Musculin and TCF21 coordinate the maintenance of myogenic regulatory factor expression levels during mouse craniofacial development

Natalia Moncaut[1], Joe W. Cross[1], Christine Siligan[1,*], Annette Keith[1], Kevin Taylor[1,‡], Peter W. J. Rigby[1] and Jaime J. Carvajal[1,2,§]

SUMMARY

The specification of the skeletal muscle lineage during craniofacial development is dependent on the activity of MYF5 and MYOD, two members of the myogenic regulatory factor family. In the absence of MYF5 or MYOD there is not an overt muscle phenotype, whereas in the double *Myf5;MyoD* knockout branchiomeric myogenic precursors fail to be specified and skeletal muscle is not formed. The transcriptional regulation of *Myf5* is controlled by a multitude of regulatory elements acting at different times and anatomical locations, with at least five operating in the branchial arches. By contrast, only two enhancers have been implicated in the regulation of *MyoD*. In this work, we characterize an enhancer element that drives *Myf5* expression in the branchial arches from 9.5 days post-coitum and show that its activity in the context of the entire locus is dependent on two highly conserved E-boxes. These binding sites are required in a subset of *Myf5*-expressing cells including both progenitors and those which have entered the myogenic pathway. The correct levels of expression of *Myf5* and *MyoD* result from activation by musculin and TCF21 through direct binding to specific enhancers. Consistent with this, we show that in the absence of musculin the timing of activation of *Myf5* and *MyoD* is not affected but the expression levels are significantly reduced. Importantly, normal levels of *Myf5* expression are restored at later stages, which might explain the absence of particular muscles in the *Msc;Tcf21* double-knockout mice.

KEY WORDS: Craniofacial, Msc, Myf5, MyoD, Myogenesis, Tcf21

INTRODUCTION

During embryonic development, precursor cells of four different mesodermal origins follow distinct routes to form skeletal muscles (reviewed by Carvajal and Rigby, 2010). These progenitors can activate the myogenic program in situ, before or after initiating migration to their final position. Muscles from the trunk derive from segmented mesoderm. The progenitor cells from the most anterior somites will form muscles of the tongue and larynx. Whereas the extraocular muscles originate from cranial mesodermal cells located in the periocular region, head mesodermal cells that migrate into the branchial arches (BAs) will give rise to the muscles of mastication (derived from the first BA) and of facial expression (derived from the second BA) (reviewed by Noden and Francis-West, 2006). It has recently been shown that splanchnic mesodermal precursors contribute to some lower jaw muscles (Nathan et al., 2008).

Irrespective of their mesodermal origin, the determination, specification and differentiation of the myogenic phenotype are driven by the activation of a regulatory cascade involving the four myogenic regulatory factors (MRFs): MYF5, MYF6 (MRF4), MYOD (MYOD1 – Mouse Genome Informatics) and myogenin. Consistent with the distinct embryological origins of skeletal muscles, the genetic hierarchies operating to activate the myogenic cascade are different in facial, extraocular and trunk muscles (Sambasivan et al., 2009). The first evidence for this came from the analysis of the double *Pax3;Myf5(Mrf4)* knockout (KO) animals, in which *Mrf4* expression is also disrupted in cis. These mutant mice lack all trunk muscles but show no abnormalities in craniofacial musculature (Tajbakhsh et al., 1997). In addition, in double *Myf5;MyoD* KO embryos in which *Mrf4* expression is not compromised, limb and facial muscles fail to develop, whereas some trunk musculature is present (Kassar-Duchossoy et al., 2004). Head and trunk myogenic programs also exhibit different outcomes in response to individual signaling molecules. For example, whereas trunk myogenesis is promoted by the action of WNT and inhibited by BMPs, antagonists to these molecules have the same effect in cranial myogenesis (Tzahor et al., 2003). The timing of muscle differentiation in the head is also distinct, taking place later than in the trunk (Noden et al., 1999).

Several studies have identified factors controlling the myogenic cascade specifically in the head: TBX1 has been shown to regulate the onset of branchiomeric myogenesis and its inactivation results in the sporadic development of first arch-derived muscles (Kelly et al., 2004). Inactivation of *Pitx2* results in increased cell death in the arch mesodermal cores and loss of musculin (*Msc* or *MyoR*)-positive cells in the first BA (Dong et al., 2006; Shih et al., 2007). MSC and TCF21 (epicardin/capsulin/POD-1) are basic helix-loop-helix (bHLH) transcription factors, known to be markers of undifferentiated muscle precursor cells (von Scheven et al., 2006). Although mice lacking either *Msc* or *Tcf21* show normal facial musculature development, the compound *Msc;Tcf21* mutant fails to

[1]Division of Cancer Biology, The Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, UK. [2]Centro Andaluz de Biologia del Desarrollo, CSIC-Universidad Pablo de Olavide-Junta de Andalucía, Ctra Utrera km1, 41013 Sevilla, Spain.

*Present address: Institute of Biophysics, University of Linz, Altenbergerstr 69, A-4040 Linz, Austria
‡Present address: Australian BioResources, PO Box 322, Moss Vale, NSW 2577, Australia
§Author for correspondence (j.carvajal@csic.es)

activate MRF expression in the first BA and the major muscles of mastication are missing (Lu et al., 2002). The analyses of the mutants mentioned above show a clear distinction in the regulation of myogenesis and survival of first and second arch-derived muscles, suggesting that these events might be differentially regulated.

In order to characterize fully the regulatory networks involved in skeletal myogenesis, we and others have embarked on the characterization of regions involved in the transcriptional regulation of the MRFs. The identification of cognate transcription factors will ultimately result in the naming of the main activators of the myogenic cascade. At the genomic level, *Mrf4* and *Myf5* are in close proximity, being separated by 8.7 kb in mouse. We have previously shown that a bacterial artificial chromosome (BAC) carrying 195 kb upstream of the *Myf5* translational start site is able to fully recapitulate the expression patterns of both genes (Carvajal et al., 2001). Within the genomic interval encompassed by this BAC, a large number of regulatory elements acting on *Mrf4* or *Myf5* have been identified (Buchberger et al., 2003; Carvajal et al., 2001; Carvajal et al., 2008; Fomin et al., 2004; Hadchouel et al., 2000; Hadchouel et al., 2003; Patapoutian et al., 1993; Summerbell et al., 2000; Zweigerdt et al., 1997). Although only a few of these elements are required to recapitulate the endogenous pattern of *Mrf4* expression (Carvajal et al., 2001), many of them are involved in the regulation of *Myf5*. Some of them, like the early epaxial enhancer (Summerbell et al., 2000; Teboul et al., 2002) are modular in nature and do not require the action of further enhancers for activity. Others, by contrast, require the input of additional enhancers in order to establish the full expression pattern in a particular subset of muscle progenitors (J.J.C. and P.W.J.R., unpublished data). This complex array of regulatory elements is able to interpret the different networks of signals from surrounding embryonic tissues and integrate them to give rise to the skeletal muscle phenotype.

Using deletion series of plasmid and BAC clones in transgenic experiments, five separate elements have been shown to be involved in the control of *Myf5* expression in BA progenitors (Carvajal et al., 2001; Patapoutian et al., 1993; Summerbell et al., 2000): (1) an intragenic element, located in intron 1 of *Myf5*, drives weak arch expression from 9.0 days post-coitum (dpc) to 10.0 dpc; (2) two distal elements, situated in the −59 kb to −23 kb and the −88 kb to −81 kb intervals, are required for consistent second and first BA expression, respectively; (3) an element overlapping the early epaxial enhancer acts selectively on the second BA from 9.5 dpc to 10.5 dpc; and finally, (4) the proximal arch element (PAE), located in the −1824 bp to −715 bp interval upstream of *Myf5*, drives second BA expression from 9.0 dpc and first BA from 10.0 dpc, and is required for the maintenance of expression in both arches until 12.5 dpc (Carvajal et al., 2008). This multiplicity of regulatory elements and expression outputs presumably reflects the highly complex signaling environment behind craniofacial muscle development.

The regulation of *MyoD* in the BAs is simpler, with two enhancers identified: the distal regulatory region (DRR) (Asakura et al., 1995; Tapscott et al., 1992) and the core enhancer (CE) (Chen et al., 2001; Faerman et al., 1995; Goldhamer et al., 1995; Kucharczuk et al., 1999), located in the −6.5 kb to −2.5 kb and the −22 kb to −18 kb intervals upstream of *MyoD*, respectively. In isolation, the CE drives faint expression in the first BA from around 9.0 dpc and also in the second BA from 10.5 dpc, although this expression is not as strong as that driven by a larger construct containing the −22 kb to 0 kb *MyoD* interval (construct −24*lacZ*) (Chen et al., 2001). The same base construct from which the CE has been removed (construct −24ΔF3*lacZ*) (Chen et al., 2001) is

able to drive early expression (9.5 dpc to 10.5 dpc) in both BAs and normal expression levels at later stages (11.5 dpc to 12.5 dpc). By contrast, the DRR in isolation only drives BA expression from 12.0 dpc, expression that seems mainly restricted to the first BA (Asakura et al., 1995). Interestingly, deletion of the CE (Chen and Goldhamer, 2004) or the DRR (Chen et al., 2002) from the genome by targeted mutagenesis reveals another layer of complexity as in the absence of the CE *MyoD* arch expression is abolished at 9.5 dpc, very faintly recovered at 10.5 dpc and fully recovered from 11.5 dpc; in the absence of the DRR the timing of *MyoD* expression is not affected, but levels of expression in the BAs are reduced. These results indicate that the CE is essential for the early activation of the gene in the BAs and that the two enhancers cooperate in order to generate the correct expression levels for *MyoD*. Importantly, analyses of the two targeted alleles also reveal that neither of them is required for late BA expression, opening the possibility that another, yet uncharacterized, element is involved in this aspect of the expression.

