Making muscle: skeletal myogenesis in vivo and in vitro

Jérôme Chal1,2,3 and Olivier Pourquié1,2,3,4,*

ABSTRACT

Skeletal muscle is the largest tissue in the body and is essential for motion and support. Skeletal muscles are distinct from both cardiac and smooth muscles in that they can be voluntarily controlled by the organism. They are composed of bundles of striated myofibers that consist of elongated multinucleated syncytia. These fibers are surrounded by a basal lamina and are filled with a highly organized cytoskeleton composed of myofibrils. Skeletal muscles of the body arise from the somites, transient embryonic structures that originate from the paraxial mesoderm. By contrast, muscles of the head and neck derive from the anterior paraxial mesoderm, which does not form somites, and which will not be discussed in this Review.

Despite its regenerative capability, compromised muscle function is a hallmark of a number of conditions, ranging from developmental disorders to rhabdomyosarcoma and muscular dystrophies (Emery, 2002). In the adult, acute muscle injury such as volumetric loss, late-onset muscle dystrophies, neuromuscular degenerative diseases, cachexia as well as aging (sarcopenia) can result in significant muscular impairment leading to a severely compromised quality of life. There is an acute need for novel therapies to treat muscle diseases, and understanding the developmental basis of skeletal muscle formation and function holds value for the elucidation and possible treatment of muscle pathologies. Notably, it is well known that rhabdomyosarcomas exhibit undifferentiated myogenic features (Dagher and Helman, 1999), and that during adult muscle regeneration, developmental programs are partially reactivated in the injured tissue (Dumont et al., 2015). Therefore, a better understanding of myogenesis can also shed light on the mechanisms of muscle disease and regeneration.

Understanding how muscle is formed in vivo will also help pave the way for recreating muscle tissue in vitro from pluripotent stem cells (PSCs), either embryonic (ESCs) or induced (iPSCs). Although this field has lagged behind other lineages such as cardiac, neural and endodermal, the in vitro transposition of early signaling events as they occur during paraxial mesoderm specification in the embryo has led to recent success in this area (Chal et al., 2015; Crist, 2017; Hicks and Pyle, 2015). This was made possible thanks to the wealth of knowledge accumulated over decades of developmental studies in model organisms. The ability to generate skeletal muscle in vitro opens up new avenues for deciphering essential but poorly understood aspects of skeletal myogenesis such as myoblast fusion and satellite cell differentiation, and might also lead to important breakthroughs in disease modeling, drug screening and cell therapeutics. In this Review, we first provide a comprehensive overview of skeletal myogenesis in vivo, from its earliest developmental origin in the paraxial mesoderm to the formation of mature myofibers. We then focus on recent efforts to recapitulate muscle specification and skeletal myogenesis from PSCs in vitro, covering both directed differentiation and direct reprogramming approaches.

Introduction

With more than 600 individual muscles, skeletal muscle represents the largest tissue mass of the body and is essential for motion and support. Skeletal muscles are distinct from both cardiac and smooth muscles in that they can be voluntarily controlled by the organism. They are composed of bundles of striated myofibers that consist of elongated multinucleated syncytia. These fibers are surrounded by a basal lamina and are filled with a highly organized cytoskeleton composed of myofibrils. Skeletal muscles of the body arise from the somites, transient embryonic structures that originate from the paraxial mesoderm. By contrast, muscles of the head and neck derive from the anterior paraxial mesoderm, which does not form somites, and which will not be discussed in this Review.

Despite its regenerative capability, compromised muscle function is a hallmark of a number of conditions, ranging from developmental disorders to rhabdomyosarcoma and muscular dystrophies (Emery, 2002). In the adult, acute muscle injury such as volumetric loss, late-onset muscle dystrophies, neuromuscular degenerative diseases, cachexia as well as aging (sarcopenia) can result in significant muscular impairment leading to a severely compromised quality of life. There is an acute need for novel therapies to treat muscle diseases, and understanding the developmental basis of skeletal muscle formation and function holds value for the elucidation and possible treatment of muscle pathologies. Notably, it is well known that rhabdomyosarcomas exhibit undifferentiated myogenic features (Dagher and Helman, 1999), and that during adult muscle regeneration, developmental programs are partially reactivated in the injured tissue (Dumont et al., 2015). Therefore, a better understanding of myogenesis can also shed light on the mechanisms of muscle disease and regeneration.

Understanding how muscle is formed in vivo will also help pave the way for recreating muscle tissue in vitro from pluripotent stem cells (PSCs), either embryonic (ESCs) or induced (iPSCs). Although this field has lagged behind other lineages such as cardiac, neural and endodermal, the in vitro transposition of early signaling events as they occur during paraxial mesoderm specification in the embryo has led to recent success in this area (Chal et al., 2015; Crist, 2017; Hicks and Pyle, 2015). This was made possible thanks to the wealth of knowledge accumulated over decades of developmental studies in model organisms. The ability to generate skeletal muscle in vitro opens up new avenues for deciphering essential but poorly understood aspects of skeletal myogenesis such as myoblast fusion and satellite cell differentiation, and might also lead to important breakthroughs in disease modeling, drug screening and cell therapeutics. In this Review, we first provide a comprehensive overview of skeletal myogenesis in vivo, from its earliest developmental origin in the paraxial mesoderm to the formation of mature myofibers. We then focus on recent efforts to recapitulate muscle specification and skeletal myogenesis from PSCs in vitro, covering both directed differentiation and direct reprogramming approaches.

From the beginning: the developmental origin of skeletal muscle

Skeletal muscles originate from the paraxial mesoderm, a tissue that forms in the primitive streak/blastopore during gastrulation and later in the tail bud during embryonic axis elongation (Fig. 1). The nascent paraxial mesoderm constitutes the presomitic mesoderm at the posterior tip of the embryo. The presomitic mesoderm is a transient tissue that can be further subdivided into an immature posterior and a committed anterior region, the latter of which segments to form the somites. It is within the somites that skeletal myogenesis is initiated with the specification of the premyogenic progenitors and skeletal myoblasts. Several phases of proliferation and differentiation lead to the formation of multinucleated myofibers from the fusion of mononucleated myocytes. In the following sections, we discuss the key cellular and molecular events that regulate the progression through these developmental steps in the embryo.

Specification of paraxial mesoderm progenitors

The paraxial mesoderm is composed of two bilateral strips of tissue flanking the neural tube and notochord. Posteriorly, these strips are unsegmented and form the presomitic mesoderm, while anteriorly they are composed of somites which define the embryonic segments (Fig. 1A). In amniotes, such as mouse or chicken, the paraxial mesoderm forms by ingestion of the epiblast at the level of the anterior streak and later on from the tail bud. Formation of the primitive streak and activation of Wnt3 and the early mesoderm

1Department of Pathology, Brigham and Women’s Hospital, 77 Avenue Louis Pasteur, Boston, MA 02115, USA. 2Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA. 3Harvard Stem Cell Institute, 77 Avenue Louis Pasteur, Boston, MA 02115, USA. 4Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS (UMR 7104), Inserm U964, Université de Strasbourg, 67400 Illkirch-Graffenstaden, France.

*Author for correspondence (pourqui@hms.harvard.edu)
marker brachyury (T) depend on several signaling factors including Nodal and BMP4 (Beddington and Robertson, 1999; Liu et al., 1999; Ramkumar and Anderson, 2011; Tam and Loebel, 2007). Detailed fate mapping and grafting studies have identified several types of progenitors that give rise to the paraxial mesoderm. One type corresponds to a resident cell population able to give rise to both paraxial mesoderm and neural tube derivatives, the so-called neuromesodermal progenitors, which co-express the genes T and Sox2 (Garriock et al., 2015; Takemoto et al., 2011; Tzouanacou et al., 2009). A second type of progenitor gives rise only to paraxial mesoderm, while a third type can give rise to both paraxial mesoderm and lateral plate derivatives (Iimura et al., 2007; Stern and Canning, 1990; Wymeersch et al., 2016), and a fourth type can give rise to paraxial mesoderm and notochord (Selleck and Stern, 1991).

