

## Article

# Promoter-Adjacent DNA Hypermethylation Can Downmodulate Gene Expression: *TBX15* in the Muscle Lineage

Kenneth C. Ehrlich<sup>1</sup>, Michelle Lacey<sup>2</sup>, Carl Baribault<sup>3</sup>, Sagnik Sen<sup>4</sup> , Pierre Olivier Esteve<sup>4</sup>, Sriharsa Pradhan<sup>4</sup> and Melanie Ehrlich<sup>1,5,\*</sup> 

<sup>1</sup> Biomedical Informatics and Genomics Center, Tulane University Health Sciences Center, New Orleans, LA 70112, USA

<sup>2</sup> Department of Mathematics, Tulane University, New Orleans, LA 70118, USA

<sup>3</sup> Information Technology, Tulane University, New Orleans, LA 70118, USA

<sup>4</sup> Genome Biology Division, New England Biolabs, Ipswich, MA 01938, USA

<sup>5</sup> Tulane Cancer Center and Hayward Genetics Center, Tulane University Health Sciences Center, New Orleans, LA 70112, USA

\* Correspondence: ehrlich@tulane.edu; Tel.: +1-504-988-2449

**Abstract:** *TBX15*, which encodes a differentiation-related transcription factor, displays promoter-adjacent DNA hypermethylation in myoblasts and skeletal muscle (psoas) that is absent from non-expressing cells in other lineages. By whole-genome bisulfite sequencing (WGBS) and enzymatic methyl-seq (EM-seq), these hypermethylated regions were found to border both sides of a constitutively unmethylated promoter. To understand the functionality of this DNA hypermethylation, we cloned the differentially methylated sequences (DMRs) in CpG-free reporter vectors and tested them for promoter or enhancer activity upon transient transfection. These cloned regions exhibited strong promoter activity and, when placed upstream of a weak promoter, strong enhancer activity specifically in myoblast host cells. In vitro CpG methylation targeted to the DMR sequences in the plasmids resulted in 86–100% loss of promoter or enhancer activity, depending on the insert sequence. These results as well as chromatin epigenetic and transcription profiles for this gene in various cell types support the hypothesis that DNA hypermethylation immediately upstream and downstream of the unmethylated promoter region suppresses enhancer/extended promoter activity, thereby downmodulating, but not silencing, expression in myoblasts and certain kinds of skeletal muscle. This promoter-border hypermethylation was not found in cell types with a silent *TBX15* gene, and these cells, instead, exhibit repressive chromatin in and around the promoter. *TBX18*, *TBX2*, *TBX3* and *TBX1* display *TBX15*-like hypermethylated DMRs at their promoter borders and preferential expression in myoblasts. Therefore, promoter-adjacent DNA hypermethylation for downmodulating transcription to prevent overexpression may be used more frequently for transcription regulation than currently appreciated.

**Keywords:** DNA hypermethylation; myoblasts; skeletal muscle; enhancers; *TBX15*; T-box transcription factors; whole-genome bisulfite sequencing; EM-seq; transfection; DMRs



**Citation:** Ehrlich, K.C.; Lacey, M.; Baribault, C.; Sen, S.; Esteve, P.O.; Pradhan, S.; Ehrlich, M. Promoter-Adjacent DNA Hypermethylation Can Downmodulate Gene Expression: *TBX15* in the Muscle Lineage. *Epigenomes* **2022**, *6*, 43. <https://doi.org/10.3390/epigenomes6040043>

Academic Editor: Ernesto Guccione

Received: 15 November 2022

Accepted: 6 December 2022

Published: 9 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The transcription factors (TFs) that play key roles during embryonic differentiation often also have important tissue-specific postnatal functions [1]. *TBX15* encodes a TF in the T-box family involved in embryogenesis, regulation of various postnatal developmental processes, and homeostasis [2–6]. T-box domain transcription factors can act as repressors and, less frequently, as activators of transcription [3]. Inactivation of these TFs by inherited mutations in humans leads to congenital defects in skeletal/craniofacial bone formation (*TBX15*, *TBX18*, *TBX1*, *TBX3*, *TBX4*, *TBX5*, *TBX6*, *TBX22*, and *TBXT* proteins) and heart formation (*TBX1*, *TBX3*, *TBX5*, and *TBX20* proteins) [7,8]. Studies of animal models and human cell cultures indicate additional functions of T-box TFs, such as regulation of skin

pigmentation (murine *Tbx15*), lung or liver hypoplasia (human *TBX2* and *TBX3*), early embryogenesis (murine *Tbr2* and *Tbxt*), and skeletal muscle (SkM) physiology (murine *Tbx15*, frog/bird-specific *Tbx16*, and *C. elegans* *TBX-2*) [3,9–12]. Studies of mice indicated a strong association of *Tbx15* expression with muscle fiber type [10] and a positive association of *Tbx15* expression with lean mass and several metabolic phenotypes [13].

Evidence for the importance of *TBX15* to the human SkM lineage was seen in our previous comparison of transcriptomic and epigenomic profiles of *Tbx15/TBX15* in SkM tissue and myoblasts [14,15]. Myoblasts are SkM progenitor cells involved in embryogenesis and postnatal repair of muscle damage [16]. We found that human myoblast primary cultures and their differentiation product, multinucleated myotubes, express moderate levels of *TBX15* RNA while there is little or no expression of this gene in five diverse cell cultures that are not derived from mesoderm. Moreover, SkM had the highest expression of *TBX15* in a comparison of 52 studied human tissues [14,17].

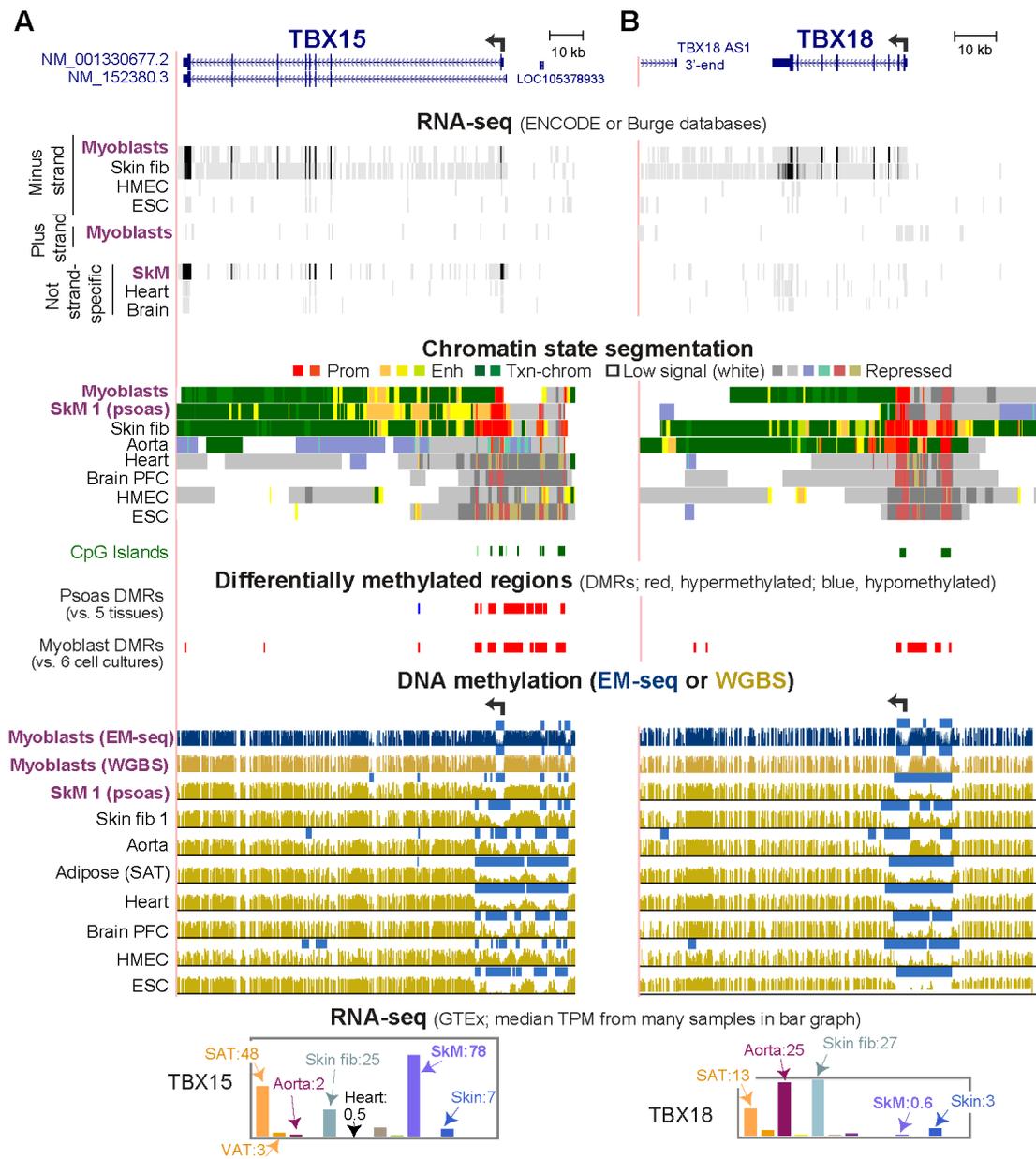
Analysis of methylomes generated by reduced representation bisulfite sequencing (RRBS) surprisingly revealed that transcriptionally active *TBX15* is strongly hypermethylated both immediately upstream and downstream of the unmethylated promoter region in myoblasts and myotubes compared with 15 types of cell cultures not expressing the gene. Association of intragenic DNA hypermethylation with actively transcribed gene bodies is a frequent, but not universal, finding that could be due to gene-body DNA methylation modulating the rate of movement of the transcription complex, regulating alternative splicing, or repressing cryptic promoters, retrotransposons, enhancers, and silencers [18]. However, gene-body DNA methylation that is positively correlated with transcription is observed most strongly in DNA sequences considerably downstream of the transcription start site (TSS) [18,19], unlike the promoter-adjacent intragenic hypermethylation that we observed in *TBX15* in the SkM lineage [14].

To elucidate the role of the DNA hypermethylation around the constitutively unmethylated *TBX15* promoter in *TBX15*-expressing myoblasts and SkM, we cloned DNA sequences from several of these differentially methylated regions (DMRs) and tested their ability to act as promoters or enhancers in reporter gene transfection assays with or without in vitro CpG methylation targeted to these sequences. We found that the hypermethylated DNA sequences immediately upstream and downstream of the promoter have strong promoter or enhancer activity when unmethylated and transfected into myoblasts. These results and our in-depth analysis of epigenomic vs. transcriptomic profiles of many cell and tissue types support the hypothesis that the myoblast-associated promoter-adjacent DNA hypermethylation at *TBX15* fine-tunes expression of this gene by downmodulation. Our findings have implications for better understanding differentiation within the SkM lineage and the transcriptional and epigenetic changes in SkM that occur with exercise and aging [20].

## 2. Results

### 2.1. Myoblast DNA Hypermethylation around an Unmethylated Promoter Region in Five T-Box Genes Was Positively Associated with Their Expression

RNA-seq databases show that human *TBX15* is preferentially transcribed in SkM tissue (highest expression), myoblasts, myocytes, fibroblast-type cells in SkM, as well as in some other cell types in postnatal tissues (smooth muscle cells and adipocytes; Figures 1A and S1; Table S1). Moreover, among fetal tissues, SkM shows the highest expression of *TBX15* (Table S2). However, there are considerable differences in transcript levels for *TBX15* in human SkM depending on the anatomical origin of the tissue (Table S2). *TBX15*'s closest related T-box encoding gene is *TBX18* (Figure 1B) [3], which has less of a preference for expression in myoblasts than does *TBX15* (Table S3).



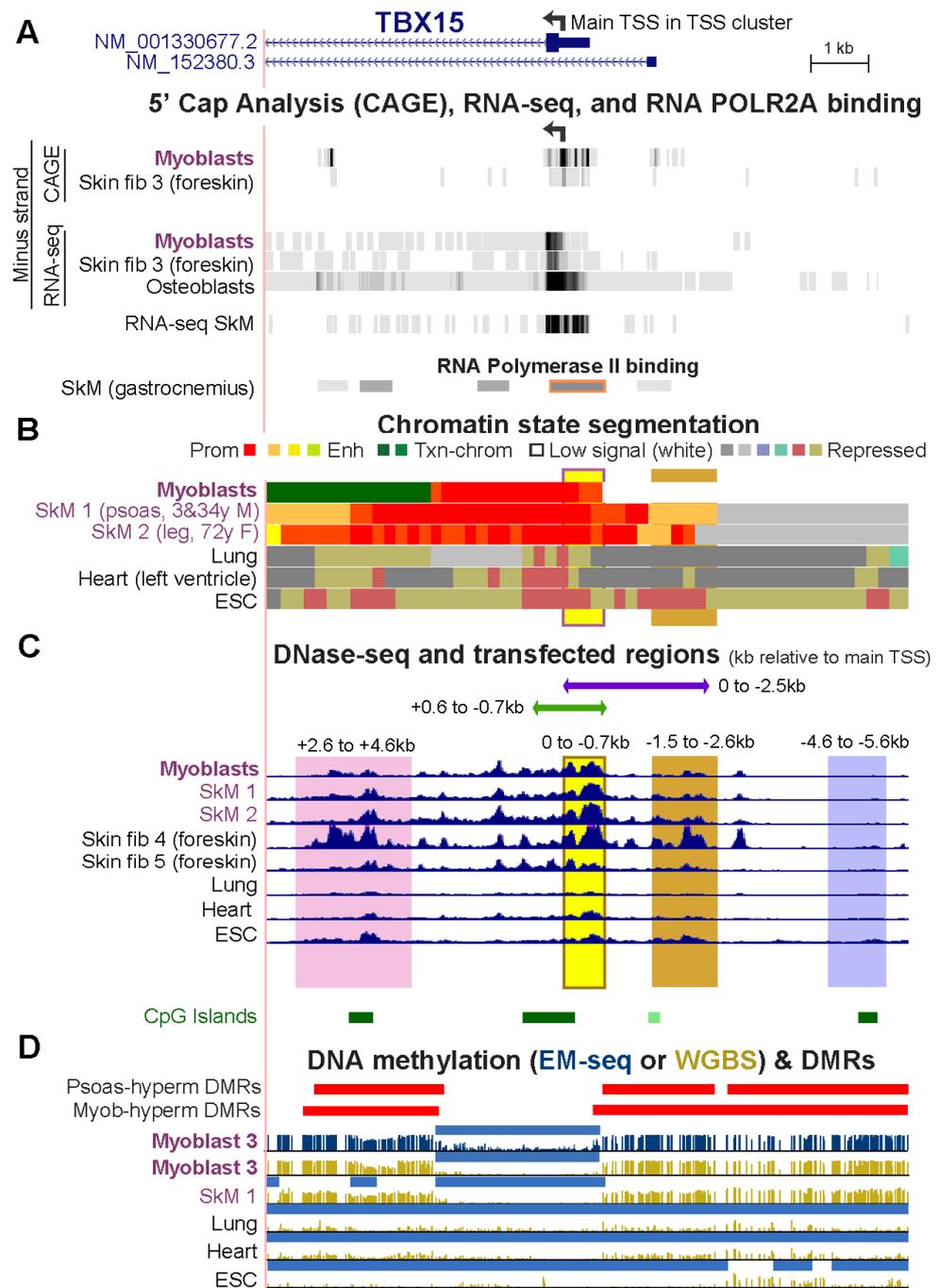
**Figure 1.** DNA hypermethylation in myoblasts borders the active, unmethylated promoter regions of *TBX15* and *TBX18*. (A) *TBX15* (chr1: 119,423,765–119,554,608) and (B) *TBX18* RefSeq gene structures (chr6: 85,410,477–85,505,693). Strand-specific RNA-seq for myoblasts, foreskin fibroblasts (Fib, Fib 3), mammary epithelial cells (HMEC), and H1 embryonic stem cells (ESC) and not strand-specific RNA-seq for skeletal muscle (SkM), heart and brain. Color-coded chromatin state segmentation indicates promoter or mixed promoter/enhancer (prom), enhancer (enh), or repressed types of chromatin, actively transcribed chromatin (txn chrom), or chromatin with little or no signal for H3K27ac, H3K27me3, and H3K4/H3K9 methylation. Significant tissue-specific (SkM, psoas) DMRs and cultured cell-type specific (myoblast) DMRs are shown. Methylome profiles are depicted in gold for WGBS and, in dark blue, for EM-seq; regions having significantly lower methylation relative to the same genome [21] are shown by light blue bars. Myoblast 3 cell strain was used for EM-seq and WGBS and Skin fib 1 (foreskin fibroblasts) for WGBS. GTEx RNA-seq expression profiles are displayed as linear-scale TPM bar graphs with some of the median TPM values from biological replicates indicated. SAT, subcutaneous adipose; VAT, visceral adipose; heart, left ventricle. All tracks are from the UCSC Genome Browser (hg19) and, except for the GTEx bar graphs, are aligned.

