

## ***Dial M(RF) for Myogenesis***

***Natalia Moncaut*** <sup>(1,2)</sup>, ***Peter W. J. Rigby*** <sup>(1)</sup> and ***Jaime J. Carvajal*** <sup>(3,4)</sup>

1: Division of Cancer Biology, The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, United Kingdom.

2: Present address: University College London, Department of Cell and Developmental Biology, London WC1E 6BT, United Kingdom.

3: Molecular Embryology Team, Centro Andaluz de Biología del Desarrollo, CSIC-UPO-JA, Ctra de Utrera km1, Seville, Spain.

4: Corresponding author. Phone No.: (+34) 954978210; email address: [j.carvajal@csic.es](mailto:j.carvajal@csic.es)

**Running title:** Dial M(RF) for Myogenesis

**Keywords:** Myogenesis, Myogenic Regulatory Factor, MRF, Myf5, Mrf4, MyoD, Myogenin, mouse embryo, miRNA, GRNs

**Abbreviations:** MRF, Myogenic Regulatory Factor; bHLH, basic helix-loop-helix; miRNA (or miR), micro-RNA; DM, dermomyotome; VLL, ventro-lateral lip; BA, branchial arch; NC, neural crest; EOM, extra-ocular muscle; GRN, gene regulatory network; MEF, mouse embryonic fibroblast; mRNP, messenger ribonucleoprotein particles; dpc, days post coitum;

### ***Abstract***

The transcriptional regulatory network that controls the determination and differentiation of skeletal muscle cells in the embryo has at its core the four Myogenic Regulatory Factors (MRFs), Myf5, MyoD, Mrf4 and Myogenin. These bHLH transcription factors act by binding, as obligate heterodimers with the ubiquitously expressed E proteins, to the E-box sequence CANNTG. While all skeletal muscle cells have the same underlying function their progenitors arise at many sites in the embryo and it has become apparent that the upstream activators of the cascade differ in these various populations so that it can be switched on by a variety of inductive signals, some of which act by initiating transcription, some by maintaining it. The application of genome wide approaches has provided important new information as to how the MRFs function to activate the terminal differentiation programme and some of these data provide significant mechanistic insights into questions which

have exercised the field for many years. We also consider the emerging roles played by miRNAs in the regulation of both upstream activators and terminal differentiation genes.

### ***Introduction***

The formation of skeletal muscle provides one of the best models for studying the processes of cellular specification and differentiation and of organogenesis. Our extensive knowledge of this system was greatly facilitated by the early development of the C2 cell line (1), which differentiates to form contractile myotubes *in vitro*, and by the discovery of a key transcription factor, MyoD, on the basis of its ability to induce the myogenic programme when introduced into non-muscle cells (2). Three related proteins, Myf5 (3), Myogenin (4, 5) and Mrf4 (also known as Myf6 or herculin: 6-8) were discovered shortly after. All four belong to the basic helix-loop-helix super family of proteins that bind to the E-box sequence, CANNTG, and all induce myogenic conversion when transfected into fibroblasts. These Myogenic Regulatory Factors (MRFs) form the core of the transcriptional cascade that leads to the skeletal muscle phenotype and therefore key questions in the field have been how these genes are turned on at the onset of myogenesis in the embryo and how do the encoded proteins function to trigger the terminal differentiation programme and subsequent organogenesis.

All skeletal muscle cells have the ability to produce a contractile force that changes the length and shape of the cell and can thus generate motion. While this general property is shared by all of the skeletal muscles of our body, the origins of these muscles during embryogenesis vary widely (Figure 1). In this review we will consider our current knowledge of the regulation and function of the MRFs including the distinct genetic networks that control myogenesis at various locations in the embryo.

### ***Origins of skeletal muscle***

The origins of the various skeletal muscle groups were initially studied by grafting experiments in avian embryos although more recently genetic approaches in the mouse have been used. The muscles of the trunk include those that move the vertebral column, the muscles that form the thoracic and abdominal body walls, and those that cover the pelvic area. Trunk muscles derive from somites, which are transient paraxial mesodermal structures that form pairwise by the sequential segmentation of the presomitic mesoderm. The dorsal-most layer of the somite, the dermomyotome (DM), is the source of both myogenic and dermal progenitors of the

trunk. During development the myotome, the first muscle in the embryo, forms below the DM by cell delamination from the dorso-medial lip to produce a post-mitotic, differentiated muscle layer termed the epaxial myotome (9, 10). Shortly after, myoblasts from the ventro-lateral lip (VLL) also begin migrating, contributing to the hypaxial myotome followed by delamination of cells from the lateral (rostral and caudal) lips of the DM (11-13). Like the cells of the hypaxial trunk muscles, limb muscle precursor cells originate from the VLL of the DM at the level of the developing limb (14, 15). They delaminate and migrate into the limb bud where they re-associate and form primary myofibres.

The skeletal muscles of the head are responsible for the motion of the jaw, face, eyes and ears and are therefore critical for activities such as feeding, sensing and vocalisation. According to their function, muscles of the head and neck can be divided into four groups: branchiomic, laryngoglossal, extraocular and axial.

The branchial arches (BAs) contribute to the development of craniofacial muscles and produce the majority of the muscles of mastication and facial expression. The branchiomic mesenchyme, the core of the BAs, is formed by cranial paraxial mesoderm and neural crest (NC) cells that migrate around 9.5dpc (16, 17). It is from the central cell population in the core that skeletal muscle arises. Recent work by Grenier et al. (18) suggests that some mixing of the NC-derived and mesodermal populations takes place, with NC cells within the mesodermal core being essential for the formation of the tendons and other connective elements of the head. It has recently been shown that splanchnic mesodermal precursors contribute to some lower jaw muscles (19), while extensive clonal analysis reveals common progenitors contributing to both extraocular and branchiomic muscles (20). After migration into the BA core, myogenic progenitor cells leave the arches and move to their final positions within the head mesenchyme where they differentiate into functional muscle blocks.

The extraocular muscles (EOMs) surround the eyeball and are responsible for the fine-tuned movements of the eyes. During development, the EOMs are formed by migratory cells from the first branchial arch (20) and cells from the cranial mesoderm located in the periocular region, the prechordal mesoderm (21, 22). The glossal muscles combine muscles of the tongue and the muscles anchoring the tongue to the jaw. These muscles are derived from the four anterior-most somites, the occipital somites, with migratory populations delaminating from the VLL and migrating in a manner similar to those of the limb, although not as individual cells but as a coherent group of migrating cells.

In the transition zone between the head and the trunk are the axial neck muscles. Due to their location it has not been clear if the origin of this group of muscles was cranial or trunk mesoderm but recent data confirm that they are formed from myoblasts of the occipital lateral plate mesoderm (23).

