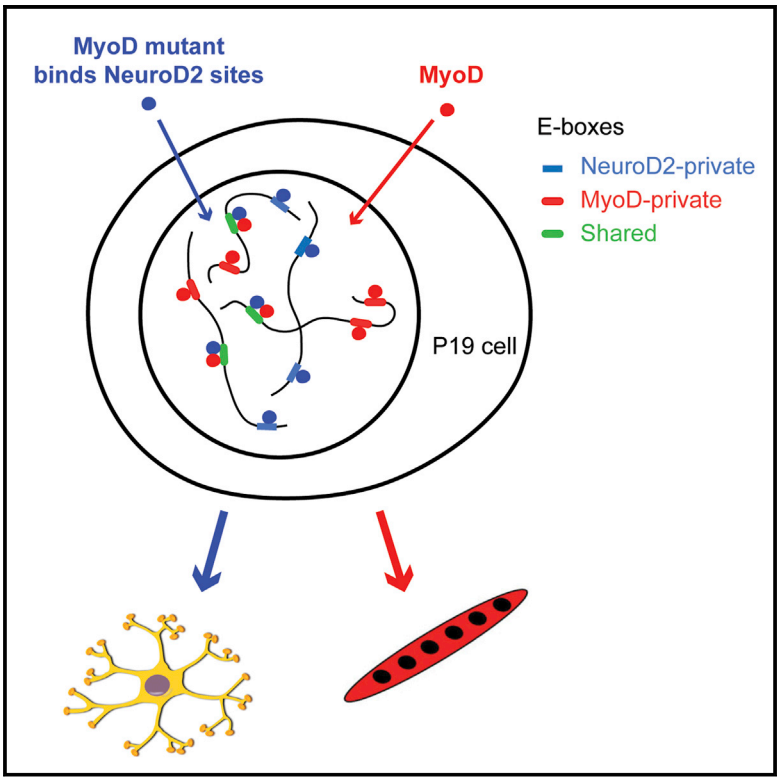
Cell Reports

Report

Conversion of MyoD to a Neurogenic Factor: Binding Site Specificity Determines Lineage

Graphical Abstract



Authors

Abraham P. Fong, Zizhen Yao, ...,
Lisa Maves, Stephen J. Tapscott

Correspondence

apfong@fhcrc.org (A.P.F.),
stapscot@fhcrc.org (S.J.T.)

In Brief

MyoD is a master regulator of muscle differentiation and controls the myogenic gene regulatory network through direct DNA binding. Fong et al. show that MyoD can be converted to a neuronal differentiation factor simply through retargeting of its DNA binding to NeuroD2 sites.

Highlights

- Mutations that redirect MyoD to NeuroD2 binding sites make it a neurogenic factor
- A MyoD chimera with a NeuroD2 bHLH domain redirects MyoD to NeuroD2 E-boxes
- Mutation of the MyoD PBX interacting residues is necessary to lose myogenic activity
- The lineage determined by a master regulatory factor can be rationally reprogrammed

Accession Numbers

GSE64627
GSE64626

Fong et al., 2015, Cell Reports *10*, 1937–1946
March 31, 2015
http://dx.doi.org/10.1016/j.celrep.2015.02.055

CellPress

Cell Reports
Report

Conversion of MyoD to a Neurogenic Factor: Binding Site Specificity Determines Lineage

Abraham P. Fong,[1,3,*] Zizhen Yao,[2] Jun Wen Zhong,[2] Nathan M. Johnson,[5] Gist H. Farr III,[5] Lisa Maves,[4,5] and Stephen J. Tapscott[1,2,6,*]
[1]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
[2]Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
[3]Department of Pediatrics, Division of Hematology-Oncology, University of Washington School of Medicine, Seattle, WA 98105, USA
[4]Department of Pediatrics, Division of Cardiology, University of Washington, Seattle, WA 98105, USA
[5]Center for Developmental Biology and Regenerative Medicine, Seattle Children's Research Institute, Seattle, WA 98101, USA
[6]Department of Neurology, University of Washington School of Medicine, Seattle, WA 98105, USA
*Correspondence: apfong@fhcrc.org (A.P.F.), stapscot@fhcrc.org (S.J.T.)
http://dx.doi.org/10.1016/j.celrep.2015.02.055

SUMMARY

MyoD and NeuroD2, master regulators of myogenesis and neurogenesis, bind to a "shared" E-box sequence (CAGCTG) and a "private" sequence (CAGGTG or CAGATG, respectively). To determine whether private-site recognition is sufficient to confer lineage specification, we generated a MyoD mutant with the DNA-binding specificity of NeuroD2. This chimeric mutant gained binding to NeuroD2 private sites but maintained binding to a subset of MyoD-specific sites, activating part of both the muscle and neuronal programs. Sequence analysis revealed an enrichment for PBX/MEIS motifs at the subset of MyoD-specific sites bound by the chimera, and point mutations that prevent MyoD interaction with PBX/MEIS converted the chimera to a pure neurogenic factor. Therefore, redirecting MyoD binding from MyoD private sites to NeuroD2 private sites, despite preserved binding to the MyoD/NeuroD2 shared sites, is sufficient to change MyoD from a master regulator of myogenesis to a master regulator of neurogenesis.

INTRODUCTION

Conversion of fibroblasts, and many other cell types, to skeletal muscle cells through expression of MyoD was the first demonstration of programmed transdifferentiation, with a single transcription factor able to specify a cell lineage (Davis et al., 1987; Fong and Tapscott, 2013; Tapscott et al., 1988). Subsequently, other transcription factors have been identified with similar lineage determining abilities, such as the ability of the NeuroD transcription factors to transdifferentiate cells to neurons (Farah et al., 2000; Lee et al., 1995; Vierbuchen et al., 2010). Both MyoD and the NeuroD factors have a basic-helix-loop-helix domain, form dimers with the E-protein family, and bind a core CANNTG motif (Berkes and Tapscott, 2005; Chae et al., 2004); yet, they activate distinct programs of differentiation with only about 10% of regulated genes in common (Fong et al., 2012). Besides the shared bHLH motif and the core DNA binding motif, MyoD and NeuroD2 proteins have little similarity, suggesting that their lineage specificity might be encoded in divergent regions of the proteins. As a possible alternative to this model, our recent chromatin immunoprecipitation sequencing (ChIP-seq) study showed that while both factors bind to a shared E-box sequence (CAGCTG), NeuroD2 and MyoD also had a strong binding preference for a distinct, or "private," E-box binding sequences (CAGATG and CAGGTG, respectively), and that these private binding sites were more strongly associated with lineage-specific gene transcription (Fong et al., 2012). To assess the relative contribution of binding-site specificity versus other regions of the proteins, we redirected the genome-wide binding of MyoD to the sites bound by NeuroD2 to determine whether binding site affinity was sufficient to convert MyoD from a myogenic-lineage determination factor into a neurogenic-lineage determination factor.

