nature portfolio

Peer Review File

Human Skeletal Muscle Fiber Heterogeneity Beyond Myosin Heavy Chains

Corresponding Author: Professor Atul Deshmukh

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The article by Roger Moreno-Justicia is a very interesting paper on the heterogeneity of muscle fibers in the vastus lateralis (VL), a human thigh muscle in healthy and in Nemaline myopathy patients. Much state-of-the-art work has been done with human samples, for which single muscle fibers were manually isolated.
The authors compare high-throughput transcriptomic and proteomic data on genes/proteins co-expressed in muscle fibers. Using numerous bioinformatic approaches, they show that there are only two main fiber types in the VL, a slow type and a fast type, and that the expression of genes encoding myosin heavy chains (Myh) are not the main determinants in distinguishing the different fiber types, but that there is a continuum in the expression of a certain number of genes, in particular those encoding ribosomal proteins and those involved in cell junctions between muscle fibers. In Nemaline myopathy samples, the authors reveal a shift towards less oxidative fibers independent of MYH-based fiber type.
Unfortunately, apart from the methods used to analyze the data, the authors did not validate their results with basic experiments (immunocytochemistry for example) that would have provided a rigorous demonstration of their findings that lack statistical values.

Main comment :
I am a bit confused by the presence of pure 2X fibers at the RNA level, suggesting that MYH2 and MYH7 are off, while fibers expressing 2X protein are only hybrid fibers that also express 2A. Are the authors suggesting that the 2A gene could be on/off depending on e.g. day/night, and while the 2A protein is stable, its mRNA has a short half-life? This should be discussed as to how a fiber that does not express the gene accumulates the corresponding protein.

It should be stated in the introduction that the human vastus lateralis is known to have a low percentage of 2X fibers compared to other muscles with a higher percentage of 2X positive fibers, or that all human muscles contain very few 2X fibers and VL is thus a good archetype of human muscles; the results should be discussed in light of this knowledge. Furthermore, I suspect that the percentage of oxidative/glycolytic positive myofibers may vary depending on the depth at which the biopsy was taken in the VL.

In their sentence lines 179-183, do the authors mean that the expression of MYH1 and genes/proteins co-expressed with MYH1 is not discriminatory? This needs to be better explained. See for example the genes/proteins co-expressed with Myh 2X in the manuscripts of Murgia et al.

In the chapter starting on line 187, the authors establish transcriptomic and proteomic correlations with several genes. Heatmaps of co-expressed genes can be shown for clarity. The heat map allows the unbiased identification of the genes co-expressed with the better correlation coefficient. The idea that there is a large heterogeneity of genes expressed with Myh7 or with Myh2 or with Myh1 has been well illustrated in several mouse snRNA-seq publications in 2020. This heterogeneity illustrates well that there is not a unique transcriptomic/proteomic signature of Myh7- or Myh2-positive myofibers, but that there is a large variability of co-expressed genes depending on the muscle analyzed.

Line 240-241, I find it a bit provocative to state that "no distinct subtypes of human skeletal muscle fibers beyond a slow-fast classification should be considered" with the study of a single limb muscle, although I agree with the authors that their data show that the slow/fast fiber type-independent heterogeneity across fibers is of a "continuous nature".

Line 335; SREBPF1 should be replaced with SREBF1. It is difficult to understand the significance of the histograms in Figure 4S1 D-K. Is there a way to get statistical values for e.g. SREBF1 enrichment in fast versus slow myofibers? Otherwise, without statistical values, the histograms do not improve the manuscript. Otherwise, explain better how these data were generated and why no statistical value can be provided.

The paragraph from lines 351-362 should be written with more explanation. How were the binding sites for specific transcription factors detected in the list of slow/fast enriched genes? In the promoter regions of selected genes? In their 100kb vicinity? The legend of Figure 4E is not very informative. If AP2 (TFAP2A, AP2, AP2a and AP2g, Figure 4E) and Tieg1/Klf10 are the putative drivers of fast and slow myofiber fate, respectively, immunocytochemistry with corresponding antibodies must be performed to confirm the prediction, or at least their expression profile must be provided in feature plots. The enrichment of MEF2 in fast fibers is counterintuitive since it has been shown to control the slow/oxidative phenotype, at least in the mouse. MEF2 enrichment is not shown in Figure 4E.
How many fibers were counted for the representative 4G image?
It is important to include the common name for the TF indicated (Klf10 for Tieg1, AP2alpha for TFAP2A, TEAD2 for ETF, ZNF263 for FPM-315). Note that in the mouse, TEAD proteins, like MEF2 proteins, are known to control the slow/oxidative phenotype. This should be discussed in the context of the muscle analyzed here, which has very few Myh2X myofibers, and the results presented or immunocytochemistry provided showing the accumulation of these TFs preferentially in fast nuclei. This should also be discussed in the light of previously published proteomic and transcriptomic data on single human myofibers from different muscles where MyH2X myofibers are identified.
In Figure 4I (and more generally in each histogram) each point should be represented.
Figure 5S2J. I am a bit surprised by the fold change in expression of ANXA1, A2, A5 and A6 between control and mutant samples. For example, for ANXA1, the LFQ (label-free quantitation, please specify in your paper) intensity is about 12 for the control and about 14 for the Acta1-NM, which means an increased expression of 1.2-fold, is this correct? Is this difference significant? Are the authors convinced that this 1.2 fold difference has a physiological relevance?
The TPPP3 (Tubulin Polymerization Promoting Protein Family Member 3) seems to be more differentially expressed between the samples, with a good p-value. Is tubulin polymerization increased in nemaline myopathies? The position of the S100A11 protein is not shown in the Volcano plots.

Please modify the title of the Supplementary tables according to their number. 488870_0_data_set_4288201_s4txxx is for example is table 27 488870_0_data_set_4288185_s4txxw is table 11.
Not easy presently to identify the tables with their name. The txxx could be modified for example by the number of the table (t1, t2..), it would thus be more easy to find the information.

In the Discussion section, I am not convinced by paragraph lines 519-530. The authors use a muscle with a very low Myh2X content. It is therefore difficult to write: "The existence of pure type 2X fibers should be questioned...". It would be more rigorous to write "....questioned in the vastus lateralis studied here". It is known that the percentage of 1/2A/2X and mixed myofibers varies depending on the muscle studied. This percentage also varies during development and aging. A reference should be made to whether 2X myofibres are very rare in all human muscles.

Reviewer #2

(Remarks to the Author)
The manuscript from Roger Moreno-Justicia and colloquies measured a tremendous amount of human single muscle fibers on the transcriptome and proteome level. The unbiased analysis allowed them to identify that MYHs are not the only drivers of skeletal muscle fiber type specification. Here, they described novel classifiers, including contractile, adhesion, metabolic and ribosomal proteins. Another interesting finding was that type 2X are phenotypically distinct to other fibers and that they do not represent not a distinct fiber type. Moreover, the authors investigated single fibers from patients with nemaline myopathy, which show a fibers shifting towards a non-oxidative phenotype independently of MYH. Overall, the data are quite interesting, the manuscript is well-written and follows a coherent flow. The statistical and bioinformatic visualisation is very clear and helpful for the interpretation of the data.
Nevertheless, there are some points that need to be discussed.
1. A weakness of the manuscript is that it contains too many different aspects that make it somewhat difficult to find the focus. In the first figures the difference between transcriptome and proteome is shown. The differences are interesting but is it mandatory to use both datasets to find the new classifiers? Is the overlap of RNA and Protein information important to investigate the human patients?
2. Furthermore, the authors present the regulation of ribosomal proteins between fibers, their structural localization and their important function as new classifiers. This is also a very interesting result. However, since the authors have emphasised this aspect, the question arises as to whether this has an influence on ribosomal activity in individual fiber types. There are some earlier studies from Alfred Goldberg discussing different metabolic activities of muscles with different fiber types.

