Two Enhancers Control Transcription of *Drosophila muscleblind* in the Embryonic Somatic Musculature and in the Central Nervous System

Ariadna Bargiela¹,²,*¹, Beatriz Llamusi¹,²,*¹, Estefanı́a Cerro-Herreros¹,², Ruben Artero¹,²*¹

¹ Translational Genomics Group, Department of Genetics, University of Valencia, Valencia, Spain, ²INCLIVA Health Research Institute, Valencia, Spain

Abstract

The phylogenetically conserved family of Muscleblind proteins are RNA-binding factors involved in a variety of gene expression processes including alternative splicing regulation, RNA stability and subcellular localization, and miRNA biogenesis, which typically contribute to cell-type specific differentiation. In humans, sequestration of Muscleblind-like proteins MBNL1 and MBNL2 has been implicated in degenerative disorders, particularly expansion diseases such as myotonic dystrophy type 1 and 2. *Drosophila muscleblind* was previously shown to be expressed in embryonic somatic and visceral muscle subtypes, and in the central nervous system, and to depend on Mef2 for transcriptional activation. Genomic approaches have pointed out candidate gene promoters and tissue-specific enhancers, but experimental confirmation of their regulatory roles was lacking. In our study, luciferase reporter assays in S2 cells confirmed that regions P1 (515 bp) and P2 (573 bp), involving the beginning of exon 1 and exon 2, respectively, were able to initiate RNA transcription. Similarly, transgenic *Drosophila* embryos carrying enhancer reporter constructs supported the existence of two regulatory regions which control embryonic expression of *muscleblind* in the central nerve cord (NE, neural enhancer; 830 bp) and somatic (skeletal) musculature (ME, muscle enhancer; 3.3 kb). Both NE and ME were able to boost expression from the Hsp70 heterologous promoter. In S2 cell assays most of the ME enhancer activation could be further narrowed down to a 1200 bp subregion (ME.3), which contains predicted binding sites for the Mef2 transcription factor. The present study constitutes the first characterization of *muscleblind* enhancers and will contribute to a deeper understanding of the transcriptional regulation of the gene.

Citation: Bargiela A, Llamusi B, Cerro-Herreros E, Artero R (2014) Two Enhancers Control Transcription of *Drosophila muscleblind* in the Embryonic Somatic Musculature and in the Central Nervous System. PLoS ONE 9(3): e93125. doi:10.1371/journal.pone.0093125

Editor: Efthimios M. C. Skoulakis, Alexander Fleming Biomedical Sciences Research Center, Greece

Received December 3, 2013; Accepted March 1, 2014; Published March 25, 2014

Copyright: © 2014 Bargiela et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by research grants from Genoma España Foundation, Generalitat Valenciana (Prometeo/2010/081), Ministerio de Ciencia e Innovacion (BFU2009-10940, including funds from Plan E), and Ministerio de Economia y Competitividad (SAF2012-36854, including funds from FEDER) to R.A. A.B. was supported by a Prometeo contract and a predoctoral fellowship from Generalitat Valenciana. B.L. was funded through Genoma España. E.C.H. was supported by a Geronimo Forteza grant from Generalitat Valenciana. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ruben.artero@uv.es

† These authors contributed equally to this work.

Introduction

Muscleblind proteins were initially identified in *Drosophila* and associated to the development of the embryonic peripheral nervous system [1], the muscles [2] and the adult photoreceptors [3]. They were later found to regulate alternative splicing of defined pre-mRNAs by binding to specific consensus sequences and to hairpins containing pyrimidine mismatches through conserved zinc finger motifs of the CCCH type ([4,5] and reviewed in [6]). Muscleblind target transcripts encode cell adhesion and cytoskeleton components, proteins involved in muscle excitation and contraction, structural proteins in muscle sarcomere and signalling molecules, among others [5,7,8,9,10,11]. Through alternative splicing *muscleblind* transcripts themselves generate at least fourteen transcript isoforms. Most of them share common 5′ sequences but differ at the 3′-ends, encoding proteins of different lengths and carboxyl termini. The *muscleblind* transcriptional unit is large and has a complex organization with ten exons distributed over about thirty times more than the average gene length in *Drosophila* [5,12].

In contrast to *Drosophila*, which has a single gene, three Muscleblind-like homologs (MBNL1, MBNL2, and MBNL3) exist in humans and mice [13,14]. Although recent results have highlighted MBNL proteins as regulators of messenger RNA (mRNA) stability [15,16,17], localization [10,18] or miRNA biogenesis [19] in the cytoplasm, these proteins are particularly well-known for their nuclear function as alternative splicing regulators. MBNL1 plays a primary role in alternative splicing allowing the fetal-to-adult splicing transitions needed for development of skeletal and cardiac muscle whereas MBNL2 seems to perform a similar function in the central nervous system [20,21]. Similarly, MBNL1 and MBNL2 are direct negative regulators of a large program of cassette exon alternative splicing events that are differentially regulated between embryonic stem cells and other cell types [22]. In contrast, MBNL3 has been reported as a member of the family with unusual functions. MBNL3 antagonizes muscle differentiation by promoting exclusion of the alternatively...
spliced β-exon of Myocyte enhancer factor 2D (Mef2D) [23] and also by the inhibition of myogenesis by maintaining myoblasts in a proliferative state [24,25]. As a result of this regulation a negative correlation exists between MBNL1 and MBNL3 expression levels in muscle during development when MBNL3 is mainly detected during embryonic development, but also transiently during injury-induced adult skeletal muscle regeneration [13,26]. MBNL1 and MBNL2 have a similar expression pattern in skeletal and heart muscle, kidney, liver, lung, intestine, brain and placenta. However, MBNL1 expression in skeletal muscle is higher than MBNL2 [13,25].