RESULTS

The DNA binding specificity of MyoD and NeuroD2 is encoded by the conserved bHLH domain (Longo et al., 2008; Ma et al., 1994) that is 39% identical between the two factors. We therefore generated a MyoD chimeric protein substituted with the NeuroD2 bHLH domain (Figure 1A). The swapped region contained all of the DNA binding and dimerization surfaces as defined by crystal structures for MyoD (Ma et al., 1994) and NeuroD1 (Longo et al., 2008), a NeuroD2 homolog that differs by only two amino acids within the bHLH domain. Similar to NeuroD2, the chimeric MD(ND2bHLH) protein, henceforth referred to as M(N), showed increased relative activation of a reporter containing a pair of NeuroD2 private (GA) E-boxes and weak activation of both the MyoD private (GG) and the shared (GC) E-box reporters (Figure 1B), whereas MyoD showed stronger activation

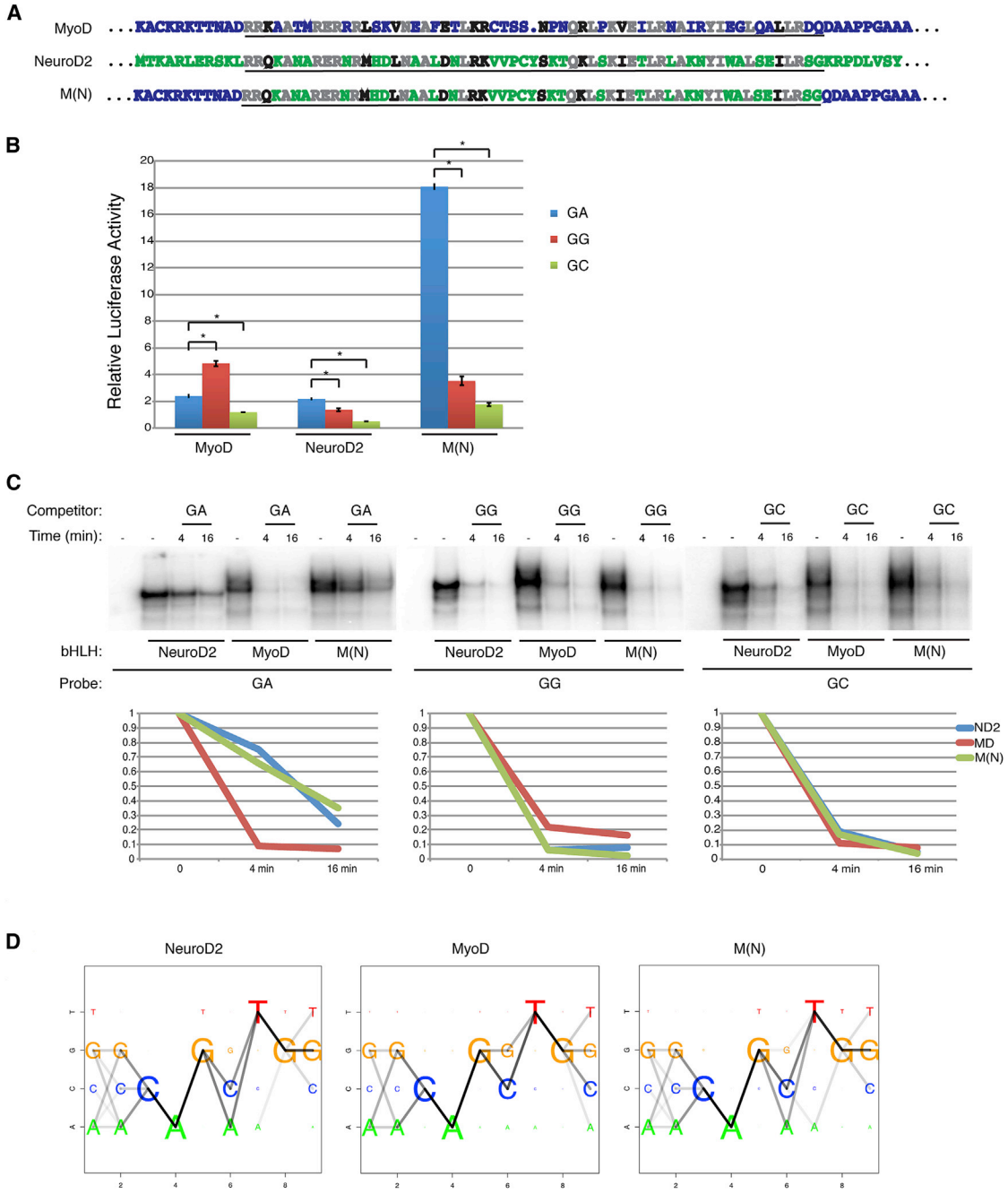


Figure 1. Replacement of the MyoD bHLH Domain with the NeuroD2 bHLH Domain Confers Binding to NeuroD2 Private E-boxes and Loss of Binding to MyoD Private E-boxes

(A) Amino acid sequences of bHLH domains (underlined) are depicted for MyoD, NeuroD2, and M(N) chimera. Residues are colored as conserved (gray), similar (black), MyoD-specific (blue), or NeuroD2-specific (green).

(B) MyoD, NeuroD2, or M(N) chimera were co-transfected with a reporter construct containing a pair of NeuroD2 private E-boxes (GA), MyoD private E-boxes (GG), or shared E-boxes (GC). M(N) activates the GA reporter, but only weakly activates the GG and GC reporters, similar to wild-type NeuroD2. The stronger activity of the M(N) chimera compared to NeuroD2 suggests a stronger activation domain provided by the MyoD portion. Error bars represent 1 SD.

(C) In vitro dissociation assay was performed using in-vitro-translated MyoD, NeuroD2, or M(N) chimera. Radiolabeled probe containing a NeuroD2 private E-box (GA), MyoD private E-box (GG), or a shared (GC) E-box, was competed with excess unlabeled oligo for the indicated amount of time. NeuroD2 and M(N) display a slower off-rate from GA compared to both GG and GC, shown graphically for NeuroD2 (blue), M(N) chimera (green), and MyoD (red).

(D) Binding site motif enrichment logo derived from ChIP-seq in P19 cells transduced with lentivirus expressing NeuroD2, MyoD, or M(N) chimera demonstrates a NeuroD2-like E-box preference for the chimera.

See also Figure S1.