3. The authors identified micro-proteins and the regulation between fiber types. Why have the authors mentioned this aspect? Is it important later to characterize the fibers from patients? of course the finding of new proteins is interesting, however since the authors have not performed any follow-up experiments to clarify the function of these proteins, this aspect might be better to discuss in another manuscript? If there is a link to the patients this might be interesting to discuss.

4. Finally, the authors describe single fibers from human patients. Have the authors tested whether the changes are also observable in intact muscle biopsies? Since the authors mentioned the great effort to isolate single fibers it might be easier

to measure the intact proteins and later deconvolute the protein regulation based on the known fiber distribution if necessary. In addition, in case of a common down- or up-regulation it might be not necessary to measure single fibers at least in this muscle disease? It remains a little unclear to what extent the detected protein changes affect the muscle disease. Is it an inflammatory or metabolic process that promotes degeneration or is it a stability problem including a reduced regeneration?

5. The authors mentioned/postulate that Myh1 positive fiber might be not a distinct fiber type. Have the authors overlapped their finding with previous data (including type 2x fiber types) as mentioned in the introduction. What about the expression with myosin light chain kinase 2 and actinin-3 in fast fibers, particularly in type 2x fibers?

6. What about muscle regeneration in nemaline myopathy? Do the authors observe a regulation of muscle stem cell markers? Can central nuclei be detected on the slices? This would indicate a regeneration of fibers.

Minor:

1. Why have the authors identified matrix proteins on single fibers?
2. Fig. 1I, J: is the legend of colour code missing?
2. Figure 3D heatmap legend is missing
3. What are the fold-changes of ribosomal proteins between fibers. This might help to judge the differences of ribosomal proteins between fibers compared to z-score normalized intensities or ratios.
4. Figure 3A. The length of the lines indicates something?
5. Figure 5E-G: the resolution is too low to recognize regulated factors in the figure
6. How are mitochondria regulated between fiber types. In addition to the abundance, do the authors recognize a specific adaption of mitochondrial proteins between types? Or in the patients?

Reviewer #3

(Remarks to the Author)
The manuscript presents an analysis of fiber level RNA and protein expression in human skeletal muscle. For RNA expression, single fibers were dissected from 14 subjects (12 males, 28 ± 6 years), barcoded and single fiber transcriptome assayed. For protein expression, timsTOF shotgun proteomics on individual dissected fibers from 5 males (29 ± 5 years) were performed. All assays were on samples obtained from vastis lateralis biopsies obtained for a different study with the subjects under differing groups of experimental conditions at the time of biopsy. More than individual 1000 fibers were assayed each for transcriptome and proteome. Additional studies were performed in nemaline myopathy samples. After a variety of informatic analyses, the authors challenge the MYH fiber type classification, specifically that Type IIx fiber is distinct and characterized by MYH type expression and that fibers show highly distinct MYH markers. The manuscript is generally clearly written and the data generated is impressive, and of value to the general research community. The analysis is sound and the study is an important contribution. However the conclusions of the study are somewhat contrived and are not convincingly supported by the data and their interpretation.

Major Points:
1. The question of fiber type in a multinucleated muscle cell is philosophically complex. Is the fiber type determined by the entire fiber (as assayed in this study), by fiber segments with transitions across the fiber, or by expression pattern driven by individual nuclei within each fiber (as assayable by single nucleus methods). With the advent of diverse single cell technologies, how to identify a cell subtype is disputable in any tissue, and the situation is even worse in a multinucleated cell with the developmental origin of muscle fibers. There are many instances of “classical” cell types not always being resolved in some single cell studies. The issue of “continuum” vs “distinct types” set up in the introduction is a straw dog--all cell types show marker continuums at single cell resolution (as well as histologically, biochemically and physiologically) . Indeed, the sole reference cited to support this “debate” about human fiber types in lines 63-65 [ref 1, Schiaffino and Reggiani, Phys. Rev., 2011] does not make that point at all (e.g. see Fig 10 in that review). Furthermore, there can be very different fiber types composition in fibers taken from different muscle or subject groups. Thus the strong, general conclusion made in this paper that the decades of histological, biochemical and physiological study of the 2X fiber should not be made based solely on the data from this one cutting edge omics study from ~30 y.o. Males (12/14 and 5/5) assayed from a single type of muscle. The contention by the authors that their study provides the “true cellular heterogeneity of human skeletal muscle fibers” is overstating what can be concluded from one, albeit novel and interesting study.

2. The data from the present study does not “challenge the concept that type 2X are phenotypically distinct” and challenge “the currently accepted model of multiple distinct fiber types defined by the expression of specific MYHs” (from the abstract) as it provides one interesting new dataset and a new window on the expression pattern of muscle fibers. The data, while impressive, does not include methylation, chromatin state, histone marks, isoforms, or other elements that may contribute to subtype identity. These results may “suggest” the view proposed by the authors, but any fiber type definition conclusions based on this single study are preliminary and warrant further investigation.

3. Specifically, the conclusion that MYH fiber types are not the main driver of heterogeneity seems overstated. The basis for this assertion seems to be that the first PC in the transcriptome PCA is driven by ribosomal and cell junction proteins and the first PC in the proteome PCA is driven by ribosomal and costamere proteins; neither are correlated with MYH classified fiber types. While this does appear to be correct, the percentage of variance in PC1 is 11.1-12.4% while the percentage of variance in PC2 (which is MYH fiber type driven) is 3.5-10.6%. This ranking does not justify considering PC1 to be the main

driver of fiber-type heterogeneity especially given that the phenotypic implications of PC1 are unclear while the phenotypic consequences of PC2 are extremely clear and very well-characterized. This plot supports the view that there are other axes of fiber-type heterogeneity, unrelated to MYH fiber types, that can be used to characterize fiber types.

4. The authors appear to reject prior human muscle physiology, histology and biochemistry research characterizing fiber subtypes, but never attempt to reconcile their results with the other fiber and single cell omics as well as the extensive pre-single cell fiber type literature. How do they explain the discrepancy between their results and the existing literature. Do they think all previous work was technically or conceptually flawed? One contribution to some of the apparent discrepancy could be that fiber level staining studies may identify fiber type based on the characteristic of the segment or cross-section of the fiber captured in the stain with the MYC or other markers varying across the length of the fiber. Alternatively could the specific RNAseq and TOF assays reflect different aspects of the expression of these markers leading to results inconsistent with other methods and protocols? In any event, the authors should discuss specific previous results and contextualize why their method appears inconsistent with diverse previous studies. Merely stating this is the results seen by “omics” begs the question of why the results appear to be different using other methods (including other omics methods).

5. I do not see any Linc and RP11 staining in Fig. 4G. Given this panel, perhaps rather than try to cherry-pick another image showing better evidence of co-localization, the RNAScope should be dropped. If the authors want to include these data, systematic quantification of these results is needed.

6. Was an LFC threshold applied for the DEGs, and if not, why?

7. A reasonable discussion of the limitations of this study (subject homogeneity, tissue homogeneity, samples obtained from an intervention study that are expected to impact the datasets, etc. etc.) is important for providing guidance in interpreting the findings in this paper.

Minor Points:
1. Language-wise, the term used to describe the heterogeneity in PC1 is somewhat awkward and unclear during a first reading - "MYH-based fiber type-independent" comes up many times in the paper and is hard to understand. A better phrasing would perhaps be "novel source of heterogeneity" or something like that, as long as it is well-defined at the beginning.