In myoblasts and SkM, expression of *TBX15* was positively correlated with DNA hypermethylation both immediately upstream and downstream of the unmethylated promoter region (Figure 1A). *TBX18* displayed a similar correlation but only for myoblasts (Figure 1B). In addition, both genes in skin fibroblasts showed a positive association between transcription and DNA hypermethylation upstream of the promoter although the hypermethylation was not as strong as in myoblasts. *TBX18* exhibited this association for aorta too. As expected, a lack of methylation at the TSS does not suffice for appreciable T-box gene expression, as seen for *TBX15* in heart and *TBX18* in SkM (Figure 1B). In tissues not expressing these two genes, the promoter regions, which contained mostly unmethylated DNA, exhibited repressive histone H3 lysine-27 trimethylation (H3K27me3). For our DNA methylation analyses, DMRs were identified by comparing methylomes from SkM (psoas) with five diverse tissues [22] and by comparing three myoblast cell strains with six types of non-cancerous cell cultures (Figure S2). Methylomes for the non-myoblast cultures and for all the tissues had been determined by whole-genome bisulfite sequencing (WGBS) [23]. The myoblast methylomes were generated by the recent enzymatic methyl-seq methodology (EM-seq) [24]. One myoblast cell strain (Myoblast 3) was analyzed by WGBS as well as EM-seq. Both methods gave similar DNA methylation profiles (Figure 1A, gold vs. blue methylome tracks for the Myoblast 3 cell strain). Cis-acting transcription-control chromatin elements, such as, repressive chromatin (H3K27me3 or H3K9me3), active promoters (H3K27 acetylation, H3K27ac, plus H3K4me3) and active enhancers (H3K27ac plus H3K4me1) had been inferred from chromatin segmentation state profiles derived from whole-genome maps of diagnostic histone modifications (Roadmap Project [19]).

All 17 members of the T-box family of genes are expressed postnatally in highly cell- or tissue-specific patterns (Table S1) consistent with their important roles in development [3,25]. Like *TBX15* and *TBX18*, three other T-box genes, *TBX2*, *TBX3*, and *TBX1*, were more highly expressed in myoblasts than most or all the five other diverse cell culture types in an ENCODE database (Table S3). In myoblasts, *TBX2*, and *TBX3* also displayed *TBX15*-like DNA hypermethylation bordering both sides of a constitutively unmethylated promoter region (Figures 1, S3 and S4). This promoter-adjacent DNA hypermethylation was not seen in non-expressing cell cultures or tissues nor was it observed in lung fibroblasts, which had by far the highest levels of expression of the six examined cell cultures and enhancer or promoter-type chromatin covering the whole gene, promoter, and promoter-upstream region. At *TBX1*, a myoblast-associated hypermethylated DMR (Myob-hyperm DMR) was adjacent to the upstream border of its unmethylated proximal promoter region (Figure S3). There was high methylation immediately adjacent to this promoter's downstream border but this methylated region was not a DMR because it was present in most examined cell types. About 1 kb downstream from the promoter, a region of myoblast-associated hypermethylation (another Myob-hyperm DMR) was seen. Although there are RefSeq isoforms of *TBX1* with an alternate distal promoter, examination of RNA-seq databases at the UCSC Genome Browser having many diverse samples provides evidence for use of only the proximal promoter (Figure S3A and data not shown).

*TBX20*, *TBX4*, and *TBX5* were among the T-box genes not expressed in myoblasts (Table S3). This transcription silencing in myoblasts correlated with Myob-hyperm DMRs covering their promoter regions instead of being only adjacent to them (dotted boxes, Figures S4 and S5). These three genes also exhibited repressive H3K27me3 at the promoter region. CpG islands (CGI) are present in the promoter regions of the five myoblast-expressed TBX genes as well as the three above-mentioned myoblast-repressed genes (Figures 1, 2 and S3–S5). In the myoblast-repressed genes, the promoter/CGI-overlapping DMRs could contribute to the gene repression. As reported for some intragenic CGIs [26], the CGI-overlapping Myob-hyperm DMRs in these T-box genes in various non-expressing cell types often overlapped bivalent (mixed H3K27me3-repressed and H3K4methylated-enhancer/promoter-type) chromatin instead of only H3K27me3 chromatin (Figures 1 and S3–S5) reflecting their potential for promoter or enhancer activity. We focused the rest of

the study on the 5' end of *TBX15*, and its strong promoter-adjacent (but not overlapping) DNA hypermethylation that was found only in expressing cells.



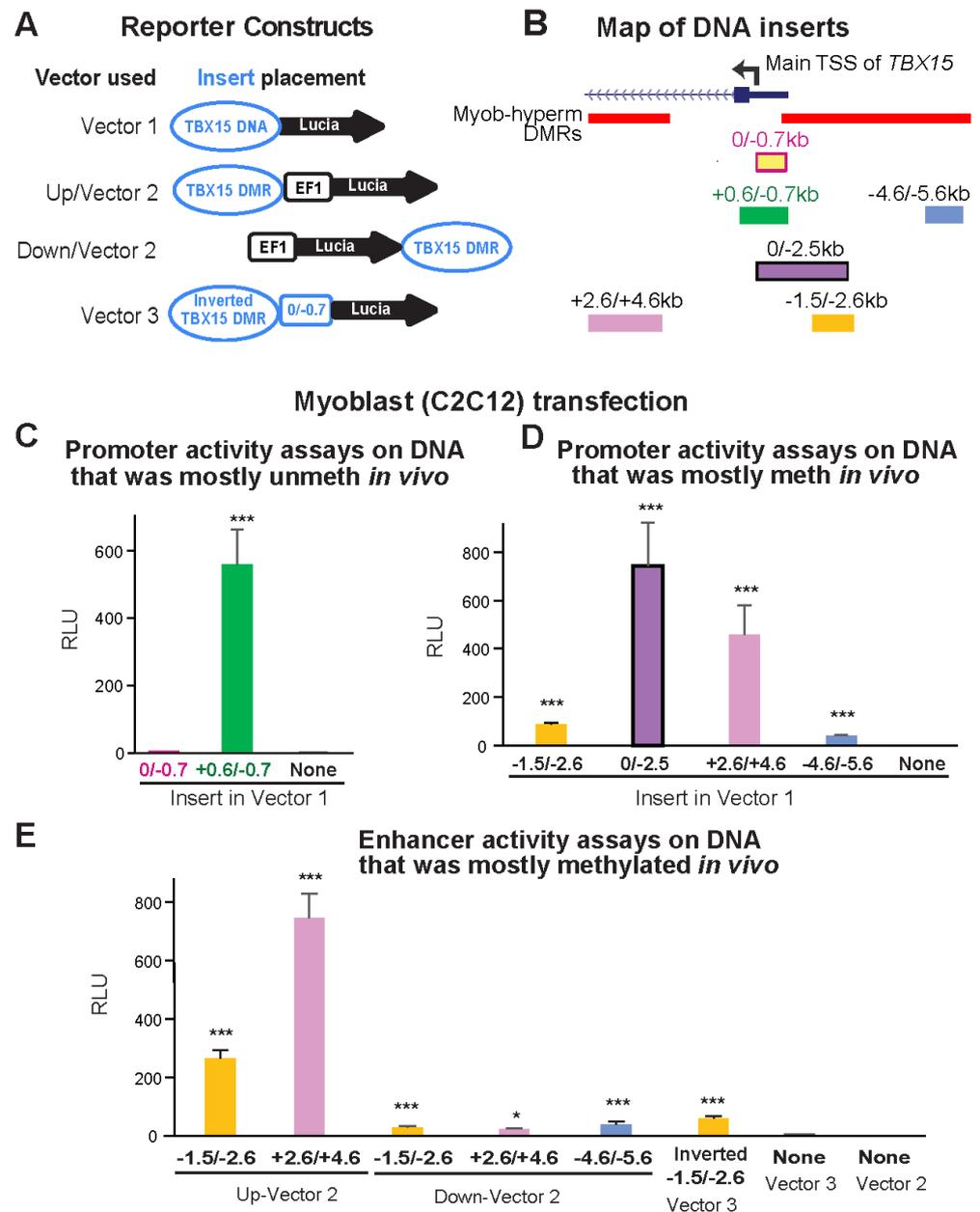
**Figure 2.** The myoblast-hypermethylated DMRs adjacent to the unmethylated *TBX15* promoter region overlap psoas-hypermethylated DMRs. (A) The 5' end of *TBX15* (chr1: 119,525,320–119,536,521) showing the minus-strand signal from 5' cap analysis of gene expression (CAGE) and/or RNA-seq of myoblasts, foreskin fibroblasts, and osteoblasts; not strand-specific RNA-seq for SkM of unknown body location (SkM 9); and the gastrocnemius muscle ChIP-seq signal for the large subunit of RNA polymerase II (ENCODE 3). (B) Chromatin state segmentation as in Figure 1. (C) DNaseI hypersensitivity profiles and the six cloned regions for transfection assays shown by horizontal bars or vertical shading. (D) Hypermethylated DMRs for SkM (psoas) vs. five heterologous tissues and for myoblasts vs. six varied cell cultures as well as methylome profiles and low methylated regions (light blue bars) as in Figure 1.

## 2.2. DNA Sequences That Were Part of Myoblast Hypermethylated DMRs near the *TBX15* Promoter Region Display Promoter or Enhancer Activity upon Transfection into Myoblasts

To understand the function of cell-type specific DNA hypermethylation around core unmethylated promoter regions, we tested the transcription regulatory activity of *TBX15* DNA sequences from Myob-hyperm DMRs in transient transfection assays using reporter gene constructs (Vector 1 or 2) for transfection into C2C12 myoblasts (Figures 2 and 3). None of the cloned sequences overlapped interspersed DNA repeats [23]. First, we ascertained the approximate location of the main *TBX15* TSS in myoblasts and skin fibroblasts so that we could use that site as a reference point for the cloned sequences. The 5' ends of the two RefSeq isoforms (RefSeq Curated, 2022) are separated by 1.2 kb (Figure 2A). However, according to cap analysis of gene expression (CAGE) and/or RNA-seq profiles from the ENCODE or RoadMap Projects, neither of these 5' ends is the best description of the *TBX15* TSS in myoblasts, skin fibroblasts, osteoblasts (which also preferentially express this gene, Table S3), and several types of SkM (Figure 2A and data not shown for SkM RNA-seq at the UCSC Genome Browser [19,23]). We used the center of the strongest CAGE signal for myoblasts and skin fibroblasts, chr1: 119,530,511 (hg19), as the nominal TSS. This site (broken arrow in Figures 1A and 2A), which we refer to as the main TSS, is inside an RNA Pol II binding site found in SkM (Figure 2A, POLR2A large subunit binding) and the unmethylated 5' region of *TBX15* in myoblasts and SkM (Figure 2D). This TSS and the TSS of isoform NM\_0011330677 are predicted to encode the same main *TBX15* protein of 602 amino acids that is observed in normal skin fibroblasts [27]. The cluster of CAGE-determined 5' ends of myoblast, skin fibroblast, and osteoblast transcripts is in promoter chromatin and adjacent to a DNaseI hypersensitive site and overlaps a CGI (Figure 2A–C).

We first tested promoter activity using a CpG-free promoter-less Lucia vector (Vector 1) and inserts from constitutively unmethylated promoter region sequences from the main *TBX15* TSS to 0.7 kb upstream (0/−0.7 insert; Figure 3A,B). Upon transient transfection into C2C12 myoblasts, luciferase reporter activity was almost undetectable from this plasmid and was not significantly greater than the background luciferase activity from Vector 1 (Figure 3C). Although the 0/−0.7 plasmid lacked promoter activity in transfected myoblasts, the endogenous sequences in human myoblasts and in SkM display tissue-specific promoter chromatin and DNaseI hypersensitivity (yellow highlighting, Figure 2B,C). Secondly, we enlarged the test insert by 0.6 kb of constitutively unmethylated sequences using a +0.6/−0.7 insert instead of the 0/−0.7 insert. With this construct, strong promoter activity was seen in the transfected myoblasts (Figure 3C). These findings suggest that endogenous sequences from the TSS to 0.7 kb upstream in myoblasts and in SkM participate in promoter activity through cooperation with neighboring, TSS-downstream unmethylated DNA sequences that lack their own promoter activity. This explanation is supported by the finding of more promoter chromatin and unmethylated DNA immediately downstream of the TSS than upstream of the TSS (Figure 2B,D).

Next, we determined the promoter activity of an enlarged TSS-upstream region from the main TSS to −2.5 kb (0/−2.5; Figure 3B). This plasmid displayed high luciferase activity in myoblast transfectants (Figure 3D) even though its constituent DNA sequences are the promoter-inactive 0/−0.7 sequence and 1.8 kb of *TBX15*-upstream DNA, a sequence that is highly methylated in endogenous human myoblast DNA and has no overlapping promoter chromatin and only low DNaseI hypersensitivity in myoblasts (brown highlighting, Figure 2B,C). Interestingly, quantitation of the WGBS profiles in this region (−0.7 to −2.5 kb, chr1: 119,531,244–119,533,033, hg19) showed that the psoas SkM sample had significantly less methylation ( $p = 2 \times 10^{-27}$ ) than the Myoblast-3 cell strain (average methylation 91% for myoblasts, based on our WGBS data, and 67% for SkM, using Roadmap WGBS data [19]). Accordingly, SkM (psoas) tissue displayed specific acquisition of enhancer or promoter chromatin in this region, which was not seen in myoblasts (Figure 2B).