In order to gain a deeper understanding of the regulation of the MRFs during branchiomeric myogenesis, we have now focused on the transcriptional regulation of *Myf5* and *MyoD* during early developmental stages. By fine-scale analysis of the *Myf5* PAE enhancer, we have identified two highly conserved E-boxes, both of which are essential for enhancer activity. By using chromatin immunoprecipitation (ChIP) analyses, we show that MSC and TCF21 bind to this enhancer in vivo, our interpretation being that this occurs through the two essential E-boxes. These binding sites are required in a subset of the expressing cells including both progenitors and those which have entered the myogenic pathway. Furthermore, we also show that these two transcription factors also bind in vivo to specific *MyoD* regulatory regions. Analyses of RNA levels show that in the absence of MSC, *Myf5* and *MyoD* expression levels in the BAs are reduced at early stages, whereas *Tbx1* transcript levels are significantly increased. Normal *Myf5* and *Tbx1* levels are recovered at later stages whereas *MyoD* remains downregulated. These results reveal the fine mechanism that operates in the spatiotemporal regulation of branchiomeric muscle development through the independent control of activation and maintenance of expression levels of the MRFs. We demonstrate the existence of a direct link between *Msc/Tcf21* and *Myf5/MyoD* and provide an insight into the causes behind the absence of some craniofacial muscles in the double *Msc;Tcf21* mutant animals.

MATERIALS AND METHODS
Comparative analysis of genomic sequences
The genomic sequences of eight species (~300 kb upstream and ~50 kb downstream of *Myf5* promoter) were submitted to the VISTA genome browser (http://genome.lbl.gov/vista) and aligned using the MLAGAN software tool. The mouse sequence was used as the base genome and conservation parameters were set as follows [common name (% of conservation identity; minimum conservation window)]: human (70%, 100 bp); chimpanzee (70%, 100 bp); cow (70%, 100 bp); opossum (70%, 100 bp); chicken (70%, 100 bp); frog (40%, 200 bp) and zebrafish (30%, 300 bp).

Modification of BAC clones by homologous recombination
We used the linear recombination method (Swaminathan et al., 2001) with the changes previously described (Carvajal et al., 2008) to modify the BAC195APZ construct (Carvajal et al., 2001). To remove the proximal arch element from B195APZ, a deletion cassette containing homology arms 140 bp upstream and downstream of the PAE was used. These homology arms were amplified using the following primers: 5′ homology arm: 5AE-F 5′-CCAATAGGGTACAGTTGGATTTTTCATTG-3′ and 5AE-R 5′-<u>TTCTGATGAGAGAGATGG</u>CTGTGTCTGTCTCTGTTCC-

CTCTACCCTGTC-3′; 3′ homology arm: 3AE-F 5′-CCATCTCTCTCA-TCAGAACTTTAC-3′ and 3AE-R 5′-AGTCAGTTTGTGTACTGGTA-ACAG-3′ and joined by fusion PCR using primers 5AE-F and 3AE-R (the underlined sequence in primer 5AE-R shows the overlap between the homology arms). The B195APZ$^{ECR-1mBS34}$ construct was generated using a cassette with the mutated E-boxes amplified with the following primers: ECR-1mBS34_F 5′-CTTAGCTGGttataaGGACCgtatcaTAACGGTG-3′ and ECR-1mBS34_R 5′-CAGTTTCTCACAGAAAGCTCCCCGTCCAG-3′. Lower case sequence indicates mutated bases.

Generation of transgenic mice
All in vivo experiments were performed according to United Kingdom Home Office Regulations.

BAC DNA was prepared using the QIAgen maxiprep kit (QIAGEN) as previously described (Carvajal et al., 2001). After dialysis against microinjection buffer (10 mM Tris-HCl pH 7.5, 0.1 mM EDTA pH 8.0 and 100 mM NaCl), DNA was diluted to 1.5 ng/µl in the same buffer and was used to inject fertilized mouse eggs from CBA/Ca × C57Bl/6 crosses as previously described (Yee and Rigby, 1993). $Msc^{tm1Eno}/Msc^{tm1Eno}$ ($Msc^{-/-}$) mice have been previously described (Lu et al., 2002) and were obtained from The International Mouse Strain Resource at The Jackson Laboratory.

Preparation of plasmid constructs
ECR-1 was cloned into a plasmid containing the *Myf5* minimal promoter as previously defined (Summerbell et al., 2000) followed by the *nlacZ* reporter gene to generate ECR-1-MZ. Mutations of ECR-1 binding sites were generated using specific primers for each sequence (supplementary material Table S1) and a somitic enhancer (ECR-111) (Ribas et al., 2011) was included as a positive control for transgenesis.

Histochemical staining and sectioning
Embryos were fixed overnight in Mirsky's fixative (National Diagnostics) at 4°C, washed three times in washing solution [PBS, 0.02% (v/v) Tween-20] for 20 minutes at room temperature and placed in β-GAL staining solution [5 mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6$, 2 mM $MgCl_2$, 0.02% (v/v) Tween-20 and 0.4 mg/ml X-gal in PBS] for 2-20 hours depending on the embryo stage. Embryos were post-fixed in Mirsky's fixative. For sectioning, embryos were embedded in 2% (w/v) agarose (SeaKem Gold, Lonza), fixed in Mirsky's fixative overnight and sectioned using a Leica VT1000S vibratome at 70 µm.

Immunohistochemistry
Mouse embryos were fixed in 2% (w/v) paraformaldehyde/0.05% (v/v) Triton X-100 for 20 minutes at room temperature, washed in PBS, passed through a sucrose gradient and embedded in OCT (Lamb). 12 µm cryostat sections were rinsed in PBS and permeabilized in PBS/0.05% (v/v) Triton X-100 for 20 minutes at room temperature. After 30 minutes incubation in blocking solution [PBS/3% (w/v) BSA/0.05% (v/v) Triton X-100/3% (v/v) goat serum], sections were incubated with primary antibodies overnight at 4°C. The antibodies used in this study were anti-β-galactosidase (1:400, Abcam ab9361), anti-MYOD (1:400, Dako M3512), anti-desmin (1:400, Dako IR606) and anti-myosin (1:800, DSHB MF20). Finally, sections were incubated with secondary antibodies (Molecular Probes, 1:1000) Alexa488-conjugated anti-mouse IgG and Alexa555-conjugated anti-chicken IgG. Immunofluorescence images were acquired with a Leica Microsystems TCS-SP2 and processed using Photoshop CS (Adobe).

Electrophoretic mobility shift assays (EMSAs) and supershifts
EMSA experiments were performed as previously described (Gilthorpe et al., 2002). Nuclear protein extracts were prepared from 10.5 dpc mouse embryos. Briefly, 20-30 embryos were collected and homogenized in hypotonic buffer (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.2 mM PMSF, 0.5 mM DTT). After centrifugation the nuclei were resuspended in low-salt buffer [20 mM HEPES pH 7.9, 25% (v/v) glycerol, 1.5 mM $MgCl_2$, 0.02 M KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT], transferred to high-salt buffer [20 mM HEPES pH 7.9, 25% (v/v) glycerol, 1.5 mM $MgCl_2$, 1.2 M KCl, 0.2 mM EDTA, 0.2 mM PMSF] and centrifuged. The nuclear extract (supernatant fraction) was stored at –20°C until use. For supershift experiments, an antibody against MSC (Santa Cruz

sc-9556) was added to the reaction 1 hour before addition of the biotin-labeled probe (5′-TAGCTGGCAGCTGGGACCCATCTGTAACGGT-GTTA-3′).

Chromatin immunoprecipitation
Chromatin immunoprecipitation (ChIP) assays were performed using isolated first and second BAs and midbrain from 10-15 10.5-dpc mouse embryos. Briefly, tissues were dissected in PBS and fixed in 1% (w/v) formaldehyde. After tissue homogenization, samples were sonicated and immunoprecipitated using antibodies against MSC and TCF21 (Abcam, ab49475) or control IgG prebound to Dynabeads Protein A (Invitrogen). Immunoprecipitated DNA was purified and PCR-amplified using specific primers (supplementary material Table S1).

Quantitative RNA analysis
Total RNA was extracted from first and second BAs using Tri Reagent (Sigma-Aldrich) according to the manufacturer's protocol. cDNA was synthesized with random primers using Superscript II (Invitrogen), amplified with specific primer pairs (supplementary material Table S1) and quantified by quantitative PCR (qPCR) using SyBr GreenER (Invitrogen).