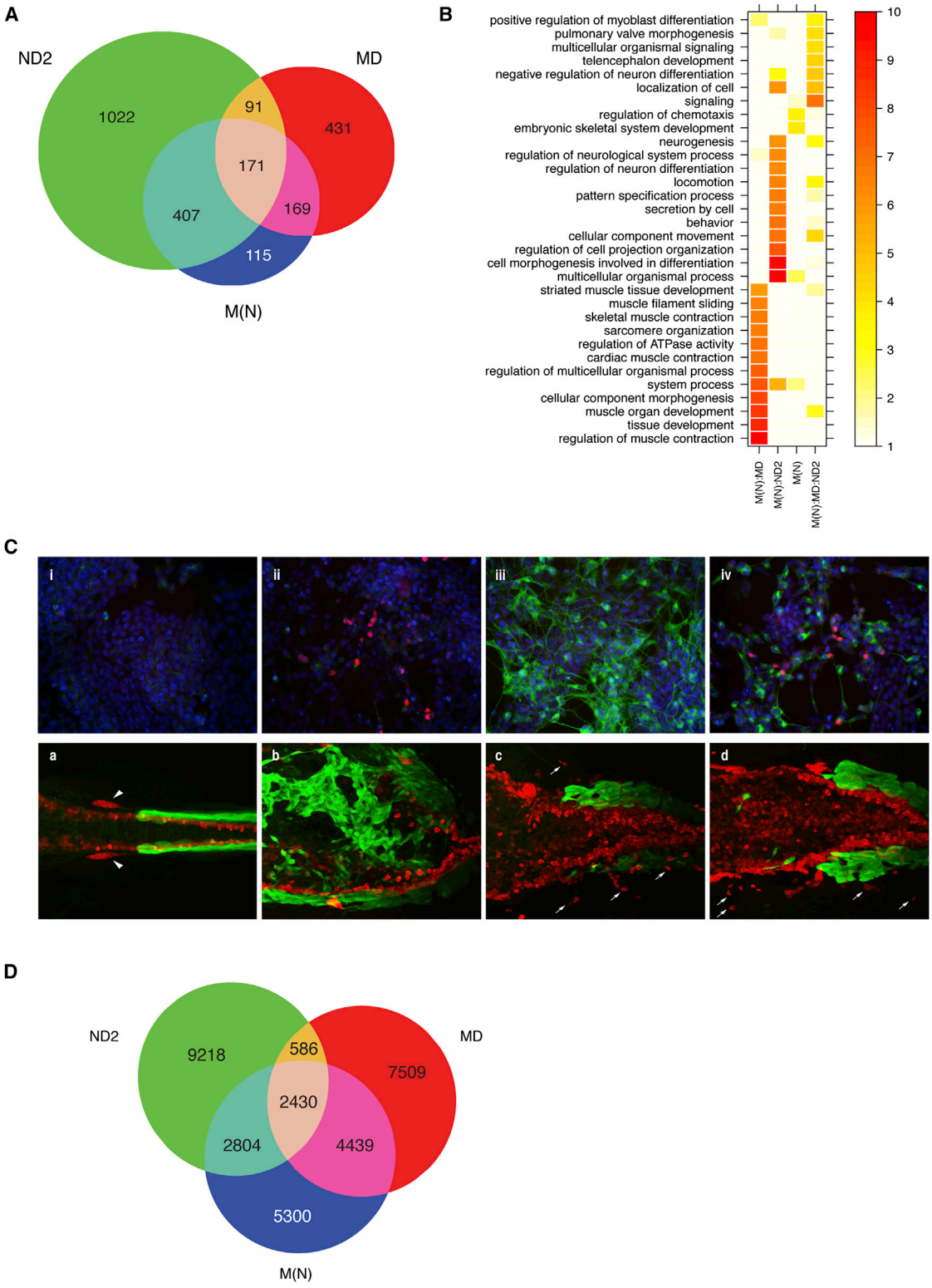


Figure 2. M(N) Chimera Activates Both Muscle and Neuronal Genes in P19 Cells

(A) RNA-seq from P19 cells transduced with MyoD, NeuroD2, or M(N) chimera lentiviruses followed by neuronal differentiation demonstrates upregulation of both MyoD and NeuroD2 target genes by the chimera. Venn diagram displays upregulated genes above log-2 fold cutoff.

(legend continued on next page)

of the reporter with the MyoD private GG E-boxes. For both NeuroD2 and the M(N) chimera, this pattern of activation correlated with a higher binding affinity to the GA E-box as demonstrated by in vitro gel shift dissociation assays (Figure 1C). The half-life for both NeuroD2 and the M(N) chimera bound to a GA E-box was ~11 min, while for MyoD bound to a GA E-box it was <2 min. In contrast, the half-life of NeuroD2, MyoD, and the chimera bound to a GC E-box was only ~2.5 min. Together, these results indicated that the preferential activation of the GA E-box by NeuroD2 correlated with relative binding affinity, and that the bHLH domain swap conferred both the affinity and preferential activation of the NeuroD2 GA E-box to the chimeric M(N) protein.

To determine whether the chimera bound to NeuroD2 E-box motifs in vivo, we performed ChIP-seq in P19 cells transduced with lentiviruses expressing NeuroD2, MyoD, or M(N). At comparable mRNA expression levels (see below), the total number of binding sites was similar for MyoD and NeuroD2, while there was roughly half the number of sites for the chimera (Table S1). A discriminative motif enrichment analysis demonstrated that the E-box sequence bound by the chimera closely matched the NeuroD2 site (Figure 1D). Therefore, the chimera acquired the E-box specificity of NeuroD2 both in vitro and in vivo.

NeuroD2 is a potent inducer of neuronal differentiation in P19 cells (Farah et al., 2000; Fong et al., 2012), while MyoD induces only partial muscle differentiation in this cell type (Fong et al., 2012; Skerjanc et al., 1994). We compared the transcriptional profiles of P19 cells transduced with NeuroD2, MyoD, or M(N) lentiviruses. Based on RNA sequencing (RNA-seq) reads, the amount of steady-state mRNA was similar between the chimera, MyoD, and NeuroD2 (data not shown); however, the amount of the transduced chimeric protein was ~50% lower than the transduced MyoD protein as determined by western blot with a monoclonal antibody to the MyoD C terminus (Figure S1A). The chimera demonstrated a partial activation of both the neurogenic and myogenic programs, with ~87% of genes upregulated by the chimera also upregulated by either NeuroD2 or MyoD (Figure 2A). The chimera induced ~34% of the genes upregulated by NeuroD2 and ~39% of genes upregulated by MyoD and demonstrated strong enrichment for both neuronal and muscle development genes (Figure 2B). Overall, the chimera had a partial activation of both the neuronal and myogenic programs.