Paraxial mesoderm specification is concomitant to precursor exit from the progenitor zone and their entry to the posterior presomitic mesoderm, which are processes largely controlled by the Wnt and fibroblast growth factor (FGF) signaling pathways (Fig. 1A) (Ciruna and Rossant, 2001; Takada et al., 1994; Yamaguchi et al., 1994). Key Wnt/FGF targets including the transcription factors T, Tbx6 and Msgn1 are expressed in partially overlapping domains of the presomitic mesoderm (Chapman et al., 1996; Ciruna and Rossant, 2001; Yamaguchi et al., 1999; Yoon et al., 2000). These transcription factors are essential for paraxial mesoderm specification.
specification and patterning (Chapman et al., 1996; Nowotschin et al., 2012; van der Velden et al., 2006; Wilkinson et al., 1990; Yoon and Wold, 2000). In the absence of Wnt or FGF signaling, the mouse embryo is truncated and ectopic neural tissue forms in place of the posterior paraxial mesoderm (Boulet and Capecechi, 2012; Ciruna and Rossant, 2001; Takada et al., 1994). Thus, Wnt and FGF signaling are required for the differentiation of neuromesodermal progenitors toward a paraxial mesoderm fate at the expense of neural fate (Garriock et al., 2015; Jurberg et al., 2014). Downregulation of FGF and Wnt signaling in the tail bud signals the arrest of paraxial mesoderm production and the end of axis elongation (Cambray and Wilson, 2007; Denans et al., 2015; Olivera-Martinez and Storey, 2007; Tenin et al., 2010). Wnt3a is required for expression of Fgf8 in the tail bud (Aulehla et al., 2003), whereas Fgf4 and Fgf8 are necessary for Wnt signaling in the nascent presomitic mesoderm (Boulet and Capecechi, 2012; Naiche et al., 2011). Thus, the Wnt and FGF pathways form a closed regulatory loop that controls the specification and formation of the paraxial mesoderm.

BMP signaling also plays an important role in the specification and fate of mesoderm, both along the anterior-posterior axis at the early primitive streak stage and later along the mediolateral axis during late primitive streak and tail bud stages (Winnier et al., 1995). BMP4 is secreted by the posterior primitive streak and lateral tissues such as the lateral plate and extraembryonic mesoderm. BMP action is counteracted by opposite gradients of noggin and other BMP antagonists produced by the axial structures of the embryo (McMahon et al., 1998; Pourquié et al., 1996; Reshef et al., 1998; Tonegawa et al., 1997). This results in the establishment of a BMP signaling gradient that controls the mediolateral fates of mesoderm. Each mesodermal type, from the notochord to the extraembryonic mesoderm, requires progressively higher levels of BMP signaling for their specification (Kishigami and Mishina, 2005). Not surprisingly, formation of the paraxial mesoderm is exquisitely sensitive to changes in BMP signaling. For example, when a bead producing the BMP inhibitor noggin is grafted in the posterior primitive streak, which normally gives rise to lateral plate, these cells are converted to a paraxial mesoderm fate, leading to the formation of ectopic somite columns (Tonegawa et al., 1997). In another study, a graft of quail posterior primitive streak, which would normally give rise to the lateral plate mesoderm, together with noggin-producing cells led to the formation of ectopic somites in a chicken embryo host (Streit and Stern, 1999). Also, mouse embryos chimeric for a BMP receptor null mutation show an expansion of the paraxial mesoderm domain, evidenced by the formation of ectopic rows of somites (Miura et al., 2006). Strikingly, this phenotype can be rescued by inhibiting FGF signaling, suggesting that BMP and FGF may antagonize each other. Together, this argues in favor of some plasticity of the paraxial mesoderm. Moreover, the parasomal mesoderm precursors of the primitive streak, which can still be signaling, suggesting that BMP and FGF may antagonize each other. This results in the establishment of a BMP signaling gradient along the paraxial mesoderm as well as muscle and cartilage formation is observed, although improperly patterned (Bessho et al., 2003; Burgess et al., 1996; Dale et al., 2003; Saga et al., 1997) (Fig. 1A). In the anterior presomitic mesoderm, the posterior Wnt/FGF gradients are counteracted by retinoic acid (RA) produced by the somitic region (Fig. 1A). The tail bud containing the paraxial mesoderm precursors is protected from the differentiating action of RA by the expression of the RA-degrading enzyme Cyp26 (Abu-Abed et al., 2001; Iulianella et al., 1999; Sakai et al., 2001).

The bilateral stripes of Mesp2 define the anterior and posterior boundaries of the future segment (Ogimuna et al., 2008; Takahashi et al., 2000). Anterior to the determination front, the newly formed segment acquires its anterior posterior identities in response to a dynamic process largely controlled by Notch signaling (Chal and Pourquié, 2009). The cells of the posterior presomitic mesoderm are mesenchymal and highly motile (Bénazéraf et al., 2010; Chalamalasetty et al., 2014; Delfini et al., 2005), whereas the anterior presomitic mesoderm undergoes a mesenchymal-to-epithelial transition (MET), becoming progressively organized into dorsal and ventral epithelial layers surrounding a mesenchymal core (Duband et al., 1987; Martins et al., 2009). This process is controlled by Tcf15 (paraxis), a transcription factor activated in the anterior presomitic mesoderm by Wnt5 from the dorsal ectoderm (Burgess et al., 1996; Linker et al., 2005; Sošić et al., 1997). In the anteriormost presomitic mesoderm, a posterior fissure forms at the boundary between Mesp2-positive and -negative cells, resulting in the formation of an epithelial block of tissue – a new somite. This tissue remodeling involves Eph-ephrins (Barrios et al., 2003; Nakajima et al., 2006; Watanabe et al., 2009), cadherins (Chal et al., 2017b; Horikawa et al., 1999) and the small G proteins Cdc42 and Rac1 (Nakaya et al., 2004). Even so, in mouse or chicken embryos with severe segmentation and/or epithelialization defects, timely differentiation and/or epithelialization defects, delaying differentiation of the paraxial mesoderm as well as muscle and cartilage formation is observed, although improperly patterned (Bessho et al., 2003; Burgess et al., 1996; Dale et al., 2003; Saga et al., 1997). Thus, the differentiation and patterning of the paraxial mesoderm can be genetically uncoupled, suggesting that the underlying molecular mechanisms are relatively independent.

Differentiation of the presomitic mesoderm
The specification of the future pairs of embryonic segments – the somites – is the result of highly dynamic molecular processes within the presomitic mesoderm. This involves a molecular oscillator known as the segmentation clock, which generates pulses of Notch, FGF and Wnt signaling to control the periodic production of somites (reviewed by Hubaud and Pourquié, 2014). The posterior domain of the paraxial mesoderm is composed of four consecutive transcriptional domains: the tail bud, the posterior presomitic mesoderm, the anterior presomitic mesoderm, and the forming somite (SO) (Chal et al., 2015). The tail bud domain contains the paraxial mesoderm progenitors and is exposed to the highest Wnt/FGF signaling activity (Aulehla et al., 2003; Chal et al., 2015; Naiche et al., 2011). More anteriorly, cells of the posterior presomitic mesoderm are characterized by the expression of genes such as Msgn1 (Fig. 1). In this domain, the Wnt/FGF pathways are still highly active and are essential for maintenance of paraxial mesoderm identity and for segmentation clock oscillations (Aulehla and Pourquié, 2008; Chal et al., 2015; Dunty et al., 2008; Jurberg et al., 2014). The Wnt/FGF activity gradient along the presomitic mesoderm is proposed to define a threshold, at which level cells become competent to respond to the clock by activating the expression of segmentation genes such as Mesp2 (Aulehla et al., 2003; Dubrulle et al., 2001). This specific threshold is called the determination front or wavefront and is roughly positioned at the posterior boundary of the anterior third of the presomitic mesoderm (Fig. 1A). Cells of the posterior presomitic mesoderm undergo abrupt signaling, metabolic and transcriptional changes as they enter the anterior presomitic mesoderm (Chal et al., 2015; Ogimuna et al., 2017; Ozbudak et al., 2010) (Fig. 1A), including downregulation of Msgn1 and activation of Mesp2, Pax5, Foxc1/2 and Meox1/2 (Goulding et al., 1991; Kume et al., 2001; Mankoo et al., 2003; Saga et al., 1997) (Fig. 1B). In the anterior presomitic mesoderm, the posterior Wnt/FGF gradients are counteracted by retinoic acid (RA) produced by the somitic region (Fig. 1A). The tail bud containing the paraxial mesoderm precursors is protected from the differentiating action of RA by the expression of the RA-degrading enzyme Cyp26 (Abu-Abed et al., 2001; Iulianella et al., 1999; Sakai et al., 2001).