RESULTS
An evolutionarily conserved region within the proximal arch element drives early *Myf5* expression in the branchial arches
We and others have shown previously that several elements, some of which have been defined as classical enhancers, regulate *Myf5* expression in the branchial arches (BAs) (Carvajal et al., 2001; Patapoutian et al., 1993; Summerbell et al., 2000). These elements are contained within the B195APZ BAC construct (Carvajal et al., 2001), which recapitulates the expression patterns of the *Mrf4* and *Myf5* genes during development. In order to gain an understanding of the contribution of these enhancers to the full facial pattern of *Myf5* expression, we deleted the proximal arch element (PAE) from B195APZ (Fig. 1D-F; construct B195APZ$^{\Delta PAE}$, $n=4$ stable lines). At early stages, B195APZ$^{\Delta PAE}$ drives faint expression in both BAs (Fig. 1D,E), although the reduction in expression levels is more marked in the second BA compared with the control B195APZ lines (Fig. 1A-C). At 11.5 dpc, expression in derivatives from the ventral branch of the second arch is absent and faint expression in the dorsal branch is visible (Fig. 1F, black arrowhead).

In order to refine the regulatory sequences responsible for *Myf5* expression within the 1.1 kb interval defined as the PAE, we used the VISTA bioinformatic tool (Dubchak and Ryaboy, 2006) and identified a 285 bp evolutionarily conserved region (ECR-1), conserved in all mammals and in chicken (Fig. 1J; supplementary material Fig. S1). To determine whether this ECR is able to recapitulate the expression pattern driven by the full PAE, we cloned it upstream of the *Myf5* minimal promoter and an *nlacZ* reporter gene (construct ECR-1-MZ). In isolation (*n*=9 transient transgenics), ECR-1 is able to drive early 9.5 dpc transgene expression in the second BA (Fig. 1G), coinciding with the previous description of the pattern driven by the PAE (Construct #11) (Summerbell et al., 2000). At 10.5 dpc, faint expression can be detected in the first BA (Fig. 1H). Transgene expression is maintained through 11.5 dpc, when the muscle precursor mass from the second BA divides into dorsal and ventral components (Fig. 1I, black arrowheads). At this stage, a subset of the expression in derivatives from the first BA can be observed (Fig. 1I, white arrowhead). From 12.5 dpc onwards, transgene expression is downregulated in both arches and only faint expression can be detected in the external ear musculature, derived from the dorsal component of the second BA (data not shown). Together, these data indicate that ECR-1 recapitulates the expression pattern driven by the entire enhancer region previously characterized as the PAE.
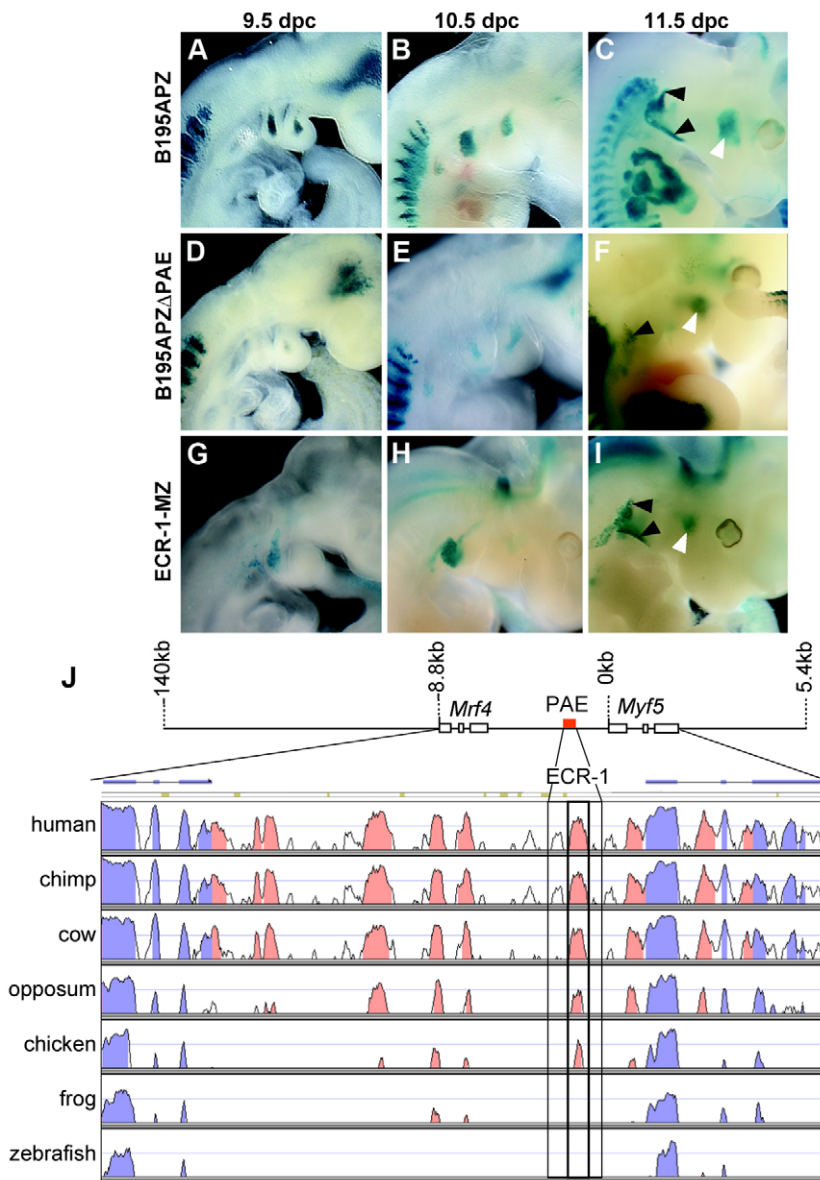
Fig. 1. The evolutionarily conserved region 1 (ECR-1) within the PAE enhancer drives *Myf5* expression in the branchial arches. (**A-I**) Branchial arch expression patterns of transgenic mouse embryos carrying B195APZ (A-C), B195APZ$^{\Delta PAE}$ (D-F) and ECR-1-MZ (G-I) stained for β-galactosidase activity at different developmental stages: 9.5 dpc (A,D,G), 10.5 dpc (B,E,H) and 11.5 dpc (C,F,I). White arrowheads indicate first BA-derivatives; black arrowheads indicate second BA-derivatives. (**J**) Schematic representation of the evolutionarily conserved regions between the *Mrf4* and *Myf5* genes determined by the Vista genome browser software followed by MLAGAN alignment after sequence comparison of eight vertebrate species. Blue and red peaks correspond to conserved coding regions (exons) and non-coding regions, respectively.

ECR-1 enhancer activity depends on two highly conserved E-boxes

In order to characterize the sequences necessary for the activity of the ECR-1, we used TFSEARCH software (Heinemeyer et al., 1998) as a tool to predict putative transcription factor binding sites (BS) within this regulatory element. We identified two putative caudal homeobox-BS (CDX-BS) in the overlapping region between oligonucleotides 1 and 2 (BS1 and BS2), two conserved E-boxes in oligonucleotide 8 (BS3 and BS4) and CDX- and heat shock factor-BS (HSF-BS) within oligonucleotide 10 (BS5 and BS6, respectively) (the full sequence of ECR-1 and the positions of the oligonucleotides are shown in supplementary material Fig. S1). Next, we assessed whether these binding sites are required for the activity of ECR-1 in vivo. For this purpose, we mutated combinations of these binding sites (Fig. 2A) and analyzed the ability of these constructs to drive BA expression in transgenic embryos at 10.5 dpc. To distinguish negative embryos due to integration site effects from those due to the introduced mutations, we used ECR-111 (Ribas et al., 2011), which drives early somitic expression, as a positive control for transgene expression (Fig. 2B,C). The combined mutation of BS1, BS2, BS5

and BS6 does not affect ECR-1 enhancer function in the BAs (Fig. 2D; construct 111-ECR-1^{mBS1256}-MZ, *n*=4), whereas double mutation of BS3 and BS4 (E-boxes; construct 111-ECR-1^{mBS34}-MZ, *n*=3) completely abolishes transgene expression in the BAs (Fig. 2E). Individual mutation of the two binding sites also abolishes reporter expression (Fig. 2F,G; 111-ECR-1^{mBS3}-MZ, *n*=6; 111-ECR-1^{mBS4}-MZ, *n*=4), indicating that both E-boxes are essential for ECR-1 enhancer function.