Single-cell analysis further demonstrated that the chimera induced both muscle and neuronal genes. Transient transfection of NeuroD2 induced robust neuronal differentiation as demonstrated by expression of a neuron-specific beta-tubulin and neuronal morphology. MyoD induced MYOG expression, while the chimera induced both neuron-specific beta-tubulin and MYOG, but usually not within the same cell (Figure 2C). Similar results were obtained staining for the neuronal marker internexin and the muscle marker desmin (Figure S1B). To further test the activity of the chimera, we expressed it in zebrafish, a model where overexpression of *myod* and *neurod* mRNAs has been shown to induce myogenesis and neurogenesis, respectively (Osborn et al., 2011; Wang et al., 2000). After injection of zebrafish embryos with mRNAs for MyoD, NeuroD2, or the M(N) chimera, MyoD induced widespread ectopic muscle differentiation (Figure 2C; 43/43 embryos), NeuroD2 induced robust ectopic neuronal differentiation (Figure 2C; 40/40 embryos), and the chimera induced ectopic neuronal differentiation and some ectopic muscle differentiation (Figure 2C; 38/38 embryos with ectopic neurons, nine of 38 with ectopic muscle), consistent with our findings in P19 cells. In these zebrafish embryo experiments, we also did not observe co-localization of neuronal and muscle markers.

The E-box motif analysis indicated that the chimera had lost binding to the MyoD private GG E-box (see Figure 1D), whereas the ChIP-seq peak locations showed that the chimera maintained binding at a substantial number of sites that were bound only by MyoD, but not NeuroD2 (Figure 2D). A discriminative motif analysis of the 200 nt region surrounding the sites bound by both the chimera and MyoD, but not NeuroD2, revealed that the MyoD/chimera-specific sites had a motif enrichment for TALE transcription factors of the PBX/MEIS family, whereas the NeuroD2-specific sites had a different PBX-like motif enrichment (Figure 3A), suggesting that different PBX/MEIS-like complexes might interact with MyoD and NeuroD2. A PBX/MEIS complex has previously been shown to recruit MyoD to a subset of binding sites through interaction with two specific domains in MyoD outside of the bHLH region, the Cys-His rich and Helix III domains that are N- and C-terminal to the bHLH domain, respectively (Figure 3B)(Berkes et al., 2004), and this interaction with PBX was required for the full activation of approximately 10% of MyoD-regulated genes.

To determine whether loss of PBX/MEIS interactions would reduce binding of the M(N) chimera to MyoD-specific sites, we introduced point mutations in the chimera previously shown to prevent MyoD interaction with PBX/MEIS: W96A/C98A in the His-Cys domain (M(N)WC), S253P in Helix 3 (M(N)S), or both

(B) Gene ontology analysis of upregulated genes identified from RNA-seq in (A) demonstrates enrichment for both muscle and neuronal development categories by the chimera. Genes are grouped as shared between the chimera and MyoD (M(N):MD), shared with NeuroD2 (M(N):ND2), shared with both MyoD and NeuroD2 (M(N):MD:ND2), or activated by the chimera alone (M(N)). The heatmap colors represent –log10 (enrichment p values).

(C) M(N) activates expression of both muscle and neuronal differentiation markers. P19 cells were transfected with control (i), MyoD (ii), NeuroD2 (iii), or M(N) chimera (iv) expression vectors and an E12 construct containing a puromycin resistance gene. Following selection and differentiation, immunostaining was performed using antibodies against neuron-specific tubulin (green) and myogenin (red), with DAPI nuclear stain (blue). (a) Control zebrafish embryo, 24 hr postfertilization (hpf), immunostained with anti-GFP (to label expression of the skeletal muscle transgene *mylpfa:GFP* (von Hofsten et al., 2008) (green) and neuron-specific anti-Elavl (Kim et al., 1996) (red). Embryo is shown in dorsal view, anterior to the left, at the level of the posterior hindbrain and anterior spinal cord. Arrowheads point to cranial ganglia. (b–d) 1-cell stage zebrafish embryos were injected with mRNAs for MyoD (b), NeuroD2 (c), or M(N) chimera (d). Embryos were allowed to develop until 24 hpf and were then immunostained as for control embryo. Embryos are shown in dorsal views as for control. Arrows in (c) and (d) point to ectopic neurons outside of the neural tube.

(D) ChIP-seq for MyoD, NeuroD2, and M(N) chimera in P19 cells transduced with each construct demonstrates significant overlap of binding sites between the chimera and both MyoD and NeuroD2. Venn diagram demonstrates degree of overlap for the top 15,000 binding sites for each factor.

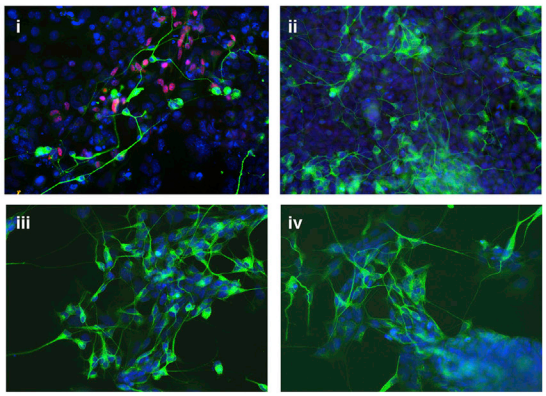
See also Figure S2. Analysis of the correlation between binding of MyoD or NeuroD2 and gene regulation has been presented previously (Fong et al., 2012).

A

Consensus	Z score	ratio	fg.frac	bg.frac	logo		
NNVYCHVTCANN	24.4	2.17	0.644	0.368		TALE (Meis)	M(N) > ND2
NNHTGTCANN	20.8	1.83	0.646	0.425		TALE (Meis)	
NNCATTCHNN	17.5	1.91	0.459	0.282		TEAD1	
NNGNCCTTGANN	12.9	4.04	0.104	0.027		Esrrb	
NNVCAGCTGNYNN	-39.8	0.39	0.578	0.948		E-box	ND2 > M(N)
NNCAGATGBBNN	-36.5	0.32	0.385	0.843		E-box	
NNRATCAATNN	-9.7	0.30	0.027	0.090		TALE (Pbx)	