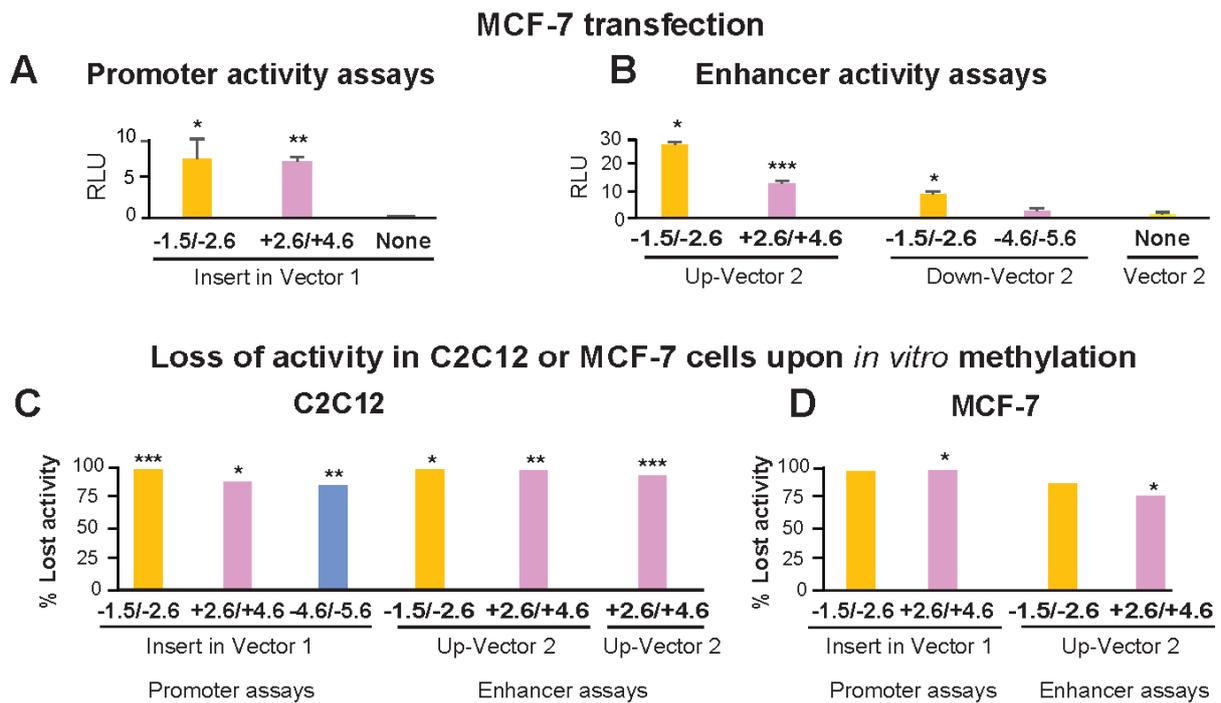
**Figure 3.** Myoblast-hypermethylated DNA sequences adjacent to the unmethylated promoter displayed strong promoter or enhancer activity from reporter constructs transfected into myoblasts. (A) The constructs used for transfection are diagrammed. Vector 1 is the promoter-less pCpGfree-Lucia plasmid; Vector 2 is the analogous plasmid with a modified human EF-1 $\alpha$ -derived minimal promoter (EF1). *TBX15* DNA, the constitutively unmethylated promoter region or DMR sequences adjacent to it. (B) Endogenous location of the *TBX15* inserts used in reporter gene assays (Figure 2). Myob-hyperm DMRs are named according to the beginning and end positions relative to the main TSS of *TBX15*. (C–E) Reference plasmid-normalized luciferase activity from C2C12 myoblasts transiently transfected with plasmids containing inserts shown in Panel B in assays for promoter (Panel C and D) or enhancer (Panel E) activity. In human myoblasts, the inserts were mostly unmethylated (C) or highly and specifically methylated (Panels D and E) *in vivo* but became unmethylated upon cloning. Results from transfections are the averages from at least three independent experiments; error bars for standard error; RLU, relative light units, bioluminescence from the transfected test construct divided by that from the cotransfected reference plasmid; *t*-tests for differences in RLU of the recombinant to the vector-only plasmid:  $p < 0.05$  (\*) or  $p < 0.001$  (\*\*). Meth, methylated; unmeth, unmethylated.

Both a 1.1-kb insert that came only from a Myob-hyperm DMR (−1.5/−2.6; Figure 3B) and a far-upstream 1.0-kb insert from the same DMR (−4.6/5.6) displayed promoter activity in transfected myoblasts relative to the background activity of Vector 1 alone (*t*-test for each construct vs. vector only,  $p < 1 \times 10^{-10}$ ). However, the promoter activity from the −1.5/−2.6 insert was much less than from the larger 0/−2.5 insert (Figure 3D). This again indicates the ability of the 0/−0.7 DNA sequence to cooperate with adjacent *TBX15* promoter region sequences to confer promoter activity, even though the 0/−0.7 sequence had no promoter activity by itself in transfected reporter constructs.

When the −1.5/−2.6 DNA sequence was assayed for enhancer activity by inserting it upstream of a minimal EF1 promoter in Vector 2 (Figure 3A,B), high luciferase activity was seen in myoblast transfectants (Figure 3E). The strongest enhancer activity was observed for a 2-kb Myob-hyperm DMR from *TBX15* intron 1 (+2.6/+4.6, Figure 3E). The genomic location of this DMR is at the downstream border of the constitutively unmethylated promoter region (Figure 2C,D). Usually, inserts tested for enhancer activity in reporter gene constructs are inserted upstream of a minimal promoter in a reporter plasmid, as was done in the experiments demonstrating high enhancer activity for the +2.6/+4.6 region (+2.6/+4.6 Up-Vector2; Figure 3E). When enhancer activity was tested more rigorously in constructs containing the insert downstream of the reporter gene, significantly enhanced reporter gene activity was observed (+2.6/+4.6 Down-Vector2, Figure 3E; *t*-test vs. Vector 2,  $p = 4 \times 10^{-5}$ ) but much less than that from the promoter-upstream insertion. The −1.5/−2.6 and −4.6/−5.6 inserts also gave low, but significant, enhancer activity when tested downstream of the reporter gene (Figure 3E; *t*-test vs. Vector 2,  $p = 1 \times 10^{-4}$  and 0.01, respectively). We previously studied the activity of the *MYOD1* core enhancer by cloning it downstream of the reporter gene in the same CpG-free, minimal EF1 promoter vector and obtained strong activity from comparable transfection assays [28], thus indicating no technical problem with the use of the downstream cloning site in this vector. These results suggest context-dependent upregulation by *TBX15* Myob-hyperm DMR sequences when tested for enhancer activity as unmethylated sequences in reporter constructs.

### 2.3. Much Lower Enhancer and Promoter Activity Was Seen for Transfected *TBX15* TSS-Upstream or Downstream Sequences in Non-Myoblast vs. Myoblast Host Cells

The above-described reporter gene constructs containing *TBX15* promoter-upstream or downstream sequences were also transfected into MCF-7 cells, a breast cancer-derived epithelial cell line. As expected from the widespread use of the MCF-7 cell line for transfection assays, these cells were highly transfectable, which we verified by testing promoter activity of DNA from the broadly expressed *IRS1* promoter region (data not shown). However, when reporter constructs containing Myob-hyperm DMR-derived inserts were used for transfection of MCF-7 cells, the transfectants had much lower reporter gene activity than that of analogous C2C12 myoblast transfectants (Figure 4A,B vs. Figure 3D,E, note different scales). The +2.6/+4.6 Myob-hyperm DMR had 109- and 69-fold lower transcription-promoting activity in transfected MCF-7 cells than in transfected myoblasts in promoter and enhancer test assays, respectively. Comparable decreases for the −1.5/−2.6 Myob-hyperm DMR sequences were 10- to 12-fold lower in MCF-7 than in myoblast transfectants.



**Figure 4.** The promoter or enhancer activity of DMRs from the 5' end of *TBX15* was strongly decreased by targeted *in vitro* methylation or transfection into MCF-7 cells instead of myoblasts. Transfection was done with promoter-assay (A) or enhancer-assay (B) constructs as in Figure 3 but with the MCF-7 breast carcinoma cell line instead of C2C12 myoblasts as the host cells. Note the very different scales for normalized luciferase activity in Figure 3 vs. Figure 4. The normalized luciferase activity lost in C2C12 myoblasts (C) or MCF-7 cells (D) transfected with *in vitro* methylated compared with mock-methylated reporter gene constructs. The difference between methylated and unmethylated DNA was significant at  $p < 0.05$  (\*),  $p < 0.01$  (\*\*) or  $p < 0.001$  (\*\*\*). *In vitro* CpG methylation was targeted only to the insert.

#### 2.4. Transfection of *M.SssI* CpG-methylated *TBX15* DMRs in Reporter Constructs

The reporter gene constructs containing inserts that were highly methylated in human myoblasts were assayed for the effects of CpG methylation on their promoter or enhancer activity. In the above-described experiments, the myoblast CpG methylation is lost upon cloning in *E. coli*. Because the reporter gene vectors that we used had been engineered to contain no CpGs, *M.SssI*-catalyzed methylation (which is CpG-specific) could only occur in the inserts; therefore, there could be no effects on reporter gene expression from methylation of the reporter gene or of the rest of the vector [29].

The  $-1.5/-2.6$  DMR sequence-containing constructs for testing promoter activity or enhancer activity lost 97 to 100% of their activity upon *M.SssI* methylation relative to mock-methylated controls in transfected myoblasts (Figure 4C). Analogous assays for promoter or enhancer activity of  $+2.6/+4.6$  or  $-4.6/-5.6$  DMR inserts in transiently transfected C2C12 cells gave losses of activity of 86–89% (Figure 4C). Methylation *in vitro* to resemble the high extent of CpG methylation of these endogenous sequences in human myoblasts also resulted in the loss of most activity of these constructs when they were transfected into MCF-7 cells (Figure 4D). In summary, the promoter activity seen for all the DMR sequences tested was largely or completely dependent on these sequences not being highly methylated, as they are in human myoblasts.

#### 2.5. Variable *TBX15* Epigenetics in Skeletal Muscle Samples, Myoblast and Skin Fibroblast Cell Strains, Adipocytes, and Cancer Cell Lines

We looked at the epigenetics *in vivo* of the cloned Myob-hyper DNA regions in additional samples of myoblast and skin fibroblast cell strains, SkM tissue, and in cancer

cell lines. We used our previous DNA methylation profiles from RRBS [30] of myoblasts supplemented with other ENCODE RRBS profiles [31]. RRBS only detects the methylation state of 5% or less of CpGs genome-wide [32] but regions of high CpG density, like the 5' end of *TBX15*, are overrepresented (Figure 5A,B) in RRBS profiles. Myoblast 3 and Myoblast 7 displayed similar methylation patterns in RRBS profiles to those seen in WGBS or EM-seq profiles of Myoblasts 1, 3, and 6 (Figures 2D and 5C). In contrast, RRBS revealed that Myoblast 8, SkM 7, and SkM 8 lacked most of the high DNA methylation observed in the other myoblast and SkM samples around the constitutively unmethylated *TBX15* promoter core (Figure 5B,C and Figure 2D). Myotubes, which we obtained by in vitro differentiation of the myoblast cell strains, shared indistinguishable *TBX15* DNA hypermethylation profiles (Figure 5B) although the myotubes had 1.8 times the RNA levels as myoblasts (RNA-seq data, not shown).

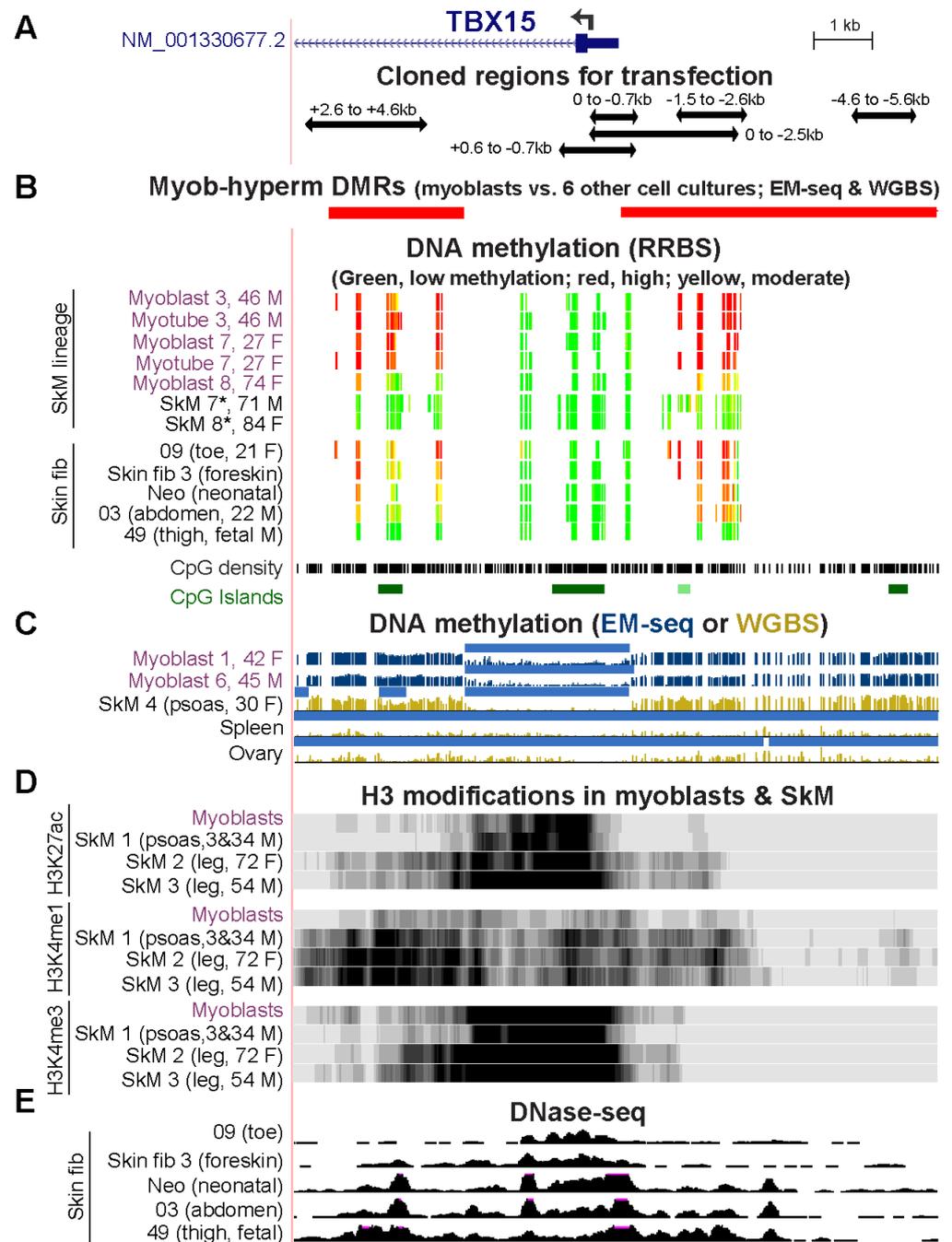
The SkM tissue samples that exhibited myoblast-like DNA hypermethylation around the 5' end of *TBX15* were from young donors (SkM 1, a mixture of 3-y M and 34-y M and SkM 4, 30-y F; Figure 1A and Figure S2) while the samples lacking most of this hypermethylation were derived from elderly individuals (71- to 84-y M or F, unknown type of SkM; Figure 5B). Another complicating factor is that Myoblast 8 was derived from a 74-y patient with inclusion-body myositis so that the disease state might have influenced *TBX15* DNA methylation. In RRBS methylomes, myoblasts and myotubes from one of two young patients with facioscapulohumeral muscular dystrophy displayed reduced amounts of methylation at the Myob-hyperm DMRs while the other sample exhibited the kind of hypermethylation seen in most myoblast and myotube samples (Figure S6). Although the RRBS profiles at the 5' end of *TBX15* reveal the Myob-hyperm DMRs (Figure 5), the WGBS and EM-seq profiles show the specificity of these DMRs for myoblasts and skin fibroblasts more clearly than do RRBS profiles (Figures 2D and 5B).