### ***Gene Regulatory Networks involved in skeletal muscle development***

The process of cell differentiation depends on the activity of a specific set of regulatory genes, that is, genes encoding transcription factors and signalling molecules. The recognition and binding of transcription factors to specific DNA target sequences is one of the central foundations of Gene Regulatory Networks (GRNs) (24). The idea of understanding and predicting the complex control systems underlying animal development has become a new field in biology and computational research. These GRNs can be defined as genomic regulatory codes, which function to determine the sets of genes that must be expressed in specific spatial and temporal patterns. Complex spatio-temporal processes of differential gene expression rule embryonic development and demand an extremely tight control of transcriptional regulation. The formation of skeletal muscle has been one of the developmental systems most studied in recent decades, initially because of the availability of muscle cell lines able to recapitulate the entire differentiation process to form myotubes. However *in vivo* this process involves a number of different steps that will result in a functional differentiated cell. The complexity of this scenario requires an intricate network of transcription factors and cis-regulatory target sequences essential for the differentiation and homeostasis of skeletal muscle progenitors.

Despite the differences in their embryological origins, all myogenic progenitor cells share the same core components of the myogenic pathway which is formed by the MRFs: Myf5, MyoD, Mrf4 and Myogenin (Figure 2). The deployment of these varies according to the particular precursor population giving rise to muscle. Expression of the MRF proteins during development is under strict temporal and spatial control. Myf5 is expressed before the adoption of the myogenic fate and, based on the information derived from Myf5 null mice, is considered a determination factor. Likewise, MyoD expression in a Myf5<sup>-/-</sup> background drives cells to the myogenic lineage, albeit with a delay in some populations (25). Mrf4 and MyoG expression is later and they were traditionally classed as genes involved in the process of differentiation. However, we now know that Mrf4 is expressed at the same time as, if not before, Myf5 in the somitic bud and that in the absence of Myf5 and MyoD, Mrf4 is able to induce myogenesis, indicating that Mrf4 should be reclassified

as having both differentiation and determination activity (26, 27). Most of the myogenic programme is severely affected only when both Myf5 and MyoD are absent. The Myf5<sup>-/-</sup>:MyoD<sup>-/-</sup> embryo fails to develop any skeletal muscle (28) although as Mrf4 expression is compromised in cis in the Myf5 allele, the necessity of both Myf5 and MyoD was not entirely demonstrated. Indeed, a new Myf5 allele in which Mrf4 expression is not affected (27) shows that in the case of the EOMs, either Myf5 or Mrf4 is required for their formation (29). The Myogenin knockout has an equally remarkable phenotype with perinatal death. While myoblasts are formed there is a complete absence of functional skeletal muscle supporting the idea that Myogenin regulates the later stages of myogenic differentiation, whilst Myf5 and MyoD (and in some cases Mrf4) are involved in the process of determination (30, 31).

There is also evidence that this distinction between MRFs is present at the molecular level. It was demonstrated that MYOD, MYF5 and MRF4, but not MYOGENIN, are able to remodel repressive chromatin environments through recruiting SWI/SNF proteins (ATP-dependent chromatin remodelling enzymes) in order to promote muscle differentiation (32-34).

The activation of the core myogenic cascade occurs in all tissues destined for the skeletal myogenic lineage; however, the factors which initiate the cascade differ radically in the different locations of the embryo. The first evidence for this came from the analysis of the double Pax3;Myf5(Mrf4) knockout animals (with Mrf4 expression disrupted in cis). These mutant mice lack all skeletal muscle in the trunk but show no abnormalities in craniofacial musculature suggesting the presence of separate specification pathways in the trunk and the head (35). Importantly, although trunk myogenesis is severely impaired in Pax3 KO animals, the early specification is unaffected and progenitors expressing Myf5 are clearly identifiable at 10.5dpc (36), indicating that Pax3 could act at the level of survival or proliferation and not as a specifying factor of early trunk myoblasts.

Deletion of the genes encoding the homeobox proteins MEOX1 and MEOX2 results in a severe disruption to the patterning of the somites and skeletal muscle of the trunk is severely compromised, indicating a role for the MEOX proteins in somitogenesis and trunk myogenesis (37). Skeletal muscles of the head were unaffected in these mutants, despite their expression of MEOX proteins. We now know that there is much greater heterogeneity in the upstream activators of the cascade.

Several studies have identified factors controlling the myogenic cascade specifically in the head: TBX1 is a T-box factor required for the formation of the BA-derived muscles (38). Its inactivation results in a severely atrophied BA2 and in the sporadic development of first arch-derived muscles (39). A mesodermal-specific knockout of the *Tbx1* gene shows that the requirement for *Tbx1* in BA myogenesis is cell-autonomous, and thus that the phenotype is not the result of a general disruption of craniofacial morphogenesis (40). PITX2 is a paired-like homeodomain transcription factor with a role in the specification of myogenic precursors in the BAs. 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 with the resulting expression of *Myf5*, *MyoD* and *Myog* being compromised (41, 42). MSC and TCF21 (epicardin/capsulin/POD-1) are bHLH transcription factors, known to be markers of undifferentiated muscle precursor cells (43). 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 (44). We have recently shown that *Msc* and *Tcf21* control the transcription levels of *Myf5* and *Myod* in the BAs through binding to specific enhancers (45). We also suggested that other regulatory factors, namely *Tbx1* and/or *Pitx2* are responsible for the onset of MRF expression in the arches.

*Isl1* (ISL1) is a marker of the splanchnic mesoderm and tracing studies using *ISL1*-cre label branchiomeric muscles to varying extents (46). Recent work introduced *Lhx2* as a new player in the GRN during branchiomeric muscle development (47). The knockout of this gene results in pharyngeal muscle specification defects and epistatic relationships between *Tbx1*, *Lhx2*, and *Myf5* were described, affecting early pharyngeal muscle specification and patterning. Furthermore, retrospective clonal analyses also show the presence of common progenitors to heart and facial muscles, indicating that these progenitors have the ability to contribute to two different muscle phenotypes. While the activation of *Lhx2* is probably linked to the skeletal muscle phenotype, it is still not clear if the "default" phenotype is cardiac muscle or if additional inputs are required for its specification, nor if the fate of those progenitors that have failed to become skeletal muscle can adopt other, non-muscle, fates or contribute exclusively to the cardiac musculature. An investigation into the putative fate changes of these dual progenitors in the different KO strains (*Lhx2*, *Isl-1*, *Pitx2*, etc) should shed new light on these questions.

Finally, and as previously mentioned, the deployment of the myogenic cascade varies according to the embryonic origin and the final location of a particular

muscle. The GRNs, responsible for the activation of this cascade, are under the control of multiple signalling inputs. It is curious that, for example, head and trunk myogenic programs also exhibit different outcomes in response to individual signalling molecules and thus, whereas trunk myogenesis is promoted by the action of WNT and inhibited by BMPs, antagonists to WNT promote cranial myogenesis (48).