Drosophila Muscleblind shows tissue-specific expression during development. In eye-antennal imaginal discs Muscleblind is required for the formation of photoreceptor rhabdomeres, identifying muscleblind as a general factor required for terminal differentiation of adult ommatidia [3]. Its expression was also reported in the embryonic central nervous system and in the somatic muscles, where disruption of muscleblind caused defects on muscle attachments to the epidermis and disrupted Z-band differentiation of adult ommatidia [3]. Its expression was also developmentally regulated, as MBNL1 expression in skeletal muscle is higher than MBNL2 during embryonic development, but also transiently during injury-induced cell proliferation state [24,25]. As a result of this regulation a negative correlation exists between MBNL1 and MBNL3 expression levels [26]. As a result of this regulation a negative correlation exists between MBNL1 and MBNL3 expression levels [26].

The analysis of available cDNA sequences and expressed sequence tags (EST) involving the muscleblind locus showed that 5'-end sequences clustered to two locations in the gene, to the beginning of exon 1 and to the beginning of exon 2 (Fig. 1B), which suggests that muscleblind might have two transcription start sites (TSS). To test this hypothesis we defined two regions as potential promoters of muscleblind. P1 ranged from −180 to +335 (515 bp long) while P2 spanned from −243 to +343 (586 bp long) (Fig. 1A), defining as +1 the first bp in exon 1 and exon 2, respectively. Although core promoter regions are typically defined as +50 to −50 of transcription start site [33], a longer region was used to include not only the core promoter but also proximal promoter sequences with potential activator binding sites.

A SCOPE database analysis of Drosophila promoter elements [34] confirmed the accumulation of known consensus sequences in P1 and P2, thus supporting their potential as promoter regions (Fig. 1 C,D). Furthermore, these motifs were phylogenetically conserved among Drosophila species in the Multiz Alignments & phastCons Scores provided by the UCSC, supporting the relevance of these non-coding regions (data not shown). To test the functional relevance of the putative promoters we generated reporter constructs in which P1 and P2 drove expression of Firefly luciferase. In addition, we also tested the activity of shorter versions, contained in the longer ones, of 220 bp long (−104 to +116; P1.1) and 397 bp long (−81 to +316; P2.1), respectively, in an attempt to define minimal promoters. Dual luciferase reporter assays in Drosophila S2 cells transfected with the resulting constructs revealed that P1 was able to boost luciferase readings more than 100 fold higher relative to the promoter-less control. This was 2.5 fold the transcription measured for P1.1, 10 fold higher than the luciferase activity driven by P2 and 7 times the activity measured for P2.1 (Fig. 1E). Thus, robust expression of luciferase was observed in P1 constructs in comparison to reporter expression driven by P2, and the higher activity obtained with P1 in comparison to P1.1 suggests that P1 contains proximal promoter elements that are not included in P1.1.

2. Identification of Putative cis-regulatory Modules

Potential cis-regulatory elements of transcription can be identified as highly conserved non-coding regions in phylogenetic footprinting analyses (as an example see [35]). A fragment of 120 kb harbouring most of the muscleblind gene, plus 20 kb upstream of the gene (complete sequence analyzed chr2R:13133058–13252891), was used as the reference DNA to align orthologous sequences from 12 Drosophilids. Using the bioinformatics tool rVista, an intronic sequence showing above 90% similarity between Drosophila melanogaster and D.najaensis was selected. This sequence, referred to as H region, was 872 bp long and was located in intron 2 (Fig. 2A and Table 1). Moreover, chromatin immunoprecipitation followed by micro-array analysis (ChIP-on-chip) data revealed putative cis-regulatory modules (CRMs) M1, M2, M3 and ML (Fig. 2A and Table 1) in the muscleblind locus that bound Drosophila Meff2 [36,37,38]. We found these results particularly relevant because Meff2 is a known activator of muscleblind expression in the Drosophila embryo [2]. Interestingly, these candidate CRMs not only bind Meff2 in ChIP-on-chip experiments but also other muscle organizing factors such as Biniou (Bin), Tinman (Tin) or Twist (Table 1). The ML region only bound Meff2 in late embryos according to [37].

3. In vivo Testing of Reporter Expression Reveals Tissue-specific Enhancers

To test the regulatory potential of the highly conserved (H) and the Meff2-bound (M1, M2, M3 and ML) genomic regions involving the muscleblind locus, we generated fusion constructs in the Drosophila transformation vector pH-Stinger. This vector

Results

1. Mapping of the Promoter Regions of Drosophila muscleblind

The analysis of available cDNA sequences and expressed sequence tags (EST) involving the muscleblind locus showed that 5'-
contains the heat shock protein 70 \((Hsp70)\) promoter and is specifically designed to avoid chromatin configuration effects (“position effects”) by flanking the eGFP reporter cDNA with two copies of insulator sequences from the \(gypsy\) transposon [39]. Embryonic expression of the eGFP was assessed either directly (green fluorescence) (Fig. 2) or immunodetecting the reporter protein with a polyclonal anti-GFP antibody (Fig. 3). Only embryos carrying reporter constructs under the control of the M2 and ML candidate CRMs revealed consistent patterns (Fig. 2G–I; Fig. 3; and not shown) and in both cases eGFP expression was restricted to nuclei, as expected by the presence of a nuclear localization signal in the vector.