2. Figure 1, panels I-J would be enhanced with feature plots that show the blended expression of MYH1 and MYH2 so that readers can visualize exactly which fibers express either or both.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
There has been a lot of improvement in the manuscript and the authors answered to the questions which were a problem. Congratulations on this fine study.

Reviewer #2

(Remarks to the Author)
The authors have addressed many of the reviewers' concerns. However, the manuscript remains a collection of different parts that do not cohesively fit together, which ultimately obscures its central focus. The authors continue to discuss a variety of topics, including RNA-protein comparisons, fiber type characterization, the existence of type IIx fibers, the identification of lncRNA-encoded microproteins, and a disease model. Given that other reviewers have raised significant concerns about the fiber type characterization and the small number of sample size, the manuscript would benefit from a sharper focus on this aspect. Additionally, the relevance of lncRNAs in this context remains unclear. In summary, concentrating on fiber typing at both the RNA and protein levels would provide a clearer narrative, while the discussions on lncRNAs and the disease model could be expanded into a separate publication, allowing for more thorough validation and follow-up of the data.

Reviewer #3

(Remarks to the Author)
The revisions and responses adequately address the points previously raised.

Reviewer #1 (Remarks to the Author)

The article by Roger Moreno-Justicia is a very interesting paper on the heterogeneity of muscle fibers in the vastus lateralis (VL), a human thigh muscle in healthy and in Nemaline myopathy patients. Much state-of-the-art work has been done with human samples, for which single muscle fibers were manually isolated. The authors compare high-throughput transcriptomic and proteomic data on genes/proteins co-expressed in muscle fibers. Using numerous bioinformatic approaches, they show that there are only two main fiber types in the VL, a slow type and a fast type, and that the expression of genes encoding myosin heavy chains (Myh) are not the main determinants in distinguishing the different fiber types, but that there is a continuum in the expression of a certain number of genes, in particular those encoding ribosomal proteins and those involved in cell junctions between muscle fibers. In Nemaline myopathy samples, the authors reveal a shift towards less oxidative fibers independent of MYH-based fiber type. Unfortunately, apart from the methods used to analyze the data, the authors did not validate their results with basic experiments (immunocytochemistry for example) that would have provided a rigorous demonstration of their findings that lack statistical values.

We thank you for taking the time to critically review this work and providing constructive suggestions to improve the manuscript.

Main comments

I am a bit confused by the presence of pure 2X fibers at the RNA level, suggesting that MYH2 and MYH7 are off, while fibers expressing 2X protein are only hybrid fibers that also express 2A. Are the authors suggesting that the 2A gene could be on/off depending on e.g. day/night, and while the 2A protein is stable, its mRNA has a short half-life? This should be discussed as to how a fiber that does not express the gene accumulates the corresponding protein.

This is an excellent question and one we have been considering ourselves for some time. mRNA expression of the major MYH isoforms is not circadian in human skeletal muscle (Perrin et al, 2018; https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5902165/). However, this does not rule out that differential protein and/or mRNA stability are apparent for MYH isoforms. We have discussed this paradox on page 23, lines 518-529:

"The discrepancy between the identification of near pure type 2X fibers at the mRNA level but only hybrid type 2A/2X fibers at the protein level is puzzling. mRNA expression of MYH isoforms is not circadian[67], indicating that it is unlikely that we have simply "missed" the on-signal of MYH2 in the apparently pure type 2X fibers at the RNA level. One possible explanation may be differences in the protein and/or mRNA stability of MYH isoforms, though this is purely hypothetical. Indeed, no fast fiber was 100% pure for any MYH

isoform, though whether levels of MYH1 mRNA expression in the range of 70-90% could result in similar MYH1 and MYH2 abundance at the protein level is unclear".

It should be stated in the introduction that the human vastus lateralis is known to have a low percentage of 2X fibers compared to other muscles with a higher percentage of 2X positive fibers, or that all human muscles contain very few 2X fibers and VL is thus a good archetype of human muscles; the results should be discussed in light of this knowledge. Furthermore, I suspect that the percentage of oxidative/glycolytic positive myofibers may vary depending on the depth at which the biopsy was taken in the VL.

We agree that there might be differences depending on which muscle is analyzed, although we believe that the vastus lateralis is a representative muscle. At the very least, the human vastus lateralis is a mixed muscle with, on average, equal proportions of slow and fast fibers. Other muscles (e.g., triceps brachii) can be more fast-twitch. We are in the process of assembling a database of single-fiber SDS-PAGE studies from the last 30 years, and these results indicate that the average proportion of 2X fibers is very low (below 1%), independent of which human muscle (vastus lateralis, gastrocnemius, deltoideus or soleus) is analyzed.

However, we do agree that it is important to add this nuance to the manuscript, which we have now done on page 4, lines 73-77, and on page 23, lines 518-522.

" The picture is further complicated by different skeletal muscles displaying specialized fiber type profiles within humans[7]. The vastus lateralis is a mixed muscle type with average (and therefore representative) MYH expression profiles[7]. Together with this, its accessibility for sampling makes it the most commonly studied muscle in humans."

"This pipeline resulted in the observation that pure type 2X fibers could not be identified at the protein level in vastus lateralis skeletal muscle of our cohort of young healthy males. This is in line with previous single-fiber studies identifying <1% pure type 2x fibers in healthy vastus lateralis[66], although this should be confirmed in other muscles in the future."

In their sentence lines 179-183, do the authors mean that the expression of MYH1 and genes/proteins co-expressed with MYH1 is not discriminatory? This needs to be better explained. See for example the genes/proteins co-expressed with Myh 2X in the manuscripts of Murgia et al.

We simplified this section (pages 8-9, lines 178-188) in an attempt to add clarity, and have added additional plots (Fig 1 S4) to substantiate our claims (see also next comment). The section now reads:

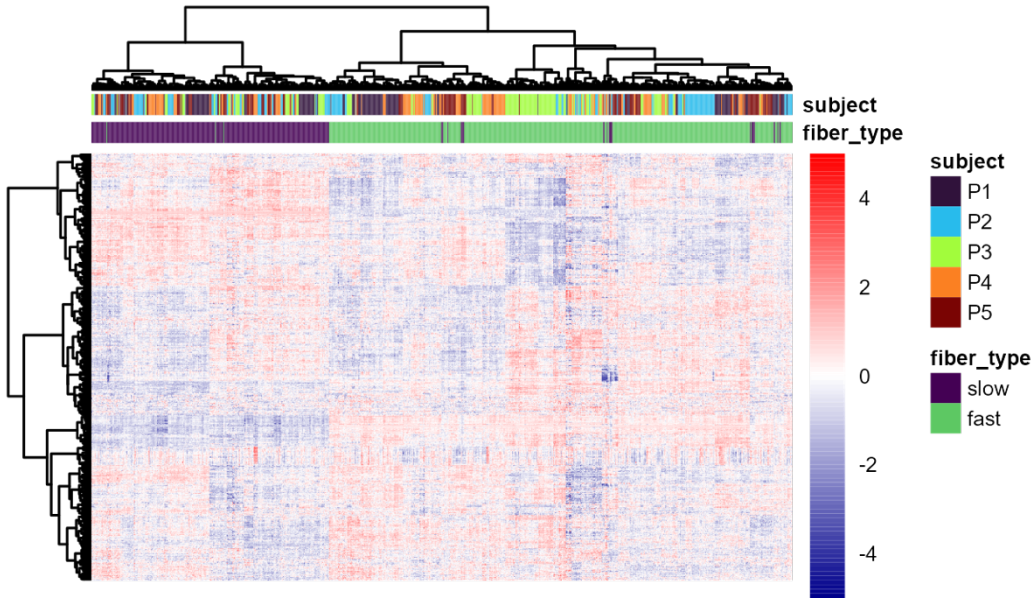
*"Two distinct clusters were apparent in the UMAP plots, which represented "fast" and "slow" fibers (**Fig 1G-H**). MYH7+ (slow) fibers clustered to the positive side of UMAP1, and MYH2+ and MYH1+ (fast) fibers clustered to the negative side of UMAP1 (**Fig 1I-J**). No distinction between the various fast MYH-based fiber types (i.e. type 2A, type 2X, or hybrid 2A/2X) could be identified, however, suggesting that when the whole transcriptome or proteome is taken into account, the expression of MYH1 (**Fig 1 I-J**), or other classical markers of type 2X fibers like ACTN3 or MYLK2 (**Fig 1 S4A-B**), does not discriminate between distinct fiber types. Furthermore, in contrast to MYH2 and MYH7, few transcripts or proteins positively correlate with MYH1 (**Fig 1 S4C-H**), suggesting that MYH1 abundance does not adequately reflect the muscle fiber transcriptome/proteome. Similar conclusions can be drawn when assessing the blended expression of the three MYH isoforms on a UMAP level (**Fig 1 S4I-J**)."*