### ***Mechanisms of transcriptional regulation in muscle development***

It is now possible to ask questions about transcription factor binding, histone modifications and transcriptional outputs on a genome wide scale and a number of recent studies have applied these new technologies to the skeletal muscle paradigm. Cao et al. (49) used ChIP-Seq to identify the sites to which MyoD is bound in the chromatin of C2C12 myoblasts and differentiated myotubes derived from them. As expected, they found the protein bound to the control elements of genes which are known to be up- or down-regulated during the process of differentiation. What was not expected were their findings that MyoD is also bound to a very large number of sites which are not obviously associated with such regulated genes and that most of the sites are the same in myoblasts and myotubes, although in myotubes there is more binding to sites associated with genes up-regulated during the differentiation process. Most of the sites which bind MyoD are not active in a classic transfection assay for enhancer function and it is thus not clear whether this widespread binding represents hitherto unknown functions of MyoD or is simply the consequence of the fact that the protein will bind to all E-boxes with some affinity. Interestingly they show that MyoD binding causes regional, rather than local, histone acetylation.

This group has also addressed the question of why different bHLH proteins drive different lineage-specific gene expression programmes (50). Just as MyoD regulates muscle-specific genes the closely related transcription factor NeuroD drives a neuronal programme. They used lentiviral transduction to introduce NeuroD into P19 cells, a pluripotential mouse embryonal carcinoma cell line, and thus convert them into neurones and MyoD into mouse embryonic fibroblasts (MEFs) converting them into skeletal muscle cells. It is important to note that NeuroD cannot convert MEFs to neurones and that MyoD cannot convert P19 cells to muscle. The data again show an unexpectedly high number of binding sites, for both factors. With that borne in mind, it is of great interest that they reveal a fascinating specificity. Both factors bind to the CAGCTG E-box but each also binds to a factor-specific E-box motif, CAGATG for NeuroD and CAGGTG for MyoD. And this binding specificity is reflected in function. Binding to the private E-box sequences is associated with the

transcriptional activation of adjacent genes, whereas binding to the shared sequence is associated with regional epigenetic modification. They further show that binding, of either factor, is constrained by chromatin accessibility, as revealed by nuclease sensitivity assays. Thus which sites are accessible is epigenetically determined in a lineage specific fashion, whereas sequence-specific binding strongly influences which factor binds where, and which genes are subsequently activated. A major question for the future is how are the lineage-specific epigenetic marks set in the first place.

Soleimani et al. (51) have reported a similar analysis although there are important differences in the procedures employed and in the conclusions reached. They used retroviral transduction to introduce TAP-tagged derivatives of the transcription factors of interest into muscle stem cells isolated from hind limb and then used the tag to recover the protein and its bound DNA. They found a much smaller number of MyoD sites and a major difference between myoblasts (1400 sites) and myotubes (9300 sites). The exogenous protein is expressed at a higher level than the endogenous one but that would be expected to lead to the occupation of a greater number of sites, not significantly fewer. They have also analysed the binding of the Snail proteins, zinc-finger transcriptional repressors that recruit the histone deacetylases HDAC1 and 2. This work leads to some interesting mechanistic conclusions. They show that Snail binds to E-boxes that have a G/C-rich central dinucleotide, and that such sites are associated with genes that are expressed in myotubes, and that it does not bind to E-boxes with A/T-rich central dinucleotides which are associated with genes expressed in myoblasts. Thus at the onset of differentiation Snail must be removed in order to allow MyoD access to the myotube genes. They present data, some of which is derived from heterologous systems, which indicate that miR-30a targets Snail1 mRNA and that miR-206 targets Snail2 mRNA and that the genes for both of these miRNAs are activated by MRFs. Furthermore, it is shown that over-expression of Snail blocks differentiation whereas siRNA against Snail induces precocious differentiation. Thus when cells receive a differentiation signal MRFs activate the miRNAs which mediate the destruction of the Snail mRNAs and as the Snail proteins turn over MyoD gains access to the enhancers of the genes expressed in myotubes. The details of this model require further work but it potentially provides a satisfying mechanistic answer to one of the major issues in the field.

Blum et al. (52) have taken a somewhat complementary approach by identifying enhancers in C2C12 myoblasts and myotubes on the basis of histone

marks and then examining their occupancy by RNA polymerase II, transcription factors and co-activators. Satisfyingly the identified enhancers that are adjacent to genes are linked to those expressed in the relevant cell type but only a small minority of the condition-specific “enhancers” are associated with transcripts, raising the possibility that the marks in question are found on elements with other functions. They use the data of Cao et al. (49) to show that MyoD binding to enhancers correlates with the binding of pol-II and p300 and that in MyoD<sup>-/-</sup> cells these proteins are not present, and that there is a marked diminution in transcription across the enhancers. Re-expression of MyoD in the null myoblasts restores pol-II binding and H3K4 monomethylation but not H3K27 acetylation while in myotubes it restores the acetylation as well. Re-expression experiments of this sort are potentially powerful in that they can help elucidate cause and effect relationships.

In each of these papers the authors examine the sequences around the sites of MyoD binding and ask, by a variety of procedures, what other transcription factors could bind there, in the expectation that such proteins might well co-operate with MyoD in the initiation of the developmental programme. They find sites for proteins that all in the field would expect, e.g. the Mef2 factors and RUNX but also sites for proteins that have not been much studied in the context of myogenesis, e.g. PPAR- $\gamma$  and c-Myb. Rhabdomyosarcomas are paediatric tumours that express many but not all skeletal muscle markers and have long been considered to be of skeletal muscle origin, although that point has not been proven. MacQuarrie et al. (53) have shown that several of the factors which may well bind adjacent to MyoD, e.g. Mef2C and RUNX are expressed at lower than expected levels in the tumour cells and that forcing their expression induces the cells to enter the terminal skeletal muscle differentiation programme, thus providing strong support for the notion that they do act co-operatively with MyoD.

In all of this it should be remembered that C2 cells were derived from the thigh muscles of a mouse that had suffered a crush injury, and they are therefore generally thought to be related to satellite cells. It will be important in the long run to acquire data of this sort from the muscle progenitor cells of the embryo at the time when fate decisions are being made but we do not under-estimate the technical challenges posed by this desire.

More recently improvements in bioinformatics, especially the use of comparative genomics, have provided high-resolution conservation maps. These analyses allow the identification of cis-regulatory elements within particular loci able to drive specific gene expression. One of the best examples is the Myf5/Mrf4 locus.

Since 1993 several studies using plasmid/BAC/YAC deletion approaches showed the presence of regulatory regions throughout the locus (reviewed in 54). Deletion experiments are time consuming and labour intensive, and thus not generally practicable for the characterisation of small regulatory sequences. The use of bioinformatic approaches comparing non-coding evolutionarily conserved regions between several species provides an advantageous tool to identify or redefine the location of those regulatory regions. Using this approach we recently reported two regulatory elements within the *Myf5/Mrf4* locus responsible for driving specific *Myf5* expression in the ventral compartment of the somites and in the branchial arches during development (45, 55). We showed that the ECR-111 somite enhancer is regulated by the transcription factors of the TEAD family, which have been implicated as effectors of the hippo signalling pathway. This enhancer is also active at later stages in the limbs, where it is regulated by Pax3 and proteins of the Six family (56), and in adult satellite cells where it is under the control of Pax7 (57). As discussed above the branchial arch enhancer is regulated by the bHLH factors Msc and Tcf21.