B

C



D

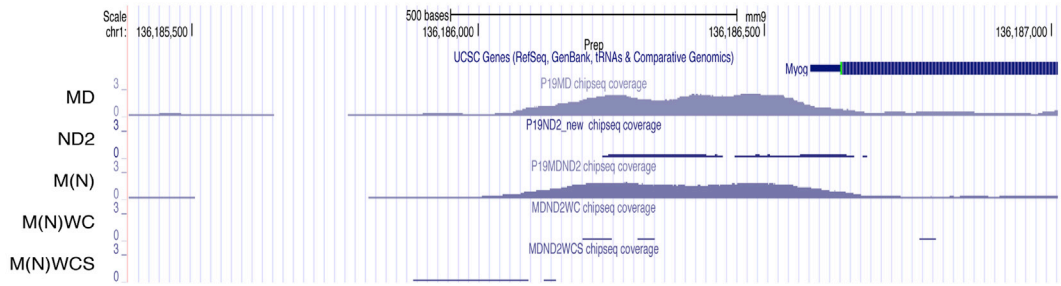


Figure 3. PBX/MEIS Co-factor Interaction Recruits M(N) Chimera to MyoD Sites and Is Required for Activation of Myogenin

(A) Discriminative motif enrichment analysis was performed comparing sites bound by M(N) but not shared with NeuroD2 against sites bound by NeuroD2 and not shared with M(N). M(N)-bound sites, which include sites shared with MyoD, are strongly enriched for a TALE factor motif (MEIS), but not the MyoD private E-box, whereas a different TALE factor motif is enriched in the NeuroD2 sites.

(B) Schematic of the M(N) chimera with the MyoD His-Cys rich (H/C) and Helix III domains located N-terminal and C-terminal to bHLH domain, respectively. Induced point mutations are depicted within the H/C region (W96A/C98A) and Helix III (S253P).

(legend continued on next page)

mutations (M(N)WCS). Single-cell analysis of P19 cells transiently transfected with these constructs showed that these mutants did not activate MYOG expression and appeared to have a more robust induction of the neuronal tubulin and phenotype (Figure 3C). In addition, the mutants lost the ability to activate a reporter driven by a myogenin promoter region containing the PBX/MEIS binding motif (Figure S2A), an assay used to determine MyoD-PBX interaction because these proteins do not co-immunoprecipitate (Berkes et al., 2004).

ChIP-seq following lentiviral transduction of P19 cells demonstrated that, whereas MyoD and the unmodified chimera both bound at the *Myog* promoter, the WC and WCS chimeras, as well as NeuroD2, did not bind to the *Myog* promoter (Figure 3D), similar to what was previously shown for MyoD (Berkes et al., 2004). Genome-wide analysis of the ChIP-seq data for the WCS chimera demonstrated a loss of binding to the MyoD-specific sites that were bound by the M(N) chimera (Figure 4A, upper panel) and enhanced binding to NeuroD2 sites. Notably, there also was loss of enrichment for the PBX/MEIS motifs at sites bound by WCS chimera (Figure S2B), confirming that these mutations were disrupting recruitment to sites with this motif, while the E-box motif of the WC and WCS chimeras remained unchanged and identical to NeuroD2 (Figure S2C).

RNA-seq demonstrated that the WC and WCS chimeras no longer activated many of the MyoD-regulated genes that were activated by the unmodified chimera, and showed more robust activation of the NeuroD2-regulated genes (Figure 4A, lower panel). The WCS chimera activated about 60% of the NeuroD2-regulated genes, but only about 5% of the MyoD-specific genes (Figure 4B), indicating that the WCS chimera was nearly a pure neurogenic transcription factor with only minor residual myogenic activity. Principal component analysis (PCA) (Figure 4C) and hierarchical clustering (Figure S3A) also demonstrated that the M(N) chimera represented an intermediate state between MyoD and NeuroD2, with weaker activation of MyoD and NeuroD2 genes, whereas the WC and WCS chimeras induced an expression pattern close to NeuroD2.

While the decreased activation of myogenic genes by the WC, S, and WCS chimeras correlated with decreased binding to MyoD-specific sites, the reason for the increased binding to NeuroD2-specific sites and activation of neurogenic genes by these mutants (see Figure 4A) was less clear. To determine whether binding to MyoD-specific sites simply decreased the amount of transcription factor available for binding to the NeuroD2 sites, we overexpressed the original M(N) chimera or the WCS chimera by increasing the lentiviral titer 5-fold. Higher amounts of either the unmodified chimera or the WCS chimera substantially increased the number of NeuroD2-regulated genes, with the chimera activating ~75% of NeuroD2-regulated genes and the WCS chimera activating ~90% of NeuroD2 genes (Figure 4D; Figure S3B). Therefore, part of the decreased activation of neurogenic genes by the chimera might reflect decreased amounts of unbound factor due to competition from MyoD-specific sites. However, at these higher titers the chimera maintained a substantial activation of MyoD-specific genes (~49%) compared to the WCS chimera (~12%). It should be noted that even at these higher expression levels, the co-expression of muscle and neuronal genes in the same cell remained rare (Figure S3C), suggesting that there is an either/or decision that might limit the abundance of neurogenic gene transcripts in a population of cells transduced by the chimera. Finally, another consideration is that the WCS chimera does not have the "myogenic code," the AT and K amino acids previously shown to be critical for muscle gene expression (Bengal et al., 1994; Davis et al., 1990; Davis and Weintraub, 1992). However, re-introducing these into the MyoD chimera did not alter its induction of either neuronal or muscle genes (Figure S4).