WGBS methylomes were publicly available for only two SkM samples, both of which were from psoas muscle (SkM 4 and SkM 1) [19]. These methylomes were similar in the vicinity of *TBX15* (Figure 2D vs. Figure 5C). From available histone methylation and chromatin state and histone H3 modification profiles for the psoas SkM 1, leg muscle SkM 2 (72-y F) and leg muscle SkM 3 (54-y M), there was more H3K27ac (indicative of active promoter or enhancer chromatin) in the cloned DMRs in the leg muscle samples than in the psoas or myoblast samples (Figure 5D). The H3K27ac profiles of SkM and myoblasts indicate the presence of a super-enhancer, a large cluster of enhancer and/or promoter chromatin regions [33]. It spanned ~32 and 46 kb from the promoter region through much of the large intron 1 for psoas and leg muscle, respectively (Figure 1 and Figure S6). This super-enhancer was not seen in myoblasts [34].

Although the small number of available SkM samples analyzed for their epigenetics at the 5' end of *TBX15* exhibited no correlation with gender, RNA-seq analysis of 543 male and 260 female gastrocnemius muscle samples (GTEx project [35]) showed that the median TPM (transcripts per million) value for females was 16% higher than that for males (top rectangle, Figure S7). A substantial gender-difference in TPM levels was also noted for *TBX1* in this database, but not for the *TBX15* neighbor *WARS2*. The statistical and biological significance of this finding is uncertain.

Like myoblasts, almost all the primary skin fibroblast cell cultures displayed DNA hypermethylation upstream and downstream of an unmethylated *TBX15* TSS-overlapping region as analyzed by RRBS or WGBS [15,36] (Figures 1A, 5B and S2). These skin fibroblast cultures were derived from various body depots (Figure 5B). Like the myoblast primary cultures, primary cultures of skin fibroblasts (from postnatal leg, temple, scalp, breast, abdomen, back, and neonatal foreskin dermis) express *TBX15* at moderate levels (Figure 1A, top and bottom panels, and data not shown from ENCODE and Roadmap databases). Skin fibroblasts varied in their extent of DNA hypermethylation in the TSS –20 to +9 kb region (Figure 5B) even among biological replicates of foreskin fibroblast primary cultures (Figures 1A and S2B). The one examined skin fibroblast cell strain (Skin fib 49) that was fetal in origin was exceptional in displaying no hypermethylation in this region. Among

the skin fibroblast cultures, there was more open chromatin (DNaseI hypersensitivity) in the regions of less DNA methylation (Figure 5B,E, e.g., toe and foreskin vs. fetal thigh).



**Figure 5.** Some sample-specific differences in epigenetic profiles at the 5' end of *TBX15*. (A) Positions of cloned regions used for transfection (chr1: 119,525,320–119,536,521; only the RefSeq Select isoform is shown). (B) WGBS- and EM-seq determined DMRs, RRBS profiles of DNA methylation, CpG density, and CpG islands. For RRBS tracks, only CpGs with  $\geq 5$  reads are shown. (C) EM-seq and WGBS profiles of DNA methylation with light blue bars denoting low methylated regions as in Figure 1. (D) H3K27ac, H3K4me1, and H3K4me3 signal and (E) DNaseI hypersensitivity for the indicated samples with vertical viewing ranges of 0–20, 0–10, 0–20, and 0–30, respectively. The ages and genders of donors of myoblast, SkM, and skin fibroblast samples are given, where known; Neo, neonatal skin fibroblasts were not further identified [23]; \*, SkM from an unspecified body region.

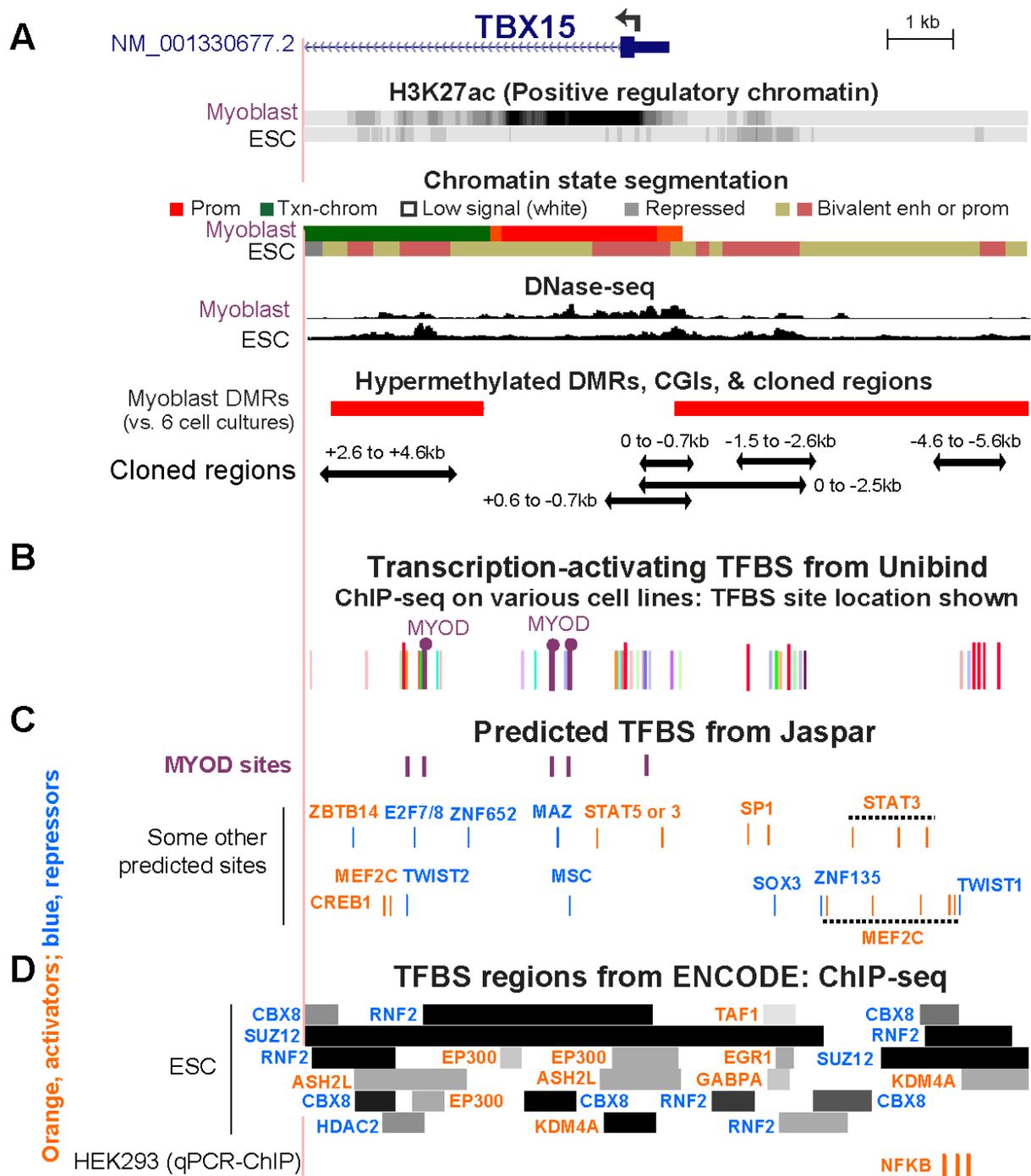
Another cell type with highly varied epigenetics according to the subtype is uncultured adipocytes. As previously reported in a WGBS and RNA-seq study of adipocytes [37], methylation at *TBX15*'s 5' end differed strongly between adipocytes from the same individual derived from either subcutaneous adipose tissue, which highly expresses *TBX15*, or visceral adipose tissue, which shows only low levels of *TBX15* RNA (Figure 1A, bottom). Promoter border DNA hypermethylation was associated with the more highly expressing adipocytes. We found that these subcutaneous adipose hypermethylated DMRs overlapped the Myob-hyperm DMRs adjacent to the constitutively unmethylated *TBX15* promoter (Figure S2B).

Myob-hyperm DMRs were located not only immediately around the unmethylated *TBX15* promoter region but also as part of an extended cluster of DMRs in myoblasts, psoas muscle, and a skin fibroblast cell strain (Figures 1A and S2B). Another prominent SkM-lineage related difference in *TBX15* epigenetics is that there was additional cell type-specific enhancer or weak enhancer chromatin within the gene body or far downstream of its promoter that was associated with *TBX15* tissue expression profiles (Figures 1, S6 and S8). This included enhancer chromatin overlapping SkM-specific DNA hypomethylation in the large gene desert downstream of the gene as far as 0.5 Mb distant from the gene in both psoas and leg muscle (Figure S8A, dotted rectangle).

Many cancer cell lines that did not express *TBX15* differed from normal cell strains in having high levels of DNA methylation throughout the region from TSS  $-20$  to  $+9$  kb, as seen in RRBS profiles [31] (Figure S6). However, other cancer cell lines without this promoter-overlapping DNA hypermethylation still did not express *TBX15* [38], like most non-transformed cell strains. We found RNA-seq [38] and methylome data [31] for one cancer cell line that expressed *TBX15*. Importantly, this cell line, U87 astrocytoma cells, was the only cancer cell line exhibiting *TBX15*-like DNA hypermethylation around an unmethylated TSS region (Figure S6). While HepG2 cells are the other *TBX15*-expressing cancer cell line, they do not initiate transcription at the canonical 5' end of *TBX15* and, instead, use a liver promoter located in exon 6 of the main isoform (Figures S6A and S8A; CAGE and GTEx data not shown [23,35]). Interestingly, like tissues which do not express *TBX15* from any promoter, liver lacks DNA hypermethylation at or around the main promoter region (Figure S6E).

## 2.6. Transcription Factor Binding Sites in the 5' *TBX15* Region

Because there are only a very small number of genome-wide profiles of TF-directed chromatin immunoprecipitation-next gen sequencing (ChIP-seq) from myoblasts (mostly POLR2A, MYOD, MYF5, CTCF), we looked for evidence of TF binding at the above-described cloned *TBX15* regions in various cell types and examined predicted transcription factor binding sites (TFBS) in these regions. From the Unibind human TFBS database, which is based upon TF ChIP-seq profiles combined with TFBS predictions [39], we found binding sites in the cloned regions at the 5' end of *TBX15* for MYOD, one of the four TFs found specifically in the SkM lineage. One of the MYOD sites in myoblasts and rhabdomyosarcoma cells (cancers derived from myogenic progenitor cells) is within the *TBX15*  $+2.6/+4.6$  Myob-hyperm DMR (Figure 6A,B; Table S4). MYOD also occupied two sites in an adjacent unmethylated region in myoblasts, one of which overlapped a large DNaseI-hypersensitive peak in myoblasts. All three MYOD sites also overlapped POLR2A binding subregions in gastrocnemius SkM (Figures 2A and 6B). The two clustered MYOD sites bind to homologous mouse DNA sequences as determined from C2C12 myoblast Myod ChIP-seq profiling [28,40]. Importantly, the mouse dataset gives the relative amount of binding of MYOD. Both C2C12 Myod binding sites in *TBX15* intron 1 were only weak sites with binding scores of 13 and 19 compared to 91–168 for strong Myod sites in enhancer chromatin far upstream of the *Myod1* gene in the same ChIP-seq profile [40].



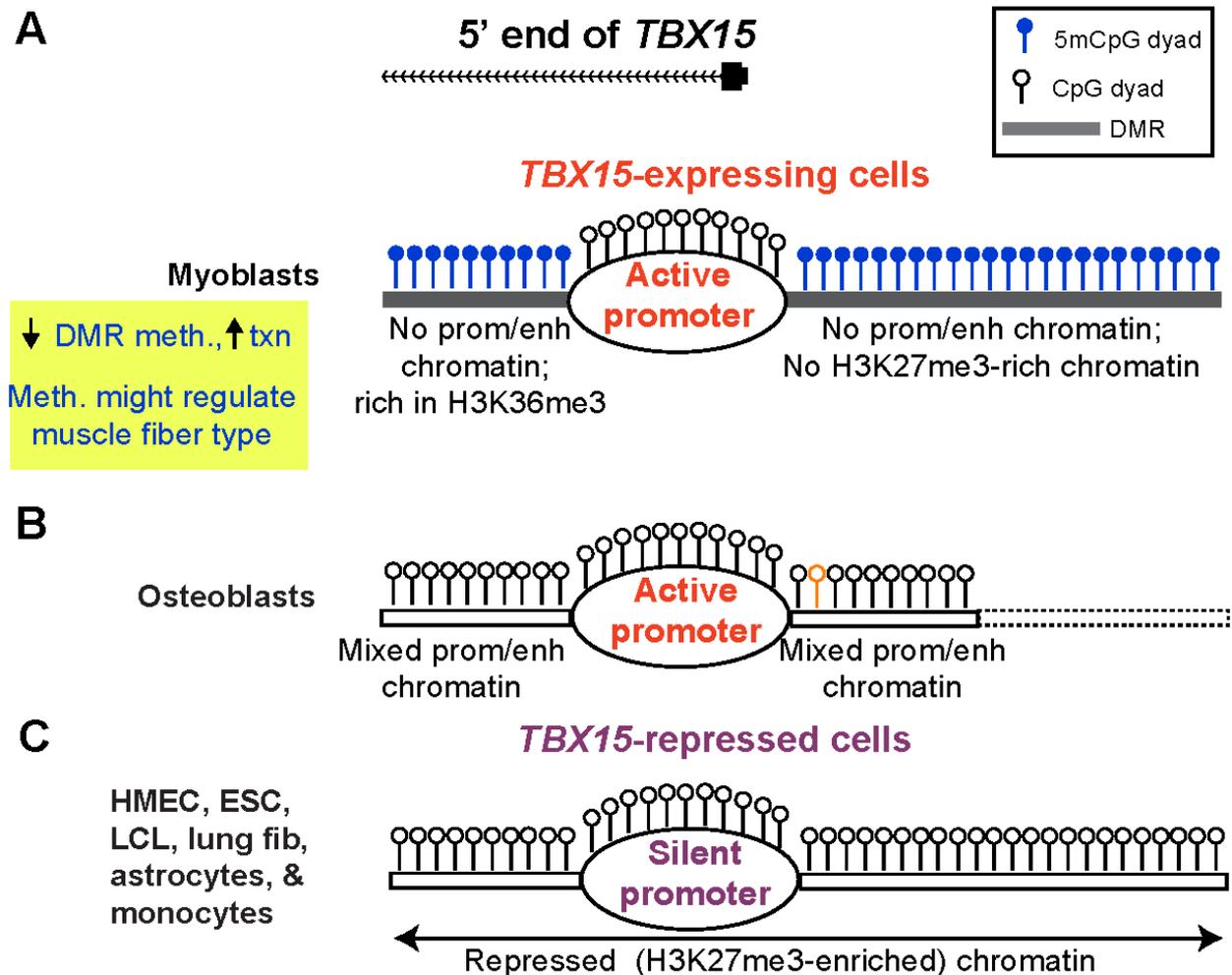
**Figure 6.** Overlap of cloned regions from the 5' end of *TBX15* with transcription factor binding sites. (A) Myoblast and embryonic stem cell (ESC) chromatin epigenetics, Myob-hyperm DMRs, and cloned regions used for reporter gene assays as in Figure 2 (chr1: 119,525,320–119,536,521). (B) Bars, the locations of representative TFBS from combined ChIP-seq data and TFBS predictions (Unibind database; see Table S4 for details). Binding sites found by ChIP-seq for MYOD, a myogenesis-associated TF, in myoblasts or rhabdomyosarcoma cells are indicated by lollipops. (C) Predicted TFBS for MYOD (JASPAR database). (D) ChIP-seq-determined regions of binding of TFBS to ESC (ENCODE 3) and sites of experimentally determined binding of NFKB to HEK293 cells [41] are shown. Binding regions indicated by dark colored segments denote strong TF binding. In Panels C and D, TF labels for transcription-activating TFs are shown in orange and for repressing TFs [8] in blue.