We have previously shown that Muscleblind is localized in the nuclei of embryonic pharyngeal, visceral and somatic muscles, in the larval photoreceptor system, and in repeated clusters of cells within the central nervous system [2,3] (Fig. 2B,C). In combination with the \(Hsp70\) promoter, M2 drove robust eGFP expression in the somatic musculature of late embryos, approximately starting from stage 13. Notably, no eGFP expression was detected in other muscle derivatives or tissues where endogenous \(muscleblind\) is normally detected, particularly the CNS (Fig. 2I). As control, we generated transgenics carrying a promoter-less M2:eGFP construct, which revealed no eGFP expression (Fig. 2D–F), thus confirming that M2 had no promoter activity by itself but requires the presence of a promoter to exert its enhancer activity. Similarly,
double eGFP and Muscleblind immunostaining of fly embryos carrying ML:eGFP constructs revealed that ML drove expression to clusters of cells in the ventral cord of late embryos that overlapped with those expressing endogenous Muscleblind (Fig. 3). First signal started at developmental stage 12 and no eGFP expression was detected in tissues other than the CNS. Consistently, ML included predicted binding sites for factors involved in nervous system development such as Ladybird early (Lbe), Ladybird late (Lbl), Kruppel (Kr) and Hunchback (Hb) [40,41,42,43] (Fig. 4G). In summary, these results support that M2 and ML are somatic muscle and CNS-specific enhancers of muscleblind, respectively, at least during embryonic development.

Figure 2. ME reproduces muscleblind expression in the embryonic somatic musculature. (A) Localization of the putative cis-regulatory modules M1, M2, ML, H and M3, indicated as orange boxes, in the context of the muscleblind genomic locus. Fluorescence confocal images of lateral (B,E,H,) and ventral (C,F,I,) views of late Drosophila embryos. (D,G) Schematic representation of the reporter constructs used to transform the germline of Drosophila. In control yw flies (B,C) an anti-Mbl antibody detects robust expression in the somatic musculature and in the CNS (green). Direct visualization of the GFP reporter under the control of the ME enhancer in the pH-Stinger vector (E,F,H,I). Promoter-less ME constructs (D–F) do not activate GFP expression and serve as negative controls. Flies carrying the ME enhancer upstream of Hsp70 (G–I) reproduce Muscleblind expression in the somatic musculature but not in the CNS. All micrographs were taken at 200× magnification. Anterior is to the left and dorsal up unless otherwise stated. doi:10.1371/journal.pone.0093125.g002

| CRM | Location | Size (bp) | TF binding sites |
|-----|----------|-----------|------------------|
| M1  | Upstream | 2278      | 1 Twi            |
|     |          |           | 6 Mef2           |
|     |          |           | 1 Bin            |
| M2  | Intron 2 | 3340      | 2 Mef2           |
|     |          |           | 3 Bin            |
| M3  | Exon 4   | 1240      | 3 Mef2           |
|     |          |           | 1 Bin            |
| ML  | Intron 2 | 830       | 3 Mef2           |
|     |          |           | 1 Bin            |
| H   | Intron 2 | 872       | 0                |

All except region H according to [37].

doi:10.1371/journal.pone.0093125.t001
We therefore renamed these candidate CRMs as “ME” (from muscle enhancer) and “NE” (from nervous system enhancer). Importantly, both ME and NE were capable of activating a heterologous promoter, a typical ability of transcriptional enhancer elements.

4. Characterization of the Muscle and Nervous Enhancer Elements of muscleblind

To test the ability of ME and NE genomic regions to enhance transcription from the putative muscleblind promoters, we used luciferase reporter assays in Drosophila S2 cells. We used S2 cells because this is a well characterized cell line whereas muscle or neuron-specific cell lines were not immediately available. Both enhancer regions were cloned upstream to each of the putative promoters P1, P1.1, P2 and P2.1 in their forward (Fig. 4A) and reverse (not shown) orientations in the pGL3 basic vector, which carries the Firefly luciferase reporter. As controls, ME and NE were tested in promoter-less constructs and no luciferase activity was detected (data not shown), thus confirming that the enhancer regions do not have any transcriptional activity by themselves.

When promoters were in the forward orientation we observed that the transcription originated from P1 and P1.1 was strongly enhanced by ME, whereas NE had no effect on P1.1, or even decreased transcription when in combination with P1 (Fig. 4B). Similarly, P2 and P2.1 promoter activity was significantly enhanced by NE (around 30% and 20%, respectively), but remained unchanged when in combination with ME that even repressed transcription from P2.1 (Fig. 4C). As control of promoter directionality, luciferase levels of constructs carrying promoters in their reverse orientation were measured. Consistently, relative luciferase readings dropped to close to background levels in constructs containing promoters in their reverse orientation (Fig. 4D compare B,C with D), although both ME and NE still managed to significantly potentiate transcription from P1 and P1.1. Thus, promoter activity is orientation-dependent, as reversed promoters expressed significantly less luciferase reporter, and the activity of the ME and NE enhancers on P1 and P2 suggests enhancer-promoter communication specificity.

ME function was further analyzed to narrow down sequences necessary for enhancer activity. This involved testing three smaller regions, approximately 1 kb each, here referred to as ME.1, ME.2 and ME.3 according to their relative position in the original region (Fig. 4A). These sequences were cloned upstream to the P1 and P1.1 promoters and were used in luciferase reporter assays in S2 cells (Fig. 4E). Compared to the luciferase activity of the promoter alone, ME.3 was the only subregion able to significantly increase transcription from P1.1, also showing the same trend on the P1 promoter. Notably, all other subregions tested either did not boost expression from P1 or P1.1 or even inhibited it. Therefore, these data, and bioinformatics analyses support that ME.3 contains sequence motifs necessary for ME enhancing activity, including consensus Mef2 binding sites (Fig. 4F), but they also suggest that for maximum enhancing activity all three subregions are required. Bioinformatics analyses in NE found an enrichment of targets for nervous system transcription factors (Fig. 4G).