The bilateral stripes of Mesp2 define the anterior and posterior boundaries of the future segment (Ogimuna et al., 2008; Takahashi et al., 2000). Anterior to the determination front, the newly formed segment acquires its anterior-posterior identities in response to a dynamic process largely controlled by Notch signaling (Chal and Pourquié, 2009). The cells of the posterior presomitic mesoderm are mesenchymal and highly motile (Bénazéraf et al., 2010; Chalamalasetty et al., 2014; Delfini et al., 2005), whereas the anterior presomitic mesoderm undergoes a mesenchymal-to-epithelial transition (MET), becoming progressively organized into dorsal and ventral epithelial layers surrounding a mesenchymal core (Duband et al., 1987; Martins et al., 2009). This process is controlled by Tcf15 (paraxis), a transcription factor activated in the anterior presomitic mesoderm by Wnt5 from the dorsal ectoderm (Burgess et al., 1996; Linker et al., 2005; Sošić et al., 1997). In the anteriormost presomitic mesoderm, a posterior fissure forms at the boundary between Mesp2-positive and -negative cells, resulting in the formation of an epithelial block of tissue – a new somite. This tissue remodeling involves Eph-ephrins (Barrios et al., 2003; Nakajima et al., 2006; Watanabe et al., 2009), cadherins (Chal et al., 2017b; Horikawa et al., 1999) and the small G proteins Cdc42 and Rac1 (Nakaya et al., 2004). Even so, in mouse or chicken embryos with severe segmentation and/or epithelialization defects, timely differentiation of the paraxial mesoderm as well as muscle and cartilage formation is observed, although improperly patterned (Bessho et al., 2003; Burgess et al., 1996; Dale et al., 2003; Saga et al., 1997). Thus, the differentiation and patterning of the paraxial mesoderm can be genetically uncoupled, suggesting that the underlying molecular mechanisms are relatively independent.

Compartmentalization of somites
Soon after their formation, somites become compartmentalized along the dorsoventral axis into a dorsal epithelial dermomyotome and a
ventral mesenchymal sclerotome. The dermomyotome gives rise to skeletal muscle, brown fat and dermis of the back, whereas the ventral sclerotome produces the axial skeleton and tendons. At the time of their formation, each somite is composed of an anterior *Tbx18* and a posterior *Uncx* compartment with distinct derivatives (reviewed by Chal and Pourquié, 2009). In the chicken embryo, newly formed somites can also be further subdivided into a medial and a lateral compartment (Olivera-Martinez et al., 2000; Ordahl and Le Douarin, 1992; Selleck and Stern, 1991), which exhibit different fates and express different sets of genes (Martins et al., 2009; Pourquié et al., 1996). Cells of the lateral somite give rise to the hyapxial muscles of the limbs or the intercostals, whereas the medial somite forms the sclerotome, dermis of the back and paraxial muscles (Fig. 2) (Olivera-Martinez et al., 2000; Ordahl and Le Douarin, 1992).

Cells of the newly formed somites are not yet committed to a specific lineage (Aoyama and Asamoto, 1988). Grafting experiments and *in vitro* explant cultures have shown that somitic domains are progressively specified in response to factors secreted by the surrounding tissues, namely the dorsal ectoderm, neural tube, notochord and lateral plate (Fig. 2A) (reviewed by Christ and Scaal, 2008). Wnt, BMP and Shh represent the major signaling pathways for the induction of different somitic fates (reviewed by Marcelle et al., 2008). Ventrally, Shh signaling from the notochord and floor plate specifies the sclerotomal compartment, which downregulates *Pax3* and upregulates *Pax1* and *Nkx3.2* expression (Fig. 2A) (Fan and Tessier-Lavigne, 1994; Johnson et al., 1994; Murtaugh et al., 1999). Ventrally, Shh signaling from the notochord and floor plate specifies the sclerotomal compartment, which downregulates *Pax3* and upregulates *Pax1* and *Nkx3.2* expression (Fig. 2A) (Fan and Tessier-Lavigne, 1994; Johnson et al., 1994; Murtaugh et al., 1999).

**Myogenesis in the embryo and the adult**

Initiation of myogenesis in the somite

In mouse and chicken embryos, the first sign of myogenesis is the activation of the myogenic factor Myf5 in cells of the dorsomedial part of the newly formed somite (Ott et al., 1991; Pownall and Emerson, 1992) (Fig. 1B). The dorsal epithelial dermomyotome,
which expresses Pax3, becomes subdivided into a central domain, a dorsomedial lip, anterior and posterior lips, and a ventrolateral lip (Gros et al., 2004). Soon after dermomyotome formation, cells in the dorsomedial lip begin to express Myf5 and to downregulate Pax3 (Bober et al., 1994; Ott et al., 1991; Williams and Ordahl, 1994) (Fig. 2B). The primary myotome forms as a cell layer sandwiched between the dermomyotome dorsally and the sclerotome ventrally (Ordahl, 1993).

The first postmitotic skeletal muscle cells formed in the embryo are the myocytes of the myotome (Fig. 3A). These cells express specialized cytoskeletal proteins including slow (type I, Myh7) and embryonic (Myh3) myosin heavy chains (MyHC), α-actins [cardiac (Actc1) and skeletal (Acta1)] and desmin (Babai et al., 1990; Furst et al., 1989; Lyons et al., 1991a, 1990; Sassoon et al., 1988), as well as the Notch ligand jagged 2 (Hayashi et al., 1996; Hirsinger et al., 2001) and metabolic enzymes such as β-enolase and carbonic anhydrase III (CAIII) (Condon et al., 1990; Lyons et al., 1991b; Tweedie et al., 1991). The newly formed mononucleated myocytes elongate along the anterior-posterior axis to span the entire somite length, a process controlled by Wnt11 signaling (Christ et al., 1983;
Denetclaw et al., 1997; Gros et al., 2009). More cells are progressively added to the myotome by the other dermomyotomal lips (Gros et al., 2004) and these cells fuse to existing myocytes leading to the formation of slow MyHC+ myofibers (Sieero-Mosti et al., 2014) (Fig. 3A). After formation of the myotome, the central dermomyotome loses its epithelial character and its Pax3+ cells translocate to populate the myotome, providing the myogenic precursors involved in later phases of myogenesis (Fig. 3) (Gros et al., 2005; Kahane et al., 2001; Relax et al., 2005). Myogenesis progresses as a rostral-to-caudal wave of maturation as the embryo elongates and as new pairs of somites are sequentially added (Cossu et al., 1995; Kato and Gurdon, 1993). Limb muscles derive from cells migrating from the lateral dermomyotome into the developing limb buds (Chevallier, 1979; Jacob et al., 1979). These cells progressively organize into muscle masses and form myofibers (reviewed by Buckingham et al., 2003). In the trunk and limbs, the myogenic program is controlled by a core network of transcription factors, including Pax3 and a set of muscle regulatory factors (MRFs) consisting of Myf5, MyoD (Myo1d), MRF4 (Myf6) and myogenin (Berkes and Taps, 2005; Pownall et al., 2002; Rudnicki et al., 1993; Taps, 2005). In the embryo, myogenin controls the terminal differentiation of myoblasts into myocytes (Hasty et al., 1993; Nabeshima et al., 1993; Venuti et al., 1995). Genetic studies have also identified a set of transcription factors that are upstream regulators of skeletal myogenesis, which include Rp58 (Zfp238 or Zbtb18), Meox1/2, Six1/4, Eya1/2 and Nfix (Biressi et al., 2007; Grifone et al., 2007, 2005; Mankoo et al., 1999, 2003; Yokoyama et al., 2009).

Primary and secondary myogenesis

Myogenesis during development can be separated into two phases: an early embryonic or primary phase (E10.5-E12.5 in mouse, E3-7 in chicken) and a later fetal or secondary phase (E14.5-17.5 in mouse, E8+ in chicken) (Biressi et al., 2007; Stockdale, 1992). The first phase results in the production of the primary myofibers, which derive from Pax3+ (mouse) or Pax3+/Pax7+ (chicken) dermomyotomal progenitors (Horst et al., 2006; Hutcheson et al., 2009; Otto et al., 2006). These primary myofibers form the early myotomes and limb muscles, providing the templates upon which adult muscles will be built (Fig. 3) (Murphy and Kardon, 2011). They express a specific set of proteins, such as the slow MyHC and myosin light chain 1 (MyLC1, My1l) (Kelly et al., 1997). During the second phase of myogenesis in mouse, a subset of the Pax3+-myogenic progenitors begins to express Pax7 and downregulates Pax3. These Pax7+ myogenic precursors fuse among themselves or to the primary fibers and give rise to secondary or fetal fibers that express specific markers such as β-enolase, Nfix or MyLC3 (My13) (Fougerousse et al., 2001; Keller et al., 1992; Kelly et al., 1997; Messina et al., 2010). At this time, the fibers also start to express fast MyHC isoforms (Van Horn and Crow, 1989). During secondary myogenesis, muscle growth is sustained essentially by cell fusion and the addition of myonuclei from proliferating Pax7+ progenitors (White et al., 2010). This is in contrast to postnatal muscle growth, which mostly results from individual fiber hypertrophy through the addition of novel myofibrils (Gokhin et al., 2008; Sparrow and Schöck, 2009). A subset of the Pax7+ progenitors will also form the pool of adult muscle stem cells – the satellite cells (Gros et al., 2005; Kassar-Duchossoy et al., 2005; Relax et al., 2005).

