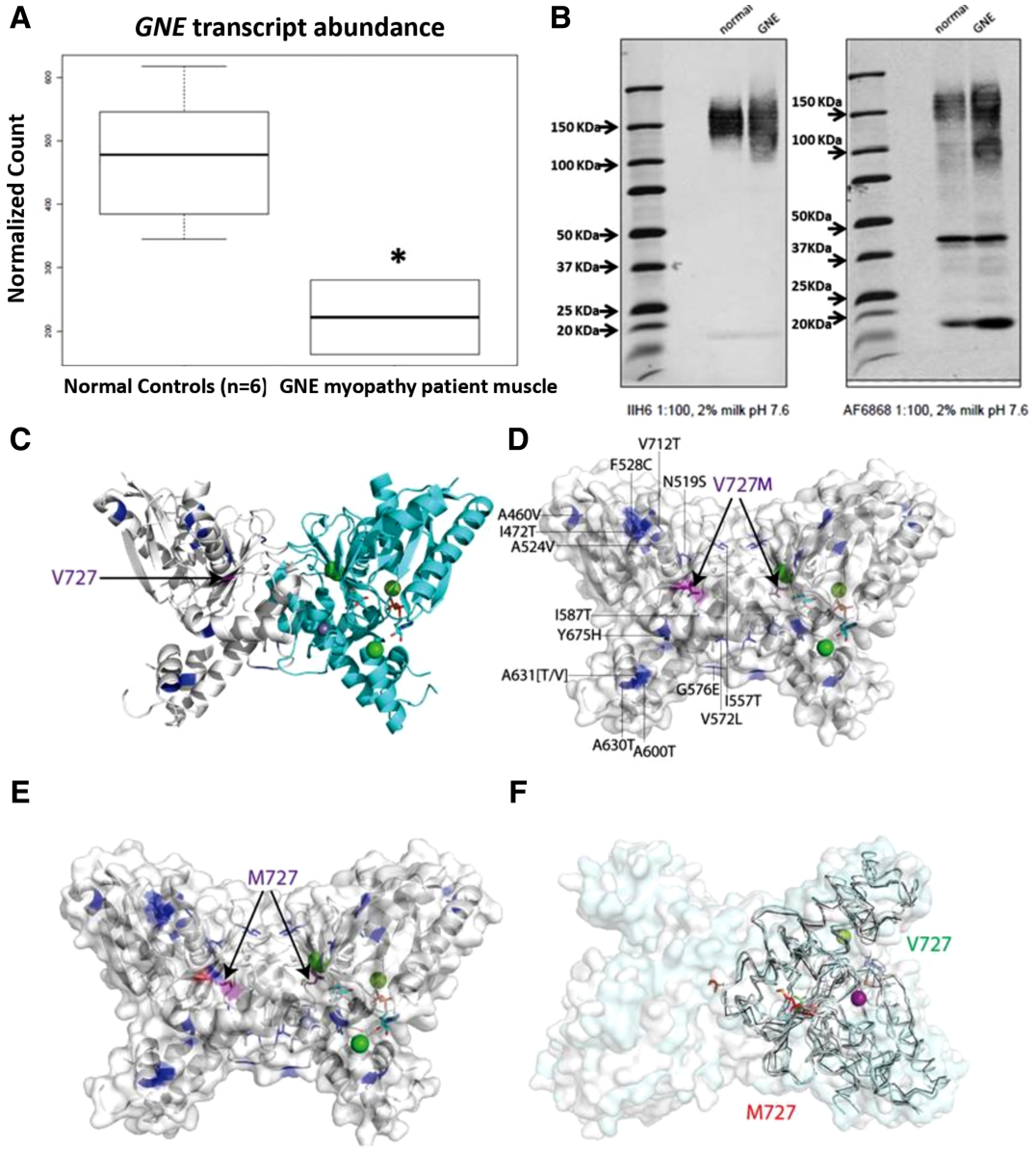


FIGURE 3. (A) Approximately 50% lower-expression ($P < 0.05$) of *GNE* in GNE myopathy patient muscle compared with that in 6 control normal muscle biopsies. **(B)** IIH6-antibody against glycosylated-α-DG shows hypoglycosylation (lighter signal, broader smear) of α-DG in patient muscle compared with control. AF6868 against core α-DG, β-DG shows different α-DG smearing patterns in patient muscle compared with control. Two predominant staining areas are 150 kDa and 100 kDa in patient sample. Roughly same β-DG-fragment expression at ~43 kDa is seen in both patient and control muscles. **(C)** HIBM-causing point mutations (blue, purple sticks) mapped onto GNE-N-acetylmannosamine-kinase-domain (PDB ID: 2YHY); N-acetylmannosamine-kinase-dimer backbone shown as cartoon. Protomers are colored white and cyan, associated divalent ions as spheres (chloride ions: green, zinc ion: lavender). Active site bound ADP and N-acetylmannosamine are shown as sticks on the right hand side GNE protomer. HIBM-causing residues are mapped on the left hand side GNE protomer. The V727M (V696M) mutation (purple sticks), other HIBM mutations (blue sticks). **(D)** Valine-727, and other HIBM mutations, are shown on both dimer subunits. **(E)** The Valine-727 sidechain is replaced by a Methionine residue on both dimer subunits and colored in Corey, Pauling, Koltun scheme (oxygen in red and sulfur in yellow). **(F)** The rotamer is changed upon replacing V727 (green sticks) by a methionine (red sticks) in MD simulations. [Color figure can be viewed at wileyonlinelibrary.com]

large deletion upstream of *GNE* in trans with the missense V727M, which causes the mono-allelic expression.

Gene Expression Analysis Showed 50% Reduced *GNE* Expression. Cluster analysis of 274 NMD-associated genes (Table S2) from next generation sequencing-based transcriptome sequencing (RNA-seq) data showed separate clustering of the patient's vastus lateralis (VL) muscle samples from the controls' VL muscles (Supplementary Fig. S2). *GNE* expression was reduced by ~50% (Fig. 3A). A total of 89 NMD-associated genes were differentially expressed and several key extracellular matrix (ECM) genes, namely collagen fiber and laminin genes (*COL6A1/A2/A3/12A1, LAMA2*) were up-regulated (Table S3).

Gene Ontology Pathway Analysis. Gene ontology-based gene set enrichment (GO-Pathway) analysis on the

differentially-regulated 89 genes (Supplementary Fig. S3; Supplementary Table S4) identified major enriched "pathways," "cellular compartments," and "molecular functions." The common biology is a predicted effect on of protein and lipid glycosylation affecting the cytoskeleton-intracellular matrix and ECM cross-talk through sarcolemmal proteins, important for the sarcomere integrity.

α-DG Hypo-glycosylation in Muscle Resembling Congenital Muscular Dystrophy. Using IIH6 antibody,[24] we detected significant hypo-glycosylation of α-DG in the patient skeletal muscle compared with normal control (Fig. 3B, left; band intensity values: 182,500 vs. 466,900) with a lighter and a broader smear from 110–180 kDa compared with a strong signal for normal muscle at 130–180 kDa. Using AF6868 antibody[24] the normal muscle showed a band at 150 kDa and β-DG at ~43 kDa, similar to the patient muscle with equivalent β-DG fragment with the exception of a second smear at 90–100 kDa (Fig. 3B, right).

No Substantial Structural Change on MD Simulation of V727M on GNE Kinase-Domain

Mutation mapping and MD simulation suggests that all known missense variants in kinase domain are away from the active site, and there is no significant structural change due to only the V727M change in the kinase domain (Fig. 3C–F). But subtle fold changes can be observed (Fig. 3D,F) when V727 and M727 structures are overlaid. Subtle fold changes in concert with another variant can cause a substantial difference in functional output.

DISCUSSION

Our study provides important insights for molecular diagnostic approaches to understand the pathological and molecular nature of unusual myopathies. We report here a family having a patient with a novel upstream promoter-region large deletion in the *GNE* gene, which abolishes expression of the respective allele. Previous reports showed that patients with compound heterozygous variants in both epimerase and kinase GNE mdomains manifest more severe phenotypes than those with both variants in 1 domain,[25] suggesting that mild pathogenicity of missense variants in each domain needed for more disease severity. Although V727M pathogenicity is uncertain given its relatively high prevalence in South Asians, the most parsimonious conclusion given many other similar reports is that this compound heterozygous state contributes to the pathology.

Generally, inflammation is not associated with HIBM, in which quadriceps muscles are relatively spared of any rimmed vacuolar pathology compared with other muscle types, as seen in studies of smaller numbers of individuals with inclusion-body myositis (IBM)[22] and in larger study cohorts.[25] The presence of both inflammation and rimmed vacuoles in the quadriceps muscle of this patient is not characteristic of either primary inflammatory or rimmed vacuolar myopathies. The second causal variant was inferred from the combination of aCGH and RNA-seq that definitively diagnosed the case as GNE-related myopathy, and led to identification of multiple gene expression perturbations. This study shows the involvement of quadriceps muscle directly with both rimmed vacuoles and inflammation, unlike previous reports of individual cases and cohorts,[22,23,26,27] enhancing the molecular-pathological spectrum of GNE-myopathy that is important to understand for patient stratification in clinical trials.

Previously, Zhu et al.[23] showed that large promoter region deletions in *GNE* are common in already clinically diagnosed GNE-myopathy patients, and Garland et al.[22] showed that a combination of such deletions and a V727M missense variant causes wa more severe reduction in *GNE* expression than the combination of V727M and another missense variant. Here, we show that such variant combinations are associated with unique GNE-related myopathy pathology and the clinical/molecular diagnostic hurdles faced. Consequently, it is likely that the combination of reduced transcription due to promoter region deletion and possible V727M-induced subtle altered kinase activity is required for the unique HIBM-like symptoms. Further functional studies are needed to classify the pathogenicity of V727M.

As per ACMG guidelines,[5] because the deletion variant causes a 50% reduction in *GNE* gene expression, we clinically classify the variant as "pathogenic." This potentially results in a significant reduction in key sarcolemmal protein α-DG glycosylation and aberrant expression of core α-DG and β-DG (Fig. 3B), which along with altered expression of genes and pathways found in GO-pathway analysis could explain the muscle wasting and weakness. Disrupted glycan metabolism and glycosyl transferase likely explains hypoglycosylation of α-DG, potentially causing de-regulation of the actin cytoskeleton, cell cortex, sarcolemma, T-tubule, and ECM (Supplementary Fig. S3).

The nature of reduced α-DG glycosylation and overexpression of β-DG fragment is a hallmark of congenital muscular dystrophies (CMDs),[28] found also by Huizing et al.[29] in GNE-related HIBM. Our study contributes to an emerging literature suggesting that GNE-related myopathy shares molecular signatures of "dystroglycanopathy" similar to CMDs, with glycosylation-defect-related muscle wasting and weakness as the primary cause and an inflammatory response as a secondary effect. The muscle structural degeneration in HIBM resembling CMD is possibly due to inability of the sarcolemmal machinery to protect the sarcomere from the load of ECM proteome dysregulation. Overall, functional assays suggest a GNE-associated inherent core muscle

glycosylation defect as the cause for this unusual GNE-related myopathy.

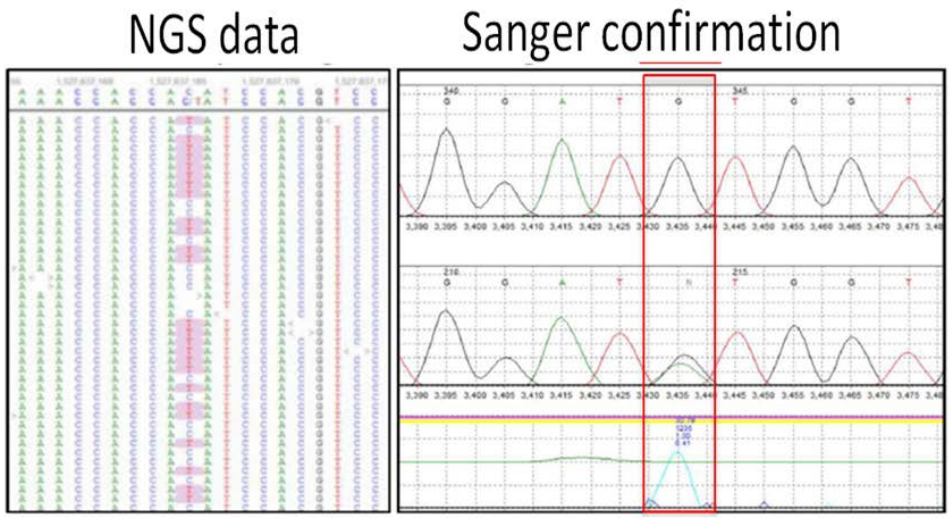
Importantly, this study shows the power of using aCGH, RNA-seq and focused functional assays on target muscle tissue following clinical/pathological clues for improving diagnostic efficiency and timeliness in the evaluation of undiagnosed myopathies. We believe that this approach will be broadly applicable to the diagnosis of NMDs, and will thus harness the advances in clinical genomics and developing precision therapies.

We thank the reported patient and his family, and the normal individuals from whom normal control muscle biopsies were used, for participating in this study. This work was also supported by Paul D. Wellstone Muscular Dystrophy Cooperative Research Center Grant (1U54NS053672). K.P.C. is an investigator of the Howard Hughes Medical Institute. Ethical Publication Statement: The authors confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines. Part of this work has been presented before in the following conferences: American Society of Human Genetics meeting 2016 Vancouver, American Academy of Neurology conference 2017, in Boston, and American College of Medical Genetics and Genomics Annual Meeting 2017, in Phoenix. Data Deposition and Access: Whole-exome sequencing data are not publicly available because consent could not be obtained. The deletion variant associated with the phenotype has been submitted to ClinVar under accession number SCV000599234 (https://www.ncbi.nlm.nih.gov/clinvar/) and will be available after publication. Author Contributions: S.C. and M.H. had full access to all data in this study and take responsibility for the integrity of the data and the accuracy of the data analysis. S.C., M.H. contributed to the study concept and design and S.C., and M.H. contributed to the overall acquisition, analysis, or interpretation of data. JG collected clinical data of patient. S.C., B.R.R.N., M.H. contributed to ES, aCGH, gDNA, cDNA Sanger sequencing analysis and to the exome sequencing in-house pipeline. S.C. and D.G. performed RNA-sequencing. S.C, K.B., H.S., G.G. contributed to RNA-seq data analysis. S.C., M.E.A., K.P.C. contributed to dystroglycan glycosylation experiment. S.J. mapped disease-causing mutations onto the GNE structure. S.C. wrote the manuscript and other authors contributed to proof-reading manuscript draft. All authors contributed to the critical revision of the manuscript for important intellectual content. S.C. and M.H. obtained funding and S.C. contributed to study supervision. The authors have declared no competing interest.