5. Functional Conservation of Human MBNL1 Promoter

Sequence conservation between Drosophila and human MBNL1 promoter sequences was patently non-existent. However, analysis of available cDNA sequences and ESTs in the MBNL1 locus suggested that human MBNL1 might also use TSS located in exon 1 and in exon 2 (Fig. 5A). Consistently, putative TSS includes...
promoter marks such as CpG islands and histone modification tracks (Fig. 5A,B). We defined 500 bp around the predicted start region from both exons to test them as putative human MBNL1 promoters; Hsa-P1 in exon 1 (chr3: 151985544–151985045) and Hsa-P2 in exon 2 (chr3: 152016823–152017382). Synthetic Hsa-P1 and Hsa-P2 sequences were designed to replace the Hsp70 promoter in the Drosophila transformation pH-Stinger vector. No eGFP expression was observed in transgenic fly embryos carrying any of the human promoters alone (data not shown and Fig. 5D–F). However, we observed a robust expression of eGFP in the somatic musculature of embryos when ME drove expression of the Hsa-P1 promoter (Fig. 5G–I compare to Fig. 2B,C). This expression was not observed in similar reporter constructs where Hsa-P2 replaced Hsa-P1 (ME-Hsa-P2; not shown). These data support that Hsa-P1 can initiate transcription, as we have also demonstrated for the muscleblind P1 promoter, and that ME is a muscle enhancer on a variety of promoters.

Figure 4. ME and NE boost P1 and P2 promoter activity, respectively. (A) Schematic representation of the Firefly luciferase reporter constructs used to transiently transfect Drosophila S2 cells. (B) ME, but not NE, potentiates P1 or P1.1 promoter activity. Basal promoter activity of P1 is significantly higher than P1.1 in these assays. Conversely, NE, but not ME, weakly enhanced P2 or P2.1 promoter function (C), being the relative luciferase activity measured approximately one tenth of that from P1 or P1.1. (D) Reversed P1 and P1.1 promoters still responded to ME and NE, although at much lower levels, whereas P2 and P2.1 did not significantly change reporter expression. (E) ME (GenBank accession number KJ201027) was subdivided into three smaller regions of approximately 1 kb (ME.1, ME.2 and ME.3). ME.3 retained most of ME ability to boost expression, although only for P1.1 it reached statistical significance. ME.3 and NE (GenBank accession number KJ201028) sequences are enriched in consensus binding sites for MeF2 (underlined) (F) and nervous system (G) transcription factors according to the following code: Hunchback sites, underlined; Ladybird early and Ladybird late, bold; Krüppel, italics respectively. Predictions used the Jaspar and rVista programs.

doi:10.1371/journal.pone.0093125.g004
Muscleblind orthologues have attracted intense research interest due to their important role in vertebrate muscle development as well as involvement in several degenerative RNA-mediated diseases including DM1 and DM2, Huntingtons disease, Huntington’s disease-like (HDL2) or spinocerebellar ataxia 8 (SCA8) [6,44,45,46]. More recently, MBNL proteins were found to repress embryonic stem cell alternative splicing patterns, uncovering an additional role in the control of the cell pluripotency [22]. Despite this, little is known about the transcriptional regulation of muscleblind genes both in Drosophila and in vertebrates.

As a means of dissecting the cis-regulation of muscleblind, we analyzed a genomic DNA fragment harboring the muscleblind locus and its upstream region, looking for CRMs that regulate basal initiation of transcription (promoters) and tissue-specific expression (enhancers). In silico and in vivo studies identified two putative promoters, P1 in exon 1 and P2 in exon 2, and two intronic tissue-specific regulatory elements, a region of 3340 bp which drives specific expression in somatic muscle (ME) and a region of 830 bp which drives expression in central nervous system (NE). Both enhancers had been selected because of their enrichment in Mef2 binding sites according to ChIP-on-chip data [37], because Mef2 is a known positive regulator of muscleblind in the embryo [2]. Nevertheless, other enhancer elements must exist in order to explain the rich embryonic expression pattern of muscleblind, which also includes expression in visceral and pharyngeal musculature, the Bolwigs organ (the larval photoreceptor system) and the imaginal discs [2,3]. Regarding this, putative CRMs that did not reproduce any embryonic pattern in our study can not be discarded as functional in other developmental stages.

Enhancer regions ME and NE were able to boost expression originating from the heterologous Hsp70 promoter in transgenic

**Discussion**

Muscleblind orthologues have attracted intense research interest due to their important role in vertebrate muscle development as well as involvement in several degenerative RNA-mediated diseases including DM1 and DM2, Huntington’s disease, Huntington's disease-like (HDL2) or spinocerebellar ataxia 8 (SCA8) [6,44,45,46]. More recently, MBNL proteins were found to repress embryonic stem cell alternative splicing patterns, uncovering an additional role in the control of the cell pluripotency [22].

As a means of dissecting the cis-regulation of muscleblind, we analyzed a genomic DNA fragment harboring the muscleblind locus and its upstream region, looking for CRMs that regulate basal initiation of transcription (promoters) and tissue-specific expression (enhancers). In silico and in vivo studies identified two putative promoters, P1 in exon 1 and P2 in exon 2, and two intronic tissue-specific regulatory elements, a region of 3340 bp which drives specific expression in somatic muscle (ME) and a region of 830 bp which drives expression in central nervous system (NE). Both enhancers had been selected because of their enrichment in Mef2 binding sites according to ChIP-on-chip data [37], because Mef2 is a known positive regulator of muscleblind in the embryo [2]. Nevertheless, other enhancer elements must exist in order to explain the rich embryonic expression pattern of muscleblind, which also includes expression in visceral and pharyngeal musculature, the Bolwigs organ (the larval photoreceptor system) and the imaginal discs [2,3].