Binding profiles are available for many cell types for the CCCTC-Binding Factor (CTCF), a protein that mediates chromatin looping and is a sequence-specific TF. A strong constitutive CTCF binding site (Figure S2C) was seen immediately upstream of the cluster of Myob/SkM-hyperm DMRs between the 5' end of *TBX15* and the 3' end of its neighbor *WARS2*, a broadly expressed gene (Table S3B). A weak CTCF site that was highly cell type-specific was found towards the 3' end of *TBX15* intron 1 in myoblasts, myotubes, skin fibroblasts, and osteoblasts, all of which express *TBX15* (Figure S2). Weaker CTCF sites were observed in myoblasts in the 0/−0.7 region and in rhabdomyosarcoma cells in the +2.6/+4.6 Myob-hyperm DMR sequences (Table S4). These two sites might facilitate only weak chromatin interactions when the +2.6/+4.6 sequence is highly methylated and stronger interactions when they are not methylated.

Given that myoblasts have only been used to test genome-wide binding of a very small number of TFs, we also looked for predicted TF binding sites in the cloned regions of *TBX15* from available data for other cell types. Many TFBS were predicted to have binding sites in the Myob-hyperm DMR regions and the promoter region in 5' end of *TBX15* (Figure 6C, JASPAR database [42]). There were two predicted additional binding sites for MYOD in the +2.6/+4.6 DMR region, for which in vivo binding was not seen, as well as multiple sites for STAT3 and MEF2C. In the ChIP-seq Unibind database, binding of various other transcription-stimulatory TFs was seen in the TSS-upstream and TSS-downstream Myob-hyperm DMRs in profiled human cancer cell lines, and endothelial cell cultures, and ESC (Figure 6B; Table S4; Unibind database [39,41]), in which *TBX15* is repressed (Figure 1A). (Table S4). The lack of expression of *TBX15* in ESC can be attributed in part to the stronger binding of repressive than of transcription-stimulatory proteins at the 5' end of *TBX15* as seen in the ENCODE database for human ChIP-seq (Figure 6D) and to the related finding of bivalent chromatin in ESC throughout the 5' end of *TBX15* (Figure 6A,D).

### 3. Discussion

Our study provides evidence that the DNA hypermethylation immediately upstream and downstream of the constitutively unmethylated *TBX15* promoter downmodulates transcription of this gene in primary myoblasts (Figure 7A). These promoter-adjacent DNA sequences were ~10 to 100 times more active in reporter gene assays for promoter or enhancer activity when transfected into myoblasts than when transfected into non-myoblast host cells. This strong activity required demethylation because 86–100% of reporter gene expression was lost upon targeting CpG methylation to the DMR sequences (Figure 4C). Unexpectedly, these DMR sequences exhibited much more reporter gene activity when they were inserted upstream of the vector's minimal promoter, as is often done (e.g., [43]), than when placed downstream of reporter gene. In the downstream position they were only 0.9 kb from the minimal promoter, a favorable distance for enhancer tests [44]. It is likely these promoter-adjacent DMRs (Myob-hyperm DMRs) are part of an extended promoter or context-sensitive enhancer in certain cell/tissue types when unmethylated (Figure 7B). This was demonstrated not only by their demethylation-dependent promoter/enhancer activity but also by in vivo correlations between less methylation and more overlap with promoter, enhancer, or open chromatin in *TBX15*-expressing osteoblasts and skin fibroblasts. Because the promoter-adjacent Myob-hyperm DMRs, when unmethylated, exhibited strong promoter/enhancer activity in transfected myoblasts, it is very unlikely that they are transcription repressors in vivo. Therefore, our findings argue against the hypothesis that the role of the hypermethylation of these DMRs is to turn off, in cis, an overlapping repressor.



**Figure 7.** A model for the role of myoblast-hypermethylated DMRs adjacent to the *TBX15* promoter in down-modulating expression of the transcriptionally active gene. (A) The high level of methylation of promoter-adjacent DMRs in myoblasts is proposed to prevent their latent enhancer and/or extended promoter activity (enh/prom chromatin) in vivo but to allow core promoter activity and downstream enhancer activity (Figures 1 and S8). Certain types of SkM with less methylation at these DMRs may have higher *TBX15* activity due to turning on these promoter-adjacent upregulatory elements; meth., methylation; txn, transcription. (B) Some cell types that specifically express *TBX15*, like osteoblasts, have low or no methylation coupled with strong enh/prom chromatin adjacent to the promoter region (Figure S6D). Dotted lines, osteoblast methylome data are from RRBS and so have limited coverage; orange lollipop, a SNP (rs1106529) that is strongly associated with bone mineral density and is located in extended promoter chromatin in osteoblasts overlapping the upstream Myob-hyperm DMR [17]. (C) Repressive histone modification (H3K27me3) is seen in many cell and tissue types that have little or no DNA methylation at the DMRs. Such cells with silent *TBX15* alleles do not need DNA hypermethylation-linked fine-tuning of expression. The top of the figure shows the main 5' end of the gene in myoblasts, SkM, osteoblasts, and skin fibroblasts, which differs from the 5' ends of the two RefSeq *TBX15* structures (Figure 2).

Cells not transcribing *TBX15* would, by definition, have no need of fine-tuning *TBX15* expression (Figure 7C). The extensive cross-talk between DNA methylation and chromatin epigenetics [45] can help explain the different DNA methylation profiles near the *TBX15* promoter in non-expressing cells vs. in myoblasts. Two caveats are that genome-wide epigenetic associations can miss important context-dependent exceptions and lesser-studied histone modifications can influence changes in DNA methylation [46]. Cell cultures with

a silent *TBX15* gene, such as HMEC, ESC, and lung fibroblasts, had H3K27me<sub>3</sub> (often as bivalent chromatin) at the constitutively unmethylated core promoter and surrounding regions (Figure 7C). High levels of this repressive histone modification suffice to silence promoters. H3K27me<sub>3</sub> enrichment displays a genome-wide association with low DNA methylation levels, but this anti-correlation is much weaker than that of H3K4 methylation and low DNA methylation [45,47]. As expected, the H3K4 methylation-rich chromatin present at the upstream and downstream DMRs adjacent to the *TBX15* promoter in osteoblasts (Figure 7B) is negatively associated with local DNA methylation. However, in this case, the unmethylated DNA sequences can help upregulate *TBX15* transcription as part of enhancers or an extended promoter [47]. The H3K36 trimethylation at the Myob-hyperm DMR immediately downstream of the promoter in myoblasts (Figure 7A) weakly favors DNA methylation [47], as might the loss of H3K27me<sub>3</sub> from the promoter-upstream Myob-hyperm DMR (Figure 7A). We propose that the use of DNA hypermethylation in myoblasts to suppress enhancer-like chromatin adjacent to the *TBX15* promoter in myoblasts allows moderate downmodulation of *TBX15* expression without causing excessive transcription repression as might result from deposition of H3K27me<sub>3</sub> in the promoter-upstream or downstream regions.

Consistent with the much higher promoter/enhancer activity of these *TBX15* DMRs in transfected myoblasts than in transfected non-myoblasts, one of these DMRs has a binding site in human myoblasts for MYOD, a central SkM lineage-specific TF, as well as a predicted site for MEF2C (Figure 6), a TF involved in various differentiation processes including myogenesis and repair of SkM [48]. Similarly, in the other *TBX15* DMR used for transfection, there were five predicted MEF2C sites interspersed with three predicted sites for STAT3, a signal transducer and transcription-activating TF involved in muscle satellite cell expansion and SkM repair, among other pathways [49]. We propose that the Myob-hyperm DMRs at the borders of the *TBX15* promoter, when unmethylated, upregulate transcription *in vivo* in certain SkM-lineage cell types. The more distal upstream Myob-hyperm DMRs may be suppressing activity of potential promoters overlapping CpG islands (Figure 1) [50].

We found less than a two-fold increase in *TBX15* RNA in human myotubes relative to their myoblast cell precursors (Table S3) and similar DMRs at the gene in both cell types. Lee et al. reported a twelve-fold increase in *Tbx15* RNA when a murine myoblast cell line (C2C12) was induced to differentiate to myotubes [10]. Differences in these results might be due to differences in the protocols used for induction of myoblasts to form myotubes. Differentiation *in vitro* of mononuclear myoblasts to elongated, broadened multinucleated myotubes involves removal of fetal bovine serum from the culture medium. The larger increase in *TBX15* RNA levels upon differentiation in the study of Lee et al. than in ours might be due to our less severe differentiation protocol in which fetal bovine serum is replaced with 1.5% horse serum (HS) for only one day followed by 4 d with 15% HS rather than the more standard procedure of Lee et al. that uses 2% HS for 4 d [10].

The enhancer and promoter chromatin at the 5' end of *TBX15* in osteoblasts (Figure 7B) is part of a super-enhancer that extends for >10 kb downstream of the TSS [14,17]. Super-enhancers strongly upregulate gene expression and are seen most frequently at developmental genes [33]. We previously reported that this super-enhancer contains a SNP (rs1106529; orange lollipop, Figure 7B and Figure S6C) which is strongly associated with bone mineral density (and obesity-risk traits) [17]. This SNP is in the Myob-hyperm DMR immediately upstream of the promoter region and in DNA sequences that are unmethylated in osteoblasts (Figure S6D). It is not in linkage disequilibrium ( $r^2 > 0.2$ ) with any other SNP, which would have complicated the determination of its biological importance [17]. Therefore, we propose that osteoblasts are an example of a cell type that benefits from DNA methylation-sensitive upregulation of *TBX15* by enhancer/promoter chromatin bordering the active promoter of *TBX15* and overlapping a Myob-hyperm DMR.

Differences in epigenomic profiles around the *TBX15* promoter in skin fibroblast cell strains from different body sites suggest differential regulation of this gene in skin. The need for fine-tuning expression of *TBX15* in skin is evidenced by dynamic position-

dependent differences in *Tbx15* expression in dermal cells during mouse embryogenesis that contribute to mouse coat patterning [51]. In a genome-wide study of human adipocytes, Bradford et al. [37] reported several subcutaneous vs. visceral adipocyte DMRs that they described as close to the 5' end of *TBX15* among the 2108 DMRs that they identified. These *TBX15* DMRs exhibited a positive correlation between hypermethylation and preferential expression in subcutaneous adipocytes compared with matched visceral adipocytes. We localized their 5' *TBX15* adipocyte DMRs to the borders of the *TBX15* promoter region and found that they overlap Myob-hyperm DMRs (Figure S2). Bradford and coworkers suggested that the function of this promoter-bordering DNA hypermethylation is to prevent spreading of repressive chromatin into the active promoter in subcutaneous adipocytes. However, our results favor the hypothesis that this *TBX15* promoter-adjacent DNA hypermethylation in both subcutaneous adipocytes and myoblasts prevents the formation of promoter-adjacent enhancer chromatin that would lead to high levels of transcription of *TBX15* in these cells.

In another study of *TBX15* in adipose cells, Ejarque et al. [52] examined the methylomes of human adipose-derived stromal/stem cells from subcutaneous adipose tissue derived from either obese or lean middle-aged females. Cells from lean individuals had approximately 2.5-fold more *TBX15* RNA than the analogous cells from the same body depot in obese individuals. They showed that upregulation of *TBX15* correlated with less methylation in the regions we identified as promoter-adjacent Myob-hyperm DMRs, a finding that is consistent with our model (Figure 7A,B). However, in that study DNA methylation changes were more moderate and seem to be behaving like a continuously adjustable, rather than an on/off, regulator of enhancer/extended promoter DNA sequences.

*TBX15* can act as both a transcription repressor and activator [53,54]. It can downregulate mitochondrial oxidation rates in conjunction with Ampk phosphorylation in SkM and myoblasts, and high *Tbx15* expression in a subfraction of murine subcutaneous adipocytes correlates with lower levels of oxidative metabolism markers [6,10]. In C2C12 myoblasts and murine embryos, *Tbx15* was implicated in indirectly upregulating *Igf2*, which controls embryonic myogenesis [10]. *Tbx15* induced proliferation of mesenchymal precursor cells and prehypertrophic chondrocytes, but only transiently during embryogenesis, as concluded from studies of *Tbx15* null mutant vs. normal mouse embryos [2]. In cancer cell lines, human *TBX15* was shown to have an anti-apoptotic function that could be partly mediated by its suppression of transcription of several apoptosis-associated *BCL2* family genes [4,55]. Therefore, while there is much more to be learned about *TBX15*'s cell type-specific regulation of transcription, clearly it can play pivotal roles in differentiation, homeostasis, and changes in cell physiology, which may necessitate careful modulation of its transcription levels partly by promoter-adjacent DNA hypermethylation.

From an examination of human RRBS methylomes, we previously reported that T-box genes are overrepresented among the genes with myoblast DNA hypermethylation [30]. In the current, much more extensive WGBS/EM-seq study, we found that *TBX18*, *TBX2*, *TBX3*, and *TBX1*, like *TBX15*, exhibited Myob-hyperm DMRs bordering the active promoter. In contrast *TBX4*, *TBX5*, and *TBX20* displayed this myoblast DNA hypermethylation bordering on and encroaching into their silent H3K27me3-enriched promoter region (Figure 1 and Figures S3–S5). The high density of CpGs around or in the promoter regions of these eight T-box genes is unlike that of most tissue-specific genes but is found at a higher frequency in genes encoding tissue-specific TFs [56]. Although we saw evidence of frequent silencing of these T-box genes in cancer cell lines by both polycomb-repressed chromatin (H3K27me3) and DNA hypermethylation (Figure S6 and data not shown [23]), one cancer cell line, U87 astrocytoma cells, expressed *TBX15* at moderate levels [38]. In contrast, astrocytes have negligible expression of this gene. Importantly, U87 cells were the only cancer cell line in the RRBS database at the UCSC Genome Browser [23,31] displaying myoblast-like hypermethylated DMRs around an unmethylated promoter region (Figure S6). These findings suggest the acquisition of expression-linked promoter-border DNA hypermethylation in

certain cancers, which might contribute to carcinogenesis through antiapoptotic effects of *TBX15* upregulation [4].

