Molecular Cell

Transcription Factor-Directed Re-Wiring of Chromatin Architecture for Somatic Cell Nuclear Reprogramming Toward Trans-differentiation

--Manuscript Draft--

Manuscript Number:	MOLECULAR-CELL-D-19-00392R1
Full Title:	Transcription Factor-Directed Re-Wiring of Chromatin Architecture for Somatic Cell Nuclear Reprogramming Toward Trans-differentiation
Article Type:	Research Article
Keywords:	Chromatin architecture, transdifferentiation, MYOD, master transcription factor, cell identity, CTCF, insulated neighborhoods
Corresponding Author:	Pier Lorenzo Puri, PhD, MD UNITED STATES
First Author:	Alessandra Dall'Agnese
Order of Authors:	Alessandra Dall'Agnese
	Luca Caputo
	Chiara Nicoletti
	Julia di Iulio
	Anthony Schmitt
	Sole Gatto
	Yarui Diao
	Zhen Ye
	Mattia Forcato
	Ranjan Perera
	Silvio Bicciato
	Amalio Telenti
	Bing Ren
	Pier Lorenzo Puri, PhD, MD
Abstract:	MYOD-directed fibroblast trans-differentiation into skeletal muscle provides a unique model to investigate how one transcription factor (TF) reconfigures the three-dimensional chromatin architecture to control gene expression, which is otherwise achieved by the combinatorial activities of multiple TFs. Integrative analysis of genome-wide high-resolution chromatin interactions, MYOD and CTCF DNA-binding profile and gene expression revealed that MYOD directs extensive re-wiring of interactions involving cis-regulatory and structural genomic elements, including promoters, enhancers and insulated neighborhoods (INs). Re-configured INs were hot-spots of differential interactions, whereby MYOD binding to highly constrained sequences of IN boundaries and/or inside INs leads to alterations of promoter-enhancer interactions to repress cell-of-origin genes and to activate muscle-specific genes. Functional evidence shows that MYOD-directed re-configuration of chromatin interactions temporally preceded the effect on gene expression and was mediated by direct MYOD-DNA binding. These data illustrate a model whereby a single TF alters multi-loop hubs to drive somatic cell trans-differentiation.
Suggested Reviewers:	Jeff Dilworth jdilworth@ohri.ca Expert in epigenetic and skeletal myogenesis
	Stephen Tapscott stapscot@fhcrc.org

	He is an expert in skeletal muscle myogenesis, in particular, MYOD's role in skeletal muscle differentiation and skeletal muscle reprogramming.
	Vittorio Sartorelli sartorev@mail.nih.gov Vittorio Sartorelli is an expert in the epigenetic regulation of skeletal myogenesis.
	Andrew Lassar andrew_lassar@hms.harvard.edu Expert in MYOD-driven skeletal myogenesis
Opposed Reviewers:	Robert Tjian University of California Berkeley Scientific conflict of interest
	Erez Aiden-Lieberman Baylor College of Medicine Scientific conflict of interest
	Leonid Mirny Massachusetts Institute of Technology Scientific conflict of interest
	Victor Corces Emory University Scientific conflict of interest
	Brian Dynlacht New York University brian.dynlacht@nyumc.org
	Denny Reinberg New York University danny.reinberg@nyumc.org
	Benoit Bruneau Gladstone Institutes bbruneau@gladstone.ucsf.edu Scientific conflict of interest
Additional Information:	
Question	**Response**
Does your manuscript report new large-scale datasets?	
Does your manuscript report custom computer code or introduce a new algorithm?	

Cover Letter

Pier Lorenzo Puri, M.D.
Professor
Development, Aging, and Regeneration Program

La Jolla, February 27th 2019

Dear Dr. Plosky,

We would like to submit to Molecular Cell our manuscript entitled "**Transcription Factor-Directed Re-Wiring of Chromatin Architecture for Somatic Cell Nuclear Reprogramming Toward Trans-differentiation**", as Research Article.

Somatic cell trans-differentiation into another somatic cell type is a key issue in regenerative medicine. While transcription factor (TF)-directed somatic cell reprogramming toward another lineage has been reported, the mechanism by which TFs re-organize the 3D genome architecture to coordinately activate and repress specific subsets of genes is currently unknown.
MYOD-directed reprograming of somatic cells into skeletal muscle is a remarkable and unique example of trans-differentiation induced by the ectopic expression of a single TF. However, MYOD-mediated regulation of gene expression has been so far studied in a linear way - that is, by associating MYOD binding at enhancers and promoters to target genes based on the linear sequence of the DNA.
We exploited the ability of MYOD, master regulator of skeletal myogenesis, to reprogram fibroblasts to skeletal muscle cells as an experimental paradigm to characterize the mechanisms used by transcription factors to reprogram somatic cells. By combining high resolution Hi-C, ChIP-seq and RNA-seq with 3C analysis and deadCas9-mediated functional challenging, we show that MYOD directs an extensive re-wiring of chromatin interactions, involving cis-regulatory and structural elements of the genome, to repress cell-of-origin genes and activate tissue-specific genes. In particular, we observed that MYOD frequently binds to and alters the boundaries of sub-TAD structures, referred to as insulated neighborhood, at highly constrained genomic sequences enriched in pathogenic variants.
Either RNAi-mediated depletion or deadCas9-directed blockade of DNA binding specifically eliminated MYOD-mediated alterations of chromatin interactions and related local changes in gene expression.
Of note, we determined that MYOD-directed re-configuration of chromatin interactions preceded the effect on gene expression, thereby establishing the temporal relationship between the DNA binding and the alterations in chromatin interactions instigated by one TF that induced somatic nuclear cell reprogramming toward a specific lineage.

As suitable reviewers of our manuscript we would like to propose:

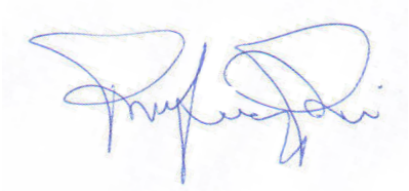
- Jeffrey Dilworth, Ottawa Hospital Research Institute, Ottawa, Canada. jdilworth@ohri.ca
- Andrew Lassar, Harvard Medical School. Boston, USA. andrew_lassar@hms.harvard.edu
- Vittorio Sartorelli, Laboratory of Muscle Stem Cells and Gene Regulation, NIAMS, Bethesda, USA sartorev@mail.nih.gov

- Stephen Tapscott, Fred Hutchinson Cancer Research Center, Human Biology and Clinical Research Divisions, Seattle, USA stapscot@fhcrc.org

We kindly ask to exclude from the review process:
Benoit Bruneau, Gladstone Institute, San Francisco, USA
Leonid Mirny, Massachussets Institute of Technology, Boston, USA
Erez Lieberman-Aiden, Baylor College of Medicine & Rice University, Houston, USA
Robert Tijan, University of California, Berkeley, USA
Danny Reinberg, NYU Langone School of Medicine, NY, USA
Brian Dynlacth, NYU, NY, USA
Victor Corces. Emory University, Atlanta, USA

We believe that the important findings described above will be interesting to a broad scientific community and hope that you will consider our work exciting and suitable for publication in *Molecular Cell.*

Your Sincerely,



Pier Lorenzo Puri, M.D.
Professor, Development, Aging and Regeneration Program

View Letter

Date:	Mar 20, 2019
To:	"Pier Lorenzo Puri" lpuri@sbpdiscovery.org
cc:	bplosky@cell.com
From:	"Molecular Cell Editorial Office" molecule@cell.com
Subject:	Editor's Decision on MOLECULAR-CELL-D-19-00392

Mar 20, 2019
RE: MOLECULAR-CELL-D-19-00392
"Transcription Factor-Directed Re-Wiring of Chromatin Architecture for Somatic Cell Nuclear Reprogramming Toward Trans-differentiation"

Dear Lorenzo,

Thank you for the opportunity to consider your manuscript at Molecular Cell. I've heard from two reviewers who both found the manuscript interesting and are generally supportive of considering a revised version for Molecular Cell. Reviewer 1's suggestions, while rather lengthy, are mainly aimed at providing greater clarity - both in the organization and presentation of the findings and to have more clarity about some of the experiments. Reviewer 2 is very positive and has one major experimental suggestion, which is to validate MyoD-dependent looping by an orthogonal approach.

We hope that the comments below will prove constructive as your work progresses, and I would be happy to discuss any potential plans for a revision once you've had a chance to consider the points raised in this letter.

Our typical timeframe for revisions is two to three months, and our general policy is that papers are considered through only one revision cycle, so any revised manuscript will be either accepted or rejected on the basis of the reviewers' comments (but not returned for further work).

Your revised paper should be no longer than **45,000 characters including spaces and figure legends, with no more than 7 figures and/or tables**. Please note that our article length guidelines have been recently adjusted to reflect that references are no longer counted towards the overall character count of the manuscript. You can visit our author's information page for more details concerning stylistic and formatting guidelines. Figures must be prepared in accordance with the Cell Press Data Processing Policy, which you can find here along with other details for preparing figures. For example, if you have digitally eliminated irrelevant or superfluous lanes from a gel or blot image, you must indicate the position of the deletion with a line or a space, and explain the manipulation in the figure legend. If you need to access your username and password, you can do so at https://www.editorialmanager.com/molecular-cell/.

As you make your revisions, I'd like to make you aware of a couple of changes that we are implementing at Molecular Cell.

1. Please keep in mind that, should we publish your paper, we will require that it be in compliance with our publication guidelines, including use of the STAR Methods format for reporting experimental procedures, methods, and analysis. At this stage, we require that your revised manuscript follows the STAR Methods format so both the editors and the reviewers are able to review the details of the experimental procedures and consider a version of the manuscript that closely reflects what would potentially be published. Cell Press introduced the STAR Methods format to help improve the rigor in reporting methods and resources for reproducibility. This section replaces the Experimental Procedures and Supplemental Experimental Procedures sections and does not count toward the main-text character limits. For detailed instructions on STAR Methods and a template for the Key Resources Table, see our STAR Methods webpage. Please contact me if you have any questions about restructuring your manuscript using the STAR Methods format.
2. To ensure best practices in data transparency and archiving, we are requiring authors to submit their unprocessed and uncompressed imaging data (microscopy as well as gels and blots) to their editor through Mendeley Data. This source data will only be required if and at the time that authors are invited to submit a revised paper for Molecular Cell. We also encourage, but do not require, submission of other non-imaging data. These original datasets will only be used for internal evaluation. However, to promote data transparency with the community, we also encourage authors to publish their source datasets alongside the paper. Click here for more detailed instructions on how to submit source data to your editor and publish it alongside your paper.
3. Should your manuscript be accepted for publication in the future, we'll encourage you to contribute any of these optional features:
    1. **Figure360:** Create a narrated, animated version of one of your figures that helps the reader zoom in on the most important take-home message in a matter of minutes. The video should contain data and panels from only one figure, and include minimal introduction. For guidelines and examples, please click here.
    2. **Methods Videos**: We encourage you to make a Methods Video for your paper if you report any methods that are challenging or nuanced, or if you have an experimental setup that is hard to describe. These videos are short (≤1 min) and are intended to improve reproducibility and transparency. For examples and guidelines, go here.

Best wishes,
Brian

Brian Plosky, Ph.D.
Editor-in-Chief, Molecular Cell

Response to Reviewers' Comments

We thank the reviewers for their valuable suggestions and comments, which have helped us to improve our manuscript. We were very pleased to see that both reviewers deemed our work as "exciting" and "solid", and that "nicely complement current hypotheses on the field" and "provides several novel aspects to our understanding of the role for transcription factors in establishing cell-specific organization of chromatin". However, while Reviewer #2 found that our results are "clearly presented", Reviewer #1 requested to improve the clarity of the data description in the text and recommended to revise length and structure of the manuscript. Reviewers also asked for additional explanations of some of the experiments performed, as well as additional analyses and experiments to help support our model.

We have modified the introduction and the result sections of the text according to reviewer indications. We have also improved the clarity of the description of specific experiments and made few modifications to the data presentation in the figures, including the inclusion of data results that were referred to as "data not shown" in the original submission, as recommended by the Reviewers. Finally, we have performed DNA-FISH to validate with an independent approach our Hi-C results, as requested by Reviewer 2.

A point-by-point response (blue font) to reviewer comments (black font) can be found below.

Reviewer comments:

Reviewer #1: The manuscript by Dall'Agnese et al., investigates the role of transcription factors (TF) in 3D chromatin organization and transcriptional control. By employing an in vitro differentiation system, the authors systematically address this question by performing Hi-C, ChIP-seq and RNA-seq at crucial timepoints of the differentiation process. The authors claim that the uniqueness of their approach relies on the fact that only one TF (MYOD) is required to promote the differentiation of human fibroblasts into myotubes. This fact allows, in principle, the elucidation of the role of that single TF in inducing major 3D-conformational and transcriptional changes.

The manuscript is broadly divided into two distinct sections. The first part focuses on a genome wide analysis of the differentiation process at the level of chromatin organization and transcription. On the second part, functional assays are performed, which validate the previous findings and provide important insights into how transcription is achieved. The authors report specific examples of different types of enhancer-promoter interactions mediated by MYOD. Specifically, one example is dependent of active transcription and another mediated by CTCF. Notably, such interactions appear to be reversible as the authors highlight after manipulation of the stimulus dosage. Furthermore, the authors extrapolate some of their findings on human cells into mouse, highlighting the

evolutionary conservation of the identified mechanisms. In general, the results nicely complement current hypotheses on the field.

Overall, the methodology and experiments of the study look solid. There are some concerns, however, related to data interpretation. Another important criticism of the manuscript relates to its density, length and its difficulty to read in certain sections.

RE: We thank the reviewer for appreciating our work. The reviewer noted that the methodology and experiments of our original submission were solid, but expressed concerns about our interpretation of some of our data and the difficulty in following certain parts of the manuscript, because of the density and sometimes redundancy of the data. To address these concerns, we modified the manuscript following the reviewer' suggestions as detailed in the answers to the reviewer's comments below.

Following, I provide a list of comments that the authors should take in consideration:

- As mentioned above, the manuscript is divided into two parts, the first mainly descriptive and the second more functional. There is a clear imbalance towards the first section, which includes a very long description of their biological system and a plethora of analysis that, in most cases, appear quite redundant.

I feel that the manuscript would clearly benefit from a more concise description of the results in certain sections and of a more careful thinking of to which audience the manuscript is targetting. Some of the findings might be of great interest for the 3D chromatin regulation community and, in that sense, there is no need to provide five dense pages of results explaining the differentiation system in detail. Similarly, an entire section is dedicated to present the results from Figure S5, although they can be easily merged with the previous section to make it more clear and concise. Another example can be found on the separate sections on insulated neighborhoods, covering chromatin contacts and transcription, which again are highly complementary and could be merged.

RE: We thank the reviewer in expressing his/her concerns regarding the text of the manuscript which helped us write a more concise, straight to the point and clear text. The reviewer correctly noted that our manuscript is divided into two sections. The first section describes the 3D chromatin changes associated to MYOD DNA binding and the associated changes in gene expression during MYOD-mediated myogenic commitment and differentiation. The second part consists of functional experiments proving that MYOD is indeed the transcription factor that mediates at least some of those changes in 3D chromatin interaction regulating gene expression.
We are indeed aware of this potential concern, as this manuscript is directed to a broad audience, which includes colleagues working in fields quite distant in terms of background and technologies used. While this clearly emphasizes the general relevance of our work, we believe that for this reason our data should be made accessible to such a broad audience, by providing enough information on background and technologies required for optimal data interpretation and overall understanding of our work. Thus, in the first part of

the manuscript an audience specialized in epigenetics and gene expression needs to be informed on “cell biology” of MYOD-directed reprogramming of somatic cells into skeletal muscle (Figure 1), while cell and developmental biologists need to be provided with an essential background of 3D chromatin interactions, Hi-C technology and related recent discoveries (i.e. TAD and INs) (Figure 2-4).
Regardless, we note that the data reported in the first part of the manuscript are novel and generated with unprecedented approaches that would invaviably require a detailed description. Indeed, we report on the first comprehensive analysis of the changes in 3D chromatin organization observed during somatic cell trans-differentiation that are associated with the binding of a master transcription factor, by exploiting a top-down approach: from the largest and most stable topological structures to the smallest and most variable chromatin interactions.
We argue that providing an accurate description of the experimental setting and data in the first part of the manuscript is necessary for a coincise description of the functional data in the second part of the manuscript. This is why the first part will necessarily be longer than the second one.
Having said that, we agree with the reviewer that some sections in the first part of the original version of our manuscript were too long and redundant, and that a more coincise description would help the readership to focus on key data. Likewise, we agree that providing a clearer presentation of the central biological question we are pursuing was an important modification to make.
Following this reviewer recommendation, we have made modifications to the text of our manuscript, by merging sections and by eliminating redundant and unnecessary sentences and figure panels.
We acknowledge that after this extensive re-hauling of the text, the revised manuscript reads more coincise and definitely accessible to the broad target audience.

In addition, the introduction section lacks a clear presentation of the biological question that the authors are pursuing. The motivations exposed at the end of paragraphs two and five are numerous and redundant. The section would benefit from exposing a more clear aim.

RE: The reviewer raised a proper point. We have replaced the numerous and redundant motivations for our work in paragraph two with one clear motivation. We have substituted the numerous and redundant motivations at the end of paragraph five with a clear presentation of the biological question we are pursuing.

- The authors state that the single expression of MYOD is sufficient to induce the differentiation into myotubes. Following that logic, the authors claim that they have a unique model to investigate the action of a single TF.
In my opinion, this is not entirely correct. The overexpression of MYOD alone commits the cells to a new fate, but they do not differentiate in growth media (GM). It is the presence of the differentiation media (DM) what ultimately triggers terminal differentiation. It is reasonable to think that MYOD is not doing the job alone, but in cooperation with

external signaling molecules that might induce the expression of other transcription factors that also contribute to the process.
That would mean that the differences on comparing "empty-GM vs MYOD-GM" can be attributed to MYOD, but not on "MYOD-GM vs MYOD-DM", where changes are resulting from components of the differentiation media (in cooperation with MYOD).
Luckily enough, all follow-up functional experiments were performed at loci where chromatin changes are observed at the commitment phase (empty-GM vs MYOD-GM), so the main authors´ conclusions hold true.
Nevertheless, it is important to make this distinction clear across the entire manuscript in introduction, results and conclusions.

RE: We agree with the reviewer on this point, and indeed we did not claim that MYOD is sufficient to induce the differentiation into myotubes. We rather claim that MYOD expression is sufficient to initiate the process of somatic cell reprogramming into another lineage and that terminal differentiation requires exposure to specific conditions (DM) that entail the cooperative activity of additional transcription factors. Still, MYOD-mediated conversion of fibroblasts into skeletal muscles provides a unique model in which one single transcription factor is sufficient to initiate and drive somatic cell reprogramming, which is otherwise achieved by the co-expression of multiple transcription factors. The epigenetic basis of this unique feature of MYOD are well accounted by the data reported in our manuscript, as they show that: 1) ~60% of the differentially interacting bins during reprogramming (considering both commitment and differentiation) differentially interacted during lineage commitment (Figure 2D); 2) changes in chromatin interactions largely precede changes in gene expression (Figure 7); 3) ~50% of changes in chromatin interactions are promoted by MYOD DNA binding, either directly or indirectly (Figure 2F and Figure S3). Importantly, we also show that the residual amount of changes in chromatin interactions that did not involve direct or indirect MYOD DNA binding, might actually involve transcription factors whose expression is upregulated by MYOD and are known to cooperate with MYOD in the activation of skeletal myogenesis (Figure 2F and S3). These data are consistent with a model, whereby MYOD operates as a master transcription factor, by instigating changes in 3D chromatin architecture that are sufficient to promote the cooperative action of other transcription factors toward completion of somatic cell transdiferentiation. We discuss this concept within current evidence showing that no other master transcription factor is able to initiate somatic cell reprogramming when expressed alone (see discussion section).

- The literature of the manuscript is not appropriate in some cases. A particular example is found in the introduction, where the authors discuss about the importance of 3D chromatin organization in transcriptional control, but they do not cite important literature on the topic. Studies by Duboule, Spitz, Furlong or Mundlos labs, among others, have contributed to a better understanding of the biological relevance of 3D chromatin organization in vivo and are not properly referenced. Important work on chromatin organization by the Aiden lab is also not cited.
The authors should make an additional effort to provide a more comprehensive and rationale citation of background work.

RE: We thank the reviewer for pointing out that we need to reference the important work from the laboratories of Duboule, Spitz, Furlong, Mundlos and Aiden, which we have not included previously, as well as other important works that have initially escaped our attention.

- There are multiple statements across the manuscript without supporting data (indicated as "data not shown"). In some instances, like in page 13, such statements hold important conclusions. In those cases, the appropriate data should be provided to facilitate a proper evaluation of the claims.

RE: We have now included all the data in the revised version of the manuscript.
They include:

a) results showing that differentially expressed genes within the same TAD tend to be all up-regulated or all down-regulated (figure 2C);

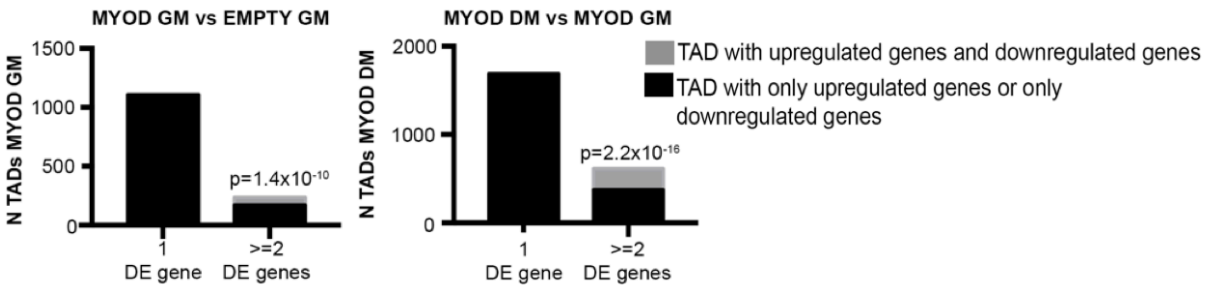
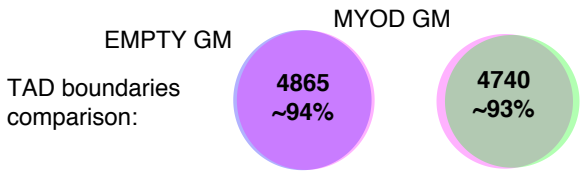


Figure 2C: Number of TADs with one or more DE genes. Black represents the TADs whose differentially expressed genes are all upregulated or all downregulated, while grey represents the TADs containing upregulated genes and downregulated genes. LEFT: gene expression comparison between EMPTY GM and MYOD GM. TADs used were identified in MYOD GM. RIGHT: gene expression comparison between MYOD GM and MYOD DM. TADs used were identified in MYOD DM. pvalue represent the significant prevalence of TADs with two or more differentially expressed genes that were either all upregulated or all downregulated compared to TADs that have both upregulated genes and downregulated genes. pvalue was calculated using the two-sided exact binomial test.

b) a more detailed sentence explaining the results that suggest there is no difference in TAD boundary location during MYOD-mediated conversion: “boundary location did not significantly differ during MYOD-mediated fibroblast conversion into myoblasts or myotubes (Figure 2B), since the percentage of overlap of TAD boundaries between samples (~90%) was similar to the percentage of overlap of TAD boundaries between biological replicates (~90%)”



Only for Reviewer Figure 1: Venn diagram of common or unique boundaries between EMPTY GM and MYOD GM (left) or MYOD GM and MYOD DM (right).

c) Results in Figure 13 that the reviewer specifically requested in his/her comments are now shown in figure S3B and S3C. We report below the section and the figures.

"To determine the extent to which the differential chromatin interactions were orchestrated by MYOD, we considered altered interactions directly bound as well as indirect events potentially generated by MYOD DNA binding, as illustrated in Figure S3A. According to this model the initial chromatin alterations are conceivably caused by MYOD binding to the DNA (Bin2, "MYOD-dependent and MYOD-bound"). MYOD binding to the DNA could promote the interaction with another bin that may be bound by MYOD ("MYOD-dependent and MYOD-bound") or not (Bin3, "MYOD-dependent and directly interacting"). MYOD binding to the DNA could also dis-engage previously interacting bins (i.e., Bin1 and Bin2) thereby generating free bins (i.e., Bin1) available for new interactions with other bins (Bin?), either bound by MYOD ("MYOD-dependent and MYOD-bound") or not (Bin?, "MYOD-dependent and indirectly interacting"). Moreover, some altered chromatin interactions can form independently on the initial chain of differential interactions triggered by MYOD DNA binding (others). These differential interactions could be mediated by other TFs, whose expression might be also regulated by MYOD (Figure S3A, co-operating TF)."

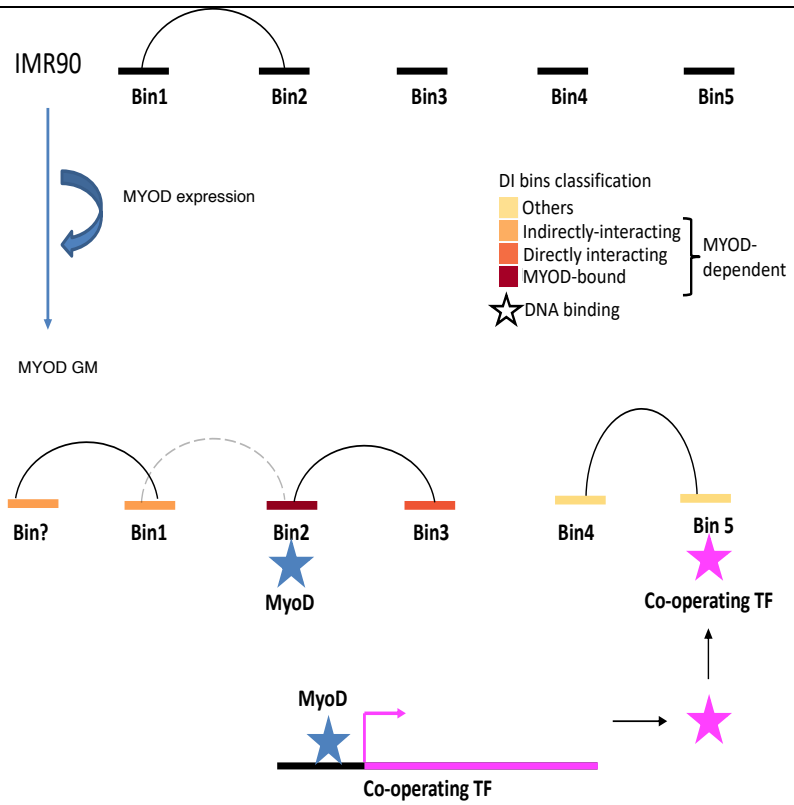


Figure S3A, Schematic representation of MYOD-domino effect on differential chromatin interactions.

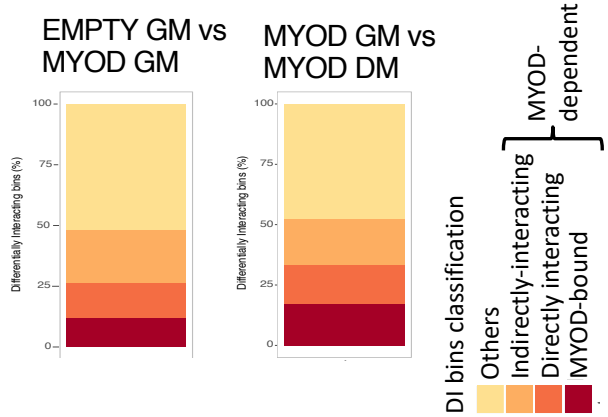


Figure 2F, Percentage of DI bins bound by MYOD (red), DI bins directly interacting with MYOD-bound bins (orange), DI bins indirectly interacting (dark yellow), others (yellow) – see figure S3 for the schematic representation of this classification.

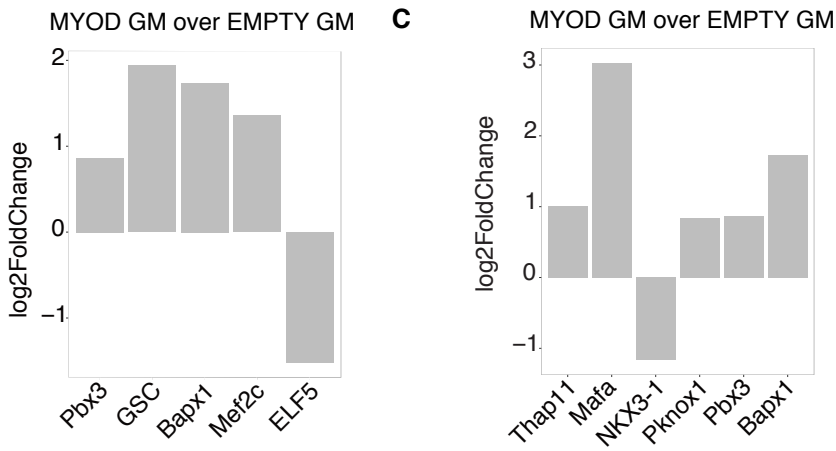


Figure S3B, Differential gene expression by RNA-seq of TFs whose DNA binding motif is present around the TSS contained within indirectly-interacting bins and other bins (right). indirectly-interacting bins and other bins refer to Fig 2F EMPTY GM vs MYOD GM.
Figure S3C, Motif analysis at the summit of MYOD peaks (+/-50bp) that mapped at differentially interacting IN boundaries during MYOD-mediated commitment or differentiation.

- An average of 470 million contacts are generated per Hi-C map. However, there is no information on the quality parameters of individual libraries, more than the maps in Fig 2. Although the maps look overall of good quality, it would be important to provide a table containing detailed statistics for each experiment, to check for consistency.

RE: We have performed systematic quality controls on our data that we did not include in the original version of the manuscript, but that is now reported in new Table S1. We initially reported the reproducibility of interaction maps between biological replicates using two independent methods: Bing Ren's published method (Dixon et al., 2015) and HiC-spector (Yan et al., 2017). However, we also performed the following analyses:

1) percentage of mapped reads (for both mates in a read pair)
2) percentage of reported pairs (removal of unmapped pairs, multiple pairs alignments, low quality pairs, not reported pairs and pairs with singleton - % considering the total number of reads)
3) percentage of valid putative interaction pairs (removal of dangling ends, fragments with no restriction site, self circles etc - % considering the number of reported pairs)
4) percentage of unique read pairs (removal of duplicates - % considering the valid putative interaction pairs)
5) number of unique read pairs
6) percentage of cis-read pairs
7) number of long-range cis-read pairs
8) percentage of trans read pairs

These additional analyses further prove the high quality of our data. We have included these statistics in a combined table – new Table S1.

Table S1: Quality control of Hi-C libraries

Sample	Replicate	Total Reads 4kb resolution analysis	R1 % mapping	R2 % mapping	% Reported Pairs*	% Valid Putative Interaction Pairs**	% Unique Read Pairs***	N Valid Unique Reads Pairs	% Cis Read Pairs	Cis Read Pairs	% Cis Long Range Read Pairs	Cis Long Range Read Pairs	% Trans Read Pairs	Trans Read Pairs
IMR90 Empty vector GM	Rep1	645,840,669	96	91.5	70.2	90.4	94.0	381,646,920	86.7	330,744,403	58.5	223,302,744	13.3	50,902,517
	Rep2	593,125,998	95.9	91.4	69	91.0	94.0	349,387,497	84.4	294,819,267	61.7	215,493,824	15.6	54,568,230
IMR90 MYOD vector GM	Rep1	755,458,186	96.1	91.9	70.6	92.1	94.0	456,813,369	83.7	382,334,279	58.1	265,620,737	16.3	74,479,090
	Rep2	686,395,066	95.7	91.6	67.9	91.8	94.0	405,134,907	82.9	335,663,651	61.9	250,819,874	17.1	69,471,256
IMR90 MYOD vector DM	Rep1	763,544,759	95.8	90.8	69.1	89.7	91.0	431,861,640	72.5	313,243,531	51.9	224,296,215	27.5	118,618,109
	Rep2	1,040,387,454	95.6	91.5	68.2	90.4	92.0	592,928,711	75.4	447,011,501	59.5	352,835,388	24.6	145,917,210
		4,484,752,132						2,617,773,044	80.9	2,103,816,632	58.6			

* Removal of unmapped pairs, multiple pairs alignments, low quality pairs, not reported pairs and pairs with singleton - % considering the total number of reads
** Removal of dangling ends, fragments with no restriction site, self circles etc - % considering the number of reported pairs
*** Removal of duplicates - % considering the valid putative interaction pairs

- While chromatin interactions are detected in Hi-C datasets using 4Kb bins, TADs are identified using 40Kb bins. Is there a specific reason for such discrepancy? The authors should justify here.