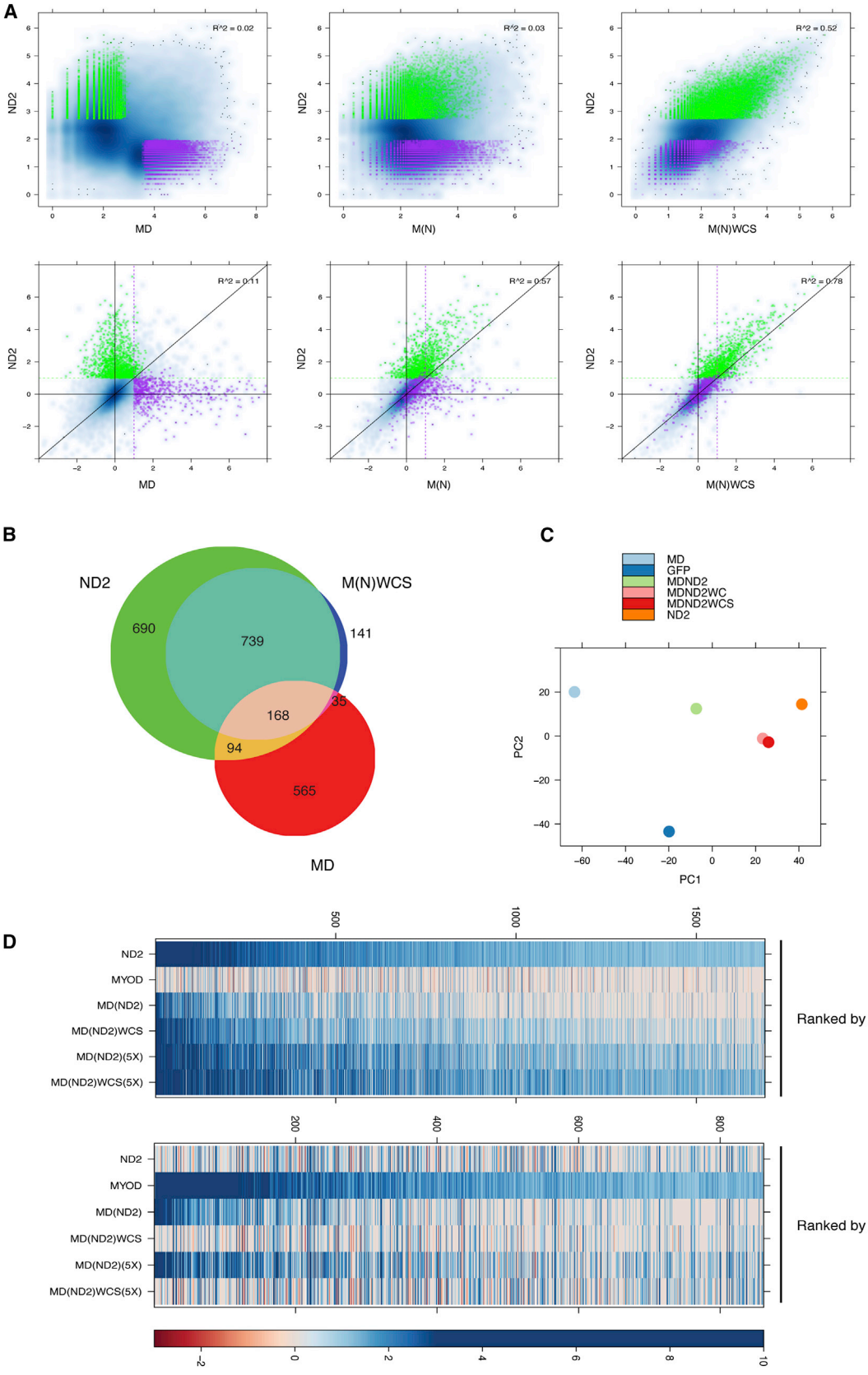
DISCUSSION

Our results demonstrate that MyoD can be converted to a potent neurogenic factor by reprogramming its binding to sites bound by NeuroD2. The chimera has gained a binding preference for the NeuroD2 "private" CAGATG E-box that is associated with gene activation, while losing binding to the MyoD "private" CAGGTG E-box and maintaining binding to the CAGCTG E-box that is shared by both MyoD and NeuroD2. Previously, we showed that the NeuroD2 private E-box sites are more strongly associated with gene activation than binding to the shared E-box, both in vivo and in reporter constructs driven by paired E-boxes (Fong et al., 2012). In this report, we demonstrated that both NeuroD2 and the MyoD chimera have a higher binding affinity for the NeuroD2 private E-box compared to the shared E-box, as determined by slower off rates. The increased dwell time at the private site might be sufficient for the increased transcriptional activity of NeuroD2 and the MyoD chimera on these sites.

MyoD and NeuroD2 have little homology outside of the bHLH region, and it is interesting that redirecting MyoD from its private CAGGTG site to the NeuroD2 CAGATG site is sufficient to convert MyoD from a muscle lineage determination factor to one that determines the neuronal lineage. This result suggests that binding location and affinity confers lineage determination rather than other features of the MyoD or NeuroD2 transcription factors. It is important to recognize that other factors might interact with the bHLH region. At this time we do not have any information regarding alternative E-protein partners, and we achieve the same results when the transfections are performed with co-transfected E12 that should drive dimerization to this common E-protein partner (A.P.F., unpublished data). Therefore, outside of the bHLH domain and targeting private binding sites, there does not appear to be any significant lineage determination function for MyoD or NeuroD2.

This study also confirms the importance of homeodomain complexes (or complexes that bind to homeodomain motifs),

(C) Mutation of H/C and/or Helix III domains leads to loss of MYOG expression and enhanced neuronal differentiation. P19 cells were transfected with expression constructs for M(N) chimera (i), M(N) with S253P mutation (ii), M(N) with W96A/C98A mutations (iii), or M(N) with W96A/C98A/S253P mutations (iv). Immunostaining was performed for neuron-specific tubulin (green) and MYOG (red), with DAPI nuclear stain (blue).
(D) Genome browser screen shot of the *Myog* promoter demonstrates ChIP-seq peaks for both MyoD and M(N) chimera, but not NeuroD2, M(N)WC, or M(N)WCS. See also Figure S3.



(legend on next page)

such as PBX/MEIS, for the recruitment of MyoD to a subset of promoters, such as myogenin. Our prior studies showed that a PBX/MEIS complex was present at the myogenin promoter in fibroblasts, prior to conversion to muscle by MyoD, and that the WC, S, or WCS mutations in MyoD that disrupt its interaction with PBX/MEIS also prevent recruitment to the myogenin promoter (Berkes et al., 2004). The findings in the present study confirm the importance of these interactions at a larger subset of genes regulated by MyoD and the importance of homeodomain proteins in marking this subset for future activation by MyoD. In addition, the computational identification of a distinct homeodomain motif associated with sites bound only by NeuroD2 and not MyoD or the chimera suggests that NeuroD2 might have its own interacting domain and homeodomain complexes that guide its binding at a subset of genes. Together, these results indicate that the ability of a cell to respond to the lineage instruction of MyoD or NeuroD2 might depend in part on the homeodomain factors that mark specific genes for lineage-specific activation.

EXPERIMENTAL PROCEDURES

Plasmid Constructs and Neuronal Differentiation

For lentiviral constructs, cDNAs for NeuroD2 and MyoD were cloned into the GFP site of the pRRL.SIN.cPPT.pGK-GFP.WPRE lentiviral plasmid. cDNA for M(N) chimera was first synthesized (GenScript) and then sub-cloned into pCS2 and pRRL as above. W96A/C98A and S253P mutations were introduced by site-directed mutagenesis (Agilent) within the pCS2 backbone per the manufacturer's protocol, followed by sub-cloning into pRRL. Replication-incompetent lentiviral particles were packaged in 293T cells by the Fred Hutchinson Cancer Research Center Virus Core Facility. P19 cells were infected in MEMα media containing polybrene at a final concentration of 8 μg/ml. After 24 hr, fresh MEMα was added, and, after 48 hr, cells were transitioned to Neurobasal media (Gibco) containing B27 supplement and 5 mM glutamine (Farah et al., 2000). Neurobasal media was replaced every 24 hr until cells were harvested. For transfections, pCS2-MyoD, NeuroD2, or M(N) were co-transfected with pCL-E12, a construct with a puromycin resistance gene, into P19 cells and 24 hr later, cells were selected with puromycin for 48 hr, followed by transition to complete Neurobasal media. For reporter assays, a pair of CAGATG, CAGGTG, or CAGCTG E-boxes were cloned upstream of the SV40 promoter in pGL3-Pro (Promega) and then co-transfected into P19 cells with the indicated bHLH construct and harvested 48 hr later. * represents p value <0.05 determined by Student's t test compared to empty vector.