To assess the requirement for the two E-boxes in the context of the entire *Mrf4/Myf5* locus, we introduced the mutations of BS3 and BS4 into B195APZ (construct B195APZ$^{ECR-1mBS34}$, *n*=5 stable lines). Although at 9.5 dpc B195APZ^{wt} drives transgene expression in both BAs (Fig. 3A,B), expression in the first BA in B195APZ$^{ECR-1mBS34}$ lines is greatly reduced and there is no consistent expression in the second BA (Fig. 3C,D), with some variation in levels seen between littermates, although always greatly reduced compared with B195APZ^{wt} BAC lines. The effects of the double mutation on the expression pattern of *Myf5-nlacZ* are maintained at 10.5 dpc, with severely reduced expression levels in both arches (compare Fig. 3E,F with 3G,H). By 12.5 dpc, the
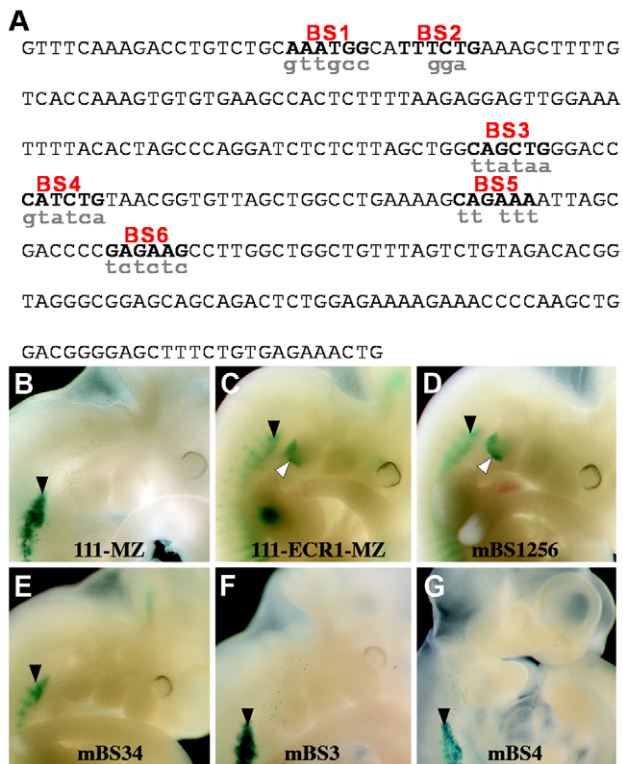
Fig. 2. The expression of *Myf5* in the branchial arches depends on two conserved E-boxes. (**A**) The 285 bp mouse sequence corresponding to ECR-1 showing the putative binding sites (in bold) as identified with the TFSEARCH software and the mutations introduced to disrupt putative binding sites (in gray). (**B-G**) Individual or grouped binding sites were functionally tested by site-directed mutagenesis using the 111-ECR-1-MZ construct (D-G). 10.5 dpc transgenic embryos carrying 111-MZ (B), 111-ECR-1-MZ (C), 111-ECR-1^{mBS1256}-MZ (D), 111- ECR-1^{mBS34}-MZ (E), 111-ECR-1^{mBS3}-MZ (F) and 111-ECR-1^{mBS4}-MZ (G) were stained for β-GAL activity. Mutations in BS3 and BS4 singly (F,G) or combined (E) abolish *nlacZ* expression in the branchial arches. The 111-MZ drives consistent *nlacZ* expression in the cervical and thoracic somites, irrespectively of the ECR-1 mutations (C-G, black arrowheads). White arrowheads indicate maintained reporter expression in the BAs.

double mutation affects expression in muscles derived from both BAs although it is more severely affected in BA2-derived muscles such as the precursors of the auricularis and the zygomaticus muscles (compare Fig. 3I,J with 3K,L). Importantly, the point mutations do not have an effect on the expression of the transgene in the extraocular muscles, which are not derived from the BAs. By 13.5 dpc, expression can be seen in both facial and mastication muscles although transgene activation in some BA1-derived muscles, such as the temporalis, is slightly delayed (compare Fig. 3M,N with 3O,P). These data indicate that the two E-boxes constitute the core of this enhancer.

The effect of the BS34 mutation is seen in a subset of cells expressing *Myf5*

In order to determine whether the effects of the BS34 mutation are specific to particular myogenic cell populations, we performed β-GAL and MYOD double staining in B195APZ^{wt} and B195APZ$^{ECR-1mBS34}$ lines at 10.5 dpc. In wild-type lines we observe at least three myoblast populations in the core of the BAs

at the time examined: those expressing only β-GAL (MYF5+/MYOD–), those expressing only MYOD (MYF5–/MYOD+) and a smaller population that co-expresses both proteins (MYF5+/MYOD+) (Fig. 4A,B). There are also cells that express neither protein but we cannot comment on whether these are destined to become myogenic at later times. The BS34 mutation results in a significant reduction in the number of β-GAL-positive cells in the cores of both BAs, and it affects both the MYF5+/MYOD+ and the MYF5+/MYOD– populations (Fig. 4C,D). The majority of the cells expressing β-GAL are precursors committed to the myogenic lineage as they also express the myogenic marker desmin (Fig. 4E,F) but they are not yet fully differentiated as they do not express myosin (supplementary material Fig. S2). The effect of the BS34 mutation is seen in both the desmin+ and desmin– populations, indicating that these binding sites are required both in progenitors and in cells which have entered the myogenic pathway. There are some cells in the BS34 lines that continue to express MYF5. We assume that in these cells expression depends upon one or more of the other arch enhancers that we have identified.

MSC and TCF21 bind to ECR-1 and to MyoD regulatory elements in vivo

It has been shown previously that *Msc* and *Tcf21* are transiently expressed in the mesodermal cores of the BAs during development and that the compound mutant of these genes lacks the major muscles of mastication (Lu et al., 2002; Robb et al., 1998). As both encode bHLH proteins, which are known to bind E-boxes, we considered them to be good regulatory candidates upstream of *Myf5* and *MyoD* in the BAs. EMSA analysis of oligonucleotide 8, containing the two E-boxes essential for enhancer activity, shows several bands when using nuclear extract from 10.5 dpc embryos (Fig. 5A). One of these shifts is lost following incubation with an antibody against MSC (Fig. 5A, arrowhead), indicating that MSC binds to oligonucleotide 8. We were not able to observe any supershifts when using a TCF21 antibody (data not shown). Next, we tested whether MSC as well as TCF21 bind this regulatory element in the BAs in vivo by using chromatin immunoprecipitation (ChIP). A region located at position +5 kb, not containing any identifiable E-box, and chromatin extracted from the midbrain were used as negative controls. Anti-MSC and anti-TCF21 antibodies efficiently immunoprecipitate the ECR-1 enhancer (Fig. 5B) from chromatin obtained from isolated BAs of 10.5 dpc embryos but not from the negative region at +5 kb or the control midbrain chromatin (supplementary material Fig. S3A,B). Together, these results demonstrate that MSC and TCF21 bind specifically to the ECR-1 enhancer element in vivo at this developmental stage.

In the trunk, inactivation of *MyoD* results in the upregulation of *Myf5* (Rudnicki et al., 1992) and vice versa (Braun et al., 1994); this might explain the absence of phenotype in the individual KOs as one can act to rescue the deficiency of the other. Indeed, in the double *Myf5;MyoD* knockout (Rudnicki et al., 1993), there is no such rescue and the animals lack all adult skeletal muscles. In the double *Msc;Tcf21* knockout animals, specific branchiomeric muscles are missing at later stages, suggesting that the expression of both *Myf5* and *MyoD* could be compromised as rescue does not take place. Therefore, we investigated whether MSC and TCF21 could also regulate the expression of *MyoD*. Three E-box-containing regulatory regions for *MyoD* have been described: the DRR (Tapscott et al., 1992), the CE (Goldhamer et al., 1995) and the proximal regulatory region (PRR, including the promoter
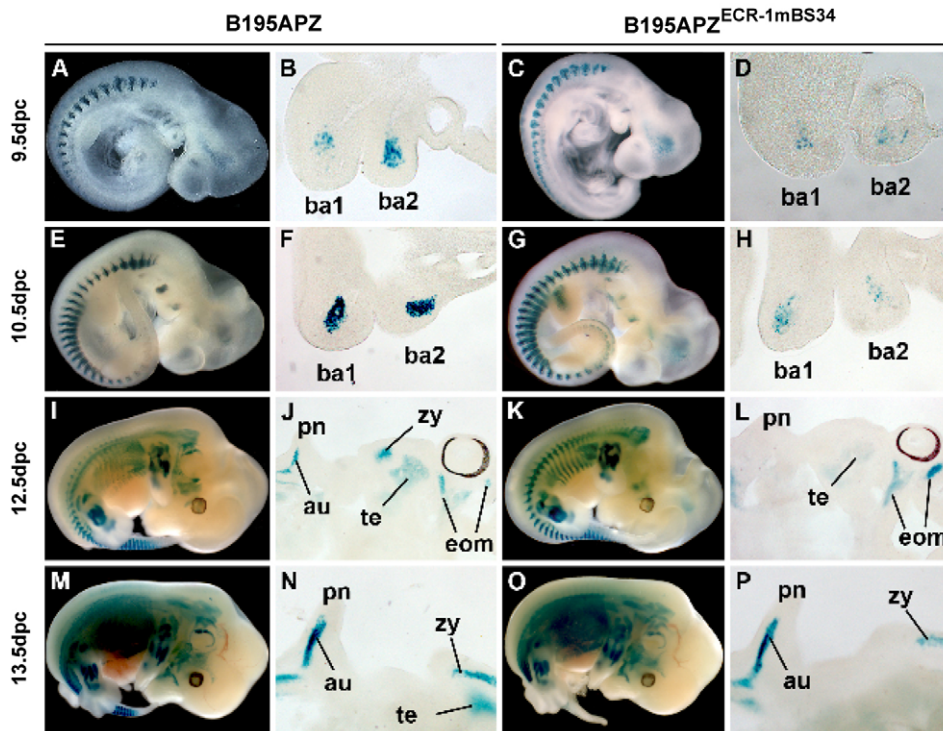
Fig. 3. Mutation of the E-boxes in the context of the *Mrf4-Myf5* locus severely compromises *Myf5* expression in the branchial arches. (**A-P**) Transgenic embryos carrying the B195APZ (A,B,E,F,I,J,M,N) and B195APZ[ECR-1mBS34] (C,D,G,H,K,L,O,P) BAC constructs were compared at different stages during development: 9.5 dpc (A-D), 10.5 dpc (E-H), 12.5 dpc (I-L) and 13.5 dpc (M-P). The combined BS3 and BS4 mutation shows that in the context of the entire locus these binding sites are necessary to drive consistent reporter expression in the branchial arches and muscle derivatives at early stages. au, auricularis muscle; eom, extraocular muscles; pn, pinna of the ear; te, temporalis muscle; zy, zygomaticus muscle.