RE: We called TADs at various resolutions, but we report data on TADs called at 40kb resolution, because 40kb was the highest resolution that gave us reproducible TAD calls between biological replicates. We have added the following sentence in the TAD calling section of the Methods.
"We called TADs at various resolutions, but we report data on TADs called at 40kb resolution, because 40kb was the highest resolution that gave us reproducible TAD calls between biological replicates."

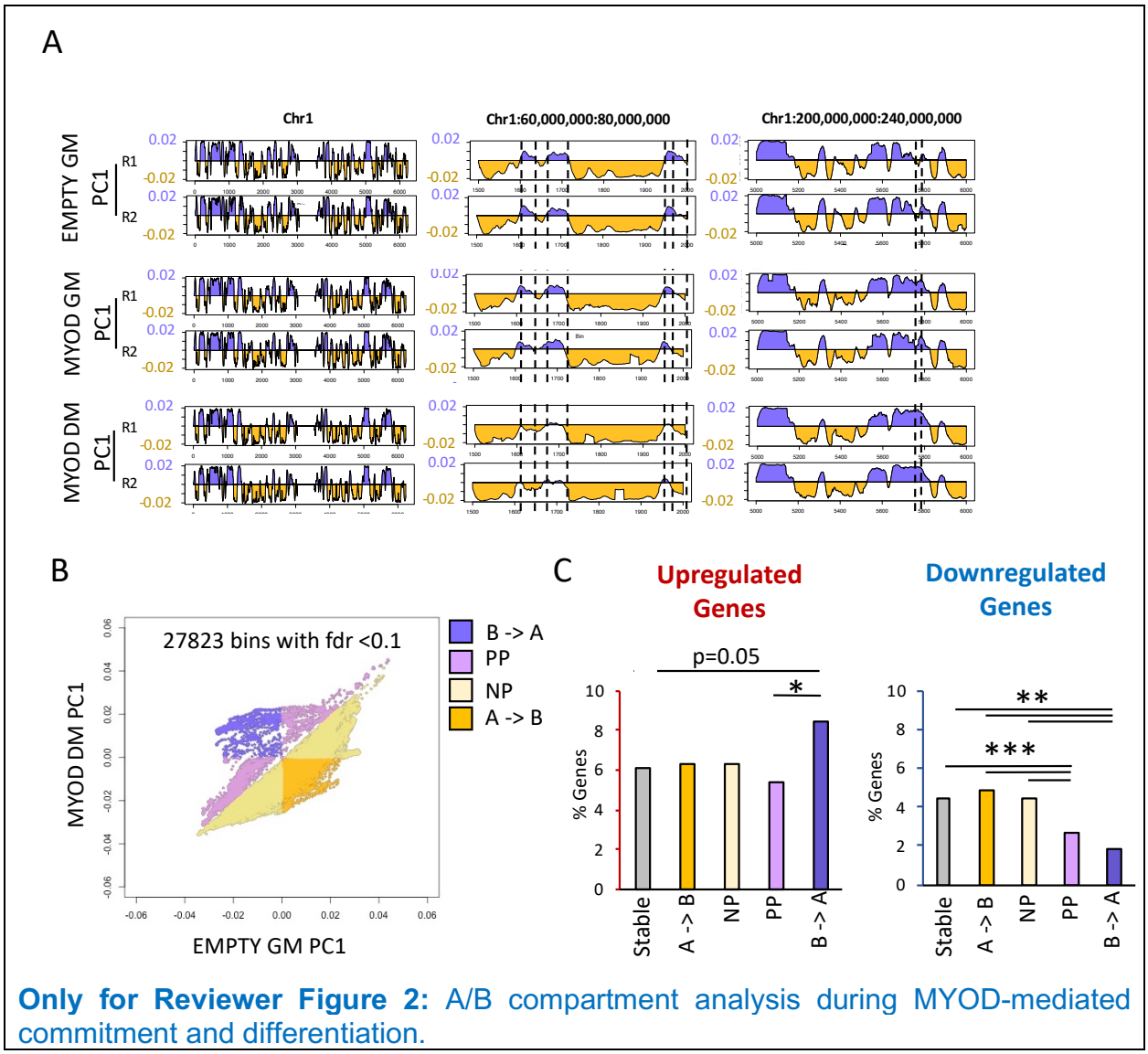
In addition, the authors state that TAD boundaries are not altered during fibroblast conversion into myoblast and myotubes. It would be useful to provide data supporting the conclusion (i.e. percentages). Indeed, the analysis depicted in Fig 2 contradicts the statement, as there are boundaries that change between conditions.

RE: We thank the reviewer pointing out that we should show the percentages of TAD boundaries in common between the experimental conditions. We have included in the rebuttal letter a Venn diagram of common or unique boundaries between EMPTY GM and MYOD GM or MYOD GM and MYOD DM in Only for Reviewer Figure 1 (see above). There is a negligible variability between experimental conditions; however, the percentage of common TAD boundaries between different experimental conditions was similar to the percentage of common boundaries between biological replicates (~90%). We therefore concluded that TAD boundaries were not altered during fibroblast conversion into myoblasts and myotubes.

For clarity we changed the following sentence:
"the location of TAD boundaries was highly not altered during MYOD-mediated fibroblast conversion into myoblasts and myotubes"
to
"boundary location did not significantly differ during MYOD-mediated fibroblast conversion into myoblasts or myotubes (Figure 2B), since the percentage of overlap of TAD boundaries between samples (~90%) was similar to the percentage of overlap of TAD boundaries between biological replicates (~90%)"

- Did the authors evaluate possible changes in A/B compartimentalization? It would be interesting to investigate this aspect.

RE: We have performed A/B compartment analysis and we detected some switches between A and B compartments (see figure below). We also found that there is a higher percentage of upregulated genes and a lower percentage of downregulated genes in the regions where the compartment switched from B to A as compared to regions whose PC1 value does not change. However, we did not find any correlation between MYOD binding and A/B compartment switch. While we decided not to include these results in the manuscript, which is already quite dense, we are happy to include these data below, for the reviewer's information.



Only for Reviewer Figure 2: A/B compartment analysis during MYOD-mediated commitment and differentiation.

A. A (purple) and B (yellow) compartments along chromosome 1 or specific portions of chromosome 1 in EMPTY GM, MYOD GM and MYOD DM. Results of each replicate (R1 and R2) is reported. Areas between dashed lines represent areas of A/B compartment switch between experimental conditions, but not biological replicates
B. Comparison of PC1 values between EMPTY GM and MYOD DM. For categories are represented in different colors: A to B switch (orange, A->B), B to A switch (purple, B->A), PC1 values decrease but no switch (yellow, NP), PC1 values increase but no compartment switch (pink, PP).
C. Percentage of genes located in the regions described in (B) that are upregulated (left) or downregulated (right). Chi-squared test was used for statistical analysis.

- The introduction would benefit from a brief explanation of TADs, subTADs and insulated neighborhoods (INs), as well as how they relate to each other.

RE: We have added a brief explanation of TADs, subTADs and INs and their relationship with each other in the Introduction, and eliminated the paragraph dedicated to the same explaination in the Result section, to avoid reduncancy.
We have however maintained the schematic drawing of this structures in the figures for clarity.

Also, the authors focus on analyzing INs extensively. It would be necessary to provide more details on how these domains are defined. How do the authors deal with nested domains? Are just the smallest possible domains considered or also larger ones containing them?

RE: As written in the original version of the manuscript "We defined altered INs as regions of DNA that contained at least one gene and whose boundaries were i) co-bound by CTCF in IMR90 and ii) showed differential interaction strength during MYOD-mediated commitment or differentiation (Figure 4A)". We did not select the altered INs to study based on size or if they are contained within other altered INs or if they contain other altered INs. We considered each altered IN as an entity on its own. We have clarified this important point in the text as follows:
"Each altered IN was considered as a separate entity, regardless its inclusion within larger altered INs or the presence of smaller altered INs inside it."

- The authors describe an interesting case at the ITGA-RDH5 locus, where MYOD seems to interact with CTCF to form a regulatory domain. Are CTCF motifs present at the summit of CTCF ChIP-seq peaks identified at the locus? An absence of them might suggest indirect binding mediated by MYOD, which might support the results of their functional experiments.

If CTCF motifs are present, do they follow the convergent-orientation, loop-extrusion theory?

RE: We find investigating CTCF motifs and their orientation of great interest. We used JASPAR[2018] to identify predicted putative CTCF binding sites with relative profile score threshold 80%. Please see below the DNA sequence of CTCF peaks at ITGA7-RDH5 locus. The summit is represented with bold white letter and blue highlight (**G**). The DNA sequences highlighted in green correspond to putative CTCF sites closest to the summit, while the DNA sequences highlighted in yellow correspond to putative CTCF sites that are relatively far from the summit.

On the DNA sequence of ITGA7 CTCF peak, JASPAR predicted three putative sites in the (-) strand:

1) **CCAGCAGGGGGTGCT**,
2) **GCTCCAGCAGGGGGTGCTC** (it overlaps with the DNA site above)
3) **TCTGAAGGTGGCCCC**

None of them overlap the summit of CTCF peak, but the first and second motifs are only ~10bp far from the summit.

On the DNA sequence of RDH5 CTCF peak, JASPAR predicted two putative sites with relative profile score threshold 80% in the (-) strand:

1. **CACATAGCTGGAGCC**
2. **TGACCTACAGGTGGCAGCC**

Interestingly **TGACCTACAGGTGGCAGCC** is located at the summit of CTCF peak and it is a partial CTCF canonical motif (CTCF canonical motif is **CCACNAGGTGGCAG**, as reported in Rao et al., 2014).

These results are of particular interest as, they imply a potential genetic code guiding cross-recruitment of MYOD and CTCF at IN boundaries, based DNA motif sequences and orientation. Although we acknowledge that this issue requires further studies (which are outside the scope of this manuscript), it is interesting to note that at the RDH5 locus, where the partial canonical CTCF motif is present, MYOD expression did not affect CTCF recruitment; however, at the ITGA7 locus, where a non-canonical CTCF motif is present, MYOD expression and MYOD DNA binding could affect CTCF recruitment (Figure 5H and Figure 6D).

The CTCF motif sequences and orientations at RDH5 and ITGA7 loci are show below:

ITGA7 CTCF peak, strand (-), summit is **G**:
CCTCCTC**TCTGAAGGTGGCCCC**TGGTTCCGTCTCCCAGAGCCAAGCTGGGGCCTT
TCCCAGAGGGCCTGACTGCCTCACCCTGCTTTT**GCTCCAGCAGGGGGTGCTC**TGC
TGCTGG**G**GGGCGGGGGGTATGTGAGAGGCCAGGCACCTGCTCAGTCCCTAGCTT
TTGAGTTGCAGGTGGCCTGCCTTAGCACTCACTGATGAAAAAAACTTCTTG

RDH5 CTCF peak, strand (-), summit is **G**:

AGAAGGCGCCCAAAGCCTGAGTCACAGCCGGTGATGAAGACAAAGGCATTGCTG
GCGGGCAGGCTCTGCCGGTCCCTGAGCAACCACAGCACTGCCCAGAGTAAGGCA
CCCAGCAGAAGAGGCAGCCACATAGCTGGAGCCCAAGTGACCTACAGGTGGCAG
CCTAGGCTGGGCAGAAAAACTTGTCCAGTTTACTGTGGCCCTCAAGCCCTTTCCCC
TAATACCCTCCCTAGCTGGAAGCATCTGGTGTGAGCATATTGAGAAATTATCTGGT
TGGCAGATAATTGG

- In their Dox-withdrawal experiments, did the authors assess the molecular identity of cells after stimuli removal? Do they go back to the initial state?

RE: The reviewer raised an interesting point. We have not assessed the molecular identity of the cells globally after Dox-withdrawal as this was beyond the aim of the presented experiments. However, given that our results show a decrease in MYOD expression (MYOD expression is not completely inhibited) and partial restoration of 3D chromatin organization, gene expression profile and CTCF recruitment at the selected loci, we assume that at the cells are not back at their initial state.
Overall, our results represent a proof of principle for the requirement of MYOD to maintain myogenic chromatin organization and myogenic identity during at the commitment stage, since decreasing its expression already affects chromatin organization, gene expression and CTCF recruitment.
We believe that this issue is of particular relevance in the field of regenerative medicine, in particular within the context of skeletal muscle regeneration. In this regard, we have referred to recent data from Goldhamer group showing that acute genetic deletion of MYOD *in vivo* alters the gene expression profile and compromises satellite cell differentiation and efficiency of muscle regeneration (Yamamoto et al., 2018).

- Across the entire manuscript, data from replicates is displayed as bars, although the number of replicates is low (3). It would be more appropriate to provide the data for each replicate as individual points.

RE: Representing data as mean +/- SEM is a well-accepted method of showing results and the reproducibility of experiments. We think that consistency in data representation is very important to help the reader follow easily the figures, without mentally switching from one data representation to another, and it improves the visual clarity of the figures, which are already full of objects. Showing each replicate as individual points throughout the paper would result in busier figures and some will be very hard to interpret when we plot the results for different conditions in the same graph. For all these reasons we prefer to keep the data representation as mean +/- SEM for all the experiments.

- Significance tests are missing in several graphs. Notably Fig 5E, 5H or 5D. As important conclusions are derived from there graphs, the data should be displayed.

RE: The reviewer is correct. We have added the results of the statistical significance in Figure 5E, 5H and 5D. While we often find a statistically significant change when comparing MYOD GM ON/ON and MYOD GM ON/OFF, in Fig 5E and Fig 5H left, we find a reproducible, but not statistically significant, change between these two conditions. However, since the difference between EMPTY GM and MYOD GM ON/ON is statistically significant, but the difference between EMPTY GM and MYOD GM ON/OFF is not, we conclude that there is a decrease of MYOD binding (Fig 5E) and CTCF binding (Fig 5H left) upon doxy removal from the media.

- Overall, the results delineate MYOD as an important factor for muscle fate commitment. It would important to discuss the findings in the context of current concepts such as pioneer TF factors or pre-established chromatin interactions.

RE: We thank the reviewer for raising this interesting point. Indeed, in the Discussion section, we discuss this issue in regard to the role of MYOD as master TF (mTF). It is currently unclear whether a mTF can also function as pioneer TF, and data from the literature would rather suggest that MYOD does not entirely fulfill the typical definition as pioneer TF – including evidence reported from Tapscot and Imbalzano groups showing that MYOD actually requires the action of the ubiquitous pioneer TF Pbx1/Meis to access chromatin at previously silent loci (Berkes et al. 2004; de la Serna et al. 2005; Fong et al. 2015). An extensive discussion of differences and similarities between master and pioneer TFs within the context of MYOD, cell lineage commitment and differentiation is reported in a review recently published in Mol Cell (Sartorelli and Puri 2019). We prefer to refer to this review, as we feel it is premature to speculate on the ability of MYOD to reconfigure 3D chromatin architecture as a pioneer TF.

- The MYOD transgene is repeatedly referred as MyoD. If it is not from mouse origin, it should be named in capital letters.

RE: We thank the reviewer for paying very close attention to our manuscript. Since the transgene is from mouse origin, we referred to it as MyoD.

Reviewer #2: This manuscript by Dall'Agnese et al. examines chromatin architectural changes directed by lineage specific master transcription factor MyoD during trans-differentiation of IMR90 fibroblasts into skeletal muscle progenitor cells. Using Dox-inducible system for MyoD expression, the authors transdifferentiated fibroblasts into skeletal muscle progenitor cells, which eventually differentiate to form myotubes. RNA-seq analysis in trans-differentiated myotubes and primary human myotubes revealed an over-lap in the modulation of key gene regulatory networks such as up regulation of myogenic regulatory genes and down regulation of fibrotic and inflammatory network genes. MyoD ChIP Seq showed enrichment of E-box motifs and a higher DNA sequence constraint with enrichment of annotated disease variant genomic bins. This study also provided a corroborating evidence for the genomic distribution of MyoD, specifically in the promoters and distal enhancers of differentially expressed genes. An intriguing observation the authors made is MyoD's enrichment at CTCF binding sites. Examining MyoD regulated chromatin interactions, the authors performed in situ Hi-C to show the existence of MyoD bound sites within TADs and at TAD boundaries. The authors also went on to prove that chromatin topological changes during IMR90 trans-differentiation is a consequence of MyoD DNA binding directly and indirectly. By integrating the MyoD ChIP Seq data with differential interactome maps, the authors provide evidence for the binding of MyoD at differentially interacting promoters and insulated neighborhoods. The authors provide evidence for these interpretations by reducing the expression of MyoD in the trans-differentiated cells, by siRNA knockdown of MyoD in C2C12 myoblasts. Using catalytically inactive Cas9 the authors prove that direct binding of MyoD is required for these chromatin interactions. Based on these findings, the authors propose a model whereby MyoD drives lineage reprograming through an alteration of chromatin loops.

This is an exciting manuscript that that demonstrates a role for lineage-specific transcription factors in mediating topological changes in chromatin architecture that establish cell specific interactions between insulated neighbors. Using an inducible expression system, they were further able to demonstrate that these chromatin changes are reversible. In addition, the use of the Cas9 system allowed them to demonstrate that the local insulated neighborhoods form independent of one another. Thus, this work provides several novel aspects to our understanding of the role for transcription factors in establishing cell-specific organization of chromatin. The authors have performed extensive analysis of the 3D-organization directed by MyoD and the results are clearly presented.

RE: We thank the reviewer for his/her detailed summary of our work. The reviewer found that our work is exciting and that it provides “several novel aspects to our understanding of the role for transcription factors in establishing cell-specific organization of chromatin”. We thank the reviewer for pointing out that we have performed extensive analyses and that our results are clearly presented.
We also thank this reviewer for having captured few small mistakes (see Other concerns/corrections) that escaped our attention.

Nevertheless several points should be addressed:

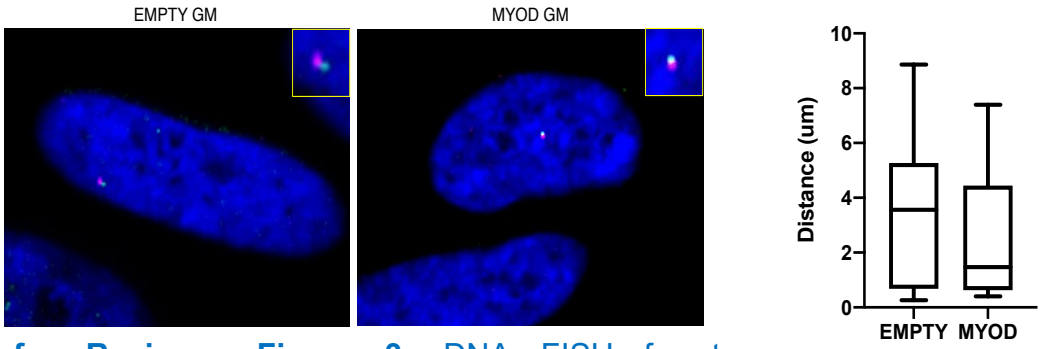
Major Points.
1) The data strongly suggests that MyoD alone can mediate a reversible change in chromatin interactions during lineage reprogramming. These results are all analyzed using the 3C technologies looking at a single view point. The authors should demonstrate the MyoD-dependent chromatin looping for a locus using a second technique such as DNA-FISH to further support the findings.

RE: The reviewer asked us to validate the changes in 3D chromatin interaction using another method, such as DNA-FISH. We could not perform DNA-FISH to investigate the interactions between the loci we deeply studied in the manuscript, because these loci are too close to each other. Given the large size of DNA-FISH probes (which usually span ~100 kb) and the limit of microscopy resolution, these loci would always appear overlapping with each other. At this time, only 3C-based technologies (as shown in our manuscript) could be used to validate and quantify differential chromatin interaction identified by Hi-C at such high resolution.

Still, we have identified an illustrative differential chromatin interaction suitable for DNA FISH, based on the following criteria:

- distance between the differentially interacting bins was higher than 100kb
- at least one bin was bound by MYOD

This example corresponds to the increased interaction between chr22 50160001 50164000 and chr22 50584001 50588000 that we identified by Hi-C when we compared EMPTY GM and MYOD GM. We designed DNA FISH probes spanning 100kb (from Agilent) for these two loci and performed DNA FISH in EMPTY GM and MYOD GM samples. We found that the distance between the selected loci was reproducibly decreased upon MYOD-mediated commitment by DNA-FISH, thereby validating our Hi-C result – see data below.



Only for Reviewer Figure 3: DNA FISH for two representative loci (green and magenta) whose interaction is higher in MYOD GM as compared to EMPTY GM based on Hi-C data. One z slice is shown. Right: quantification of distance between the loci.

While these data would in principle confirm that interactions detected by Hi-C, we also note that they are biased by limitations intrinsic to DNA FISH, which can detect differential interactions between 100kb regions, and our differentially interacting 4kb bins consist only

a small fraction of those 100kb regions. Thus, we are reluctant to include these data as formal validation of Hi-C and 3C data. Considering also the very large amount of figures and data already presented in this work, we suggest that these data are not included in the manuscript, but are shown for reviewer vision only.

2) The authors suggest that MyoD would co-regulate genes present within a TAD without providing data. The authors should present the data showing this coregulation in a supplemental figure to support their rational that MyoD would alter interactions between promoters, enhancers and insulators.

RE: The reviewer is correct in stating that we should show the tendency of genes to be co-regulated when they belong to the same TAD. We have added these results in new figure 2C.

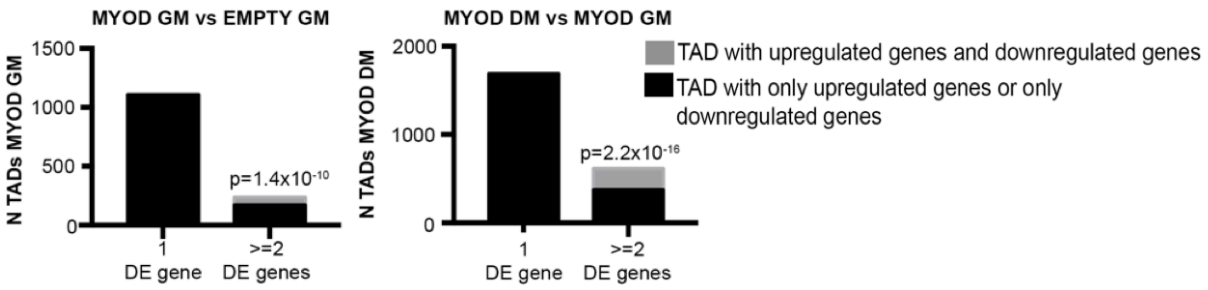


Figure 2C: Number of TADs with one or more DE genes. Black represents the TADs whose differentially expressed genes are all upregulated or all downregulated, while grey represents the TADs containing upregulated genes and downregulated genes. LEFT: gene expression comparison between EMPTY GM and MYOD GM. TADs used were identified in MYOD GM. RIGHT: gene expression comparison between MYOD GM and MYOD DM. TADs used were identified in MYOD DM. pvalue represent the significant prevalence of TADs with two or more differentially expressed genes that were either all upregulated or all downregulated compared to TADs that have both upregulated genes and downregulated genes. pvalue was calculated using the two-sided exact binomial test.

Other concerns/corrections

1) The authors should change the legend in Figure 1F from MyoD-GM and MyoD DM to MyoD-mediated commitment and MyoD-mediated differentiation (as in figure 2F).

RE: We have replaced MYOD-GM and MYOD-DM with MYOD-mediated commitment and MYOD-mediated differentiation, as correctly suggested by the reviewer, in previous Figure 1F and current Figure 1E.

2) In figure 1G, it would be informative to see the % of MyoD peaks in promoters, enhancers and TAD boundaries rather than having these different elements pooled together under the caption "Others".

RE: MYOD association to promoters, enhancers and CTCF-bound sites is presented in the initial figure 1F (now figure 1E). The goal of initial figure 1G (now figure 1D) is to show the low association of MYOD peaks to differential gene expression if we only consider the linear sequence of the DNA. This message cannot come across if we show MYOD binding to all promoters, all enhancers and all TAD boundaries, because MYOD can also bind promoters of non-differentially expressed genes as well as enhancers that have not been assigned to target genes at this stage of the manuscript progression– please, note that TAD boundaries are introduced later in the manuscript (Figure 2), so it was premature to analyze them in Figure 1. We therefore prefer to maintain the original flow of the figures, as it has been purposely designed to transition from one finding (low percentage of MYOD peaks bound at the promoter of differentially expressed genes; high amount of MYOD peaks at distal regulatory and architectural elements) to the overall purpose of the work (role of MYOD in promoting/altering long range chromatin interactions).

3) In figure 3I, it would be informative to incorporate the acetylation profile on the enhancer regions as well.

RE: We agree with the reviewer that for completeness we should show the acetylation profile on the enhancer. We have now added the H3K27ac ChIP-qPCR results on the enhancer region in figure 3I.

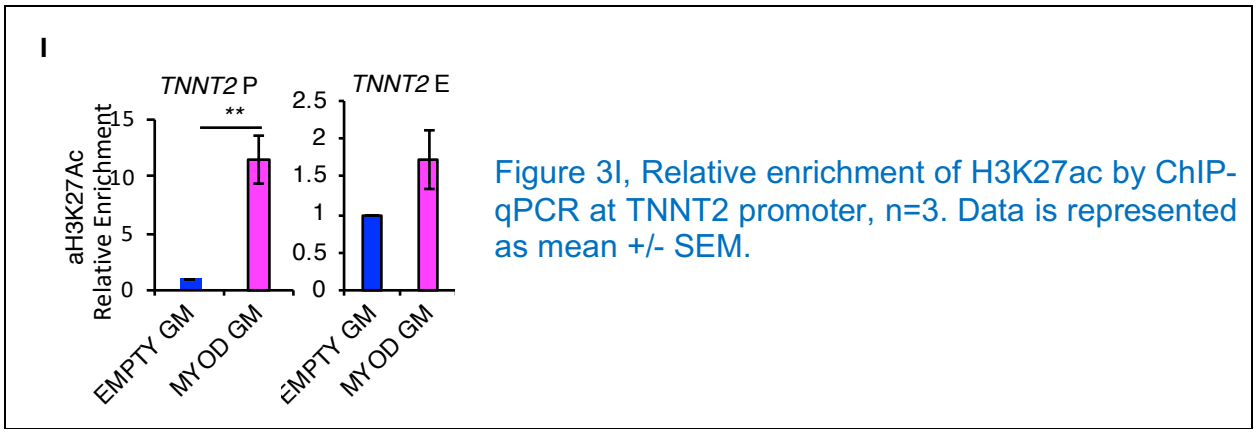


Figure 3I, Relative enrichment of H3K27ac by ChIP-qPCR at TNNT2 promoter, n=3. Data is represented as mean +/- SEM.

4) In page 16, paragraph 2, if the conclusion "2% altered interactions occurred between IN boundaries" is based on data from figure 3F. This has to be referred at the end of the sentence.

RE: The conclusion that "2% altered interactions occurred between IN boundaries" came from a mathematical calculation that did not refer to any figure panel. We realized that

this sentence is ambiguous and we therefore replaced it with the following sentence: “We found that only ~2% (1,332 or 1,595) of altered interactions accounted for changes in interaction strength between IN boundaries”. We have also added a figure panel to represent this information for clarity that is now in figure 4B (see below).

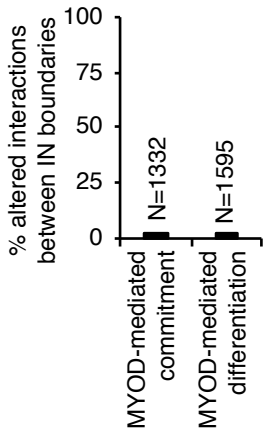


Figure 4B, Percentage of differential interactions corresponding to altered IN boundary interactions

5) In page 19, there is never figure 4L. This has to be corrected in text as well as in the figure legends of figure 5 and figure 6 where 4L is referred.

RE: We thank the reviewer for pointing out this mistake. We have corrected the references to the figure panels.

6) In page 85, the figure legend B "compared to hMB and hMB" should be corrected "compared to hMB and hMT".

RE: We appreciate that the reviewer read very carefully our manuscript. We have corrected the mistake the reviewer pointed out.

7) On page 13, listing of six categories of DIs begins with numbering for 1-3 (ie. 1)…., 2)…, 3)) but this does not continue for 4-6.

RE: We have added 4), 5), 6) to the list of the six categories of DIs.

References

Yamamoto, M., Legendre, N.P., Biswas, A.A., Lawton, A., Yamamoto, S., Tajbakhsh, S., Kardon, G., and Goldhamer, D.J. (2018). Loss of MyoD and Myf5 in Skeletal Muscle Stem Cells Results in Altered Myogenic Programming and Failed Regeneration. Stem Cell Reports *10*, 956-969.

Transcription Factor-Directed Re-Wiring of Chromatin Architecture for Somatic Cell Nuclear Reprogramming Toward Trans-differentiation

Alessandra Dall'Agnese[1,2,3,12,13], Luca Caputo[2,12], Chiara Nicoletti[2,4], Julia di Iulio[5], Anthony Schmitt[6,7], Sole Gatto[2,8], Yarui Diao[6], Zhen Ye[6], Mattia Forcato[4], Ranjan Perera[9,10], Silvio Bicciato[4], Amalio Telenti[5], Bing Ren[6,11], Pier Lorenzo Puri[2,13,14]

[1]Graduate School of Biomedical Sciences, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037, USA

[2]Development, Aging and Regeneration Program, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037, USA

[3]Whitehead Institute for Biomedical Research, Cambridge, MA, 02142, USA

[4]Department of Life Sciences, University of Modena and Reggio Emilia, Modena, 41125, Italy

[5]The Scripps Research Institute, La Jolla, CA 92037, USA

[6]Ludwig Institute for Cancer Research, La Jolla, CA 92093

[7]Current address: Arima Genomics, Inc., San Diego, CA, 92121, USA

[8]Current address: Monoceros Biosystems LLC, San Diego, USA

[9]Analytical Genomics and Bioinformatics, Sanford Burnham Prebys Medical Discovery Institute, Orlando, FL 32827, USA

[10]Current address: Johns Hopkins University School of Medicine, Department of Oncology & Sydney Kimmel Comprehensive Cancer Center 401N Broadway, Baltimore MD 21231

[11]Department of Cellular and Molecular Medicine, Moores Cancer Center and Institute of Genome Medicine, UCSD School of Medicine, La Jolla, CA 92093, USA

[12]These authors contributed equally

[13]Co-corresponding authors: Alessandra Dall'Agnese aled@wi.mit.edu and Pier Lorenzo Puri lpuri@sbpdiscovery.org

[14]Lead contact lpuri@sbpdiscovery.org

Summary

MYOD-directed fibroblast trans-differentiation into skeletal muscle provides a unique model to investigate how one transcription factor (TF) reconfigures the three-dimensional chromatin architecture to control gene expression, which is otherwise achieved by the combinatorial activities of multiple TFs. Integrative analysis of genome-wide high-resolution chromatin interactions, MYOD and CTCF DNA-binding profile and gene expression revealed that MYOD directs extensive re-wiring of interactions involving cis-regulatory and structural genomic elements, including promoters, enhancers and insulated neighborhoods (INs). Re-configured INs were hot-spots of differential interactions, whereby MYOD binding to highly constrained sequences of IN boundaries and/or inside INs leads to alterations of promoter-enhancer interactions to repress cell-of-origin genes and to activate muscle-specific genes. Functional evidence shows that MYOD-directed re-configuration of chromatin interactions temporally preceded the effect on gene expression and was mediated by direct MYOD-DNA binding. These data

illustrate a model whereby a single TF alters multi-loop hubs to drive somatic cell trans-differentiation.