In respect to the work of Murgia et al, even if a small number of specific proteins correlate with MYH1, it can be seen in the PCA of figure 1A of the manuscript *Protein profile of fiber types in human skeletal muscle: a single-fiber proteomics study* (https://skeletalmusclejournal.biomedcentral.com/articles/10.1186/s13395-021-00279-0) that type 2X fibers do not cluster distinctly from type 2A fibers when the whole proteome is considered. In this respect, we do not believe that there is any major discrepancy between our data and those of Murgia et al. To aid comparison between the works, we have added UMAP plots coloured by the gene expression/protein intensity of ACTN3 and MYLK2, two markers of type 2X fibers from Murgia et al (**Figure 1 S4 A-B**).

In the chapter starting on line 187, the authors establish transcriptomic and proteomic correlations with several genes. Heatmaps of co-expressed genes can be shown for clarity. The heat map allows the unbiased identification of the genes co-expressed with the better correlation coefficient. The idea that there is a large heterogeneity of genes expressed with Myh7 or with Myh2 or with Myh1 has been well illustrated in several mouse snRNA-seq publications in 2020. This heterogeneity illustrates well that there is not a unique transcriptomic/proteomic signature of Myh7- or Myh2-positive myofibers, but that there is a large variability of co-expressed genes depending on the muscle analyzed.

We thank you for your comment and think this heterogeneity instead of unique signatures for fiber types is an important message of our manuscript. In the revised manuscript, we provide U-plots of correlations between MYH1, MYH2, and MYH7 transcripts/proteins and the rest of the transcriptome/proteome in Figure 1 S4 C-H. These plots illustrate which transcripts/proteins most strongly correlate with each myosin heavy chain isoform as well as the heterogeneity in co-expression amongst transcripts/proteins. We chose to represent these data in this manner as we felt U-plots better highlighted co-expressing transcripts/proteins than a heatmap, in which the magnitude of difference between fast

and slow fiber types masked any heterogeneity amongst transcripts/proteins (heatmap for proteomics provided below).



Line 240-241, I find it a bit provocative to state that "no distinct subtypes of human skeletal muscle fibers beyond a slow-fast classification should be considered" with the study of a single limb muscle, although I agree with the authors that their data show that the slow/fast fiber type-independent heterogeneity across fibers is of a "continuous nature".

We have deleted the clause "and no distinct subtypes of human skeletal muscle fibers beyond a slow-fast classification should be considered." The sentence in page 11, lines 235-236 now reads as:

"These data collectively show that slow/fast fiber type independent heterogeneity across fibers is of a continual nature."

Line 335; SREBPF1 should be replaced with SREBF1. It is difficult to understand the significance of the histograms in Figure 4S1 D-K. Is there a way to get statistical values for e.g. SREBF1 enrichment in fast versus slow myofibers? Otherwise, without statistical values, the histograms do not improve the manuscript. Otherwise, explain better how these data were generated and why no statistical value can be provided.

We thank you for identifying this typographical error, we have rectified this (page 15, line 327). In addition, we identify adjusted p values for each of the comparisons in Figures 4S1 D-K.

The paragraph from lines 351-362 should be written with more explanation. How were the binding sites for specific transcription factors detected in the list of slow/fast enriched genes? In the promoter regions of selected genes? In their 100kb vicinity? The legend of Figure 4E is not very informative. If AP2 (TFAP2A, AP2, AP2a and AP2g, Figure 4E) and Tieg1/Klf10 are the putative drivers of fast and slow myofiber fate, respectively, immunocytochemistry with corresponding antibodies must be performed to confirm the prediction, or at least their expression profile must be provided in feature plots. The enrichment of MEF2 in fast fibers is counterintuitive since it has been shown to control the slow/oxidative phenotype, at least in the mouse. MEF2 enrichment is not shown in Figure 4E.

This is an excellent comment and has directed us to rethink the transcription factor enrichment analysis that we have performed. As a result, we have reanalysed these data using SCENIC – the current gold standard methodology for differential transcription factor enrichment (https://www.nature.com/articles/nmeth.4463). This approach better compares the fast and slow phenotype rather than enriching simply for muscle-specific transcription factors in both datasets – the reason why we believe MEF2 was enriched within fast fibers using the previous analysis. Using SCENIC, we find a number of known and novel transcription factors enriched within fast and slow fibers, including MAFA, PITX1, and MYF6 in fast fibers, and ZSCAN30 and EPAS1 in slow fibers.

The text on page 15, lines 332-341 now reads:
*"To get more insights into the transcriptional regulation of slow/fast fiber type signatures, we performed transcription factor enrichment analysis using SCENIC[49] (Table S16). Many transcription factors were significantly enriched between fast and slow fibers (**Fig 4E**). This included transcription factors such as MAFA, previously linked to the development of fast fibers[50], but also multiple novel transcription factors potentially involved in the fiber type specific gene program. These included PITX1, EGR and MYF6 as the most enriched transcription factors within fast fibers (**Fig 4E**). Conversely, ZSCAN30 and EPAS1 (also known as HIF2A) were the most enriched transcription factors within slow fibers (**Fig 4E**). In line with this, MAFA expression levels were higher in the UMAP area corresponding to fast muscle fibers, whereas an opposite expression pattern was observed for EPAS1 (**Fig 4F**)."*

How many fibers were counted for the representative 4G image?
Please note that this is now Fig 5B. Quantification of these data are now included in Fig 5C. Furthermore, we have added the following information in the methods section of the manuscript (page 42, lines 975-979):

"Biopsies from three individuals were used for quantification, with 187 muscle fibers being counted in total. As in RNAscope each dot corresponds to one RNA molecule, the number of dots/mm^2 was used as a measure of RNA expression. We first determined the number of dots/mm^2 within each fiber for both probes, then averaged the results by fiber type and participant. The given average was then used as an input for a two sample t-test."

It is important to include the common name for the TF indicated (Klf10 for Tieg1, AP2alpha for TFAP2A, TEAD2 for ETF, ZNF263 for FPM-315). Note that in the mouse, TEAD proteins, like MEF2 proteins, are known to control the slow/oxidative phenotype. This should be discussed in the context of the muscle analyzed here, which has very few Myh2X myofibers, and the results presented or immunocytochemistry provided showing the accumulation of these TFs preferentially in fast nuclei. This should also be discussed in the light of previously published proteomic and transcriptomic data on single human myofibers from different muscles where MyH2X myofibers are identified.