In depth analysis of the regulatory elements of this locus showed another degree of complexity during transcriptional regulation. Some of these elements, like the early epaxial enhancer are modular in nature and do not require the action of further enhancers for activity (58, 59). In isolation these regulatory elements drive the expression pattern missing when deleted from the locus. Others, by contrast, require the input of additional enhancers in order to establish the full expression pattern in a particular subset of muscle progenitors (45; 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.

It is also becoming apparent that particular signalling pathways may play different functions in the trunk and the limb during myogenesis. It has recently been shown that Shh signalling is required for the production of the ventral limb muscles and to regulate directional muscle cell migration in the distal limb (60, 61). Anderson et al. furthermore show that sites within the previously defined limb enhancer which bind Gli proteins *in vitro* are required for the activation of *Myf5* expression in the precursors of the ventral muscle masses, although they did not provide evidence that Gli proteins are bound to the enhancer *in vivo* (60). It has previously been shown that in both the mouse and the zebrafish Shh signalling is not required for the activation of *Myf5* expression in the paraxial mesoderm but rather that it acts to maintain expression levels once the gene is switched on (62, 63).

### ***Post-transcriptional regulation by miRNAs***

MicroRNAs (miRs) are small non-coding RNAs of around 20 nucleotides which post-transcriptionally regulate gene expression by binding to specific sequences in the 3'UTR of target genes (64, 65). Several lines of evidence have shown important effects of miRNAs in muscle development. Mice lacking Dicer, the enzyme responsible for miRNA maturation, specifically in skeletal muscle cells show abnormal muscle fibre development and reduced muscle mass (66). Recent miRNA transcriptome analysis in human myoblasts combined with bioinformatic-based miR-target prediction revealed the presence of at least 60 miRNAs differentially expressed during myogenic differentiation (67), either upregulated (43 out of 60) or downregulated (17 out of 60) during myogenesis. The collection of genes targeted by these miRs are involved in a wide range of cellular functions (transcriptional regulation, cell cycle progression, protein degradation, ubiquitination, apoptosis, cell motility, intracellular transport, etc.) revealing the degree of complexity linked to miR regulation of the myogenic processes.

Some of the best characterised muscle specific miRNAs are miR-1, miR-206 and miR-133 and their expression has been shown to be directly regulated by MRFs (68, 69). By combining microarray screening and target prediction, Goljanek-Whysall and co-workers (70) have identified a set of miR-1 and miR-206 targets in C2C12 cells. Interestingly the single overexpression of some of these targets disrupts the myogenic programme in C2C12s by different mechanisms which include cell-cycle exit, morphology alterations or the activation of alternative differentiation programmes. Furthermore, they show that sustained expression of anti-miR-206 in C2C12 cells grown in differentiation media also affects myogenic differentiation. Whether the targets identified are relevant in the myogenic process during embryogenesis remains to be determined, specially as some have not been detected to date in embryonic musculature. It is also possible that these newly discovered targets are involved in later myogenic processes, such as muscle regeneration, specially if we keep in mind that the C2C12 cells probably derive from limb muscle satellite cells.

Pax3 and Pax7 play several roles during embryonic myogenesis, satellite cell survival, self-renewal, and proliferation. It has been shown that miR-27 is responsible for the downregulation of Pax3 expression during terminal differentiation of skeletal myogenesis. On the other hand, Pax7 is specifically controlled by miR-1, miR-206 and miR-486 in muscle progenitors. The direct control of the expression levels of Pax3/7 through miRs plays a role in the delicate balance between the proliferation

and differentiation of satellite cells, and in the regeneration of injured muscle (71, 72, 73). It has also been shown that one isoform of the paired-related homeobox gene Pitx2 controls the expression of Pax3 through modulating miR27 expression during myogenesis (74). The maintenance of the quiescent state of muscle satellite cells is similarly controlled by miRNAs. miR-489, highly expressed in quiescent satellite cells and quickly downregulated during satellite-cell activation, post-transcriptionally suppresses the oncogene Dek, which is involved in cell proliferation (75). Most satellite cells transcribe the myogenic determination gene Myf5 without activating the myogenic program, being held in a poised state. Myf5 levels are controlled by miR-31 and it is sequestered in mRNP granules until satellite cells are activated. Once activated, dissociation of granules and reduced miR-31 levels allow immediate release of accumulated Myf5 protein (76).

miRNAs are not only essential for earlier cell fate decisions in progenitors but they are also necessary for decisions on which muscle fibre type to form. Adult skeletal muscle is formed by different type of fibres commonly referred to as 'fast' or 'slow' fibres depending on which form of myosin heavy chain (Myh) is expressed. It has been recently shown that the control of muscle myosin content, myofibre identity, and muscle performance is highly dependent on miRNAs. The transcription of specific Myh (Myh7 and Myh7b) in slow fibres is accompanied by the repressive action of miRNAs (miR-208b and miR-499) on Sox6, a gene that promotes the expression of fast fibre Myh genes (77).

### ***Concluding remarks***

Only about 1.5% of mammalian genomes is comprised of protein-coding genes with huge intergenic 'deserts' occupying the majority of the remaining material (78-80). However, it has become clear that much of this non-protein-coding DNA is transcribed into a variety of classes of RNA, both large and small (81, 82). Of particular interest in the context of this review was the report that in activated mouse cortical neurons many enhancers, defined by their epigenetic marks and protein binding profiles, produce short, bidirectional, un-spliced and un-polyadenylated RNAs called eRNAs. This transcription appears to occur specifically at enhancers which are actively engaged in promoting mRNA synthesis (83). However, there are few data indicating whether these eRNAs are functionally active or are simply a by-product of the functional engagement of enhancers with promoters. It is not obvious how one would experimentally distinguish between these possibilities. In complex developmental loci, of which the Mrf4/Myf5 locus is a prime example, such non-

coding transcripts could play key roles in achieving the exquisitely specific regulation and experiments to assess this possibility are a high priority.

The recently released ENCODE project provides a revolutionary view of the genome (84). Through examining many well-established cells lines using a range of techniques such as RNAseq, transcription factor ChIPseq, histone ChIPseq, DNase I hypersensitivity and DNA methylation, a multi-layered view of the genome has been assembled with signatures for enhancers, non-coding RNAs and promoters becoming apparent. Whilst studying a single genetic locus may appear parochial in the face of such enterprises, there is reason to believe that applying similar methods to more directed, perhaps more biologically relevant, questions could reveal much about how gene regulation operates *in vivo* and especially in that most complex of processes, development. For that reason, application of such techniques, especially ChIPseq and RNAseq, to the Mrf4/Myf5 locus will doubtless reveal much about the dynamics of transcription factor loading and interaction as different input signals converge on the master regulators of the myogenic cascade.