Introduction

Growing evidence indicates the importance of the three-dimensional (3D) genome organization for the spatiotemporal regulation of gene expression (Andrey and Mundlos, 2017; Bonev and Cavalli, 2016; Bonev et al., 2017; Franke et al., 2016; Guerreiro et al., 2016; Hnisz et al., 2016a; Kragesteen et al., 2018; Lupianez et al., 2015; Noordermeer et al., 2014; Noordermeer et al., 2011; Ong and Corces, 2014; Palstra et al., 2003; Remeseiro et al., 2016; Rodriguez-Carballo et al., 2017; Schauer et al., 2017; Spielmann et al., 2018; Symmons et al., 2016). The genome is folded into a hierarchy of chromatin domains (Dekker and Mirny, 2016; Dixon et al., 2012; Dowen et al., 2014; Hnisz et al., 2016a; Nora et al., 2012; Phillips-Cremins et al., 2013; Rao et al., 2014; Sexton et al., 2012), which facilitate and constrain interactions between regulatory elements and genes. Among these chromatin domains, topologically associating domains (TADs) and insulated neighborhoods (INs) are structural units that are largely conserved across cell types (Beagan et al., 2016; Bonev et al., 2017; Chandra et al., 2015; Dixon et al., 2015; Ji et al., 2016; Jin et al., 2013; Krijger et al., 2016; Siersbaek et al., 2017). TADs consist of genomic regions that interact more frequently within the domain than with regions outside, and are separated by boundaries across which chromatin interactions are relatively scarce (Dixon et al., 2012; Nora et al., 2012). Within each TAD there may be subTADs forming a hierarchical and nested topological organization (Phillips-Cremins et al., 2013; Schmitt et al., 2016b). TADs and subTADs are generally composed of and/or contain INs (Hnisz et al., 2016a), which are regions of the DNA that contain one or more genes and whose boundaries are co-bound by CTCF and cohesin and interact with each other (Dowen et al., 2014; Flavahan et al., 2016; Hnisz et al., 2016b; Ji et al., 2016;

Narendra et al., 2015). INs contrain gene regulation within their boundaries, by harboring interactions between cis-regulatory elements, such as promoter-enhancer communication (Sun et al., 2019).

While higher genomic structures appear to be generally conserved, chromatin interactions within TADs, subTADs and INs, are instead cell-type specific and dynamic (Bonev et al., 2017; Dixon et al., 2015; Hnisz et al., 2013; Hnisz et al., 2016a; Javierre et al., 2016; Ji et al., 2016; Phanstiel et al., 2017; Remeseiro et al., 2016; Siersbaek et al., 2017) and the role of transcription factors (TFs) in regulating these interactions at the genome-wide level has not been directly addressed yet. Previous studies have been mostly based on correlative analysis of 3D chromatin reorganization and DNA sequence motifs for multiple ubiquitous TFs (Phanstiel et al., 2017) or DNA binding of neural (Bonev et al., 2017) or pluripotency TFs (Stadhouders et al., 2018), or cMYC expression (Kieffer-Kwon et al., 2017). Furthermore, the ubiquitous TF YY1 has been shown to contribute to the formation of enhancer/promoter loops (Weintraub et al., 2017). Other studies investigated promoter-related chromatin loops formed by tissue-specific TFs at individual gene level (de Wit et al., 2013; Krijger et al., 2016). However, the causative role of a single tissue-specific TF in directing genome-wide rewiring of 3D chromatin organization during lineage commitment and differentiation has yet to be determined.

Somatic cell reprogramming into another somatic cell type (trans-differentiation) or toward pluripotency (induced pluripotency) by ectopic expression of TFs provides an experimental model to address the role of master TFs (mTFs) in re-wiring chromatin interactions to regulate gene expression during establishment of cell identity, stemness, lineage commitment and terminal differentiation. However, nuclear somatic cell

reprograming almost invariably requires the combinatorial overexpression and activity of multiple mTFs (Caiazzo et al., 2011; Chronis et al., 2017; Ieda et al., 2010; Pang et al., 2011; Pfisterer et al., 2011; Qian et al., 2012; Schaub et al., 2018; Stadhouders et al., 2018; Takahashi et al., 2007; Tsunemoto et al., 2018; Vierbuchen et al., 2010; Wada et al., 2013), which complicates the interpretation of the relative contribution of each individual mTF to this process.

The analysis of the relative contribution of individual mTFs to changes in 3D chromatin architecture is also hindered by the functional versatility, context-specificity, redundancy and cooperative activity of most mTFs. A notable exception is provided by somatic cell trans-differentiation into skeletal muscle cells through the ectopic expression of one single mTF, MYOD, which is sufficient to reprogram virtually all somatic cells into skeletal muscles (Davis et al., 1987; Weintraub et al., 1989). MYOD-mediated trans-differentiation also permits the study of two separate and sequential stages of trans-differentiation: lineage commitment and terminal differentiation. Several distinctive features of MYOD, even among other myogenic bHLH factors (Conerly et al., 2016; Gerber et al., 1997), predict that MYOD possesses unique properties that enable epigenetic and transcriptional events necessary to coordinate repression of cell-of-origin gene expression and transcription of new lineage-specific genes, a complicated task that is otherwise carried out by the concerted action of multiple mTFs (Sartorelli and Puri, 2018). As such, MYOD-mediated somatic cell trans-differentiation into skeletal muscles provides a unique experimental system to investigate whether and how one single TF can re-wire 3D chromatin architecture to orchestrate activation and repression of gene expression during lineage commitment and terminal differentiation.

Genome-wide analysis of MYOD DNA binding revealed a pervasive binding through the genome; however, only a small percentage of MYOD binding sites are associated with regional gene expression (i.e. binding to proximal promoters of target genes) (Cao et al., 2010; Fong et al., 2012). Thus, the function of most MYOD binding sites remains unknown. Previous works reporting on MYOD interactions with architectural proteins, such as CTCF (Delgado-Olguin et al. 2011), and on MYOD-regulated chromatin interactions (Harada et al. 2015; Battistelli et al., 2014; Busanello et al., 2012), suggest that MYOD could regulate gene expression also by altering the 3D genome architecture.

Here we investigated the ability of MYOD to mediate topological changes in chromatin organization, establishing cell-type specific interactions to regulate gene expression programs important for myogenic commitment and differentiation.

Results

MYOD-driven myogenic conversion of primary human fibroblasts

To investigate the impact of MYOD on 3D chromatin architecture during skeletal myogenesis, we exploited the model of MYOD-directed reprogramming of fibroblasts into skeletal muscle (Davis et al., 1987; Weintraub et al., 1989). To this purpose, we introduced a tetracycline-inducible *Myod1* transgene (MYOD) or vector control (EMPTY) in human primary IMR90 fibroblasts (Figure 1A). Upon doxycycline treatment in growth media (GM) for 24hrs, ~95% of cells transfected with *Myod1* expressed *Myod1,* but not the skeletal muscle differentiation marker myosin heavy chain (MyHC) (Figure 1B and Figures S1A-C). At this stage, MYOD-expressing cells continued to proliferate (Figure S1D), indicating that *Myod1* expression levels were compatible with proliferation and

therefore the progenitor state. Following 72hrs in differentiation medium (DM), most MYOD-expressing cells (over 90%) differentiated into MyHC-expressing multinucleated myotubes (Figure 1B, Figures S1A-C). We further validated MYOD-mediated transdifferentiation of IMR90 at the trascriptome level by RNA-seq in two biological replicates. We identified 1,446 or 2,772 differentially expressed (DE) genes (see Methods) during MYOD-mediated commitment (MYOD GM vs EMPTY GM) or differentiation (MYOD DM vs MYOD GM), respectively. Gene ontology (GO) and upstream regulator prediction analysis of the DE genes showed that MYOD committed IMR90 fibroblasts toward the skeletal muscle lineage in GM by activating myogenic transcriptional networks and repressing pro-fibrotic and pro-inflammatory transcriptional networks (Figure S1E and Figure 1C) that are typically active in fibroblasts and antagonize skeletal myogenesis (Gerber et al., 1997; Liu et al., 2001; Loell and Lundberg, 2011; Puri and Sartorelli, 2000). The exposure to DM was required for the activation of the gene expression program of terminal muscle differentiation (Figure S1E). Activation or repression of these transcriptional networks was also observed when we compared RNA-seq data between EMPTY GM and primary human myotubes (hMTs, data from ENCODE), revealing that ~40% of the DE genes between MYOD DM and EMPTY GM were in common with the DE genes between hMTs and EMPTY GM (Figure S1F). Activation or repression of these transcriptional networks was validated by RT-qPCR (Figure S1G). Taken together, these results show that this system is highly suitable for investigating MYOD activity during myogenic commitment and differentiation.

The repression of the original transcriptional program is a feature shared with fibroblast reprogramming to induced-pluripotent stem cells (iPSCs). Analysis of mouse embryonic

fibroblasts (MEFs) and MEF-derived iPSCs using available gene expression data (Chronis et al., 2017) revealed that OCT4, SOX2 and NANOG inhibited the same transcriptional networks repressed upon MYOD expression in IMR90 fibroblasts (Figure S1H). These results indicate that master TFs share the ability to coordinately activate and repress specific transcriptional programs during reprogramming, as previously suggested (Ciglar et al., 2014) and that MYOD can integrate multiple activities that are otherwise accounted for by the combinatorial activity of multiple TFs.

A small fraction of MYOD binding sites are associated with local transcription regulation

To study whether MYOD regulates gene expression by direct DNA binding, we performed ChIP-seq for MYOD in IMR90-derived myoblasts (MYOD GM) and myotubes (MYOD DM). We found that MYOD pervasively bound the genome (~50,000 and ~80,000 MYOD binding sites in MYOD GM and DM, respectively), with a large preference for the prototypical E-box motif CA(G/C)GTG (Figure S1I), as previously reported (Cao et al., 2010). Two examples are reported in Figure S1J. By integrating MYOD ChIP-Seq and RNA-Seq analyses, we found that only ~5% of MYOD binding sites were located at promoters of DE genes, both during myogenic commitment (GM) and differentiation (DM) (Figure 1D), in agreement with previous studies (Cao et al., 2010; Fong et al., 2012). Since only a small fraction of MYOD binding sites are associated to local transcription regulation, we investigated MYOD DNA binding distribution to cis-regulatory elements and insulators using publicly available H3K27ac and CTCF ChIP-seq datasets in IMR90, human myoblasts (hMBs) and myotubes (hMTs) (Consortium, 2012; Jin et al., 2013; Yue

et al., 2014). We found enrichment of MYOD binding i) at promoters of DE genes (Figure 1E, see Methods), with no preference for promoters of up or down-regulated genes (Figure S1K), ii) at distal enhancers, and iii) at CTCF-binding sites in IMR90, hMBs and hMTs during both MYOD-mediated commitment (MYOD GM vs EMPTY GM) and differentiation (MYOD DM vs MYOD GM) (Figure 1E). These results suggest that MYOD might regulate transcription by binding distal regulatory and/or structural genomic elements.

MYOD DNA binding correlates with significant alterations in chromatin interactions

The enrichment of MYOD binding at *cis*-regulatory elements and at DNA elements bound by CTCF, an architectural protein implicated in chromatin looping (Hnisz et al., 2016a; Nora et al., 2017; Ong and Corces, 2014; Tsui et al., 2016), prompted us to investigate whether MYOD could regulate transcription by re-organizing interactions between functional and/or structural genomic elements.

To study MYOD-mediated changes in chromatin structure during myogenic conversion, we conducted *in situ* Hi-C (Rao et al., 2014) in two biological replicates in EMPTY GM, MYOD GM and MYOD DM. We collectively detected 2.6 billion unique pairwise contacts (each map contained on average ~470M unique pairwise contacts). Our Hi-C libraries were of high quality (Table S1, see Methods) and were highly reproducible (Figure S2A-B, see Methods).

The genome is compartmentalized into TADs (Figure 2A) that we identified using Armatus (Filippova et al., 2014), and we called TAD boundaries as in Crane et al, 2015 (Crane et al., 2015) from Hi-C matrices binned at 40kb resolution, i.e., by partitioning the

genome in 40kb bins. Although we detected MYOD DNA binding both inside TADs and at TAD boundaries, boundary location did not significantly differ during MYOD-mediated fibroblast conversion into myoblasts or myotubes (Figure 2B), since the percentage of overlap of TAD boundaries between samples (~90%) was similar to the percentage of overlap of TAD boundaries between biological replicates (~90%). This result is consistent with previous observations showing TAD conservation across cell types (Dixon et al., 2015; Schmitt et al., 2016a; Siersbaek et al., 2017). Interestingly, we observed a general pattern of co-regulation of genes within MYOD-bound TADs (Figure 2C), suggesting that MYOD could alter chromatin interactions between promoters, enhancers and insulators within TADs.

To test this hypothesis, we analyzed the *in situ* Hi-C maps at 4kb resolution and identified differential intra-chromosomal interactions between bin-pairs using diffHic (Lun and Smyth, 2015). Around 14% of the genome differentially interacted *in cis* during MYOD-mediated myogenic commitment (MYOD GM vs EMPTY GM) and/or differentiation (MYOD DM vs MYOD GM) (Figure 2D). Examples of differential chromatin interactions are shown in Figure S2C.

By integrating the differential interactome maps with MYOD ChIP-seq profile, we found that the number of bins with altered chromatin interactions and bound by MYOD was significantly higher than expected by chance (chi-squared test, $p<2.2x10^{-16}$, circular permutations, p=0, see Methods) during myogenic commitment or differentiation (Figure 2E).

To determine the extent to which the differential chromatin interactions were orchestrated by MYOD, we considered altered interactions directly bound as well as

indirect events potentially generated by MYOD DNA binding, as illustrated in Figure S3A. According to this model the initial chromatin alterations are conceivably caused by MYOD binding to the DNA (Bin2, “MYOD-dependent and MYOD-bound”). MYOD binding to the DNA could promote the interaction with another bin that may be bound by MYOD (“MYOD-dependent and MYOD-bound”) or not (Bin3, “MYOD-dependent and directly interacting”). MYOD binding to the DNA could also dis-engage previously interacting bins (i.e., Bin1 and Bin2) thereby generating free bins (i.e., Bin1) available for new interactions with other bins (Bin?), either bound by MYOD (“MYOD-dependent and MYOD-bound”) or not (Bin?, “MYOD-dependent and indirectly interacting”). Moreover, some altered chromatin interactions can form independently on the initial chain of differential interactions triggered by MYOD DNA binding (others). These differential interactions could be mediated by other TFs, whose expression might be also regulated by MYOD (Figure S3A, co-operating TF).

When we applied this analysis to our experimental system, we found that about 50% of the differentially interacting (DI) bins could be dependent on MYOD binding to the DNA (Figure 2F). We observed that ~12% and ~18% of DI bins during MYOD-mediated commitment and differentiation, respectively, were “MYOD-dependent and MYOD-bound” bins; ~14% of the DI bins during both MYOD-mediated commitment and differentiation were “MYOD-dependent and directly interacting” bins; over 22% of the DI bins during both MYOD-mediated commitment and differentiation were “MYOD-dependent and indirectly interacting” bins (Figure 2F). Interestingly, directly interacting bins were enriched in binding motifs for TFs that are typically found at MYOD-bound promoters and/or enhancers, and reported to facilitate MYOD DNA binding (i.e., Pbx)

(Berkes et al., 2004) and potentiate MYOD activation of target genes (i.e., MEF2) (Black et al., 1998; Dodou et al., 2003) (Figure S3B). Finally, ~50% of the DI bins detected in IMR90 upon MYOD expression do not appear to derive from the "domino effect" of differential interactions triggered by direct MYOD-DNA binding. However, we note that bins involved in these interactions, as well as bins involved in MYOD-dependent indirect interactions, were enriched in motifs for endogenous TFs that were up- or down-regulated as a consequence of direct MYOD binding at their promoter (Figure S3C). This is consistent with a model whereby changes in chromatin topology during IMR90 trans-differentiation derive from initial MYOD DNA binding, with MYOD-regulated expression of TFs adding an additional layer of complexity to further expand the extent of 3D chromatin re-configuration through a cooperative action of TFs.

To determine the identity of the regulatory elements involved in the differential interactions (DIs) bound by MYOD, we divided the DIs into six categories: 1) all DIs (All), 2) DIs involving promoters (Promoter-all), 3) DIs involving enhancers (Enhancer-all), 4) DIs between promoters and enhancers (Promoter-enhancer), 5) DIs bound by CTCF (CTCF), 6) DIs between bins co-bound by CTCF (CTCF-CTCF) (Figure 2G). Interestingly, the percentage of Promoter-all DIs, Enhancer-all DIs, Promoter-enhancer DIs, CTCF DIs and CTCF-CTCF DIs that was bound by MYOD was higher than the percentage of all DIs bound by MYOD (Figure 2G), suggesting that MYOD could re-wire chromatin architecture at promoter, enhancers and insulators during fibroblast trans-differentiation into skeletal muscle.

MYOD DNA binding at regions showing differential interactions at gene promoters

We observed a significant enrichment in MYOD binding at altered chromatin interactions involving promoters (chi-squared p<2.2x10-16, Figure 3A) and at promoter-enhancer pairs (chi-squared p<2.2x10-16, Figure 3B). Of note, differential chromatin interactions anchored at promoters were more frequently associated to DE genes as compared to genomic regions not bound by MYOD during MYOD-mediated commitment or differentiation (chi-squared $p<2.2x10^{-16}$, Figure 3C-D), suggesting that MYOD re-wires chromatin interactions to regulate transcription.

GO analysis on the DE genes whose promoters were involved in MYOD-mediated differential chromatin interactions revealed that MYOD-bound differential interactions involving promoters were associated with cell proliferation and muscle contractility in GM (Figure 3E), consistent with MYOD ability to stimulate proliferation of myoblasts (Latella et al., 2017), while in DM MYOD-bound differential promoter interactions were associated with cell cycle arrest and terminal muscle differentiation (Figure 3E). GO analysis on the DE genes whose promoter was involved in differential MYOD-bound interactions with enhancers revealed activation of muscle specific genes and inhibition of anti-myogenic signaling (e.g. TGF-β) (Figure 3F).

A representative example of enhancer-promoter interactions increased by MYOD is illustrated by the interaction between *TNNT2* promoter and a pre-existing putative enhancer, marked by H3K27ac, whose target gene was not previously known (Figure 3G and 3H). Upon *Myod1* expression, MYOD bound the *TNNT2* promoter (Figure 3H), and this binding correlated with a local increase of H3K27ac (Figure 3I) and with increased interaction between *TNNT2* promoter and a pre-existing enhancer, as determined by Hi-

C and validated by 3C analyses (Figure 3G and 3H bottom panel). Importantly, these events coincided with the upregulation of *TNNT2* transcription (Figure 3J).

These results establish a functional link between MYOD-directed re-wiring of chromatin interactions among *cis*-regulatory elements and dynamic regulation of gene expression that enables fibroblast conversion into skeletal muscle cells, through a stepwise model of somatic cell reprogramming.

MYOD DNA binding at re-configured insulated neighborhoods

Changes in IN strength regulate chromatin interactions and expression of genes within INs during cell differentiation (Bonev et al., 2017); however, the molecular effectors of these events remain poorly understood. Since MYOD is known to physically and functionally interact with CTCF (Battistelli et al., 2014; Delgado-Olguin et al., 2011) (see also Figure 1E and Figure 2G), we postulated that MYOD could alter the INs present in fibroblasts by targeting CTCF-organized IN boundaries.

We defined altered INs as regions of DNA that contained at least one gene and whose boundaries were i) co-bound by CTCF in IMR90 and ii) showed differential interaction strength during MYOD-mediated commitment or differentiation (Figure 4A). Each altered IN was considered as a separate entity, regardless its inclusion within larger altered INs or the presence of smaller altered INs inside it.

We found that only ~2% (1,332 or 1,595) of altered interactions accounted for changes in interaction strength between IN boundaries during MYOD-mediated commitment or differentiation, respectively (Figure 4B). Interestingly, a large proportion (greater than 40%) of the altered genome-wide interactions, e.g. enhancer-promoter intreactions,

involved DI bins within the altered INs and between bins located inside and outside altered INs (Figure 4C). The altered INs comprised a higher percentage of bins involved in differential interactions than expected by chance during both MYOD-mediated commitment and differentiation (Figure 4D). Thus, altered INs appear to be “hot-spots” of differential chromatin contacts during myogenic conversion of fibroblasts. This is consistent with the insulation effect of INs (Hnisz et al., 2016b; Lupianez et al., 2015; Nora et al., 2017; Sanborn et al., 2015; Schuijers et al., 2018; Sun et al., 2019).

By overlaying MYOD binding profile with the map of altered INs, we observed an enrichment of MYOD binding at IN boundaries whose interaction strength significantly changed during MYOD-mediated myogenic commitment or differentiation, as compared to the genome-wide binding of MYOD at CTCF-bound bins (chi-squared test $p<2.2x10^{-16}$, Figure 4E, see Methods). Consistently, we detected an overlap between MYOD ChIP-seq signal and CTCF peak summits detected at changing IN boundaries (Figure 4F). Furthermore, MYOD-binding distribution at altered INs revealed that over 90% of altered INs during myogenic commitment or differentiation were bound by MYOD at the boundaries and/or inside INs (Figure 4G). These results suggest that MYOD alters IN strength and highly re-configures the chromatin interactions landscape at those INs.

We next set to analyze genetic determinants that could discriminate between MYOD-bound IN boundaries with increased or decreased interaction strength. DNA motif analysis indicated that MYOD-bound IN boundaries in both cases were enriched in CTCF- and MYOD-binding motifs, as expected, with no significant differences in nucleotide composition (Tables S2 and S3); however, while MYOD-bound IN boundaries with increased interaction strength were strongly enriched in AP1 (Jun/Fos dimers) motifs

(Table S2), these motifs were notably absent in MYOD-bound IN boundaries with decreased interaction strength (Table S3). Conversely, MYOD-bound IN boundaries with decreased interaction strength were enriched in motifs for a collection of TFs that did not rank in the top 20 TF binding motifs found in MYOD-bound IN boundaries with increased interaction strength, with the notable exception of TCF 12 and 21, which encode potential bHLH heterodimerization partners of MYOD, and were common to both sets (Table S3). These results suggest that the presence of MYOD and other TFs at specidic loci may determine whether the interaction strength between IN boundaries is increased or decreased.

Given the transcriptional regulatory function of INs (Dowen et al., 2014; Flavahan et al., 2016; Hnisz et al., 2016b; Ji et al., 2016; Narendra et al., 2015; Sun et al., 2019) and our evidence of a significant clustering of differential interactions at altered INs (Figure 4D), we investigated whether a functional relationship exists between altered interaction strength of MYOD-bound IN boundaries and gene expression regulation within INs. Genes were considered within an IN when at least the promoter was inside the IN. We found that increased interaction strength of IN boundaries correlated with upregulation of genes within INs, especially when both IN boundaries were bound by MYOD compared to no MYOD binding or MYOD-binding at one IN boundary (Wilcoxon rank sum test p-value = 0.05 and 0.01, respectively, Figure 4H). Interestingly, these MYOD co-bound IN boundaries were also enriched in H3K27ac in the proximity of MYOD and CTCF binding sites, as compared to control IMR90 fibroblasts (Figure 4I). These results suggest that a functional relationship exists between MYOD binding, increased H3K27ac levels, transcription and increased IN boundary interactions.

GO analysis of the DE genes in INs altered by MYOD revealed inhibition of fibrosis (*TGF-β1*), activation of function of muscle and contractility of muscle (*TNNT2, ACTC1)* during MYOD-mediated commitment; activation of muscle differentiation and contractility during MYOD-mediated differentiation (Figures 4J).

An illustrative example of MYOD binding that correlated with increased IN boundary interaction and transcription upregulation is the *ITGA7-RDH5* locus (Figures 4K-L). Upon ectopic expression, MYOD bound the promoter of *ITGA7* and *RDH5* at CTCF-bound elements in IMR90 (Figure 4L) and this coincided with increased interaction between the two CTCF-bound regions, as measured by Hi-C (Figures 4K-L) and validated by 3C (Figure 4L, bottom right). Importantly, these events correlated with an increased expression of both *ITGA7* and *RDH5* (Figure L, bottom left).

The *TGF-β1* locus is an example of MYOD binding that correlated with multiple changes in chromatin interactions – e.g. decreased IN boundary interaction strength and disruption of interactions between regulatory elements – for transcriptional repression. *TGF-β1* was downregulated by MYOD (Figure S1G and S4) and is contained within an IN whose boundaries were both bound by MYOD in GM and whose interaction intensity significantly decreased during MYOD-mediated commitment (Figure S4A-B). Furthermore, MYOD binding to *TGF-β1* promoter coincided with increased interaction strength with a putative enhancer, whose H3K27ac levels were lower in hMB than in IMR90 (Figure S4B-C). Changes in interaction strength between IN boundaries as well as between the putative enhancer and *TGF-β1* promoter were first observed by Hi-C and then validated by 3C (Figure S4A-C). *TGF-β1* repression is therefore an example in which

3D chromatin reorganization occurs at multiple levels upon MYOD binding to IN boundaries as well as inside the IN that contains the *TGF-β1* locus.

MYOD-bound differentially interacting elements are highly constraint and enriched in annotated pathogenic variants

To determine the biological significance of MYOD we performed genetic constraint analysis using context-dependent tolerance score (CDTS) (di Iulio et al., 2018), which is an estimate of sequence constraint and functional importance that is calculated as the absolute difference of the observed variation from the expected variation (di Iulio et al., 2018). This analysis revealed that differentially interacting bins (DI) were more constrained than the whole-genome (WG) (Figure S5A). Sequence constraint was even higher when considering DI bins either bound by CTCF (DI_CTCF) or MYOD (DI_MYOD) or co-bound by these two TFs (DI_MYOD_CTCF) (Figure S5A). Analysis of annotated pathogenic variants revealed that CTCF and/or MYOD-bound DI bins are enriched in single nucleotide variants associated with inflammatory and muscle diseases, with a notable preference for MYOD-bound DI bins (Figure S5B). Moreover, differentially interacting IN boundaries (Bd) were also significantly more constrained than CTCF-bound 4kb bins (WG_CTCF), DI bins (DI) and all bins (WG) (Figure S5C). The high constraint was further pronounced for differentially interacting IN boundaries bound by MYOD (Figure S5C). Of note, MYOD-anchored IN boundaries were enriched in annotated pathogenic variants, including inflammatory and skeletal muscle diseases (Figure S5D). The high level of constraint and the enrichment in disease-associated pathogenic variants

observed at CTCF/MYOD-bound altered interactions indicate the biological relevance of MYOD-altered chromatin interactions.

These results also further emphasize the importance of studying the effects of mutations outside of the coding genome in altering the 3D chromatin architecture interfering with transcriptional control as reviewed in (Spielmann and Mundlos, 2016).

MYOD expression is required for sustained MYOD re-wiring of chromatin interactions.

We investigated whether MYOD expression is required for the maintenance of MYOD effects on 3D chromatin interactions, once the myogenic commitment has been established. To address this question, we turned *Myod1* expression on with doxycycline (ON) for 24hrs and then we decreased its expression by doxycycline withdrawal (OFF) for additional 48hrs, or maintained MYOD expression ON during the whole time (72hrs) (Figure 5A). Once verified the decreased expression of *Myod1* after doxycycline withdrawal (Figure 5B-C), we investigated the effect of turning OFF *Myod1* expression on MYOD-upregulated genes – *TNNT2, ITGA7* and *RDH5* – or repressed genes – *TGF-β1.*

Upon MYOD induction in fibroblasts and its binding to *TNNT2* promoter, we observed an increase in *TNNT2* promoter-enhancer interaction together with an increase in *TNNT2* expression (Figure 5D-F). Decreasing *Myod1* expression at the commitment stage (GM) led to a reduction of MYOD binding to *TNNT2* promoter, which coincided with a decrease in promoter-enhancer interaction strength to levels similar to those detected in control IMR90 fibroblasts, and a consensual reduction of TNNT2 expression (Figure 5D-F). We obtained similar results for *ITGA7-RDH5* locus. Upon MYOD expression in fibroblasts,

we observed increased *ITGA7* and *RDH5* expression, MYOD binding to CTCF-bound elements in *ITGA7* and *RDH5* promoters, increased CTCF binding at *ITGA7* promoter and increased CTCF-CTCF interaction between *ITGA7* and *RDH5* promoters (Figure 5G-I). Decreasing *Myod1* expression drastically reduced MYOD binding to CTCF-bound elements in *ITGA7* and *RDH5* promoters (Figure 5H), which was paralleled by reduction in the expression levels of *ITGA7* and *RDH5* (Figure 5G), decreased CTCF binding at *ITGA7* promoter (Figure 5H), and reduced CTCF-CTCF interaction strength (Figure 5I). Finally, decreasing MYOD restored the original expression pattern of *TGF-β1* (Figure S6A) and interactions (Figure S6B, C and E) as well as chromatin occupancy of MYOD and CTCF (Figure S6D and F) at the regulatory elements of *TGF-β1* locus.

These results suggest that steady expression of MYOD is required for the maintenance of the 3D chromatin landscape at the stage of myogenic commitment. The reversible nature of MYOD-directed chromatin interactions to coordinate repression of fibrotic genes and activation of myogenic genes observed during lineage determination in our system could be implicated in the altered differentiation and gene expression observed in satellite cells upon acute loss of MYOD *in vivo*, as reported by Yamamoto et al (Yamamoto et al., 2018).

MYOD regulates chromatin interactions in mouse myoblasts.

To further validate the role of MYOD in regulating the 3D chromatin landscape within the physiological context of skeletal myogenesis, we analyzed, as a proof of concept, the *Tnnt2* enhancer-promoter interaction in mouse C2C12 skeletal myoblasts. Upon siRNA-mediated silencing of *Myod1* (Figure S7A), we observed a significant decrease in *Myod1*

and *Tnnt2* expression (Figure S7B-C). We then used publicly available MYOD ChIP-seq dataset in C2C12 (Yue et al., 2014) and identified a MYOD peak in the murine *Tnnt2* promoter (golden eye Figure S7D) that corresponds to the peak detected in MYOD-expressing IMR90 cells (shown in Figure 3H). By inspecting the sequence conservation between the human and murine genomes, we identified in myoblasts a DNA element that is conserved with the human *TNNT2* enhancer region (shown in Figure 3H) (Figure S7D), suggesting that it could be a conserved *Tnnt2* enhancer. Interestingly, we found that in myoblasts the MYOD-bound DNA element at *Tnnt2* promoter interacted with the putative *Tnnt2* enhancer by 3C; and the interaction strength between these two genomic regions dramatically decreased upon *Myod1* silencing, (Figure S7D). These results extend to mouse skeletal muscle cells the notion that MYOD could regulate gene expression by re-organizing the 3D chromatin architecture.

MYOD rewires chromatin structure by direct DNA binding.

To investigate whether MYOD directly rewires chromatin interactions, and whether directly interplays with CTCF to alter INs, we employed a catalytically inactive Cas9 (dCas9) to block MYOD and/or CTCF binding at specific genomic loci. Briefly, we transfected IMR90 fibroblasts with dCas9-expressing vector and guide RNAs (gRNAs) that direct dCas9 to specific MYOD-bound DNA elements in the *TNNT2* locus, in MYOD and CTCF co-bound DNA elements in *ITGA7* locus or CTCF-bound elements in *RDH5* locus (Figure 6A-C). We then monitored MYOD and CTCF DNA binding, gene expression and chromatin interaction changes after dCas9 blockade of MYOD and/or CTCF DNA binding at the target sites. When we targeted dCas9 to MYOD binding site at *TNNT2*

promoter, MYOD binding was decreased at *TNNT2* promoter, but not at *ITGA7-RDH5* locus (used as negative control) (Figure 6D), demonstrating the efficacy and specificity of this approach. We found that blocking MYOD binding at the *TNNT2* promoter caused a decreased expression of *TNNT2*, while no effect on *TNNT2* expression was observed by the same dCas9 in the absence of MYOD expression (EMPTY GM) (Figure 6E). MYOD-mediated *TNNT2* promoter-enhancer interaction also decreased upon dCAS9-gRNA-mediated E-box targeting to *TNNT2* promoter (Figure 6F). These results suggest a direct role of MYOD DNA binding in promoting *TNNT2* promoter-enhancer interaction and *TNNT2* expression.