Signaling controlling myogenesis

Embryonic tissues surrounding the somite provide key inductive signals for skeletal myogenesis. The notochord, neural tube and dorsal ectoderm, which produce Shh and Wnt signals, are crucial for this process (Münsterberg et al., 1995; Münsterberg and Lassar, 1995; Rios et al., 2011). The neural crest cells that migrate between the neural tube and somites also play a key role in the onset of myogenesis by triggering Myf5 activation in dermomyotome cells. Neural crest cells, which carry Wnt1 at their surface and also express delta-like-1 (Dll1), activate Notch signaling in the cells they contact in the dorsomedial lip. Notch activation in the lip cells in turn inhibits GSK3β activity and leads to the stabilization of Snail, resulting in the delamination of cells into the nascent myotome (Rios et al., 2011; Serralbo and Marcella, 2014; Sieiro et al., 2016) (Fig. 2B). Another signaling molecule, hepatocyte growth factor (HGF), supports different aspects of myogenesis during development. In particular, HGF produced by the lateral plate mesoderm is essential for the proper migration of myoblasts (Andermarcher et al., 1996; Bladt et al., 1995; Brand-Saberi et al., 1996; Takayama et al., 1996). Insulin signaling has been shown to act with Wnt signaling to promote myogenesis and myoblast fusion (Charge and Rudnicki, 2004; van der Velden et al., 2006), while FGF signaling promotes myoblast proliferation while blocking differentiation (Itoh et al., 1996; Milasincic et al., 1996). In the context of adult regeneration, the FGF, insulin-like growth factor (IGF), HGF and Wnt signaling pathways have all been shown to play a role in satellite cell activation (Charge and Rudnicki, 2004; Flanagan-Steet et al., 2000; von Maltzahn et al., 2012). Wnt7a/Fzd7 planar cell polarity signaling is involved in the control of satellite cell self-renewal versus differentiation (Bentzinger et al., 2014, 2013; Le Grand et al., 2009).

Another important protein for myogenesis is myostatin (Gdf8), which represses muscle hypertrophy (Amthor et al., 2009; McPherron et al., 1997). In the absence of myostatin, mutants show a striking hypertrophic phenotype (Lee, 2004; reviewed by Rodriguez et al., 2014). Additionally, a number of microRNAs (miRNAs) are highly expressed, often specifically, in muscle cells (Chen et al., 2006; Kim et al., 2006; Sweetman et al., 2008). Although the identification of the signaling upstream and downstream of these miRNAs is still in its infancy (Motohashi et al., 2013), miRNAs regulate almost every aspect of myogenesis (reviewed by Ge and Chen, 2011). Finally, two additional pathways have been shown to play key roles in regulating myogenesis. In vitro work has revealed an important role for p38 MAPK intracellular signaling in the control of myogenesis, whereby inhibition of p38α/β (Mapk14/11) was shown to block myogenic differentiation and cell fusion (Cuenda and Cohen, 1996; Milasincic et al., 1996). In the context of adult regeneration, the FGF, insulin-like growth factor (IGF), HGF and Wnt signaling pathways have all been shown to play a role in satellite cell activation (Charge and Rudnicki, 2004; Flanagan-Steet et al., 2000; von Maltzahn et al., 2012). Wnt7a/Fzd7 planar cell polarity signaling is involved in the control of satellite cell self-renewal versus differentiation (Bentzinger et al., 2014, 2013; Le Grand et al., 2009).

Formation of the skeletal myofibers

Myoblast fusion, myofibrillogenesis and basal lamina

The mature skeletal myofiber contains a highly organized cytoskeleton composed of aligned myofibrils. These fibers form by fusion of myoblasts to produce multinucleated myotubes, which further mature into myofibers (reviewed by Abmayr and Pavlath, 2012). The processes of myoblast-myoblast fusion and myoblast-myotube fusion are highly regulated during development but are still poorly understood in vertebrates. In flies, muscle fusion is orchestrated by two distinct types of myoblasts, with a limited number of founder cells initiating fusion with surrounding fusion-competent myoblasts (Rochlin et al., 2010). Fusion-competent
myofibrillar proteins and associated organelles, including mitochondria, are organized into distinct sarcomeric structures. These structures are critical for the rhythmic contraction of muscles and are essential for their function.

**Myofibrillogenesis** refers to the formation of the myofibrils that fill most of the fiber sarcoplasm, extending to both extremities of the myofiber. Each myofiber is divided into repeating structures called sarcomeres, which are the basic functional units of skeletal muscle. The sarcomeres are composed of myofilaments, including actin and myosin, and are closely associated with the sarcolemma, the specialized plasma membrane of the myofiber. The sarcomeres are organized into sarcomeric units, forming the myofibrils that are integral to myofiber function.

**Myofibers and Myotubes**

Myofibers are the primary contractile units of skeletal muscle and are derived from myoblasts, which are myogenic precursors that give rise to myotubes. Myotubes are the multinucleated cells that form and differentiate into mature myofibers. Myoblasts can be isolated from skeletal muscle in vivo or in vitro, and myotubes can be generated from myoblasts in vitro by differentiation protocols. Myoblasts produce actin-based podosome-like structures that invade the muscle founder cell and form pores, allowing the transfer of the cytoplasm and nucleus from the myoblast to the founder cell (Kim et al., 2015b). In vertebrates, myoblast fusion starts with a recognition and an adhesion phase involving surface receptors. There is evidence to suggest that actin dynamics might play an important role (Laurin et al., 2008), and proteins involved in endocytosis and in membrane repair pathways have also been implicated (Demonbreun et al., 2015; Kim et al., 2015a). In vivo, elongation of the myofibers takes place by fusion of new myoblasts at their extremities and is dependent on TGFβ signaling (Gu et al., 2016; Williams and Goldspink, 1971). The cytokine IL4 has also been shown to control myoblast fusion, acting downstream of NFATC2 to control the fusion of myoblasts to myotubes (Horsley et al., 2003).

**Skeletal Myogenesis**

Skeletal myogenesis refers to the formation of muscle fibers that fill most of the fiber sarcoplasm, extending to both extremities of the fibers where they anchor to the myotendinous junction (reviewed by Sparrow and Schöck, 2009; Lenke and Schnorrer, 2017). Myofibers are composed of a regular array of contractile modules—the sarcomeres. Each myofiber is also individually surrounded by a specialized basal lamina, called the endomysium (Bowman, 1840; Sanes, 2003), and harbors a specialized plasma membrane called the sarcolemma, which provides structural stability and allows for neuronal signal transduction. The sarcolemma anchors to the basal lamina through a complex of transmembrane proteins called the dystrophin-associated glycoprotein complex (DGC), which connects the myofiber cytoskeleton to the extracellular matrix (ECM) (Rahimov and Kunkel, 2013). The excitation-contraction coupling that is integral to myofiber function is achieved through the formation of a network of plasma membrane invaginations called T-tubules. In the sarcoplasm, the apposition of the cisternae, a specialized sarcoplasmic reticulum that acts as a calcium store, to the T-tubules, forms the triads (Flucher, 1992; Flucher et al., 1992). The triads, which are established during perinatal development (Flucher et al., 1993), are closely associated to myofibrils and allow for the transduction of the sarcolemmal depolarization upon neural excitation.

**Myofiber Types through Developmental Stages**

Adult myofibers express specific isoforms of MyHC, which are associated with distinct electrophysiological properties and metabolism. Thus, oxidative slow twitch fibers express slow MyHC (type I, Myh7), whereas glycolytic fast twitch fibers express fast MyHC (types IIa (Myh2), IIb (Myh4) and IIx (Myh1)). The expression of MyHC isoforms follows a developmental sequence, with the embryonic and slow MyHC being the first to be expressed (reviewed by Schiaffino and Reggiani, 2011). Fetal and neonatal fibers transiently express perinatal MyHC (Myh8), while the fast isoforms start to be expressed during late fetal myogenesis (Fig. 3B). Fiber type specificity is controlled in part by the transcription factors Six and Eya, which promote fast-type fiber diversity (Grifone et al., 2004; Richard et al., 2011). Fiber type also depends on neural input, which is established during fetal stages—as early as E14.5 in mouse limb (Hurren et al., 2015; Martin, 1990). Thus, the final fiber type of a given muscle is the product of its developmental history, its innervation and, during postnatal life, the physiological demands placed upon it (Schiaffino and Reggiani, 2011).