REFERENCES

1. Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Sci Transl Med 2017;9:eaal5209.
2. Lappalainen T, Sammeth M, Friedlander MR, t Hoen PA, Monlong J, Rivas MA, et al. Transcriptome and genome sequencing uncovers functional variation in humans. Nature 2013;501:506–511.
3. Chakravorty S, Hegde M. Gene and variant annotation for Mendelian disorders in the era of advanced sequencing technologies. Annu Rev Genomics Hum Genet 2017;18:229–256.
4. Chakravorty S, Hegde M. Clinical utility of transcriptome sequencing: toward a better diagnosis for mendelian disorders. Clin Chem 2018;64: 882–884.
5. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–424.
6. Eisenberg I, Avidan N, Potikha T, Hochner H, Chen M, Olender T, et al. The UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase gene is mutated in recessive hereditary inclusion body myopathy. Nat Genet 2001;29:83–87.
7. Huizing M, Krasnewich DM. Hereditary inclusion body myopathy: a decade of progress. Biochim Biophys Acta 2009;1792:881–887.
8. Huizing M, Malicdan MCV, Krasnewich DM, Manoli I, Carrillo-Carrasco N. GNE myopathy. In: Beaudet AL, Vogelstein B, Kinzler KW, Antonarakis SE, Ballabio A, Gibson KM, Mitchell G, editors. *The online metabolic and molecular bases of inherited disease*. New York, NY: The McGraw-Hill Companies, Inc; 2014.
9. Varki A. Selectins and other mammalian sialic acid-binding lectins. Curr Opin Cell Biol 1992;4:257–266.
10. Varki A. Diversity in the sialic acids. Glycobiology 1992;2:25–40.
11. Varki A. Sialic acids as ligands in recognition phenomena. FASEB J 1997;11:248–255.
12. Nishino I, Malicdan MC, Murayama K, Nonaka I, Hayashi YK, Noguchi S. Molecular pathomechanism of distal myopathy with rimmed vacuoles. Acta Myol 2005;24:80–83.
13. Malicdan MC, Noguchi S, Nonaka I, Hayashi YK, Nishino I. A Gne knockout mouse expressing human GNE D176V mutation develops features similar to distal myopathy with rimmed vacuoles or hereditary inclusion body myopathy. Hum Mol Genet 2007;16:2669–2682.
14. Berg AT, Chakravorty S, Koh S, Grinspan ZM, Shellhaas RA, Saneto RP, et al. Why West? Comparisons of clinical, genetic and molecular features of infants with and without spasms. PLoS One 2018;13:e0193599.
15. Voermans NC, Guillard M, Doedee R, Lammens M, Huizing M, Padberg GW, et al. Clinical features, lectin staining, and a novel GNE frameshift mutation in hereditary inclusion body myopathy. Clin Neuropathol 2010;29:71–77.
16. Nalini A, Gayathri N, Nishino I, Hayashi YK. GNE myopathy in India. Neurol India 2013;61:371–374.
17. Liewluck T, Pho-Iam T, Limwongse C, Thongnoppakhun W, Boonyapisit K, Raksadawan N, et al. Mutation analysis of the GNE gene in distal myopathy with rimmed vacuoles (DMRV) patients in Thailand. Muscle Nerve 2006;34:775–778.
18. Chaouch A, Brennan KM, Hudson J, Longman C, McConville J, Morrison PJ, et al. Two recurrent mutations are associated with GNE myopathy in the North of Britain. J Neurol Neurosurg Psychiatry 2014; 85:1359–1365.
19. Nishino I, Carrillo-Carrasco N, Argov Z. GNE myopathy: current update and future therapy. J Neurol Neurosurg Psychiatry 2015;86:385–392.
20. Celeste FV, Vilboux T, Ciccone C, de Dios JK, Malicdan MC, Leoyklang P, et al. Mutation update for GNE gene variants associated with GNE myopathy. Hum Mutat 2014;35:915–926.
21. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–291.
22. Garland J, Stephen J, Class B, Gruber A, Ciccone C, Poliak A, et al. Identification of an Alu element-mediated deletion in the promoter region of GNE in siblings with GNE myopathy. Mol Genet Genomic Med 2017;5:410–417.
23. Zhu W, Mitsuhashi S, Yonekawa T, Noguchi S, Huei JC, Nalini A, et al. Missing genetic variations in GNE myopathy: rearrangement hotspots encompassing 5'UTR and founder allele. J Hum Genet 2017;62: 159–166.
24. Briggs DC, Yoshida-Moriguchi T, Zheng T, Venzke D, Anderson ME, Strazzulli A, et al. Structural basis of laminin binding to the LARGE glycans on dystroglycan. Nat Chem Biol 2016;12:810–814.
25. Pogoryelova O, Cammish P, Mansbach H, Argov Z, Nishino I, Skrinar A, et al. Phenotypic stratification and genotype-phenotype correlation in a heterogeneous, international cohort of GNE myopathy patients: first report from the GNE myopathy Disease Monitoring Program, registry portion. Neuromuscul Disord 2018;28:158–168.
26. Yabe I, Higashi T, Kikuchi S, Sasaki H, Fukazawa T, Yoshida K, et al. GNE mutations causing distal myopathy with rimmed vacuoles with inflammation. Neurology 2003;61:384–386.
27. Argov Z, Eisenberg I, Grabov-Nardini G, Sadeh M, Wirguin I, Soffer D, et al. Hereditary inclusion body myopathy: the Middle Eastern genetic cluster. Neurology 2003;60:1519–1523.
28. Mendell JR, Boue DR, Martin PT. The congenital muscular dystrophies: recent advances and molecular insights. Pediatr Dev Pathol 2006;9:427–443.
29. Huizing M, Rakocevic G, Sparks SE, Mamali I, Shatunov A, Goldfarb L, et al. Hypoglycosylation of alpha-dystroglycan in patients with hereditary IBM due to GNE mutations. Mol Genet Metab 2004;81:196–202.

Figure S1



Heterozygous variant c.2179 C>T: G>A (p.V727M) identified in genomic DNA sequence analysis

Figure S2

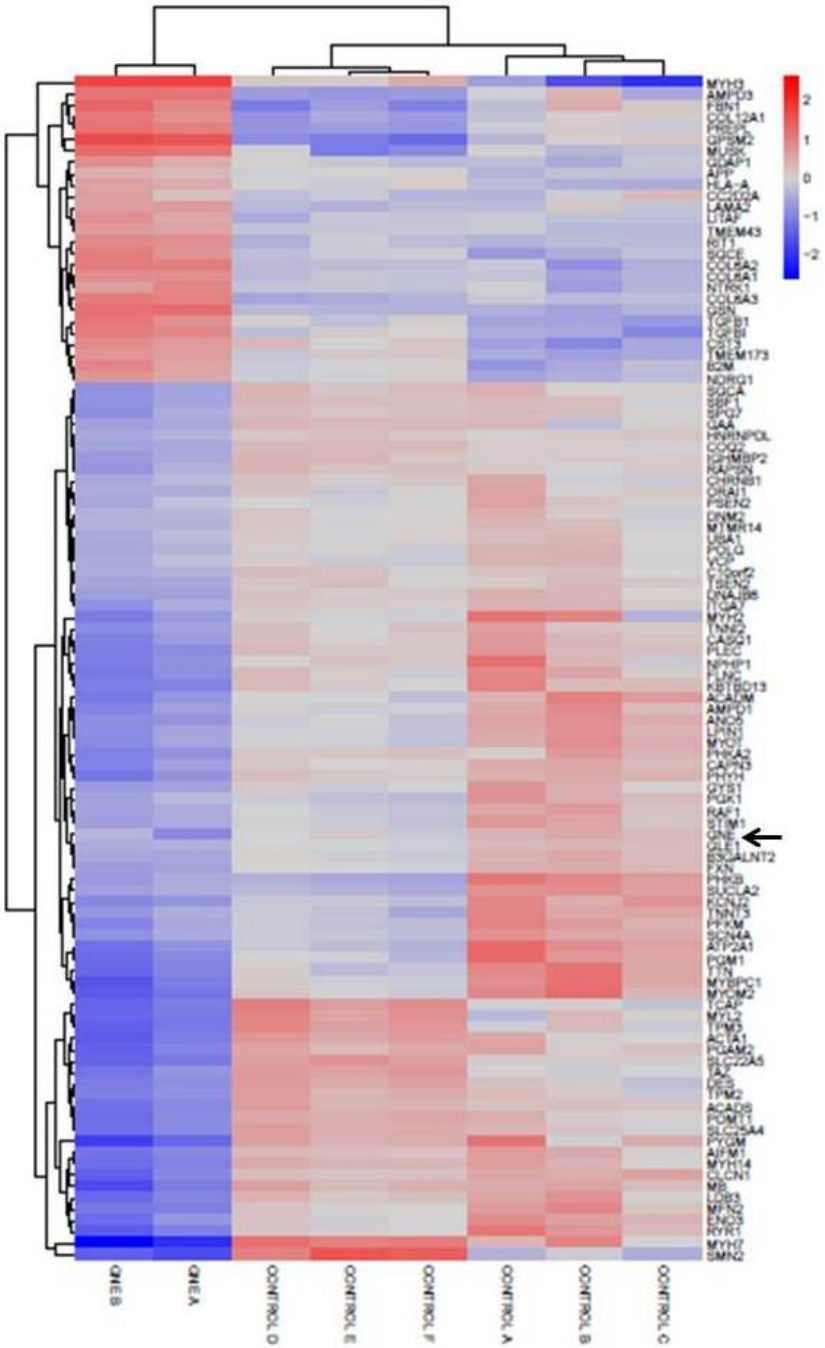

2
1
0
-1
-2
QNEB
QNEA
OCNITROLD
OCNITROLE
OCNITROLF
OCNITROLA
OCNITROLB
OCNITROLC

Figure S3

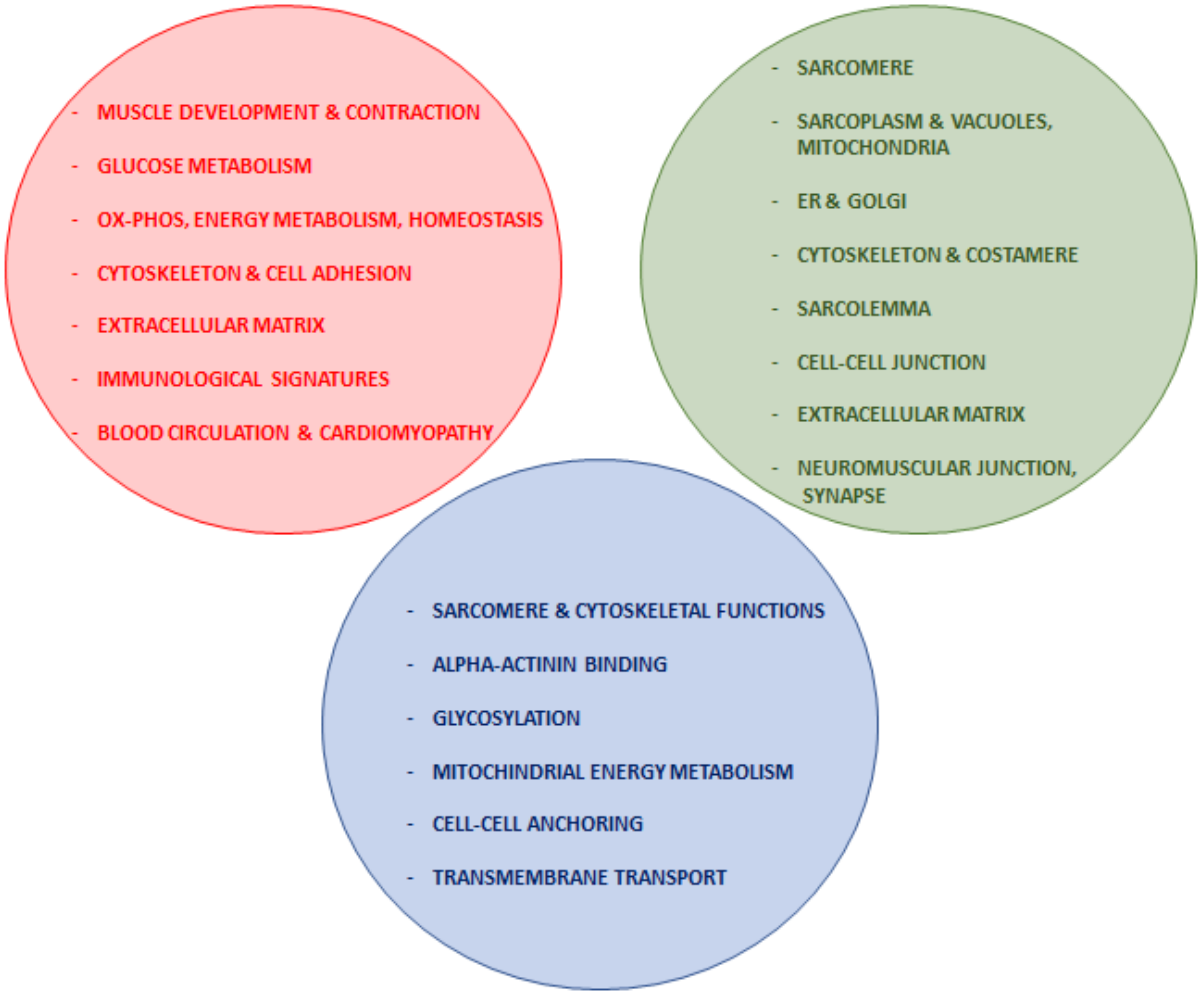

- MUSCLE DEVELOPMENT & CONTRACTION
- GLUCOSE METABOLISM
- OX-PHOS, ENERGY METABOLISM, HOMEOSTASIS
- CYTOSKELETON & CELL ADHESION
- EXTRACELLULAR MATRIX
- IMMUNOLOGICAL SIGNATURES
- BLOOD CIRCULATION & CARDIOMYOPATHY
- SARCOMERE
- SARCOPLASM & VACUOLES, MITOCHONDRIA
- ER & GOLGI
- CYTOSKELETON & COSTAMERE
- SARCOLEMMA
- CELL-CELL JUNCTION
- EXTRACELLULAR MATRIX
- NEUROMUSCULAR JUNCTION, SYNAPSE
- SARCOMERE & CYTOSKELETAL FUNCTIONS
- ALPHA-ACTININ BINDING
- GLYCOSYLATION
- MITOCHINDRIAL ENERGY METABOLISM
- CELL-CELL ANCHORING
- TRANSMEMBRANE TRANSPORT

Table S1.

Gene symbol	Chromosome: position start	Chromosome: position end	Reference DNA	Alternate DNA	ClinVar Accession Number	HGVSc	HGVSp	Predicted Effect	*dbSNP ID*	Variant type	Genotype (hetero-/homo-zygous)
GNE (NM_001128227.2)	9:36,217,445		C	T	VCV000006028.3	c.2179C>T	p.Val727Met	Deleterious	rs12190867	Missense	Heterozygous
GNE (NM_001128227.2)	9:36,259,402	9:36,266,483			SCV000599234			Deleterious		Large Deletion	Heterozygous

HGVSc, Human Genome Variation Society coding sequence name; HGVSp, Human Genome Variation Society protein sequence name; ExAC, Exome Aggregation Consortium.

Table S2: 274 NMD-associated genes with muscle expression as per GTeX (https://gtexportal.org/home/) for focused gene expression analysis

Ensembl ID	Gene_ID
ENSG00000143632	*ACTA1*
ENSG00000198125	*MB*
ENSG00000175084	*DES*
ENSG00000198467	*TPM2*
ENSG00000196091	*MYBPC1*
ENSG00000022267	*FHL1*
ENSG00000111245	*MYL2*
ENSG00000155657	*TTN*
ENSG00000125414	*MYH2*
ENSG00000092054	*MYH7*
ENSG00000108515	*ENO3*
ENSG00000196296	*ATP2A1*
ENSG00000130595	*TNNT3*
ENSG00000105048	*TNNT1*
ENSG00000239474	*KLHL41*
ENSG00000183091	*NEB*

ENSG00000068976	*PYGM*
ENSG00000143549	*TPM3*
ENSG00000173991	*TCAP*
ENSG00000130598	*TNNI2*
ENSG00000148180	*GSN*
ENSG00000109846	*CRYAB*
ENSG00000122367	*LDB3*
ENSG00000086967	*MYBPC2*
ENSG00000143318	*CASQ1*
ENSG00000120729	*MYOT*
ENSG00000128591	*FLNC*
ENSG00000151729	*SLC25A4*
ENSG00000164708	*PGAM2*
ENSG00000136717	*BIN1*
ENSG00000134333	*LDHA*
ENSG00000152556	*PFKM*
ENSG00000166710	*B2M*
ENSG00000196218	*RYR1*
ENSG00000106211	*HSPB1*
ENSG00000092529	*CAPN3*
ENSG00000169714	*CNBP*
ENSG00000079739	*PGM1*
ENSG00000165410	*CFL2*
ENSG00000142173	*COL6A2*
ENSG00000036448	*MYOM2*

ENSG00000152137	*HSPB8*
ENSG00000162688	*AGL*
ENSG00000142156	*COL6A1*
ENSG00000177469	*PTRF*
ENSG00000116748	*AMPD1*
ENSG00000163359	*COL6A3*
ENSG00000234745	*HLA-B*
ENSG00000142798	*HSPG2*
ENSG00000081248	*CACNA1S*
ENSG00000102144	*PGK1*
ENSG00000135424	*ITGA7*
ENSG00000163754	*GYG1*
ENSG00000101439	*CST3*
ENSG00000178209	*PLEC*
ENSG00000054523	*KIF1B*
ENSG00000166147	*FBN1*
ENSG00000206503	*HLA-A*
ENSG00000165280	*VCP*
ENSG00000177666	*PNPLA2*
ENSG00000104812	*GYS1*
ENSG00000168477	*TNXB*
ENSG00000111716	*LDHB*
ENSG00000108823	*SGCA*
ENSG00000054654	*SYNE2*
ENSG00000172037	*LAMB2*

ENSG00000160789	*LMNA*
ENSG00000107537	*PHYH*
ENSG00000075785	*RAB7A*
ENSG00000169567	*HINT1*
ENSG00000152795	*HNRNPDL*
ENSG00000142192	*APP*
ENSG00000197102	*DYNC1H1*
ENSG00000151929	*BAG3*
ENSG00000117054	*ACADM*
ENSG00000130985	*UBA1*
ENSG00000116688	*MFN2*
ENSG00000111046	*MYF6*
ENSG00000104763	*ASAH1*
ENSG00000105993	*DNAJB6*
ENSG00000173402	*DAG1*
ENSG00000015479	*MATR3*
ENSG00000157119	*KLHL40*
ENSG00000132155	*RAF1*
ENSG00000007314	*SCN4A*
ENSG00000140374	*ETFA*
ENSG00000136003	*ISCU*
ENSG00000198947	*DMD*
ENSG00000163069	*SGCB*
ENSG00000135636	*DYSF*
ENSG00000102893	*PHKB*