### **Figure Legends**

**Figure 1. Origins of skeletal muscles.** Vertebrate skeletal muscles originate from four main mesodermal progenitor populations. The presomitic mesoderm (PSM) gives rise to most of the body musculature through the differential contribution of the somitic derivatives; ventral somitic compartments give rise to body wall, limb, tongue and pharyngeal muscles, the later three involve a process of cell delamination and cell migration, while dorsal somitic compartments are thought to give rise to deep back muscles; the lateral plate mesoderm (LPM) contributes to some mastication muscles following migration through the pharyngeal arches, and to the neck musculature; the posterior head mesoderm (PHM) generates progenitors that also migrate into the pharyngeal arches: those migrating through the second arch give rise to the muscles of facial expression while those contributing to the first arch will give rise to muscles of mastication and with some contribution to the extraocular muscles; finally, the anterior head mesoderm (not shown) contributes progenitors that migrate directly to the future eye area and give rise to extraocular muscles.

**Figure 2. Schematic representation of Gene Regulatory Networks in the trunk, head and EOMs.** Gene Regulatory Networks (GRNs) controlling the processes of determination and differentiation of skeletal muscle share a functional core formed by

the MRFs: Mrf4, Myf5, MyoD and Myogenin. The activators of the cascade and the deployment of the different members of the MRF core varies depending on the mesodermal origin of the cells giving rise to the different muscle groups.

### **References**

1. Yaffe D & Saxel O (1977) A myogenic cell line with altered serum requirements for differentiation. *Differentiation* **7**, 159-166.
2. Davis RL, Weintraub H & Lassar AB (1987) Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987-1000.
3. Braun T, Buschhausen-Denker G, Bober E, Tannich E & Arnold HH (1989) A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. *EMBO J* **8**, 701-709.
4. Edmondson DG & Olson EN (1989) A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. *Genes Dev* **3**, 628-640.
5. Wright WE, Sassoon DA & Lin VK (1989) Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* **56**, 607-617.
6. Braun T, Bober E, Winter B, Rosenthal N & Arnold HH (1990) Myf-6, a new member of the human gene family of myogenic determination factors: evidence for a gene cluster on chromosome 12. *EMBO J* **9**, 821-831.
7. Miner JH & Wold B (1990) Herculin, a fourth member of the MyoD family of myogenic regulatory genes. *Proc Natl Acad Sci U S A* **87**, 1089-1093.
8. Rhodes SJ & Konieczny SF (1989) Identification of MRF4: a new member of the muscle regulatory factor gene family. *Genes Dev* **3**, 2050-2061.
9. Denetclaw WF Jr, Berdugo E, Venters SJ & Ordahl CP (1997) Morphogenetic cell movements in the middle region of the dermomyotome dorsomedial lip associated with patterning and growth of the primary epaxial myotome. *Development* **128**, 1745-1755.
10. Denetclaw WF Jr, Christ B & Ordahl CP (1997) Location and growth of epaxial myotome precursor cells. *Development* **124**, 1601-1610.
11. Cinnamon Y, Kahane N & Kalcheim C (1999) Characterization of the early development of specific hypaxial muscles from the ventrolateral myotome. *Development* **126**, 4305-4315.
12. Gros J, Scaal M & Marcelle C (2004) A two-step mechanism for myotome formation in chick. *Dev Cell* **6**, 875-882.
13. Kahane N, Cinnamon Y & Kalcheim C (2002) The roles of cell migration and myofiber intercalation in patterning formation of the postmitotic myotome. *Development* **129**, 2675-2687.

14. Christ B, Jacob HJ & Jacob M (1974) Origin of wing musculature. Experimental studies on quail and chick embryos. *Experientia* **30**, 1446-1449.
15. Christ B, Jacob HJ & Jacob M (1977) Experimental analysis of the origin of the wing musculature in avian embryos. *Anat Embryol (Berl)* **150**, 171-186.
16. Trainor PA, Tan SS & Tam PP (1994) Cranial paraxial mesoderm: regionalisation of cell fate and impact on craniofacial development in mouse embryos. *Development* **120**, 2397-2408.
17. Trainor PA & Tam PP (1995) Cranial paraxial mesoderm and neural crest cells of the mouse embryo: co-distribution in the craniofacial mesenchyme but distinct segregation in branchial arches. *Development* **121**, 2569-2582.
18. Grenier J, Teillet MA, Grifone R, Kelly RG & Duprez D (2009) Relationship between neural crest cells and cranial mesoderm during head muscle development. *PLoS One* **4**, e4381. doi: 10.1371/journal.pone.0004381.
19. Nathan E, Monovich A, Tirosh-Finkel L, Harrelson Z, Rousso T, Rinon A, Harel I, Evans SM & Tzahor E (2008) The contribution of Islet1-expressing splanchnic mesoderm cells to distinct branchiomic muscles reveals significant heterogeneity in head muscle development. *Development* **135**, 647-657.
20. Lescroart F, Kelly RG, Le Garrec JF, Nicolas JF, Meilhac SM & Buckingham M (2010) Clonal analysis reveals common lineage relationships between head muscles and second heart field derivatives in the mouse embryo. *Development* **137**, 3269-3279.
21. Jacob M, Jacob HJ, Wachtler F & Christ B (1984) Ontogeny of avian extrinsic ocular muscles. I. A light- and electron-microscopic study. *Cell Tissue Res* **237**, 549-557.
22. Evans DJ & Noden DM (2006) Spatial relations between avian craniofacial neural crest and paraxial mesoderm cells. *Dev Dyn* **235**, 1310-1325.
23. Theis S, Patel K, Valasek P, Otto A, Pu Q, Harel I, Tzahor E, Tajbakhsh S, Christ B & Huang R (2010) The occipital lateral plate mesoderm is a novel source for vertebrate neck musculature. *Development* **137**, 2961-2971.
24. Davidson EH (2010) Emerging properties of animal gene regulatory networks. *Nature* **468**, 911-920.
25. Braun T, Bober E, Rudnicki MA, Jaenisch R & Arnold HH (1994) MyoD expression marks the onset of skeletal myogenesis in Myf-5 mutant mice. *Development* **120**, 3083-3092.
26. Summerbell D, Halai C & Rigby PWJ (2002) Expression of the myogenic regulatory factor Mrf4 precedes or is contemporaneous with that of Myf5 in the somitic bud. *Mech Dev* **117**, 331-335.
27. Kassar-Duchossoy L, Gayraud-Morel B, Gomès D, Rocancourt D, Buckingham M, Shinin V & Tajbakhsh S (2004) Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature* **431**, 466-471.