The origin of satellite cells

Satellite cells, which make up the main population of tissue-specific stem cells found in adult muscle, can be traced back to the Pax7+ myogenic progenitors of the dermomyotome central domain (Gros et al., 2005; Kassar-Duchossoy et al., 2005; Relaix et al., 2005). These Pax7+ embryonic progenitors become the source of the adult Pax7+ satellite cells (Gros et al., 2005; Lepper and Fan, 2010; Seale et al., 2000), although in some specific muscles, such as the diaphragm, prospective satellite cells also maintain Pax3 expression (Relaix et al., 2004, 2005). Notch signaling is essential to maintain the pool of Pax3+/+ progenitors, and in its absence no satellite cells can form (Vasyutina et al., 2007). During the peak of muscle mass growth, which is during the fetal and perinatal stages, the pool of satellite cell progenitors actively divides and represents up to 30% of the mononucleated cells in the muscle mouse tissue (Allbrook et al., 1971; Hellmuth and Allbrook, 1971). However, within 2 months after birth, these progenitors are reduced in number to a small pool of quiescent Pax7+ satellite cells, which account for only a few percent of the mononucleated cells in adult muscle (Allbrook et al., 1971; Cardasis and Cooper, 1975; Schmalbruch and Hellhammer, 1976). In mouse, satellite cells become enclosed under the basal lamina of myofibers during late fetal stages, at about E18.5 (Relaix et al., 2005). Although Pax7+ is not essential for the specification of adult satellite cells per se, these cells are progressively lost in Pax7+ mutant mice (Günther et al., 2013; Oustanina et al., 2004; von Maltzahn et al., 2013). A majority of the adult Pax7+ satellite cells also expresses Myf5+, but these are distinct from proliferating Myf5+ myoblasts which also express MyoD (Beauchamp et al., 2000; Cornelison and Wold, 1997; Zammit et al., 2004) (Fig. 3).

**Skeletal Myogenesis in the Dish: Learning from Development**

Skeletal muscle has been generated in vitro from PSCs using two main approaches: directed differentiation and direct reprogramming, and sometimes a combination thereof. Direct reprogramming is achieved by the overexpression of selected transcription factors to reprogram cells to the myogenic lineage (Table S1). Conversely, directed differentiation approaches have employed several methodologies, each aimed at recapitulating the early differentiation stages that take place in the embryo to generate skeletal muscle (Table S2). Reviews of the field have so far been mostly comparative, focusing on the different methods used to generate skeletal muscles in vitro and their relative success (Abujamra and Valamere, 2015; Baker and Lyons, 1996; Smith et al., 2016). In the following sections, we will discuss these studies in the context of developmental and cell biology and consider the extent to which they recapitulate in vivo myogenesis.

**Directed Differentiation of PSCs into Skeletal Muscle**

The first report of skeletal myogenesis in vitro was provided by Rohwedel et al. (1994), who differentiated mouse ESCs (mESCs) into embryonic bodies (EBs) and observed the expression of MRFs and the formation of myogenin+ skeletal myocytes. EBs are three-dimensional (3D) cell aggregates that differentiate spontaneously into the three germ layers; however, they are usually very heterogeneous and lineage specification cannot be properly controlled (Doetschman et al., 1985; Robbins et al., 1990; Yamada et al., 1994). More recently, adherent monolayer cultures of PSCs have been used to generate more homogenous differentiation. The general principle of directed differentiation approaches is to present specific signaling molecules to the differentiating cells in vitro, as they would experience them in vivo, thus recapitulating normal development.

**Primitive Streak and Mesoderm Induction in Vitro**

The earliest stage of muscle differentiation in vivo occurs when paraxial mesoderm precursors in the epiblast activate the paraxial mesoderm induction axis (Huchard et al., 2014; Shigekawa et al., 2015). This activation results in the expression of proneural and mesodermal markers, such as Pax3 and Mixl1, which are essential for the specification of myogenic progenitors. In vitro, these markers are induced by canonical Wnt signaling, which activates the transcription of myogenic genes (Zhu et al., 2016). However, the differentiation of PSCs into skeletal muscle remains a challenging task, and further research is needed to improve the efficiency and homogeneity of these approaches.
mesodermal marker T in the primitive streak. Several studies aimed at recapitulating these early stages of mesoderm induction and specification using PSCs have led to the identification of the BMP, activin/Nodal, FGF, insulin (PI3K) and Wnt pathways as major signals able to induce the specification of various T⁺ mesodermal subpopulations (reviewed by Murrey and Keller, 2008). Initial studies focused mostly on the differentiation of posterior primitive streak derivatives such as lateral plate and extraembryonic mesoderm, which produce hematopoietic and cardiogenic progenitors (Kaufman et al., 2001; Mummery et al., 2003). These studies identified BMP signaling as a key requirement for hematopoietic progenitor specification and they established the role of Wnt signaling for mesoderm induction in vitro (Chadwick et al., 2003; Gadue et al., 2006; Johansson and Wiles, 1995; Kattman et al., 2011; Lindley et al., 2006; Nakashishi et al., 2009; Yang et al., 2008; Zhang et al., 2008). More recent studies have shown that treatment of PSCs with activin favors the differentiation of anterior primitive streak derivatives such as endoderm, whereas treatment with BMP triggers the differentiation of posterior streak derivatives such as lateral plate and extraembryonic mesoderm (Bernardo et al., 2011; Mendjan et al., 2014). Differentiation of late primitive streak that generates paraxial mesoderm could be achieved by treating PSCs first with activin, FGF and PI3K inhibitors to induce anterior primitive streak fate and then with GSK3 inhibitors (to activate Wnt signaling) and FGF to recapitate the late primitive streak environment (Mendjan et al., 2014). Alternatively, by treating both mouse and human PSCs with BMP and Wnt activators sequentially, several groups were able to generate paraxial mesoderm progenitors (Chal et al., 2015; Mendjan et al., 2014; Sakurai et al., 2009, 2012). Recent studies also highlighted the importance of cell-cell and cell-ECM interactions in controlling early stages of mesoderm differentiation: human PSCs (hPSCs) exposed to identical media will adopt distinct cell fate depending on their precise in vitro organization and volumetric cell density (Kempf et al., 2016; Warmflash et al., 2014). Early stages of gastrulation have also been recapitulated in vitro in 3D mouse PSC cultures, termed gastruloids, where Tbx6 expression is observed (ten Berge et al., 2008; van den Brink et al., 2014).

Induction of a presomitic mesoderm fate in vitro

Production of mouse or human paraxial mesoderm progenitors characterized by expression of the presomitic mesoderm markers Msgn1 or Tbx6 can be achieved by treating PSC cultures with Wnt activators alone or in combination with other factors, such as FGF, without prior primitive streak induction (Borchin et al., 2013; Gouti et al., 2014; Mendjan et al., 2014; Shelton et al., 2014). Xu et al. (2013) also added forskolin to the Wnt/FGF combination, whereas others used only FGF [in a 3D system (Hosoyama et al., 2014)], only Wnt (Hwang et al., 2014), or Wnt followed by a Notch inhibition step (Choi et al., 2016). In most instances, these protocols require purification of the paraxial mesodermal progenitors by flow cytometry using cell surface markers or lineage-specific fluorescent reporters (for more details, see Box 1). In the mouse tail bud, Wnt activates expression of Fgfg8 (Aulehla et al., 2003), and thus treating cultures with a Wnt pathway activator is also likely to result in FGF activation. Accordingly, in mouse FGF treatment is dispensable for paraxial mesoderm induction from mESCs (Chal et al., 2015). Wnt is also required for the induction of the neuromesodermal (T⁺/Sox2⁺) progenitors from mouse and human PSCs in vitro (reviewed by Henrique et al., 2015).

The generation of cells with a posterior presomitic mesoderm fate from PSCs in vitro was analyzed using mESC reporter lines for Msgn1 expression (Chal et al., 2015; Sudheer et al., 2016) (Box 1). Optimal induction of posterior presomitic mesoderm fate was observed when mESCs were pretreated with BMP4 to differentiate them into epiblast-like cells, followed by treatment combining a Wnt activator (R-spondin 3 or the GSK3β inhibitor CHIR99021) and a BMP inhibitor [LDN-193189 (Cuny et al., 2008)], which prevented the induced presomitic mesoderm cells from drifting to a lateral plate mesoderm fate (Chal et al., 2015). In the absence of BMP inhibition, Msgn1⁺ cells started to activate Bmp4 and to upregulate lateral plate markers in vitro (Chal et al., 2017a preprint).

Paraxial mesoderm and lateral plate mesoderm share common precursors in the primitive streak, and grafts of the anterior streak (territory fated to give rise to the paraxial mesoderm) into the mid-streak level (territory fated to give rise to the lateral plate) is enough to change the fate of the cells according to their new position (Garcia-Martinez and Schoenwolf, 1992). These data indicate that the fate of the precursors is not determined at the primitive streak stage. Furthermore, exposure of paraxial mesoderm precursors to BMP4 in vivo leads them to switch to a lateral plate mesoderm fate, indicating that BMP4 acts as a lateral plate inducer (Tonegawa et al., 1997). BMP signaling also has a well-established role in promoting the formation of lateral plate mesoderm derivatives such as hematopoietic and cardiovascular cell types both in vivo and in vitro (Adelman et al., 2002; Tonegawa et al., 1997; reviewed by Murrey and Keller, 2008; Orlova et al., 2015). Sakurai et al. (2012) also reported that high levels of BMP4 inhibit presomitic mesoderm and myogenic differentiation from PSCs.