region) (Tapscott et al., 1992). Using MSC and TCF21 antibodies and chromatin from isolated 10.5 dpc BAs, we were able to immunoprecipitate the PRR and the DRR but not the CE (Fig. 5B). These results show that MSC and TCF21 are loaded in vivo in regulatory regions that are responsible for driving expression of both *Myf5* and *MyoD* in the BAs.

MSC regulates MRF and *Tbx1* expression in the branchial arches

Next, we tested whether the physical interaction of MSC with the regulatory regions of *Myf5* and *MyoD* has a physiological output in the levels of these MRFs. Although in situ hybridization (ISH) does not reveal any obvious differences in the expression patterns of *Myf5* or *MyoD* between $Msc^{-/-}$, $Msc^{+/-}$ and wild-type embryos (supplementary material Fig. S4), qPCR analysis shows significant changes in the expression levels of these genes in the BAs. In the absence of MSC, there is a 50% decrease in the levels of *Myf5* and *MyoD* expression at 9.5 dpc (Fig. 6A). Reduction in *Myf5* expression is partially restored by 10.5 dpc (~70% of the wild type; Fig. 6B) whereas *MyoD* expression remains downregulated. Although there seems to be an increase in the expression levels of *Tcf21* and *Pitx2* in the absence of MSC at 9.5 dpc, this difference is not statistically significant, suggesting that the absence of MSC is not necessarily rescued by an increase of *Tcf21* or *Pitx2*. Remarkably, in the absence of MSC, *Tbx1* transcript levels are significantly higher at 9.5 dpc compared with wild-type embryos but return to wild-type values by 10.5 dpc.

In order to test whether MYF5 could be responsible for autoregulating its expression in the BAs, we analyzed *lacZ* expression levels driven by B195APZ in hetero- and homozygous knockout embryos for *Myf5* (Kaul et al., 2000) but no changes in expression levels were observed, indicating that MYF5 does not participate in early autoregulation in the BAs (supplementary material Fig. S5A). *MyoD* transcript levels in the arches were not

changed in $Myf5^{-/-}$ embryos at the developmental stage analyzed (supplementary material Fig. S5B). We also analyzed the endogenous *Myf5* expression levels in the BAs in the absence of MYOD and found it to be upregulated twofold in $MyoD^{-/-}$ embryos, supporting the idea that an increase of *Myf5* levels acts to rescue the *MyoD* KO phenotype (supplementary material Fig. S5C). No other transcription factor gene analyzed (*Pitx2*, *Msc*, *Tbx1* or *Tcf21*) was found to be affected by the loss of one or both alleles for *MyoD* (supplementary material Fig. S5D-G).

DISCUSSION
Diverse environmental signals regulate expression of the MRFs at different locations in the embryo during development generating transcriptional cascades that orchestrate muscle formation. Interestingly, although the genes essential for specification of premyogenic progenitors in the trunk and the head are different, the myogenic cascade that ultimately forms mature muscle, comprising the ordered activation of MYF5, MYOD, MRF4 and myogenin, appears to be quite similar in head and trunk skeletal muscle. It has been shown that PAX3 (Bajard et al., 2006) and members of the TEAD family (Ribas et al., 2011) regulate *Myf5* expression by binding directly to specific enhancers. In the head, genetic studies have revealed that TBX1, PITX2, MSC and TCF21 are major players upstream of the myogenic determination genes. Although in $Pitx2^{-/-}$ (Dong et al., 2006), $Tbx1^{-/-}$ (Kelly et al., 2004) and $Msc^{-/-};Tcf21^{-/-}$ mutants (Lu et al., 2002) *Myf5* and *MyoD* expression levels are downregulated in the BAs, a direct interaction between these factors and individual enhancers for the MRFs has not been shown. In the case of $Pitx2^{-/-}$, which has also been implicated in extraocular muscle specification, direct association with the *Myf5* and *MyoD* minimal promoters has been shown in C2C12 cells and in a mouse extraocular muscle primary cell line (Zacharias et al., 2011). Our data allow us to place MSC (and probably TCF21) directly upstream of both *Myf5* and *MyoD* in the genetic network controlling branchiomeric myogenesis. With regard to the regulation
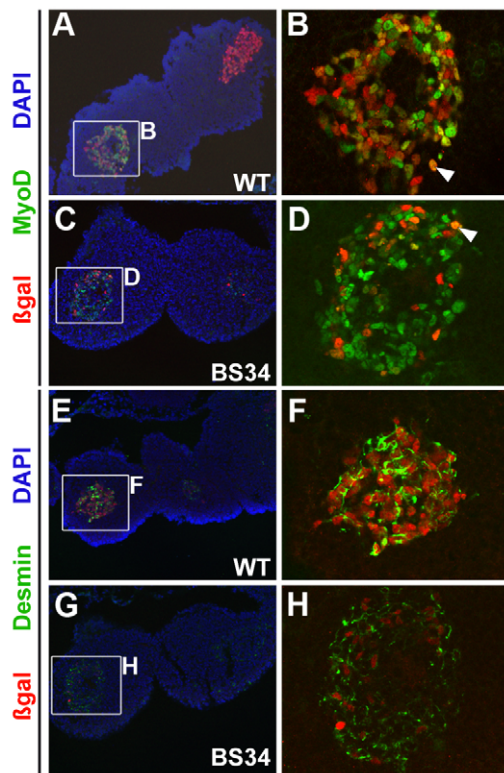
Fig. 4. The effect of the BS34 mutation is seen in a subset of cells expressing *Myf5*. (**A-H**) Sections through the branchial arches of 10.5 dpc B195APZ^{wt} (A,B,E,F) and B195APZ$^{ECR-1mBS34}$ embryos (C,D,G,H) immunostained with anti-β-GAL (A-H), anti-MYOD (A-D) and anti-desmin (E-H) antibodies. At 10.5 dpc, three progenitor cell populations coexist [MYF5$^+$/MYOD$^-$ (red), MYF5$^-$/MYOD$^+$ (green) and MYF5$^+$/MYOD$^+$ (yellow, white arrowheads)] in the mesodermal core of the BAs (B,D). The BS34 mutation results in a reduction of the number of expressing cells but does not specifically affect any of the expressing populations (D,H). Most of the cells in the mesodermal core express desmin (E-H), indicating that the progenitors are committed and the BS34 mutation does not specifically target uncommitted or committed cells (G,H). Note the presence of some MYF5$^+$ cells in animals carrying the BS34 mutant transgene (D,H).

of *Myf5*, we show that ECR-1, located within an experimentally characterized *Myf5* enhancer element (PAE) (Patapoutian et al., 1993; Summerbell et al., 2000), drives *Myf5* expression in the BAs during development and that its function is dependent on two highly conserved E-boxes where MSC and TCF21 bind. With regard to the regulation of *MyoD*, we show that both MSC and TCF21 bind in vivo to two previously defined regulatory regions in the locus. The data also show that at 9.5 dpc, when the MRFs are first expressed in the BAs, both *Myf5* and *MyoD* are downregulated in the absence of MSC, and only *Myf5* expression levels are restored from 10.5 dpc. Interestingly, the onset of expression is not affected, suggesting that MSC plays a crucial role in the maintenance of expression levels of these two MRFs during craniofacial development. Finally, we show the existence of at least three different populations of mesodermal progenitors in the BAs (MYF5$^+$/MYOD$^-$, MYF5$^-$/MYOD$^+$ and MYF5$^+$/MYOD$^+$). This raises the question of whether the rescue of the facial muscle phenotype in the *Myf5*$^{-/-}$ or the *MyoD*$^{-/-}$ mutants takes place by the upregulation of the remaining MRF in all cell precursors or by the expansion of the cell population in which that particular MRF is expressed.