ENSG00000109099	*PMP22*
ENSG00000136143	*SUCLA2*
ENSG00000060237	*WNK1*
ENSG00000154153	*FAM134B*
ENSG00000104936	*DMPK*
ENSG00000196569	*LAMA2*
ENSG00000111799	*COL12A1*
ENSG00000067177	*PHKA1*
ENSG00000171714	*ANO5*
ENSG00000005893	*LAMP2*
ENSG00000114480	*GBE1*
ENSG00000179295	*PTPN11*
ENSG00000170624	*SGCD*
ENSG00000131018	*SYNE1*
ENSG00000102683	*SGCG*
ENSG00000105379	*ETFB*
ENSG00000104419	*NDRG1*
ENSG00000167323	*STIM1*
ENSG00000134324	*LPIN1*
ENSG00000109063	*MYH3*
ENSG00000171503	*ETFDH*
ENSG00000133020	*MYH8*
ENSG00000079805	*DNM2*
ENSG00000211592	*IGKC*
ENSG00000122971	*ACADS*

ENSG00000182533	*CAV3*
ENSG00000170876	*TMEM43*
ENSG00000197912	*SPG7*
ENSG00000100836	*PABPN1*
ENSG00000143337	*TOR1AIP1*
ENSG00000184640	*SEPT9*
ENSG00000120708	*TGFBI*
ENSG00000170175	*CHRNB1*
ENSG00000124164	*VAPB*
ENSG00000105357	*MYH14*
ENSG00000100241	*SBF1*
ENSG00000198836	*OPA1*
ENSG00000138078	*PREPL*
ENSG00000085998	*POMGNT1*
ENSG00000162430	*SEPN1*
ENSG00000065154	*OAT*
ENSG00000151835	*SACS*
ENSG00000188037	*CLCN1*
ENSG00000171867	*PRNP*
ENSG00000116001	*TIA1*
ENSG00000115904	*SOS1*
ENSG00000169083	*AR*
ENSG00000144061	*NPHP1*
ENSG00000072163	*LIMS2*
ENSG00000102119	*EMD*

ENSG00000138777	*PPA2*
ENSG00000165917	*RAPSN*
ENSG00000189067	*LITAF*
ENSG00000171298	*GAA*
ENSG00000124788	*ATXN1*
ENSG00000122591	*FAM126A*
ENSG00000165029	*ABCA1*
ENSG00000148290	*SURF1*
ENSG00000140521	*POLG*
ENSG00000156709	*AIFM1*
ENSG00000106105	*GARS*
ENSG00000064419	*TNPO3*
ENSG00000178537	*SLC25A20*
ENSG00000000419	*DPM1*
ENSG00000048392	*RRM2B*
ENSG00000171109	*MFN1*
ENSG00000138435	*CHRNA1*
ENSG00000133424	*LARGE1 (LARGE)*
ENSG00000014919	*COX15*
ENSG00000276045	*ORAI1*
ENSG00000104133	*SPG11*
ENSG00000166548	*TK2*
ENSG00000101596	*SMCHD1*
ENSG00000163719	*MTMR14*
ENSG00000134684	*YARS*

ENSG00000198380	*GFPT1*
ENSG00000133812	*SBF2*
ENSG00000135902	*CHRND*
ENSG00000080815	*PSEN1*
ENSG00000044446	*PHKA2*
ENSG00000185963	*BICD2*
ENSG00000070061	*IKBKAP*
ENSG00000206561	*COLQ*
ENSG00000147224	*PRPS1*
ENSG00000100288	*CHKB*
ENSG00000102125	*TAZ*
ENSG00000149311	*ATM*
ENSG00000130714	*POMT1*
ENSG00000177000	*MTHFR*
ENSG00000171680	*PLEKHG5*
ENSG00000119392	*GLE1*
ENSG00000146282	*RARS2*
ENSG00000121957	*GPSM2*
ENSG00000168000	*BSCL2*
ENSG00000184584	*TMEM173*
ENSG00000173085	*COQ2*
ENSG00000173432	*SAA1*
ENSG00000120725	*SIL1*
ENSG00000115170	*ACVR1*
ENSG00000157184	*CPT2*

ENSG00000139132	*FGD4*
ENSG00000174684	*B3GNT1*
ENSG00000143801	*PSEN2*
ENSG00000011198	*ABHD5*
ENSG00000090054	*SPTLC1*
ENSG00000125741	*OPA3*
ENSG00000110090	*CPT1A*
ENSG00000162885	*B3GALNT2*
ENSG00000165996	*PTPLA / HACD1*
ENSG00000198642	*KLHL9*
ENSG00000139116	*KIF21A*
ENSG00000204842	*ATXN2*
ENSG00000133805	*AMPD3*
ENSG00000130816	*DNMT1*
ENSG00000160131	*VMA21*
ENSG00000188157	*AGRN*
ENSG00000144647	*POMGNT2*
ENSG00000109536	*FRG1*
ENSG00000140199	*SLC12A6*
ENSG00000119401	*TRIM32*
ENSG00000087053	*MTMR2*
ENSG00000123700	*KCNJ2*
ENSG00000165060	*FXN*
ENSG00000197375	*SLC22A5*
ENSG00000196811	*CHRNG*

ENSG00000135541	*AHI1*
ENSG00000101986	*ABCD1*
ENSG00000159921	*GNE*
ENSG00000175920	*DOK7*
ENSG00000109689	*STIM2*
ENSG00000106692	*FKTN*
ENSG00000171100	*MTM1*
ENSG00000107371	*EXOSC3*
ENSG00000112367	*FIG4*
ENSG00000140650	*PMM2*
ENSG00000170892	*TSEN34*
ENSG00000105329	*TGFB1*
ENSG00000179085	*DPM3*
ENSG00000009830	*POMT2*
ENSG00000139131	*YARS2*
ENSG00000181027	*FKRP*
ENSG00000138379	*MSTN*
ENSG00000046651	*OFD1*
ENSG00000134569	*LRP4*
ENSG00000156873	*PHKG2*
ENSG00000143622	*RIT1*
ENSG00000173540	*GMPPB*
ENSG00000119523	*ALG2*
ENSG00000148384	*INPP5E*
ENSG00000006530	*AGK*

ENSG00000132740	*IGHMBP2*
ENSG00000136908	*DPM2*
ENSG00000145794	*MEGF10*
ENSG00000106080	*FKBP14*
ENSG00000118600	*TMEM5*
ENSG00000109618	*SEPSECS*
ENSG00000048342	*CC2D2A*
ENSG00000154743	*TSEN2*
ENSG00000198707	*CEP290*
ENSG00000172269	*DPAGT1*
ENSG00000256525	*POLG2*
ENSG00000112357	*PEX7*
ENSG00000107815	*TWNK (C10orf2)*
ENSG00000187049	*TMEM216*
ENSG00000182173	*TSEN54*
ENSG00000104381	*GDAP1*
ENSG00000100749	*VRK1*
ENSG00000006788	*MYH13*
ENSG00000018236	*CNTN1*
ENSG00000172062	*SMN1*
ENSG00000169379	*ARL13B*
ENSG00000144681	*STAC*
ENSG00000234438	*KBTBD13*
ENSG00000205571	*SMN2*
ENSG00000103494	*RPGRIP1L*

ENSG00000164953	*TMEM67*
ENSG00000214960	*ISPD*
ENSG00000030304	*MUSK*
ENSG00000198400	*NTRK1*
ENSG00000108556	*CHRNE*
ENSG00000185900	*POMK*
ENSG00000185960	*SHOX*
ENSG00000111199	*TRPV4*

Table S3: Significantly differentially regulated genes (89 genes) (padj<0.05) in patient muscle compared to 6 normal control muscles. These genes are used as input in MSigDB analysis for gene ontology-pathway analysis.

Genes	baseMean	log_2FoldChange	lfcSE	stat	pvalue	padj
PYGM	215662.939	-3.4612982	0.4398334	-7.8695665	3.56E-15	8.48E-13
KBTBD13	834.350077	-2.5521899	0.3278288	-7.7851303	6.96E-15	1.54E-12
PLEC	34084.2202	-2.1363056	0.28215694	-7.5713381	3.69E-14	7.19E-12

GSN	44403.78	2.06903531	0.27972333	7.39672048	1.40E-13	2.43E-11
NPHP1	1750.16841	-2.7026669	0.36989836	-7.3065123	2.74E-13	4.60E-11
RYR1	100128.526	-2.9433339	0.46996072	-6.262936	3.78E-10	3.01E-08
AIFM1	2142.10284	-2.1807285	0.35164465	-6.2015119	5.59E-10	4.26E-08
MYH14	7237.25266	-2.20039	0.36294605	-6.0625814	1.34E-09	9.51E-08
CASQ1	95810.2009	-1.9512235	0.32853484	-5.9391677	2.86E-09	1.82E-07
MYH7	923013.539	-3.4582813	0.59405385	-5.8214946	5.83E-09	3.37E-07
SGCE	631.897075	2.07762162	0.3603512	5.76554653	8.14E-09	4.52E-07
FLNC	113199.118	-2.3586583	0.40940377	-5.7612032	8.35E-09	4.60E-07
TNNI2	172860.853	-1.7268595	0.30167164	-5.7243017	1.04E-08	5.57E-07
PGAM2	43850.9319	-2.209193	0.39037558	-5.6591474	1.52E-08	7.76E-07
MB	479777.95	-2.3027802	0.41951718	-5.4891202	4.04E-08	1.81E-06
CLCN1	3279.4856	-2.1748322	0.4049899	-	7.87E-08	3.19E-06

			4	5.3700894		
PGM1	29406.7725	-2.9588745	0.56551406	-5.2321855	1.68E-07	6.27E-06
ENO3	191400.056	-2.3100247	0.46521427	-4.965507	6.85E-07	2.06E-05
RIT1	419.950758	1.61743507	0.32905436	4.91540385	8.86E-07	2.52E-05
COL6A3	9742.82978	1.41302185	0.29368034	4.81142814	1.50E-06	3.99E-05
B2M	21182.2707	1.89871954	0.39670841	4.7861843	1.70E-06	4.45E-05
PHYH	7662.70884	-1.6982238	0.3619008	-4.6925119	2.70E-06	6.53E-05
GYS1	14872.3899	-1.7507199	0.37422308	-4.6782789	2.89E-06	6.88E-05
ACTA1	2023320.57	-2.2579906	0.4845262	-4.6602034	3.16E-06	7.40E-05
LDB3	109962.948	-2.0660101	0.45048757	-4.5861644	4.51E-06	0.00010032
POMT1	1866.39333	-1.9232486	0.42033627	-4.5755	4.75E-06	0.00010493
DNAJB6	8070.63876	-1.3186894	0.29817697	-4.4225058	9.76E-06	0.00019102
ORAI1	1372.91648	-1.3330951	0.30673161	-4.3461289	1.39E-05	0.00025765

SGCA	12815.5575	-1.4173678	0.32652417	-4.3407743	1.42E-05	0.00026293
ACADS	5475.57591	-1.5979215	0.36836652	-4.3378576	1.44E-05	0.0002659
MFN2	13010.3006	-2.2178764	0.51265305	-4.3262718	1.52E-05	0.00027772
ATP2A1	239784.606	-2.8213432	0.66156058	-4.2646785	2.00E-05	0.0003489
MYOM2	33043.5445	-2.8007561	0.66448866	-4.2149043	2.50E-05	0.00042226
SLC25A4	44891.1127	-1.7545127	0.42730966	-4.1059516	4.03E-05	0.00062435
MYH2	611457.911	-2.5538693	0.63146883	-4.0443316	5.25E-05	0.00076393
MYBPC1	448829.438	-2.6462044	0.66094789	-4.0036506	6.24E-05	0.00087808
CAPN3	39210.2503	-1.5918617	0.39916361	-3.987993	6.66E-05	0.00092335
NDRG1	2317.49837	1.5120461	0.37988546	3.9802684	6.88E-05	0.00094735
COL6A1	15122.9572	1.28880064	0.32953369	3.9109829	9.19E-05	0.00119855
SBF1	4205.44285	-1.2675074	0.32805732	-3.8636767	0.00011169	0.00140553
COL6A2	18952.899	1.44891949	0.37858503	3.8271970	0.0001296	0.0015903

	7		3	1	1	7
KCNJ2	456.942499	-2.3100857	0.6055750 6	-3.8146975	0.0001363 5	0.0016545 6
TGFB1	482.043015	2.00426008	0.5323858 4	3.7646758 1	0.0001667 7	0.0019336
PFKM	43890.3519	-2.1303397	0.5669386 6	-3.7576194	0.0001715 4	0.0019675 2
ITGA7	13962.8101	-1.2742624	0.3397428 4	-3.7506674	0.0001763 6	0.0020126 9
PSEN2	1138.16652	-1.2270766	0.3311619 8	-3.7053667	0.0002110 9	0.0023113 4
TTN	631210.184	-2.5484081	0.6900353	-3.6931562	0.0002214 9	0.0024049
DNM2	4180.33663	-1.0307984	0.2827690 9	-3.6453715	0.0002670 1	0.0027887 8
SLC22A5	680.62025	-1.7969364	0.4985081 1	-3.6046282	0.0003126	0.0031663 9
CHRNB1	4568.29833	-1.1837816	0.3285762 8	-3.6027605	0.0003148 6	0.0031874 5
STIM1	4900.35031	-1.6128928	0.4546325 8	-3.5476843	0.0003886 3	0.0037857 1
SPG7	3859.12342	-1.1782469	0.3345299 5	-3.5220968	0.0004281 5	0.0040919 2
SCN4A	13396.1667	-1.7862696	0.5126258 7	-3.4845483	0.0004929 7	0.0046045 1

SMN2	88.1590371	-6.1722247	1.79617707	-3.4363119	0.00058969	0.00528464
TGFBI	1583.74013	1.84820871	0.54062073	3.41867892	0.00062926	0.00555928
TPM2	573542.734	-1.5178652	0.44752241	-3.3917078	0.00069458	0.00601539
FXN	595.139701	-1.3917789	0.41597859	-3.3457945	0.00082047	0.00685088
TNNT3	251250.408	-1.9370015	0.58981673	-3.2840736	0.00102318	0.00817778
TMEM43	2196.71385	0.90742838	0.27656051	3.28112055	0.00103396	0.00824569
GPSM2	602.5591	2.46213357	0.76727739	3.20892233	0.00133233	0.00998802
LAMA2	3102.98638	1.19758626	0.37364158	3.20517392	0.00134981	0.01010599
PGK1	13193.7964	-1.5201907	0.47931156	-3.1716128	0.00151595	0.01113808
PHKA2	1096.6922	-1.2837079	0.4069878	-3.1541679	0.00160956	0.01166604
ANO5	3125.54431	-1.7523947	0.55796989	-3.1406618	0.00168567	0.01210031
APP	6162.87764	0.92508247	0.29457053	3.14044476	0.00168692	0.01210449
TMEM173	982.40931	1.51348168	0.5060310	2.9908868	0.0027816	0.0177453

	9		8	3	9	7
AMPD1	23890.9063	-1.6100815	0.54535775	-2.95234	0.00315375	0.01958767
LITAF	1161.61026	1.0220765	0.34644392	2.95019317	0.00317575	0.0196838
MTMR14	1282.30568	-0.9269699	0.31773191	-2.9174592	0.00352896	0.02134119
POLG	1974.56401	-1.0687552	0.36706357	-2.9116353	0.00359542	0.02162627
DES	583290.769	-1.4781212	0.51026873	-2.8967504	0.0037705	0.02244083
MYH3	4250.11097	2.9450246	1.01950043	2.88869382	0.00386846	0.02290343
TPM3	190146.127	-1.510777	0.52298666	-2.8887487	0.00386778	0.02290343
GAA	2696.46772	-1.1674104	0.40703854	-2.8680585	0.00412999	0.02411292
CST3	10732.2211	1.73300964	0.60881784	2.84651586	0.00442005	0.02544551
LPIN1	4328.6256	-1.5336107	0.54668151	-2.8053093	0.00502683	0.02817733
PHKB	5505.96056	-2.1515463	0.76909049	-2.7975203	0.00514965	0.02870587
PREPL	2116.95356	1.47858538	0.53481497	2.76466713	0.00569809	0.03097542