Posterior presomitic mesoderm identity of mESCs induced with Wnt activation and BMP inhibition was confirmed using microarrays comparing their transcriptome with that of microdissected posterior presomitic mesoderm from E9.5 mouse embryos (Chal et al., 2015). Remarkably, these posterior presomitic

---

**Box 1. Tools for characterizing PSC-derived myogenic progenitors**

**Surface markers.** A number of surface markers have been used to purify myogenic precursors from PSCs induced to differentiate to the muscle lineage by various protocols. Among them, PDGFRα (CD140a) and VEGFR2 (Flk1, KDR, CD309) have been used to enrich for various subtypes of early mesoderm produced in vitro (Nishikawa et al., 1998). Notably, PDGFRα⁺/VEGFR2⁻ cells isolated from differentiating PSC cultures have been proposed to correspond to putative paraxial mesoderm progenitors, including premyogenic cells (Darabi et al., 2012, 2008; Sakurai et al., 2009, 2012). Myogenic progenitors have also been isolated from mouse or human PSC cultures using CD106 (VCAM1, SM/C2.6) (Chang et al., 2009; Mizuno et al., 2010), CD34 (Demestre et al., 2015), CD56 (NCAM1) (Young et al., 2016), CD56 and CD73 (Barberi et al., 2007), CD82 (Uezumi et al., 2016), CXCR4 (CD184) and c-Met (Borchin et al., 2013). However, the usefulness of these markers is limited by their overall poor lineage specificity. Notably, PDGFRα and VEGFR2 are dynamically expressed on several mesodermal populations including paraxial mesoderm and lateral plate mesoderm and thus are not lineage-specific markers (Ding et al., 2013; Ema et al., 2006; Motoke et al., 2003; Takebe et al., 2006).

**Fluorescent reporters.** Lineage-specific promoters have also been used to drive the expression of a fluorescent protein (chiefly GFP) to generate reporter cell lines. One of the first examples of this was the brachyury-GFP mESC reporter line, which was used to analyze mesoderm induction in vitro (Fehling et al., 2003; Gadue et al., 2006). More recently, ESC lines that harbor fluorescent reporters for the expression of key paraxial mesoderm/myogenic lineage markers, namely Msgn1, Pax3, myogenin (Myog) and Pax7, have been used to track the sequential differentiation of PSCs toward paraxial mesoderm and skeletal muscle (Chal et al., 2015; Sudheer et al., 2016).
mesoderm cells induced in vitro also expressed cyclic genes such as \( \text{Hes7} \), suggesting that the segmentation machinery is activated in vitro (Chal et al., 2015). Efficient differentiation to a posterior presomitic mesoderm fate was also achieved by exposing hPSCs to a medium that included the small molecules CHIR99021 and LDN-193189, followed by FGF addition to the medium after 3 days of differentiation (Chal et al., 2016, 2015). In these conditions, cells acquired a posterior presomitic mesoderm fate with over 90% of cells positive for Tbx6 (Chal et al., 2016). Faster presomitic mesoderm induction was also obtained by pretreating cells with activators of the Wnt, FGF and TGF pathways combined with PI3K inhibition for 1 day prior to treatment with Wnt activator and BMP inhibitor (Loh et al., 2016).

In mouse embryos, Msgn1 expression is downregulated at the determination front where embryonic segments are first specified. This marks the level where the first stripes of segmentation gene expression, such as Mesp2 or Ripply1/2, are observed and where Pax3 is first activated. Differentiation of mouse ESCs to anterior presomitic mesoderm was studied using a Pax3 reporter line, demonstrating that expression of segmentation genes such as Ripply2 or those involved in somite rostrocaudal patterning such as \( \text{Ucnx} \) or \( \text{Tbx18} \) can be activated by maintaining Wnt activation and BMP inhibition in cultured cells (Chal et al., 2015). Similar activation of the segmentation program was also reported in hPSCs using Wnt and FGF inhibition (Loh et al., 2016).

Generation of skeletal myoblasts

The proliferating myoblasts that fuse to form adult myofibers in vivo are derived from Pax3+ precursors found in the dermomyotome. Induction of such Pax3+ precursors has been achieved by maintaining mESC cultures in a medium containing a Wnt activator and a BMP inhibitor (Chal et al., 2015). The RA biosynthetic enzyme Aldh1a2 (Raldh2) is expressed in mouse Pax3+ cells induced in vitro (Chal et al., 2015), suggesting that the cultures are able to produce RA in sufficient amounts to promote presomitic mesoderm-like cell maturation without the need for external supply (Kennedy et al., 2009; Ryan et al., 2012). Subsequent treatment of the cultures with HGF, IGF and FGF led to efficient induction of myogenin+ myocytes both from mouse and human PSCs after 1 to 2 weeks of differentiation (Chal et al., 2015, 2016) (Fig. 4). In mESC cultures, these myocytes appear as elongated mononucleated cells expressing slow MyHC and primary myotome cells. The rapid increase in myogenin+ myocytes was accompanied by a concomitant decrease of Pax3+ myogenic progenitors. These were progressively replaced by a pool of Pax7+ myogenic progenitors, as observed in vivo (Chal et al., 2015; Kassar-Duchossoy et al., 2005; Relaix et al., 2005). Several studies have aimed at deciphering the respective contribution of individual transcription factors including Pax3/7 and the MRFs during myogenic differentiation of mESCs using engineered mutations or siRNA-based knockdown (Table S2). Other studies with hPSCs have reported different induction kinetics of the Pax3/7+ progenitors. Awaya et al. (2012) described the late differentiation (50 days) of Pax3+ and Pax7+ cells that can be isolated with CD56 (NCAM1) and CD73 (NT5E) surface markers. Myosphere culture of differentiating hPSCs has also been described, which produces Pax7 expression after 6 weeks and myotube formation after 8 weeks of culture (Hosoyama et al., 2014). Recently, Caron et al. (2016) described a serum-based protocol to differentiate human myoblasts in vitro.

![Fig. 4. Skeletal myogenesis from pluripotent stem cells.](image-url)

Comparison of strategies to generate skeletal muscles from PSCs. The sequence of the differentiation stages of mESCs and hPSCs into skeletal muscle and their corresponding markers is shown from left to right (top). A comparison of the approaches to generating skeletal muscle from PSCs is shown in the middle. Directed differentiation approaches aim to recapitulate the developmental stages of paraxial mesoderm specification and differentiation by manipulating signaling pathways such as Wnt and BMP. The sequence of developmental stages and important signaling pathways associated with each stage are shown. Direct reprogramming approaches aim to bypass early developmental stages by overexpressing a myogenic regulator, chiefly Pax3/7 or MyoD (PAX3/7, iMYOD). In some instances, the cellular events occurring during differentiation remain largely unclear (dashed line). Directed differentiation into skeletal muscle occurs according to slightly different timelines [days (d) of \textit{in vitro} differentiation are indicated] using mESCs as compared with hPSCs (bottom).
within 26 days; although some cells in these cultures express Pax3, no Pax7+ cells were generated.

**Direct reprogramming strategies using transcription factors**

The discovery that treatment of fibroblast cultures with 5-azacytidine, a demethylating agent, could induce cells to differentiate into myoblasts (Constantinides et al., 1977; Taylor and Jones, 1979) led to the demonstration that MyoD, whose locus is a target of 5-azacytidine, could reprogram fibroblasts to a muscle fate (Davis et al., 1987; Weintraub et al., 1989). This in turn led to direct reprogramming strategies whereby overexpression of a myogenic factor – often MyoD – in PSCs forces their differentiation toward the skeletal muscle lineage, bypassing de facto early developmental stages (Dekel et al., 1992; Shani et al., 1992; reviewed by Comai and Tajbakhsh, 2014) (Fig. 4, Table S1). MyoD overexpression is often triggered after EB formation, and MyoD expression directly in PSCs is inefficient unless BAF60c (SMARCD3) is co-expressed (Albini et al., 2013). MyoD-reprogrammed iPSCs can generate myotubes within 10 days (Rao et al., 2012) and are also able to engraft when transplanted in mouse muscle (Goudenege et al., 2012; Otsara et al., 2007; Zheng et al., 2006) (Table S1). Moreover, MyoD-reprogrammed myotubes have been used to model muscle diseases such as Duchenne or Miyoshi muscular dystrophies (Shoji et al., 2007; Yasuno et al., 2014; Young et al., 2016) (Table S1).