28. Rudnicki MA, Schnegelsberg PN, Stead RH, Braun T, Arnold HH & Jaenisch R (1993) MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75**, 1351-1359.
29. Sambasivan R, Gayraud-Morel B, Dumas G, Cimper C, Paisant S, Kelly RG & Tajbakhsh S (2009) Distinct regulatory cascades govern extraocular and pharyngeal arch muscle progenitor cell fates. *Dev Cell* **16**, 810-821 (erratum appears in *Dev Cell* **17**, 150).
30. Hasty P, Bradley A, Morris JH, Edmondson DG, Venuti JM, Olson EN & Klein WH (1993) Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* **364**, 501-506.
31. Nabeshima Y, Hanaoka K, Hayasaka M, Esumi E, Li S, Nonaka & Nabeshima Y (1993) Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **364**, 532-535.
32. Gerber AN, Klesert TR, Bergstrom DA & Tapscott SJ (1997) Two domains of MyoD mediate transcriptional activation of genes in repressive chromatin: a mechanism for lineage determination in myogenesis. *Genes Dev* **11**, 436-450.
33. Roy K, de la Serna IL & Imbalzano AN (2002) The myogenic basic helix-loop-helix family of transcription factors shows similar requirements for SWI/SNF chromatin remodeling enzymes during muscle differentiation in culture. *J Biol Chem* **277**, 33818-33824.
34. de la Serna IL, Ohkawa Y, Berkes CA, Bergstrom DA, Dacwag CS, Tapscott SJ & Imbalzano AN (2005) MyoD targets chromatin remodeling complexes to the myogenin locus prior to forming a stable DNA-bound complex. *Mol Cell Biol* **25**, 3997-4009.
35. Tajbakhsh S, Rocancourt D, Cossu G & Buckingham M (1997) Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* **89**, 127-138.
36. Relaix F, Rocancourt D, Mansouri A & Buckingham M (2005) A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**, 948-953.
37. Mankoo BS, Skuntz S, Harrigan I, Grigorieva E, Candia A, Wright CV, Arnheiter H & Pachnis V (2003) The concerted action of Meox homeobox genes is required upstream of genetic pathways essential for the formation, patterning and differentiation of somites. *Development* **130**, 4655-4664.
38. Jerome LA & Papaioannou VE (2001) DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nat Genet* **27**, 286-291.
39. Kelly RG, Jerome-Majewska LA & Papaioannou VE (2004) The del22q11.2 candidate gene Tbx1 regulates branchiomeric myogenesis. *Hum Mol Genet* **13**, 2829-2840.
40. Dastjerdi A, Robson L, Walker R, Hadley J, Zhang Z, Rodriguez-Niedenführ M, Ataliotis P, Baldini A, Scambler P & Francis-West P (2007) Tbx1 regulation of myogenic differentiation in the limb and cranial mesoderm. *Dev Dyn* **236**, 353-363.

41. Dong F, Sun X, Liu W, Ai D, Klysik E, Lu MF, Hadley J, Antoni L, Chen L, Baldini A, Francis-West P & Martin JF (2006) Pitx2 promotes development of splanchnic mesoderm-derived branchiomeric muscle. *Development* **133**, 4891-4899.
42. Shih HP, Gross MK & Kioussi C (2007) Cranial muscle defects of Pitx2 mutants result from specification defects in the first branchial arch. *Proc Natl Acad Sci U S A* **104**, 5907-5912.
43. von Scheven G, Bothe I, Ahmed MU, Alvares LE & 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.
44. Lu JR, Bassel-Duby R, Hawkins A, Chang P, Valdez R, Wu H, Gan L, Shelton JM, Richardson JA & Olson EN (2002) Control of facial muscle development by MyoR and capsulin. *Science* **298**, 2378-2381.
45. Moncaut N, Cross JW, Siligan C, Keith A, Taylor K, Rigby PWJ & Carvajal JJ (2012) Musculin and TCF21 coordinate the maintenance of myogenic regulatory factor expression levels during mouse craniofacial development. *Development* **139**, 958-967.
46. Harel I, Nathan E, Tirosh-Finkel L, Zigdon H, Guimarães-Camboa N, Evans SM & Tzahor E (2009) Distinct origins and genetic programs of head muscle satellite cells. *Dev Cell* **16**, 822-832.
47. Harel I, Maezawa Y, Avraham R, Rinon A, Ma HY, Cross JW, Leviatan N, Hegesh J, Roy A, Jacob-Hirsch J, Rechavi G, Carvajal J, Tole S, Kioussi C, Quaggin S & Tzahor E (2012) Pharyngeal mesoderm regulatory network controls cardiac and head muscle morphogenesis. *Proc Natl Acad Sci U S A* **109**, 18839-18844.
48. Tzahor E, Kempf H, Mootoosamy RC, Poon AC, Abzhanov A, Tabin CJ, Dietrich S & Lassar AB (2003) Antagonists of Wnt and BMP signaling promote the formation of vertebrate head muscle. *Genes Dev* **17**, 3087-3099.
49. Cao Y, Yao Z, Sarkar D, Lawrence M, Sanchez GJ, Parker MH, MacQuarrie KL, Davison J, Morgan MT, Ruzzo WL, Gentleman RC & Tapscott SJ (2010) Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. *Dev Cell* **18**, 662-674.
50. Fong AP, Yao Z, Zhong JW, Cao Y, Ruzzo WL, Gentleman RC & Tapscott SJ (2012) Genetic and epigenetic determinants of neurogenesis and myogenesis. *Dev Cell* **22**, 721-735.
51. Soleimani VD, Yin H, Jahani-Asl A, Ming H, Kockx CE, van Ijcken WF, Grosveld F & Rudnicki MA (2012) Snail regulates MyoD binding-site occupancy to direct enhancer switching and differentiation-specific transcription in myogenesis. *Mol Cell* **47**, 457-468.
52. Blum R, Vethantham V, Bowman C, Rudnicki M & Dynlacht BD (2012) Genome-wide identification of enhancers in skeletal muscle: the role of MyoD1. *Genes Dev* **26**, 2763-2779.
53. MacQuarrie KL, Yao Z, Fong AP, Diede SJ, Rudzinski ER, Hawkins DS and Tapscott SJ (2013) Comparison of genome-wide binding of MyoD in normal human

myogenic cells and rhabdomyosarcomas identifies regional and local suppression of promyogenic transcription factors. *Mol Cell Biol* **33**, 773-784.

54. Carvajal JJ & Rigby PWJ (2010) Regulation of gene expression in vertebrate skeletal muscle. *Exp Cell Res* **316**, 3014-3018.

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

56. Daubas P & Buckingham ME (2013) Direct molecular regulation of the myogenic determination gene Myf5 by Pax3, with modulation by Six1/4 factors, is exemplified by the -111kb-Myf5 enhancer. *Dev Biol* **376**, 236-244.

57. Soleimani VD, Punch VG, Kawabe Y, Jones AE, Palidwor GA, Porter CJ, Cross JW, Carvajal JJ, Kockx CE, van IJcken WF, Perkins TJ, Rigby PW, Grosveld F & Rudnicki MA (2012b) Transcriptional dominance of Pax7 in adult myogenesis is due to high-affinity recognition of homeodomain motifs. *Dev Cell* **22**, 1208-1220.