Regarding this, putative CRMs that did not reproduce any embryonic pattern in our study can not be discarded as functional in other developmental stages.

Enhancer regions ME and NE were able to boost expression originating from the heterologous Hsp70 promoter in transgenic...
embryos. However, in luciferase S2 assays, ME preferentially activated the P1 promoter while NE showed preference for P2, although the enhancer activity of NE was smaller than that of ME and transcription arising from P2 was on average one tenth of that from P1 (Fig. 4B,C). Our results in transgenic flies carrying the ME in combination with the human MBNL1 P1 promoter also suggest enhancer promoter specificity as ME only induced reporter expression from Hsa-P1, but not from Hsa-P2. The use of alternative promoters is a known mechanism of transcriptional regulation, which has been reported to influence levels of transcription, turnover or translation efficiency of mRNA isoforms with different leader exons, tissue specificity [47] and to generate protein isoforms differing in the amino termini (reviewed in [48]). Furthermore, different core promoters have been found to possess distinct regulatory activities driven by the same enhancer in the Drosophila embryo ([49]). In muscleblind, the potential in vivo use of P1 and P2 as alternative promoters would have no consequences as for the encoded protein since the start codon is located in exon 2, which is downstream of both. However, the in vivo relevance of the internal promoter P2 remains to be specifically addressed. It is also worth mentioning that whereas the identified enhancers provided strong activation of the reporter in vivo, particularly ME, the measured activation in S2 cells was discrete, reaching some 6-fold increase over the promoter alone condition (Fig. 4B). This may stem from the particular combination of transcription factors that S2 cells express, which may not be particularly favourable to activate myogenic enhancers. Mef2, for example, is weakly expressed in S2 cells according to modENCODE data [50]. Consistently, Muscleblind is only barely detectable in this cell line [51]. Nevertheless, the low expression of Muscleblind in S2 cells offers an opportunity to test the activating potential of candidate regulatory transcription factors.

Despite that sequences homologous to ME or NE in human MBNL1 are not obvious, ENCODE Chip-seq data confirms that there is a high concentration of transcription factor binding sites in the first intron of MBNL1, including multiple MEF2A and MEF2C binding regions, thus suggesting that MEF2, a central regulator of diverse developmental programs [52], is also involved in the regulation of human MBNL1 transcription. Indeed, detailed information on the multiple transcription factors converging on the Drosophila ME and NE enhancer elements would help in the identification of the functionally equivalent enhancer regions that integrate inputs from the same factors in humans. Although hypothetical, the functional conservation among fly and human muscleblind enhancers is conceivable. Deformed enhancers, for example, drive meaningful spatial expression patterns in vivo in a mouse context [53]. In any event, this initial characterization of cis-regulatory regions is the first step towards the understanding of the transcriptional regulation of muscleblind.

Materials and Methods

Constructs

P1 and P2, and their shorter versions P1.1 and P2.1, promoter regions were synthesized by GenScript with XhoI/XbaI terminal adapters and were provided cloned into the pUC57 vector. High fidelity PCR (KAPA HIF DNA Polymerase, KAPA biosystems) was used to subclone into the pGL3-Basic vector (Promega) previously linearized with the same enzymes. ME, ME.1, ME.2, ME.3 (spanning from 13170622–13171543, ME.1; 13171544–13172759, ME.2; 13172760–13173908, ME.3) and NE were obtained from Drosophila genomic DNA by high fidelity PCR and were cloned into KpnI/SalI digested pGL3-Basic vectors already including different promoter regions. To generate transgenic flies, M1, M2, M3, H and ML fragments were PCR amplified from genomic DNA and were cloned into the BglII and XhoI sites of the Drosophila expression vector pH-Stinger [54]. Synthetic Hsa-P1 and Hsa-P2 promoter regions, including PAd, BglII, XhoI (5’) and HindIII, PAd (5’) adapter sites, were cloned into the pUC57 vector and subsequently transferred to the Pst site of the pH-Stinger vector. To generate ME-Hsa-P1 and ME-Hsa-P2 constructs, M2 was amplified with specific oligos containing BglII/XhoI adapters, digested, and cloned into the corresponding sites of the pH-Stinger vector already containing the candidate human promoters. All constructs were confirmed by sequencing. Description of used primers is in Table 2.

Cell Culture and Dual Luciferase Assays

Drosophila melanogaster Schneider 2 cells (S2) were cultured at 27°C in growing media containing 90% Schneider insect media (Gibco), 10% heat inactivated fetal bovine serum (FBS), 100 units/ml of penicillin and 100 mg/ml of streptomycin. 48 h before transfection, 106 log-phase cells were transferred onto 24-well plates (300 μl per well). 4 μl of Cellfectin reagent (Invitrogen) in 200 μl of serum and antibiotic-free medium were used to co-transfect 450 ng of the pGL3 reporter plasmid of interest and 25 ng of Renilla luciferase. A GFP expressing vector served as transfection efficiency control. Cells were incubated 16 h with transfection mix, then media was replaced by Schneider complete medium and the culture was additionally maintained 24 h at 27°C. Luciferase expression was monitored using the Dual-Luciferase Reporter Assay System (Promega). Briefly, this involved adding 100 μl of lysis buffer per well, shaking for 15 min and transferring the lysate to a white 96-well plate with 40 μl of luciferase substrate. After 10 s of luminescence detection, Stop&Glo buffer was added and luminescence measurement was repeated. Luminescence readings used an EnVision plate reader (PerkinElmer). In cell culture luciferase reporter assays, all graphs show the average of three independent experiments with three technical replicates each. P-values were obtained using a two-tailed, non-paired t-test (α = 0.05). Welch’s correction was applied when variances were significantly different.