Please see previous response relating to the updated transcription factor analysis.

In Figure 4I (and more generally in each histogram) each point should be represented.

We thank the reviewer for this comment. All of our boxplots are now showing the individual data points.

Figure 5S2J. I am a bit surprised by the fold change in expression of ANXA1, A2, A5 and A6 between control and mutant samples. For example, for ANXA1, the LFQ (label-free quantitation, please specify in your paper) intensity is about 12 for the control and about 14 for the Acta1-NM, which means an increased expression of 1.2-fold, is this correct? Is this difference significant? Are the authors convinced that this 1.2 fold difference has a physiological relevance?

Please note that this is now figure 6 S3B. These data are log transformed and therefore there is a log2 fold change of approximately 2 (or an absolute four-fold difference). This was not originally outlined on the figure, so we apologize for any confusion caused. We have rectified this.

The TPPP3 (Tubulin Polymerization Promoting Protein Family Member 3) seems to be more differentially expressed between the samples, with a good p-value. Is tubulin polymerization increased in nemaline myopathies? The position of the S100A11 protein is not shown in the Volcano plots.

As far as we know, tubulin polymerization has never been studied in the context of nemaline myopathy. Only one study reports a heterogenous localization of beta-tubulin within muscle fibres of a mouse model of nemaline myopathy (expressing nebulin mutants) (Ross JA et al, 2019; https://pubmed.ncbi.nlm.nih.gov/31218456/).

We have added S100A11 to all volcano plots where it was significantly different between conditions (see figure 6 S2D, E and G).

Please modify the title of the Supplementary tables according to their number. 488870_0_data_set_4288201_s4txxx is for example is table 27 488870_0_data_set_4288185_s4txxw is table 11.
Not easy presently to identify the tables with their name. The txxx could be modified for example by the number of the table (t1, t2..), it would thus be more easy to find the information.

We have not named the Supplementary Tables in this manner. This must be a result of the processing performed by the manuscript submission system after submission. We believe Supplementary table numbers will be displayed as provided by us once the manuscript is available online.

In the Discussion section, I am not convinced by paragraph lines 519-530. The authors use a muscle with a very low Myh2X content. It is therefore difficult to write: "The existence of pure type 2X fibers should be questioned...". It would be more rigorous to write "....questioned in the vastus lateralis studied here". It is known that the percentage of 1/2A/2X and mixed myofibers varies depending on the muscle studied. This percentage also varies during development and aging. A reference should be made to whether 2X myofibres are very rare in all human muscles.

We agree with you that this conclusion extrapolated further than our data supports. We have edited this sentence to page 23, lines 518-520:

“This pipeline resulted in the observation that pure type 2X fibers could not be identified at the protein level in vastus lateralis skeletal muscle of our cohort of young healthy males.”

Reviewer #2 (Remarks to the Author)

The manuscript from Roger Moreno-Justicia and colloquies measured a tremendous amount of human single muscle fibers on the transcriptome and proteome level. The unbiased analysis allowed them to identify that MYHs are not the only drivers of skeletal muscle fiber type specification. Here, they described novel classifiers, including

contractile, adhesion, metabolic and ribosomal proteins. Another interesting finding was that type 2X are phenotypically distinct to other fibers and that they do not represent not a distinct fiber type. Moreover, the authors investigated single fibers from patients with nemaline myopathy, which show a fibers shifting towards a non-oxidative phenotype independently of MYH. Overall, the data are quite interesting, the manuscript is well-written and follows a coherent flow. The statistical and bioinformatic visualisation is very clear and helpful for the interpretation of the data. Nevertheless, there are some points that need to be discussed.

Thank you for taking the time to review and improve this manuscript.

1. A weakness of the manuscript is that it contains too many different aspects that make it somewhat difficult to find the focus. In the first figures the difference between transcriptome and proteome is shown. The differences are interesting but is it mandatory to use both datasets to find the new classifiers? Is the overlap of RNA and Protein information important to investigate the human patients?

The information gathered by both technologies is important and complementary, as they each have distinct advantages. For example, by integrating the datasets we have identified novel fiber-type specific non-coding RNAs and microproteins that could not be identified using either technique independently (Figure 5). Furthermore, in Figure 2 we show concordant and discordant features of the two datasets, yielding new insights into muscle fiber signatures at the mRNA and protein level. We therefore pose that the integration of both technologies is a strength, rather than a weakness, of our manuscript.

2. Furthermore, the authors present the regulation of ribosomal proteins between fibers, their structural localization and their important function as new classifiers. This is also a very interesting result. However, since the authors have emphasised this aspect, the question arises as to whether this has an influence on ribosomal activity in individual fiber types. There are some earlier studies from Alfred Goldberg discussing different metabolic activities of muscles with different fiber types.

Our data certainly raises this question, however the technology required to investigate ribosomal activity (e.g. ribo-seq) requires as of yet a sample input that is incompatible with single-fiber analyses. Therefore, the technology is not yet sensitive enough for application to individual skeletal muscle fibers, although we believe this will become available in the coming years. We allude to these possibilities on page 24, lines 548-553, and on page 26, lines 585-589.

“The underlying reason for such vast slow/fast-fiber type independent heterogeneity amongst fibers is not immediately obvious, though this could allude to a specialized

spatial organization within a muscle fascicle to enable an optimal response to specific forces and loads[72], specialized cellular or organ communication with other cell types within the muscle microenvironment[73-75], or differential ribosomal activity in individual fibers."

"As we and others continue to optimize single-cell and single muscle fiber omics-analyses towards ultra-low sample input (as demonstrated here with the analysis of fibers from nemaline myopathy patients), the possibility of combining multi-omics (and functional) approaches in a single muscle fiber is tantalizingly close."

3. The authors identified micro-proteins and the regulation between fiber types. Why have the authors mentioned this aspect? Is it important later to characterize the fibers from patients? of course the finding of new proteins is interesting, however since the authors have not performed any follow-up experiments to clarify the function of these proteins, this aspect might be better to discuss in another manuscript? If there is a link to the patients this might be interesting to discuss.

We believe these data highlight important novel information about skeletal muscle physiology *per se*, independent of nemaline myopathy. Given that studying healthy muscle physiology is a major component of this manuscript we believe that it is important. However, we agree that this aspect could be strengthened by providing a link to the myopathy dataset. Therefore, in the revised manuscript, we searched the myopathy proteome using the custom microprotein FASTA file. In this dataset we identified 5 microproteins. It is to be expected that the number of microproteins identified in these samples would be substantially lower than in the 1000 fiber dataset, given the low protein input from the small fibers of nemaline myopathy patients. Three of these microproteins displayed differential abundance in nemaline myopathy patients, with effects consistent with our overall finding that fibers from nemaline myopathy patients resemble healthy fast skeletal muscle fibers independent of their MYH composition.

These data have been included in Fig 6 S3 A and on page 19, lines 437-444.