We next investigated the direct interplay between MYOD and CTCF in mediating IN boundary interaction between *ITGA7* and *RDH5* promoters (Figure 6C). Interestingly, we found that dCas9-mediated blockade of MYOD and CTCF DNA binding at *ITGA7* gene decreased MYOD and CTCF binding at both *ITGA7* and *RDH5* promoters (Figure 6D), but not at a distal gene *TNNT2*. Likewise, dCas9-mediated blockade of CTCF DNA binding at *RDH5* gene resulted in decreased binding for MYOD and CTCF at both *ITGA7* and *RDH5* promoters, but not at a distal gene *TNNT2* (Figure 6D). Blocking CTCF and MYOD binding at *ITGA7* promoter or blocking CTCF binding at *RDH5* promoter invariably decreased MYOD-mediated CTCF-CTCF interactions and decreased *ITGA7* and *RDH5* expression (Figure 6G-H). These results show that CTCF and MYOD cooperate in recruiting each other at specific DNA elements, directly altering chromatin interactions that spatially regulate tissue-specific gene expression.

The absolute requirement of MYOD-DNA binding for changes in chromatin interactions in the above loci indicates a direct role of MYOD in re-configuring 3D chromatin architecture.

Relationship between MYOD-mediated chromatin interactions and transcription

It has been previously shown that transcription can be implicated in the formation of chromatin interactions (Isoda et al., 2017). We therefore investigated the dependency of MYOD-driven chromatin interactions on transcription in our system. We first performed a time-course experiment, in which we monitored the expression of *Myod1*, *TNNT2* and *ITGA7* and MYOD-driven interactions. We found that *Myod1* expression and MYOD-mediated interactions preceded *TNNT2* and *ITGA7* upregulation (Figure 7A-D). We detected chromatin interactions already 3hrs after *Myod1* induction (Figure 7C-D), while upregulation of *TNNT2* and *ITGA7* became apparent after 12hrs (Figure 7B). These results suggest that chromatin interactions can be dissociated temporally from the transcriptional regulation of their target genes. To address whether the interactions depend on active transcription, 6 hours after inducing *Myod1* expression we inhibited transcription with the Polymerase II inhibitor Actinomycin D (ActD) for 30min. ActD treatment reduced the levels of *GAPDH* nascent RNA, but not of *GAPDH* mRNA, as compared to DMSO control, thus confirming that ActD effectively blocked transcription (Figure 7E). We then investigated the effect of ActD on MYOD-mediated chromatin interactions. Interestingly, we found that while ActD prevented MYOD-dependent enhancer-promoter interaction at *TNNT2* locus (Figure 7F), it did not affect the MYOD-promoted CTCF-CTCF chromatin interaction at *ITGA7-RDH5* locus (Figure 7G). These

results indicate that, at least in some instances, MYOD-mediated chromatin interactions occur independently on active transcription.

Discussion

Our data prompt a model for master-transcription factor (mTF)-driven re-wiring of chromatin interactions during two sequential stages of somatic cell nuclear reprogramming toward trans-differentiation: lineage commitment and terminal differentiation. Upon expression in human fibroblasts, the pervasive genome-wide binding of the mTF MYOD leads to significant alteration of chromatin interactions to direct the erasure of the cell-of-origin identity and the commitment toward the skeletal muscle lineage. At this stage, MYOD-driven chromatin interaction changes were reversed after *Myod1* downregulation. Upon differentiation stimuli, which promote MYOD hetero-dimerization with E-proteins and activation of terminal differentiation, we observe further re-configuration of the chromatin 3D interactions, indicating a stepwise and signal-regulated, MYOD-directed alteration of the genome architecture.

MYOD-mediated somatic cell trans-differentiation provides a unique model to investigate whether and how one single TF re-wires the 3D genome architecture in order to disengage chromatin interactions responsible for the gene expression pattern of the cell of origin (i.e., fibroblasts) and at the same time establish new genomic contacts that promote the expression of new lineage-specific genes (i.e., skeletal muscle cells). Somatic cell nuclear reprogramming toward either trans-differentiation or pluripotency is a multi-step task that is typically achieved by the combinatorial activities of multiple TFs (Caiazzo et al., 2011; Chronis et al., 2017; Ieda et al., 2010; Pang et al., 2011; Pfisterer

et al., 2011; Qian et al., 2012; Schaub et al., 2018; Stadhouders et al., 2018; Takahashi et al., 2007; Tsunemoto et al., 2018; Vierbuchen et al., 2010; Wada et al., 2013). This is consistent with a model whereby defined TFs complement each other activity, which is otherwise not sufficient to drive the entire program. Our data suggest that the unique property of MYOD to initiate a successful program of somatic cell trans-differentiation, upon ectopic expression, might rely on its ability to re-configure 3D chromatin architecture, via binding to its consensus DNA motifs – the myogenic E-boxes – at structural and *cis*-regulatory elements. Importantly, our data show that MYOD-mediated changes in nuclear architecture temporally precede the changes in the expression of target genes, as also predicted bioinformatically (Liu et al., 2018) and shown previously for some genes (Ghavi-Helm et al., 2014). Moreover, while MYOD-mediated alteration of IN boundary interaction is independent on active transcription, at least some MYOD-mediated alteration of enhancer-promoter interactions appear to depend on active transcription, as it is sensitive to the PolII inhibitor ActD. It is possible that short-lived RNAs may cooperate with MYOD in looping *cis*-regulatory elements, as recently proposed by Sartorelli and colleagues (Mousavi et al., 2013; Tsai et al., 2018).

DNA binding of pluripotency factors has also been shown to correlate with chromatin topological changes that occurred prior to transcriptional regulation (Stadhouders et al., 2018). Indeed, many analogies could be found between somatic cell nuclear reprogramming by MYOD alone vs multiple “defined factors”, including the genome reorganization at multiple architectural levels, in order to coordinate the erasure of the pre-existing transcriptional program and the selective activation of the gene expression pattern that defines the new cell identity. In this regard, we argue that our data provide an

initial model for TF-driven re-configuration of 3D chromatin architecture for somatic cell nuclear reprogramming. In support to this model, a recent work that exploited the ectopic expression of just one of the pluripotency factors, KLF4, in mouse embryonic fibroblasts (MEFs), has shown striking analogies with MYOD, including the ability to promote enhancer-promoter interactions enriched with H3K27ac that activate the expression of pluripotency genes (Di Giammartino D., 2018). However, KLF4 is not able to drive the entire somatic cell nuclear reprogramming toward pluripotency without the co-expression of other defined factors (i.e. NANOG, OCT4, MYC). We speculate that MYOD integrates multiple architectural and transcriptional properties into one TF, thereby providing a general paradigm for TF-directed re-wiring of chromatin interactions to instruct somatic cells toward a specific lineage.

Figure 2 shows that the large majority of changes in chromatin interactions identified in fibroblasts upon the ectopic expression of MYOD are actually orchestrated by a single TF (i.e. MYOD), either directly or indirectly. TF-directed processive re-configuration of the 3D chromatin interactions illustrates how the effects of MYOD on 3D chromatin interactions is amplified by events secondary to the direct formation of initial interactions. It shows that MYOD-bound DIs regulate the expression of downstream TFs that can mediate new chromatin interactions. This model is consistent with the current view of cooperative activity of TFs in nuclear re-programming and accounts for the expansion of the architectural repertoire of master TFs, such as MYOD.

Our study shows that MYOD promotes or impairs interactions between *cis* regulatory elements, i.e., enhancers and promoters. Moreover, we discovered that MYOD can affect the topological organization of the genome at the subTAD level, by altering INs structure,

via binding at CTCF-anchored boundaries as well as by targeting interactions inside INs. This finding further extends recent observations reported in mESCs (Sun et al., 2019).

The reported MYOD-dependent changes in chromatin contacts suggest that multiple levels of 3D chromatin organization imposed by mTFs might provide robustness (and possibly redundancy) to newly erased or acquired interactions, which might account for tolerance of gene expression to genetic mutations. This finding will inspire future studies to determine whether there is a functional hierarchy or reciprocal connectivity between MYOD-directed chromatin interactions at IN boundaries and within INs. At the same time, the high constraint of sequences implicated in MYOD-directed genomic interactions in the human population and their enrichment in disease-associated single nucleotide variants indicate that TF-altered INs could be “hotspots” for re-configuration of nuclear architecture for lineage determination during developmental and post-natal skeletal myogenesis.

Of note, we found a strong association between MYOD-mediated increased strength of IN boundaries, enrichment in H3K27ac and activation of genes within the INs. Co-bound of MYOD to IN boundaries and presence of AP1 binding sites appear as two major determinants of transcriptional activation of genes within INs. Enrichment of AP1 binding sites flanking MYOD peaks has been consistently observed in ChIP-seq studies (Cao et al. 2010) and was anticipated by earlier studies reporting on physical interactions between MYOD and cJUN (Bengal et al., 1992). However, previous reports, based on overexpression experiments, have suggested a functional antagonism between MYOD and AP1 (Li et al., 1992). Our finding that AP1 binding motifs selectively associate with MYOD-mediated alteration of INs that activate gene expression suggests that AP1 could

represent a major genetic determinant of MYOD-directed control of local gene expression. This is consistent with the role currently assigned to AP1 as frequent component of super-enhancers (Phanstiel et al., 2017).

Another interesting aspect of MYOD-mediated somatic nuclear cell reprogramming concerns its ability to repress the expression of cell-of-origin genes by altering pre-existing chromatin interactions through binding to E-box sequences. This is well illustrated by the MYOD-mediated alterations of promoter-enhancer interactions at the insulated neighborhood that harbors the *TGFβ* locus. MYOD has been known since its discovery as a sequence-specific transcriptional activator (Weintraub et al., 1991), with no structural and functional features that can account for its ability to repress gene expression (Puri and Sartorelli 2000). Although transient interactions with co-repressors has been reported (Puri et al., 2001; Singh et al., 2015), this mechanism has been implicated in the temporal regulation of target gene activation, rather than the stable repression of other lineage genes. Thus, our data suggest that MYOD-mediated re-wiring of chromatin interactions can account for its ability to stably repress gene expression, via direct DNA binding to E-box motifs.

Overall, our work revealed previously unappreciated features and mechanistic insights on alterations in 3D genome architecture by a single TF that allow significant changes in gene expression, leading to coordinated repression of cell of origin gene networks and activation of tissue-specific genes during somatic cell reprogramming, therefore significantly extending our knowledge on TF-mediated lineage activation and terminal differentiation (Heinz et al., 2010; Natoli, 2010; Spitz and Furlong, 2012).

STAR METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal MYOD	Santa Cruz	M-318
Rabbit polyclonal H3K27Ac	Active Motif	
Mouse monoclonal MYOD	BD Bioscience	554130
Mouse monoclonal MyHC	DSHB	MF20
Goat anti-mouse IgG, Fc subclass 1 specific	Jackson Immuno	NC0469362
Goat anti-mouse IgG, Fc subclass 2b specific	Jackson Immuno	NC0266980
Mouse monoclonal GAPDH	Abcam	ab9485
Anti-bActin-HRP	Abcam	ab20272
Anti-human TNNT2	Abcam	ab10214
Bacterial and Virus Strains		
Biological Samples		
Chemicals, Peptides, and Recombinant Proteins		
Doxycycline	Sigma-Aldrich	D3072-1ML
Puromycin dihydrochloride	MP Biomedicals	ICN10055210
ITS	Sigma-Aldrich	I2146
Hoechst 33258	Life Technologies	H3569
PMSF	Sigma-Aldrich	93482
Protease inhibitors	Roche	11697498001
DpnII	NEB	R0543
T4 DNA Ligase	NEB	M0202L
Proteinase K	NEB	P8107
BSA	NEB	B9000
biotin-14-dATP	Life Technology	19524-016
Critical Commercial Assays		
Neon Transfection System	Invitrogen	MPK5000, MPK10025
BCA Protein Assay Kit	Invitrogen	23235
TruSeq Stranded mRNA Library Prep Kit set A	Illumina	RS-122-2101
Deposited Data		
RNA-seq, ChIP-seq, Hi-C	This paper	GEO: TBD

Other ChIP-seq data	(Consortium, 2012; Jin et al., 2013; Yue et al., 2014)	GEO: GSM935404, GSM733762, GSM733783, GSM1055818, GSM915188, GSM733666, GSM733755, GSM915165
Other RNA-seq data	(Chronis et al., 2017)	GEO: GSM2417196, GSM2417197, GSM2417198, GSM2417204, GSM2417203
Experimental Models: Cell Lines		
IMR90	Coriell	I90-83
C2C12	ATCC	CRL-1772
Experimental Models: Organisms/Strains		
Oligonucleotides		
PCR primers – see list at the end of STAR Methods	This paper	N/A
siScr	Dharmacon	D-001210-01-05
siMyod	Ambion	s232596
Recombinant DNA		
Helper plasmid	Provided by Dr. Alessandro Rosa	N/A
epB-Puro-TT plasmid EMPTY	Provided by Dr. Alessandro Rosa	N/A
epB-Puro-TT plasmid EMPTY mouse Myod1 cDNA	Provided by Dr. Alessandro Rosa	N/A
Software and Algorithms		
TopHat2.1.1	(Kim et al., 2013)	https://ccb.jhu.edu/software/tophat/index.shtml
HTSeq-0.6.1	(Anders et al., 2015)	http://www-huber.embl.de/HTSeq/doc/overview.html
DESeq2	(Love et al., 2014)	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Ingenuity Pathway Analysis	Qiagen	https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/

Bowtie2-2.0.5/bowtie2	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Samtools1.3	(Li et al., 2009)	http://samtools.sourceforge.net/
Macs2v2.1.1	(Zhang et al., 2008)	http://liulab.dfci.harvard.edu/MACS/
Bedtoolsv2.26.0	(Quinlan and Hall, 2010)	http://bedtools.readthedocs.io/en/latest/
HiCPro-v2.7.7	(Servant et al., 2015)	https://github.com/nservant/HiC-Pro
HiTC	(Servant et al., 2012)	https://bioconductor.org/packages/release/bioc/html/HiTC.html
HiCPlotter	(Akdemir and Chin, 2015)	https://github.com/kcakdemir/HiCPlotter
Armatus	(Filippova et al., 2014)	https://www.cs.cmu.edu/~ckingsf/software/armatus/
DiffHic	(Lun and Smyth, 2015)	https://bioconductor.org/packages/release/bioc/html/diffHic.html
Jaspar	(Mathelier et al., 2014)	http://jaspar.genereg.net/
MEME	(Bailey et al., 2009)	http://meme-suite.org/tools/meme-chip
HOMER	(Heinz et al., 2010)	http://homer.ucsd.edu/homer/
Other		
Olympus IX71 microscope		

Sequences. Primers sequences for expression analysis, ChIP-qPCR and 3C experiments are provided in Supplementary Table 4,5 and 6.

Antibodies and recombinant proteins. The following primary antibodies were used in this study: rabbit polyclonal anti-MYOD (Santa Cruz, sc-760), mouse monoclonal anti-MYOD (Santa Cruz, sc-377460 and BD Bioscience, 554130), rabbit polyclonal anti-

H3K27ac (Active Motif, 39135), mouse monoclonal anti-MyHC (DSHB, MF-20), mouse monoclonal anti-GAPDH (Abcam, ab9485), mouse monoclonal anti-beta Actin (Abcam, ab20272) and mouse monoclonal anti-TNNT2 (Abcam, ab10214). The secondary antibodies were anti-mouse IgG HRP conjugated (Thermo Fisher Scientific), goat anti-mouse IgG, Fc subclass 1 specific Cy3-conjugated (Jackson ImmunoResearch, 115-545-207) and goat anti-mouse IgG, Fc subclass 2b specific 488-conjugated (Jackson ImmunoResearch, 115-165-205). DpnII (R0543), T4 DNA Ligase (M0202L), Proteinase K (P8107) and BSA (B9000) were from NEB. Biotin-14-dATP from Life Technology (19524-016).

Cell Culture Experiments. IMR90 cells (Coriell) were grown in growth media (GM) consisting of EMEM (ATCC) supplemented with 10% FBS (Omega Scientific). Electroporation was performed in proliferating cells at passage 11-15. All other experiments were performed in proliferating cells at passage 23-28. Doubling passage is crucial for success of myogenic conversion. C2C12 cells (ATCC) were grown in growth media (GM) consisting of DMEM/High Glucose (HyClone) supplemented with 10% FBS.

Myogenic conversion. IMR90 cells were electroporated using the Neon Transfection System (Invitrogen, MPK5000, MPK10025) with helper plasmid and epB-Puro-TT containing or not murine MYOD cDNA. Cells were then selected with 2 ug/ml of puromycin dihydrochloride (MP Bio). When cells were 60% confluent, *Myod1* was induced with 200 ng/ml doxycycline (Sigma) in GM for 24 hr and cells were collected for the GM point. When cells were 95-100% confluent, MYOD was induced with 200 ng/ml doxycycline

(Sigma) in GM for 24 hr and then cells were differentiated in EMEM supplemented with 2% horse serum (Gibco), 1% ITS (Sigma), 200 ng/ml doxycycline for three days for the DM time point. Media with doxycycline was refreshed every 2 days.

MYOD Time-course. When cells were 60% confluent, *Myod1* was induced with 200 ng/ml doxycycline (Sigma) in GM and cells were collect for IF, RNA or 3C after 3, 6, 12 and 24 hrs.

Transcription inhibition. When cells were 60% confluent, *Myod1* was induced with 200 ng/ml doxycycline (Sigma) in GM for 6 hours and treated with 1µg/ml of Actinomycin D (Sigma) for 30 minutes at 37C. DMSO was used as vehicle control. Following the treatment cells were collected for gene expression and 3C analyses.

siRNA transfection. C2C12 cells were transfected with 12.5 pmol of siScr (Dharmacon) or siMyod (Ambion) using Lipofectamine RNAiMAX (Life Technologies) according the manufacturer's instructions. 48hrs post transfection media containing transfection mix has been replaced with 2 ml fresh GM media. Cells have been collected after additional 24hrs in culture.

Generation of gRNAs expressing plasmid. gRNA plasmids have been generated according to Kabadi et al (2014)(Kabadi et al., 2014). Briefly oligos DNA, with the appropriate overhangs have been annealed and cloned into the appropriate donor plasmid and subsequently cloned into pLV hUbC-dCas9-T2A-GFP. phU6-gRNA; pmU6-

gRNA; phH1-gRNA; p7SK-gRNA and pLV hUbC-dCas9-T2A-GFP are gift from Charles Gersbach (Addgene # 53187, 53187, 53186, 53189, 53191 respectively).

Plasmid Transfection. IMR90 were grown in GM media until approximately 60-70% confluency and transfected with gRNA expressing vectors using Lipofectamine 2000 (Life Technologies) according to the manufacturer's instructions. 36 hrs after transfections media was changed with GM media containing 200 ng/ml doxycycline to induce MYOD expression and cells were grown for additional 24hrs before being collected for *in-situ* 3C and RNA expression analysis.

Immunofluorescence. Cells were fixed with 4% PFA in PBS, permeabilized with 0.5% TX100 and blocked with 5% BSA in PBS. Cells were stained with anti-MYOD (BD Bioscience, 554130) and anti-myosin heavy chain (DSHB, MF20) for 3 hrs or O/N at RT followed by anti-mouse IgG, Fc-subclass 2b 488 conjugate (Jackson ImmunoResearch) and anti-mouse IgG, Fc-subclass 1 Cy3 conjugated (Jackson ImmunoResearch) for 1hr at RT in the dark. Nuclei were then counterstained with 2 ug/ml Hoechst 33258 pentahydrate (bis-benzimide) (Life Technologies). Images were acquired with fluorescence microscope. Fields reported in figures are representative of all examined fields.

Western Blot. Cells were lysed in RIPA Buffer (50 mM Tris-HCl, 0.1M NaCl, 0.5% sodium deoxycholate, 1% IGEPAL CA630, 0.5% SDS, 1mM EDTA) supplemented with 1mM PMSF (Sigma) and protease inhibitors cocktail (Roche). Protein concentration was

measured by BCA Protein Assay Kit (Invitrogen). 5-20ug of proteins were run on a 4%-12% or 10% tris-glycine gel (Novex) and transferred to a 0.45 µm nitrocellulose membrane. Membrane was blocked with 2.5% skim milk (BD) in PBS-Tween (PBS with 0.1% Tween 20) for 1 hr at RT. Membrane was incubated with primary antibodies anti-MYOD (1:1000 BD Bioscience, 554130) and anti-myosin heavy chain (DSHB, MF20), anti-TNNT2 (1:1000 Abcam, ab10214) O/N at 4C or with anti-GAPDH (1:1000 Abcam, ab9485) anti-bACTIN (1:1000 Abcam ab20272) for 1hr at RT. After three washes in PBS-Tween, membrane was incubated O/N with anti-mouse IgG HRP (Thermo Fisher Scientific). For detection, ECL (Thermo Scientific, 32106) was used.

mRNA expression analysis. Cells were lysed in Trizol (Ambion) and RNA was extracted following manufacture's recommendation. RNA concentration was measured on Qubit (Invitrogen). 100-500 ng of RNA was reverse transcribed using QuantiTek Reverse Transcrition Kit (Qiagen). Real-time quantitative PCR (qPCR) was performed using Power SYBR Green Master Mix (Life Technologies) following manufacture's indications. Expression was normalized to *Gadph* for IMR90 cells or *b-actin* for C2C12 using $2^{-\Delta\Delta Ct}$ method.

RNA Sequencing and data analysis. Cells were collected from the plate using trypsin, that was then inhibited by adding the media cells were in before trypsinization. Spike-in were added based on number of nuclei, but not used for the analysis. PolyA RNA-seq Libraries were prepared using TruSeq RNA Library Prep Kit (Illumina) and deep sequenced on the HiSeq2500 ~50 million reads per conditions. Read quality was

determined using FASTQC. Reads were mapped to the female Homo sapiens hg19 genome using TopHat2.1.1{Kim:2013eo} using the following options: : -p 8 –g 1 –segment-length 17 –library-type fr-firststrand. Over 84% of the reads successfully mapped. HTSeq-0.6.1p1173 with –stranded=reverse option was used to assign mapped reads to Homo Sapiens GRCh37.75 genes. Differential expression analysis was performed using DESeq2 (Love et al., 2014). Genes were considered differentially expressed if $p<0.05$ and fold change was lower than 0.5 or higher than 2. For integrated analysis with ChIP-seq and Hi-C, we considered differentially expressed if $p<0.05$ and fold change was lower than 0.5 or higher than 2 and gene transcript per million was higher or equal to 1 in at least one of the conditions compared. Ingenuity pathway analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity) was used for gene ontology. Human skeletal myotubes RNASeq was taken from ENCODE database (SRR307932.sra and SRR307933.sra). Reads were trimmed to 50 bases using fastx_trimmer. Reads were mapped to the male Homo sapiens hg19 genome using TopHat2.1.1{Kim:2013eo} using the following options: -p 8 -g 1 --segment-length 17. HTSeq-0.6.1p1173 with –stranded=no option was used to assign mapped reads to Homo Sapiens GRCh37.75 genes. Differential analysis was performed as described above.

ChIP and ChIP-seq. Cells were fixed in 1% formaldehyde (Sigma, F8775) in PBS for 15 min at RT. Formaldehyde was then quenched with 125mM Glycine for 5 min at RT. Cells were washed in PBS and harvested in PBS supplemented with 1mM PMSF and protease inhibitors. Dry cell pellet was stored at -80C. Nuclei were then extracted and then lysed in lysis buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM EDTA,

pH 8.0, 0.5% SDS, 0.5% NP-40, 1 mM PMSF and a protease inhibitor. Chromatin was sheared with sonicator (ColeParmer, Misonix 3000) to an average DNA fragment length of 200-500bp. Chromatin was then diluted 5 times in lysis buffer without SDS. DNA amount was measured with the Qubit (Invitrogen Q32854). DNA was immunoprecipated with either rabbit anti-MYOD (Santa Cruz), or rabbit anti-H3K27ac (Active Motif) O/N at 4C. The immunocomplexes were captured with protein A magnetic beads (Life Technologies) for 3-4 hrs at 4C. After four washes with buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM EDTA, pH 8.0, 0.1% SDS, 1% NP-40, 0.5% sodium deoxycholate, one wash with a buffer containing 250 mM LiCl, 100 mM NaCl, 5 mM EDTA, pH 8.0, 1% NP-40, 1% sodium deoxycholate and two washes with TE buffer (10mM Tris-HCl pH=8, 1mM EDTA) chromatin was then eluted and decrossliked with 1% SDS in TE O/N at 65C 600RPM rotation. Also, the input is decrosslinked with 1% SDS in TE O/N at 65C 600RPM rotation. After 2 hrs digestion at 37C with 0.2 mg/ml proteinase K, DNA was extracted with phenol/chloroform and ethanol precipitated O/N at -20C. Prior to sequencing, DNA was then suspended in mQ water. The DNA was then analysed by qPCR calculating the amount of immunoprecipitated DNA relative to the input DNA (percentage of input). Library preparation and sequencing of immunoprecipitated and input DNA were performed as described http://bioinformatics-renlab.ucsd.edu/RenLabLibraryProtocolV1.pdf.

ChIP-seq analysis. Read quality was determined using FASTQC. Reads were mapped using bowtie2-2.0.5/bowtie2 to the female *Homo sapiens* hg19 genomes with options: --very-sensitive-local. Over 85% of the reads successfully mapped. Duplicate reads were

removed using samtools1.3. Peaks were called using macs2 2.1.1.20160309 with qvalue=0.01, macs2 2.1.1.20160309 was also used for differential peak calling among samples. Reads were extended based on the fragment size predicted with macs2. Heatmap of ChIP-seq signal was generated using Seqminer. We also analyzed previously published ChIP-seq data. To compare H3K27ac levels between IMR90, hMB and hMT we started from the same number of reads. These data were analyzed following the same workflow as our MYOD ChIP-seq data. Motif analysis was performed using MEME Suite (Bailey et al., 2009), Jaspar (Mathelier et al., 2014) or HOMER (Heinz et al., 2010).

***In situ* Hi-C.** Hi-C was performed as previously described (Rao et al., 2014) with the following modifications. Cells were cross-linked with 2% formaldehyde in media. Formaldehyde was then quenched with 200mM of glycine for 5 min at RT. Cells were then washed in PBS and pelleted. Cell pellet was then saved at -80C. $2x10^6$ cells were then lysed with lysis buffer (10mM Tris-HCl pH 8.0, 10mM NaCl, 0.2% Igepal CA630). Incubation of cells in 0.5% SDS in mQ at 62C for 10 min is followed by SDS quenching with TritonX-100 for 15 min at 37C. NEB3 buffer was added to reach 1X final concentration. DNA was then digested with DPNII O/N at 37C at 900RPM. Inactivation of DPNII was performed by incubating the samples at 62C for 20 min. Fill in of the digested end was performed by adding biotin-14-dATP (Life Technology, 19524-016), dCTP, dGTP, dTTP (Invitrogen) and Klenow (NEB, M0210). Mixture was incubated at 37C for 90 min 500RPM. Ligation was performed in 1.2ml by adding mQ water, T4 DNA ligase buffer to concentration 1X (NEB, B0202), 0.083% TritionX-100, 0.01mg/mL BSA,

2000U/uL T4 DNA Ligase (NEB, M0202) for 4 hr at RT with slow rotation. DNA is then ethanol precipitated, resuspended in 10mM Tris-HCl pH=8 and sheared using Covaris sonicator. Size selection of DNA (200-600bp) was performed using AmpureXP beads. DNA ends were then repaired and biotin removed from un-ligated samples by incubating the DNA at 37C for 30 min in 1X T4 DNA ligase buffer with 0.5mM dNTPs (Invitrogen), 50U T4 PNK (NEB, M0201), 12U T4 DNA Polymerase (NEB, M0203) and 5U Klenow. Biotin-labelled DNA was pulled down using Dynabeads My One T1 Streptavidin beads (Life Tech). Illumina Indexed adapter are then ligated with NEB DNA Quick Ligase (NEB, M2200). Beads were then washed and dissolved in 10mM Tris-HCl pH=8. KAPA qPCR assay was then performed to estimate concentration and cycle number needed for final PCR.

Hi-C analyses. Read quality was determined using FASTQC. **For interaction matrix**, HiCPro-v2.7.7 was used for read mapping, detection of valid ligation products, quality control, and sparse chromosomal interaction maps(Servant et al., 2015) using the following settings: BOWTIE2_GLOBAL_OPTIONS = --very-sensitive -L 30 --score-min L,-0.6,-0.2 --end-to-end –reorder, BOWTIE2_LOCAL_OPTIONS = --very-sensitive -L 20 --score-min L,-0.6,-0.2 --end-to-end –reorder, REFERENCE_GENOME = hg19_XX, LIGATION_SITE = GATCGATC, MIN_FRAG_SIZE = 100, MAX_FRAG_SIZE = 100000, MIN_INSERT_SIZE = 100, MAX_INSERT_SIZE = 600, MAX_ITER = 100, FILTER_LOW_COUNT_PERC = 0.02, FILTER_HIGH_COUNT_PERC = 0, EPS = 0.1. Quality of the libraries based on percentage of mapped reads (for both ends), percentage of reported pairs (removal of unmapped pairs, multiple pairs alignments, low quality pairs,

not reported pairs and pairs with singleton - % considering the total number of reads), percentage of valid putative interaction pairs (removal of dangling ends, fragments with no restriction site, self-circles etc - % considering the number of reported pairs), percentage of unique read pairs (removal of duplicates - % considering the valid putative interaction pairs), number of unique read pairs, percentage of cis-read pairs, number of long-range cis-read pairs, percentage of trans read pairs was determined using HiCPro-v2.7.7 (Servant et al., 2015). HiTC was used to transform sparse matrices to NxN matrices(Servant et al., 2012). For Hi-C library quality analysis presented in Table S1 please refer to (Servant et al., 2015). **Hi-C data reproducibility** between replicates per chromosome was calculated in two ways a) as previously described by Dixon et al, 2012 (Dixon et al., 2012) and b) using HiC-spector (Yan et al., 2017). For the first method, the set of all possible intra-chromosomal interactions for two replicates were correlated by comparing each point in interaction matrix at 4kb resolution from one replicate with the same point from the second replicate. We restricted the correlation to a maximum distance between points of 2Mb (500 bins), since Hi-C data is skewed toward proximal interactions (Dixon et al., 2012). We used the *cor* function of R (version 3.2.3) to calculate the Pearson correlation between the two vectors. For the second method, we used the python script HiC-spector (Yan et al., 2017) on 4kb raw intra-chromosomal triple sparse format matrices. **Hi-C heatmap** in Figure 2B and Figure S6A were generated using HiCPlotter (Akdemir and Chin, 2015). **TAD calling:** TADs were called on *NxN* ICE-normalized matrices using Armatus (Filippova et al., 2014) (v2.1), with gamma-max (-g) set to 0.3, resolution (-r) set to 40kb and the remaining parameters left as default. We called TADs at various resolutions, but we report data on TADs called at 40kb resolution,

because 40kb was the highest resolution that gave us reproducible TAD calls between biological replicates. **Boundaries calling:** TAD boundaries were called following a previously described method (Crane et al., 2015) using a window size of 400kb, which was selected based on the reproducibility across biological replicates. We chose the smallest window size that would give high (>=90%) reproducibility. **Differential interaction calling:** differential interactions were called with diffHic (Lun and Smyth, 2015) v1.4.2, in R (v3.3.1), on raw matrices at 4kb resolution, with two biological replicates for each condition. Triplet sparse format matrices generated by HiCPro were converted to InteractionSet objects with the GInteractions function from the InteractionSet package (v1.0.4) and then organized in a counts matrix with the InteractionSet function. Differential analysis was performed for each chromosome separately, considering only intra-chromosomal bin pairs where the average logCPM > 1 across samples. Data were normalized using Loess normalization, applied separately to regions near (<=2L) and far from the Hi-C matrix diagonal. InteractionSet was converted to a DGEList object to be used as input for EdgeR (Robinson et al., 2010) v3.14.0. Once calculated data dispersion with the estimateDisp and glmQLFit functions, differential analysis is performed using a Quasi-Likelihood F-Test and Benjamini-Hochberg FDR correction. Requirement for significant differential interaction: fold change lower than 0.5 or higher than 2 and p-value lower than 0.05. All differential interactions falling on the diagonal were excluded. For fold change distribution analysis, we converted all fold changes between 0 and 1 as -1/(Fold change).