MSC and TCF21 are direct upstream activators of *Myf5* and *MyoD* in branchiomeric myogenesis

We have shown that the bHLH transcription factors MSC and TCF21 regulate the expression of *Myf5* and *MyoD* in the mesodermal core that gives rise to the main branchiomeric muscles. With regard to the regulation of *Myf5*, MSC and TCF21 directly bind ECR-1, presumably occurring through the two essential E-boxes within the enhancer. With regard to the regulation of *MyoD*, we show that both bHLH factors are found bound in vivo to two previously defined regulatory elements, the DDR and the PRR, both of which contain E-boxes (Asakura et al., 1995; Tapscott et al., 1992), although we could not find any evidence of binding to the third characterized *MyoD* element, the CE. Transcriptional regulation might occur through the formation of MSC and/or TCF21 homodimers (Eben-Massari et al., 1998) or heterodimers, either with each other or with other factors such as E-proteins, which have been shown to interact with MSC (Eben-Massari et al., 1998; Lu et al., 1999) and TCF21 (Lu et al., 1998).

MSC has been shown to act as a repressor of the skeletal muscle program in non-myogenic lineages (Yu et al., 2003; Yu et al., 2004). Indeed, MSC strongly represses MYOD-E12 (TCF3 – Mouse Genome Informatics)-mediated activation of an MCK-promoter reporter construct in vitro but fails to activate transcription of the same construct in isolation or in the presence of E12 protein (Lu et al., 1999), suggesting that MSC-E12 competes with MYOD-E12 for the same binding sites or that MSC sequesters the E12 co-activator.

TCF21, which is broadly expressed, also acts as a direct transcriptional repressor in vitro and in vivo (Familari et al., 2009; Funato et al., 2003; Tamura et al., 2001). Interestingly, it has been shown that the activity of TCF21 as an activator or as a repressor is cell type-dependent (Miyagishi et al., 2000) and, although transfection of HeLa cells with *Tcf21* is sufficient to repress the expression of the androgen receptor, this repression is abolished by *Tcf3* co-transfection (Hong et al., 2005), strongly indicating that the TCF21-binding partner has a profound effect on its activity as a repressor or activator.

Nonetheless, in the double *Msc;Tcf21* knockout allele, *Myf5* and *MyoD* fail to be upregulated in the first BA at early developmental stages (Lu et al., 2002), suggesting a transcriptional activation role for both genes. Consistent with these findings, our data indicate that the absence of MSC results in the downregulation of *Myf5* and *MyoD* in the BAs, demonstrating that, at least in branchiomeric myogenesis, MSC (and possibly TCF21) are positive regulators.

MSC controls expression levels of *Myf5* and *MyoD* in the branchial arches

With the exceptions of the element required for expression in the mandibular arch (Carvajal et al., 2001), which remains to be isolated, and the upstream intergenic element (UIE, also known as NAE) (Summerbell et al., 2000), the remaining three *Myf5* arch elements in isolation are able to drive expression from 9.0 dpc (Summerbell et al., 2000) (this work; P.W.J.R. and J.J.C., unpublished). Interestingly, deletion of the PAE from B195APZ does not interfere with the timing of expression but does have an important effect on the expression levels of *Myf5* in the BAs.

This duality in activation and control of expression levels is also observed in the regulation of *MyoD*. The CE in isolation drives *MyoD* expression in first and second BAs from 9.0 dpc and 10.5 dpc, respectively (Faerman et al., 1995; Goldhamer et al., 1995). In the context of a 24 kb construct also containing the DRR and the PRR, expression in the first BA starts at 9.5 dpc
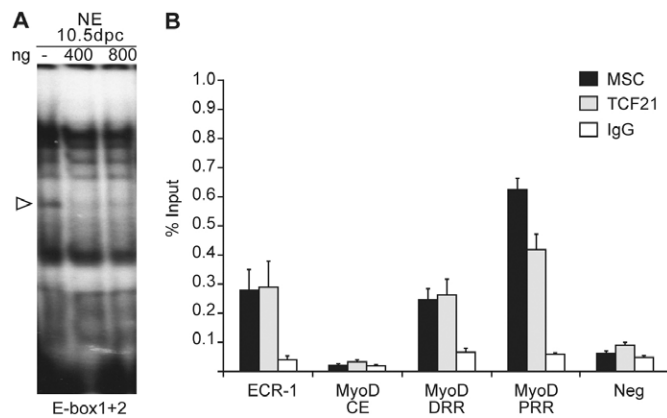
Fig. 5. MSC and TCF21 bind to ECR-1 in vivo. (**A**) Electrophoretic mobility shift assay shows that an antibody against MSC abolishes the shift (arrowhead), indicating that MSC binds to the ECR-1 sequence in vitro. (**B**) Chromatin immunoprecipitation (ChIP) of first and second BA of 10.5 dpc embryos using MSC (black bars), TCF21 (gray bars) antibodies or control IgG (white bars). Specific primers were used to amplify different regulatory regions within the *Myf5* and *MyoD* loci. ChIP analyses indicate that both antibodies immunoprecipitate *Myf5* ECR-1, *MyoD* DRR and *MyoD* PRR, but fail to immunoprecipitate the CE and the negative region. Enrichment is shown as a percentage of input. Error bars represent s.e.m.

Fig. 6. MSC regulates MRF and *Tbx1* expression levels in the branchial arches. (**A**,**B**) *Myf5*, *MyoD*, *Tbx1* and *Pitx2* mRNA levels of first and second BAs were quantified by real-time qPCR at 9.5 dpc (A) and 10.5 dpc (B). Whereas at 9.5 dpc *Myf5* and *MyoD* transcripts are significantly reduced and *Tbx1* levels increased in the Msc$^{-/-}$ (white bars) compared with the wild type (black bars), no significant changes were observed in the levels of *Tcf21* and *Pitx2*. At 10.5 dpc, the expression levels of *Myf5* and *Tbx1* reach those in the wild type whereas *MyoD* remains downregulated. ns, not significant. *$P<0.01$; **$P<0.001$; ***$P<0.0001$. Error bars represent s.e.m. $n=6$.

whereas the timing of activation in the second BA remains at 10.5 dpc (Chen et al., 2001). As these are the correct times of activation for endogenous *MyoD*, this indicates that although the CE is able to interpret myogenic signals at 9.0 dpc in the BAs, in the context of the locus a second element must be restricting the activation until 9.5 dpc.

Deletion of the CE in the context of this 24 kb construct alters the intensity of expression in the BAs, but not the timing of activation (Chen et al., 2001). Interestingly, deletion of the CE from the genome abolishes all early arch expression (therefore altering timing) although normal expression is obtained from 11.5 dpc (Chen and Goldhamer, 2004), indicating that in the context of the entire locus the CE is essential to establish the timing of expression. By contrast, deletion of the DRR from the genome results in a reduction in the expression levels of *MyoD* in the BAs from 9.5 dpc to 10.5 dpc without impacting on timing (Chen et al., 2002), even if in isolation it is only able to drive transgene expression in the BAs from 12.0 dpc. It is important to note that, as in the case with the CE, normal expression levels are attained from 11.5 dpc, suggesting that the late *MyoD* BA expression is driven by a third element in the locus. In summary, these data indicate that: (1) the CE is essential for establishing the correct timing of *MyoD* expression in the BAs; (2) the DRR is essential for establishing the correct levels of expression; (3) another element is also involved in the initiation of early expression; and (4) another (or the latter) element is involved in the regulation of transcription from 11.5 dpc onwards.

We now show that MSC and TCF21 are bound in vivo to both the DRR and the PAE. As deletion of either of these two elements results in a defect in the expression levels of their respective gene without affecting the timing, and because in the absence of MSC, which we show binds in vivo to both elements, the levels of their respective genes are reduced, we propose that MSC (and possibly TCF21) are involved in the establishment of the correct levels of expression of the MRFs during branchiomeric myogenesis.
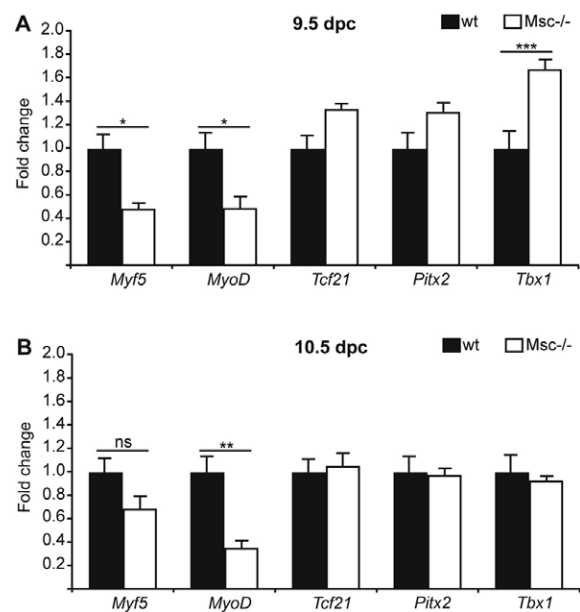
Multiple populations of myogenic precursors in the branchial arches

We show that at 10.5 dpc there are at least three populations of myogenic precursors in the arches: MYF5+/MYOD–, MYF5–/MYOD+ and MYF5+/MYOD+. The BS34 mutation significantly reduces MYF5 expression in both of the populations that normally express the protein. The majority of MYF5+ cells express desmin but not myosin at this time and are thus committed precursors but not yet fully differentiated. The effect of the BS34 mutation is seen in both the desmin+ and desmin– populations, indicating that these binding sites are required in both progenitors and in cells that have entered the myogenic pathway. It is interesting to note that there are some cells in the BS34 lines that continue to express MYF5 at readily detectable levels. The simplest explanation for this observation is that in these cells expression depends upon one or more of the other arch enhancers that we have identified. Whether the enhancer(s) is also regulated by MSC and/or TCF21 remains to be investigated. However, the presence of muscles in the *Msc;Tcf21* KO indicates that these transcription factors might not be key regulators of these enhancers.