*"Three microproteins were also regulated by either ACTA1- or TNNT1-nemaline myopathies. Two of those microproteins, ENSG00000215483_TR14_ORF67 (also known as LINC00598 or Lnc-FOXO1) and ENSG00000229425_TR25_ORF40 (lnc-NRIP1-2) displayed differential abundance only within type 1 fibers, with ENSG00000215483_TR14_ORF67 having previously reported to play a role in cell cycle regulation[56]. On the other hand, ENSG00000232046_TR1_ORF437 (corresponding to LINC01798) was upregulated in both type 1 and type 2 fibers from ACTA1-nemaline myopathy when compared to healthy controls (**Fig 6 S3A, Table S27**)."*

4. Finally, the authors describe single fibers from human patients. Have the authors tested whether the changes are also observable in intact muscle biopsies? Since the authors mentioned the great effort to isolate single fibers it might be easier to measure the intact proteins and later deconvolute the protein regulation based on the known fiber distribution if necessary. In addition, in case of a common down- or up-regulation it might be not necessary to measure single fibers at least in this muscle disease? It remains a little unclear to what extent the detected protein changes affect the muscle disease. Is it an inflammatory or metabolic process that promotes degeneration or is it a stability problem including a reduced regeneration?

We have now measured the bulk proteome of muscle biopsies from TNNT1-nemaline myopathy and healthy control individuals and have included these data in Fig 6 S4 and on page 21 lines 476-485 (text also included below). ACTA1-nemaline myopathy samples could not be measured in a bulk-manner due to lack of material. The results from the bulk proteome validate our findings from single muscle fibers. However, the heterogeneity and variance between fibers can only be identified using a single-fiber approach, and potential fiber type specific proteome remodeling would have been missed with a bulk-only approach.

*"In order to assess whether the observed effects of nemaline myopathy were also present at the whole muscle-level, we conducted a bulk proteome analysis on muscle biopsies from the same TNNT1-nemaline myopathy patients and compared them against control individuals (n = 3 per group) (**Fig 6 S4A, Table S28**). As expected, upon PCA, control individuals tightly clustered together whereas, similarly to the single fiber analysis, and TNNT1-nemaline myopathy patients displayed a higher inter-sample variance (**Fig 6 S4B**). The bulk analysis was able to recapitulate the differentially expressed proteins (**Fig 6 S4C, Table S29**) and biological processes (**Fig 6 S4D, Table S30**) highlighted in the single fiber comparison, although lost the potential to discriminate between fiber types, and did not account for the heterogeneous effect of the disease across different fibers."*

We agree with the reviewer in that deconvolution algorithms are becoming more powerful and are a potential solution to this particular question, as they allow inferring fiber type composition out of a given reference dataset. Nevertheless, these algorithms are extremely tissue and/or organism –specific and in order to be applied successfully the algorithm should be provided a reference dataset consisting of dissected single fibers (one way how our dataset could be used in the future). In this particular case, such a dataset should include both ACTA1- and TNNT1- nemaline myopathy fibers, which was not yet available. Moreover, most deconvolution tools simplify their algorithms to use only a handful of highly abundant protein markers as a reference for deconvolution, which in this case would have made it impossible to separate the influence of MYHs from

fiber metabolism. Overall, we think that the workflow presented in this manuscript allowed us to characterize nemaline myopathy in an unprecedented resolution, and adds to the (patho)physiological understanding of the disease.

5. The authors mentioned/postulate that Myh1 positive fiber might be not a distinct fiber type. Have the authors overlapped their finding with previous data (including type 2x fiber types) as mentioned in the introduction. What about the expression with myosin light chain kinase 2 and actinin-3 in fast fibers, particularly in type 2x fibers?

Regarding the markers suggested by the reviewer, we have now added UMAP plots colored by the RNA expression/protein abundance of both ACTN3 and MYLK2 (figure 1 S4 A-B; page 9 lines 183-184). We do not observe any particular pattern in the clustering that could indicate that these markers are discriminatory within fast fibers when the whole transcriptome or proteome is considered. This is further supported by only limited correlations of transcripts/proteins with MYH1 expression/abundance (Fig 1 S4C-H).

6. What about muscle regeneration in nemaline myopathy? Do the authors observe a regulation of muscle stem cell markers? Can central nuclei be detected on the slices? This would indicate a regeneration of fibers.

Nemaline myopathy is one of the most common congenital myopathies. This genetic disorder is not considered a dystrophy as centrally located nuclei (marker of regeneration) are not usually/routinely found (Fisher G et al, 2022; https://pubmed.ncbi.nlm.nih.gov/36524401/). In line with this, we do not see any up-regulation of major stem cell markers in our analysis.

<u>Minor</u>

1. Why have the authors identified matrix proteins on single fibers?

We identify extracellular matrix proteins because these are structurally attached to the sarcolemma and therefore purified as a part of single muscle fibers.

2. Fig. 1I, J: is the legend of colour code missing?

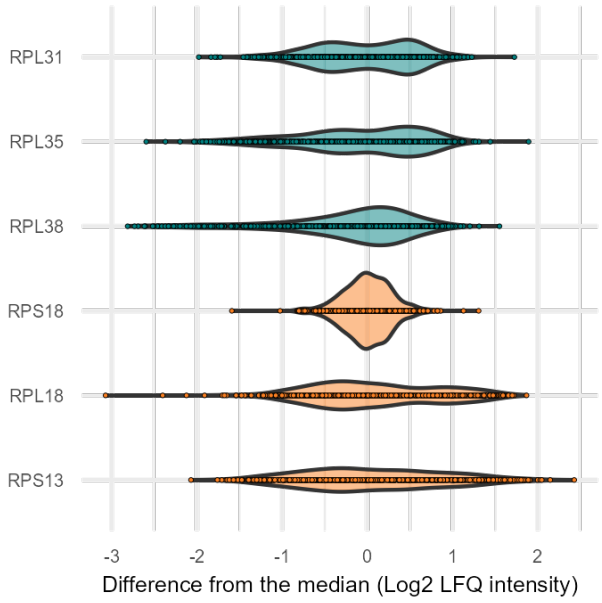
We have updated the plots to include the legend title.

2. Figure 3D heatmap legend is missing

We have now updated that plot to include the legend title.

3. What are the fold-changes of ribosomal proteins between fibers. This might help to judge the differences of ribosomal proteins between fibers compared to z-score normalized intensities or ratios.

When performing clustering analysis (both UMAP or PCA), the differential abundance of ribosomal proteins does not seem to generate distinct clusters of fibers. Rather, our results show that the fiber type-independent heterogeneity observed is of a continual nature and thus it is not possible to perform differential expression analysis across distinct fiber groups. This continual nature is further shown in the new plot below, which has been added in Figure 3 S1C.



4. Figure 3A. The length of the lines indicates something?

The arrows in figure 3A indicate the direction towards which the proteins are driving the variance in the PCA plot. The length of the lines represents the principal component score, which is how much of the variance of that particular component is explained by a given protein. We have now edited the figure 3 caption to include this information. Please see page 55 lines 1282-1284:

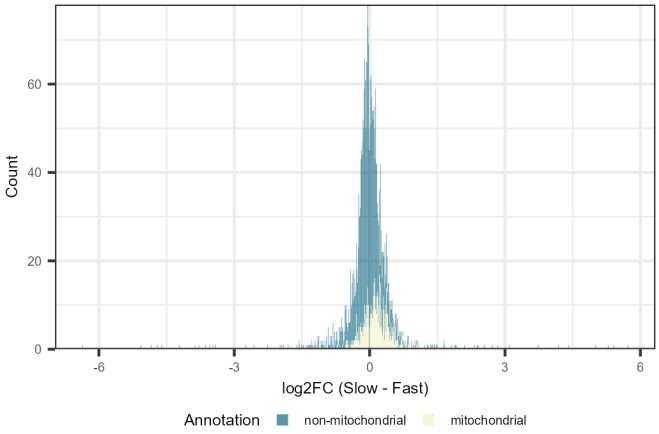
"Figure 3. Ribosomal heterogeneity drives fiber type-independent heterogeneity. (A) *Principal component analysis (PCA) plot colored by cytosolic ribosome gene ontology (GO) term in the proteome. Arrows indicate the direction towards which the proteins are driving the variance in the PCA plot. The length of the lines represents the principal component score for a given protein. "*

5. Figure 5E-G: the resolution is too low to recognize regulated factors in the figure

We have now modified this figure (now Figure 6E-G) to have a larger font size.