Immunohistochemistry of Drosophila Embryos

Embryos were fixed for 20 min using 4% paraformaldehyde in PBS, devitellinized with a heptane:methanol 1:1 mixture, and blocked with 0.1% Triton-X-100 in PBS (PBT) with 1% BSA for 15 min and later with 2% BSA in PBT for 30 min. Subsequently they were incubated with rabbit anti-GFP (1:200 Torrey Pines Biologs) antibody diluted in blocking solution containing 1% donkey serum for 2 h at room temperature. After washes embryos were incubated with an anti-rabbit-FITC (1:200 Calbiochem) secondary antibody for 45 min. Muscleblind detection used sheep anti-Muscleblind (1:500 [55]) for 2 h followed by washes and primary antibody recognition with sheep biotin-conjugated secondary antibody (1:100, Sigma) for 2 h. Then, washed embryos were incubated with ABC solution (ABC kit, VECTASTAIN) for 30 min at room temperature, and were washed and incubated with streptavidin-Texas Red (1:1000, Vector) for 45 min. In all cases embryos were washed 3× with 1% BSA in PBT and were mounted in Vectashield (Vector) with 2 μg/ml DAPI. Images were taken on an Olympus Fluoview FV100 confocal microscope. At least 10–15 embryos of the desired stage, and showing the relevant expression patterns, were analyzed.

Drosophila Strains and Transgenics

Genomic DNA was extracted from the Drosophila genome project strain y++; Gt2281 Gt2282 cn c1 CG33964 and bal sp1; LysC1 project strain.
MstProx1 GstD51 Rh61 w1118 BDGP R5/dm3 and human GRCh37 releases. 

Predictions of transcription factor binding sites used JASPAR [56]. Phylogenetic conservation employed the Whole Genome vRISTA browser [57]. EST mapping was according to the UCSC Genome Browser database [58]. Predicted promoter consensus sequences used Suite for Computational Identification of Promoters (SCOPE) [34]. Genomes of reference were flies were also from Bloomington IN. All constructs were injected into w1118 embryos by BestGene, typically resulting in 2–6 independent transgenic lines.

Web Resources

Predictions of transcription factor binding sites used JASPAR [56]. Phylogenetic conservation employed the Whole Genome vRISTA browser [57]. EST mapping was according to the UCSC Genome Browser database [58]. Predicted promoter consensus sequences used Suite for Computational Identification of Promoters (SCOPE) [34]. Genomes of reference were Drosophila BDGP R5/dm3 and human GRCh37 releases.

References

1. Kania A, Salberg A, Bhat M, D’Evelyn D, He Y, et al. (1995) P-element mutations affecting embryonic peripheral nervous system development in Drosophila melanogaster. Genetics 139: 1663–1678.

2. Artero R, Prokop A, Paricio N, Begemann G, Pueyo I, et al. (1998) The muscleblind gene participates in the organization of Z-bands and epidermal attachments of Drosophila muscles and is regulated by Dmef2. Dev Biol 195: 131–143.

3. Begemann G, Paricio N, Artero R, Kiss I, Perez-Alonso M, et al. (1997) Muscleblind isoforms are functionally distinct and regulate alpha-actinin attachments of Drosophila muscles and is regulated by Dmef2. Dev Biol 195: 131–143.

4. Artero R, Prokop A, Paricio N, Begemann G, Pueyo I, et al. (1998) The muscleblind gene participates in the organization of Z-bands and epidermal attachments of Drosophila muscles and is regulated by Dmef2. Dev Biol 195: 131–143.

5. Irion U (2012) Drosophila muscleblind codes for proteins with one and two Cys2H-type zinc-finger-containing proteins. Development 124: 4321–4331.

6. Goens ES, Voelker RB, Gates DP, Berghard JA (2008) RNA binding specificity of Drosophila muscleblind. Biochemistry 47: 7294–7299.

7. Irion U (2012) Drosophila muscleblind codes for proteins with one and two tandem zinc finger motifs. PLoS One 7: e34248.

8. Fernandez-Costa JM, Llamusi MB, Garcia-Lopez A, Artero R (2011) Alternative splicing regulation by Muscleblind proteins: from development to disease. Biol Rev Camb Philos Soc 86: 947–958.

9. Vicente-Crespo M, Pascual M, Fernandez-Costa JM, Garcia-Lopez A, Monferrer L, et al. (2008) Drosophila muscleblind is involved in troponin T alternative splicing and apoptosis. PLoS ONE 3: e1613.

10. Vicente M, Monferrer L, Poulos MG, Houlemy JM, Moneckson DG, et al. (2007) Muscleblind isoforms are functionally distinct and regulate alpha-actinin splicing. Differentiation 75: 427–440.

11. Charizanis K, Lee KY, Batra R, Goodwin M, Zhang C, et al. (2012) Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron 75: 437–450.