58. Summerbell D, Ashby PR, Coutelle O, Cox D, Yee S & Rigby PWJ (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.

59. Teboul L, Hadchouel J, Daubas P, Summerbell D, Buckingham M & Rigby PWJ (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.

60. Anderson C, Williams VC, Moyon B, Daubas P, Tajbakhsh S, Buckingham ME, Shiroishi T, Hughes SM & Borycki AG (2012) Sonic hedgehog acts cell-autonomously on muscle precursor cells to generate limb muscle diversity. *Genes Dev* **26**, 2103-2117.

61. Hu JK, McGlenn E, Harfe BD, Kardon G & Tabin CJ (2012) Autonomous and nonautonomous roles of Hedgehog signaling in regulating limb muscle formation. *Genes Dev* **26**, 2088-2102.

62. Coutelle O, Blagden CS, Hampson R, Halai C, Rigby PW & Hughes SM (2001) Hedgehog signalling is required for maintenance of myf5 and myoD expression and timely terminal differentiation in zebrafish adaxial myogenesis. *Dev Biol* **236**, 136-150.

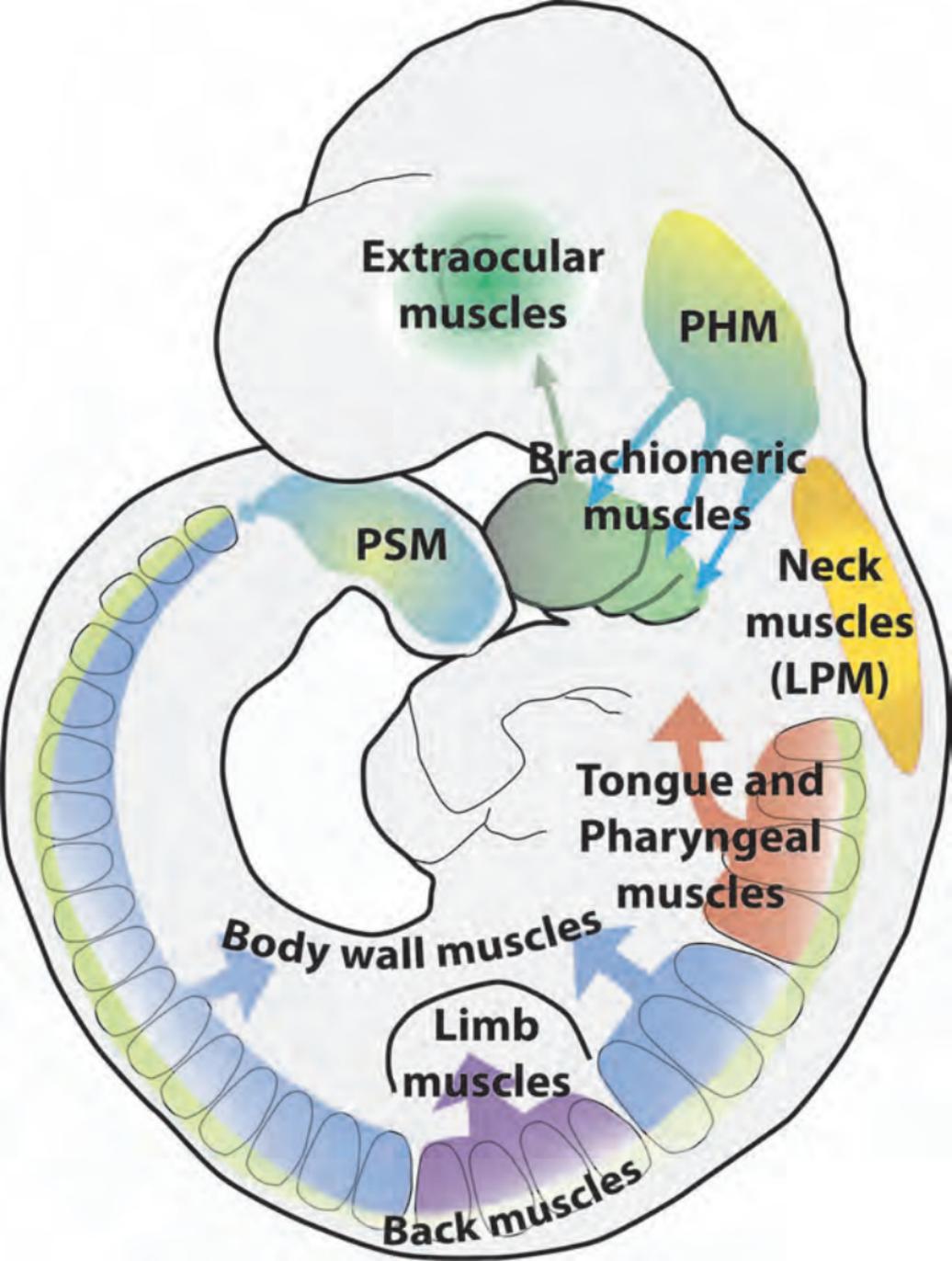
63. Teboul L, Summerbell D & Rigby PWJ (2003) The initial somitic phase of Myf5 expression requires neither Shh signaling nor Gli regulation. *Genes Dev* **17**, 2870-2874.

64. He L & Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* **5**, 522-531 (erratum appears in *Nat Rev Genet* 5, 631).