Similar to myoblasts, two SkM muscle samples for which there are available WGBS profiles, displayed *TBX15* DMRs similar to the Myob-hyperm DMRs although with less extensive methylation (Figures 1 and 5). These SkM samples were both psoas muscle. However, two other SkM samples of unknown muscle type were largely unmethylated around the *TBX15* promoter (RRBS profiles, Figure 5). Analyses of epigenomics and transcriptomics in SkM are complicated by many factors including cell heterogeneity, muscle fiber type composition differences and other SkM subtype differences and can be influenced by exercise, muscle disuse, aging, gender, and diet [57–63]. In mice, scRNA-seq indicated that only 68% of the nuclei in soleus and quadriceps are the myocyte nuclei [64]. Varying proportions of non-myocyte cells can be found in SkM tissues from different parts of the body [60]. Despite these complicating factors, murine *Tbx15* was found to be predominantly expressed in SkM types enriched in fast-twitch glycolytic muscle fibers rather than in slow-twitch myofibers or fast oxidative myofibers [10], and its expression has been used as a marker of fast glycolytic muscle fibers [65]. Different muscle types are mixtures of slow and fast myofibers, including some hybrid slow/fast myofibers. Lee et al. [10] reported and Terry et al. [60] confirmed that mice had about twice as high *Tbx15* RNA levels in gastrocnemius, tibialis anterior, and extensor digitorum longus muscle (all of which are enriched in glycolytic myofibers) than in soleus muscle, which has a higher percentage of oxidative myofibers.

Human psoas SkM is mostly a body-support muscle. Accordingly, it has a low content of fast glycolytic myofibers [66]. Psoas SkM exhibited less enhancer/extended-promoter chromatin at the 5' end of *TBX15* than did two leg muscle samples for which chromatin epigenomic profiles, but not methylomes, were available (Figure 5D). Although the donors for the leg muscle samples (54 y M and 72 y F) were much older than those for the examined psoas sample (mixed 3-y M and 34-y M), we favor the explanation that the observed chromatin epigenetic differences between the psoas and leg samples at *TBX15*'s 5' end resulted from differences in myofiber composition rather than age differences. A meta-analysis of 908 SkM samples did not identify *TBX15* among the genes with age-associated DMRs [61]. Furthermore, fast-twitch glycolytic fibers in humans, which in mice are associated with high *Tbx15* expression [10], have been found to decrease, rather than increase, with age [58]. We propose that the absence of *TBX15* promoter-adjacent DNA hypermethylation in two RRBS-analyzed SkM samples from unspecified parts of the body (Figure 5C) is due to their derivation from muscle types with especially high expression of *TBX15*. Because myofiber type plays major and complex roles in muscle performance, muscle formation, muscle repair, and sarcopenia [67,68], the epigenetic regulation of postnatal *TBX15* expression is likely to be important for normal muscle function and maintenance.

*TBX15* exhibits higher expression in both postnatal SkM (gastrocnemius, a lower leg muscle) and fetal human SkM than in other examined tissues and is expressed in both SkM myocytes and muscle satellite cells (Tables S1 and S2). Nonetheless, the major phenotype of homozygous loss of function of *TBX15* in humans (Cousin syndrome) or in mouse knock-out models is major deformities in the skeletal system reflecting a critical role for its encoded protein in embryonic skeletal bone formation [2,69]. In a mouse strain with homozygous loss-of-function of *Tbx15*, a few changes in the musculature were noted [3,70]. The major limb malformations in humans or mice associated with loss of *TBX15* activity could largely mask muscular defects given the interrelations of SkM and bone functionality. Interestingly, a T-box gene called *Tbx15/18/22* in *Ciona intestinalis*, an aquatic invertebrate, is essential for normal transcription of many muscle structural genes [71]. Of all the human T-box genes, only *TBX15* has strong specificity for fetal human SkM cells (Table S2).

The biological importance of epigenetic fine-tuning of *TBX15* transcription may be related to the phenotypic effects of partial loss of T-box TFs, in general. Deleterious mutations in all but one (*TBX18*) of the 17 T-box family genes cause disease phenotypes or

prenatal lethality when homozygous in humans, and mutations in 10 of the genes also give a phenotype when heterozygous [3]. Such heterozygosity usually decreases the wild-type levels of the corresponding gene product approximately two-fold. Therefore, this type of heterozygote phenotype is evidence for the importance of close regulation of expression of T-box genes. Although *TBX15* heterozygosity for loss-of-function mutations in humans and mice has not been reported to confer overt skeletal abnormalities [3,9], careful examination of heterozygous knock-out mice revealed SkM [10] and facial phenotypic differences from wild-type mice [72]. Lee et al. [10] found that these heterozygotes, which had a ~40% decrease in *TBX15* mRNA and protein, had a significant (~10%) decrease in muscle mass in tibialis anterior, a type of muscle consisting predominantly of glycolytic fibers, that was not seen in soleus muscle, a mostly oxidative type of muscle. The analogous homozygotes had a yet larger decrease in muscle mass (~25%) due to changes in muscles enriched in glycolytic myofibers. Our results suggest that large increases in DNA methylation can repress or maintain repression of enhancer-like activity surrounding the unmethylated *TBX15* promoter in muscle fibers. Furthermore, graded changes in this methylation might give partial enhancer-like activity to these promoter-adjacent regions. In both cases, promoter border hypermethylation in *TBX15* is ideally suited for the dynamic regulation of muscle physiology.

#### 4. Materials and Methods

##### 4.1. Preparation of DNA Constructs, Transfection, and In Vitro DNA Methylation

Reporter gene constructs were prepared by overlap extension PCR (Table S5) or by using the Gibson assembly kit (NEBuilder HiFi Assembly, New England Biolabs, Ipswich, MA, USA) as previously described [28]. The vectors (InvivoGen, San Diego, CA, USA/Invitrogen, Waltham, MA, USA) were pCpGfree-Lucia or pCpGfree-promoter-Lucia (Vectors 1 and 2, with or without a human EF-1 $\alpha$ -derived minimal promoter, respectively, Figure 3). These vectors have a Lucia luciferase reporter gene and no CpGs. The inserts for cloning were obtained by PCR on mixed human brain and placenta DNAs using the primers shown in Table S5. Recombinant plasmid structure was checked by partial DNA sequencing and restriction site analysis. Transfection into C2C12 or MCF-7 cells utilized a lipid-based reagent (Fast-forward protocol, Effectene reagent, Qiagen, Hilden, Germany). As a reference for transfection efficiency, pCMV-CLuc 2 (New England Biolabs) encoding the Cypridina luciferase was co-transfected with the test construct. About 48 h after the transfection, Lucia and Cypridina luciferase activity was quantified by bioluminescence from aliquots of the cell supernatant (BioLux Cypridina Luciferase assay kit, New England Biolabs; Quanti-Luc, InvivoGen). Reference plasmid-normalized luciferase activity was from the average of three independent transfections. Methylation of the plasmids was targeted just to the *TBX15* inserts, which were the only CpG-containing sequences, by incubating the DNA construct (1  $\mu$ g) with 4 units of SssI methylase and 160  $\mu$ M S-adenosylmethionine (New England Biolabs) for 4 h at 37 °C or mock-methylating by incubating in the absence of S-adenosylmethionine. A similar plasmid construct that contained three BstUI CGCG sites was methylated as above and shown thereafter to be fully resistant to BstUI cleavage.

##### 4.2. EM-Seq and WGBS on Myoblast DNA and Determination of DMRs and LMRs

The myoblasts used for DNA isolation were non-transformed cultures derived from quadriceps biopsies of control individuals [73]; Myoblast 1, 42-y F; Myoblast 3, 46-y M; Myoblast 6, 45-y M. Although primary myoblasts, especially from commercial sources, are often contaminated with large numbers of fibroblast-like cells, which can provide misleading results in DNA methylation analyses, we demonstrated that all of our batches of myoblasts contained >90% desmin-positive cells. WGBS of myoblast cell line Myoblast 3 [30] was performed by standard methods [74]. Methylation profiling by EM-seq of Myoblast 3 and two additional myoblast cell strains (Myoblasts 1 and 6) was done as previously described [75]. This involved the enzymatic oxidation of 5mC (TET2) to

5hmC residues and then to 5-carboxylcytosine (5caC) residues followed by glucosylation (T4-phage  $\beta$ -glucosyltransferase) of any remaining 5hmC, conversion of C residues to U residues (APOBEC3A), and PCR [24]. In brief, 0.2  $\mu$ g of DNA was used for EM-seq library preparation using the NEBNext<sup>®</sup> Enzymatic Methyl-seq kit for Myoblasts 1, 3 and 6 in duplicate. The resultant libraries were cleaned (NEBNext<sup>®</sup> sample selection beads) and duplicates were pooled. The final library pool was diluted to 1.5 nM for NovaSeq (illumina) sequencing.

For determining myoblast DMRs, the EM-seq data for the three examined myoblast cell strains were compared to WGBS profiles of foreskin fibroblasts (Skin Fib 2) [76], adipose-derived mesenchymal stem cells induced to differentiate to adipocytes [36], prostate epithelial cells [77], human mammary epithelial cells (HMEC [78]), prenatal lung fibroblasts (IMR90) and ESC [79]; the last three were cell lines established from non-malignant cells and the others are cell strains. We had previously determined SkM (psoas) DMRs by comparing WGBS profiles [19] from psoas to those of heart (left ventricle), aorta, monocytes, lung, and subcutaneous adipose tissue [22,80,81]. To verify that differences were not associated with technical effects, EM-seq myoblast methylation profiles were initially compared to a WGBS methylation profile from one of the three cell strains. While there was significant biological variation among the three cell strains, results indicated that the proportion of differentially methylated sites between the EM-seq and WGBS profiles for the same cell line was consistent with random variation (data not shown). DMRs between the three EM-seq profiles and the group of five cell cultures were determined using a two-phase process, with significantly differentially methylated sites identified via generalized linear models and aggregated into DMRs based on the Uniform Product distribution for p-values as previously described [81]. Low methylated regions (LMRs) shown in the figures refer to regions with significantly lower DNA methylation than in the rest of the same genome as determined using the method of Song et al. [21].

#### 4.3. Bioinformatics

Most of the bioinformatic profiles were from the UCSC Genome Browser using the hg19 (mainly) or hg38 reference genomes and are shown with hg19 coordinates in the figures [23]. RefSeq Curated gene isoforms are shown unless otherwise specified. WGBS profiles of genome-wide CpG methylation of tissues and cell cultures other than myoblasts (see above) were used for most DNA methylation comparisons and were complemented with RRBS data for *TBX15* methylation, including some previously described cell or tissue profiles [14,31]. Unless otherwise stated, the SkM sample for WGBS was a mixture of psoas DNA from a 3-y male and a 34-y male. Human transcription data was from the following UCSC Browser tracks or hubs: cultured cells (strand-specific RNA-seq, ENCODE/Cold Spring Harbor Lab, or non-strand specific RNA-seq, Transcription Levels Assayed by RNA-seq on 9 cell lines/ENCODE [15]); GTEx (medium TPM from RNA-seq from hundreds of samples for each tissue [35]; Table S1), nine tissues (Burge Lab RNA-seq [82]), Fetal Gene Atlas binned by organ or cell type from Cao et al. (scRNA-seq on samples from ~72 to 129 days post-conception fetuses [83]; Table S2), Muscle RNA binned by biosample from De Micheli et al. (scRNA-seq on postnatal muscle tissues [57]; Table S2); RNA-seq analysis of multiple types of SkM shown in the hg38 reference genome [19,23], and 5' Cap Analysis of Gene Expression (CAGE; RIKEN Omics Science Center [84]). The quantitation of cell culture-derived RNA-seq data for poly(A)<sup>+</sup> RNA was previously described [30]. In addition, for extensive scRNA-seq, the poly(A)<sup>+</sup> RNA Human Protein Atlas (scRNA-seq on tissues [38]; Table S1; Figure S1) was employed. For comparisons of RNA levels in myoblasts and myotubes, we used previously generated RNA-seq data for poly(A)<sup>+</sup> RNA from myoblast cell strains from our lab [85]. CGI identification followed the definitions at the UCSC Genome Browser with islands between only 200 and 300 bp identified by their light green color in figures.

The 18-state chromatin state segmentation analysis (chromHMM, AuxilliaryHMM, Roadmap Epigenomics [19]) was used for determination of chromatin states. The color

coding in figures of gene regions is as follows: red, promoter or mixed promoter/enhancer chromatin (States 1–4); light or dark green, H3K36me3-enriched chromatin (States 5 and 6); orange or yellow-green, enhancer chromatin (States 7–10); light yellow, weak enhancer chromatin (State 11); blue-green, H3K9me3- and H3K36me3-enriched chromatin and ZNF-gene associated chromatin (State 12); light blue, H3K9me3-associated (State 13); reddish brown or gray-green, bivalent poised promoter and enhancer (States 14 and 15, respectively); light or dark gray, H3K27me3-associated repressed chromatin (States 16 and 17); and white, low signal for H3K27 acetylation or methylation, H3K4me1, H3K4me3, H3K36me3, and H3K9me3 (State 18). Also available at the UCSC Genome Browser [23] are DNase-seq profiles of various skin fibroblast cell strains (Roadmap Epigenetics Project for Figure 2 and ENCODE project/University of Washington for Figure 5), CTCF binding profiles for cell cultures (ENCODE), predicted TFBS from the JASPAR database [42], ChIP-seq profiles combined with TFBS prediction from UniBind [39], and ChIP-seq profiles from ENCODE 3 Transcription Factor ChIP-seq Clusters. The presence of super-enhancers was determined using the SEDb tool [34] although the sizes of the super-enhancers were determined by visual examination of H3K27ac and H3K4me1 tracks (vertical viewing range, 0–10).