Data integration analysis. To integrate data from RNA-seq, ChIP-seq and Hi-C experiments, we used bedtools 2.26.0(Quinlan and Hall, 2010) and R packages. Promoters were defined based on the location of the TSS of Homo sapiens GRCh37.75 genes (+/- 4kb from annotated TSS). Enhancers were defined as DNA regions containing H3K27ac that did not overlap with promoters.

Circular permutations. Chromosome-bound circular permutations test described in Cabrera et al., 2012(Cabrera et al., 2012) was applied to evaluate if the observed overlap between Hi-C differentially interacting bins found in the different comparisons (i.e., EMPTY GM vs. MYOD GM and MYOD GM vs. MYOD DM) and MYOD ChIP-seq peaks was significantly enriched compared to the expected overlap. For each permutation, differentially interacting bins obtained from the comparison in analysis were shifted of a randomly generated number of bins (comprised between 1 and the maximum number of bins of the smallest chromosome in analysis, chr21), for a total of 10,000 permutations. When the shift exceeded the end of the chromosome, the permutation continued from the chromosome start, thus regarding chromosomes as circularized. Randomly generated genomic intervals were then overlapped with MYOD ChIP-seq peaks found in MYOD GM and MYOD DM, respectively. The number of permuted datasets, n, having a number of bins overlapping MYOD peaks greater than or equal to the observed number were noted and used to estimate approximate p-values (n/10,000) for enrichment.

Gene expression analysis inside differential INs. For Figure S5A, the list of INs whose boundaries show a higher interaction strength between boundaries in MYOD GM

as compared to EMPTY GM, and the list of MYOD ChIP-seq peaks in MYOD GM, were overlapped with bedtools (intersectbed, with parameters –wa and –u) in order to partition the differential INs in three categories: i) INs whose boundaries are not bound by MYOD, ii) INs where only one boundary is bound by MYOD and iii) INs where MYOD binds both boundaries. These sub-lists, together with the initial list of all INs whose boundaries interaction is strengthen upon MYOD expression, were overlapped with the list of human promoters (NB. genes were considered within an IN when at least the promoter overlapped the IN). Promoters were defined based on the location of the TSS of Homo sapiens GRCh37.75 genes (+/- 4kb from annotated TSS). Genes were subsequently filtered in order to retain only those with p-val<0.01 from the RNA-seq differential analysis (EMPTY GM *vs* MYOD GM comparison) and mapping exclusively to only one of the subgroups, in order to exclude the potential confounding effects due to the presence of nested INs. The bed intervals of the INs whose boundary interaction strength increases upon MYOD expression and that are bound by MYOD at both boundaries were used to plot the ChIP-seq signal of CTCF, MYOD and H3K27 acetylation (the latter both in EMPTY GM and in human myoblasts) with the NGSplot R s, with parameters -G hg19 -R bed -L 2500 -MW 7 -YAS 0,0.7 (Shen et al., 2014).

Calculations of "expected number" of events for each figure panel. For Figure 1F, observed/expected ratios were calculated as in Chronis et al., 2017 (Chronis et al., 2017), as the ratio between the observed and the expected overlap for each feature based on their sizes and the size of the hg19 human genome:

$$\frac{F \cap S}{F * S/G}$$

where F is the number of base pairs annotated for the feature F (e.g. MYOD peaks), S is the size of the feature S (e.g. promoters of differentially expressed genes) and G is the length of the human genome.

For Figure 2D, expected frequency of MYOD binding at differentially interacting bins (f_{ex}) =

= (total number of 4kb bins bound by MYOD)/(total number of 4kb bins genome-wide)

Expected number of bins that differentially interacted that are bound by MYOD =

= f_{ex}* (number of 4kb bins that differentially interacted).

For Figure 3A, expected frequency of MYOD binding at differentially interacting bins (f_{ex}) =

= (total number of 4kb bins bound by MYOD)/(total number of 4kb bins genome-wide)

Expected number of bins that differentially interacted that are bound by MYOD =

= f_{ex}* (number of 4kb bins that differentially interacted, in this case one of the differentially interacting partner bin has to be in a promoter).

For Figure 3B, expected frequency of MYOD binding at differentially interacting bins (f_{ex}) =

= (number of 4kb bins that coincide with promoters or enhancers that are bound by MYOD)/(number of 4kb bins genome-wide that coincide with promoters or enhancers)

Expected number of bins that differentially interacted that are bound by MYOD =

= f_{ex}* (number of differentially interacting 4kb bins involved in differential enhancer-promoter interaction).

For Figure 4E, expected frequency of MYOD binding at differentially interacting bins co-bound by CTCF (f_{ex}) =

= (number of 4kb bins that are co-bound by MYOD and CTCF)/(number of 4kb bins that are bound by CTCF)

Expected number of bins that differentially interacted that are bound by MYOD =

= f_{ex}* (number of 4kb bins that differentially interacted with each other and that are both bound by CTCF)

To determine statistical significance between observed and expected, chi-squared test was performed using R version 3.2.3, function chisq.test().

Genetic constraint analysis. The hg19 human reference genome was downloaded from ftp://hgdownload.cse.ucsc.edu/goldenPath/hg19/chromosomes/ and split into 4kb bins, which corresponds to the experimental resolution. The bins were further classified into separate categories, using bedops (v2.4.30)(Neph et al., 2012), depending on whether they contained MYOD and/or CTCF peaks, were present in differentially interacting regions or boundaries. Of note, the categories are not mutually exclusive. The names and number of bins per categories are the following: DI (N=71,501 in EMPTY_GMvsMYOD_GM and N=74,327 in MYOD_GMvsMYOD_DM); DI_CTCF (N=7,414 in EMPTY_GMvsMYOD_GM and N=7,327 in MYOD_GMvsMYOD_DM); DI_MYOD (N=9,096 in EMPTY_GMvsMYOD_GM and N= 13,206 in MYOD_GMvsMYOD_DM); DI_MYOD_CTCF (N=2,666 in EMPTY_GMvsMYOD_GM and N=3,447 in MYOD_GMvsMYOD_DM); WG (N=759,086); WG_CTCF (N=47,625); WG_MYOD (N=44,741 MYOD_GM and N=68,989 MYOD_DM); Bd (N=2,241 in EMPTY_GMvsMYOD_GM and N=2,596 in MYOD_GMvsMYOD_DM); Bd_MYOD (N=933 in EMPTY_GMvsMYOD_GM and N=1,373 in MYOD_GMvsMYOD_DM). The

mean CDTS value (di Iulio et al., 2018) of every 4kb bin was extracted using bedops and the cumulative distribution function of all genomic bins in a given category were plotted using R version 3.4.3. The CDTS used was computed with whole genome sequencing data obtained from the gnomAD dataset (N=15,496). The explanation is provided here: http://www.hliopendata.com/noncoding/Pipeline/README_compute_CDTS_fromPublic Dataset.txt

The difference in the cumulative distribution function of different categories was assessed using two-sample Kolmogorov-Smirnov test.

Pathogenic variants distribution analysis. Variants from HGMD (Stenson et al., 2003) were filtered to retain only “High DM” flagged variants. The pathogenic variants were categorized into three non-mutually exclusive groups (All, Skeletal Muscle and Inflammation). All (N=154,503) – encompasses all pathogenic variants. Skeletal Muscle (N=5,888) – encompasses variants retrieved with the following key word extraction using *grep*: > grep "[Mm]usc\|[Mm]yopath\|[Mm]yogen\|[Mm]yasten" and grep -v "[Cc]ardio\|[Ss]tatin\|[Cc]ardiac\|[Cc]orne\|[Ss]mooth\|[Aa]drenoleukodystrophy". Inflammation (N=2,737) – includes variants retrieved with the following key word extraction using *grep*: > grep "[Ii]nflam\|itis" and grep -v "British_Columbia\|Pseudohermaphroditism\|[Hh]epatitis [ABC]". The fraction of genomic bins containing at least one pathogenic variant was then extracted. The difference in the fraction of bins with pathogenic variant of different categories was assessed using one sided Fisher’s Exact Test, with the following assumptions of expected fraction of variants

per category: WG < DI; WG < WG_CTCF; WG < Bd; WG < Bd_MYOD; WG_CTCF < Bd; DI < Bd.

***In situ* 3C.** 3C was performed as described for the *in-situ* Hi-C omitting the biotinylation step. Following Ethanol precipitation DNA is resuspended in 10mM Tris-HCl pH=8 and diluted to 25ng/µl. 1µl of the diluted DNA is then analysed by qPCR. Primers were designed using Primer3 and blasted to hg19 genome to ensure specificity for the fragment analyzed. For positive control, equimolar amount of BAC DNA containing the locus of interested were digested and ligated in the same way of sample DNA.

Acknowledgements

The authors would like to give special thanks to Dr. Alessandro Rosa for sharing Helper and epB-Puro-TT plasmids, Dr. David Huhta for his help with the computer cluster, Dr. Nicolas Servant for his help with HiC-Pro, Samantha Kuan and Bin Li for operation of the sequencing instruments and data processing, Amy Cortez for help with cell cycle analysis and Jonathan E. Henninger for helping with data analysis. We would like to thank Dr. Michael Walker for statistics guidance and the members of Puri's laboratory for helpful discussion. We are also grateful to Dr. Richard Young for his helpful suggestions on the manuscript. This work was supported by: NIAMS R01 AR056712; R01AR052779; AR061303 to PLP; Epigen Project Progetto Bandiera Epigenomica to PLP and SB; Ellison Medical Foundation AG-NS-0843-11; AFAR G16294 AD. Ludwig Institute for Cancer Research and a pilot project from the SAN DIEGO MUSCLE RESEARCH CENTER (SDMRC) to BR. YL is funded by a postdoctoral fellowship from the Human Frontier Program.

We apologize to authors whose work could not be cited because of space limitations.

Author Contribution

A.D. and P.L.P. conceptualized and organized the project, designed experiments and data analysis. A.D. set up the myogenic conversion of IMR90, conducted ChIP-seq, Hi-C and validation experiments as well as RNA-seq, ChIP-seq and Hi-C data analysis. L.C. conducted validation experiments and transcriptional inhibition experiments. C.N.

identified TADs and differential interactions from Hi-C matrices, performed permutation analysis and integrative analysis. A.S. shared helpful protocols, codes and suggestions for this study. J.d.I. performed genomic constraint and pathogenic variant analyses. S.G. and M.F. shared useful codes and suggestions. Y.D. prepared MYOD ChIP-seq libraries. Z.Y prepared RNA-seq libraries. R.P., A.T., S.B. and B.R provided input into experimental design and interpretation. A.D. and P.L.P. wrote the manuscript with inputs from all other authors. P.L.P. supervised this study. All authors have read and approved the manuscript. Correspondence and requests for materials should be addressed to P.L.P. (lpuri@sbpdiscovery.org) and A.D. (aled@wi.mit.edu).

PLP dedicates this work to his unforgettable colleague, teacher and friend Elisabetta "Betta" De Marzio, who will reamain for ever in his heart.

The accession number for the data for RNA-seq, ChIP-seq and Hi-C reported in this paper in NCBI GEO: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98530

Please use the following token: mnanyaakhzazvgj

The authors declare no conflict of interest.

References

Akdemir, K.C., and Chin, L. (2015). HiCPlotter integrates genomic data with interaction matrices. Genome Biol *16*, 198.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Andrey, G., and Mundlos, S. (2017). The three-dimensional genome: regulating gene expression during pluripotency and development. Development *144*, 3646-3658.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res *37*, W202-208.

Battistelli, C., Busanello, A., and Maione, R. (2014). Functional interplay between MyoD and CTCF in regulating long-range chromatin interactions during differentiation. J Cell Sci *127*, 3757-3767.

Beagan, J.A., Gilgenast, T.G., Kim, J., Plona, Z., Norton, H.K., Hu, G., Hsu, S.C., Shields, E.J., Lyu, X., Apostolou, E.*, et al.* (2016). Local Genome Topology Can Exhibit an Incompletely Rewired 3D-Folding State during Somatic Cell Reprogramming. Cell Stem Cell *18*, 611-624.

Bengal, E., Ransone, L., Scharfmann, R., Dwarki, V.J., Tapscott, S.J., Weintraub, H., and Verma, I.M. (1992). Functional antagonism between c-Jun and MyoD proteins: a direct physical association. Cell *68*, 507-519.

Berkes, C.A., Bergstrom, D.A., Penn, B.H., Seaver, K.J., Knoepfler, P.S., and Tapscott, S.J. (2004). Pbx marks genes for activation by MyoD indicating a role for a homeodomain protein in establishing myogenic potential. Mol Cell *14*, 465-477.

Black, B.L., Molkentin, J.D., and Olson, E.N. (1998). Multiple roles for the MyoD basic region in transmission of transcriptional activation signals and interaction with MEF2. Mol Cell Biol *18*, 69-77.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nat Rev Genet *17*, 661-678.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A.*, et al.* (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell *171*, 557-572 e524.

Cabrera, C.P., Navarro, P., Huffman, J.E., Wright, A.F., Hayward, C., Campbell, H., Wilson, J.F., Rudan, I., Hastie, N.D., Vitart, V.*, et al.* (2012). Uncovering networks from genome-wide association studies via circular genomic permutation. G3 (Bethesda) *2*, 1067-1075.

Caiazzo, M., Dell'Anno, M.T., Dvoretskova, E., Lazarevic, D., Taverna, S., Leo, D., Sotnikova, T.D., Menegon, A., Roncaglia, P., Colciago, G.*, et al.* (2011). Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. Nature *476*, 224-227.

Cao, Y., Yao, Z., Sarkar, D., Lawrence, M., Sanchez, G.J., Parker, M.H., MacQuarrie, K.L., Davison, J., Morgan, M.T., Ruzzo, W.L.*, et al.* (2010). Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. Dev Cell *18*, 662-674.

Chandra, T., Ewels, P.A., Schoenfelder, S., Furlan-Magaril, M., Wingett, S.W., Kirschner, K., Thuret, J.Y., Andrews, S., Fraser, P., and Reik, W. (2015). Global reorganization of the nuclear landscape in senescent cells. Cell Rep *10*, 471-483.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative Binding of Transcription Factors Orchestrates Reprogramming. Cell *168*, 442-459 e420.
Ciglar, L., Girardot, C., Wilczynski, B., Braun, M., and Furlong, E.E. (2014). Coordinated repression and activation of two transcriptional programs stabilizes cell fate during myogenesis. Development *141*, 2633-2643.
Conerly, M.L., Yao, Z., Zhong, J.W., Groudine, M., and Tapscott, S.J. (2016). Distinct Activities of Myf5 and MyoD Indicate Separate Roles in Skeletal Muscle Lineage Specification and Differentiation. Dev Cell *36*, 375-385.
Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.
Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature *523*, 240-244.
Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987-1000.
de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M.*, et al.* (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227-231.
Dekker, J., and Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. Cell *164*, 1110-1121.
Delgado-Olguin, P., Brand-Arzamendi, K., Scott, I.C., Jungblut, B., Stainier, D.Y., Bruneau, B.G., and Recillas-Targa, F. (2011). CTCF promotes muscle differentiation by modulating the activity of myogenic regulatory factors. J Biol Chem *286*, 12483-12494.
Di Giammartino D., L.Y., Kloetgen A., Polyzos A, Kim D, Stadtfeld M, Tsirigos A, Apostolou E (2018). KLF4 binding during reprogramming is involved in 3D architectural rewiring and transcriptional regulation of enhancer hubs. Biorxiv.
di Iulio, J., Bartha, I., Wong, E.H.M., Yu, H.C., Lavrenko, V., Yang, D., Jung, I., Hicks, M.A., Shah, N., Kirkness, E.F.*, et al.* (2018). The human noncoding genome defined by genetic diversity. Nat Genet *50*, 333-337.
Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W.*, et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331-336.
Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.
Dodou, E., Xu, S.M., and Black, B.L. (2003). mef2c is activated directly by myogenic basic helix-loop-helix proteins during skeletal muscle development in vivo. Mech Dev *120*, 1021-1032.
Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K.*, et al.* (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell *159*, 374-387.
Filippova, D., Patro, R., Duggal, G., and Kingsford, C. (2014). Identification of alternative topological domains in chromatin. Algorithms Mol Biol *9*, 14.
Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suva, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature *529*, 110-114.

Fong, A.P., Yao, Z., Zhong, J.W., Cao, Y., Ruzzo, W.L., Gentleman, R.C., and Tapscott, S.J. (2012). Genetic and epigenetic determinants of neurogenesis and myogenesis. Dev Cell *22*, 721-735.
Franke, M., Ibrahim, D.M., Andrey, G., Schwarzer, W., Heinrich, V., Schopflin, R., Kraft, K., Kempfer, R., Jerkovic, I., Chan, W.L.*, et al.* (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. Nature *538*, 265-269.
Gerber, A.N., Klesert, T.R., Bergstrom, D.A., and Tapscott, S.J. (1997). Two domains of MyoD mediate transcriptional activation of genes in repressive chromatin: a mechanism for lineage determination in myogenesis. Genes Dev *11*, 436-450.
Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., and Furlong, E.E. (2014). Enhancer loops appear stable during development and are associated with paused polymerase. Nature *512*, 96-100.
Guerreiro, I., Gitto, S., Novoa, A., Codourey, J., Nguyen Huynh, T.H., Gonzalez, F., Milinkovitch, M.C., Mallo, M., and Duboule, D. (2016). Reorganisation of Hoxd regulatory landscapes during the evolution of a snake-like body plan. Elife *5*.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.
Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934-947.
Hnisz, D., Day, D.S., and Young, R.A. (2016a). Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. Cell *167*, 1188-1200.
Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A.*, et al.* (2016b). Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science *351*, 1454-1458.
Ieda, M., Fu, J.D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell *142*, 375-386.
Isoda, T., Moore, A.J., He, Z., Chandra, V., Aida, M., Denholtz, M., Piet van Hamburg, J., Fisch, K.M., Chang, A.N., Fahl, S.P.*, et al.* (2017). Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. Cell *171*, 103-119 e118.
Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Varnai, C., Thiecke, M.J.*, et al.* (2016). Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell *167*, 1369-1384 e1319.
Ji, X., Dadon, D.B., Powell, B.E., Fan, Z.P., Borges-Rivera, D., Shachar, S., Weintraub, A.S., Hnisz, D., Pegoraro, G., Lee, T.I.*, et al.* (2016). 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. Cell Stem Cell *18*, 262-275.
Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290-294.
Kabadi, A.M., Ousterout, D.G., Hilton, I.B., and Gersbach, C.A. (2014). Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. Nucleic Acids Res *42*, e147.

Kieffer-Kwon, K.R., Nimura, K., Rao, S.S.P., Xu, J., Jung, S., Pekowska, A., Dose, M., Stevens, E., Mathe, E., Dong, P.*, et al.* (2017). Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. Mol Cell *67*, 566-578 e510.
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.
Kragesteen, B.K., Spielmann, M., Paliou, C., Heinrich, V., Schopflin, R., Esposito, A., Annunziatella, C., Bianco, S., Chiariello, A.M., Jerkovic, I.*, et al.* (2018). Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. Nat Genet *50*, 1463-1473.
Krijger, P.H., Di Stefano, B., de Wit, E., Limone, F., van Oevelen, C., de Laat, W., and Graf, T. (2016). Cell-of-Origin-Specific 3D Genome Structure Acquired during Somatic Cell Reprogramming. Cell Stem Cell *18*, 597-610.
Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.
Latella, L., Dall'Agnese, A., Boscolo, F.S., Nardoni, C., Cosentino, M., Lahm, A., Sacco, A., and Puri, P.L. (2017). DNA damage signaling mediates the functional antagonism between replicative senescence and terminal muscle differentiation. Genes Dev *31*, 648-659.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.
Li, L., Chambard, J.C., Karin, M., and Olson, E.N. (1992). Fos and Jun repress transcriptional activation by myogenin and MyoD: the amino terminus of Jun can mediate repression. Genes Dev *6*, 676-689.
Liu, D., Black, B.L., and Derynck, R. (2001). TGF-beta inhibits muscle differentiation through functional repression of myogenic transcription factors by Smad3. Genes Dev *15*, 2950-2966.
Liu, S., Chen, H., Ronquist, S., Seaman, L., Ceglia, N., Meixner, W., Chen, P.Y., Higgins, G., Baldi, P., Smale, S.*, et al.* (2018). Genome Architecture Mediates Transcriptional Control of Human Myogenic Reprogramming. iScience *6*, 232-246.
Loell, I., and Lundberg, I.E. (2011). Can muscle regeneration fail in chronic inflammation: a weakness in inflammatory myopathies? J Intern Med *269*, 243-257.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
Lun, A.T., and Smyth, G.K. (2015). diffHic: a Bioconductor package to detect differential genomic interactions in Hi-C data. BMC Bioinformatics *16*, 258.
Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R.*, et al.* (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012-1025.
Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C.Y., Chou, A., Ienasescu, H.*, et al.* (2014). JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res *42*, D142-147.
Mousavi, K., Zare, H., Dell'orso, S., Grontved, L., Gutierrez-Cruz, G., Derfoul, A., Hager, G.L., and Sartorelli, V. (2013). eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. Mol Cell *51*, 606-617.

Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science *347*, 1017-1021.
Natoli, G. (2010). Maintaining cell identity through global control of genomic organization. Immunity *33*, 12-24.
Neph, S., Kuehn, M.S., Reynolds, A.P., Haugen, E., Thurman, R.E., Johnson, A.K., Rynes, E., Maurano, M.T., Vierstra, J., Thomas, S.*, et al.* (2012). BEDOPS: high-performance genomic feature operations. Bioinformatics *28*, 1919-1920.
Noordermeer, D., Leleu, M., Schorderet, P., Joye, E., Chabaud, F., and Duboule, D. (2014). Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. Elife *3*, e02557.
Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W., and Duboule, D. (2011). The dynamic architecture of Hox gene clusters. Science *334*, 222-225.
Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944 e922.
Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.
Ong, C.T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nat Rev Genet *15*, 234-246.
Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet *35*, 190-194.
Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiano, V., Marro, S., Sudhof, T.C.*, et al.* (2011). Induction of human neuronal cells by defined transcription factors. Nature *476*, 220-223.
Pfisterer, U., Kirkeby, A., Torper, O., Wood, J., Nelander, J., Dufour, A., Bjorklund, A., Lindvall, O., Jakobsson, J., and Parmar, M. (2011). Direct conversion of human fibroblasts to dopaminergic neurons. Proc Natl Acad Sci U S A *108*, 10343-10348.
Phanstiel, D.H., Van Bortle, K., Spacek, D., Hess, G.T., Shamim, M.S., Machol, I., Love, M.I., Aiden, E.L., Bassik, M.C., and Snyder, M.P. (2017). Static and Dynamic DNA Loops form AP-1-Bound Activation Hubs during Macrophage Development. Mol Cell *67*, 1037-1048 e1036.
Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y.*, et al.* (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281-1295.
Puri, P.L., Iezzi, S., Stiegler, P., Chen, T.T., Schiltz, R.L., Muscat, G.E., Giordano, A., Kedes, L., Wang, J.Y., and Sartorelli, V. (2001). Class I histone deacetylases sequentially interact with MyoD and pRb during skeletal myogenesis. Mol Cell *8*, 885-897.
Puri, P.L., and Sartorelli, V. (2000). Regulation of muscle regulatory factors by DNA-binding, interacting proteins, and post-transcriptional modifications. J Cell Physiol *185*, 155-173.
Qian, L., Huang, Y., Spencer, C.I., Foley, A., Vedantham, V., Liu, L., Conway, S.J., Fu, J.D., and Srivastava, D. (2012). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. Nature *485*, 593-598.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.
Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.
Remeseiro, S., Hornblad, A., and Spitz, F. (2016). Gene regulation during development in the light of topologically associating domains. Wiley Interdiscip Rev Dev Biol *5*, 169-185.
Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.
Rodriguez-Carballo, E., Lopez-Delisle, L., Zhan, Y., Fabre, P.J., Beccari, L., El-Idrissi, I., Huynh, T.H.N., Ozadam, H., Dekker, J., and Duboule, D. (2017). The HoxD cluster is a dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory landscapes. Genes Dev *31*, 2264-2281.
Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J.*, et al.* (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A *112*, E6456-6465.
Sartorelli, V., and Puri, P.L. (2018). Shaping Gene Expression by Landscaping Chromatin Architecture: Lessons from a Master. Mol Cell *71*, 375-388.
Schaub, J.R., Huppert, K.A., Kurial, S.N.T., Hsu, B.Y., Cast, A.E., Donnelly, B., Karns, R.A., Chen, F., Rezvani, M., Luu, H.Y.*, et al.* (2018). De novo formation of the biliary system by TGFbeta-mediated hepatocyte transdifferentiation. Nature *557*, 247-251.
Schauer, T., Ghavi-Helm, Y., Sexton, T., Albig, C., Regnard, C., Cavalli, G., Furlong, E.E., and Becker, P.B. (2017). Chromosome topology guides the Drosophila Dosage Compensation Complex for target gene activation. EMBO Rep.
Schmitt, A.D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C.L., Li, Y., Lin, S., Lin, Y., Barr, C.L.*, et al.* (2016a). A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. Cell Rep *17*, 2042-2059.
Schmitt, A.D., Hu, M., and Ren, B. (2016b). Genome-wide mapping and analysis of chromosome architecture. Nat Rev Mol Cell Biol.
Schuijers, J., Manteiga, J.C., Weintraub, A.S., Day, D.S., Zamudio, A.V., Hnisz, D., Lee, T.I., and Young, R.A. (2018). Transcriptional Dysregulation of MYC Reveals Common Enhancer-Docking Mechanism. Cell Rep *23*, 349-360.
Servant, N., Lajoie, B.R., Nora, E.P., Giorgetti, L., Chen, C.J., Heard, E., Dekker, J., and Barillot, E. (2012). HiTC: exploration of high-throughput 'C' experiments. Bioinformatics *28*, 2843-2844.
Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol *16*, 259.
Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.
Siersbaek, R., Madsen, J.G.S., Javierre, B.M., Nielsen, R., Bagge, E.K., Cairns, J., Wingett, S.W., Traynor, S., Spivakov, M., Fraser, P.*, et al.* (2017). Dynamic Rewiring of Promoter-Anchored Chromatin Loops during Adipocyte Differentiation. Mol Cell *66*, 420-435 e425.

Singh, K., Cassano, M., Planet, E., Sebastian, S., Jang, S.M., Sohi, G., Faralli, H., Choi, J., Youn, H.D., Dilworth, F.J., *et al.* (2015). A KAP1 phosphorylation switch controls MyoD function during skeletal muscle differentiation. Genes Dev *29*, 513-525.
Spielmann, M., Lupianez, D.G., and Mundlos, S. (2018). Structural variation in the 3D genome. Nat Rev Genet *19*, 453-467.
Spielmann, M., and Mundlos, S. (2016). Looking beyond the genes: the role of non-coding variants in human disease. Hum Mol Genet *25*, R157-R165.
Spitz, F., and Furlong, E.E. (2012). Transcription factors: from enhancer binding to developmental control. Nat Rev Genet *13*, 613-626.
Stadhouders, R., Vidal, E., Serra, F., Di Stefano, B., Le Dily, F., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y., *et al.* (2018). Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. Nat Genet *50*, 238-249.
Stenson, P.D., Ball, E.V., Mort, M., Phillips, A.D., Shiel, J.A., Thomas, N.S., Abeysinghe, S., Krawczak, M., and Cooper, D.N. (2003). Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat *21*, 577-581.
Sun, F., Chronis, C., Kronenberg, M., Chen, X.F., Su, T., Lay, F.D., Plath, K., Kurdistani, S.K., and Carey, M.F. (2019). Promoter-Enhancer Communication Occurs Primarily within Insulated Neighborhoods. Mol Cell *73*, 250-263 e255.
Symmons, O., Pan, L., Remeseiro, S., Aktas, T., Klein, F., Huber, W., and Spitz, F. (2016). The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. Dev Cell *39*, 529-543.
Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.
Tsai, P.F., Dell'Orso, S., Rodriguez, J., Vivanco, K.O., Ko, K.D., Jiang, K., Juan, A.H., Sarshad, A.A., Vian, L., Tran, M., *et al.* (2018). A Muscle-Specific Enhancer RNA Mediates Cohesin Recruitment and Regulates Transcription In trans. Mol Cell *71*, 129-141 e128.
Tsui, S., Wang, J., Wang, L., Dai, W., and Lu, L. (2016). CTCF-Mediated and Pax6-Associated Gene Expression in Corneal Epithelial Cell-Specific Differentiation. PLoS One *11*, e0162071.
Tsunemoto, R., Lee, S., Szucs, A., Chubukov, P., Sokolova, I., Blanchard, J.W., Eade, K.T., Bruggemann, J., Wu, C., Torkamani, A., *et al.* (2018). Diverse reprogramming codes for neuronal identity. Nature *557*, 375-380.
Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Sudhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035-1041.
Wada, R., Muraoka, N., Inagawa, K., Yamakawa, H., Miyamoto, K., Sadahiro, T., Umei, T., Kaneda, R., Suzuki, T., Kamiya, K., *et al.* (2013). Induction of human cardiomyocyte-like cells from fibroblasts by defined factors. Proc Natl Acad Sci U S A *110*, 12667-12672.
Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., *et al.* (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell *171*, 1573-1588 e1528.
Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T.K., Turner, D., Rupp, R., Hollenberg, S., *et al.* (1991). The myoD gene family: nodal point during specification of the muscle cell lineage. Science *251*, 761-766.