65. Lai EC (2002) Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet* **30**, 363-364.

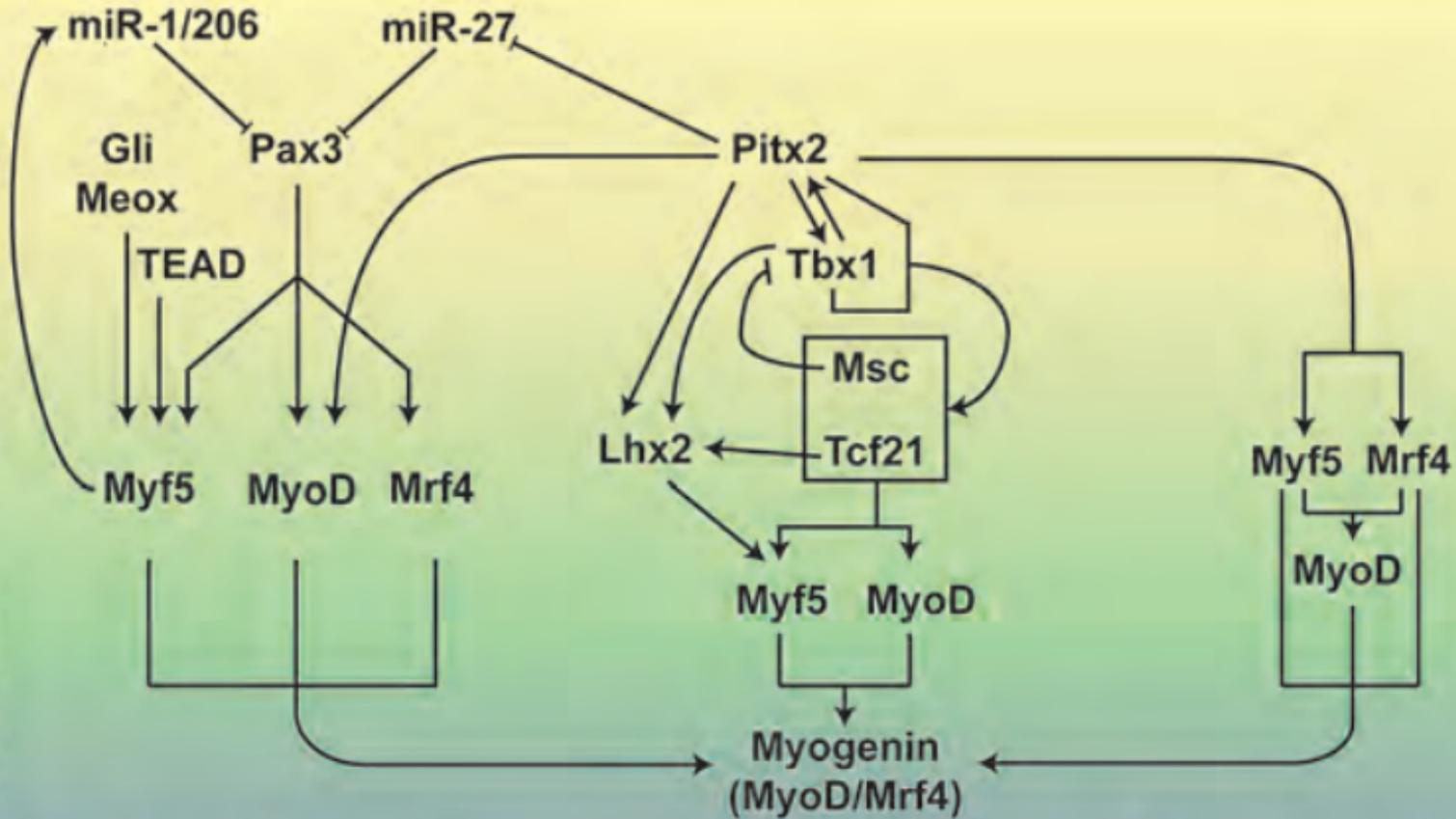
66. O'Rourke JR, Georges SA, Seay HR, Tapscott SJ, McManus MT, Goldhamer DJ, Swanson MS & Harfe BD (2007) Essential role for Dicer during skeletal muscle development. *Dev Biol* **311**, 359-368.
67. Dmitriev P, Barat A, Poleskaya A, O'Connell MJ, Robert T, Dessen P, Walsh TA, Lazar V, Turki A, Carnac G, Laoudj-Chenivresse D, Lipinski M & Vassetzky YS (2103) Simultaneous miRNA and mRNA transcriptome profiling of human myoblasts reveals a novel set of myogenic differentiation-associated miRNAs and their target genes. *BMC Genomics* **14**, 265.
68. McCarthy JJ (2008) MicroRNA-206: the skeletal muscle-specific myomiR. *Biochim Biophys Acta* **1779**, 682-691.
69. Sweetman D, Goljanek K, Rathjen T, Oustanina S, Braun T, Dalmay T & Münsterberg A (2008) Specific requirements of MRFs for the expression of muscle specific microRNAs, miR-1, miR-206 and miR-133. *Dev Biol* **321**, 491-499.
70. Goljanek-Whysall K, Pais H, Rathjen T, Sweetman D, Dalmay T & Münsterberg A (2012) Regulation of multiple target genes by miR-1 and miR-206 is pivotal for C2C12 myoblast differentiation. *J Cell Sci* **125**, 3590-3600
71. Crist CG, Montarras D, Pallafacchina G, Rocancourt D, Cumano A, Conway SJ & Buckingham M (2009) Muscle stem cell behavior is modified by microRNA-27 regulation of Pax3 expression. *Proc Natl Acad Sci U S A* **106**, 13383-13387.
72. Chen JF, Tao Y, Li J, Deng Z, Yan Z, Xiao X & Wang DZ (2010) microRNA-1 and microRNA-206 regulate skeletal muscle satellite cell proliferation and differentiation by repressing Pax7. *J Cell Biol* **190**, 867-879.
73. Dey BK, Gagan J & Dutta A (2011) miR-206 and -486 induce myoblast differentiation by downregulating Pax7. *Mol Cell Biol* **31**, 203-214.
74. Lozano-Velasco E, Contreras A, Crist C, Hernández-Torres F, Franco D & Aránega AE (2011) Pitx2c modulates Pax3+/Pax7+ cell populations and regulates Pax3 expression by repressing miR27 expression during myogenesis. *Dev Biol* **357**, 165-178.
75. Cheung TH, Quach NL, Charville GW, Liu L, Park L, Edalati A, Yoo B, Hoang P & Rando TA (2012) Maintenance of muscle stem-cell quiescence by microRNA-489. *Nature* **482**, 524-528.
76. Crist CG, Montarras D & Buckingham M (2012) Muscle satellite cells are primed for myogenesis but maintain quiescence with sequestration of Myf5 mRNA targeted by microRNA-31 in mRNP granules. *Cell Stem Cell* **11**, 118-126.
77. van Rooij E, Quiat D, Johnson BA, Sutherland LB, Qi X, Richardson JA, Kelm RJ Jr & Olson EN (2009) A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev Cell* **17**, 662-673.
78. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, Weinstock GM, Wilson RK, Gibbs RA, Kent WJ, Miller W & Haussler D (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* **15**, 1034-1050.

79. ENCODE Project Consortium (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799-816.
80. Margulies EH, Cooper GM, Asimenos G, Thomas DJ, Dewey CN, Siepel A, Birney E, Keefe D, Schwartz AS, Hou M, Taylor J, Nikolaev S, Montoya-Burgos JI, Löytynoja A, Whelan S, Pardi F, Massingham T, Brown JB, Bickel P, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Stone EA, Rosenbloom KR, Kent WJ, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VV, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Hinrichs A, Trumbower H, Clawson H, Zweig A, Kuhn RM, Barber G, Harte R, Karolchik D, Field MA, Moore RA, Matthewson CA, Schein JE, Marra MA, Antonarakis SE, Batzoglou S, Goldman N, Hardison R, Haussler D, Miller W, Pachter L, Green ED & Sidow A (2007) Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. *Genome Res* **17**, 760-774.
81. Qureshi IA & Mehler MF (2012) Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. *Nat Rev Neurosci* **13**, 528-541.
82. Aalto AP & Pasquinelli AE (2012) Small non-coding RNAs mount a silent revolution in gene expression. *Curr Opin Cell Biol* **24**, 333-340.
83. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G & Greenberg ME (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182-187.
84. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74.



determination

differentiation



Trunk/Limb

Head

Extraocular