B3GALNT2	2086.75774	-1.1710371	0.42463445	-2.7577535	0.00582001	0.03141763
IGHMBP2	630.393851	-0.9320425	0.34504484	-2.7012214	0.00690853	0.03584086
RAF1	7484.67053	-1.3281564	0.49373842	-2.69	0.0071452	0.03667991
GLE1	1076.75474	-0.9773555	0.36457519	-2.6808063	0.0073445	0.03736596
GDAP1	145.006888	1.24674629	0.46891197	2.6588067	0.00784179	0.03925806
GNE	411.244399	-1.3122427	0.4937026	-2.6579618	0.00786148	0.03932043
MYOT	53803.5583	-1.5024566	0.57137421	-2.6295492	0.00854982	0.0417475
COL12A1	2107.53845	1.46670231	0.55922747	2.62272937	0.00872285	0.04246763
TSEN2	277.62777	-0.9178723	0.35527207	-2.5835757	0.0097782	0.04634729
ACADM	6594.66807	-1.6631506	0.64479386	-2.5793524	0.00989858	0.04680999
RAPSN	2602.37132	-0.8788375	0.34451307	-2.5509556	0.0107428	0.04969688

Table S4:

42

Biological Process/Pathways							
Compiled Pathways	**Gene Set Name**	**# Genes in Gene Set (K)**	**Description**	**# Genes in Overlap (k)**	**k/K**	**p-value**	**FDR q-value**
MUSCLE DEVELOPMENT AND CONTRACTION	HALLMARK_MYOGENESIS	200	Genes involved in development of skeletal muscle (myogenesis).	32	0.16	2.55E-53	2.77E-49
	GO_MUSCLE_SYSTEM_PROCESS	282	A organ system process carried out at the level of a muscle. Muscle tissue is composed of contractile cells or fibers.	26	0.0922	1.38E-36	7.50E-33
	GO_MUSCLE_CONTRACTION	233	A process in which force is generated within muscle tissue, resulting in a change in muscle geometry. Force generation involves a chemo-mechanical energy conversion step that is carried out by the actin/myosin complex activity, which generates force through ATP hydrolysis.	23	0.0987	4.72E-33	1.71E-29
	GO_ACTIN_MYOSIN_FILAMENT_SLIDING	38	The sliding movement of actin thin filaments and myosin thick filaments past each other.	11	0.2895	8.72E-22	2.37E-18
	GO_SYSTEM_PROCESS	1785	A multicellular organismal process carried out by any of the organs or tissues in an organ system. An organ system is a regularly interacting or interdependent group of organs or tissues that work together to carry out a biological objective.	31	0.0174	1.38E-21	3.01E-18
	GO_ACTIN_FILAMENT_BASE	93	Movement of organelles or other particles along actin filaments, or sliding of actin	13	0.1398	4.94E-21	8.95E-18

	D_MOVEMENT		filaments past each other, mediated by motor proteins.				
	GO_ACTIN_MEDIATED_CELL_CONTRACTION	74	The actin filament-based process in which cytoplasmic actin filaments slide past one another resulting in contraction of all or part of the cell body.	12	0.1622	2.56E-20	3.77E-17
	GO_ACTIN_FILAMENT_BASED_PROCESS	450	Any cellular process that depends upon or alters the actin cytoskeleton, that part of the cytoskeleton comprising actin filaments and their associated proteins.	19	0.0422	2.77E-20	3.77E-17
	GO_STRIATED_MUSCLE_CONTRACTION	99	A process in which force is generated within striated muscle tissue, resulting in the shortening of the muscle. Force generation involves a chemo-mechanical energy conversion step that is carried out by the actin/myosin complex activity, which generates force through ATP hydrolysis. Striated muscle is a type of muscle in which the repeating units (sarcomeres) of the contractile myofibrils are arranged in registry throughout the cell, resulting in transverse or oblique striations observable at the level of the light microscope.	11	0.1111	8.31E-17	1.00E-13
	GO_MUSCLE_STRUCTURE_DEVELOPMENT	432	The progression of a muscle structure over time, from its formation to its mature state. Muscle structures are contractile cells, tissues or organs that are found in multicellular organisms.	16	0.037	2.66E-16	2.89E-13
	GO_SKELETAL_MUSCLE_CONTRACTION	31	A process in which force is generated within skeletal muscle tissue, resulting in a change in muscle geometry. Force generation involves a chemo-mechanical energy conversion step that is carried out by the actin/myosin complex	8	0.2581	1.09E-15	1.08E-12

			activity, which generates force through ATP hydrolysis. In the skeletal muscle, the muscle contraction takes advantage of an ordered sarcomeric structure and in most cases it is under voluntary control.				
	GO_MUSCLE_ORGAN_DEVELOPMENT	277	The process whose specific outcome is the progression of the muscle over time, from its formation to the mature structure. The muscle is an organ consisting of a tissue made up of various elongated cells that are specialized to contract and thus to produce movement and mechanical work.	13	0.0469	9.77E-15	8.85E-12
	GO_MULTICELLULAR_ORGANISMAL_MOVEMENT	41	Any physiological process involved in changing the position of a multicellular organism or an anatomical part of a multicellular organism.	8	0.1951	1.30E-14	1.01E-11
	GO_MUSCLE_CELL_DEVELOPMENT	128	The process whose specific outcome is the progression of a muscle cell over time, from its formation to the mature structure. Muscle cell development does not include the steps involved in committing an unspecified cell to the muscle cell fate.	9	0.0703	3.99E-12	1.97E-09
	REACTOME_STRIATED_MUSCLE_CONTRACTION	27	Genes involved in Striated Muscle Contraction	6	0.2222	1.27E-11	5.33E-09
	GO_STRIATED_MUSCLE_CELL_DIFFERENTIATION	173	The process in which a relatively unspecialized cell acquires specialized features of a striated muscle cell; striated muscle fibers are divided by transverse bands into striations, and cardiac and voluntary muscle are types of striated muscle.	9	0.052	6.04E-11	2.19E-08

	GO_ACTOMYOSIN_STRUCTURE_ORGANIZATION	77	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly of cytoskeletal structures containing both actin and myosin or paramyosin. The myosin may be organized into filaments.	7	0.0909	1.73E-10	5.53E-08
	GO_MOVEMENT_OF_CELL_OR_SUBCELLULAR_COMPONENT	1275	The directed, self-propelled movement of a cell or subcellular component without the involvement of an external agent such as a transporter or a pore.	17	0.0133	3.56E-10	9.91E-08
	GO_MYOFIBRIL_ASSEMBLY	48	Formation of myofibrils, the repeating units of striated muscle.	6	0.125	5.11E-10	1.32E-07
	REACTOME_MUSCLE_CONTRACTION	48	Genes involved in Muscle contraction	6	0.125	5.11E-10	1.32E-07
	GO_TISSUE_DEVELOPMENT	1518	The process whose specific outcome is the progression of a tissue over time, from its formation to the mature structure.	18	0.0119	6.51E-10	1.61E-07
	GO_MUSCLE_CELL_DIFFERENTIATION	237	The process in which a relatively unspecialized cell acquires specialized features of a muscle cell.	9	0.038	9.84E-10	2.23E-07
	GO_SARCOMERE_ORGANIZATION	27	The myofibril assembly process that results in the organization of muscle actomyosin into sarcomeres. The sarcomere is the repeating unit of a myofibril in a muscle cell, composed of an array of overlapping thick and thin filaments between two adjacent Z discs.	5	0.1852	1.90E-09	3.97E-07
	GO_REGULATION_OF_SKELETAL_MUSCLE_CONTRACTIO	11	Any process that modulates the frequency, rate or extent of skeletal muscle contraction.	4	0.3636	4.29E-09	8.64E-07

	N						
	GO_REGULATION_OF_SYSTEM_PROCESS	507	Any process that modulates the frequency, rate or extent of a system process, a multicellular organismal process carried out by any of the organs or tissues in an organ system.	11	0.0217	4.45E-09	8.81E-07
	GO_MUSCLE_FIBER_DEVELOPMENT	47	The process whose specific outcome is the progression of the muscle fiber over time, from its formation to the mature structure. In skeletal muscle, fibers are formed by the maturation of myotubes. They can be classed as slow, intermediate/fast or fast.	5	0.1064	3.50E-08	5.36E-06
	GO_CELL_DEVELOPMENT	1426	The process whose specific outcome is the progression of the cell over time, from its formation to the mature structure. Cell development does not include the steps involved in committing a cell to a specific fate.	15	0.0105	9.82E-08	1.37E-05
	GO_CELLULAR_COMPONENT_MORPHOGENESIS	900	The process in which cellular structures, including whole cells or cell parts, are generated and organized.	12	0.0133	1.77E-07	2.11E-05
	GO_PROTEIN_COMPLEX_BIOGENESIS	1132	A cellular process that results in the biosynthesis of constituent macromolecules, assembly, and arrangement of constituent parts of a protein complex. Includes the synthesis of non-protein components, and those protein modifications that are involved in synthesis or assembly of the complex.	13	0.0115	2.86E-07	3.27E-05
	GO_ANATOMICAL_STRUCTURE_FORMATION_INVOLVED_IN_MORPH	957	The developmental process pertaining to the initial formation of an anatomical structure from unspecified parts. This process begins with the specific processes that contribute to the appearance of the discrete structure and	12	0.0125	3.39E-07	3.76E-05

	OGENESIS		ends when the structural rudiment is recognizable. An anatomical structure is any biological entity that occupies space and is distinguished from its surroundings. Anatomical structures can be macroscopic such as a carpel, or microscopic such as an acrosome.				
	GO_REGULATION_OF_MUSCLE_CONTRACTION	147	Any process that modulates the frequency, rate or extent of muscle contraction.	6	0.0408	4.51E-07	4.91E-05
	GO_PROTEIN_COMPLEX_SUBUNIT_ORGANIZATION	1527	Any process in which macromolecules aggregate, disaggregate, or are modified, resulting in the formation, disassembly, or alteration of a protein complex.	18	0.0118	7.15E-10	1.65E-07
GLUCOSE METABOLISM	REACTOME_GLUCOSE_METABOLISM	69	Genes involved in Glucose metabolism	9	0.1304	1.30E-14	1.01E-11
	GO_SINGLE_ORGANISM_CATABOLIC_PROCESS	957	A catabolic process - chemical reactions and pathways resulting in the breakdown of substances - which involves a single organism.	18	0.0188	3.51E-13	2.01E-10
	GO_CARBOHYDRATE_CATABOLIC_PROCESS	113	The chemical reactions and pathways resulting in the breakdown of carbohydrates, any of a group of organic compounds based of the general formula Cx(H2O)y.	9	0.0796	1.28E-12	6.63E-10
	GO_POLYSACCHARIDE_CATABOLIC_PROCESS	24	The chemical reactions and pathways resulting in the breakdown of a polysaccharide, a polymer of many (typically more than 10) monosaccharide residues linked glycosidically.	6	0.25	5.82E-12	2.75E-09
	GO_GLUCAN_METABOLIC_PROCESS	58	The chemical reactions and pathways involving glucans, polysaccharides consisting only of glucose residues.	7	0.1207	2.23E-11	8.98E-09

	GO_CELLULAR_CARBOHYDRATE_CATABOLIC_PROCESS	33	The chemical reactions and pathways resulting in the breakdown of carbohydrates, any of a group of organic compounds based of the general formula Cx(H2O)y, as carried out by individual cells.	6	0.1818	4.72E-11	1.77E-08
	GO_POLYSACCHARIDE_METABOLIC_PROCESS	80	The chemical reactions and pathways involving a polysaccharide, a polymer of many (typically more than 10) monosaccharide residues linked glycosidically.	7	0.0875	2.27E-10	6.51E-08
	GO_CELLULAR_CARBOHYDRATE_METABOLIC_PROCESS	144	The chemical reactions and pathways involving carbohydrates, any of a group of organic compounds based of the general formula Cx(H2O)y, as carried out by individual cells.	8	0.0556	4.35E-10	1.18E-07
	GO_CARBOHYDRATE_METABOLIC_PROCESS	662	The chemical reactions and pathways involving carbohydrates, any of a group of organic compounds based of the general formula Cx(H2O)y. Includes the formation of carbohydrate derivatives by the addition of a carbohydrate residue to another molecule.	13	0.0196	5.28E-10	1.34E-07
	GO_CATABOLIC_PROCESS	1773	The chemical reactions and pathways resulting in the breakdown of substances, including the breakdown of carbon compounds with the liberation of energy for use by the cell or organism.	19	0.0107	1.08E-09	2.36E-07
	REACTOME_METABOLISM_OF_CARBOHYDRATES	247	Genes involved in Metabolism of carbohydrates	9	0.0364	1.41E-09	3.02E-07
	REACTOME_GLYCOGEN_BREAKDOWN_GLYCOGENOLYSIS	18	Genes involved in Glycogen breakdown (glycogenolysis)	4	0.2222	3.94E-08	5.95E-06

	GO_HEXOSE_CATABOLIC_PROCESS	49	The chemical reactions and pathways resulting in the breakdown of hexose, any monosaccharide with a chain of six carbon atoms in the molecule.	5	0.102	4.33E-08	6.46E-06
	GO_GLYCOSYL_COMPOUND_METABOLIC_PROCESS	368	The chemical reactions and pathways involving glycosyl compound.	9	0.0245	4.44E-08	6.53E-06
	GO_MONOCARBOXYLIC_ACID_METABOLIC_PROCESS	503	The chemical reactions and pathways involving monocarboxylic acids, any organic acid containing one carboxyl (COOH) group or anion (COO-).	10	0.0199	5.30E-08	7.69E-06
	GO_ORGANONITROGEN_COMPOUND_METABOLIC_PROCESS	1796	The chemical reactions and pathways involving organonitrogen compound.	17	0.0095	5.63E-08	8.06E-06
	GO_MONOSACCHARIDE_CATABOLIC_PROCESS	59	The chemical reactions and pathways resulting in the breakdown of monosaccharides, polyhydric alcohols containing either an aldehyde or a keto group and between three to ten or more carbon atoms.	5	0.0847	1.12E-07	1.54E-05
	GO_CARBOHYDRATE_DERIVATIVE_METABOLIC_PROCESS	1047	The chemical reactions and pathways involving carbohydrate derivative.	13	0.0124	1.18E-07	1.58E-05
	GO_GLUCOSE_METABOLIC_PROCESS	119	The chemical reactions and pathways involving glucose, the aldohexose gluco-hexose. D-glucose is dextrorotatory and is sometimes known as dextrose; it is an important source of energy for living organisms and is found free as	6	0.0504	1.29E-07	1.71E-05

			well as combined in homo- and hetero-oligosaccharides and polysaccharides.				
	KEGG_GLYCOLYSIS_GLUCONEOGENESIS	62	Glycolysis / Gluconeogenesis	5	0.0806	1.44E-07	1.78E-05
	GO_ORGANOPHOSPHATE_METABOLIC_PROCESS	885	The chemical reactions and pathways involving organophosphates, any phosphate-containing organic compound.	12	0.0136	1.48E-07	1.80E-05
	GO_PYRUVATE_METABOLIC_PROCESS	65	The chemical reactions and pathways involving pyruvate, 2-oxopropanoate.	5	0.0769	1.83E-07	2.17E-05
	GO_GLUCOSE_CATABOLIC_PROCESS	29	The chemical reactions and pathways resulting in the breakdown of glucose, the aldohexose gluco-hexose.	4	0.1379	3.01E-07	3.37E-05
	REACTOME_GLYCOLYSIS	29	Genes involved in Glycolysis	4	0.1379	3.01E-07	3.37E-05
OXIDATIVE PHOSPHORYLATION & ENERGY METABOLISM, HOMEOSTASIS	GO_NUCLEOTIDE_PHOSPHORYLATION	61	The process of introducing one or more phosphate groups into a nucleotide to produce a phosphorylated nucleoside.	5	0.082	1.33E-07	1.71E-05
(OX-PHOS, ENERGY METABOLISM, HOMEOSTASIS)	GO_RIBONUCLEOSIDE_DIPHOSPHATE_METABOLIC_PROCESS	64	The chemical reactions and pathways involving a ribonucleoside diphosphate, a compound consisting of a nucleobase linked to a ribose sugar esterified with diphosphate on the sugar.	5	0.0781	1.69E-07	2.05E-05
	GO_CELLULAR	799	Any process that results in a change in state or	11	0.0138	4.37E-07	4.80E-05