6. How are mitochondria regulated between fiber types. In addition to the abundance, do the authors recognize a specific adaption of mitochondrial proteins between types? Or in the patients?

We thank the reviewer for this question. In our analysis we see a higher abundance of mitochondrial transcripts/proteins (mitochondrial ribosomes, OXPHOS proteins, etc.) in slow fibers when compared against their fast counterpart (e.g., Figure 4D). Nonetheless, we have not measured mitochondrial abundance in this study, so it is impossible to rule out if the differences in gene expression/protein abundance are due to a higher amount of mitochondria in slow fibers or if the mitochondria are different between fiber types *per se*. We can see on the plot on the right that not all mitochondrial proteins are increased in slow fibers, and some are even higher in fast fibers. A notable example of a mitochondrially-localised protein that is upregulated in fast fibers is LDHA (Figure 4B).



Reviewer #3 (Remarks to the Author)

The manuscript presents an analysis of fiber level RNA and protein expression in human skeletal muscle. For RNA expression, single fibers were dissected from 14 subjects (12 males, 28 ± 6 years), barcoded and single fiber transcriptome assayed. For protein expression, timsTOF shotgun proteomics on individual dissected fibers from 5 males (29 ± 5 years) were performed. All assays were on samples obtained from vastis lateralis biopsies obtained for a different study with the subjects under differing groups of experimental conditions at the time of biopsy. More than individual 1000 fibers were assayed each for transcriptome and proteome. Additional studies were performed in nemaline myopathy samples. After a variety of informatic analyses, the authors challenge the MYH fiber type classification, specifically that Type IIx fiber is distinct and characterized by MYH type expression and that fibers show highly distinct MYH markers. The manuscript is generally clearly written and the data generated is impressive, and of value to the general research community. The analysis is sound and the study is an important contribution. However the conclusions of the study are somewhat contrived

and are not convincingly supported by the data and their interpretation.

Thank you for this fair and constructive review.

Major Points

1. The question of fiber type in a multinucleated muscle cell is philosophically complex. Is the fiber type determined by the entire fiber (as assayed in this study), by fiber segments with transitions across the fiber, or by expression pattern driven by individual nuclei within each fiber (as assayable by single nucleus methods). With the advent of diverse single cell technologies, how to identify a cell subtype is disputable in any tissue, and the situation is even worse in a multinucleated cell with the developmental origin of muscle fibers. There are many instances of "classical" cell types not always being resolved in some single cell studies. The issue of "continuum" vs "distinct types" set up in the introduction is a straw dog--all cell types show marker continuums at single cell resolution (as well as histologically, biochemically and physiologically) . Indeed, the sole reference cited to support this "debate" about human fiber types in lines 63-65 [ref 1, Schiaffino and Reggiani, Phys. Rev., 2011] does not make that point at all (e.g. see Fig 10 in that review). Furthermore, there can be very different fiber types composition in fibers taken from different muscle or subject groups. Thus the strong, general conclusion made in this paper that the decades of histological, biochemical and physiological study of the 2X fiber should not be made based solely on the data from this one cutting edge omics study from ~30 y.o. Males (12/14 and 5/5) assayed from a single type of muscle. The contention by the authors that their study provides the "true cellular heterogeneity of human skeletal muscle fibers" is overstating what can be concluded from one, albeit novel and interesting study.

Thank you for this interesting perspective. It is one we agree with and we have therefore edited the first paragraph of the introduction to remove reference to the "discrete vs continuous debate" and instead highlight that the field is increasingly viewing skeletal muscle fibers as falling on a continuum:

"However, skeletal muscle fibers are increasingly being viewed along a continuum, as opposed to discrete fiber types[6], fueled by the identification and subsequent acceptance of the existence of "hybrid" fibers that express varying proportions of multiple MYHs simultaneously." (Page 4, lines 67-70)

In addition, we refer to the different fiber type profiles across different muscles on page 4, lines 73-77 and on lines 518-520 on page 23.

"The picture is further complicated by different skeletal muscles displaying specialized fiber type profiles within humans[7]. The vastus lateralis is a mixed muscle type with

average (and therefore representative) MYH expression profiles[7]. Together with this, its accessibility for sampling makes it the most commonly studied muscle in humans."

"This pipeline resulted in the observation that pure type 2X fibers could not be identified at the protein level in vastus lateralis skeletal muscle of our cohort of young healthy males. This is in line with previous single-fiber studies identifying <1% pure type 2x fibers in healthy vastus lateralis[66], although this should be confirmed in other muscles in the future."

Finally in the introduction we have removed the word "true" so that the sentence referred to above now states:
"Our methodological development of transcriptome and proteome profiling of single skeletal muscle fibers manually isolated from human biopsy specimens, and their application to over a thousand fibers each, enabled us to investigate the cellular heterogeneity of human skeletal muscle fibers." (page 5, lines 93-96).

2. The data from the present study does not "challenge the concept that type 2X are phenotypically distinct" and challenge "the currently accepted model of multiple distinct fiber types defined by the expression of specific MYHs" (from the abstract) as it provides one interesting new dataset and a new window on the expression pattern of muscle fibers. The data, while impressive, does not include methylation, chromatin state, histone marks, isoforms, or other elements that may contribute to subtype identity. These results may "suggest" the view proposed by the authors, but any fiber type definition conclusions based on this single study are preliminary and warrant further investigation.

We thank you for this suggestion. We have edited the sentences referred to above to now read

"Furthermore, whilst slow and fast fiber clusters can be identified, described by their contractile and metabolic profiles, our data suggests that type 2X fibers are not phenotypically distinct to other fast fibers at an omics level." (Page 3, lines 50-53)
and
"Overall, our data indicates that skeletal muscle fiber heterogeneity is multi-dimensional with sources of variation beyond myosin heavy chain isoforms." (Page 3, lines 59-60).

Moreover, we have added a paragraph in the discussion section where we address the limitations of the study and highlight next steps in single muscle fiber profiling:

"Whilst this study represents the largest omic analysis of intact human single muscle fibers to date, it is not without limitations. We isolated skeletal muscle fibers from a

relatively small and homogenous sample of participants and from a single muscle (vastus lateralis). In this respect, it is impossible to rule out the existence of specific fiber populations in different muscle types and at the extremes of muscle physiology. For example, we cannot rule out that a sub type of ultra-fast fibers (e.g. pure type 2X fibers) may become apparent in highly trained sprint and/or power athletes[79] or during muscle disuse[66,80]. Furthermore, our limited participant pool did not allow us to investigate sex differences in fiber heterogeneity, in which there are known differences in fiber type proportions between males and females. Furthermore, we were unable to perform transcriptomics and proteomics on the same muscle fibers or in samples from the same participants. As we and others continue to optimize single-cell and single muscle fiber omics-analyses towards ultra-low sample input (as demonstrated here with the analysis of fibers from nemaline myopathy patients), the possibility of combining multi-omics (and functional) approaches in a single muscle fiber is tantalizingly close."
(Pages 25-26, lines 575-589).