## 5. Conclusions

DNA methylation near the promoter region is usually associated with silenced genes. In contrast, *TBX15* had strong hypermethylation bordering its unmethylated CGI-containing promoter in myoblasts and psoas skeletal muscle that correlates with its preferential expression in skeletal muscle. Results from our reporter gene assays and bioinformatic comparisons of many cell and tissue types indicate that DNA hypermethylation at the upstream and downstream borders of the *TBX15* promoter helps to prevent overexpression of *TBX15*. Previous functions suggested for this *TBX15* promoter-adjacent hypermethylation were silencing of putative repressor elements, protecting against the expansion of nearby repressed chromatin into the promoter, or directing the use of alternate promoters. Our results suggest that the loss of this methylation or a decrease in the extent of methylation in vivo is associated with higher expression of *TBX15* by removing repression from the silenced enhancer-like DNA sequence elements in these hypermethylated DMRs. Our results also present a cautionary tale about how cancer-related DNA hypermethylation at a CpG island near the 5' end of a gene need not correlate with silencing of the gene. While high DNA methylation levels within a CpG-rich promoter are well known to strongly repress transcription, such methylation in promoter-adjacent regions, may only downmodulate, rather than silence expression and may not be found in cell types in which the gene is otherwise silenced by repressive chromatin. These findings reinforce the importance of understanding the genetic and chromatin context of regions being examined for DNA hypermethylation associated with differentiation, physiological changes, or disease.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/epigenomes6040043/s1>, Figure S1: *TBX15* is preferentially expressed in myocytes, fibroblasts, and smooth muscle cells in skeletal muscle (SkM) tissue. Figure S2: Whole-genome methylome profiles of *TBX15* for the three myoblast cell strains and six non-myogenic cell cultures used for determining myoblast DMRs: Similarities of myoblast vs. non-myoblast DMRs (this study) to subcutaneous vs. visceral adipocyte DMRs from Bradford et al. 2019. Figure S3: Myoblast-hypermethylated DMRs adjacent to promoters in *TBX1* and *TBX2*, which are preferentially expressed in myoblasts. Figure S4: Myoblast-hypermethylated DMRs surround the promoter in *TBX3*, which is preferentially expressed in myoblasts but overlap the promoter of *TBX20*, which is repressed in myoblasts. Figure S5: Myoblast-hypermethylated DMRs overlap promoters of *TBX4* and *TBX5*, which are repressed in myoblasts. Figure S6: The intergenic region between *TBX15* (cell-type specific expression) and *WARS2* (broad expression) displays multiple myoblast-hypermethylated DMRs but less osteoblast hypermethylation. Figure S7: *TBX15* and *TBX1*, but not *TBX15*'s neighbor *WARS2*, display more expression in female skeletal muscle than in male skeletal muscle. Figure S8: *TBX15* is the only gene preferentially expressed in the skeletal muscle lineage in its gene neighborhood but there is a very far downstream enhancer chromatin region overlapping a hypomethylated

DMR in skeletal muscle. Table S1: RNA-seq profiles showing that *TBX15* is the T-box gene with the strongest preference for expression in both skeletal muscle myocytes and skeletal muscle tissues. Table S2: Expression of *TBX15* among human fetal organs and cells and postnatal skeletal muscle from different anatomical sites. Table S3: Cell type-specificity of RNA levels from T-box family genes among different types of human primary cell cultures. Table S4: Transcription factor binding sites demonstrated for various cell types in the 5' end of *TBX15*. Table S5. Oligonucleotides used for fusion cloning of 5' *TBX15* sequences into CpG-free Luciferase reporter vectors.

**Author Contributions:** K.C.E. and M.E. conceived the study, made the reporter constructs, did the transfection experiments and wrote the manuscript. M.L. determined the DMRs for myoblasts and C.B. determined the LMRs from myoblast methylomes generated by S.S. and P.O.E., who were under the direction of S.P. The myoblast DNA was previously isolated from primary cells grown and characterized by immunocytochemistry under the direction of M.E. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded in part by grants from the National Institutes of Health (NS04885) and the Louisiana Cancer Center.

**Data Availability Statement:** Data is contained within the article and supplementary material.

**Acknowledgments:** This research was also supported in part using high performance computing (HPC) resources and services provided by Information Technology at Tulane University, New Orleans, LA.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. L'Honore, A.; Drouin, J.; Buckingham, M.; Montarras, D. Pitx2 and Pitx3 transcription factors: Two key regulators of the redox state in adult skeletal muscle stem cells and muscle regeneration. *Free Radic. Biol. Med.* **2014**, *75*, S37. [[CrossRef](#)] [[PubMed](#)]
2. Singh, M.K.; Petry, M.; Haenig, B.; Lescher, B.; Leitges, M.; Kispert, A. The T-box transcription factor Tbx15 is required for skeletal development. *Mech. Dev.* **2005**, *122*, 131–144. [[CrossRef](#)] [[PubMed](#)]
3. Papaioannou, V.E. The T-box gene family: Emerging roles in development, stem cells and cancer. *Development* **2014**, *141*, 3819–3833. [[CrossRef](#)] [[PubMed](#)]
4. Arribas, J.; Gimenez, E.; Marcos, R.; Velazquez, A. Novel antiapoptotic effect of TBX15: Overexpression of TBX15 reduces apoptosis in cancer cells. *Apoptosis* **2015**, *20*, 1338–1346. [[CrossRef](#)] [[PubMed](#)]
5. Pan, D.Z.; Miao, Z.; Comencho, C.; Rajkumar, S.; Koka, A.; Lee, S.H.T.; Alvarez, M.; Kaminska, D.; Ko, A.; Sinsheimer, J.S.; et al. Identification of TBX15 as an adipose master trans regulator of abdominal obesity genes. *Genome Med.* **2021**, *13*, 123. [[CrossRef](#)]
6. Lee, K.Y.; Sharma, R.; Gase, G.; Ussar, S.; Li, Y.; Welch, L.; Berryman, D.E.; Kispert, A.; Bluher, M.; Kahn, C.R. Tbx15 defines a glycolytic subpopulation and white adipocyte heterogeneity. *Diabetes* **2017**, *66*, 2822–2829. [[CrossRef](#)] [[PubMed](#)]
7. de Araujo, T.K.; Secolin, R.; Félix, T.M.; de Souza, L.T.; Fontes, M.; Monlleó, I.L.; de Souza, J.; Fett-Conte, A.C.; Ribeiro, E.M.; Xavier, A.C.; et al. A multicentric association study between 39 genes and nonsyndromic cleft lip and palate in a Brazilian population. *J. Craniomaxillofac. Surg.* **2016**, *44*, 16–20. [[CrossRef](#)] [[PubMed](#)]
8. Stelzer, G.; Rosen, N.; Plaschkes, I.; Zimmerman, S.; Twik, M.; Fishilevich, S.; Stein, T.I.; Nudel, R.; Lieder, I.; Mazon, Y.; et al. The GeneCards Suite: From gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinform.* **2016**, *54*, 1–30. [[CrossRef](#)]
9. Candille, S.I.; Van Raamsdonk, C.D.; Chen, C.; Kuijper, S.; Chen-Tsai, Y.; Russ, A.; Meijlink, F.; Barsh, G.S. Dorsoventral patterning of the mouse coat by Tbx15. *PLoS Biol.* **2004**, *2*, e3. [[CrossRef](#)]
10. Lee, K.Y.; Singh, M.K.; Ussar, S.; Wetzel, P.; Hirshman, M.F.; Goodyear, L.J.; Kispert, A.; Kahn, C.R. Tbx15 controls skeletal muscle fibre-type determination and muscle metabolism. *Nat. Commun.* **2015**, *6*, 8054. [[CrossRef](#)]
11. Osborn, D.P.S.; Li, K.; Cutty, S.J.; Nelson, A.C.; Wardle, F.C.; Hinits, Y.; Hughes, S.M. Fgf-driven Tbx protein activities directly induce myf5 and myod to initiate zebrafish myogenesis. *Development* **2020**, *147*, dev184689. [[CrossRef](#)] [[PubMed](#)]
12. Hubert, J.; Bourgeois, C.A. The nuclear skeleton and the spatial arrangement of chromosomes in the interphase nucleus of vertebrate somatic cells. *Hum. Genet.* **1986**, *74*, 1–15. [[CrossRef](#)] [[PubMed](#)]
13. Williams, K.; Ingerslev, L.R.; Bork-Jensen, J.; Wohlwend, M.; Hansen, A.N.; Small, L.; Ribel-Madsen, R.; Astrup, A.; Pedersen, O.; Auwerx, J.; et al. Skeletal muscle enhancer interactions identify genes controlling whole-body metabolism. *Nat. Commun.* **2020**, *11*, 2695. [[CrossRef](#)] [[PubMed](#)]
14. Chandra, S.; Baribault, C.; Lacey, M.; Ehrlich, M. Myogenic differential methylation: Diverse associations with chromatin structure. *Biology* **2014**, *3*, 426–451. [[CrossRef](#)]
15. Myers, R.M.; Stamatoyannopoulos, J.; Snyder, M.; Dunham, I.; Hardison, R.C.; Bernstein, B.E.; Gingeras, T.R.; Kent, W.J.; Birney, E.; Wold, B.; et al. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* **2011**, *9*, e1001046.
16. Massenet, J.; Gardner, E.; Chazaud, B.; Dilworth, F.J. Epigenetic regulation of satellite cell fate during skeletal muscle regeneration. *Skelet. Muscle* **2021**, *11*, 4. [[CrossRef](#)]

17. Zhang, X.; Ehrlich, K.C.; Yu, F.; Hu, X.; Meng, X.H.; Deng, H.W.; Shen, H.; Ehrlich, M. Osteoporosis- and obesity-risk interrelationships: An epigenetic analysis of GWAS-derived SNPs at the developmental gene TBX15. *Epigenetics* **2020**, *15*, 728–749. [[CrossRef](#)]
18. Dahlet, T.; Argueso Lleida, A.; Al Adhami, H.; Dumas, M.; Bender, A.; Ngondo, R.P.; Tanguy, M.; Vallet, J.; Auclair, G.; Bardet, A.F.; et al. Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. *Nat. Commun.* **2020**, *11*, 3153. [[CrossRef](#)]
19. Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; Ziller, M.J.; et al. Integrative analysis of 111 reference human epigenomes. *Nature* **2015**, *518*, 317–330. [[CrossRef](#)]
20. Min, J.L.; Hemani, G.; Hannon, E.; Dekkers, K.F.; Castillo-Fernandez, J.; Luijk, R.; Carnero-Montoro, E.; Lawson, D.J.; Burrows, K.; Suderman, M.; et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat. Genet.* **2021**, *53*, 1311–1321. [[CrossRef](#)]
21. Song, Q.; Decato, B.; Hong, E.E.; Zhou, M.; Fang, F.; Qu, J.; Garvin, T.; Kessler, M.; Zhou, J.; Smith, A.D. A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. *PLoS ONE* **2013**, *8*, e81148. [[CrossRef](#)] [[PubMed](#)]
22. Ehrlich, K.C.; Lacey, M.; Ehrlich, M. Epigenetics of skeletal muscle-associated genes in the ASB, LRRC, TMEM, and OSBPL gene families. *Epigenomes* **2020**, *4*, 1. [[CrossRef](#)] [[PubMed](#)]
23. Rosenbloom, K.R.; Armstrong, J.; Barber, G.P.; Casper, J.; Clawson, H.; Diekhans, M.; Dreszer, T.R.; Fujita, P.A.; Guruvadoo, L.; Haeussler, M.; et al. The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res.* **2015**, *43*, D670–D681. [[CrossRef](#)]
24. Vaisvila, R.; Ponnaluri, V.K.C.; Sun, Z.; Langhorst, B.W.; Saleh, L.; Guan, S.; Dai, N.; Campbell, M.A.; Sexton, B.S.; Marks, K.; et al. Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. *Genome Res.* **2021**, *31*, 1280–1289. [[CrossRef](#)]
25. Niu, G.; Hao, J.; Sheng, S.; Wen, F. Role of T-box genes in cancer, epithelial-mesenchymal transition, and cancer stem cells. *J. Cell. Biochem.* **2022**, *123*, 215–230. [[CrossRef](#)]
26. Lee, S.M.; Lee, J.; Noh, K.M.; Choi, W.Y.; Jeon, S.; Oh, G.T.; Kim-Ha, J.; Jin, Y.; Cho, S.W.; Kim, Y.J. Intragenic CpG islands play important roles in bivalent chromatin assembly of developmental genes. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E1885–E1894. [[CrossRef](#)]
27. Lausch, E.; Hermanns, P.; Farin, H.F.; Alanay, Y.; Unger, S.; Nikkel, S.; Steinwender, C.; Scherer, G.; Spranger, J.; Zabel, B.; et al. TBX15 mutations cause craniofacial dysmorphism, hypoplasia of scapula and pelvis, and short stature in Cousin syndrome. *Am. J. Hum. Genet.* **2008**, *83*, 649–655. [[CrossRef](#)]
28. Ehrlich, K.C.; Paterson, H.L.; Lacey, M.; Ehrlich, M. DNA hypomethylation in intragenic and intergenic enhancer chromatin of muscle-specific genes usually correlates with their expression. *Yale J. Biol. Med.* **2016**, *89*, 441–455.
29. Klug, M.; Rehli, M. Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. *Epigenetics* **2006**, *1*, 127–130. [[CrossRef](#)]
30. Tsumagari, K.; Baribault, C.; Terragni, J.; Varley, K.E.; Gertz, J.; Pradhan, S.; Baddoo, M.; Crain, C.M.; Song, L.; Crawford, G.E.; et al. Early de novo DNA methylation and prolonged demethylation in the muscle lineage. *Epigenetics* **2013**, *8*, 317–332. [[CrossRef](#)]
31. Varley, K.E.; Gertz, J.; Bowling, K.M.; Parker, S.L.; Reddy, T.E.; Pauli-Behn, F.; Cross, M.K.; Williams, B.A.; Stamatoyannopoulos, J.A.; Crawford, G.E.; et al. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.* **2013**, *23*, 555–567. [[CrossRef](#)] [[PubMed](#)]
32. Smith, Z.D.; Gu, H.; Bock, C.; Gnirke, A.; Meissner, A. High-throughput bisulfite sequencing in mammalian genomes. *Methods* **2009**, *48*, 226–232. [[CrossRef](#)] [[PubMed](#)]
33. Hnisz, D.; Abraham, B.J.; Lee, T.I.; Lau, A.; Saint-Andre, V.; Sigova, A.A.; Hoke, H.A.; Young, R.A. Super-enhancers in the control of cell identity and disease. *Cell* **2013**, *155*, 934–947. [[CrossRef](#)] [[PubMed](#)]
34. Jiang, Y.; Qian, F.; Bai, X.; Liu, Y.; Wang, Q.; Ai, B.; Han, X.; Shi, S.; Zhang, J.; Li, X.; et al. SEDb: A comprehensive human super-enhancer database. *Nucleic Acids Res.* **2019**, *47*, D235–D243. [[CrossRef](#)]
35. The GTEx Consortium. Human genomics. The genotype-tissue expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **2015**, *348*, 648–660. [[CrossRef](#)] [[PubMed](#)]
36. Lister, R.; Pelizzola, M.; Kida, Y.S.; Hawkins, R.D.; Nery, J.R.; Hon, G.; Antosiewicz-Bourget, J.; O'Malley, R.; Castanon, R.; Klugman, S.; et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **2011**, *471*, 68–73. [[CrossRef](#)]
37. Bradford, S.T.; Nair, S.S.; Statham, A.L.; van Dijk, S.J.; Peters, T.J.; Anwar, F.; French, H.J.; von Martels, J.Z.H.; Sutcliffe, B.; Maddugoda, M.P.; et al. Methylome and transcriptome maps of human visceral and subcutaneous adipocytes reveal key epigenetic differences at developmental genes. *Sci. Rep.* **2019**, *9*, 9511. [[CrossRef](#)]
38. Karlsson, M.; Zhang, C.; Méar, L.; Zhong, W.; Digre, A.; Katona, B.; Sjöstedt, E.; Butler, L.; Odeberg, J.; Dusart, P.; et al. A single-cell type transcriptomics map of human tissues. *Sci. Adv.* **2021**, *7*, eab2169. [[CrossRef](#)]
39. Puig, R.R.; Boddie, P.; Khan, A.; Castro-Mondragon, J.A.; Mathelier, A. UniBind: Maps of high-confidence direct TF-DNA interactions across nine species. *BMC Genom.* **2021**, *22*, 482. [[CrossRef](#)]
40. 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. Genome-wide MyoD binding in skeletal muscle cells: A potential for broad cellular reprogramming. *Dev. Cell* **2010**, *18*, 662–674. [[CrossRef](#)]