In the chicken embryo, Pax3 overexpression can lead to ectopic activation of the myogenic program in lateral plate or neural tube cells (Maroto et al., 1997). Using both mouse and human PSCs, Perlingeiro and colleagues developed direct reprogramming methods using Pax3/7 overexpression combined with EB formation to induce myogenic differentiation (Darabi et al., 2012, 2008, 2011; Filareto et al., 2013) (Fig. 4). As Pax3/7 are also expressed in neural derivatives in vivo (Fougerousse et al., 2002; Gerard et al., 1995; Goulding et al., 1991; Jostes et al., 1990), a purification step using surface markers such as PDGFRα (CD140a) or CD56 was needed to enrich in myogenic cells (Darabi et al., 2008, 2011; Quattrocelli et al., 2011) (Box 1). Pax3 overexpression acts on the early PDGFrα+ population, upregulating transcription factors Tcf15 and Mef1 and preventing the cells from undergoing cardiomyogenesis (Magli et al., 2013, 2014). Both Pax3- and Pax7-reprogrammed mouse progenitors were shown to engraft when transplanted in a mouse model. Pax3- or Pax7-reprogrammed hPSCs can generate myogenic progenitors that are CD29 (ITGα1)+/CD44+/CD56+ and CXCR4+/CD106 (VCAM1)+ and which can engraft into mouse muscle (Darabi et al., 2012; Filareto et al., 2015) (Box 1).

Overexpression of Meq1 has also been used to produce putative cardio-pharyngeal progenitors with the capacity to differentiate into both cardiac and cranial paraxial mesoderm (Chan et al., 2013; Lescroart et al., 2014). Other factors used for direct reprogramming of PSCs into myogenic cells include other MRFs and a mutated version of HMGA2, as well as the expression of growth factors such as IGF2 or a modified HGF (Table S1). Despite some success, direct reprogramming methods are nevertheless often undefined, as they rely on EB formation in serum-containing media and often require further cell sorting (Table S1, Box 1).

**Myofiber maturation in vitro**

Sarcomeres and structural assembly

Maturation of PSC-derived cell types has proven a major challenge for most cell lineages, with differentiated cells often retaining a fetal-like phenotype (Robertson et al., 2013). Mature striated

---

Fig. 5. Cellular features of PSC-derived skeletal myofibers. Schematic of a myofiber (center) and the corresponding representative features detectable by immunohistochemistry on skeletal myofibers differentiated from mESCs. In the central scheme, the elongated myofiber contains aligned striated myofibril bundles composed of various MyHCs (yellows), dystrophin in a subsarcolemmal position (thin irregular black line), aligned mitochondria (green), myonuclei (gray) and nascent T-tubule network (light blue lines). The fiber extremities form a specialized contact surface reminiscent of a myotendinous junction (MTJ), while at the equatorial level the sarcolemma (solid black line) can harbor a postsynaptic element (brown) indicating post-neuromuscular junction (NMJ) assembly. Satellite cells (purple) can be located under the basal lamina (black dotted line). Images show the various markers (green) and the location of the features illustrated in the central scheme. In some cases, nuclei are counterstained with DAPI (blue). In the NMJ/post-synapse image, fibers are counterstained for MyHC (red). In the satellite-like cells panel, fibers are visible in a merged phase-contrast image. Scale bars: 100 μm.
myotubes have been generated in vitro from mESCs spontaneously differentiated from EBs (Hirsch et al., 1998; Rohwedel et al., 1998b), as well as from Pax7-reprogrammed hPSCs (Skoglund et al., 2014). Demestre et al. (2015) compared these two methods in their ability to produce human actinin+ myotubes and reported that myotubes generated by Pax7-mediated reprogramming were slower to mature. Millimeter-long myofibers exhibiting MyHC+ and titin+ striations and containing large numbers of nuclei have also been generated from mouse and human PSCs in adherent, serum-free conditions (Chal et al., 2015) (Fig. 5). The maturation of the myofibers was evidenced by the transition from slow to peripheral MyHC expression, the formation of highly organized myofibrils, the expression of Nfix, which is a secondary myogenesis marker (Messina et al., 2010), and the assembly of individual basal lamina (Fig. 5). Based on their morphological characteristics, the authors suggested that these myofibers resemble perinatal fibers (Chal et al., 2015; White et al., 2010). Remarkably, the formation of such long striated myofibers is rarely observed when primary myoblasts, myogenic lines (such as C2C12) or satellite cells are differentiated in vitro, as they tend to form poorly organized multinucleated myotubes.

Electron microscopy analysis of mouse and human PSC-derived myofibers revealed that they exhibit highly organized sarcomeric units, reminiscent of the cytoarchitecture found in vivo (Hirsch et al., 1998; Skoglund et al., 2014). Myofibrils generated in these fibers can exhibit the appropriate organization of MyHC, troponin T, nebulin and titin, and undergo the stage-specific developmental transitions through α-actins (cardiac to skeletal) and MyHC (embryonic, fetal, and perinatal/fast) isoforms (Chal et al., 2015; Hirsch et al., 1998; Mizuno et al., 2009; Rohwedel et al., 1998b). Furthermore, desmin null mESCs have been used to demonstrate the important role of desmin in the early formation of myotubes (Höllerig et al., 2002; Weitzer et al., 1995). While these studies are encouraging, ultrastructural analyses have shown that hPSC-derived myofibers have myofibrils that are less regularly organized than those found in vivo, featuring irregular and misaligned Z-disks. They also fail to form a mature T-tubule network, a hallmark of adult myofibers (Skoglund et al., 2014). Although triad-like structures could be identified, they were limited to the subsarcolemmal space, lacking the orthogonal alignment to myofibrils observed in vivo. This incomplete maturation could be explained by the immature character of the fibers and the lack of innervation in these cultures.

The formation of a basal lamina (endomysium) has been shown to occur in vitro, as evidenced by the deposition of a continuous sheath of laminin along individual fibers (Chal et al., 2015). The dystrophin-associated glycoprotein complex (DGC), which links the myofiber cytoskeleton to the basal lamina through the sarcolemma, is also assembled in vitro. DGC components, including dystrophin (Chal et al., 2015; Ozasa et al., 2007), dystroglycan and integrins (Hirsch et al., 1998; Jacobson et al., 2001; Tremblay and Carbonetto, 2006), were found distributed at the sarcolemma of PSC-derived myofibers (Fig. 5, Tables S1 and S2). During development, integrins α4 and α7 play a key role in skeletal myogenesis and associate with integrin β1 to form receptors for ECM proteins such as laminin (Menko and Boettiger, 1987; Rosen et al., 1992). Intriguingly, differentiation of HgbI null mESCs showed that mutant myoblasts can still fuse and form myotubes with apparently normal sarcomeres and costameres (Hirsch et al., 1998; Rohwedel et al., 1998a). Moreover, myofibers differentiated from DagI null mESCs exhibit defects in the assembly of postsynaptic elements including acetylcholine receptor (AchR) clusters (Jacobson et al., 2001). Dmdmdx mPSC lines were shown to exhibit defects in myogenic differentiation and increased apoptosis (Chen et al., 2012), while Chal et al. (2015) reported a branching defect characterized by an increased number of split fibers, a phenotype also seen in Duchenne muscular dystrophy (DMD) patients and the mdx mouse (Chan and Head, 2011). Finally, myoblasts derived from DMD human iPSCs were shown to exhibit fusion defects that could be corrected by dual inhibition of TGFβ and BMP signalings (Choi et al., 2016).

Excitation-contraction coupling

During development, motoneuron input is important for proper myofiber maturation and viability. However, in vitro, cultures of PSC-derived muscle that do not receive motoneuron input nevertheless achieve a relatively high level of functional maturation, as evidenced by myofibrils and spontaneous contractions (Caron et al., 2016; Chal et al., 2015; Choi et al., 2016; Shelton et al., 2014). PSC-derived myofibers have been further characterized by electrophysiological studies focusing on calcium handling, ions currents and AchR cluster properties. In vivo, early embryonic myofibers exhibit transient, low voltage-activated calcium currents (T-type, ICaL), which are progressively replaced during fetal stages by slow, high voltage-activated currents (L-type, ICaL) (Beam et al., 1986; Cognard et al., 1986). Myofibers derived from Pax7-reprogrammed hPSCs exhibit excitable properties (Skoglund et al., 2014) but have immature ion handling characterized by slow calcium buffering and a small potassium conductance, which could explain the spontaneous fiber twitches observed in these cultures (Skoglund et al., 2014).

In vivo, as myotubes mature, they upregulate AchR expression and switch from the fetal γ-subunit to the adult ε-subunit (Buonanno and Merlie, 1986; Mishina et al., 1986). Patch clamp and gene expression analyses of mESC-derived myotubes have shown that this developmental transition can also occur in vitro, as these cells start to express the adult AchR after 2 weeks of differentiation (Guaj et al., 1999; Jacobson et al., 2001; Rohwedel et al., 1998b, 1994; Tremblay and Carbonetto, 2006) (Fig. 5B). However, it has been shown that postsynaptic elements can form in vitro on myotubes in the absence of neural input, and that these cells remain immature (reviewed by Sanes and Lichtman, 1999). For this reason, there have been several efforts to establish a co-culture system of PSC-derived myogenic cells with rat embryonic spine or with PSC-derived motoneurons. This system improved the maturation of PSC-derived myofibers, as evidenced by AchR clustering and electrophysiological recordings (Demestre et al., 2015; Puttonen et al., 2015; Rohwedel et al., 1998b); however, electrophysiological analysis suggested immaturity of neuromuscular junction activity (Puttonen et al., 2015; Vyskočil and Vrbová, 1993).