	_RESPONSE_TO_OXYGEN_CONTAINING_COMPOUND		activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an oxygen-containing compound stimulus.				
	GO_GENERATION_OF_PRECURSOR_METABOLITES_AND_ENERGY	292	The chemical reactions and pathways resulting in the formation of precursor metabolites, substances from which energy is derived, and any process involved in the liberation of energy from these substances.	12	0.0411	5.30E-13	2.88E-10
	GO_NUCLEOSIDE_MONOPHOSPHATE_METABOLIC_PROCESS	239	The chemical reactions and pathways involving a nucleoside monophosphate, a compound consisting of a nucleobase linked to a deoxyribose or ribose sugar esterified with phosphate on the sugar.	10	0.0418	4.25E-11	1.65E-08
	GO_HOMEOSTATIC_PROCESS	1337	Any biological process involved in the maintenance of an internal steady state.	18	0.0135	8.57E-11	3.01E-08
	GO_ENERGY_RESERVE_METABOLIC_PROCESS	72	The chemical reactions and pathways by which a cell derives energy from stored compounds such as fats or glycogen.	7	0.0972	1.07E-10	3.63E-08
	GO_OXIDATION_REDUCTION_PROCESS	898	A metabolic process that results in the removal or addition of one or more electrons to or from a substance, with or without the concomitant removal or addition of a proton or protons.	15	0.0167	2.05E-10	6.36E-08
	HALLMARK_HYPOXIA	200	Genes up-regulated in response to low oxygen levels (hypoxia).	9	0.045	2.20E-10	6.46E-08
	GO_CELLULAR_HOMEOSTASIS	676	Any process involved in the maintenance of an internal steady state at the level of the cell.	13	0.0192	6.80E-10	1.64E-07
	GO_NUCLEOSIDE_TRIPHOSP	228	The chemical reactions and pathways involving a nucleoside triphosphate, a compound	9	0.0395	7.00E-10	1.65E-07

	HATE_METABOLIC_PROCESS		consisting of a nucleobase linked to a deoxyribose or ribose sugar esterified with triphosphate on the sugar.				
	GO_PURINE_CONTAINING_COMPOUND_METABOLIC_PROCESS	394	The chemical reactions and pathways involving a purine-containing compound, i.e. any compound that contains purine or a formal derivative thereof.	10	0.0254	5.33E-09	1.04E-06
	GO_PHOSPHATE_CONTAINING_COMPOUND_METABOLIC_PROCESS	1977	The chemical reactions and pathways involving the phosphate group, the anion or salt of any phosphoric acid.	19	0.0096	6.39E-09	1.22E-06
	GO_SMALL_MOLECULE_METABOLIC_PROCESS	1767	The chemical reactions and pathways involving small molecules, any low molecular weight, monomeric, non-encoded molecule.	18	0.0102	6.99E-09	1.29E-06
	GO_RESPONSE_TO_OXYGEN_CONTAINING_COMPOUND	1381	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an oxygen-containing compound stimulus.	16	0.0116	9.10E-09	1.62E-06
	GO_RESPONSE_TO_OXYGEN_LEVELS	311	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus reflecting the presence, absence, or concentration of oxygen.	9	0.0289	1.05E-08	1.84E-06
	GO_ENERGY_DERIVATION_BY_OXIDATION_OF_ORGAN	217	The chemical reactions and pathways by which a cell derives energy from organic compounds; results in the oxidation of the compounds from which energy is released.	8	0.0369	1.10E-08	1.91E-06

	IC_COMPOUNDS						
	GO_ATP_GENERATION_FROM_ADP	39	The process of introducing a phosphate group into ADP, adenosine diphosphate, to produce ATP.	5	0.1282	1.33E-08	2.22E-06
	GO_CHEMICAL_HOMEOSTASIS	874	Any biological process involved in the maintenance of an internal steady state of a chemical.	13	0.0149	1.45E-08	2.39E-06
	GO_SMALL_MOLECULE_CATABOLIC_PROCESS	328	The chemical reactions and pathways resulting in the breakdown of small molecules, any low molecular weight, monomeric, non-encoded molecule.	9	0.0274	1.66E-08	2.69E-06
	GO_ADP_METABOLIC_PROCESS	47	The chemical reactions and pathways involving ADP, adenosine 5'-diphosphate.	5	0.1064	3.50E-08	5.36E-06
	GO_NUCLEOBASE_CONTAINING_SMALL_MOLECULE_METABOLIC_PROCESS	535	The cellular chemical reactions and pathways involving a nucleobase-containing small molecule: a nucleobase, a nucleoside, or a nucleotide.	10	0.0187	9.39E-08	1.33E-05
CYTOSKELETON & CELL ADHESION	GO_RESPONSE_TO_INORGANIC_SUBSTANCE	479	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an inorganic substance stimulus.	13	0.0271	9.98E-12	4.52E-09
	GO_RESPONSE_TO_ABIOTIC_STIMULUS	1024	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an abiotic (non-living) stimulus.	17	0.0166	1.22E-11	5.29E-09
	GO_BIOLOGIC	1032	The attachment of a cell or organism to a	16	0.0155	1.42E-10	4.69E-08

	AL_ADHESION		substrate, another cell, or other organism. Biological adhesion includes intracellular attachment between membrane regions.				
	GO_RESPONSE_TO_METAL_ION	333	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a metal ion stimulus.	10	0.03	1.07E-09	2.36E-07
	PID_INTEGRIN1_PATHWAY	66	Beta1 integrin cell surface interactions	6	0.0909	3.68E-09	7.55E-07
	GO_CYTOSKELETON_ORGANIZATION	838	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly of cytoskeletal structures.	13	0.0155	8.84E-09	1.60E-06
	GO_ION_TRANSPORT	1262	The directed movement of charged atoms or small charged molecules into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore.	15	0.0119	2.00E-08	3.20E-06
	PID_SYNDECAN_1_PATHWAY	46	Syndecan-1-mediated signaling events	5	0.1087	3.13E-08	4.93E-06
	KEGG_FOCAL_ADHESION	201	Focal adhesion	7	0.0348	1.40E-07	1.75E-05
	GO_REGULATION_OF_ION_TRANSPORT	592	Any process that modulates the frequency, rate or extent of the directed movement of charged atoms or small charged molecules into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore.	10	0.0169	2.38E-07	2.79E-05
	GO_REGULATION_OF_METAL_ION_TRANSPORT	325	Any process that modulates the frequency, rate, or extent of metal ion transport. Metal ion transport is the directed movement of metal ions, any metal ion with an electric charge,	8	0.0246	2.47E-07	2.85E-05

			into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore.				
	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	200	Genes defining epithelial-mesenchymal transition, as in wound healing, fibrosis and metastasis.	7	0.035	1.36E-07	1.71E-05
EXTRACELLULAR MATRIX	GO_EXTRACELLULAR_STRUCTURE_ORGANIZATION	304	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly of structures in the space external to the outermost structure of a cell. For cells without external protective or external encapsulating structures this refers to space outside of the plasma membrane, and also covers the host cell environment outside an intracellular parasite.	13	0.0428	3.22E-14	2.34E-11
IMMUNOLOGICAL SIGNATURES	GSE32533_WT_VS_MIR17_OVEREXPRESS_ACT_CD4_TCELL_DN	200	Genes down-regulated in activated CD4 [GeneID=920] T cells: wildtype versus over-expressing MIR17 [GeneID=406952].	11	0.055	2.18E-13	1.32E-10
	GSE18804_SPLEEN_MACROPHAGE_VS_COLON_TUMORAL_MACROPHAGE_DN	199	Genes down-regulated in macrophages: control versus colorectal adenocarcinoma conditioned.	9	0.0452	2.10E-10	6.36E-08
	GSE22935_WT_VS_MYD88_KO_MACROPHAGE_48H_MBOVIS_BCG_S	199	Genes up-regulated in macrophages 48h after M. bovis BCG infection: wildtype versus MYD88 [GeneID=4615] knockout.	7	0.0352	1.31E-07	1.71E-05

	TIM_UP						
	GSE22886_NAIVE_CD4_TCELL_VS_12H_ACT_TH2_UP	200	Genes up-regulated in comparison of naive CD4 [GeneID=920] T cells versus stimulated CD4 [GeneID=920] Th2 cells at 12 h.	7	0.035	1.36E-07	1.71E-05
BLOOD CIRCULATION & CARDIOMYOPATHY	KEGG_HYPERTROPHIC_CARDIOMYOPATHY_HCM	85	Hypertrophic cardiomyopathy (HCM)	9	0.1059	9.23E-14	6.27E-11
	KEGG_DILATED_CARDIOMYOPATHY	92	Dilated cardiomyopathy	9	0.0978	1.93E-13	1.23E-10
	KEGG_VIRAL_MYOCARDITIS	73	Viral myocarditis	6	0.0822	6.82E-09	1.28E-06
	GO_REGULATION_OF_HEART_CONTRACTION	221	Any process that modulates the frequency, rate or extent of heart contraction. Heart contraction is the process in which the heart decreases in volume in a characteristic way to propel blood through the body.	8	0.0362	1.27E-08	2.16E-06
	GO_REGULATION_OF_BLOOD_CIRCULATION	295	Any process that modulates the frequency, rate or extent of blood circulation.	8	0.0271	1.18E-07	1.58E-05
Cellular Components							
Compiled Cellular Components	**Gene Set Name**	**# Genes in Gene Set (K)**	**Description**	**# Genes in Overlap (k)**	**k/K**	**p-value**	**FDR q-value**
SARCOMERIC CONTRACTILE	GO_CONTRACTILE_FIBER	211	Fibers, composed of actin, myosin, and associated proteins, found in cells of smooth or striated muscle.	23	0.109	4.40E-34	2.55E-31

APPARATUS							
	GO_I_BAND	121	A region of a sarcomere that appears as a light band on each side of the Z disc, comprising a region of the sarcomere where thin (actin) filaments are not overlapped by thick (myosin) filaments; contains actin, troponin, and tropomyosin; each sarcomere includes half of an I band at each end.	13	0.1074	1.79E-19	2.59E-17
	GO_MYOSIN_FILAMENT	22	A protein complex containing myosin heavy chains, plus associated light chains and other proteins, in which the myosin heavy chains are arranged into a filament.	6	0.2727	3.23E-12	2.34E-10
	GO_MYOFILAMENT	24	Any of the smallest contractile units of a myofibril (striated muscle fiber).	6	0.25	5.82E-12	3.75E-10
	GO_MYOSIN_COMPLEX	67	A protein complex, formed of one or more myosin heavy chains plus associated light chains and other proteins, that functions as a molecular motor; uses the energy of ATP hydrolysis to move actin filaments or to move vesicles or other cargo on fixed actin filaments; has magnesium-ATPase activity and binds actin. Myosin classes are distinguished based on sequence features of the motor, or head, domain, but also have distinct tail regions that are believed to bind specific cargoes.	7	0.1045	6.35E-11	3.69E-09
	GO_MYOSIN_II_COMPLEX	25	A myosin complex containing two class II myosin heavy chains, two myosin essential light chains and two myosin regulatory light chains. Also known as classical myosin or conventional myosin, the myosin II class includes the major muscle myosin of vertebrate and invertebrate muscle, and is characterized by alpha-helical	5	0.2	1.25E-09	5.59E-08

			coiled coil tails that self-assemble to form a variety of filament structures.				
	GO_MUSCLE_MYOSIN_COMPLEX	19	A filament of myosin found in a muscle cell of any type.	4	0.2105	4.98E-08	1.60E-06
	GO_ACTIN_FILAMENT_BUNDLE	57	An assembly of actin filaments that are on the same axis but may be oriented with the same or opposite polarities and may be packed with different levels of tightness.	4	0.0702	4.80E-06	8.43E-05
	GO_ACTOMYOSIN	62	Any complex of actin, myosin, and accessory proteins.	4	0.0645	6.73E-06	1.08E-04
	GO_A_BAND	34	The dark-staining region of a sarcomere, in which myosin thick filaments are present; the center is traversed by the paler H zone, which in turn contains the M line.	3	0.0882	4.02E-05	5.18E-04
	GO_M_BAND	21	The midline of aligned thick filaments in a sarcomere; location of specific proteins that link thick filaments. Depending on muscle type the M band consists of different numbers of M lines.	2	0.0952	7.60E-04	6.58E-03
CYTOSKELETON & COSTAMERE	GO_CYTOSKELETON	1967	Any of the various filamentous elements that form the internal framework of cells, and typically remain after treatment of the cells with mild detergent to remove membrane constituents and soluble components of the cytoplasm. The term embraces intermediate filaments, microfilaments, microtubules, the microtrabecular lattice, and other structures characterized by a polymeric filamentous nature and long-range order within the cell. The various elements of the cytoskeleton not only serve in the maintenance of cellular shape	32	0.0163	1.82E-21	3.51E-19

			but also have roles in other cellular functions, including cellular movement, cell division, endocytosis, and movement of organelles.				
	GO_ACTIN_CYTOSKELETON	444	The part of the cytoskeleton (the internal framework of a cell) composed of actin and associated proteins. Includes actin cytoskeleton-associated complexes.	15	0.0338	9.38E-15	1.09E-12
	GO_CYTOSKELETAL_PART	1436	Any constituent part of the cytoskeleton, a cellular scaffolding or skeleton that maintains cell shape, enables some cell motion (using structures such as flagella and cilia), and plays important roles in both intra-cellular transport (e.g. the movement of vesicles and organelles) and cellular division. Includes constituent parts of intermediate filaments, microfilaments, microtubules, and the microtrabecular lattice.	21	0.0146	3.67E-13	3.04E-11
	GO_CELL_CORTEX	238	The region of a cell that lies just beneath the plasma membrane and often, but not always, contains a network of actin filaments and associated proteins.	5	0.021	1.04E-04	1.19E-03
	GO_COSTAMERE	19	Regular periodic sub membranous arrays of vinculin in skeletal and cardiac muscle cells, these arrays link Z-discs to the sarcolemma and are associated with links to extracellular matrix.	2	0.1053	6.21E-04	5.63E-03
	GO_MICROTUBULE_CYTOSKELETON	1068	The part of the cytoskeleton (the internal framework of a cell) composed of microtubules and associated proteins.	8	0.0075	1.11E-03	8.85E-03
	GO_CLUSTER_OF_ACTIN_BASED_CELL_PROJECTIONS	139	A cell part consisting of multiple, closely packed actin-based cell projections.	3	0.0216	2.54E-03	1.71E-02
	GO_ACTIN_BA	181	A cell projection supported by an assembly of	3	0.0166	5.32E-03	3.15E-02

	SED_CELL_PROJECTION		actin filaments, and which lacks microtubules.				
ER & GOLGI	GO_ENDOPLASMIC_RETICULUM	1631	The irregular network of unit membranes, visible only by electron microscopy, that occurs in the cytoplasm of many eukaryotic cells. The membranes form a complex meshwork of tubular channels, which are often expanded into slitlike cavities called cisternae. The ER takes two forms, rough (or granular), with ribosomes adhering to the outer surface, and smooth (with no ribosomes attached).	19	0.0116	2.72E-10	1.31E-08
	GO_ENDOPLASMIC_RETICULUM_PART	1163	Any constituent part of the endoplasmic reticulum, the irregular network of unit membranes, visible only by electron microscopy, that occurs in the cytoplasm of many eukaryotic cells. The membranes form a complex meshwork of tubular channels, which are often expanded into slitlike cavities called cisternae.	15	0.0129	6.80E-09	2.82E-07
	GO_GOLGI_APPARATUS	1445	A compound membranous cytoplasmic organelle of eukaryotic cells, consisting of flattened, ribosome-free vesicles arranged in a more or less regular stack. The Golgi apparatus differs from the endoplasmic reticulum in often having slightly thicker membranes, appearing in sections as a characteristic shallow semicircle so that the convex side (cis or entry face) abuts the endoplasmic reticulum, secretory vesicles emerging from the concave side (trans or exit face). In vertebrate cells there is usually one such organelle, while in invertebrates and plants, where they are known usually as dictyosomes, there may be several scattered in	16	0.0111	1.71E-08	6.62E-07

			the cytoplasm. The Golgi apparatus processes proteins produced on the ribosomes of the rough endoplasmic reticulum; such processing includes modification of the core oligosaccharides of glycoproteins, and the sorting and packaging of proteins for transport to a variety of cellular locations. Three different regions of the Golgi are now recognized both in terms of structure and function: cis, in the vicinity of the cis face, trans, in the vicinity of the trans face, and medial, lying between the cis and trans regions.				
	GO_ENDOPLASMIC_RETICULUM_LUMEN	201	The volume enclosed by the membranes of the endoplasmic reticulum.	7	0.0348	1.40E-07	4.07E-06
	GO_SMOOTH_ENDOPLASMIC_RETICULUM	33	The smooth endoplasmic reticulum (smooth ER or SER) has no ribosomes attached to it. The smooth ER is the recipient of the proteins synthesized in the rough ER. Those proteins to be exported are passed to the Golgi complex, the resident proteins are returned to the rough ER and the lysosomal proteins after phosphorylation of their mannose residues are passed to the lysosomes. Glycosylation of the glycoproteins also continues. The smooth ER is the site of synthesis of lipids, including the phospholipids. The membranes of the smooth ER also contain enzymes that catalyze a series of reactions to detoxify both lipid-soluble drugs and harmful products of metabolism. Large quantities of certain compounds such as phenobarbital cause an increase in the amount of the smooth ER.	4	0.1212	5.15E-07	1.30E-05