Weintraub, H., Tapscott, S.J., Davis, R.L., Thayer, M.J., Adam, M.A., Lassar, A.B., and Miller, A.D. (1989). Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. Proc Natl Acad Sci U S A *86*, 5434-5438.
Yamamoto, M., Legendre, N.P., Biswas, A.A., Lawton, A., Yamamoto, S., Tajbakhsh, S., Kardon, G., and Goldhamer, D.J. (2018). Loss of MyoD and Myf5 in Skeletal Muscle Stem Cells Results in Altered Myogenic Programming and Failed Regeneration. Stem Cell Reports *10*, 956-969.
Yan, K.K., Yardimci, G.G., Yan, C., Noble, W.S., and Gerstein, M. (2017). HiC-spector: a matrix library for spectral and reproducibility analysis of Hi-C contact maps. Bioinformatics *33*, 2199-2201.
Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., Pope, B.D.*, et al.* (2014). A comparative encyclopedia of DNA elements in the mouse genome. Nature *515*, 355-364.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Supplementary Table 4: Primers for gene expression

Primer Name	Primer Sequence
mouse MYOD mRNA For	AGCACTACAGTGGCGACTCA
mouse MYOD mRNA Rev	GGCCGCTGTAATCCATCAT
hGAPDH mRNA For	GAAGGTGAAGGTCGGAGTC
hGAPDH mRNA Rev	GAAGATGGTGATGGGATTTC
hGAPDH premRNA Rev	CCATACGACTGCAAAGACCC
hbeta-actin mRNA For	CCTGGGCATGGAGTCCTGTGG
hbeta-actin mRNA Rev	CTGTGTTGGCGTACAGGTCTT
hMYOG mRNA For	AATGCAGCTCTCACAGCGCCTC
hMYOG mRNA Rev	TCAGCCGTGAGCAGATGATCC
hMYHC1 mRNA For	CGAAGCTGGAGCTACTGTAA
hMYHC1 mRNA Rev	CCATGTCCTCGATCTTGTCATA
hCKM mRNA For	TGGAGAAGCTCTCTGTGGAAG
hCKM mRNA Rev	TCCGTCATGCTCTTCAGAGGGT
hTGF-β1 mRNA For	GCCTGAGGCCGACTACTA
hTGF-β1 mRNA Rev	CTGTGTGTACTCTGCTTGAACT
hCTGF mRNA For	TGTGCACCGCCAAAGAT
hCTGF mRNA Rev	GCACGTGCACTGGTACTT
hPOSTN mRNA For	CTGCTTATTGTTAACCCTATAAACGC
hPOSTN mRNA Rev	CAGACATTTGGGCCTTGGT
hIL6 mRNA For	CGGGAACGAAAGAGAAGCTCTA
hIL6 mRNA Rev	GGCGCTTGTGGAGAAGGAG
hVCAM1 mRNA For	CCGGATTGCTGCTCAGATTG
hVCAM1 mRNA Rev	AGCGTGGAATTGGTCCCCTCA
hIL1b mRNA For	CATCTACGAATCTCCGACCAC
hIL1b mRNA Rev	GGCAGGGAACCAGCATCTTC
hTNNT2 mRNA For	TCAAAGTCCACTCTCTCTCCATC
hTNNT2 mRNA Rev	GGAGGAGTCCAAACCAAAGCC
hCTCF mRNA For	GTGGCGCGGAGAATGATTA
hCTCF mRNA Rev	AATAGAACCCAGCTCTCAAGC

hITGA7 mRNA For	CGCCTTCAATCTGGACGTGA
hITGA7 mRNA Rev	CCCACCAGCAGCCAGC
hRDH5 mRNA For	TTAGTGTGGGAGGCTGGGAA
hRDH5 mRNA Rev	GTGGCAGCCTCCTGTGG
hBLOC1S1 mRNA For	CAGGCACTCAAGGAAATTGGG
hBLOC1S1 mRNA Rev	GACTGCAGCTGCCCTTTGTA
hLAD1 mRNA For	CACGGCCATACGGAGATCAG
hLAD1 mRNA Rev	TTCTCAAAGAGGTGGCGCTT
hTNNI1 mRNA For	GAGGTGTTCCAACCTGGGAG
hTNNI1 mRNA Rev	ACGCATACACACCATGCTCA
hPKP1 mRNA For	ACCGCGTCATCATTCCCTTC
hPKP1 mRNA Rev	AGCTCAGGTTCCTCAAGCAG
hIGFN1 mRNA For	CAGCTGTACACCCAGGGTAAT
hIGNF1 mRNA Rev	TCACTCCAGGGATGTGGGAC
hSARNP mRNA For	TCTTGCTCGTGGTTTGGAGA
hSARNP mRNA Rev	TTTGCCTCCTCTTCAGCATGT
mouse b-actin mRNA For	GGCTGTATTCCCCTCCATCG
mouse b-actin mRNA Rev	CCAGTTGGTAACAATGCCATGT
mouse Tnnt2 mRNA For	CAGAGGAGGCCAACGTAGAAG
mouse Tnnt2 mRNA Rev	CTCCATCGGGGATCTTGGGT

Supplementary Table 5: Primers for ChIP-qPCR

Primer Name	Primer Sequence
hTNNT2 promoter For	GCAGAGGGCAAGAGTTATGT
hTNNT2 promoter Rev	AAGGCTGCTCAGTCCATTAG
hTNNT2 enhancer For	TCACATGCTGTTCCCTCTATTT
hTNNT2 enhancer Rev	GGTGTCATGGAGCAATTGATTT
hIL6 promoter For	CTTCGGTCCAGTTGCCTTC
hIL6 promoter Rev	GCGGCTACATCTTTGGAATCT
hTGF-β1 promoter For	GTCACCAGAGAAAGAGGACCAG
hTGF-β1 promoter Rev	CTACCTTGTTTCCCAGCCTGA
hTGF-β1 enhancer For	CAAAACAGACCCACACAACTCC
hTGF-β1 enhancer Rev	CAGTGTTTGTCTGTGTGAGGG
hTGF-β1 IN5 For	CGTCAAAATGTCAAAATGCCGT
hTGF-β1 IN5 Rev	TTACATTCTCAAGCAGGCAGTG
hTGF-β1 IN3 For	TGAGCTTCAAAACGAGACACAG
hTGF-β1 IN3 Rev	GGTTGCCCTATTTCTTTGTGGA
hITGA7 promoter For	ACCCATCACGTCCAGATTGAAG
hITGA7 promoter Rev	CTCCGGGATTTGCTACCTTTT
hITGA7 enhancer For	CTGTAGGTCACTTGGGCTCC
hITGA7 enhancer For	CGGTGATGAAGACAAAGGCATT

Supplementary Table 6: Primers for 3C

Primer Name	Primer Sequence	Genomic Coordinates DpnII Fragment
Human hg19		

3C_LoadingCtrl_F	AATGGGCTTAATGGAAGACAAAT	chr12 6632801 6634701
3C_LoadingCtrl_R	TGTGCTTAAAATTTCGGTCATCT	chr12 6632801 6634701
TNNT2_201347539 Viewpoint MYOD peak	GTCATCCCCTAACGGCTTTAAAA	chr1 201346773 201347562
TNNT2_201346763 Viewpoint CTCF peak	GAGCCTTACCTCAGAACAGCAG	chr1 201345693 201346776
TNNT2_201354973	AGTGTCACATAAGAGGTCTAGGG	chr1 201354133 201355158
TNNT2_201356497	AAGGTATCATCTGAGCAAAGCAC	chr1 201355155 201356621
TNNT2_201352732	TTCCAAGAAACGTCCCATTCAAT	chr1 201352335 201352895
TNNT2_201351662	CTCCTGAAATGACTCCTCCTCTC	chr1 201350893 201351795
TNNT2_201349564	AACTCTCGGCTATGAAGTTGGAA	chr1 201348865 201349739
TNNT2_201357190	CCCCACCTCATAATCCCAAAATC	chr1 201356659 201357298
TNNT2_201354065	AAGACATGTCCTCCTCCTCTTC	chr1 201353778 201354136
TNNT2_201353395	AGTAAATGTTCATGGGGCAGATG	chr1 201352950 201353451
TNNT2_201353646	CCTTTCCCTGTCCTAACCTCAAG	chr1 201353448 201353713
TNNT2_201352338	ACAGGAGCCACGAATAACTCAG	chr1 201351817 201352338
TGF-β1_41946123 Viewpoint IN	TCTCCACAAAGAAATAGGGCAAC	chr19 41946101 41946123
TGF-β1_41827968	TTGAGAATGTAAGCCCATGTGTC	chr19 41827946 41827968
TGF-β1_41827078	CATCAGGAAAGTGGGAATTGTGA	chr19 41827056 41827078
TGF-β1_41825348	CCTGCATGTTTTCCTGTGTTTT	chr19 41825327 41825348
TGF-β1_41829197	TTAGGTTCCCATCTCTCCTTCTG	chr19 41829175 41829197
TGF-β1_41830037	CAACCAGAATTAGAGCCAGACAG	chr19 41830015 41830037
TGF-β1_41859165 Viewpoint Prom	TCCCACGGAAATAACCTAGATG	chr19 41858791 41859165
TGF-β1_41832499	CTGAGAACCACAGGAAGCATG	chr19 41830487 41832499
TGF-β1_41834210	GATGTCAGACTGAGTGGGTAAGA	chr19 41832786 41834210
TGF-β1_41835463	CACCACAATGCCAGACTAATTTT	chr19 41834207 41835463
TGF-β1_41838117	CCAACTCACCTCTCTGACTTTAC	chr19 41837557 41838117
ITGA7_56100255 Viewpoint	ACAGTAGAGGCAGAAGACAGTC	chr12 56099397 56100255
ITGA7_56112320	TAACCACAGCCCATACCTTGAT	chr12 56111586 56112320
ITGA7_56114416	GGAACTGAGAAGACTAGCCAGA	chr12 56112971 56114416
ITGA7_56115224	CAGAGAGGCTTCCGAGTCC	chr12 56114729 56115224
ITGA7_56115639	CGGGTGCTGAATGTGAACAC	chr12 56115321 56115639
Mouse mm10		
3C_loadingCtrl_mm_F	GAACGGTACCACACTCAGTTTAC	chr5 142911269 142911402
3C_loadingCtrl_mm_R	AGACCTGACTCTTCAAGCTATCA	chr5 142911269 142911402
Tnnt2_135836700 Viewpoint Myod	TGTTTCCAGGACAGACTCTAACA	chr1 135835646 135836700
Tnnt2_135835151	CACACATGTCTCAGGATAAGCC	chr1 135834524 135835151
Tnnt2_135834527 Viewpoint CTCF	GTGTAGGAGGTGACATTTGAGTG	chr1 135833823 135834527
Tnnt2_135833826	CCCTTTGTCTTAGGCCCTATAGT	chr1 135833115 135833826
Tnnt2_135832869	AGTGCTAGACCTATTTACCGCTC	chr1 135832431 135832869
Tnnt2_135832434	TCCCTTTCCTTGTTCTAATTGGC	chr1 135831572 135832434
Tnnt2_135831575	CGTGACATCAAGACTTAACCTGT	chr1 135831231 135831575
Tnnt2_135831234	GTACTGAAGTCCTTGTCACACCT	chr1 135830616 135831234
Tnnt2_135830422	GCAAGAGGAGAACAATGACAAGA	chr1 135828430 135830422
Tnnt2_135828433	GGTAGCCTAGACTCCTGTTTTCT	chr1 135826900 135828433
Tnnt2_135825883	CCTGAGTTCAACCACAGATGATG	chr1 135825088 135825883

Supplementary Table 7: guide RNAs

Name	Sequence	PAM	Coordinates
TNNT2_M1	CAGCAGCTGCCGACAGATCC	TGG	chr1 201346758 201346777
TNNT2_ M2	CCACATGGGCTTATATGGCG	TGG	chr1 201346994 201347013
TNNT2_ M3	TAGCTTATCTGAGCAGCTGG	AGG	chr1 201347018 201347037
TNNT2_ M4	AGGGCTTTAAGCAGGCATGT	GGG	chr1 201346829 201346848
ITGA7_CM1	TGGCTCTGGGAGACGGAACC	AGG	chr12 56099833 56099852
ITGA7_CM2	CCCCCTGCTGGAGCAAAAGC	AGG	chr12 56099774 56099793
ITGA7_CM3	CCACCTGCAACTCAAAAGCT	AGG	chr12 56099697 56099716
RDH5_CTCF1	CTGCCACCTGTAGGTCACTT	GGG	chr12 56114938 56114957
RDH5_CTCF2	TTCTGCTGGGTGCCTTACTC	TGG	chr12 56114982 56115001
RDH5_CTCF3	GCTTGAGGGCCACAGTAAAC	TGG	chr12 56114891 56114910

Figure

Dall'Agnese et al., Figure 1

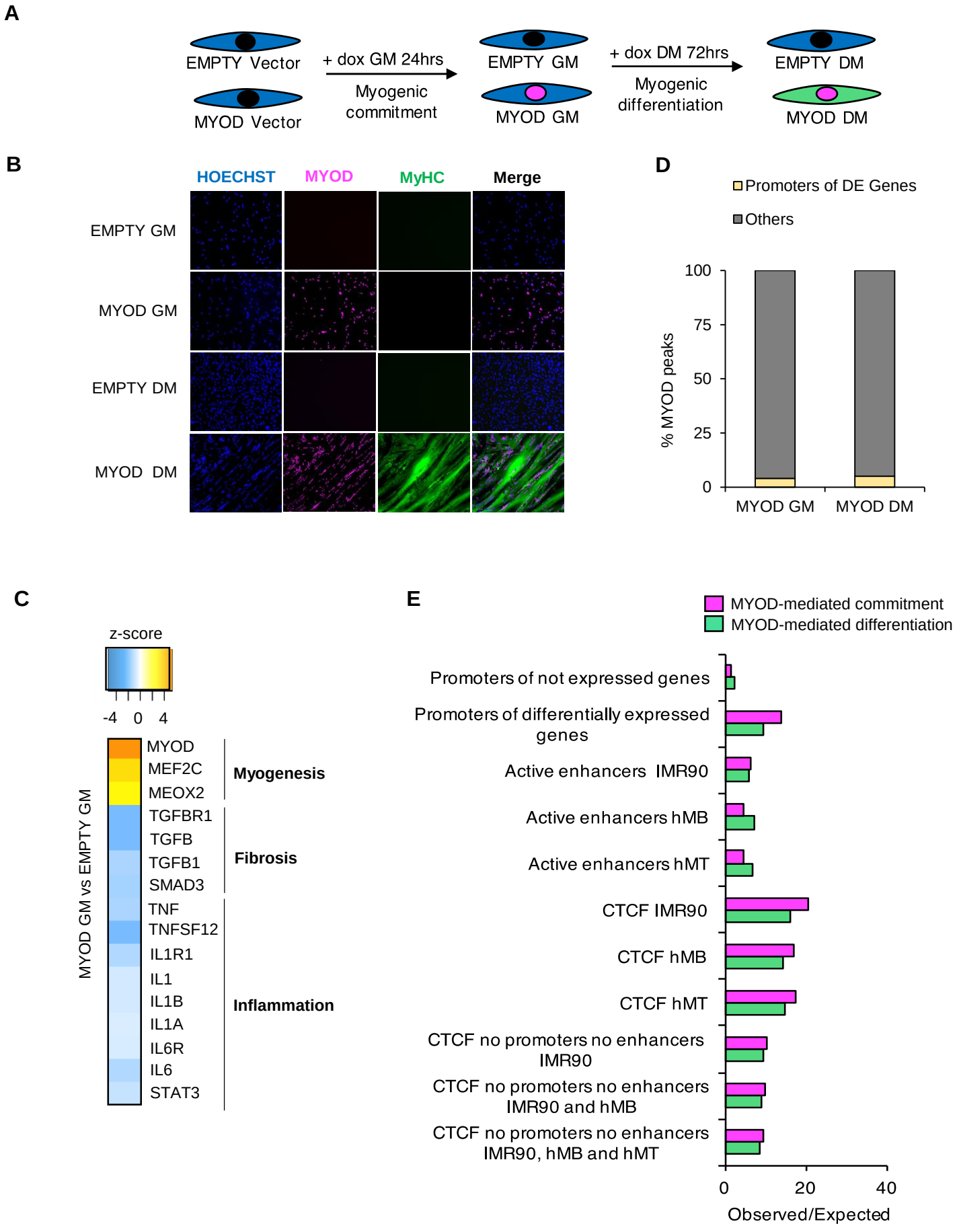


Figure 1: MYOD regulation of gene expression in the linear sequence of the DNA.

A, Experimental design. IMR90 fibroblasts were electroporated with doxycycline-inducible MYOD or EMPTY vector, cells were treated with doxycycline (dox) for 24hrs in growth media (GM) prior differentiation stimuli (DM) with doxycycline for 72hrs. Experiments were always performed at these time points, unless otherwise stated.
B, Representative immunofluorescence images of IMR90 cells stained for MYOD (magenta) and MyHC antibody (green). The nuclei were stained with Hoechst (blue).
C, Transcriptional networks predicted to be altered by MYOD comparing MYOD GM vs EMPTY GM (Ingenuity Pathway Analysis, IPA, Qiagen). For all predictions, $p<0.001$.
D, Percentage of MYOD peaks at promoters of differentially expressed (DE) genes (yellow) or not (grey).
E, Observed/expected ratio of MYOD binding at the genomic regions listed in the y axes, as described in Chronis et al, 2017 (see Methods).

Dall'Agnese et al., Figure 2

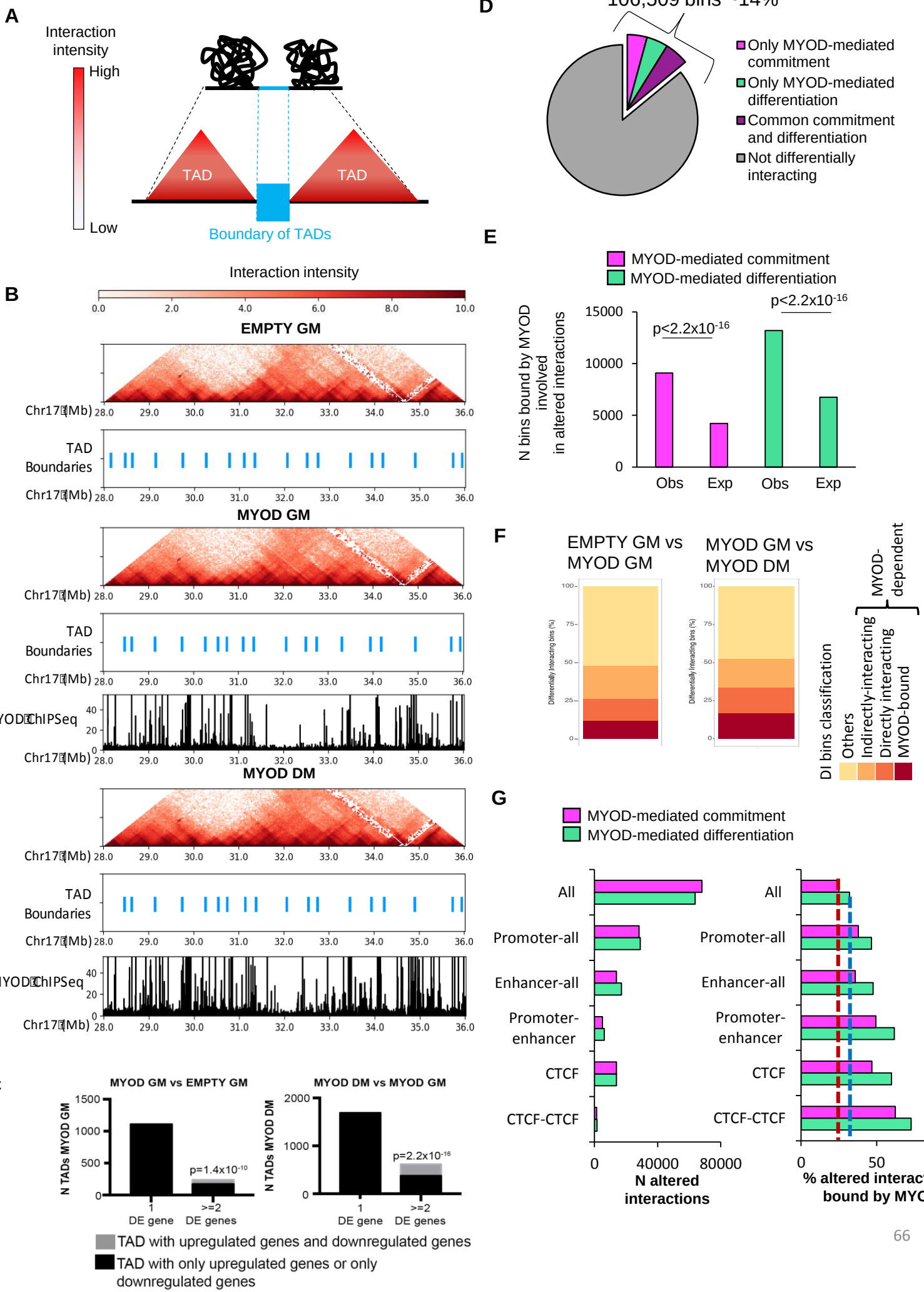
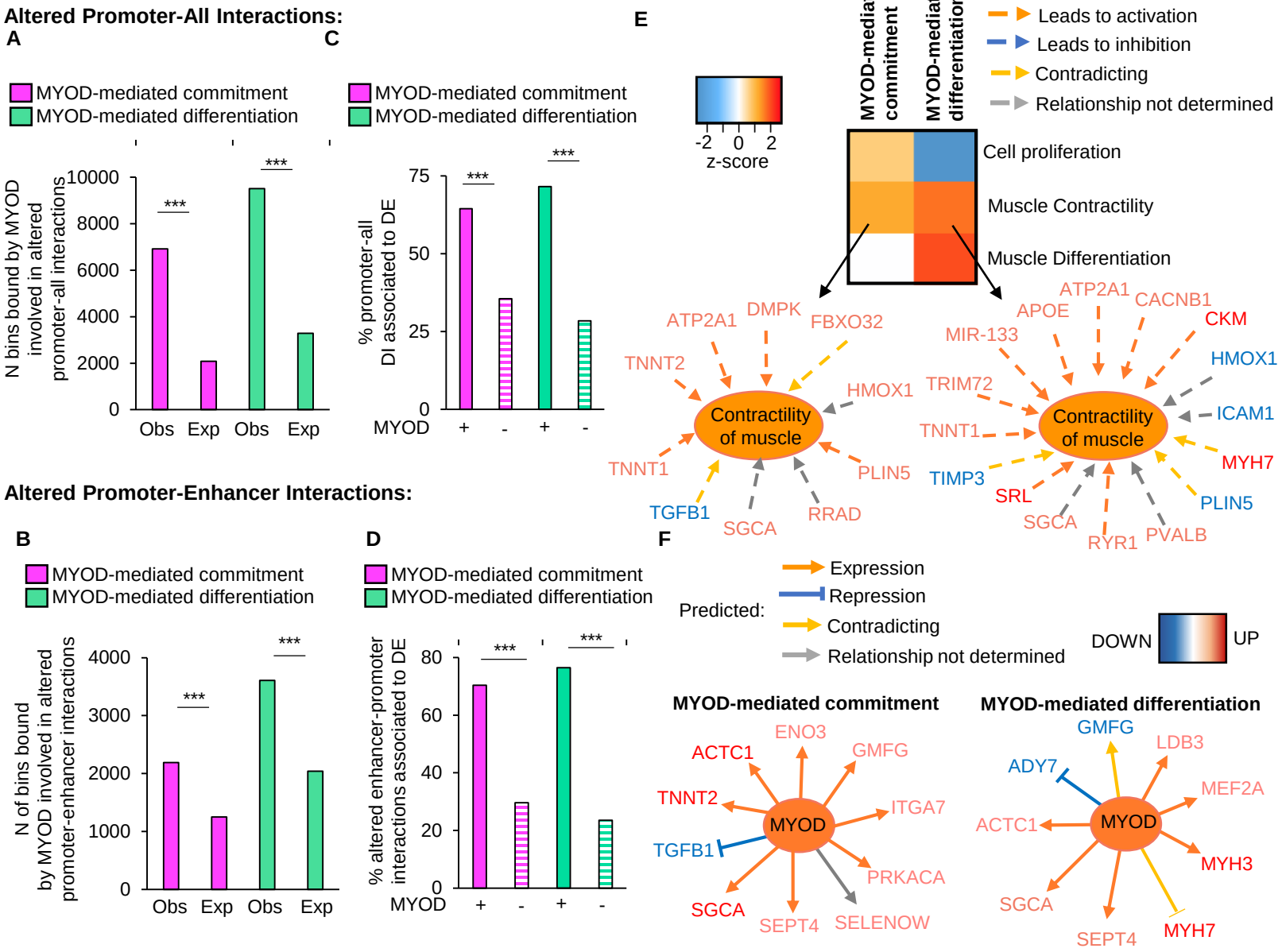
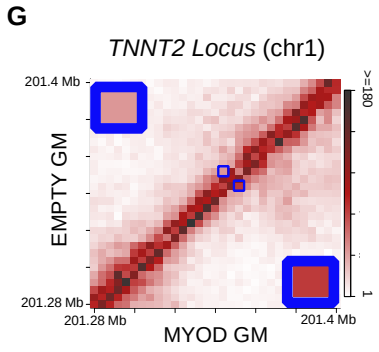


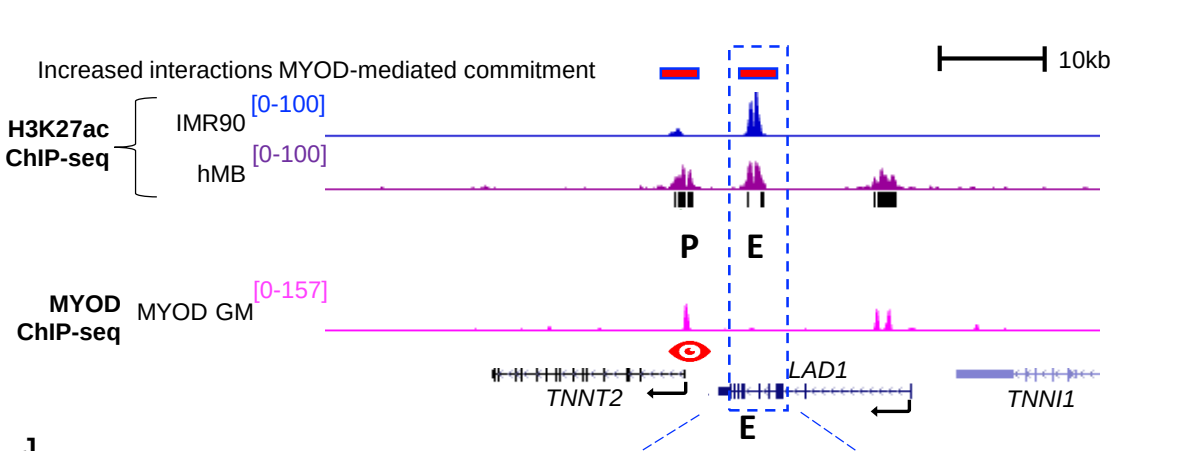
Figure 2: Profound alteration of chromatin contacts by MYOD during myogenic conversion

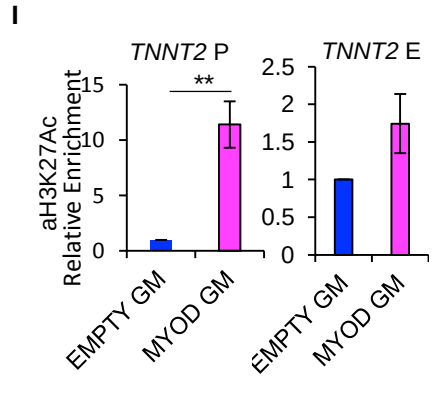
A, Graphical representation of TADs and TAD boundaries.
B, Hi-C interaction pattern (red heat map) and TAD boundaries (light blue) in EMPTY GM, MYOD GM and MYOD DM, MYOD ChIP-seq in MYOD GM and MYOD DM (black).
C, Number of TADs with one or more DE genes. Black represents the TADs whose differentially expressed genes are all upregulated or all downregulated, while grey represents the TADs containing upregulated genes and downregulated genes. LEFT: gene expression comparison between EMPTY GM and MYOD GM. TADs used were identified in MYOD GM. RIGHT: gene expression comparison between MYOD GM and MYOD DM. TADs used were identified in MYOD DM. pvalue represent the significant prevalence of TADs with two or more differentially expressed genes that were either all upregulated or all downregulated compared to TADs that have both upregulated genes and downregulated genes. pvalue was calculated using the two-sided exact binomial test.
D, Percentage of 4kb bins involved in at least one differential interaction only during MYOD-mediated commitment (magenta), only during MYOD-mediated differentiation (green), or at both stages (violet).
E, Number (N) of bins involved in altered chromatin interactions during MYOD-mediated commitment (magenta) or differentiation (green) that were observed or expected to be bound by MYOD. Expected bin number was calculated based on the number of bins bound by MYOD genome-wide. Chi-squared test was used for statistical analysis.
F, Percentage of DI bins bound by MYOD (red), DI bins directly interacting with MYOD-bound bins (orange), DI bins indirectly interacting (dark yellow), others (yellow) – see figure S3A for the schematic representation of this classification
G, Left: number (N) of differential interactions, including all differential interactions (All) Promoters-all (indicating interactions between promoters and any other genomic region), Enhancers-all (indicating interactions between enhancers and any other genomic region), Promoter-enhancers and CTCF-bound regions during MYOD-mediated commitment (magenta) or differentiation (green). **Right:** percentage of the differential interactions described on the left that were bound by MYOD during MYOD-mediated commitment (magenta) or differentiation (green) (right). Dashed lines represent the percentage of all differential interactions that are bound by MYOD during MYOD-mediated commitment (red) or differentiation (blue).

Dall'Agnese et al., Figure 3









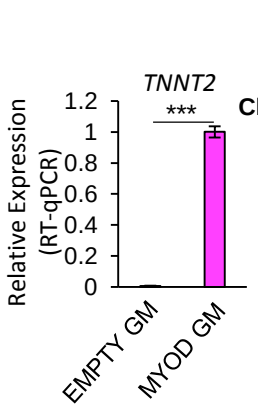


Figure 3: Characterization of MYOD-altered *cis*-regulatory interactions

A, Number (N) of MYOD-bound bins involved in altered interactions between promoters and other genomic elements during MYOD-mediated commitment (magenta) or differentiation (green). The expected number of MYOD-bound bins was calculated based on MYOD-binding genome-wide (see methods).
B, Number (N) of MYOD-bound bins involved in altered enhancer-promoter interactions during MYOD-mediated commitment (magenta) or differentiation (green). The expected number of MYOD-bound bins was calculated based on MYOD-binding to enhancer or promoter (see methods).
C,D, Percentage (%) of MYOD-bound or unbound (dashed) differential interactions between promoters of DE genes and (C) other genomic elements or (D) enhancers.
E, Heatmap representing biological functions activated (orange) or inhibited (blue) based on the DE genes whose promoter was involved in MYOD-bound differential interactions during MYOD-mediated commitment or differentiation. Analysis performed using IPA.
F, DE genes whose promoter was involved in MYOD-bound differential interactions with enhancers during MYOD-mediated commitment or differentiation. Analysis performed using IPA.
G, Normalized contact heatmap at *TNNT2* locus in EMPTY GM (top left) or MYOD GM (bottom right). The region in blue box corresponds to *TNNT2* enhancer-promoter interaction. Enlargement of the interaction of interest in the corners.
H, From top to bottom: Magenta bars represent bins whose interaction increased during MYOD-mediated commitment determined by Hi-C. UCSC snapshots of: H3K27ac ChIP-seq in IMR90 (blue), hMB (violet), P indicates the TNNT2 promoter, E represents an enhancer, MYOD ChIP-seq in MYOD GM (magenta), RefSeq genes, black bars represent regions with increased H3K27ac levels in hMBs compared to IMR90. Close up representation of the enhancer region in the dashed blue box H3K27ac ChIP-seq in hMB (violet), and DpnII sites. Relative crosslinking frequencies (RCF) by in situ 3C using as view point MYOD peak at *TNNT2* promoter (red eye) (n=3).
I, Relative enrichment of H3K27ac by ChIP-qPCR at *TNNT2* promoter and enhancer, n=3. Data is represented as mean +/- SEM.
J, Relative mRNA expression of *TNNT2* (n=3).
In A-D chi-squared test was used for statistical analysis, *** $p<2.2x10-16$
In H-J data is represented as mean + SEM. T-test was used for statistical analysis, * $p<0.05$, ** $p<0.01$.

Dall'Agnese et al., Figure 4

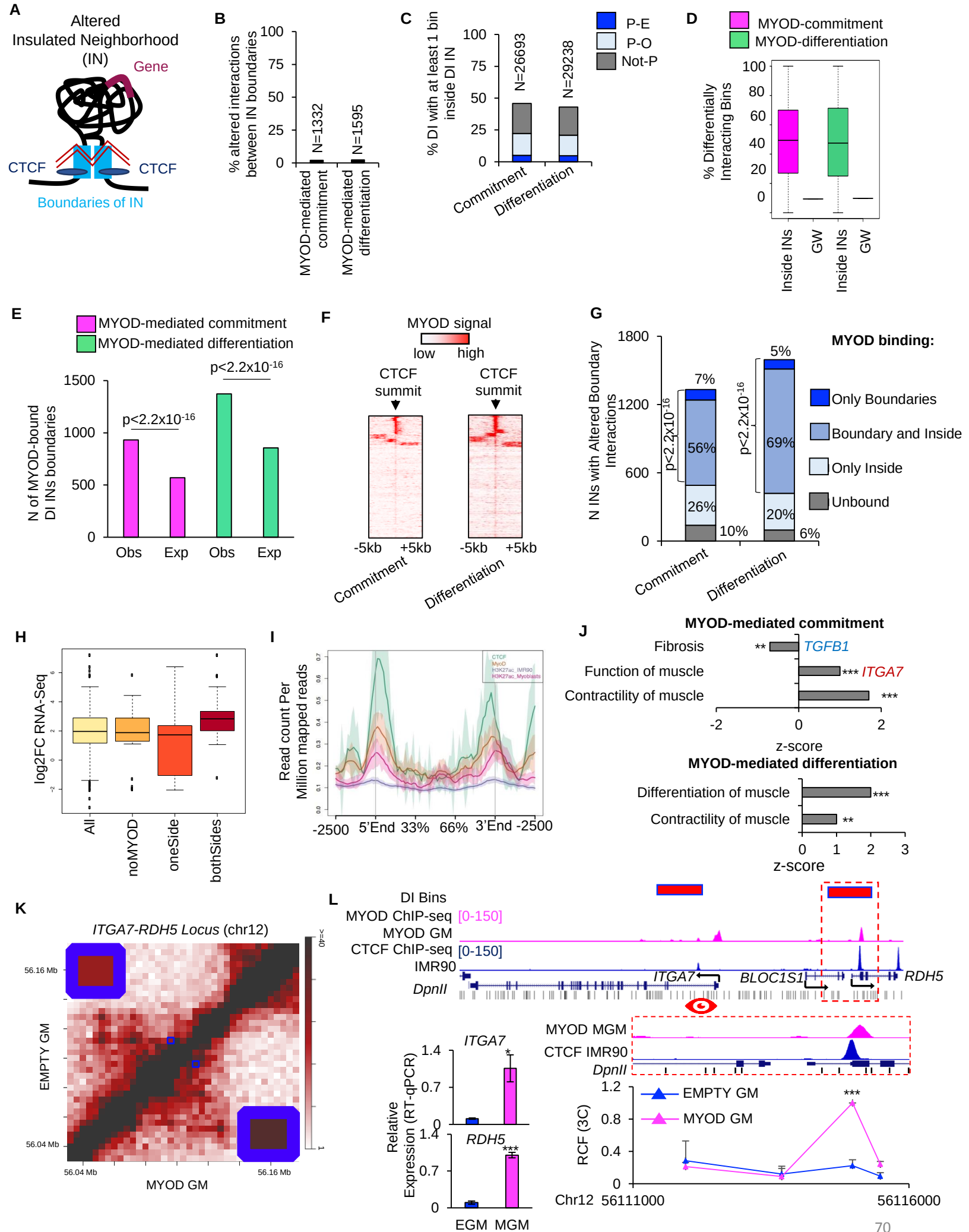


Figure 4: MYOD alters insulated neighborhoods to regulate myogenesis.