Generation of satellite-like cells in vitro

In adult muscle, the pool of quiescent Pax7+ satellite cells is closely associated with the myofibers (reviewed by Brack and Rando, 2012; Dumont et al., 2015), an arrangement that aids muscle repair during regeneration (Lepper et al., 2011; Murphy et al., 2011; Sambasivan et al., 2011). However, satellite cells are found in very limited numbers and they cannot be amplified in vitro as they differentiate and lose their regenerative properties in culture (Gilbert et al., 2010; Montarras et al., 2005). The generation of Pax7+ cells from PSCs has been reported in several of the protocols discussed previously but in most cases their characterization has remained preliminary (Tables S1 and S2). Notably, the association of Pax7+ cells with mature myofibers and the location of the cells under the
basal lamina, as it is seen in vivo, has only been reported by Chal et al. (2015) (Fig. 5). One of the key characteristics of the Pax7+ satellite cells is their capacity to regenerate myofibers while being able to self-renew and reconstitute the satellite stem cell compartment when grafted in adult injured muscles. Several of the myogenic differentiation protocols discussed above have reported the production of myofibers when the progenitors obtained in vitro were transplanted into injured adult muscle (Tables S1 and S2). In some cases these produced donor-derived Pax7+ cells resembling satellite cells; however, in all cases the transplanted cells were either the result of 'bulk' differentiation or were isolated using surface markers resulting in a mixed population of cells containing a fraction of putative Pax7+ cells. The in vivo potential of pure populations of PSC-derived Pax7+ cells produced in vitro was evaluated using a Pax7-fluorescent reporter mESC line (Fig. 5). In differentiated cultures of this line, Pax7+ cells first experienced a proliferative phase between 2 and 3 weeks of differentiation, followed by a reduction in the number of Pax7+ progenitors, with the remaining cells found closely associated to mature fibers, as expected for satellite cells (Chal et al., 2015). Similar dynamics could be observed in differentiated hPSCs, where Pax7+ cells could also be found closely associated to myofibers. Furthermore, a significant proportion of the Pax7+ cells differentiated in vitro were found to be Ki67+, suggesting that they are quiescent (Chal et al., 2015). When activated, satellite cells are able to divide asymmetrically, thus allowing for self-renewal and for the generation of a myoblast that will proliferate and contribute to muscle regeneration. Pax7+ cells generated in vitro from mESCs can give rise to both myofibers and Pax7+ satellite cells when grafted in vivo in an mdx mouse model, suggesting that these cells can behave like bona fide satellite cells (Chal et al., 2015). The myofibers differentiated in vitro from mESCs exhibit characteristics of perinatal fibers, suggesting that the associated Pax7+ cells are likely to correspond to perinatal satellite cells (Tierney et al., 2016).

Biological relevance and open questions
The potential for PSCs to serve as developmental models, including for skeletal muscle development, has been recognized for more than 20 years (Baker and Lyons, 1996; Keller, 1995; Rohwedel et al., 2001; Wobus et al., 2001). However, these systems have been criticized for their artificial nature and their simplicity compared with the full complexity of organ systems (Moretti et al., 2013). Moreover, it is still unclear if in vitro differentiation encounters the same developmental constraints as in vivo. A similar concern has been raised about the physiological relevance and the proper maturation of the cell types generated in vitro, which in most cases reach only a fetal state (Robertson et al., 2013; Satin et al., 2008). For skeletal muscle, this is evidenced by the neonatal properties of the in vitro myofibers, including smaller size, immature metabolism and a weak physiological response (Bursac et al., 2015). Parameters that control the maturation of PSC-derived myofibers remain largely unexplored and the use of integrated co-culture systems or biomaterials has been seldom investigated (Chal et al., 2016; Leung et al., 2013). Nevertheless, the recent availability of more efficient myogenic differentiation protocols for PSCs will allow the study of skeletal myogenesis with unprecedented accessibility and resolution. A number of questions remain to be answered about the nature of the muscle cell generated in vitro, in particular about their axial/anatomical identities and fiber composition, and how these relate to what is seen in vivo (Donoghue et al., 1992; Jarad and Miner, 2009; Kiency et al., 1972; Nowicki and Burke, 2000; Rosenthal et al., 1989).

A number of translational applications for PSC-derived skeletal muscle cells can be envisioned. Although satellite cells represent the ideal target population for muscle regenerative medicine (Collins et al., 2005), their accessibility, number and manipulability remain very limited. Furthermore, cell therapy approaches still face significant challenges, including transplantation routes, scale up, host immune response and functional integration (Tremblay and Skuk, 2008). If these limitations can be overcome, then PSC-derived myogenic cultures that recreate the satellite cell niche in vitro may be a valuable source of cells for transplantation (Cosgrove et al., 2009). In the near future, hPSC-based myogenic cultures are likely to become an important preclinical model with which to study disease mechanisms in place of primary cultures generated from patient biopsies, which are subject to variability and limited accessibility. Combined with the relative ease in deriving novel hPSCs and progress in genome engineering, disease-specific models can now be generated for a large spectrum of muscle pathologies (Avior et al., 2016; Merkle and Eggan, 2013; Rüegg and Glass, 2011), facilitating the development of in vitro therapeutic approaches (Kazuki et al., 2010; Rohwedel et al., 2001; Tran et al., 2013; van Deutkom et al., 2007).

Concluding remarks
Decades of developmental studies have provided a compelling picture of skeletal muscle formation in vertebrate embryos; however, important myogenic processes remain poorly studied due to limited in vivo accessibility. PSC-based systems herald a new area of ‘development-in-a-dish’, with the disciplines of developmental and stem cell biology now merging closer than ever. Directed differentiation methods can recapitulate many aspects of bona fide skeletal myogenesis, allowing us to probe key developmental mechanisms in an unprecedented manner. One of the most exciting opportunities that PSC-based systems provide is to understand the regulation of skeletal myogenesis as it occurs during human development. In turn, this will help to provide more relevant tissue models for diseases and therapeutic investigations. Several key challenges remain on this front, the first of which is the generation of physiologically relevant myofibers. This is likely to be solved by rational design in combining the factors that control skeletal muscle maturation with advances in bioengineering. The second key challenge is the recreation of the satellite cell niche. While there is evidence that coexistence of myofibers and Pax7+ progenitors allows for the occurrence of Pax7+ satellite-like cells in vitro (Chal et al., 2015), a precise understanding of the niche environment and the molecular mechanisms at play remains to be achieved. A sustained effort to better characterize the regulation of skeletal myogenesis during development and in the adult is therefore crucial for success in the clinic. Although PSC-derived skeletal myogenesis is still a nascent field, this is a time of exciting opportunities for the developmental biology, drug discovery and regenerative medicine communities.

Acknowledgements
We are grateful to members of the O.P. laboratory for helpful discussions.

Competing interests
O.P. and J.C. are co-founders and share-holders of Anagenesis Biotechnologies, a start-up company specialized in the production of muscle cells in vitro for cell therapy and drug screening.

Funding
This work was supported by advanced grant ERC-2009-AdG 249831 from the European Research Council to O.P., by the EU Seventh Framework Programme grant Plurimes (agreement no. 602423), by a strategic grant from the French
Muscular Dystrophy Association (AFM-Téléthon) to O.P., and by the INSTRACT project (Agence Nationale de la Recherche).

Supplementary information
Supplementary information available online at http://dev.biologists.orglookup/toi.1242/dev.151035.supplemental

References


Dysfunction enables the genetic correction of dystrophic satellite cells. J. Biol. Chem. 283, 11188-11197.


Weitzer, G., Milner, D. J., Kim, J. U., Bradley, A. and Capetanaki, Y. van der Velden, J. L. J., Langen, R. C., Kelders, M. C., Wouters, E. F., Janssen-
2122
van den Brink, S. C., Baillie-Johnson, P., Balayo, T., Hadjantonakis, A.-K., 2122
Yamaguchi, A., Yamada, H., Kasai, T., Masuda, S., Narita, A. et al. (2016). Cell-
van der Velden, J. L. J., Langen, R. C., Kelders, M. C., Wouters, E. F., Janssen-
von Deutekom, J. C., Janson, A. A., Ginjaar, I. B., Frankhuizen, W. S., Aartsma-
Van Horn, R. and Crow, M. T. van Deutekom, J. C., Janson, A. A., Ginjaar, I. B., Frankhuizen, W. S., Aartsma-