	GO_SARCOPLASMIC_RETICULUM_MEMBRANE	38	The lipid bilayer surrounding the sarcoplasmic reticulum.	4	0.1053	9.22E-07	2.23E-05
	GO_ENDOPLASMIC_RETICULUM_SUBCOMPARTMENT	16	A distinct region of the endoplasmic reticulum	3	0.1875	3.86E-06	7.22E-05
	GO_ENDOSOME_LUMEN	26	The volume enclosed by the membrane of an endosome.	2	0.0769	1.17E-03	9.17E-03
	GO_GOLGI_APPARATUS_PART	893	Any constituent part of the Golgi apparatus, a compound membranous cytoplasmic organelle of eukaryotic cells, consisting of flattened, ribosome-free vesicles arranged in a more or less regular stack.	7	0.0078	1.77E-03	1.29E-02
	GO_ORGANELLE_SUBCOMPARTMENT	311	A compartment that consists of a lumen and an enclosing membrane, and is part of an organelle.	6	0.0193	3.32E-05	4.48E-04
	GO_NUCLEAR_OUTER_MEMBRANE_ENDOPLASMIC_RETICULUM_MEMBRANE_NETWORK	1005	The continuous network of membranes encompassing the nuclear outer membrane and the endoplasmic reticulum membrane.	9	0.009	1.47E-04	1.61E-03
EXTRACELLULAR MATRIX	GO_EXTRACELLULAR_MATRIX	426	A structure lying external to one or more cells, which provides structural support for cells or tissues; may be completely external to the cell (as in animals and bacteria) or be part of the cell (as in plants).	8	0.0188	1.87E-06	3.88E-05
	GO_EXTRACELLULAR_MATRI	125	Any constituent part of the extracellular matrix, the structure lying external to one or more	5	0.04	4.75E-06	8.43E-05

	X_COMPONENT		cells, which provides structural support for cells or tissues; may be completely external to the cell (as in animals) or be part of the cell (as often seen in plants).				
	GO_PROTEINACEOUS_EXTRACELLULAR_MATRIX	356	A layer consisting mainly of proteins (especially collagen) and glycosaminoglycans (mostly as proteoglycans) that forms a sheet underlying or overlying cells such as endothelial and epithelial cells. The proteins are secreted by cells in the vicinity. An example of this component is found in Mus musculus.	7	0.0197	6.31E-06	1.05E-04
	GO_EXTRACELLULAR_SPACE	1376	That part of a multicellular organism outside the cells proper, usually taken to be outside the plasma membranes, and occupied by fluid.	12	0.0087	1.40E-05	2.14E-04
	GO_COLLAGEN_TRIMER	88	A protein complex consisting of three collagen chains assembled into a left-handed triple helix. These trimers typically assemble into higher order structures.	4	0.0455	2.71E-05	3.83E-04
	GO_COMPLEX_OF_COLLAGEN_TRIMERS	23	A complex of collagen trimers such as a fibril or collagen network.	2	0.087	9.14E-04	7.68E-03
NEUROMUSCULAR JUNCTION, SYNAPSE	GO_NEURON_PROJECTION	942	A prolongation or process extending from a nerve cell, e.g. an axon or dendrite.	12	0.0127	2.87E-07	7.92E-06
	GO_NEURON_PART	1265	Any constituent part of a neuron, the basic cellular unit of nervous tissue. A typical neuron consists of a cell body (often called the soma), an axon, and dendrites. Their purpose is to receive, conduct, and transmit impulses in the nervous system.	13	0.0103	9.92E-07	2.30E-05
	GO_AXON	418	The long process of a neuron that conducts	7	0.0167	1.78E-05	2.64E-04

			nerve impulses, usually away from the cell body to the terminals and varicosities, which are sites of storage and release of neurotransmitter.				
	GO_POSTSYNAPSE	378	The part of a synapse that is part of the post-synaptic cell.	6	0.0159	9.74E-05	1.15E-03
	GO_SYNAPSE	754	The junction between a nerve fiber of one neuron and another neuron or muscle fiber or glial cell; the site of interneuronal communication. As the nerve fiber approaches the synapse it enlarges into a specialized structure, the presynaptic nerve ending, which contains mitochondria and synaptic vesicles. At the tip of the nerve ending is the presynaptic membrane; facing it, and separated from it by a minute cleft (the synaptic cleft) is a specialized area of membrane on the receiving cell, known as the postsynaptic membrane. In response to the arrival of nerve impulses, the presynaptic nerve ending secretes molecules of neurotransmitters into the synaptic cleft. These diffuse across the cleft and transmit the signal to the postsynaptic membrane.	8	0.0106	1.11E-04	1.24E-03
	GO_NEUROMUSCULAR_JUNCTION	54	The junction between the axon of a motor neuron and a muscle fiber. In response to the arrival of action potentials, the presynaptic button releases molecules of neurotransmitters into the synaptic cleft. These diffuse across the cleft and transmit the signal to the postsynaptic membrane of the muscle fiber, leading to a change in post-synaptic potential.	3	0.0556	1.62E-04	1.74E-03
	GO_SYNAPSE_PART	610	Any constituent part of a synapse, the junction between a nerve fiber of one neuron and	7	0.0115	1.89E-04	1.99E-03

			another neuron or muscle fiber or glial cell.				
	GO_SOMATODENDRITIC_COMPARTMENT	650	The region of a neuron that includes the cell body (cell soma) and the dendrite, but excludes the axon.	7	0.0108	2.77E-04	2.82E-03
	GO_SMN_COMPLEX	13	A protein complex that contains the survival motor neuron (SMN) protein and at least eight additional integral components, including the Gemin2-8 and Unrip proteins; the complex is found in the cytoplasm and in nuclear Gems, and is involved in spliceosomal snRNP assembly in the cytoplasm and in pre-mRNA splicing in the nucleus.	2	0.1538	2.85E-04	2.85E-03
	GO_MYELIN_SHEATH	168	An electrically insulating fatty layer that surrounds the axons of many neurons. It is an outgrowth of glial cells: Schwann cells supply the myelin for peripheral neurons while oligodendrocytes supply it to those of the central nervous system.	4	0.0238	3.30E-04	3.25E-03
	GO_NEURON_SPINE	121	A small membranous protrusion, often ending in a bulbous head and attached to the neuron by a narrow stalk or neck.	3	0.0248	1.71E-03	1.26E-02
	GO_CELL_BODY	494	The portion of a cell bearing surface projections such as axons, dendrites, cilia, or flagella that includes the nucleus, but excludes all cell projections.	5	0.0101	2.78E-03	1.84E-02
SARCOLEMMA	GO_SARCOLEMMA	125	The outer membrane of a muscle cell, consisting of the plasma membrane, a covering basement membrane (about 100 nm thick and sometimes common to more than one fiber), and the associated loose network of collagen fibers.	15	0.12	4.46E-23	3.07E-04

	GO_ENVELOPE	1090	A multilayered structure surrounding all or part of a cell; encompasses one or more lipid bilayers, and may include a cell wall layer; also includes the space between layers.	12	0.011	1.32E-06	1.29E-20
	GO_T_TUBULE	45	Invagination of the plasma membrane of a muscle cell that extends inward from the cell surface around each myofibril. The ends of T-tubules make contact with the sarcoplasmic reticulum membrane.	4	0.0889	1.84E-06	3.88E-05
	GO_INTRINSIC_COMPONENT_OF_PLASMA_MEMBRANE	1649	The component of the plasma membrane consisting of the gene products and protein complexes having either part of their peptide sequence embedded in the hydrophobic region of the membrane or some other covalently attached group such as a GPI anchor that is similarly embedded in the membrane.	14	0.0085	3.42E-06	6.61E-05
	GO_MEMBRANE_REGION	1134	A membrane that is a part of a larger membrane. Examples include the apical region of the plasma membrane of an epithelial cell and the various regions of the endoplasmic reticulum membrane.	11	0.0097	1.24E-05	1.95E-04
	GO_MEMBRANE_PROTEIN_COMPLEX	1020	Any protein complex that is part of a membrane.	10	0.0098	2.90E-05	4.00E-04
	GO_OUTER_MEMBRANE	190	The external membrane of Gram-negative bacteria or certain organelles such as mitochondria and chloroplasts; freely permeable to most ions and metabolites.	5	0.0263	3.59E-05	4.73E-04
	GO_PLASMA_MEMBRANE_PROTEIN_COMPLEX	510	Any protein complex that is part of the plasma membrane.	7	0.0137	6.26E-05	7.71E-04

	GO_PLASMA_MEMBRANE_REGION	929	A membrane that is a (regional) part of the plasma membrane.	8	0.0086	4.52E-04	4.23E-03
	GO_TRANSPORTER_COMPLEX	321	A protein complex facilitating transport of molecules (proteins, small molecules, nucleic acids) into, out of or within a cell, or between cells.	5	0.0156	4.14E-04	3.94E-03
	GO_BRUSH_BORDER	102	The dense covering of microvilli on the apical surface of a epithelial cells in tissues such as the intestine, kidney, and choroid plexus; the microvilli aid absorption by increasing the surface area of the cell.	3	0.0294	1.05E-03	8.57E-03
	GO_PLASMA_MEMBRANE_RECEPTOR_COMPLEX	175	Any protein complex that is part of the plasma membrane and which functions as a receptor.	3	0.0171	4.85E-03	2.99E-02
	GO_RECEPTOR_COMPLEX	327	Any protein complex that undergoes combination with a hormone, neurotransmitter, drug or intracellular messenger to initiate a change in cell function.	4	0.0122	3.82E-03	2.46E-02
	GO_CATION_CHANNEL_COMPLEX	167	An ion channel complex through which cations pass.	3	0.018	4.26E-03	2.71E-02
SARCOPLASM & VACUOLES	GO_SARCOPLASM	68	The cytoplasm of a muscle cell; includes the sarcoplasmic reticulum.	9	0.1324	1.13E-14	1.10E-12
	GO_PERINUCLEAR_REGION_OF_CYTOPLASM	642	Cytoplasm situated near, or occurring around, the nucleus.	11	0.0171	4.95E-08	1.60E-06
	GO_CYTOPLASMIC_REGION	287	Any (proper) part of the cytoplasm of a single cell of sufficient size to still be considered	6	0.0209	2.12E-05	3.07E-04

			cytoplasm\				
	GO_NUCLEAR_ENVELOPE	416	The double lipid bilayer enclosing the nucleus and separating its contents from the rest of the cytoplasm; includes the intermembrane space, a gap of width 20-40 nm (also called the perinuclear space).	5	0.012	1.32E-03	9.95E-03
	GO_VACUOLAR_LUMEN	115	The volume enclosed within the vacuolar membrane.	3	0.0261	1.48E-03	1.10E-02
	GO_VACUOLE	1180	A closed structure, found only in eukaryotic cells, that is completely surrounded by unit membrane and contains liquid material. Cells contain one or several vacuoles, that may have different functions from each other. Vacuoles have a diverse array of functions. They can act as a storage organelle for nutrients or waste products, as a degradative compartment, as a cost-effective way of increasing cell size, and as a homeostatic regulator controlling both turgor pressure and pH of the cytosol.	8	0.0068	2.08E-03	1.49E-02
	GO_VACUOLAR_PART	694	Any constituent part of a vacuole, a closed structure, found only in eukaryotic cells, that is completely surrounded by unit membrane and contains liquid material.	6	0.0086	2.34E-03	1.61E-02
	GO_BLOOD_MICROPARTICLE	143	A phospholipid microvesicle that is derived from any of several cell types, such as platelets, blood cells, endothelial cells, or others, and contains membrane receptors as well as other proteins characteristic of the parental cell. Microparticles are heterogeneous in size, and are characterized as microvesicles free of nucleic acids.	3	0.021	2.76E-03	1.84E-02
	GO_MICROBO	134	Cytoplasmic organelles, spherical or oval in	3	0.0224	2.29E-03	1.60E-02

	DY		shape, that are bounded by a single membrane and contain oxidative enzymes, especially those utilizing hydrogen peroxide (H2O2).				
	GO_NUCLEAR_INNER_MEMBRANE	54	The inner, i.e. lumen-facing, lipid bilayer of the nuclear envelope.	2	0.037	4.97E-03	3.00E-02
	GO_PLATELET_ALPHA_GRANULE_LUMEN	55	The volume enclosed by the membrane of the platelet alpha granule.	2	0.0364	5.15E-03	3.08E-02
	GO_VACUOLAR_MEMBRANE	587	The lipid bilayer surrounding the vacuole and separating its contents from the cytoplasm of the cell.	5	0.0085	5.76E-03	3.37E-02
	GO_CYTOPLASMIC_VESICLE_PART	601	Any constituent part of cytoplasmic vesicle, a vesicle formed of membrane or protein, found in the cytoplasm of a cell.	5	0.0083	6.34E-03	3.65E-02
	GO_GLYCOPROTEIN_COMPLEX	21	A protein complex containing at least one glycosylated protein, may be held together by both covalent and noncovalent bonds.	2	0.0952	7.60E-04	6.58E-03
	GO_CATALYTIC_COMPLEX	1038	A protein complex which is capable of catalytic activity.	8	0.0077	9.28E-04	7.69E-03
	GO_TRANSFERASE_COMPLEX_TRANSFERRING_PHOSPHORUS_CONTAINING_GROUPS	237	A transferase complex capable of catalysis of the transfer of a phosphorus-containing group from one compound (donor) to another (acceptor).	4	0.0169	1.19E-03	9.24E-03
	GO_TRANSFERASE_COMPLEX	703	A protein complex capable of catalyzing the transfer of a group, e.g. a methyl group, glycosyl group, acyl group, phosphorus-containing, or other groups, from one compound (generally regarded as the donor) to	6	0.0085	2.49E-03	1.70E-02

			another compound (generally regarded as the acceptor).				
MITOCHONDRIA	GO_MITOCHONDRIAL_PART	953	Any constituent part of a mitochondrion, a semiautonomous, self replicating organelle that occurs in varying numbers, shapes, and sizes in the cytoplasm of virtually all eukaryotic cells. It is notably the site of tissue respiration.	13	0.0136	3.98E-08	1.44E-06
	GO_MITOCHONDRION	1633	A semiautonomous, self replicating organelle that occurs in varying numbers, shapes, and sizes in the cytoplasm of virtually all eukaryotic cells. It is notably the site of tissue respiration.	16	0.0098	9.17E-08	2.80E-06
	GO_MITOCHONDRIAL_ENVELOPE	691	The double lipid bilayer enclosing the mitochondrion and separating its contents from the cell cytoplasm; includes the intermembrane space.	8	0.0116	6.08E-05	7.66E-04
	GO_INTRINSIC_COMPONENT_OF_MITOCHONDRIAL_OUTER_MEMBRANE	22	The component of the mitochondrial outer membrane consisting of the gene products and protein complexes having either part of their peptide sequence embedded in the hydrophobic region of the membrane or some other covalently attached group such as a GPI anchor that is similarly embedded in the membrane.	2	0.0909	8.35E-04	7.12E-03
	GO_MITOCHONDRIAL_MATRIX	412	The gel-like material, with considerable fine structure, that lies in the matrix space, or lumen, of a mitochondrion. It contains the enzymes of the tricarboxylic acid cycle and, in some organisms, the enzymes concerned with fatty acid oxidation.	5	0.0121	1.27E-03	9.66E-03
	GO_INTRINSIC_COMPONENT_OF_ORGAN	281	The component of the organelle membrane consisting of the gene products and protein complexes having either part of their peptide	4	0.0142	2.22E-03	1.57E-02

