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

Natalia Moncaut1, Joe W. Cross1, Christine Siligan1, Annette Keith1, Kevin Taylor1,2, Peter W. J. Rigby1 and Jaime J. Carvajal1,2,3

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 periorbital 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 (Kassa-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...
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, Myf4 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 Myf4 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 Myf4 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 other 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; Tapscoot 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 contrast, 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 –24lacZ) (Chen et al., 2001). The same base construct from which the CE has been removed (construct –24ΔF3lacZ) (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 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 branchiomyogenic 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 branchiomyogenic 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'-CCAATAGGTACAGTTGGATTTTTCATTG-3' and 5AE-R 5'-TCTCTGATGAGAGATGCGTGTGCTGCTCTGCTGTCC-
Generation of transgenic mice

All in vivo experiments were performed according to United Kingdom Home Office Regulations.

Preparation of plasmid constructs

ECR-1 was cloned in a plasmid containing the Myf5 minimal promoter as previously described (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 Misky’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 K3Fe(CN)6, 5 mM K4Fe(CN)6, 2 mM MgCl2, 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 Misky’s fixative. For sectioning, embryos were embedded in 2% (w/v) agarose (SeaKem Gold, Lonza), fixed in Misky’s fixative overnight and sectioned using a Leica VT1000S vibratome at 70 μm.

Immunohistochemistry

Mouse embryos were fixed in 2% (v/v) parafomaldehyde/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 (1:1000, 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, 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).

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 (BA) (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 B195APZPAE, n=4 stable lines). At early stages, B195APZPAE 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 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.

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).

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 MgCl2, 10 mM KCl, 0.2 mM PMSE, 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 MgCl2, 0.02% M KCl, 0.2 mM EDTA, 0.2 mM PMSE, 0.5 mM DTT], transferred to high-salt buffer [20 mM HEPES pH 7.9, 25% (v/v) glycerol, 1.5 mM MgCl2, 12 M KCl, 0.2 mM EDTA, 0.2 mM PMSE] 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′-TAGCTGGCACGCTGGAGCCATCTGTAACGGTT-GTTA-3′).
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-1mBS1256-MZ, n=4), whereas double mutation of BS3 and BS4 (E-boxes; construct 111-ECR-1mBS34-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-1mBS3-MZ, n=6; 111-ECR-1mBS4-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 B195APZECR-1mBS34, n=5 stable lines). Although at 9.5 dpc B195APZwt drives transgene expression in both BAs (Fig. 3A,B), expression in the first BA in B195APZECR-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 B195APZwt 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...
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. 3IJ with 3KL). 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 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 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 chromatin extracted from the midbrain (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 branchiomer 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. 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 corresponding to ECR-1 showing the putative binding sites (in bold) as 111-EKR-1mBS34-MZ (E), 111-EKR-1mBS3-MZ (F) and 111-EKR-1mBS4-MZ (G) were stained for \( \beta \)-GAL activity. Mutations in BS3 and BS4 singly (F,G) or combined (E) abolish \( \beta \)-GAL expression in the branchial arches. The 111-MZ drives consistent \( \beta \)-GAL 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.

**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 \( \beta \)-GAL and MYOD double staining in B195APZ\(^{34}\) 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 \( \beta \)-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 \( \beta \)-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 \( \beta \)-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 branchiomer 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...
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, MRFP4 and myogenin, different locations in the embryo during development generating transcriptional cascades that orchestrate muscle formation.
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 (TF3 – 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 (Familiari 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 Tcf21 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 also failed to 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.
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 branchiomeriic myogenesis.

### 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.

### 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
first BA. This might explain the mastication muscle-specific defect; the only BA-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.

Acknowledgements
We would like to thank Eric Olson for providing the Msc knockout line and Thomas Braun for providing the Myf5 and MyoD knockout lines. We thank David Robertson from the Breakthrough Breast Cancer Research Centre for help with microscopy. We are grateful to all members of the Rigby laboratory for helpful discussions, and the anonymous referees for their insightful comments.

Funding
This work was supported by The Institute of Cancer Research, a Medical Research Council Programme Grant [G040119] and the European Muscle Development Network (MYORES) Network of Excellence (contract 511978) from the European Commission 6th Framework Programme. J.W.C. is supported by a studentship from The Institute of Cancer Research. J.J.C. was partly supported by a Ministry of Science and Innovation (MICINN) grant [BFU2011-22928].

Competing interests statement
The authors declare no competing financial interests.

Supplementary material
Supplementary material available online at http://dev.biologists.orglookup/suppl/doi:10.1242/dev.068015/-DC1

References