A, Graphical representation of altered IN: black line represents the DNA, light-blue boxes represent IN boundaries, blue ovals represent CTCF, violet line represent gene, zig-zagged red lines represent altered interaction.
B, Percentage (%) and number (N) of differential interactions corresponding to altered IN boundary interactions.
C, Percentage (%) and number (N) of DIs with at least one bin that mapped inside altered INs during MYOD-mediated myogenic commitment or differentiation.
D, Percentage of DI bins genome-wide (GW) and distribution of percentages of DI bins inside altered INs.
E, Number (N) of IN boundaries which differentially interacted during myogenic commitment (magenta) or differentiation (green) that were observed (Obs) or expected (Exp) to be bound by MYOD. Expected number of MYOD-bound IN boundaries was calculated based on MYOD binding at bins containing CTCF genome-wide (see Methods). For statistical analysis Chi-squared test was used.
F, MYOD ChIP-seq signal over CTCF-summit +/-5kb at IN boundaries which differentially interacted during MYOD-mediated myogenic commitment or differentiation.
G, MYOD binding distribution at altered INs. For statistical analysis Chi-squared test was used.
H, Boxplots of the gene expression changes EMPTY GM vs MYOD GM of DE genes ($p<0.01$) in DI INs with strengthen interaction between boundaries (All); among these, DI INs not bound by MYOD at the boundaries (noMYOD), at only one boundary (oneSide), and at both boundaries (bothSides).
I, NGSplot of CTCF, MYOD and H3K27ac signal ChIP-seq from IMR90 (CTCF, H3K27ac EMPTY GM) and myoblast (MYOD, H3K27ac_myoblast). 167 regions
J, IPA-based GO analysis of the DE genes within MYOD-bound altered INs.
K, Normalized contact heatmap for EMPTY GM (top left) and MYOD GM (bottom right) at *ITGA7-RDH5* locus. Interaction under investigation is highlighted by blue boxes. Magnification of the blue boxes is shown in the corners.
L, From top to bottom: Magenta bars represent bins whose interaction increased during MYOD-mediated commitment. UCSC snapshots of: MYOD ChIP-seq in MYOD GM (magenta) and CTCF in IMR90 (blue), RefSeq genes from UCSC browser, DpnII sites (black). Close up representation of the region in the dashed red box. Relative crosslinking frequencies (RCF) by in situ 3C using as view point MYOD-CTCF peak at ITGA7 promoter (red eye) (n=3). 3C data is represented as mean + SEM. Relative mRNA expression of *ITGA7* and *RDH5* (n=3). Data is represented as mean +/- SEM. T-test was used for statistical analysis, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

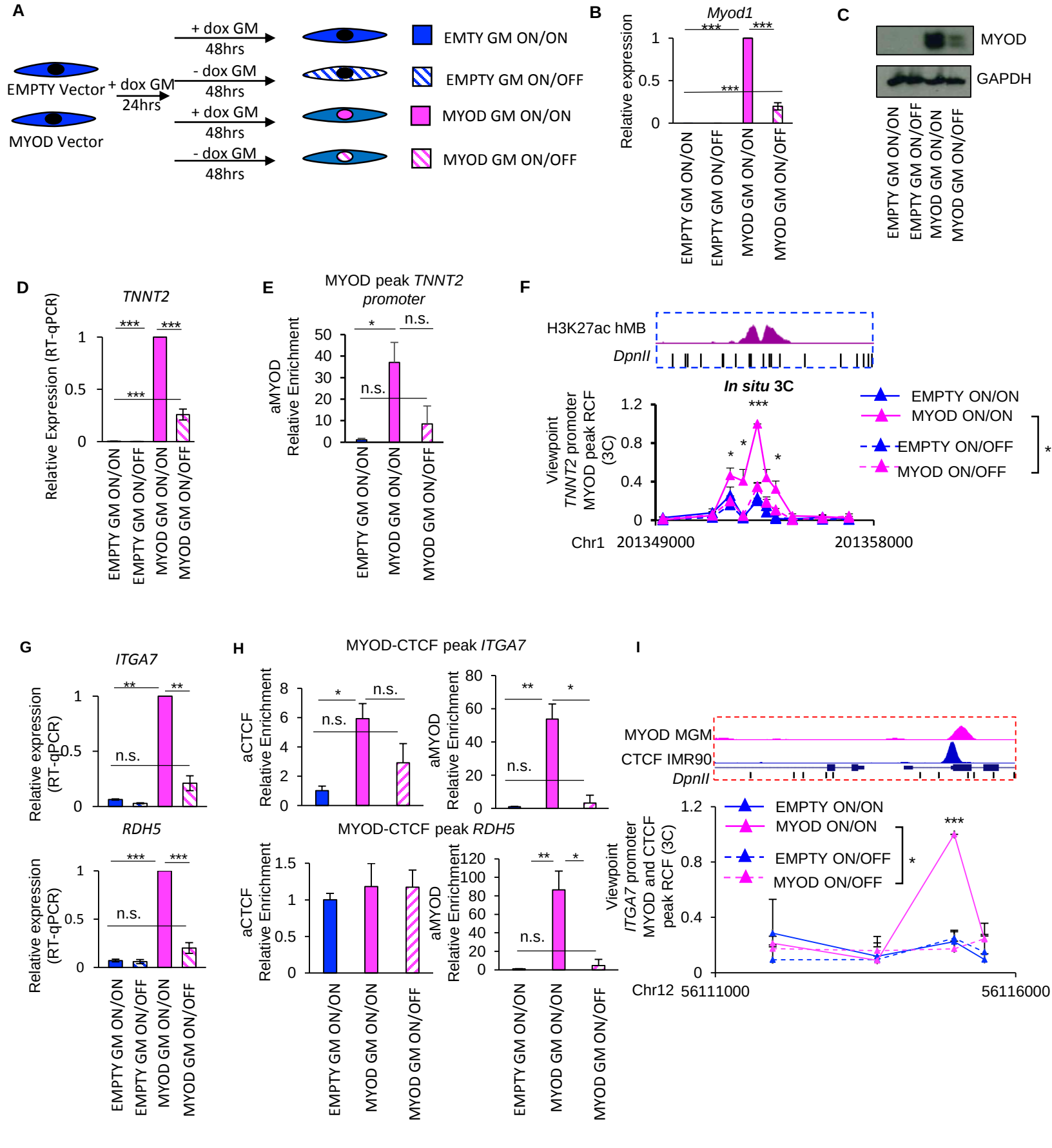

A
EMPTY Vector
MYOD Vector
+ dox GM 24hrs
+ dox GM 48hrs
- dox GM 48hrs
+ dox GM 48hrs
- dox GM 48hrs
EMTY GM ON/ON
EMPTY GM ON/OFF
MYOD GM ON/ON
MYOD GM ON/OFF
B
Myod1
Relative expression
EMPTY GM ON/ON
EMPTY GM ON/OFF
MYOD GM ON/ON
MYOD GM ON/OFF
C
MYOD
GAPDH
D
TNNT2
Relative Expression (RT-qPCR)
E
MYOD peak TNNT2 promoter
aMYOD Relative Enrichment
n.s.
F
H3K27ac hMB
DpnII
In situ 3C
Viewpoint TNNT2 promoter MYOD peak RCF (3C)
EMPTY ON/ON
MYOD ON/ON
EMPTY ON/OFF
MYOD ON/OFF
Chr1 201349000 201358000
G
ITGA7
RDH5
Relative expression (RT-qPCR)
H
MYOD-CTCF peak ITGA7
MYOD-CTCF peak RDH5
aCTCF Relative Enrichment
aMYOD Relative Enrichment
I
MYOD MGM
CTCF IMR90
DpnII
Viewpoint ITGA7 promoter MYOD and CTCF peak RCF (3C)
Chr12 56111000 56116000

Figure 5: MYOD expression is necessary for the maintenance of MYOD-regulated chromatin interactions

A, Scheme of the experimental approach used for all experiment in Fig. 5, EMPTY or MYOD IMR90 were exposed to doxycycline for 24h in GM followed by additional 48h of with/out doxycycline (ON/ON, ON/OFF).
B, Relative mRNA expression of *Myod1* compared to EMPTY ON/ON (n=3). Data is represented as mean +/- SEM.
C, Immunoblot analysis of the whole cell lysate. GAPDH is used as loading control
D, Relative mRNA expression of *TNNT2* compared to EMPTY ON/ON (n=3). Data is represented as mean +/- SEM
E, ChIP-qPCR for MYOD at *TNNT2* promoter relative to EMPTY ON/ON (n=3).
F, Relative crosslink frequency (RCF) values between MYOD peak at *TNNT2* promoter (view point – red eye – see Fig. 3H) and the enhancer. Data is represented as mean + SEM (n=3).
G, Relative mRNA expression of ITGA7 and RDH5 compared to EMPTY ON/ON (n=3). Data is represented as mean +/- SEM
H, ChIPqPCR for CTCF and MYOD at regulatory elements in the locus relative to EMPTY ON/ON (n=3).
I, Relative crosslink frequency (RCF) values between CTCF MYOD peak at ITGA7 promoter (view point – red eye – see Fig. 4L) and the CTCF MYOD peak in RDH5. Data is represented as mean + SEM (n=3).
T-test was used for statistical analysis, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Dall'Agnese et al., Figure 6

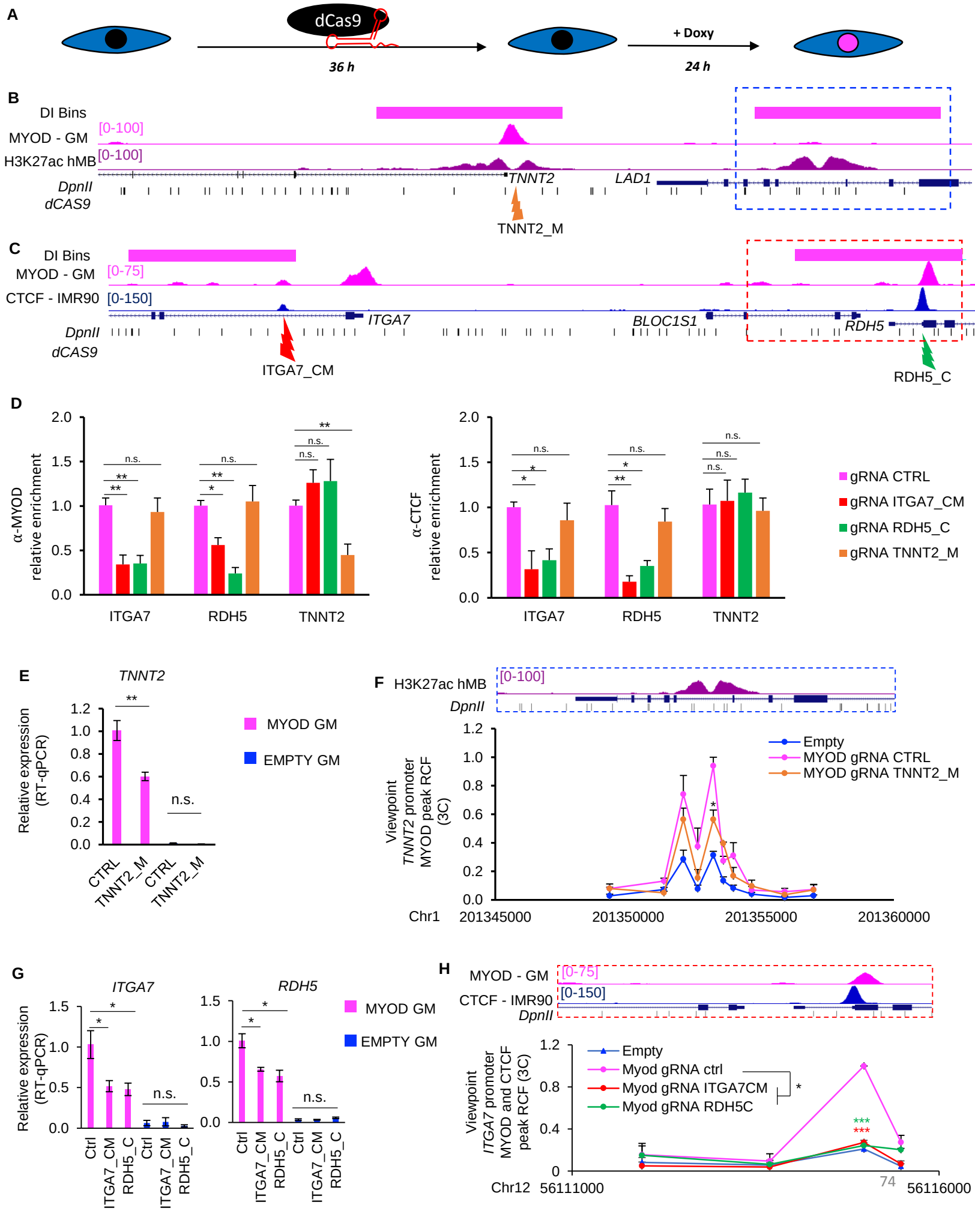


Figure 6: Direct MYOD binding is required for MYOD-directed changes in the 3D chromatin structure

A, Scheme of the experimental approach used for all experiments in Fig. 6, EMPTY or MYOD IMR90 were transfected with plasmid encoding dCAS9 and specific gRNAs 36hours, then cells were treated with doxycycline for 24h in GM.
B, From top to bottom: *TNNT2* locus - magenta bars represent bins whose interaction increased during MYOD-mediated commitment, UCSC genome browser snapshots of MYOD ChIP-seq in MYOD GM (magenta), H3K27ac ChIP-seq in hMB (violet), RefSeq genes, DpnII sites (black). Orange arrow indicates the region targeted by the gRNAs, which is MYOD and CTCF peak at *TNNT2* promoter (TNNT2_M).
C, From top to bottom: *ITGA7-RDH5* locus - magenta bars represent bins with increased interaction between during MYOD-mediated commitment, UCSC genome browser snapshots of MYOD ChIP-seq in MYOD GM (magenta), CTCF ChIP-seq in IMR90 (blue), RefSeq genes, DpnII sites (black). Red arrow indicates gRNAs targeting the MYOD-CTCF in the *ITGA7* promoter (ITGA7_CM), green arrow indicates gRNAs targeting the CTCF in the *RDH5* promoter (RDH5_C).
D, ChIP-qPCR for MYOD (left) or CTCF (right) at regulatory elements in *ITGA7, RDH5* or *TNNT2* loci. Data is represented as relative enrichment over MYOD expressing IMR90 transfected with CTRL gRNAs (n=3) Data is represented as mean + SEM.
E, Relative mRNA expression of *TNNT2*. Data is represented as mean +/- SEM
F, Close up representation of the enhancer region in the dashed blue box H3K27ac ChIP-seq in hMB (violet), and DpnII sites. Relative crosslinking frequencies (RCF) by in situ 3C using as view point MYOD peak at TNNT2 promoter (red eye, see Fig 3H) (n=3). 3C data is represented as mean + SEM.
G, Relative mRNA expression of *ITGA7* and *RDH5*. Data is represented as mean +/- SEM
H, Close up representation of the enhancer region in the dashed red box MYOD ChIP-seq in MYOD GM (magenta), CTCF ChIP-seq in IMR90 (blue) and DpnII sites. Relative crosslinking frequencies (RCF) by in situ 3C using as view point CTCF-MYOD peak at ITGA7 promoter (red eye, see Fig 4L) (n=3). 3C data is represented as mean + SEM.
T-test was used for statistical analysis, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

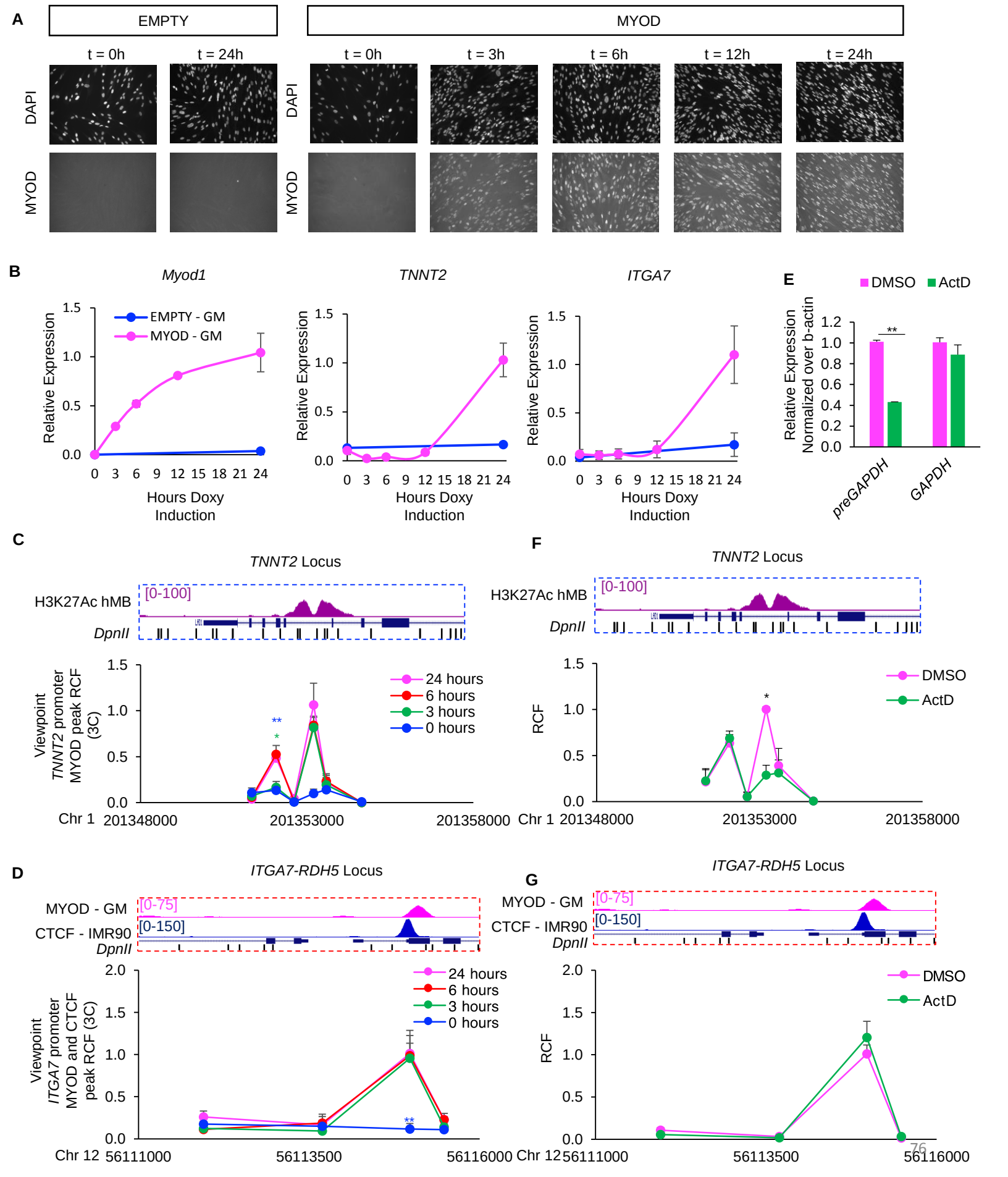

A
EMPTY
MYOD
t = 0h
t = 24h
t = 0h
t = 3h
t = 6h
t = 12h
t = 24h
DAPI
MYOD
B
Myod1
TNNT2
ITGA7
EMPTY - GM
MYOD - GM
Relative Expression
Hours Doxy Induction
E
DMSO
ActD
Relative Expression Normalized over b-actin
preGAPDH
GAPDH
C
TNNT2 Locus
H3K27Ac hMB
DpnII
[0-100]
Viewpoint TNNT2 promoter MYOD peak RCF (3C)
24 hours
6 hours
3 hours
0 hours
Chr 1
201348000
201353000
201358000
F
TNNT2 Locus
RCF
D
ITGA7-RDH5 Locus
MYOD - GM
CTCF - IMR90
[0-75]
[0-150]
Viewpoint ITGA7 promoter MYOD and CTCF peak RCF (3C)
Chr 12
56111000
56113500
56116000
G
ITGA7-RDH5 Locus

Figure 7: MYOD loop formation and transcription

A, Time course analysis of *Myod1* expression in doxycycline-treated IMR90 cells. Representative immunofluorescence images of IMR90 cells stained for MYOD. The nuclei were stained with DAPI.

B, Relative expression of exogenous *Myod1,* endogenous *TNNT2, ITGA7* (n=3). Data is represented as mean +/- SEM.

C,D, *In situ* 3C analysis of the *TNNT2* (C)*, ITGA7* (D) loci at different time points of MYOD inductions. View point for *TNNT2* locus is MYOD peak at *TNNT2* promoter (Fig. 3H red eye). View point for *ITGA7-RDH5* locus is MYOD and CTCF co-peak at *ITGA7* promoter (Fig. 4L, red eye).

E, Relative expression of pre-*GAPDH* or *GAPDH* mRNA after treatment with 1µg/ml of Actinomycin D (ActD) or DMSO for 30 minutes at 37°C after 6 hours of *Myod1* induction.

F,G, *In situ* 3C analysis of the *TNNT2, ITGA7* loci after treatment with 1µg/ml of Actinomycin D (ActD) or DMSO for 30 minutes at 37°C after 6 hours of MYOD induction.

3C data is represented as mean + SEM. View point for *TNNT2* locus is MYOD peak at *TNNT2* promoter (Fig. 3H red eye). View point for *ITGA7-RDH5* locus is MYOD and CTCF co-peak at *ITGA7* promoter (Fig. 4L, red eye).

T-test was used for statistical analysis, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. p-values have been calculated comparing 0hr vs 24hrs time point (blue), 3hrs vs 24hrs time point (green) in C and D and DMSO in F and G.

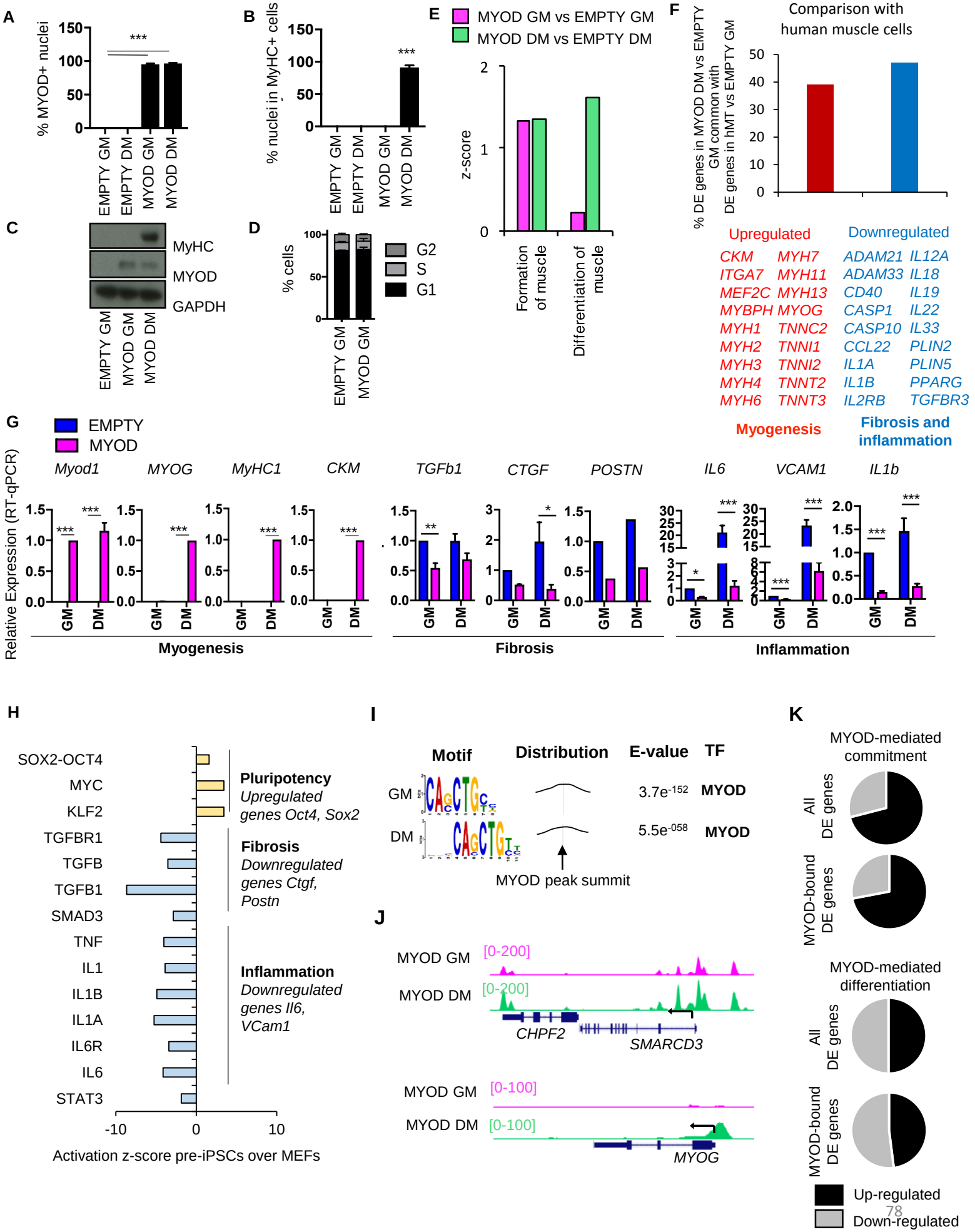

A
% MYOD+ nuclei
EMPTY GM
EMPTY DM
MYOD GM
MYOD DM
B
% nuclei in MyHC+ cells
C
MyHC
MYOD
GAPDH
D
% cells
G2
S
G1
E
MYOD GM vs EMPTY GM
MYOD DM vs EMPTY DM
z-score
Formation of muscle
Differentiation of muscle
F
Comparison with human muscle cells
% DE genes in MYOD DM vs EMPTY GM common with DE genes in hMT vs EMPTY GM
Upregulated
Downregulated
CKM MYH7 ADAM21 IL12A
ITGA7 MYH11 ADAM33 IL18
MEF2C MYH13 CD40 IL19
MYBPH MYOG CASP1 IL22
MYH1 TNNC2 CASP10 IL33
MYH2 TNNI1 CCL22 PLIN2
MYH3 TNNI2 IL1A PLIN5
MYH4 TNNT2 IL1B PPARG
MYH6 TNNT3 IL2RB TGFBR3
Myogenesis
Fibrosis and inflammation
G
EMPTY
MYOD
Relative Expression (RT-qPCR)
Myod1
MYOG
MyHC1
CKM
TGFb1
CTGF
POSTN
IL6
VCAM1
IL1b
GM
DM
Myogenesis
Fibrosis
Inflammation
H
SOX2-OCT4
MYC
KLF2
TGFBR1
TGFB
TGFB1
SMAD3
TNF
IL1
IL1B
IL1A
IL6R
IL6
STAT3
Pluripotency
Upregulated genes Oct4, Sox2
Fibrosis
Downregulated genes Ctgf, Postn
Inflammation
Downregulated genes Il6, VCam1
Activation z-score pre-iPSCs over MEFs
I
Motif
Distribution
E-value
TF
GM
DM
3.7e-152
MYOD
5.5e-058
MYOD
MYOD peak summit
J
MYOD GM
MYOD DM
[0-200]
[0-100]
CHPF2
SMARCD3
MYOG
K
MYOD-mediated commitment
All DE genes
MYOD-bound DE genes
MYOD-mediated differentiation
Up-regulated
Down-regulated

Figure S1. Related to Figure 1. Master transcription factor-driven transcription reprogramming and chromatin binding

A, Percentage of MYOD-positive nuclei. (n=3). Data are represented as mean +/- standard error of the mean (SEM). Two-way ANOVA was used for statistical analysis, corrected for multiple testing (Tukey), *** $p<0.001$.

B, Percentage of nuclei in MyHC positive cells. (n=3). Data are represented as mean +/- standard error of the mean (SEM). Two-way ANOVA was used for statistical analysis, corrected for multiple testing (Tukey), *** $p<0.001$.

C, Immunoblot analysis of the whole cell lysate (left). Data is represented as mean +/- SEM.

D, Cell cycle analysis (n=3). Data is represented as mean +/- SEM.

E, Selected biological functions enriched by MYOD comparing MYOD GM vs EMPTY GM and MYOD DM vs EMPTY DM (IPA).

F, Percentage of up- or down-regulated genes in MYOD DM vs EMPTY GM that are common to up- or down-regulated genes in hMT vs EMPTY GM.

G, Relative expression of muscle, fibrotic and inflammatory genes (n=2-4) by RT-qPCR. Data is represented as mean +/- SEM. Two-way ANOVA was used for statistical analysis, corrected for multiple testing (Tukey), * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

H, Upstream prediction analysis by IPA based on the gene expression profile of MEFs and MEF-derived iPSCs. All categories have $p< 1x10^{-7}$.

I, E-box motif enrichment (left) and distribution (right) at +/-50bp from MYOD peak summit, in GM (top) and DM (bottom).

J, UCSC snapshot of MYOD ChIP-seq in MYOD GM and in MYOD DM on *SMARCD3* (top) or *MYOG* (bottom) loci.

K, Percentage of up- (black) or down- (grey) regulated genes during commitment and differentiation (All DE genes). Percentage of MYOD-bound up- or down-regulated genes during commitment and differentiation (MYOD-bound DE genes).