3. Specifically, the conclusion that MYH fiber types are not the main driver of heterogeneity seems overstated. The basis for this assertion seems to be that the first PC in the transcriptome PCA is driven by ribosomal and cell junction proteins and the first PC in the proteome PCA is driven by ribosomal and costamere proteins; neither are correlated with MYH classified fiber types. While this does appear to be correct, the percentage of variance in PC1 is 11.1-12.4% while the percentage of variance in PC2 (which is MYH fiber type driven) is 3.5-10.6%. This ranking does not justify considering PC1 to be the main driver of fiber-type heterogeneity especially given that the phenotypic implications of PC1 are unclear while the phenotypic consequences of PC2 are extremely clear and very well-characterized. This plot supports the view that there are other axes of fiber-type heterogeneity, unrelated to MYH fiber types, that can be used to characterize fiber types.

We have toned down this assertion throughout the manuscript. For example:

"In doing so, we demonstrate the power of muscle fiber phenotyping at the transcriptomic and proteomic level, identifying that metabolic, ribosomal, and cell junction proteins are important sources of variation among muscle fibers." (Line 96-98 on page 5)

"Unexpectedly, MYH-based fiber type explained only the second greatest degree of variability (PC2), indicating that other biological factors (PC1) independent of MYH-based fiber type, have a substantial role in regulating skeletal muscle fiber heterogeneity." (Line 218-220 on page 10)

"Furthermore, we identify metabolic, ribosomal, and cell junction proteins as major determinants of skeletal muscle fiber heterogeneity." (Line 503-504 on page 22)

4. The authors appear to reject prior human muscle physiology, histology and biochemistry research characterizing fiber subtypes, but never attempt to reconcile their results with the other fiber and single cell omics as well as the extensive pre-single cell fiber type literature. How do they explain the discrepancy between their results and the existing literature. Do they think all previous work was technically or conceptually flawed? One contribution to some of the apparent discrepancy could be that fiber level staining studies may identify fiber type based on the characteristic of the segment or cross-section of the fiber captured in the stain with the MYC or other markers varying across the length of the fiber. Alternatively could the specific RNAseq and TOF assays reflect different aspects of the expression of these markers leading to results inconsistent with other methods and protocols? In any event, the authors should discuss specific previous results and contextualize why their method appears inconsistent with diverse previous studies. Merely stating this is the results seen by "omics" begs the question of why the results appear to be different using other methods (including other omics methods).

We thank the reviewer for this comment. The discussion section of the manuscript has been modified to consider these comparisons. We do not think that our results are inconsistent with previous reports using targeted approaches or omics technologies. Antibody-based single muscle fiber typing consistently identifies <1% pure type 2x fibers in vastus lateralis and several other limb muscles from healthy individuals. Single nuclei sequencing consistently identifies just two distinct clusters (slow and fast) of myonuclei, whilst PCA analyses of previous proteomics data identify fibers along a continuum. Our data builds upon these findings to provide more evidence that skeletal muscle fibers exist along a continuum, with more analyzed fibers and in human muscle, as you argue in comment 1. (Please see page 23, lines 518 –522 and lines 529-536).

"This pipeline resulted in the observation that pure type 2X fibers could not be identified at the protein level in vastus lateralis skeletal muscle of our cohort of young healthy males. This is in line with previous single-fiber studies identifying <1% pure type 2x fibers in healthy vastus lateralis[66], although this should be confirmed in other muscles in the future."

"Nonetheless, when the whole transcriptome or proteome is considered, clustering-based analyses could only confidently identify two distinct clusters, which represented slow and fast skeletal muscle fibers regardless of their exact MYH composition. This is consistent with analyses using single nuclei transcriptomics approaches which commonly identify only two distinct clusters of myonuclei[68-70]. Furthermore, whilst previous proteomics-based studies have identified type 2X fibers, these fibers did not

cluster away from the rest of the fast fibers and only displayed a handful of differentially abundant proteins compared to other MYH-based fiber types[14].”

Furthermore, when compared with traditional antibody-based approaches, we can demonstrate that the higher sensitivity of mass spectrometry provides greater resolution in typifying a fiber (see Figure 1 S2 for a comparison of mass spectrometry and the antibody-based dot blotting technique) (Please see pages 22-23 lines 512 –515).

“As demonstrated herein, omics approaches also have the advantage of increased sensitivity in the quantification of fiber-type markers over traditional antibody-based approaches, whilst also not relying on the quantification of a single (or few) markers for determining skeletal muscle fiber type.”

5. I do not see any Linc and RP11 staining in Fig. 4G. Given this panel, perhaps rather than try to cherry-pick another image showing better evidence of co-localization, the RNAScope should be dropped. If the authors want to include these data, systematic quantification of these results is needed.

In the revised manuscript, these data have been quantified and included in the figure (figure 5 panel B-C). We also included the number of fibers used for the analysis in the methods section (Page 42, lines 975-979):
“Biopsies from three individuals were used for quantification, with 187 muscle fibers being counted in total. As in RNAscope each dot corresponds to one RNA molecule, the number of dots/mm^2 was used as a measure of RNA expression. We first determined the number of dots/mm^2 within each fiber for both probes, then averaged the results by fiber type and participant. The given average was then used as an input for a two sample t-test.

6. Was an LFC threshold applied for the DEGs, and if not, why?

We chose to have a log fold change threshold of 1 for both proteome and transcriptome in the volcano plots of figure 4 A and B. This threshold is only applied at the volcano plot-level, in order to limit the reported genes/proteins in the figure. Significantly different genes/proteins with a log fold change smaller than 1 are also colored in a lighter shade. All significantly different genes/proteins can be found in supplementary tables S13-14.

7. A reasonable discussion of the limitations of this study (subject homogeneity, tissue homogeneity, samples obtained from an intervention study that are expected to impact the datasets, etc. etc.) is important for providing guidance in interpreting the findings in this paper.

We have now added a limitations paragraph in the discussion of the manuscript. Pages 25-26, lines 575-589.

"Whilst this study represents the largest omic analysis of intact human single muscle fibers to date, it is not without limitations. We isolated skeletal muscle fibers from a relatively small and homogenous sample of participants and from a single muscle (vastus lateralis). In this respect, it is impossible to rule out the existence of specific fiber populations in different muscle types and at the extremes of muscle physiology. For example, we cannot rule out that a sub type of ultra-fast fibers (e.g. pure type 2X fibers) may become apparent in highly trained sprint and/or power athletes[78] or during muscle disuse[66,79]. Furthermore, our limited participant pool did not allow us to investigate sex differences in fiber heterogeneity, in which there are known differences in fiber type proportions between males and females. Furthermore, we were unable to perform transcriptomics and proteomics on the same muscle fibers or in samples from the same participants. As we and others continue to optimize single-cell and single muscle fiber omics-analyses towards ultra-low sample input (as demonstrated here with the analysis of fibers from nemaline myopathy patients), the possibility of combining multi-omics (and functional) approaches in a single muscle fiber is tantalizingly close."

Minor Points

1. Language-wise, the term used to describe the heterogeneity in PC1 is somewhat awkward and unclear during a first reading - "MYH-based fiber type-independent" comes up many times in the paper and is hard to understand. A better phrasing would perhaps be "novel source of heterogeneity" or something like that, as long as it is well-defined at the beginning.

We agree that the terminology "MYH-based fiber type-independent" might be awkward. However, we wanted the term to be as descriptive as possible throughout the manuscript, and while it might be long, we believe it is the least confusing option to refer to the heterogeneity observed.

2. Figure 1, panels I-J would be enhanced with feature plots that show the blended expression of MYH1 and MYH2 so that readers can visualize exactly which fibers express either or both.

We have included UMAP plots coloured by the blended expression of MYH1, MYH2, and MYH7 in figure 1 S4I-J (page 9, lines 187-188).