Regulatory networks controlling craniofacial development

A regulatory mechanism can be proposed for the branchial arches with both myogenic determination genes, *Myf5* and *MyoD*, being direct downstream targets of MSC and TCF21 (Fig. 7). The incomplete penetrance of the *Tbx1*$^{-/-}$;*Myf5*$^{-/-}$ double mutant phenotype where only a bilateral loss of first BA-derived muscles is observed (Sambasivan et al., 2009), together with the sporadic
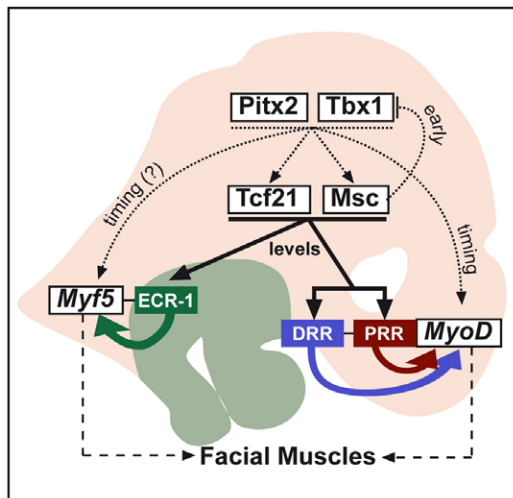
Fig. 7. **Proposed model for the transcriptional regulation of *Myf5* and *MyoD* during arch myogenesis.** In the BAs, PITX2 and TBX1 act as upstream factors in the myogenic cascade regulating the expression of *Msc* and *Tcf21* in a direct or indirect fashion. The expression levels of the MRFs during early BA development are controlled by the direct interaction of MSC and TCF21 with the regulatory regions of *Myf5* (ECR-1) and *MyoD* (DRR and PRR). Although PITX2 and TBX1 could control the timing of activation of expression of *Myf5* and *MyoD*, the control of their expression levels probably takes place through the activation of MSC and TCF21. Dashed arrows represent direct or indirect interactions; solid arrows represent direct interactions.

development of the first arch-derived muscles in the *Tbx1*$^{-/-}$ mouse (Kelly et al., 2004), reveal that parallel pathways are involved in arch myogenesis. PITX2 is a potential candidate to act together with TBX1 in this pathway during arch development (Nowotschin et al., 2006; Shih et al., 2007). Previous data show that in the *Tbx1*$^{-/-}$ as well as in the *Pitx2*$^{-/-}$ mice the levels of *Tcf21* and *Msc* transcripts are decreased whereas we show that *Tbx1* expression is upregulated in the *Msc*$^{-/-}$ mice at early stages. This allows us to propose a regulatory model where MSC (and probably TCF21) is located together with TBX1 and PITX2 upstream of *Myf5* and *MyoD*. In this model, MSC regulates the levels of expression of *Myf5* and *MyoD* by direct interactions through ECR-1 and the DRR, respectively. The low expression levels of *Myf5* and *MyoD* observed in the *Tbx1* and *Pitx2* mutants might be a consequence of a direct downregulation of the two MRFs, the reduction of *Msc* and *Tcf21* levels, or a combination of both (Fig. 7). Low levels of MRF expression might be sufficient to trigger the sporadic development of some craniofacial muscles. By contrast, in the double *Msc;Tcf21* knockout all mastication muscles derived from the first BA are missing, whereas facial expression muscles originating from the second BA and non-mastication muscles from the first BA develop normally, possibly through the arch-specific upregulation of *Myf5* which we have shown to recover its normal expression levels in the *Msc*$^{-/-}$ mice at later stages. These data indicate that, at least in the first BA, the MRF activation by TBX1 and PITX2, if any, is not sufficient to initiate the myogenic cascade. But when only MSC is missing, higher *Tbx1* expression levels might be able to maintain normal levels of *Tcf21* and *Pitx2* and finally re-establish MRF levels (at least for *Myf5*), resulting in a facial musculature without phenotype. In addition, we also show that craniofacial expression of the transgene in the double E-box mutation recovers earlier in muscles derived from the second BA than those derived from the

first BA. This might explain the mastication muscle-specific defect; the only BA1-derived muscles present in this mutant correspond to those with a contribution from splanchnic mesoderm (Nathan et al., 2008) indicating that not only are the premyogenic activating signals different but also those maintaining the myogenic cascade.

References
Asakura, A., Lyons, G. E. and Tapscott, S. J. (1995). The regulation of MyoD gene expression: conserved elements mediate expression in embryonic axial muscle. *Dev. Biol.* **171**, 386-398.

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P. and Buckingham, M. E. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. *Genes Dev.* **20**, 2450-2464.

Braun, T., Bober, E., Rudnicki, M. A., Jaenisch, R. and Arnold, H. H. (1994). MyoD expression marks the onset of skeletal myogenesis in Myf-5 mutant mice. *Development* **120**, 3083-3092.

Buchberger, A., Nomokonova, N. and Arnold, H. H. (2003). Myf5 expression in somites and limb buds of mouse embryos is controlled by two distinct distal enhancer activities. *Development* **130**, 3297-3307.

Carvajal, J. J. and Rigby, P. W. J. (2010). Regulation of gene expression in vertebrate skeletal muscle. *Exp. Cell Res.* **316**, 3014-3018.

Carvajal, J. J., Cox, D., Summerbell, D. and Rigby, P. W. J. (2001). A BAC transgenic analysis of the Mrf4/Myf5 locus reveals interdigitated elements that control activation and maintenance of gene expression during muscle development. *Development* **128**, 1857-1868.

Carvajal, J. J., Keith, A. and Rigby, P. W. J. (2008). Global transcriptional regulation of the locus encoding the skeletal muscle determination genes Mrf4 and Myf5. *Genes Dev.* **22**, 265-276.

Chen, J. C. and Goldhamer, D. J. (2004). The core enhancer is essential for proper timing of MyoD activation in limb buds and branchial arches. *Dev. Biol.* **265**, 502-512.

Chen, J. C., Love, C. M. and Goldhamer, D. J. (2001). Two upstream enhancers collaborate to regulate the spatial patterning and timing of MyoD transcription during mouse development. *Dev. Dyn.* **221**, 274-288.

Chen, J. C., Ramachandran, R. and Goldhamer, D. J. (2002). Essential and redundant functions of the MyoD distal regulatory region revealed by targeted mutagenesis. *Dev. Biol.* **245**, 213-223.

Dong, F., Sun, X., Liu, W., Ai, D., Klysik, E., Lu, M. F., Hadley, J., Antoni, L., Chen, L., Baldini, A., Francis-West, P. and Martin, J. F. (2006). Pitx2 promotes development of splanchnic mesoderm-derived branchiomeric muscle. *Development* **133**, 4891-4899.

Dubchak, I. and Ryaboy, D. V. (2006). VISTA family of computational tools for comparative analysis of DNA sequences and whole genomes. *Methods Mol. Biol.* **338**, 69-89.

Eben-Massari, M., Rivera, R. R., Voland, J. R., Quong, M. W., Breit, T. M., van Dongen, J. J., de Smit, O. and Murre, C. (1998). Characterization of ABF-1, a novel basic helix-loop-helix transcription factor expressed in activated B lymphocytes. *Mol. Cell. Biol.* **18**, 3130-3139.

Faerman, A., Goldhamer, D. J., Puzis, R., Emerson, C. P., Jr and Shani, M. (1995). The distal human myoD enhancer sequences direct unique muscle-specific patterns of lacZ expression during mouse development. *Dev. Biol.* **171**, 27-38.

Familari, M., Vu, D. and Parry, L. J. (2009). Regulation of Rxfp2 (Lgr8) expression in the mouse fetal kidney by the transcription factor Pod1 (Tcf 21). *Ann. N. Y. Acad. Sci.* **1160**, 317-319.

Fomin, M., Nomokonova, N. and Arnold, H. H. (2004). Identification of a critical control element directing expression of the muscle-specific transcription factor MRF4 in the mouse embryo. *Dev. Biol.* **272**, 498-509.

Funato, N., Ohyama, K., Kuroda, T. and Nakamura, M. (2003). Basic helix-loop-helix transcription factor Epicardin/Capsulin/Pod-1 suppresses differentiation by negative regulation of transcription. *J. Biol. Chem.* **278**, 7486-7493.

Gilthorpe, J., Vandromme, M., Brend, T., Gutman, A., Summerbell, D., Totty, N. and Rigby, P. W. J. (2002). Spatially specific expression of Hoxb4 is dependent on the ubiquitous transcription factor NFY. *Development* **129**, 3887-3889.