Table 2. Sequence and melting temperature (Tm) of PCR primers used for cloning genomic DNA.

| Primer pair | Forward 5’→3’ | Tm | Reverse 5’→3’ | Tm |
|-------------|---------------|----|---------------|----|
| M1 pH-Stinger | GAGAATTCCTGGTCAGGAGCATTGC | 70.1 | GCTCTAGAATAATAGGAAATAATGG | 52.2 |
| M2 pH-Stinger | GAGAATTCCTGGTCAGGAGCATTGC | 66.8 | GCTCTAGAATAATAGGAAATAATGG | 70.6 |
| M3 pH-Stinger | GAGAATTCCTGGTCAGGAGCATTGC | 66.9 | GCTCTAGAATAATAGGAAATAATGG | 62.0 |
| H pH-Stinger | GAGAATTCCTGGTCAGGAGCATTGC | 52.3 | GCTCTAGAATAATAGGAAATAATGG | 67.5 |
| M2 pH-Stinger | GAGAATTCCTGGTCAGGAGCATTGC | 66.8 | GCTCTAGAATAATAGGAAATAATGG | 70.6 |
| M3 pH-Stinger | GAGAATTCCTGGTCAGGAGCATTGC | 66.9 | GCTCTAGAATAATAGGAAATAATGG | 62.0 |
| H pH-Stinger | GAGAATTCCTGGTCAGGAGCATTGC | 52.3 | GCTCTAGAATAATAGGAAATAATGG | 67.5 |

Restriction sites are highlighted in bold.

doi:10.1371/journal.pone.0093125.t002

Acknowledgments

We thank J. L. Gómez-Skarmeta and A. Fernández-Miñán for advice with the bioinformatics analysis, and EE. Furlong for providing detailed ChIP-on-chip data.

Author Contributions

Conceived and designed the experiments: RA BL AB. Performed the experiments: AB ECH BL. Analyzed the data: AB BL RA. Wrote the paper: AB RA BL ECH.

Two New Enhancer of Drosophila muscleblind Gene

Web Resources

Predictions of transcription factor binding sites used JASPAR [56]. Phylogenetic conservation employed the Whole Genome rVISTA browser [57]. EST mapping was according to the UCSC Genome Browser database [58]. Predicted promoter consensus sequences used Suite for Computational Identification of Promoter Elements (SCOPE) [34]. Genomes of reference were Drosophila BDGP R5/dm3 and human GRCh37 releases.

References

1. Kania A, Salberg A, Bhat M, D’Evelyn D, He Y, et al. (1995) P-element mutations affecting embryonic peripheral nervous system development in Drosophila melanogaster. Genetics 139: 1663–1678.

2. Artero R, Prokop A, Paricio N, Begemann G, Pueyo I, et al. (1998) The muscleblind gene participates in the organization of Z-bands and epidermal attachments of Drosophila muscles and is regulated by Dmef2. Dev Biol 195: 131–143.

3. Begemann G, Paricio N, Artero R, Kiss I, Perez-Alonso M, et al. (1997) muscleblind, a gene required for photoreceptor differentiation in Drosophila, encodes novel nuclear Cys2H-type zinc-finger-containing proteins. Development 124: 4321–4331.

4. Goens ES, Voelker RB, Gates DP, Berghard JA (2008) RNA binding specificity of Drosophila muscleblind. Biochemistry 47: 7294–7299.

5. Irion U (2012) Drosophila muscleblind codes for proteins with one and two tandem zinc finger motifs. PLoS One 7: e34248.

6. Fernandez-Costa JM, Llamusi MB, Garcia-Lopez A, Artero R (2011) Alternative splicing regulation by Muscleblind proteins: from development to disease. Biol Rev Camb Philos Soc 86: 947–958.

7. Vicente-Crespo M, Pascual M, Fernandez-Costa JM, Garcia-Lopez A, Monferrer L, et al. (2008) Drosophila muscleblind is involved in troponin T alternative splicing and apoptosis. PLoS ONE 3: e1613.

8. Vicente M, Monferrer L, Poulos MG, Houlemy JM, Moneckson DG, et al. (2007) Muscleblind isoforms are functionally distinct and regulate alpha-actinin splicing. Differentiation 75: 427–440.

9. Machaca-Tiizli L, Thorpe H, Robinson TE, Sewry C, Brook JD (2006) Flies deficient in Muscleblind protein model features of myotonic dystrophy with altered splice forms of Z-band associated transcripts. Hum Genet 120: 487–499.

10. Wang ET, Cody NA, Jorg S, Biancoella M, Wang TT, et al. (2012) Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. Cell 150: 710–724.

11. Picchio L, Plantiez E, Renaud Y, Poweshumkadavil P, Jagla K (2013) Novel Drosophila model of myotonic dystrophy type 1: phenotypic characterization and genome-wide view of altered gene expression. Hum Mol Genet 22: 2795–2805.

12. Alman A, Solovyev VV (2000) Ab initio gene finding in Drosophila genomic DNA. Genome Res 10: 516–522.

13. Kanadia RN, Urbani CR, Crusselle VJ, Luo D, Lee YJ, et al. (2003) Developmental expression of mouse muscleblind genes Mbnl1, Mbnl2 and Mbnl3. Gene Expr Patterns 3: 459–462.

14. Miller JW, Urbani CR, Teng-Ummay P, Steinberg MG, Byrne BJ, et al. (2000) Recruitment of human muscleblind proteins to (CUG)n expansions associated with myotonic dystrophy. Endo J 19: 4439–4448.

15. Du H, Cline MS, Osborne RJ, Torti DL, Clark TA, et al. (2010) Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. Nat Struct Mol Biol 17: 187–193.

16. Osborne RJ, Lin X, Welle S, Sobczak K, O’Rourke JR, et al. (2009) Transcriptional and post-transcriptional impact of toxic RNA in myotonic dystrophy. Hum Mol Genet 18: 2795–2810.

17. Masuda A, Andersen HS, Doktor TK, Okamoto T, Ito M, et al. (2012) CUGBP1 and MBNL1 preferentially bind to 3’ UTRs and facilitate mRNA decay. Sci Rep 2: 209.

18. Aderseth Y, Damvald V, Kose N, Li R, Hsu T (2003) RNA-dependent integrin α3 protein localization regulated by the Muscleblind-like protein MBLP1. Nature Cell Biology 7: 1240–1247.