Dall'Agnese et al., Figure S2

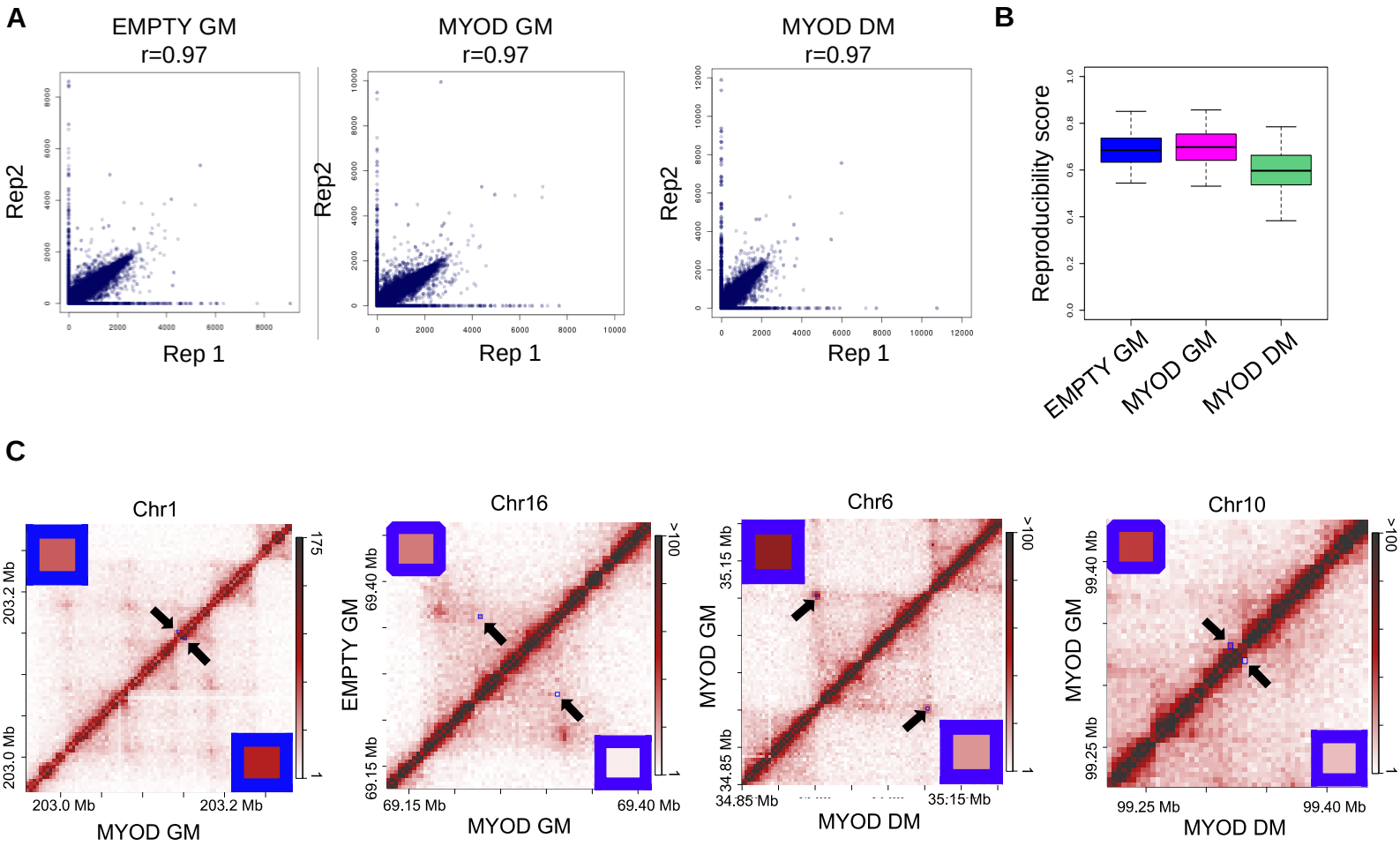
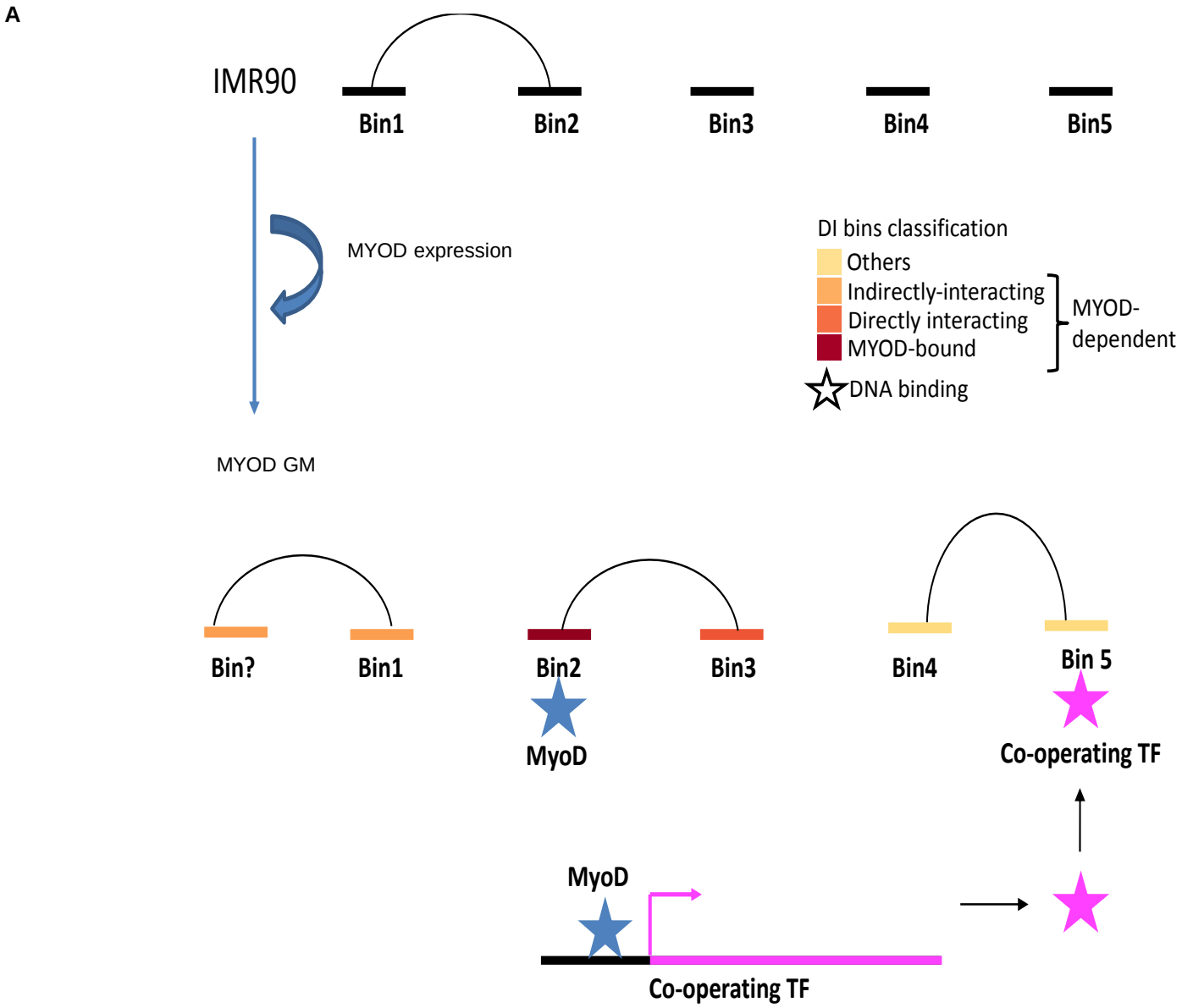
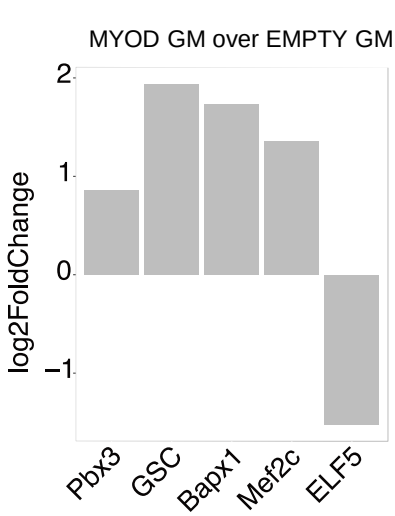


Figure S2. Related to Figure 2. Hi-C Reproducibility and examples of differential interactions

A, Scatter plot representing reproducibility of two biological replicates as in Dixon et al, 2012. **B.** Reproducibility score of intra-chromosomal Hi-C contact matrices among biological replicates calculated using HiC-spector. Each chromosome is represented as a single circle. **C,** Normalized Hi-C contact heatmaps for EMPTY GM (top left) vs MYOD GM (bottom right), and for MYOD GM (top left) vs MYOD DM (bottom right). Interaction under investigation is highlighted by blue boxes. Magnification of the blue boxes is shown on the corners of each heatmap.

Dall'Agnese et al., Figure S3





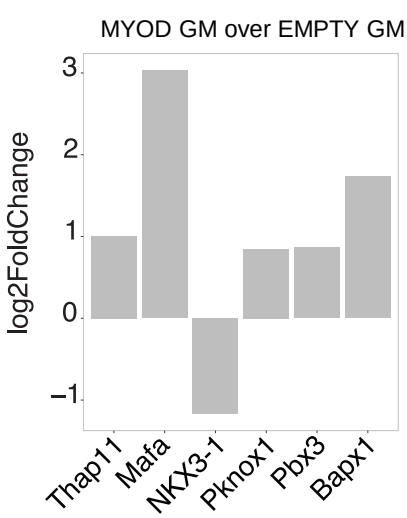


Figure S3. Related to Figure 2. MYOD-domino effect of differential interactions

A, Schematic representation of MYOD-domino effect on differential chromatin interactions.
B, Differential gene expression by RNA-seq of TFs whose DNA binding motif is present around the TSS located inside directly interacting bins (left). The directly interacting bins refer to Fig 2F EMPTY GM vs MYOD GM.
C, Differential gene expression by RNA-seq of TFs whose DNA binding motif is present around the TSS contained within indirectly-interacting bins and other bins (right). indirectly-interacting bins and other bins refer to Fig 2F EMPTY GM vs MYOD GM.

Dall'Agnese et al., Fig. S4

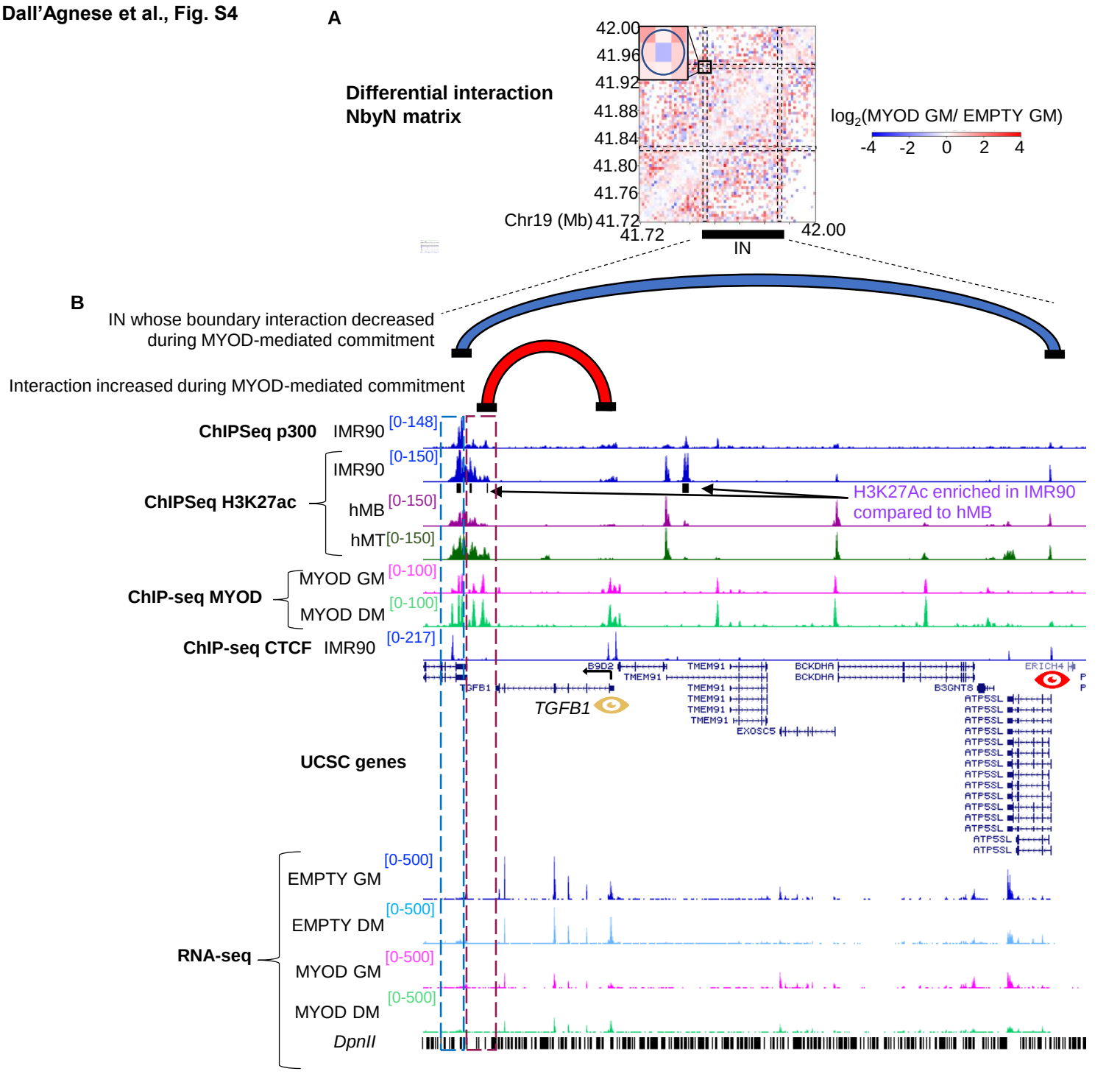


Figure S4. Related to Figure 4. 3D Regulation of *TGF-β1* Expression.

A, Differential interaction NbyN map of *TGF*-β1 locus and location of altered IN during MYOD-mediated commitment (black bar).

B, From top to bottom: blue arch representing decreased interaction between two IN boundaries (CTCF-bound bins) during MYOD-mediated commitment, red arch representing an increased interaction during MYOD-mediated commitment, UCSC snapshot of p300 ChIP-seq in IMR90 (blue), H3K27ac ChIP-seq in IMR90 (blue), H3K27ac ChIP-seq peaks enriched in IMR90 as compared to hMB and hMT (black bars), H3K27Ac ChIP-Seq in hMB, H3K27Ac ChIP-Seq in hMT, MYOD ChIP-seq in MYOD GM (magenta) and MYOD DM (light green), CTCF ChIP-seq in IMR90 (blue), RefSeq genes from UCSC browser, RNASeq of EMPTY GM (blue), EMPTY DM (light blue), MYOD GM (magenta), MYOD DM (light green). DpnII digestion sites (black)

C, Left: Close up representation of the DNA regions delimited by violet dashed box in (B), from top to bottom: H3K27ac ChIP-seq in IMR90 (blue) and hMB (violet), MYOD ChIP-seq in MYOD GM (magenta), DpnII sites, relative crosslinking frequencies (RCF) by in situ 3C using as view point MYOD peak at *TGF*-β1 promoter (gold eye) and measuring RCF with DpnII fragments above. Right: Close up representation of the DNA regions delimited by blue dashed box in (B), from top to bottom: MYOD ChIP-seq in MYOD GM (magenta), CTCF ChIP-seq in IMR90 (blue) and DpnII sites. Relative crosslinking frequencies (RCF) by in situ 3C using as view point CTCF peak at IN boundaries containing *TGF*-β1 (red eye) and measuring RCF with DpnII fragments above (n=3). 3C data is represented as mean + SEM. T-test was used for statistical analysis, * $p<0.05$, ** $p<0.01$.

Dall'Agnese et al., Figure S5

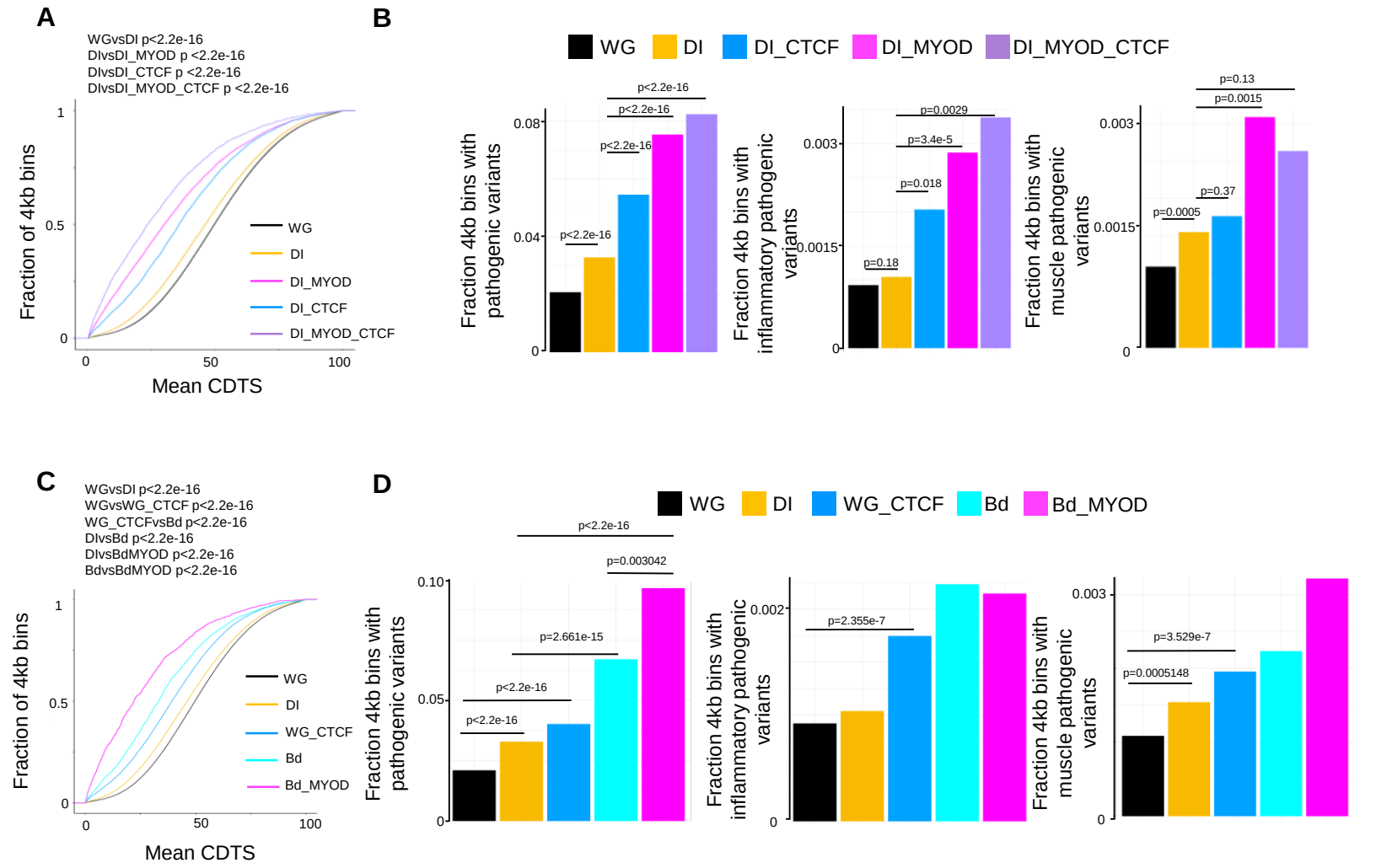


Figure S5. Related to Figure 2 and 4. MYOD-bound differentially interacting elements are highly constraint and enriched in annotated pathogenic variants

A, DNA sequence constraint of all bins genome-wide bins (WG, black), differentially interacting bins (DI, yellow), MYOD-bound bins DI bins (DI_MYOD, magenta), CTCF-bound DI bins (DI_CTCF, light blue) or MYOD and CTCF bound DI bins (DI_MYOD_CTCF, purple). Two-sample Kolmogorov-Smirnov test was used for statistical analysis.
B, Fraction of genome-wide bins (WG, black), differentially interacting bins (DI, yellow), MYOD-bound bins DI bins (DI_MYOD, magenta), CTCF-bound DI bins (DI_CTCF, light blue) or MYOD and CTCF bound DI bins (DI_MYOD_CTCF, purple) harboring variants associated to all diseases (left), inflammatory diseases (middle) or muscle diseases (right). One sided Fisher's Exact Test was used for all statistical analysis in the figure.
C, DNA sequence constraint of all bins genome-wide (WG), differentially interacting bins (DI), CTCF-bound bins (WG_CTCF), differentially interacting bins co-bound by CTCF (Bd), and differentially interacting bins co-bound by CTCF and bound by MYOD (Bd_MYOD). Two-sample Kolmogorov-Smirnov test was used for statistical analysis.
D, Fraction of genome-wide (WG), differentially interacting bins (DI), CTCF-bound bins (WG_CTCF), differentially interacting bins co-bound by CTCF (Bd), and differentially interacting bins co-bound by CTCF and bound by MYOD (Bd_MYOD) harboring annotated pathogenic variants associated to all annotated diseases (left), inflammatory (middle) or muscle diseases (right).One sided Fisher's Exact Test was used for all statistical analysis in the figure.

Dall'Agnese et al., Figure S6

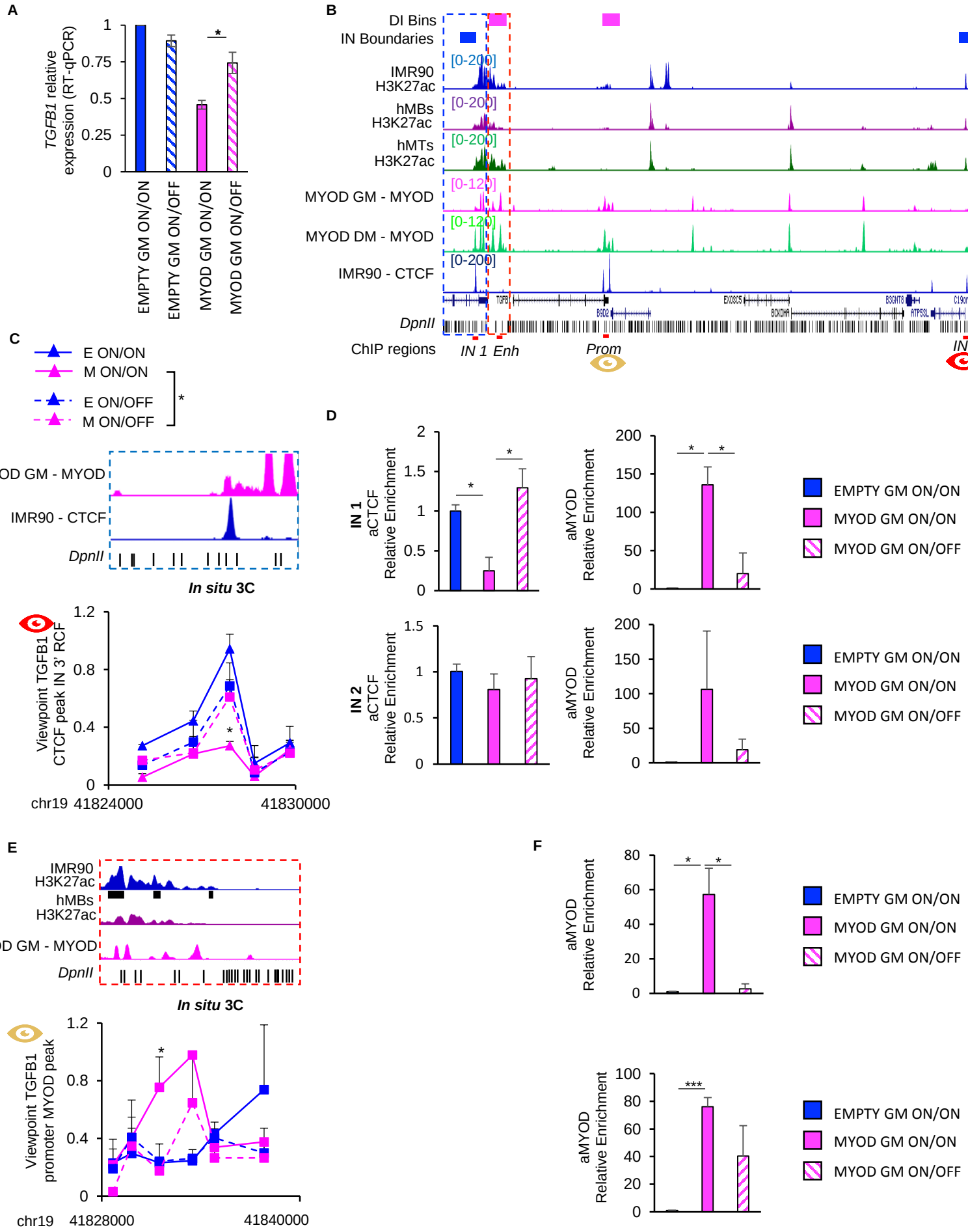


Figure S6. Related to Figure 5. Requirement of *Myod1* expression for maintaining myogenic 3D interaction

A, Relative mRNA expression of *TGF-β1* compared to EMPTY ON/ON (n=3). Data is represented as mean +/- SEM.
B, From top to bottom: magenta boxes representing bins with an increased interaction during MYOD-mediated commitment; blue boxes representing bins with decreased interaction between two IN boundaries (CTCF-bound bins) during MYOD-mediated commitment; UCSC snapshot of H3K27ac ChIP-seq in IMR90 (blue), in hMBs (purple) and in hMTs (dark green); MYOD ChIP-seq in MYOD GM (magenta) and MYOD DM (light green); CTCF ChIP-seq in IMR90 (blue); RefSeq genes; DpnII digestion sites (black); regions analyzed by ChIP-qPCR; eyes are viewpoints for 3C experiments.
C, From top to bottom: UCSC snapshot of the same region that in panel (B) within dashed blue box, MYOD ChIP-seq in MYOD GM; CTCF ChIP-seq in IMR90; DpnII sites; RCF values between CTCF peak at 5' the region of an altered IN containing *TGF*-β1 (view point – red eye – see panel B) and the CTCF peak at 3' region after 1 day of doxycycline treatment and 2 days with/out doxycycline treatment (ON/ON, ON/OFF). Data is represented as mean + SEM (n=3).
D, ChIP-qPCR analysis for CTCF, MYOD enrichment on the IN1 region (top panels) and on the IN2 region (bottom panels). Data are shown as relative enrichment to EMPTY ON/OFF, as mean +/- SEM (n=2).
E, From top to bottom: UCSC snapshot of the same region that in panel (B) within dashed red box, H3K27ac ChIP-seq in IMR90; H3K27ac ChIP-seq in hMBs; MYOD ChIP-seq in MYOD GM; DpnII sites; RCF values between a DpnII fragment containing MYOD peak in the promoter of *TGF*-β1 (view point – gold eye – see panel B) and an enhancer located at the 3' of the gene after 1 day of doxycycline treatment and 2 days with/out doxycycline treatment (ON/ON, ON/OFF). (Right panel) Data is represented as mean + SEM (n=3).
F, ChIP-qPCR analysis for MYOD enrichment on the *TGF*-β1 promoter region (top panel) and on the Enhancer region (bottom panel). Data are shown as relative enrichment to EMPTY ON/OFF, as mean +/- SEM (n=2).
T-test was used for statistical analysis, * $p<0.05$, ** $p<0.01$.

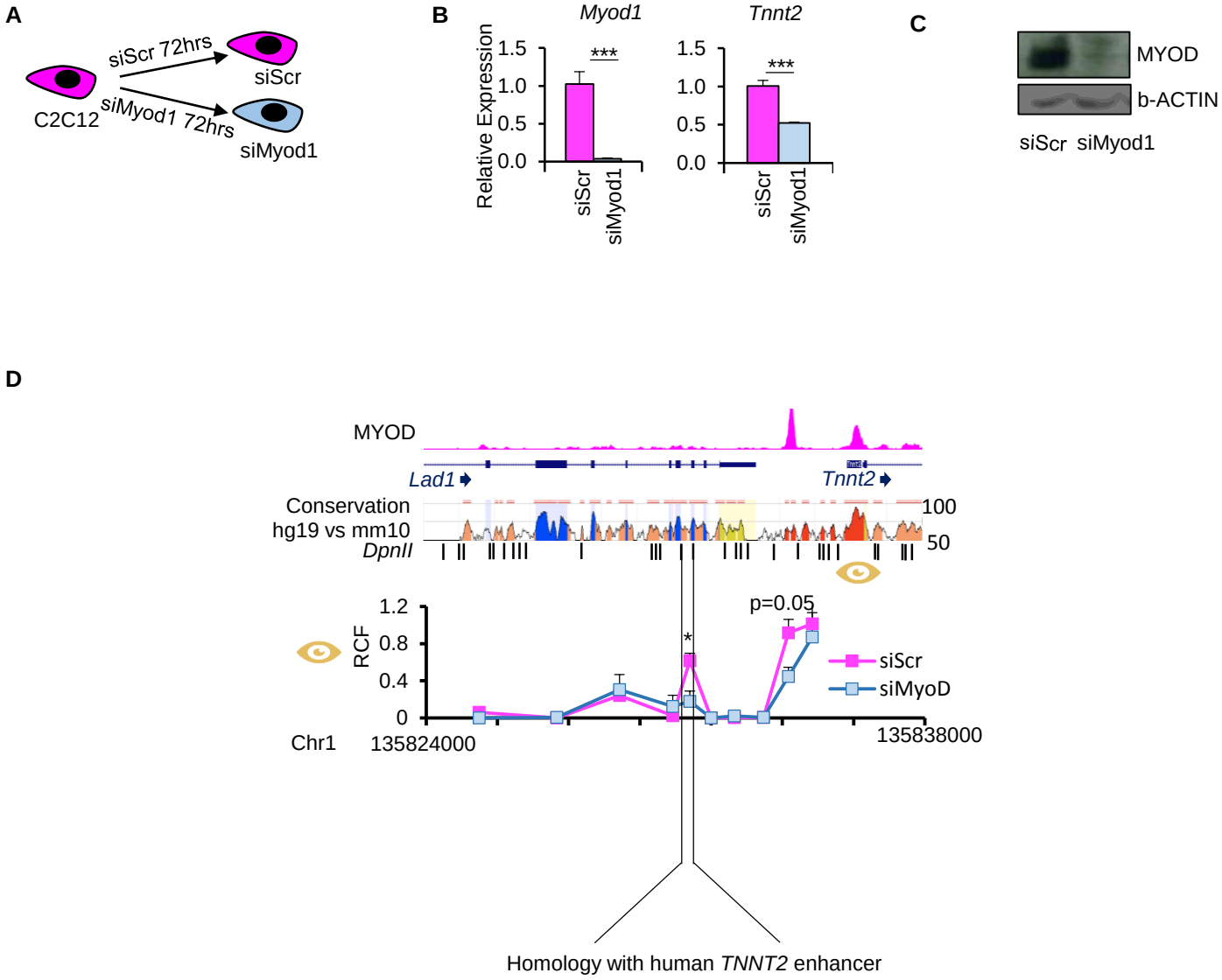


Figure S7. Related to Fig. 5. MYOD mechanisms to regulate gene expression conserved between human and mouse

A, Scheme of the experimental approach in C2C12. C2C12 cells were kept non-confluent and transfected with siRNA scramble (siScr) or siRNA targeting *Myod1* (siMyod1) for 72h.
B, Relative mRNA expression of *Myod1*, *Tnnt2* (n=3). Data is represented as mean +/- SEM. *** p<0.001
C, Immunoblot analysis of the whole cell lysate. b-ACTIN is used as loading control.
D, From top to bottom: UCSC snapshot of MYOD ChIP-seq in C2C12 myoblasts, UCSC genes (arrow indicates transcription direction), conservation levels between human and mouse genomes, DpnII sites, RCF values between MYOD peak at *Tnnt2* promoter (view point - gold eye symbol) and mouse homolog region to the human *TNNT2* enhancer. Data is represented as mean + SEM (n=3).

Motif	TF	p-value
	CTCF	1e-52
	Boris/CTCFL	1e-23
	Myf5	1e-9
	MyoG	1e-8
	MyoD	1e-8
	Atoh1	1e-7
	Tcf12	1e-7
	Fli1	1e-7
	NeuroD1	1e-6
	Ap4	1e-5
	Fra2	1e-5
	Tcf21	1e-4
	TEAD2	1e-4
	Jun-AP1	1e-4
	Olig2	1e-4
	Fosl2	1e-4
	Fra1	1e-4
	Ascl1	1e-4
	Bach2	1e-4
	NeuroG2	1e-3

Table S2. Related to Fig. 4

HOMER motif analysis at the center of MYOD peaks (+/-500bp) that mapped at IN boundaries with strengthen interaction bound by MYOD at both sides during MYOD-mediated commitment

Motif	TF	p-value
	CTCF	1e-37
	Boris/CTCFL	1e-23
	Ap4	1e-4
	KLF6	1e-4
	Elk1	1e-4
	MyoG	1e-3
	Ets1-distal	1e-3
	MyoD	1e-3
	Egr1	1e-3
	KLF14	1e-2
	E2F3	1e-2
	NeuroD1	1e-2
	Sp5	1e-2
	EWS:FLI1-fusion	1e-2
	GABPA	1e-2
	Tcf12	1e-2
	Tcf21	1e-2
	Foxo1	1e-2
	Foxa3	1e-2
	Myf5	1e-2

Table S3. Related to Fig. 4

HOMER motif analysis at the center of MYOD peaks (+/-500bp) that mapped at IN boundaries with weakened interaction bound by MYOD at both sides during MYOD-mediated commitment