Goldhamer, D. J., Brunk, B. P., Faerman, A., King, A., Shani, M. and Emerson, C. P., Jr (1995). Embryonic activation of the myoD gene is regulated by a highly conserved distal control element. *Development* **121**, 637-649.

Hadchouel, J., Tajbakhsh, S., Primig, M., Chang, T. H., Daubas, P., Rocancourt, D. and Buckingham, M. (2000). Modular long-range regulation of Myf5 reveals unexpected heterogeneity between skeletal muscles in the mouse embryo. *Development* **127**, 4455-4467.

Hadchouel, J., Carvajal, J. J., Daubas, P., Bajard, L., Chang, T., Rocancourt, D., Cox, D., Summerbell, D., Tajbakhsh, S., Rigby, P. W. J. and Buckingham, M. (2003). Analysis of a key regulatory region upstream of the Myf5 gene reveals multiple phases of myogenesis, orchestrated at each site by a combination of elements dispersed throughout the locus. *Development* **130**, 3415-3426.

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolodny, N. L. and Kolchanov, N. A. (1998). Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acids Res.* **26**, 362-367.

Hong, C. Y., Gong, E. Y., Kim, K., Suh, J. H., Ko, H. M., Lee, H. J., Choi, H. S. and Lee, K. (2005). Modulation of the expression and transactivation of androgen receptor by the basic helix-loop-helix transcription factor Pod-1 through recruitment of histone deacetylase 1. *Mol. Endocrinol.* **19**, 2245-2257.

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomès, D., Rocancourt, D., Buckingham, M., Shinin, V. and Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature* **431**, 466-471.

Kaul, A., Köster, M., Neuhaus, H. and Braun, T. (2000). Myf-5 revisited: loss of early myotome formation does not lead to a rib phenotype in homozygous Myf-5 mutant mice. *Cell* **102**, 17-19.

Kelly, R. G., Jerome-Majewska, L. A. and Papaioannou, V. E. (2004). The del22q11.2 candidate gene Tbx1 regulates branchiomeric myogenesis. *Hum. Mol. Genet.* **13**, 2829-2840.

Kucharczuk, K. L., Love, C. M., Dougherty, N. M. and Goldhamer, D. J. (1999). Fine-scale transgenic mapping of the MyoD core enhancer: MyoD is regulated by distinct but overlapping mechanisms in myotomal and non-myotomal muscle lineages. *Development* **126**, 1957-1965.

Lu, J., Richardson, J. A. and Olson, E. N. (1998). Capsulin: a novel bHLH transcription factor expressed in epicardial progenitors and mesenchyme of visceral organs. *Mech. Dev.* **73**, 23-32.

Lu, J., Webb, R., Richardson, J. A. and Olson, E. N. (1999). MyoR: a muscle-restricted basic helix-loop-helix transcription factor that antagonizes the actions of MyoD. *Proc. Natl. Acad. Sci. USA* **96**, 552-557.

Lu, J. R., Bassel-Duby, R., Hawkins, A., Chang, P., Valdez, R., Wu, H., Gan, L., Shelton, J. M., Richardson, J. A. and Olson, E. N. (2002). Control of facial muscle development by MyoR and capsulin. *Science* **298**, 2378-2381.

Miyagishi, M., Hatta, M., Ohshima, T., Ishida, T., Fujii, R., Nakajima, T. and Fukamizu, A. (2000). Cell type-dependent transactivation or repression of mesoderm-restricted basic helix-loop-helix protein, POD-1/Capsulin. *Mol. Cell. Biochem.* **205**, 141-147.

Nathan, E., Monovich, A., Tirosh-Finkel, L., Harrelson, Z., Rousso, T., Rinon, A., Harel, I., Evans, S. M. and Tzahor, E. (2008). The contribution of Islet1-expressing splanchnic mesoderm cells to distinct branchiomeric muscles reveals significant heterogeneity in head muscle development. *Development* **135**, 647-657.

Noden, D. M., Marcucio, R., Borycki, A. G. and Emerson, C. P., Jr (1999). Differentiation of avian craniofacial muscles: I. Patterns of early regulatory gene expression and myosin heavy chain synthesis. *Dev. Dyn.* **216**, 96-112.

Nowotschin, S., Liao, J., Gage, P. J., Epstein, J. A., Campione, M. and Morrow, B. E. (2006). Tbx1 affects asymmetric cardiac morphogenesis by regulating Pitx2 in the secondary heart field. *Development* **133**, 1565-1573.

Patapoutian, A., Miner, J. H., Lyons, G. E. and Wold, B. (1993). Isolated sequences from the linked Myf-5 and MRF4 genes drive distinct patterns of muscle-specific expression in transgenic mice. *Development* **118**, 61-69.

Ribas, R., Moncaut, N., Siligan, C., Taylor, K., Cross, J. W., Rigby, P. W. J. and Carvajal, J. J. (2011). Members of the TEAD family of transcription factors regulate the expression of Myf5 in ventral somitic compartments. *Dev. Biol.* **355**, 372-380.

Robb, L., Hartley, L., Wang, C. C., Harvey, R. P. and Begley, C. G. (1998). musculin: a murine basic helix-loop-helix transcription factor gene expressed in embryonic skeletal muscle. *Mech. Dev.* **76**, 197-201.

Rudnicki, M. A., Braun, T., Hinuma, S. and Jaenisch, R. (1992). Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell* **71**, 383-390.

Rudnicki, M. A., Schnegelsberg, P. N., Stead, R. H., Braun, T., Arnold, H. H. and Jaenisch, R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75**, 1351-1359.

Sambasivan, R., Gayraud-Morel, B., Dumas, G., Cimper, C., Paisant, S., Kelly, R. G. and Tajbakhsh, S. (2009). Distinct regulatory cascades govern extraocular and pharyngeal arch muscle progenitor cell fates. *Dev. Cell* **16**, 810-821.

Shih, H. P., Gross, M. K. and Kioussi, C. (2007). Cranial muscle defects of Pitx2 mutants result from specification defects in the first branchial arch. *Proc. Natl. Acad. Sci. USA* **104**, 5907-5912.

Summerbell, D., Ashby, P. R., Coutelle, O., Cox, D., Yee, S. and Rigby, P. W. J. (2000). The expression of Myf5 in the developing mouse embryo is controlled by discrete and dispersed enhancers specific for particular populations of skeletal muscle precursors. *Development* **127**, 3745-3757.

Swaminathan, S., Ellis, H. M., Waters, L. S., Yu, D., Lee, E. C., Court, D. L. and Sharan, S. K. (2001). Rapid engineering of bacterial artificial chromosomes using oligonucleotides. *Genesis* **29**, 14-21.

Tajbakhsh, S., Rocancourt, D., Cossu, G. and Buckingham, M. (1997). Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* **89**, 127-138.

Tamura, M., Kanno, Y., Chuma, S., Saito, T. and Nakatsuji, N. (2001). Pod-1/Capsulin shows a sex- and stage-dependent expression pattern in the mouse gonad development and represses expression of Ad4BP/SF-1. *Mech. Dev.* **102**, 135-144.

Tapscott, S. J., Lassar, A. B. and Weintraub, H. (1992). A novel myoblast enhancer element mediates MyoD transcription. *Mol. Cell. Biol.* **12**, 4994-5003.

Teboul, L., Hadchouel, J., Daubas, P., Summerbell, D., Buckingham, M. and Rigby, P. W. J. (2002). The early epaxial enhancer is essential for the initial expression of the skeletal muscle determination gene Myf5 but not for subsequent, multiple phases of somitic myogenesis. *Development* **129**, 4571-4580.

Tzahor, E., Kempf, H., Mootoosamy, R. C., Poon, A. C., Abzhanov, A., Tabin, C. J., Dietrich, S. and Lassar, A. B. (2003). Antagonists of Wnt and BMP signaling promote the formation of vertebrate head muscle. *Genes Dev.* **17**, 3087-3099.

von Scheven, G., Bothe, I., Ahmed, M. U., Alvares, L. E. and Dietrich, S. (2006). Protein and genomic organisation of vertebrate MyoR and Capsulin genes and their expression during avian development. *Gene Expr. Patterns* **6**, 383-393.

Yee, S. P. and Rigby, P. W. J. (1993). The regulation of myogenin gene expression during the embryonic development of the mouse. *Genes Dev.* **7**, 1277-1289.

Yu, L., Mikloucich, J., Sangster, N., Perez, A. and McCormick, P. J. (2003). MyoR is expressed in nonmyogenic cells and can inhibit their differentiation. *Exp. Cell Res.* **289**, 162-173.

Yu, L., Sangster, N., Perez, A. and McCormick, P. J. (2004). The bHLH protein MyoR inhibits the differentiation of early embryonic endoderm. *Differentiation* **72**, 341-347.

Zacharias, A. L., Lewandoski, M., Rudnicki, M. A. and Gage, P. J. (2011). Pitx2 is an upstream activator of extraocular myogenesis and survival. *Dev. Biol.* **349**, 395-405.

Zweigerdt, R., Braun, T. and Arnold, H. H. (1997). Faithful expression of the Myf-5 gene during mouse myogenesis requires distant control regions: a transgene approach using yeast artificial chromosomes. *Dev. Biol.* **192**, 172-180.