Immunohistochemistry and Antibodies

Cells were fixed for 10 min with 4% paraformaldehyde in phosphate-buffered saline (PBS) and permeabilized using 0.5% Triton X-100 in PBS. For immunostaining and ChIP, antibodies used include (1) two different NEUROD2 antibodies raised in rabbits by our lab (1240J, 1168J); (2) anti-MYOD generated in our lab and previously characterized (Tapscott, et. al., 1988); (3) Tuj1, anti-neuron-specific tubulin (Covance); (4) Myogenin M-225 (Santa Cruz Biotechnology); (5) Desmin D93F5 XP (Cell Signaling Technology); (6) Beta-III tubulin EPR1568Y (Abcam); (7) myosin heavy chain MF20 (DHB); (8) Internexin 2E3 (Santa Cruz). Secondary antibodies (Jackson ImmunoResearch Laboratories) were used at 1:500. For western blot, MyoD 5.8A (Santa Cruz) raised against the C terminus of MyoD was used.

Electromobility Shift Assays

The following sequences represent the forward strands of the oligonucleotides: GA, 5-CAGAGTGACAGATGGCGGCGGG; GG, 5-CAGAGTGACAGGT GGCGGCGGG; GC, CAGAGTGACAGCTGGCGGCGGG. For dissociation assays, radiolabeled oligonucleotide probes were end labeled using PNK and [γ-^{32}P]ATP and cleaned on a Sephadex G-50 spin column. NEUROD2, MYOD, M(N), and E12 were in vitro translated using the TnT Coupled SP6 Reticulocyte Lysate kit (Promega). In-vitro-translated product was mixed with probe for 15 min at room temperature. The typical electromobility shift assay (EMSA) mixture contained 12.5 mM Tris (pH 7.9), 50 mM KCl, 5 mM MgCl 27.5% glycerol, 0.1 mM EDTA, 1 mM DTT, 0.5 pg of poly(dI-dC). Following formation of complexes, 100-fold excess of unlabeled oligo was added for the indicated amount of time, followed by resolution on 6% polyacrylamide gel at 200 V for 3 hr at 4°C.

ChIP-Seq

ChIP was performed as previously described (Cao et al., 2010). ChIP samples were validated by qPCR and prepared for sequencing with the NuGEN Ovation Ultralow Multiplex System kit. All samples were sequenced on the Illumina HiSeq 2500 by the Fred Hutchinson Cancer Research Center Genomics Resource Core Facility.

RNA-Seq

Total RNA was harvested from lentivirally transduced P19 cells that were neuronally differentiated as described above. RNA libraries were subsequently prepared using oligo-dT selection and sequenced on the Illumina HiSeq 2500. Differentially expressed genes were detected using Bioconductor package DESeq (Anders and Huber, 2010). GO enrichment analysis was performed using Bioconductor package GOstats (Falcon and Gentleman, 2007).

ChIP-Seq Analysis

ChIP-seq reads were aligned using Burrows-Wheeler Aligner to the mouse genome (mm9). Quality control and peak calling were performed using an in house-developed tool as previously described (Geng et al., 2012).

Motif Analysis

We performed discriminative de novo motif discovery using Bioconductor package motifRG (Cao et al., 2010; Palii et al., 2011; Yao et al., 2014) to find motifs that were specifically enriched in MyoD, NeuroD2, M(N) chimera, and M(N)WCS chimera binding sites. Predicted motifs were annotated based on matches to Jaspar, Uniprobe, and house-curated motif databases using

Figure 4. Disruption of the PBX/MEIS Myogenic Co-factor Interactions Results in Enhanced Binding to NeuroD2 Private Sites and Neuronal Gene Activation by the M(N) Chimera

(A) Scatterplots were generated for ChIP-seq (top row) or RNA-seq (log fold change relative to GFP, bottom row) binding site intensities in P19 cells transduced with lentivirus expressing NeuroD2, MyoD, or M(N) chimeras). The NeuroD2-specific upregulated genes or binding sites are green, and the MyoD-specific upregulated genes or binding sites are purple. There is increased upregulation of NeuroD2 genes and binding of NeuroD2 sites and decreased upregulation of MyoD genes and binding of MyoD sites by the chimera with H/C and Helix III mutations (M(N)WCS) compared with the unmodified chimera (M(N)). R^2 represents fit to the linear regression line.

(B) Venn diagram of log-2 fold upregulated genes demonstrates increased overlap between NeuroD2 and M(N)WCS and decreased overlap between MyoD and M(N)WCS compared with M(N) (see Figure 2A).

(C) Principal component analysis of RNA expression profiles demonstrates increased similarity with NeuroD2 for M(N)WCS relative to both MyoD and M(N).

(D) RNA-seq profiles demonstrate increased activation of both MyoD and NeuroD2 targets with higher expression of M(N) chimera, but increased expression only of NeuroD2 targets with higher expression of M(N)WCS. Upregulated genes are ranked from left to right for NeuroD2 (top panel) or MyoD (bottom panel).

See also Figure S4.

software Tomtom in MEME suite (Gupta et al., 2007). To show motif significance, we included motif *Z* scores, which follow normal distribution under null distribution (e.g., *Z* score of 5 translates to a p value of 5.7×10^{-7}, and an adjusted p value of 10^{-3} after Bonferroni correction for multiple hypotheses) (Yao et al., 2014).

Zebrafish Experiments

All experiments involving live zebrafish (*Danio rerio*) were carried out in compliance with IACUC guidelines. Zebrafish were raised and staged as previously described (Westerfield, 2007). The *Tg(−2.2mylpfa:GFP) i135* line that labels differentiated fast-twitch skeletal muscle has been described (von Hofsten et al., 2008). Time refers to hours postfertilization (hpf) at 28.5°C. mRNAs for injections into zebrafish embryos were synthesized in vitro, using the mMessage mMachine kit (Ambion), from the same pCS2-MyoD, NeuroD2, and M(N) expression constructs used for P19 cell transfections. mRNA injections were performed using a Narishige IM 300 Microinjector. Embryos were collected from the *Tg(−2.2 mylpfa:GFP)* strain and injected at the 1-cell stage with 2 nl of 125 ng/μl mRNAs. Sibling control embryos were not injected with any mRNAs. Embryos were processed for whole-mount immunostaining as previously described (Bird et al., 2012) with the following primary antibodies: anti-Elavl, 1:250 (clone 16A11) (Henion et al., 1996); anti-GFP, 1:300 (Torrey Pines TP401). Secondary antibodies used were goat anti-mouse Alexa Fluor 568 and goat anti-rabbit Alexa Fluor 488 (1:250, Life Technologies/Molecular Probes). Embryos were mounted in 70% glycerol in PBS under a coverslip and imaged using a Leica TCS SP5 confocal microscope.