	ELLE_MEMBRANE		sequence embedded in the hydrophobic region of the membrane or some other covalently attached group such as a GPI anchor that is similarly embedded in the membrane.				
	GO_INTRINSIC_COMPONENT_OF_MITOCHONDRIAL_MEMBRANE	52	The component of the mitochondrial membrane consisting of the gene products and protein complexes having either part of their peptide sequence embedded in the hydrophobic region of the membrane or some other covalently attached group such as a GPI anchor that is similarly embedded in the membrane.	2	0.0385	4.62E-03	2.88E-02
CELL-CELL JUNCTION	GO_CELL_PROJECTION_PART	946	Any constituent part of a cell projection, a prolongation or process extending from a cell, e.g. a flagellum or axon.	11	0.0116	2.25E-06	4.50E-05
	GO_CELL_LEADING_EDGE	350	The area of a motile cell closest to the direction of movement.	6	0.0171	6.38E-05	7.71E-04
	GO_SITE_OF_POLARIZED_GROWTH	149	Any part of a cell where non-isotropic growth takes place.	4	0.0268	2.09E-04	2.17E-03
	GO_LAMELLIPODIUM	172	A thin sheetlike process extended by the leading edge of a migrating cell or extending cell process; contains a dense meshwork of actin filaments.	4	0.0233	3.61E-04	3.49E-03
	GO_PSEUDOPODIUM	17	A temporary protrusion or retractile process of a cell, associated with flowing movements of the protoplasm, and serving for locomotion and feeding.	2	0.1176	4.95E-04	4.56E-03
	GO_APICAL_PART_OF_CELL	361	The region of a polarized cell that forms a tip or is distal to a base. For example, in a polarized epithelial cell, the apical region has an exposed surface and lies opposite to the basal lamina	5	0.0139	7.04E-04	6.28E-03

			that separates the epithelium from other tissue.				
	GO_CELL_PROJECTION	1786	A prolongation or process extending from a cell, e.g. a flagellum or axon.	20	0.0112	1.70E-10	8.94E-09
	GO_CELL_JUNCTION	1151	A cellular component that forms a specialized region of connection between two or more cells or between a cell and the extracellular matrix. At a cell junction, anchoring proteins extend through the plasma membrane to link cytoskeletal proteins in one cell to cytoskeletal proteins in neighboring cells or to proteins in the extracellular matrix.	13	0.0113	3.45E-07	9.10E-06
	GO_ANCHORING_JUNCTION	489	A cell junction that mechanically attaches a cell (and its cytoskeleton) to neighboring cells or to the extracellular matrix.	8	0.0164	5.16E-06	8.80E-05
	GO_CELL_SUBSTRATE_JUNCTION	398	A cell junction that forms a connection between a cell and the extracellular matrix.	5	0.0126	1.09E-03	8.75E-03
	GO_INTERCALATED_DISC	51	A complex cell-cell junction at which myofibrils terminate in cardiomyocytes; mediates mechanical and electrochemical integration between individual cardiomyocytes. The intercalated disc contains regions of tight mechanical attachment (fasciae adherentes and desmosomes) and electrical coupling (gap junctions) between adjacent cells.	2	0.0392	4.45E-03	2.80E-02
	GO_CELL_CELL_ADHERENS_JUNCTION	54	An adherens junction which connects a cell to another cell.	2	0.037	4.97E-03	3.00E-02
	GO_RUFFLE	156	Projection at the leading edge of a crawling cell; the protrusions are supported by a microfilament meshwork.	3	0.0192	3.52E-03	2.29E-02

Molecular Functions							
Compiled Molecular Functions	**Gene Set Name**	**# Genes in Gene Set (K)**	**Description**	**# Genes in Overlap (k)**	**k/K**	**p-value**	**FDR q-value**
SARCOMERE & CYTOSKELETAL FUNCTIONS	GO_CYTOSKELETAL_PROTEIN_BINDING	819	Interacting selectively and non-covalently with any protein component of any cytoskeleton (actin, microtubule, or intermediate filament cytoskeleton).	22	0.0269	3.29E-19	2.97E-16
	GO_ACTIN_BINDING	393	Interacting selectively and non-covalently with monomeric or multimeric forms of actin, including actin filaments.	14	0.0356	3.79E-14	1.71E-11
	GO_IDENTICAL_PROTEIN_BINDING	1209	Interacting selectively and non-covalently with an identical protein or proteins.	20	0.0165	1.54E-13	4.61E-11
	GO_STRUCTURAL_CONSTITUENT_OF_MUSCLE	41	The action of a molecule that contributes to the structural integrity of a muscle fiber.	7	0.1707	1.71E-12	3.85E-10
	GO_CALMODULIN_BINDING	179	Interacting selectively and non-covalently with calmodulin, a calcium-binding protein with many roles, both in the calcium-bound and calcium-free states.	10	0.0559	2.43E-12	4.38E-10
	GO_STRUCTURAL_MOLECULE_ACTIVITY	732	The action of a molecule that contributes to the structural integrity of a complex or assembly within or outside a cell.	13	0.0178	1.77E-09	2.66E-07
	GO_MICROFILAMENT_MOTOR_ACTIVITY	21	Catalysis of movement along a microfilament, coupled to the hydrolysis of a nucleoside triphosphate (usually ATP).	4	0.1905	7.67E-08	7.68E-06
	GO_ENZYME_BINDING	1737	Interacting selectively and non-covalently with any enzyme.	15	0.0086	1.19E-06	9.72E-05

	GO_PROTEASE_BINDING	104	Interacting selectively and non-covalently with any protease or peptidase.	5	0.0481	1.92E-06	1.44E-04
	GO_ATPASE_ACTIVITY_COUPLED	313	Catalysis of the reaction: ATP + H2O = ADP + phosphate; this reaction directly drives some other reaction, for example ion transport across a membrane.	7	0.0224	2.72E-06	1.88E-04
	GO_ATPASE_ACTIVITY	427	Catalysis of the reaction: ATP + H2O = ADP + phosphate + 2 H+. May or may not be coupled to another reaction.	7	0.0164	2.04E-05	1.15E-03
	GO_ACTIN_FILAMENT_BINDING	121	Interacting selectively and non-covalently with an actin filament, also known as F-actin, a helical filamentous polymer of globular G-actin subunits.	4	0.0331	9.39E-05	3.53E-03
	GO_MOTOR_ACTIVITY	131	Catalysis of the generation of force resulting either in movement along a microfilament or microtubule, or in torque resulting in membrane scission, coupled to the hydrolysis of a nucleoside triphosphate.	4	0.0305	1.28E-04	4.60E-03
	GO_MYOSIN_BINDING	59	Interacting selectively and non-covalently with any part of a myosin complex; myosins are any of a superfamily of molecular motor proteins that bind to actin and use the energy of ATP hydrolysis to generate force and movement along actin filaments.	3	0.0508	2.11E-04	6.18E-03
	GO_ACTIN_DEPENDENT_ATPASE_ACTIVITY	13	Catalysis of the reaction: ATP + H2O = ADP + phosphate. This reaction requires the presence of an actin filament to accelerate release of ADP and phosphate.	2	0.1538	2.85E-04	8.03E-03
	GO_TITIN_BINDING	14	Interacting selectively and non-covalently with titin, any of a family of giant proteins found in striated and smooth muscle. In striated muscle, single titin molecules span half the sarcomere,	2	0.1429	3.32E-04	8.81E-03

			with their N- and C-termini in the Z-disc and M-line, respectively.				
	GO_CALCIUM_ION_BINDING	697	Interacting selectively and non-covalently with calcium ions (Ca2+).	8	0.0115	6.45E-05	2.64E-03
	GO_PROTEIN_DIMERIZATION_ACTIVITY	1149	The formation of a protein dimer, a macromolecular structure consists of two noncovalently associated identical or nonidentical subunits.	10	0.0087	7.82E-05	3.06E-03
	GO_KINASE_BINDING	606	Interacting selectively and non-covalently with a kinase, any enzyme that catalyzes the transfer of a phosphate group.	7	0.0116	1.81E-04	6.05E-03
	GO_INTRAMOLECULAR_TRANSFERASE_ACTIVITY_PHOSPHOTRANSFERASES	11	Catalysis of the transfer of a phosphate group from one position to another within a single molecule.	2	0.1818	2.02E-04	6.18E-03
	GO_PROTEIN_SELF_ASSOCIATION	44	Interacting selectively and non-covalently with a domain within the same polypeptide.	2	0.0455	3.33E-03	5.00E-02
	GO_WW_DOMAIN_BINDING	31	Interacting selectively and non-covalently with a WW domain of a protein, a small module composed of 40 amino acids and plays a role in mediating protein-protein interactions via proline-rich regions.	2	0.0645	1.66E-03	2.67E-02
	GO_CALMODULIN_DEPENDENT_PROTEIN_KINASE_ACTIVITY	28	Catalysis of the reactions: ATP + a protein serine = ADP + protein serine phosphate; and ATP + a protein threonine = ADP + protein threonine phosphate. These reactions require the presence of calcium-bound calmodulin.	2	0.0714	1.36E-03	2.29E-02
	GO_MOLECUL	1353	A molecular function that modulates the	9	0.0067	1.25E-03	2.24E-02

	AR_FUNCTION_REGULATOR		activity of a gene product or complex. Examples include enzyme regulators and channel regulators.				
ALPHA-ACTININ BINDING	GO_ALPHA_ACTININ_BINDING	21	Interacting selectively and non-covalently with alpha-actinin, one of a family of proteins that cross-link F-actin as antiparallel homodimers. Alpha-actinin has a molecular mass of 93-103 KDa; at the N-terminus there are two calponin homology domains, at the C-terminus there are two EF-hands. These two domains are connected by the rod domain. This domain is formed by triple-helical spectrin repeats.	3	0.1429	9.10E-06	5.86E-04
	GO_ACTININ_BINDING	29	Interacting selectively and non-covalently with actinin, any member of a family of proteins that crosslink F-actin.	3	0.1034	2.47E-05	1.24E-03
GLYCOSYLATION	GO_CARBOHYDRATE_BINDING	277	Interacting selectively and non-covalently with any carbohydrate, which includes monosaccharides, oligosaccharides and polysaccharides as well as substances derived from monosaccharides by reduction of the carbonyl group (alditols), by oxidation of one or more hydroxy groups to afford the corresponding aldehydes, ketones, or carboxylic acids, or by replacement of one or more hydroxy group(s) by a hydrogen atom. Cyclitols are generally not regarded as carbohydrates.	5	0.0181	2.11E-04	6.18E-03
	GO_MONOSACCHARIDE_BINDING	70	Interacting selectively and non-covalently with any monosaccharide. Monosaccharides are the simplest carbohydrates; they are polyhydroxy aldehydes H[CH(OH)]nC(=O)H or polyhydroxy ketones H[CHOH]nC(=O)[CHOH]mH with three or more carbon atoms. They form the	3	0.0429	3.50E-04	9.01E-03

			constitutional repeating units of oligo- and polysaccharides.				
	GO_ENZYME_REGULATOR_ACTIVITY	959	Binds to and modulates the activity of an enzyme.	8	0.0083	5.56E-04	1.19E-02
	GO_HYDROLASE_ACTIVITY_ACTING_ON_ESTER_BONDS	739	Catalysis of the hydrolysis of any ester bond.	7	0.0095	5.95E-04	1.22E-02
	GO_TRANSFERASE_ACTIVITY_TRANSFERRING_HEXOSYL_GROUPS	203	Catalysis of the transfer of a hexosyl group from one compound (donor) to another (acceptor).	4	0.0197	6.73E-04	1.32E-02
	GO_CARBOHYDRATE_KINASE_ACTIVITY	20	Catalysis of the transfer of a phosphate group, usually from ATP, to a carbohydrate substrate molecule.	2	0.1	6.89E-04	1.32E-02
	GO_PHOSPHORIC_ESTER_HYDROLASE_ACTIVITY	368	Catalysis of the reaction: RPO-R' + H2O = RPOOH + R'H. This reaction is the hydrolysis of any phosphoric ester bond, any ester formed from orthophosphoric acid, O=P(OH)3.	5	0.0136	7.67E-04	1.44E-02
	GO_MOLECULAR_FUNCTION_REGULATOR	1353	A molecular function that modulates the activity of a gene product or complex. Examples include enzyme regulators and channel regulators.	9	0.0067	1.25E-03	2.24E-02
	GO_INTRAMOLECULAR_TRANSFERASE_ACTIVITY	27	Catalysis of the transfer of a functional group from one position to another within a single molecule.	2	0.0741	1.26E-03	2.24E-02
	GO_KINASE_ACTIVITY	842	Catalysis of the transfer of a phosphate group, usually from ATP, to a substrate molecule.	7	0.0083	1.27E-03	2.24E-02
	GO_TRANSFE	282	Catalysis of the transfer of a glycosyl group	4	0.0142	2.25E-03	3.50E-02

	RASE_ACTIVITY_TRANSFERRING_GLYCOSYL_GROUPS		from one compound (donor) to another (acceptor).				
	GO_COFACTOR_BINDING	263	Interacting selectively and non-covalently with a cofactor, a substance that is required for the activity of an enzyme or other protein. Cofactors may be inorganic, such as the metal atoms zinc, iron, and copper in certain forms, or organic, in which case they are referred to as coenzymes. Cofactors may either be bound tightly to active sites or bind loosely with the substrate.	6	0.0228	1.30E-05	7.78E-04
	GO_MACROMOLECULAR_COMPLEX_BINDING	1399	Interacting selectively and non-covalently with any macromolecular complex.	10	0.0071	3.81E-04	9.53E-03
MITOCHINDRIAL ENERGY METABOLISM	GO_RIBONUCLEOTIDE_BINDING	1860	Interacting selectively and non-covalently with a ribonucleotide, any compound consisting of a ribonucleoside that is esterified with (ortho)phosphate or an oligophosphate at any hydroxyl group on the ribose moiety.	19	0.0102	2.37E-09	3.06E-07
	GO_ADENYL_NUCLEOTIDE_BINDING	1514	Interacting selectively and non-covalently with adenyl nucleotides, any compound consisting of adenosine esterified with (ortho)phosphate.	16	0.0106	3.26E-08	3.67E-06
	GO_HYDROLASE_ACTIVITY_ACTING_ON_ACID_ANHYDRIDES	820	Catalysis of the hydrolysis of any acid anhydride.	11	0.0134	5.64E-07	5.08E-05
	GO_FLAVIN_ADENINE_DINU	74	Interacting selectively and non-covalently with FAD, flavin-adenine dinucleotide, the coenzyme	3	0.0405	4.12E-04	1.00E-02

	CLEOTIDE_BINDING		or the prosthetic group of various flavoprotein oxidoreductase enzymes, in either the oxidized form, FAD, or the reduced form, FADH2.				
	GO_PROTEIN_COMPLEX_BINDING	935	Interacting selectively and non-covalently with any protein complex (a complex of two or more proteins that may include other nonprotein molecules).	8	0.0086	4.71E-04	1.12E-02
	GO_ACYL_COA_DEHYDROGENASE_ACTIVITY	17	Catalysis of the reaction: acyl-CoA + acceptor = 2,3-dehydroacyl-CoA + reduced acceptor.	2	0.1176	4.95E-04	1.14E-02
	GO_PROTEIN_HOMODIMERIZATION_ACTIVITY	722	Interacting selectively and non-covalently with an identical protein to form a homodimer.	7	0.0097	5.19E-04	1.15E-02
	GO_ELECTRON_CARRIER_ACTIVITY	112	Any molecular entity that serves as an electron acceptor and electron donor in an electron transport chain. An electron transport chain is a process in which a series of electron carriers operate together to transfer electrons from donors to any of several different terminal electron acceptors to generate a transmembrane electrochemical gradient.	3	0.0268	1.37E-03	2.29E-02
	GO_FATTY_ACYL_COA_BINDING	31	Interacting selectively and non-covalently with acyl-CoA, any derivative of coenzyme A in which the sulfhydryl group is in thiolester linkage with a fatty acyl group.	2	0.0645	1.66E-03	2.67E-02
TRANSMEMBRANE TRANSPORT, CELL-CELL ANCHORING	GO_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	997	Enables the transfer of a substance from one side of a membrane to the other.	10	0.01	2.39E-05	1.24E-03

	GO_PASSIVE_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	464	Enables the transfer of a solute from one side of the membrane to the other, down the solute's concentration gradient.	7	0.0151	3.45E-05	1.64E-03
	GO_TRANSPORTER_ACTIVITY	1276	Enables the directed movement of substances (such as macromolecules, small molecules, ions) into, out of or within a cell, or between cells.	11	0.0086	3.66E-05	1.65E-03
	GO_GATED_CHANNEL_ACTIVITY	325	Enables the transmembrane transfer of a solute by a channel that opens in response to a specific stimulus.	6	0.0185	4.24E-05	1.82E-03
	GO_TRANSFERASE_ACTIVITY_TRANSFERRING_PHOSPHORUS_CONTAINING_GROUPS	992	Catalysis of the transfer of a phosphorus-containing group from one compound (donor) to another (acceptor).	9	0.0091	1.33E-04	4.62E-03
	GO_CATION_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	622	Enables the transfer of cation from one side of the membrane to the other.	7	0.0113	2.13E-04	6.18E-03
	GO_CATION_CHANNEL_ACTIVITY	298	Enables the energy-independent passage of cations across a lipid bilayer down a concentration gradient.	5	0.0168	2.95E-04	8.05E-03
	GO_VOLTAGE_GATED_ION_CHANNEL_ACTIVITY	190	Enables the transmembrane transfer of an ion by a voltage-gated channel. An ion is an atom or group of atoms carrying an electric charge by virtue of having gained or lost one or more electrons. A voltage-gated channel is a channel	4	0.0211	5.25E-04	1.15E-02

			whose open state is dependent on the voltage across the membrane in which it is embedded.				
	GO_INORGANIC_CATION_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	527	Enables the transfer of inorganic cations from one side of a membrane to the other. Inorganic cations are atoms or small molecules with a positive charge that do not contain carbon in covalent linkage.	6	0.0114	5.75E-04	1.20E-02
	GO_ANKYRIN_BINDING	20	Interacting selectively and non-covalently with ankyrin, a 200 kDa cytoskeletal protein that attaches other cytoskeletal proteins to integral membrane proteins.	2	0.1	6.89E-04	1.32E-02
	GO_CALCIUM_ION_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	128	Enables the transfer of calcium (Ca) ions from one side of a membrane to the other.	3	0.0234	2.01E-03	3.18E-02
	GO_METAL_ION_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	417	Enables the transfer of metal ions from one side of a membrane to the other.	5	0.012	1.34E-03	2.29E-02
	GO_LIGAND_GATED_CHANNEL_ACTIVITY	142	Enables the transmembrane transfer of a solute by a channel that opens when a specific ligand has been bound by the channel complex or one of its constituent parts.	3	0.0211	2.70E-03	4.12E-02
	GO_MOLECULAR_FUNCTION_REGULATOR	1353	A molecular function that modulates the activity of a gene product or complex. Examples include enzyme regulators and channel regulators.	9	0.0067	1.25E-03	2.24E-02

43

44

45

SUPPLEMENTARY METHODS

Skeletal Muscle Biopsy and Histochemistry

A single (2.0 x 0.75 x 0.75 cm) piece of tissue from the vastus lateralis of the proband was obtained under local anesthesia via needle biopsy after taking written informed consent approved by the Institutional Review Board for Human Subject Research at Emory University School of Medicine. The biopsy measuring 2.0 x 0.75 x 0.75 cm was wrapped in saline moistened gauze and immediately dissected for frozen and paraffin sections used for immunohistochemistry analysis and staining which included H&E, modified Gomori Trichrome, myosin ATPase (pH 10.4 and 4.2), NADH, SDH, acid phosphatase, cytochrome oxidase, Oil Red O, and PAS. A piece embedded in paraffin for cross section and longitudinal section was stained with H&E, Congo red and Trichrome. After quick dissection, the tissue was immediately flash frozen in isopentane precooled in liquid nitrogen at −80°C.