41. Arribas, J.; Cajuso, T.; Rodio, A.; Marcos, R.; Leonardi, A.; Velázquez, A. NF- $\kappa$ B Mediates the Expression of TBX15 in Cancer Cells. *PLoS ONE* **2016**, *11*, e0157761. [[CrossRef](#)] [[PubMed](#)]
42. Castro-Mondragon, J.A.; Riudavets-Puig, R.; Rauluseviciute, I.; Lemma, R.B.; Turchi, L.; Blanc-Mathieu, R.; Lucas, J.; Boddie, P.; Khan, A.; Manosalva Pérez, N.; et al. JASPAR 2022: The 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **2022**, *50*, D165–D173. [[CrossRef](#)] [[PubMed](#)]
43. Kheradpour, P.; Ernst, J.; Melnikov, A.; Rogov, P.; Wang, L.; Zhang, X.; Alston, J.; Mikkelsen, T.S.; Kellis, M. Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay. *Genome Res.* **2013**, *23*, 800–811. [[CrossRef](#)]
44. Bergman, D.T.; Jones, T.R.; Liu, V.; Ray, J.; Jagoda, E.; Siraj, L.; Kang, H.Y.; Nasser, J.; Kane, M.; Rios, A.; et al. Compatibility rules of human enhancer and promoter sequences. *Nature* **2022**, *607*, 176–184. [[CrossRef](#)]
45. Li, Y.; Chen, X.; Lu, C. The interplay between DNA and histone methylation: Molecular mechanisms and disease implications. *EMBO Rep.* **2021**, *22*, e51803. [[CrossRef](#)]
46. Weinberg, D.N.; Rosenbaum, P.; Chen, X.; Barrows, D.; Horth, C.; Marunde, M.R.; Popova, I.K.; Gillespie, Z.B.; Keogh, M.C.; Lu, C.; et al. Two competing mechanisms of DNMT3A recruitment regulate the dynamics of de novo DNA methylation at PRC1-targeted CpG islands. *Nat. Genet.* **2021**, *53*, 794–800. [[CrossRef](#)]
47. Fu, K.; Bonora, G.; Pellegrini, M. Interactions between core histone marks and DNA methyltransferases predict DNA methylation patterns observed in human cells and tissues. *Epigenetics* **2020**, *15*, 272–282. [[CrossRef](#)]
48. Potthoff, M.J.; Arnold, M.A.; McAnally, J.; Richardson, J.A.; Bassel-Duby, R.; Olson, E.N. Regulation of skeletal muscle sarcomere integrity and postnatal muscle function by Mef2c. *Mol. Cell. Biol.* **2007**, *27*, 8143–8151. [[CrossRef](#)]
49. Tierney, M.T.; Aydogdu, T.; Sala, D.; Malecova, B.; Gatto, S.; Puri, P.L.; Latella, L.; Sacco, A. STAT3 signaling controls satellite cell expansion and skeletal muscle repair. *Nat. Med.* **2014**, *20*, 1182–1186. [[CrossRef](#)]
50. Ponnaluri, V.K.; Ehrlich, K.C.; Zhang, G.; Lacey, M.; Johnston, D.; Pradhan, S.; Ehrlich, M. Association of 5-hydroxymethylation and 5-methylation of DNA cytosine with tissue-specific gene expression. *Epigenetics* **2016**, *12*, 123–138. [[CrossRef](#)]
51. Johansson, J.A.; Headon, D.J. Regionalisation of the skin. *Semin. Cell Dev. Biol.* **2014**, *25–26*, 3–10. [[CrossRef](#)] [[PubMed](#)]
52. Ejarque, M.; Ceperuelo-Mallafre, V.; Serena, C.; Maymo-Masip, E.; Duran, X.; Diaz-Ramos, A.; Millan-Scheiding, M.; Núñez-Álvarez, Y.; Núñez-Roa, C.; Gama, P.; et al. Adipose tissue mitochondrial dysfunction in human obesity is linked to a specific DNA methylation signature in adipose-derived stem cells. *Int. J. Obes.* **2019**, *43*, 1256–1268. [[CrossRef](#)] [[PubMed](#)]
53. Farin, H.F.; Bussen, M.; Schmidt, M.K.; Singh, M.K.; Schuster-Gossler, K.; Kispert, A. Transcriptional repression by the T-box proteins Tbx18 and Tbx15 depends on Groucho corepressors. *J. Biol. Chem.* **2007**, *282*, 25748–25759. [[CrossRef](#)] [[PubMed](#)]
54. Zhang, Y.; Li, J.; Ji, Y.; Cheng, Y.; Fu, X. Mutations in the TBX15-ADAMTS2 pathway associate with a novel soft palate dysplasia. *Hum. Mutat.* **2022**. [[CrossRef](#)]
55. Nagel, S.; Pommerenke, C.; Meyer, C.; MacLeod, R.A.F.; Drexler, H.G. Establishment of the TALE-code reveals aberrantly activated homeobox gene PBX1 in Hodgkin lymphoma. *PLoS ONE* **2021**, *16*, e0246603. [[CrossRef](#)]
56. Tian, H.; He, Y.; Xue, Y.; Gao, Y.Q. Expression regulation of genes is linked to their CpG density distributions around transcription start sites. *Life Sci. Alliance* **2022**, *5*, e202101302. [[CrossRef](#)]
57. De Micheli, A.J.; Spector, J.A.; Elemento, O.; Cosgrove, B.D. A reference single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated muscle stem cell populations. *Skelet. Muscle* **2020**, *10*, 19. [[CrossRef](#)]
58. Ronaldson, S.M.; Stephenson, D.G.; Head, S.I. Calcium and strontium contractile activation properties of single skinned skeletal muscle fibres from elderly women 66–90 years of age. *J. Muscle Res. Cell Motil.* **2022**, *43*, 173–183. [[CrossRef](#)]
59. Turner, D.C.; Seaborne, R.A.; Sharples, A.P. Comparative transcriptome and methylome analysis in human skeletal muscle anabolism, hypertrophy and epigenetic memory. *Sci. Rep.* **2019**, *9*, 4251. [[CrossRef](#)]
60. Terry, E.E.; Zhang, X.; Hoffmann, C.; Hughes, L.D.; Lewis, S.A.; Li, J.; Wallace, M.J.; Riley, L.A.; Douglas, C.M.; Gutierrez-Monreal, M.A.; et al. Transcriptional profiling reveals extraordinary diversity among skeletal muscle tissues. *Elife* **2018**, *7*, e34613. [[CrossRef](#)]
61. Voisin, S.; Jacques, M.; Landen, S.; Harvey, N.R.; Haupt, L.M.; Griffiths, L.R.; Gancheva, S.; Ouni, M.; Jähnert, M.; Ashton, K.J.; et al. Meta-analysis of genome-wide DNA methylation and integrative omics of age in human skeletal muscle. *J. Cachexia Sarcopenia Muscle* **2021**, *12*, 1064–1078. [[CrossRef](#)] [[PubMed](#)]
62. Plotkin, D.L.; Roberts, M.D.; Haun, C.T.; Schoenfeld, B.J. Muscle Fiber Type Transitions with Exercise Training: Shifting Perspectives. *Sports* **2021**, *9*, 127. [[CrossRef](#)] [[PubMed](#)]
63. Sarchielli, E.; Comeglio, P.; Filippi, S.; Cellai, I.; Guarnieri, G.; Guasti, D.; Rapizzi, E.; Rastrelli, G.; Bani, D.; Vannelli, G.; et al. Testosterone improves muscle fiber asset and exercise performance in a metabolic syndrome model. *J. Endocrinol.* **2020**, *245*, 259–279. [[CrossRef](#)] [[PubMed](#)]
64. Dos Santos, M.; Backer, S.; Saintpierre, B.; Izac, B.; Andrieu, M.; Letourneur, F.; Relaix, F.; Sotiropoulos, A.; Maire, P. Single-nucleus RNA-seq and FISH identify coordinated transcriptional activity in mammalian myofibers. *Nat. Commun.* **2020**, *11*, 5102. [[CrossRef](#)] [[PubMed](#)]
65. Delezie, J.; Weihrauch, M.; Maier, G.; Tejero, R.; Ham, D.J.; Gill, J.F.; Karrer-Cardel, B.; Rüegg, M.A.; Tabares, L.; Handschin, C. BDNF is a mediator of glycolytic fiber-type specification in mouse skeletal muscle. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 16111–16120. [[CrossRef](#)]
66. Arbanas, J.; Klasan, G.S.; Nikolic, M.; Jerkovic, R.; Miljanovic, I.; Malnar, D. Fibre type composition of the human psoas major muscle with regard to the level of its origin. *J. Anat.* **2009**, *215*, 636–641. [[CrossRef](#)]

67. Mendieta-Serrano, M.A.; Dhar, S.; Ng, B.H.; Narayanan, R.; Lee, J.J.Y.; Ong, H.T.; Toh, P.J.Y.; Röllin, A.; Roy, S.; Saunders, T.E. Slow muscles guide fast myocyte fusion to ensure robust myotome formation despite the high spatiotemporal stochasticity of fusion events. *Dev. Cell* **2022**, *57*, 2095–2110.e5. [[CrossRef](#)]
68. Coletti, C.; Acosta, G.F.; Keslacy, S.; Coletti, D. Exercise-mediated reinnervation of skeletal muscle in elderly people: An update. *Eur. J. Transl. Myol.* **2022**, *32*, 10416. [[CrossRef](#)]
69. Dikoglu, E.; Simsek-Kiper, P.O.; Utine, G.E.; Campos-Xavier, B.; Boduroglu, K.; Bonafe, L.; Superti-Furga, A.; Unger, S. Homozygosity for a novel truncating mutation confirms TBX15 deficiency as the cause of Cousin syndrome. *Am. J. Med. Genet. A* **2013**, *161*, 3161–3165. [[CrossRef](#)]
70. Curry, G.A. Genetical and developmental studies on Droop-eared mice. *Development* **1959**, *7*, 39–65. [[CrossRef](#)]
71. Yu, D.; Iwamura, Y.; Satou, Y.; Oda-Ishii, I. Tbx15/18/22 shares a binding site with Tbx6-r.b to maintain expression of a muscle structural gene in ascidian late embryos. *Dev. Biol.* **2021**, *483*, 1–12. [[CrossRef](#)] [[PubMed](#)]
72. Qian, Y.; Xiong, Z.; Li, Y.; Kayser, M.; Liu, L.; Liu, F. The effects of Tbx15 and Pax1 on facial and other physical morphology in mice. *FASEB Bioadv.* **2021**, *3*, 1011–1019. [[CrossRef](#)] [[PubMed](#)]
73. Tsumagari, K.; Chang, S.-C.; Lacey, M.; Baribault, C.; Chittur, S.V.; Sowden, J.; Tawil, R.; Crawford, G.E.; Ehrlich, M. Gene expression during normal and FSHD myogenesis. *BMC Med. Genom.* **2011**, *4*, 67. [[CrossRef](#)] [[PubMed](#)]
74. Chen, F.; Zhang, G.; Yu, L.; Feng, Y.; Li, X.; Zhang, Z.; Wang, Y.; Sun, D.; Pradhan, S. High-efficiency generation of induced pluripotent mesenchymal stem cells from human dermal fibroblasts using recombinant proteins. *Stem Cell Res. Ther.* **2016**, *7*, 99. [[CrossRef](#)] [[PubMed](#)]
75. Sun, Z.; Vaisvila, R.; Hussong, L.M.; Yan, B.; Baum, C.; Saleh, L.; Samaranyake, M.; Guan, S.; Dai, N.; Corrêa, I.R., Jr.; et al. Nondestructive enzymatic deamination enables single-molecule long-read amplicon sequencing for the determination of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Genome Res.* **2021**, *31*, 291–300. [[CrossRef](#)] [[PubMed](#)]
76. Laurent, L.; Wong, E.; Li, G.; Huynh, T.; Tsigos, A.; Ong, C.T.; Low, H.M.; Kin Sung, K.W.; Rigoutsos, I.; Loring, J.; et al. Dynamic changes in the human methylome during differentiation. *Genome Res.* **2010**, *20*, 320–331. [[CrossRef](#)] [[PubMed](#)]
77. Pidsley, R.; Zotenko, E.; Peters, T.J.; Lawrence, M.G.; Risbridger, G.P.; Molloy, P.; Van Dijk, S.; Muhlhausler, B.; Stirzaker, C.; Clark, S.J. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol.* **2016**, *17*, 208. [[CrossRef](#)] [[PubMed](#)]
78. Hon, G.C.; Hawkins, R.D.; Caballero, O.L.; Lo, C.; Lister, R.; Pelizzola, M.; Valsesia, A.; Ye, Z.; Kuan, S.; Edsall, L.E.; et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res.* **2012**, *22*, 246–258. [[CrossRef](#)]
79. Lister, R.; Pelizzola, M.; Downen, R.H.; Hawkins, R.D.; Hon, G.; Tonti-Filippini, J.; Nery, J.R.; Lee, L.; Ye, Z.; Ngo, Q.M.; et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **2009**, *462*, 315–322. [[CrossRef](#)]
80. Lacey, M.; Baribault, C.; Ehrlich, K.C.; Ehrlich, M. Data showing atherosclerosis-associated differentially methylated regions are often at enhancers. *Data Brief* **2019**, *23*, 103812. [[CrossRef](#)]
81. Lacey, M.R.; Baribault, C.; Ehrlich, M. Modeling, simulation and analysis of methylation profiles from reduced representation bisulfite sequencing experiments. *Stat. Appl. Genet. Mol. Biol.* **2013**, *12*, 723–742. [[CrossRef](#)] [[PubMed](#)]
82. Wang, E.T.; Sandberg, R.; Luo, S.; Khrebtkova, I.; Zhang, L.; Mayr, C.; Kingsmore, S.F.; Schroth, G.P.; Burge, C.B. Alternative isoform regulation in human tissue transcriptomes. *Nature* **2008**, *456*, 470–476. [[CrossRef](#)] [[PubMed](#)]
83. Cao, J.; O’Day, D.R.; Pliner, H.A.; Kingsley, P.D.; Deng, M.; Daza, R.M.; Zager, M.A.; Aldinger, K.A.; Blecher-Gonen, R.; Zhang, F.; et al. A human cell atlas of fetal gene expression. *Science* **2020**, *370*, eaba7721. [[CrossRef](#)]
84. Kodzius, R.; Kojima, M.; Nishiyori, H.; Nakamura, M.; Fukuda, S.; Tagami, M.; Sasaki, D.; Imamura, K.; Kai, C.; Harbers, M.; et al. CAGE: Cap analysis of gene expression. *Nat. Methods* **2006**, *3*, 211–222. [[CrossRef](#)] [[PubMed](#)]
85. Terragni, J.; Zhang, G.; Sun, Z.; Pradhan, S.; Song, L.; Crawford, G.E.; Lacey, M.; Ehrlich, M. Notch signaling genes: Myogenic DNA hypomethylation and 5-hydroxymethylcytosine. *Epigenetics* **2014**, *9*, 842–850. [[CrossRef](#)] [[PubMed](#)]