ACCESSION NUMBERS

Sequence information has been deposited in GEO (GSE64627, GSE64626).

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.02.055.

AUTHOR CONTRIBUTIONS

A.P.F. designed, executed, interpreted, and supervised all experiments. Z.Y. performed all computational analyses, including ChIP-seq peak calling, RNA-seq analysis, and motif enrichment studies and assisted with data interpretation. J.Z. participated in experimental design and executed and analyzed reporter assays, immunofluorescence, and RNA expression experiments. N.M.J., G.H.F., and L.M. performed zebrafish experiments and analyzed data. S.J.T. provided study oversight and supervision. Manuscript was written by A.P.F., J.Z., and S.J.T.

ACKNOWLEDGMENTS

S.J.T. was supported by NIH NIAMS R01AR045113, and A.P.F was supported by NIH NCI 1K08CA169014. N.M.J., G.H.F., and L.M. were supported by NIH NIAMS R03AR065760 and the Seattle Children's Myocardial Regeneration Initiative. We acknowledge the help of S. Stone with the ATK myogenic code experiments. We thank P. Ingham and the University of Sheffield for the transgenic zebrafish.

Received: September 18, 2014
Revised: January 16, 2015
Accepted: February 23, 2015
Published: March 19, 2015

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Bengal, E., Flores, O., Rangarajan, P.N., Chen, A., Weintraub, H., and Verma, I.M. (1994). Positive control mutations in the MyoD basic region fail to show cooperative DNA binding and transcriptional activation in vitro. Proc. Natl. Acad. Sci. USA *91*, 6221–6225.

Berkes, C.A., and Tapscott, S.J. (2005). MyoD and the transcriptional control of myogenesis. Semin. Cell Dev. Biol. *16*, 585–595.

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.

Bird, N.C., Windner, S.E., and Devoto, S.H. (2012). Immunocytochemistry to study myogenesis in zebrafish. Methods Mol. Biol. *798*, 153–169.

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.

Chae, J.H., Stein, G.H., and Lee, J.E. (2004). NeuroD: the predicted and the surprising. Mol. Cells *18*, 271–288.

Davis, R.L., and Weintraub, H. (1992). Acquisition of myogenic specificity by replacement of three amino acid residues from MyoD into E12. Science *256*, 1027–1030.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987–1000.

Davis, R.L., Cheng, P.F., Lassar, A.B., and Weintraub, H. (1990). The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. Cell *60*, 733–746.

Falcon, S., and Gentleman, R. (2007). Using GOstats to test gene lists for GO term association. Bioinformatics *23*, 257–258.

Farah, M.H., Olson, J.M., Sucic, H.B., Hume, R.I., Tapscott, S.J., and Turner, D.L. (2000). Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. Development *127*, 693–702.

Fong, A.P., and Tapscott, S.J. (2013). Skeletal muscle programming and reprogramming. Curr. Opin. Genet. Dev. *23*, 568–573.

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.

Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J.M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R., and Tapscott, S.J. (2012). DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. Dev. Cell *22*, 38–51.

Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. Genome Biol. *8*, R24.

Henion, P.D., Raible, D.W., Beattie, C.E., Stoesser, K.L., Weston, J.A., and Eisen, J.S. (1996). Screen for mutations affecting development of Zebrafish neural crest. Dev. Genet. *18*, 11–17.

Kim, C.H., Ueshima, E., Muraoka, O., Tanaka, H., Yeo, S.Y., Huh, T.L., and Miki, N. (1996). Zebrafish elav/HuC homologue as a very early neuronal marker. Neurosci. Lett. *216*, 109–112.

Lee, J.E., Hollenberg, S.M., Snider, L., Turner, D.L., Lipnick, N., and Weintraub, H. (1995). Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. Science *268*, 836–844.

Longo, A., Guanga, G.P., and Rose, R.B. (2008). Crystal structure of E47-NeuroD1/beta2 bHLH domain-DNA complex: heterodimer selectivity and DNA recognition. Biochemistry *47*, 218–229.

Ma, P.C., Rould, M.A., Weintraub, H., and Pabo, C.O. (1994). Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. Cell *77*, 451–459.

Osborn, D.P., Li, K., Hinits, Y., and Hughes, S.M. (2011). Cdkn1c drives muscle differentiation through a positive feedback loop with Myod. Dev. Biol. *350*, 464–475.

Palii, C.G., Perez-Iratxeta, C., Yao, Z., Cao, Y., Dai, F., Davison, J., Atkins, H., Allan, D., Dilworth, F.J., Gentleman, R., et al. (2011). Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages. EMBO J. *30*, 494–509.

Skerjanc, I.S., Slack, R.S., and McBurney, M.W. (1994). Cellular aggregation enhances MyoD-directed skeletal myogenesis in embryonal carcinoma cells. Mol. Cell. Biol. *14*, 8451–8459.

Tapscott, S.J., Davis, R.L., Thayer, M.J., Cheng, P.F., Weintraub, H., and Lassar, A.B. (1988). MyoD1: a nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. Science *242*, 405–411.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035–1041.

von Hofsten, J., Elworthy, S., Gilchrist, M.J., Smith, J.C., Wardle, F.C., and Ingham, P.W. (2008). Prdm1- and Sox6-mediated transcriptional repression specifies muscle fibre type in the zebrafish embryo. EMBO Rep. *9*, 683–689.

Wang, X., Wan, H., Korzh, V., and Gong, Z. (2000). Use of an IRES bicistronic construct to trace expression of exogenously introduced mRNA in zebrafish embryos. Biotechniques *29*, 814–816, 818, 820.

Westerfield, M. (2007). The Zebrafish Book, Fifth Edition: A guide for the laboratory use of zebrafish (Danio rerio) (University of Oregon Press).

Yao, Z., Macquarrie, K.L., Fong, A.P., Tapscott, S.J., Ruzzo, W.L., and Gentleman, R.C. (2014). Discriminative motif analysis of high-throughput dataset. Bioinformatics *30*, 775–783.