19. Rau F, Freyermuth F, Fugier C, Villemin JP, Fischer MC, et al. (2011) Muscular dystrophy with myotonic dystrophy is caused by mutations in the muscleblind-like 2 gene MLP1. Cell 150: 710–724.
21. Lin X, Miller JW, Mankodi A, Kanadia RN, Yuan Y, et al. (2006) Failure of MBNL1-dependent postnatal splicing transitions in myotonic dystrophy. Hum Mol Genet.

22. Han H, Irimia M, Ross PJ, Sung HK, Alipanahi B, et al. (2013) MBNL proteins repress E-box-specific alternative splicing and reprogramming. Nature 490: 241–245.

23. Lee KS, Wang EH (2002) Inhibition of muscle differentiation by the novel muscleblind-related protein CHCR. Dev Biol 250: 218–230.

24. Juni N, Yamamoto D (2009) Genetic analysis of chaste, a new mutation of Drosophila melanogaster characterized by extremely low female sexual receptivity. J Neurogenet 23: 329–340.

25. Tabebordbar M, Wang ET, Wagers AJ (2013) Skeletal muscle degenerative diseases and strategies for therapeutic muscle repair. Annu Rev Pathol 8: 441–475.

26. Poules MG, Batra R, Li M, Yuan Y, Zhang C, et al. (2013) Progressive impairment of muscle regeneration in muscleblind-like 3 isofrom knockout mice. Hum Mol Genet 22: 3547–3558.

27. Poulos MG, Batra R, Li M, Yuan Y, Zhang C, et al. (2013) Progressive impairment of muscle regeneration in muscleblind-like 3 isofrom knockout mice. Hum Mol Genet 22: 3547–3558.

28. Soullairme RM, Chenault DM, Wang EH (2007) Expression pattern of muscleblind-like 3 protein Muscleblind-like 3 (MBNL3) disrupts myocyte enhancer factor 2 (Mef2) (horizion) splicing. J Biol Chem 282: 33779–33787.

29. Lee KS, Squillace RM, Wang EH (2007) Expression pattern of muscleblind-like proteins differs in differentiating myoblasts. Biochem Biophys Res Commun 361: 151–153.

30. Ohler U (2006) Identification of core promoter modules in Drosophila and their application in accurate transcription start site prediction. Nucleic Acids Res 34: 5943–5950.

31. Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshere MG, et al. (2002) Impaired expression of muscleblind-related proteins at all stages of muscle development. Dev Cell 10: 797–807.

32. Houseley JM, Wang Z, Brock GJ, Soloway J, Artero R, et al. (2005) Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind and deregulated alternative splicing in neurons. Hum Mol Genet 14: 873–883.

33. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, et al. (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res 22: 1813–1831.

34. Jagla K, Georgel P, Bellard M, Dretzen G, Bellard F, et al. (1993) A novel homeobox gene expresses a muscleblind-related protein CHCR. Development 118: 401–415.

35. Houseley JM, Brock GJ, Soloway J, Artero R, et al. (2005) Identification of functional elements and regulatory circuits by Drosophila modENCODE. Science 310: 1767–1797.

36. Fernandez-Costa JM, Artero R (2010) A conserved motif controls nuclear localization of Drosophila Muscleblind. Mol Cells 30: 65–70.

37. Jagla K, Georgel P, Bellard M, Dretzen G, Bellard F, et al. (1993) A novel homeobox gene expresses a muscleblind-related protein CHCR. Development 118: 401–415.

38. Mykowska A, Sobczak K, Wójcieszewska M, Kołodzki P, Krzyżosiak WJ (2011) CAG repeats mimic CUG repeats in the misregulation of alternative splicing. Nucleic Acids Res 39: 8938–8951.

39. Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, et al. (2010) Identification of functional elements and regulatory circuits by Drosophila modENCODE. Genome Res 20: 2418–2428.

40. Awgulewitsch A, Jacobs D (1992) Deformed autoregulatory element from Drosophila muscleblind. Gene 13: 3079–3088.

41. Fiszer A, Krzyżosiak WJ (2013) RNA toxicity in polyglutamine disorders: concepts, models, and progress of research. J Mol Med (Berl) 91: 683–691.

42. Fiszer A, Krzyżosiak WJ (2013) RNA toxicity in polyglutamine disorders: concepts, models, and progress of research. J Mol Med (Berl) 91: 683–691.

43. Maykowska A, Sobczak K, Wójcieszewska M, Kołodzki P, Krzyżosiak WJ (2011) CAG repeats mimic CUG repeats in the misregulation of alternative splicing. Nucleic Acids Res 39: 8938–8951.

44. Houseley JM, Wang Z, Brock GJ, Soloway J, Artero R, et al. (2005) Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind and deregulated alternative splicing in neurons. Hum Mol Genet 14: 873–883.

45. Fiszer A, Krzyżosiak WJ (2013) RNA toxicity in polyglutamine disorders: concepts, models, and progress of research. J Mol Med (Berl) 91: 683–691.

46. Romani S, Jimenez F, Hoch M, Patel NH, Taubert H, et al. (1996) Kruppel, a Drosophila segmentation gene, participates in the specification of neurons and glial cells. Dev Biol 60: 95–107.

47. Bermingham JR, Jr., Martinez-Arias A, Petit MG, Scott MP (1990) Different patterns of transcription from the two Antennapedia promoters during Drosophila embryogenesis. Development 109: 553–566.

48. Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, et al. (2010) Identification of functional elements and regulatory circuits by Drosophila modENCODE. Genome Res 20: 2418–2428.

49. Awgulewitsch A, Jacobs D (1992) Deformed autoregulatory element from Drosophila muscleblind. Gene 13: 3079–3088.

50. Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, et al. (2010) Identification of functional elements and regulatory circuits by Drosophila modENCODE. Genome Res 20: 2418–2428.