Exome Sequencing and Analysis

Peripheral blood was collected into EDTA tubes from the patient proband and his parents. Genomic DNA was extracted using GenElute™ Blood Genomic DNA (Sigma-Aldrich NA2020) according to the manufacturer's protocol. Exome Sequencing (ES) was performed on genomic DNA using the NimbleGen (Madison, WI) v3.0 targeted sequence capture method to enrich for the exome. These targeted regions were then sequenced using the Illumina (San Diego, CA) HiSeq 2000 sequencing system with 100 base pair (bp) paired-end reads at an average coverage of 100X in the target region. The targeted regions included the exon and 10 bp of flanking intronic sequence. In general, ES assays performed at Emory Genetics Laboratory (EGL) have an overall coverage of 92.9%, with as high as 94.8% coverage in the coding region.

Bioinformatics Analysis and Variant Classification: Bioinformatic analysis was performed using NextGENe software from SoftGenetics (State College, PA). The NextGENe output was customized to mine variants from EGL's internal variant database EmVClass [1] other public databases, and variant prediction tools, such as SIFT [2] and PolyPhen [3]. All variants detected were classified using population frequency data available from the Exome Variant Server (http://evs.gs.washington.edu/EVS/) and previous reports of disease association and pathogenicity available through the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php), National Center for Biotechnology Information PubMed (http://www.ncbi.nlm.nih.gov/pubmed/), and Google (https://www.google.com/). Variant classification was performed based on American College of Medical Genetics and Genomics- Association for Molecular Pathology (ACMG-AMP) guidelines [4]. A detailed overview of the bioinformatics pipeline and variant annotation protocol is described elsewhere [5]. All variants were curated and reviewed by board-certified laboratory directors and maintained as an in-house variant database; they were made publicly accessible via EGL's online tool, EmVClass[1].

Sanger Confirmation of Genomic DNA: PCR-amplified genomic DNA and RT-PCR amplified cDNA products of the *GNE* gene were Sanger sequenced for confirmation of the variants. An automated primer design script, developed and validated in-house, was used to design primers[6].

Next Generation Sequencing (NGS)-based RNA-sequencing

High quality (RNA Integrity Number; RIN>7) RNA was extracted from right thigh muscle biopsies of the patient and six control individuals using Aurum™ Total RNA Fatty and Fibrous Tissue Kit

(cat# 7326870) following the manufacturer's protocol. Library preparation was performed using Agilent Sure Select XT RNA Reagent kit (cat# G9692B) following manufacturer's protocol. Patient samples in replicate and six control normal individual muscle biopsy mRNA samples were sequenced in the Georgia Tech Molecular Evolution Core on an Illumina Next Seq instrument to obtain high output paired-end by 150bp reads at a depth of more than 150 million reads. From the whole transcriptome sequence data, we focused our analysis on 274 genes (see Table S3) that are known to be NMD-associated and are known to have skeletal muscle expression as retrieved from GTEX portal[7]. These 274 genes were curated initially based on the associations to NMDs listed in http://www.muscle.ca/wp-content/uploads/2012/10/Disorder_List_ENG_May2017.pdf.

Bioinformatics workflow

Raw sequencing reads were checked for quality using FastQC[8]. Reads were not trimmed beyond removal of adapter sequences [9]. Human reference genome GRCh38 (NCBI) and NCBI *Homo sapiens* Annotation Release 106 were obtained from Illumina iGenomes and sequenced reads were aligned using the splice-aware alignment program STAR version 2.5.2b in 2-pass mode [10].

Quality metrics for all samples were obtained by running Picard RNASeqMetrics (*Available online at: http://broadinstitute.github.io/picard)* and principal component analysis (PCA) on gene expression was performed to check for outlier status based on tissue composition or contamination. Initial unsupervised clustering of samples by relative gene expression levels clearly differentiated the patient sample from six controls.

For analysis of the variant of interest, alignment files were loaded into Integrated Genomics Viewer (IGV) [11] to confirm positive identification of variants found in ES and analyze the

transcript structure of the candidate gene for potential variant effects such as abnormal splicing or allele specific expression.

To identify possible downstream effects, differential gene expression analysis, including both normalization to account for sequencing depth and RNA composition as well as identification of surrogate variables to correct for un-modeled variation, was performed using the R packages DESeq2 version 1.16.1 [12] and SVA version 3.24.3 [13]. Read counts mapped per gene were obtained using HTSeq-Count [14] and the analysis was conducted following the Bioconductor DESeq2 vignette and the RNA-Seq Gene Expression Analysis workflow [15]. To maintain consistency, the same annotation file used for alignment was used for all downstream analysis. The result file was filtered to include only significantly up or down regulated genes using Benjamini-Hochberg adjusted[16] $p<0.05$, from our curated panel of 274 NMD-associated genes.

Array CGH

Array design: The targeted gene high-resolution oligonucleotide CGH array was custom designed on Oxford Gene Technologies (OGT) 180 K platform to detect deletions and duplications. Long oligonucleotides (~45–60 mer) were used to design the array, with repeat sequence masking implemented to ensure greater sensitivity and specificity. The *GNE* gene and its upstream was covered by 720,000 probes including covering the *GNE* 13 exons at an average spacing of 15 bp between probes and the intronic region was covered at an average spacing of 25 bp between probes. Use of intronic oligonucleotide probes allows detection of dosage changes within the entire genomic region of the gene and determination on the approximate breakpoints.

aCGH Experimental set up: DNA extraction was performed on patient DNA using a Gentra Puregene DNA extraction kit according to the manufacturer's instructions. Male and female wild-type control DNA was obtained from Promega, Inc. Each patient and reference DNA sample

was sonicated, such that the DNA fragment size was between 200-5,000 bases and verified on a 1% agarose gel. Patient and reference DNA samples were labeled using Klenow enzyme (NEB) and Cy3 or Cy5 9mer wobble primers (TriLink Technologies), respectively. After labeling, each sample was purified by isopropanol precipitation and reconstituted in ultra-pure water. 4 ug each of labeled patient and reference DNA were combined. The products were desiccated in a vacufuge (Savant DNA 120), and resuspended in appropriate hybridization buffer along with Cy3 and Cy5 control CPK6 50mer oligonucleotides. This mixture was hybridized to a NimbleGen targeted gene CGH array for 16-20 hours at 42 °C in a Maui Hybridization system (BioMicro Systems). The custom array allows for detection of copy number changes and breakpoints as small as 100-bp within the entire coding region. Arrays were then washed according to the manufacturer's recommendation and immediately scanned on a GenePix 4000 scanner (Molecular Devices). After scanning, data were extracted from images, and within-array normalization was accomplished using manufacturer-provided software (NimbleScan). Normalized log(2) ratio data were analyzed using two different analysis programs: SegMNT and DNA copy NimbleScan (NimbleGen Systems, Inc.). Both software programs report breakpoints for predicted deletions or duplications in the patient or test sample relative to the reference and also display results in a bar graph where the y-axis indicates gain or loss of material (1 = gain, 0 = normal, -1 = loss), while the x-axis indicates the position of each feature on the chromosome. Data files (.GFF) generated from different averaging windows using NimbleScan software were parsed using a custom program (Nimkit) that was developed in-house. Nimkit enables the laboratory to select and analyze only the *GNE* locus. Nimkit generates a gene-specific report summarizing breakpoints detected in the gene of interest and its up or downstream, the respective log(2) ratios, and the exons present at each region. All other regions are masked and not analyzed by Nimkit, preventing genetic analysis of genes for which clinical testing was not requested, in compliance with HIPAA requirements.

Array quality was assessed by control resequencing oligonucleotides on each array that correspond to synthetic sequences designed to have no cross-hybridization potential to any known sequence. This sequence was designed to have three distinct sequencing domains with different characteristics: A, B, and C domains. Resequencing was performed on both the forward and reverse strands, so that the resequencing report has six different scores for the Cy3 channel and six distinct scores for the Cy5 channel: A-forward and A-reverse, B-forward and B-reverse, C-forward and C-reverse. The "A" domain contained long runs of G nucleotides that can be difficult to synthesize. The "B" domain contained a large perfect hairpin sequence. The "C" domain contained a straightforward domain that should always resequence. Failure of domain "C" indicated a catastrophic failure. Control DNA was spiked into each experiment for both CGH and resequencing arrays. A score from 0-100% was obtained that indicating sequence fidelity and correlated well with the overall performance of a microarray experiment [17].

Mapping HIBM mutations and molecular dynamics simulation onto the N-acetylmannosamine kinase domain of GNE

Sixteen point substitutions (LEU-36; PHE-200; ASN-225; GLN-246; VAL-303; TYR-378; VAL-460; CYS-528; THR-557; LEU-572; GLU-576; THR-587; THR-631; VAL-631; MET-727 and THR-712) that cause HIBM listed in UNIPROT database entry Q9Y223 were mapped onto the crystal structure of N-acetylmannosamine kinase from the protein data bank (PDB ID: 2YHY) using PyMOL Valine-727, which is numbered 696 in the crystal structure, was mutated *in silico* to methionine using PyMOL. The rotamer was not altered and the protein was not energy minimized. YASARA was used to generate the dimer of N-acetylmannosamine kinase (PDB ID: 2YHY). It was also used to generate missing sidechains and an internal loop. The sugar and ADP were removed from the PDB file. Molecular dynamics simulations were conducted in GROMACS using the AMBER09 force-field. The simulation was run for one nanosecond.

Western Blot testing alpha-Dystroglycan Glycosylation

One hundred milligrams of normal human controls and the GNE mutant patient muscle biopsies were minced with a scalpel on ice and placed in a 1.5 mL screw cap tube and solubilized in 1 ml of Tris-buffered saline (TBS) containing 1% Triton X-100 with protease inhibitors and a portion of zinc oxide beads. The samples were run through two rounds on the Next Advance bullet blender at setting 7 for 3 minutes. The samples were rotated for 45 min at 4C followed by centrifugation for 10 minutes. The sample solubilized supernatant was added to 200uL of WGA (wheat germ agglutinin) slurry (Vector Labs), rotating overnight at 4C. The pelleted WGA beads were washed 3 times with 1 ml TBS containing 0.1% Triton X-100. After the final wash, 250 uL loading buffer (Loading Sample buffer (LSB)) was added to the beads. The samples were heated to 99˚C for 10 minutes before loading into 3–15% SDS–PAGE gels and transferred to polyvinylidene fluoride (PVDF) membranes and probed with dystroglycan antibodies. IIH6 antibody against the laminin-binding glycoepitope of glycosylated α-dystroglycan, and AF6868 antibody against core α-dystroglycan and β-dystroglycan were used, described previously for specificity. Densitometry quantification of the band intensities in the western blot images was performed using LI-COR Image Studio Lite software (https://www.licor.com/bio/products/software/image_studio_lite/)

Gene Ontology-Pathway Analysis: We performed a tiered approach to investigate the genetic and pathway networks affected by genes that are statistically significantly differentially regulated in the proband muscle biopsy compared to the six control biopsies (Table S4). We performed a gene-ontology-based analysis with MSigDB (Molecular Signature Database) [18,19] that curates gene ontologies (GOs), biological processes or pathways, molecular functions, and cellular compartments separately that are significantly associated with the gene clustering ($p<0.01$;

FDR<0.05) based on greater than 1325 biologically defined gene sets (see Table S5 for raw statistical values of significant clustered gene sets), similar to our recent study on infantile spasm[20]. We retrieved the top 100 enriched gene set pathways/cellular compartment/molecular functions that are affected in the patient muscle. The individual gene sets were then manually collapsed with biological evidence (see Table S5 for compiled functions and corresponding individual gene set categories). The criteria behind manual compiling are based on a) established hierarchical superfamily of the GO functions [21] and external links with ontology and hierarchy for non-GO gene sets from MSigDB database, and b) biological similarity of the individual functions and pathways (eg. “HALLMARK_MYOGENESIS” and “GO_MUSCLE_CONTRACTION” were compiled into “Muscle Development and Contraction"). The compilation criterion was consistently followed and performed. The value k is the number of significant genes from the patient muscle that are differentially regulated and found in a particular enriched gene set, and value K is the total number of significant genes in a gene set. The ratio k/K is the proportion of significant genes in the patient muscle found in an enriched gene set.

SUPPLEMENTARY REFERENCES

1. Bean LJ, Tinker SW, da Silva C, Hegde MR. Free the data: one laboratory's approach to knowledge-based genomic variant classification and preparation for EMR integration of genomic data. Hum Mutat 2013;34(9):1183-1188.

2. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31(13):3812-3814.

3. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods 2010;7(4):248-249.

4. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, Committee ALQA. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine : official journal of the American College of Medical Genetics 2015;17(5):405-424.

5. Chin EL, da Silva C, Hegde M. Assessment of clinical analytical sensitivity and specificity of next-generation sequencing for detection of simple and complex mutations. BMC Genet 2013;14:6.

6. Jones MA, Bhide S, Chin E, Ng BG, Rhodenizer D, Zhang VW, Sun JJ, Tanner A, Freeze HH, Hegde MR. Targeted polymerase chain reaction-based enrichment and next generation sequencing for diagnostic testing of congenital disorders of glycosylation. Genetics in medicine : official journal of the American College of Medical Genetics 2011;13(11):921-932.

7. Consortium GT. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45(6):580-585.

8. Wingett SW, Andrews S. FastQ Screen: A tool for multi-genome mapping and quality control. F1000Res 2018;7:1338.

9. Williams CR, Baccarella A, Parrish JZ, Kim CC. Trimming of sequence reads alters RNA-Seq gene expression estimates. BMC Bioinformatics 2016;17(1):103.

10. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29(1):15-21.
11. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotech 2011;29(1):24-26.
12. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 2014;15(12):550.
13. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28(6):882-883.
14. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31(2):166-169.
15. Love MI, Anders S, Kim V, Huber W. RNA-Seq workflow: gene-level exploratory analysis and differential expression. F1000Research 2015;4:1070.
16. Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics 2003;19(3):368-375.
17. Hegde MR, Chin EL, Mulle JG, Okou DT, Warren ST, Zwick ME. Microarray-based mutation detection in the dystrophin gene. Hum Mutat 2008;29(9):1091-1099.
18. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 2005;102(43):15545-15550.
19. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34(3):267-273.

20. Berg AT, Chakravorty S, Koh S, Grinspan ZM, Shellhaas RA, Saneto RP, Wirrell EC, Coryell J, Chu CJ, Mytinger JR, Gaillard WD, Valencia I, Knupp KG, Loddenkemper T, Sullivan JE, Poduri A, Millichap JJ, Keator C, Wusthoff C, Ryan N, Dobyns WB, Hegde M. Why West? Comparisons of clinical, genetic and molecular features of infants with and without spasms. PLoS One 2018;13(3):e0193599.

21. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25(1):25-29.