5.3: Discussion

5.3.1: The enhancers controlling *lama1* expression in the zebrafish embryo are located within intron 1 of the gene

Data presented in this chapter demonstrate that intron 1 of *lama1*, in combination with 3291 bases upstream of the transcriptional start site (Δ3.3kb*lama1*GFPΔex2), contains enhancers capable of driving GFP reporter gene expression in a pattern that is identical to the *lama1* BAC transgene (FL), and recapitulates fully *lama1* expression pattern.

Overall, the deletion analyses I performed established that *lama1* transcription requires the presence of several regulatory elements acting in a concerted manner to control initiation and maintenance of *lama1* expression in the developing zebrafish embryo. For instance, upon removal of all DNA sequences downstream of exon 2 (Δ3.3kb*lama1*GFPΔex2), an increased number of embryos express GFP in the muscle fibres, eye, neural tube, and the vasculature/pro-nephric tubule region in comparison to *lama1* BAC (FL)-injected embryos (*Table 5.9*). This suggests that regulatory elements downstream of exon 2 may be required to repress the expression of *lama1* in these tissues (*Figure 5.32*).

I have also shown that the first 4415 bases of intron 1, in combination with 878 bases upstream of the transcriptional start site that is likely to contain a minimal promoter (Δmin.prom*lama1*GFPΔ5k), contains enhancers capable of driving GFP reporter gene expression in the muscle fibres, notochord and PSM, but not in the anterior CNS, eye, otic vesicles, jaw musculature, neural tube, heart, pectoral fins, vasculature/pro-nepric tubules, and the uro-genital region. Consistent with the idea that repressive regulatory elements are also located downstream of exon 2, I have observed that removal of all DNA sequences downstream of base 4416 within intron 1 of *lama1*, in combination with the removal of the sequence that is -3291 to -878 bases upstream of the *lama1* transcription start site (Δmin.prom*lama1*GFPΔ5k), also results in an increased number of embryos which express GFP in the muscle fibres at every stage analysed, in comparison to *lama1* BAC (FL)-injected embryos (*Table 5.9*). Unlike Δ3.3kb*lama1*GFPΔex2-injected embryos, an increased expression in the eye and neural tube is not observed in Δmin.prom*lama1*GFPΔ5k-injected embryos. This is likely because the removal of bases 4416 to 9779 within intron 1, together with the region -3291 to -878bp, causes a specific loss of reporter gene expression in these tissues. Therefore, the enhancers controlling *lama1* expression in these tissues are located within 5363bp immediately upstream of exon 2 (*Figure 5.32*). Expression of GFP in the vasculature and uro-genital region may also be reduced in embryos injected with the construct Δmin.prom*lama1*GFPΔ5k (*Table 5.9*), although I cannot rule out that the reduction observed could be a consequence of experimental variation, because these tissues are very small in size in comparison to others such as the eyes. For example, no GFP expression was observed in the vasculature/pro-nephric tubules at 49hpf in Δmin.prom*lama1*GFPΔ5k-injected embryos.

However, GFP-expressing cells were detected in this tissue in only 1% of *lama1* BAC (FL)-injected embryos. At earlier stages of development, 3% of Δmin.prom*lama1*GFPΔ5k-injected embryos contained GFP-positive cells in this tissue, in comparison to 12% of *lama1* BAC (FL)-injected embryos (*Table 5.9*). Taken together, results indicate that enhancers controlling the expression of *lama1* in the vasculature/pro-nephric tubules and the uro-genital region are likely to be lost upon removal of bases 4416 to 9779 within intron 1.

Whilst the Δmin.prom*lama1*GFPΔ5k construct is sufficient to initiate and maintain expression of GFP in the muscle fibres, bases 1 to 4415 of intron 1 appear to drive the initial transcription of the GFP reporter gene, but fail to maintain it in the notochord and the PSM/tail. At 49hpf, 20% of *lama1* BAC (FL)-injected embryos display GFP expression in the tail, whilst no embryo injected with Δmin.prom*lama1*GFPΔ5k or Δ3.3kb*lama1*GFPΔex2 express GFP in the tail (*Table 5.9*). This indicates that enhancers which maintain *lama1* expression in the PSM/tail are likely to be located downstream of exon 2 (*Figure 5.32*). In contrast, there is a significant down-regulation in the number of embryos expressing GFP in the notochord in Δmin.prom*lama1*GFPΔ5k-injected embryos at 25hpf and 49hpf, yet expression in Δ3.3kb*lama1*GFPΔex2-injected embryos at 49hpf remains comparable to that of *lama1* BAC (FL)-injected embryos (*Table 5.9*). These results suggest that an enhancer which maintains expression of *lama1* in the notochord is located between bases 4416 to 9779 within intron 1, possibly in or around the conserved region of DNA at position 8859 bases, whilst enhancers controlling the initiation of *lama1* in both the PSM and notochord are located in the first 4415 bases of intron 1 (*Figure 5.32*).

Construct	*Lama1* BAC (FL)			Δmin.prom*Lama1*GFPΔ 5kb			Δ3.3kb*Lama1* GFPΔex2
Tissue	19-20s	25hpf	49hpf	19-20s	25hpf	49hpf	49hpf
Muscle	60	76	50	**85**	**86**	**62**	**63**
Eye	17	44	21	**1**	**2**	**1**	**49**
Ant. CNS	47	67	47	**13**	**0**	**0**	46
Notochord	37	30	2	23	**2**	**0**	3
PSM/tail	30	21	20	14	**0**	**0**	**0**
Va/Pn	7	12	1	**0**	**3**	**0**	**14**
Uro-genital	3	11	3	**0**	**1**	**1**	0
Neural tube	3	12	7	**0**	**0**	**0**	**14**
Heart	0	0	12	0	0	**1**	11
Pectoral fin	na	na	5	na	na	**0**	6

Table 5.9: *A comparison of the percentage of embryos with tissue-specific GFP expression at the 19-20-somite stage, 25hpf, or 49hpf, that have been injected with three different BAC DNA constructs*. The first 4415 bases of intron 1, in combination with 878 bases upstream of the transcriptional start site, are sufficient for *lama1* expression in the muscle fibres, notochord, and the PSM. Green and red numbers represent an increase or decrease, respectively, in the number of embryos with tissue-specific GFP expression relative to *lama1* BAC (FL)-injected embryos. na = not applicable.

These findings strongly suggest that the enhancers that initiate and maintain *lama1* transcription are spread throughout intron 1 of the gene, and likely downstream of exon 2. Enhancer elements may also be located in the region -3291 bases upstream of the *lama1* translational start site (*Figure 5.32*), which is sufficient for very weak activation of GFP in the muscle fibres, notochord, eye, and the anterior CNS (sections 5.2.5, 5.2.9). These enhancers could also be involved in the maintenance of *lama1* expression in the PSM, but do not function very efficiently without intron 1. A complex combinatorial mode of regulation between these different regulatory elements is likely to control tissue-specific *lama1* expression.

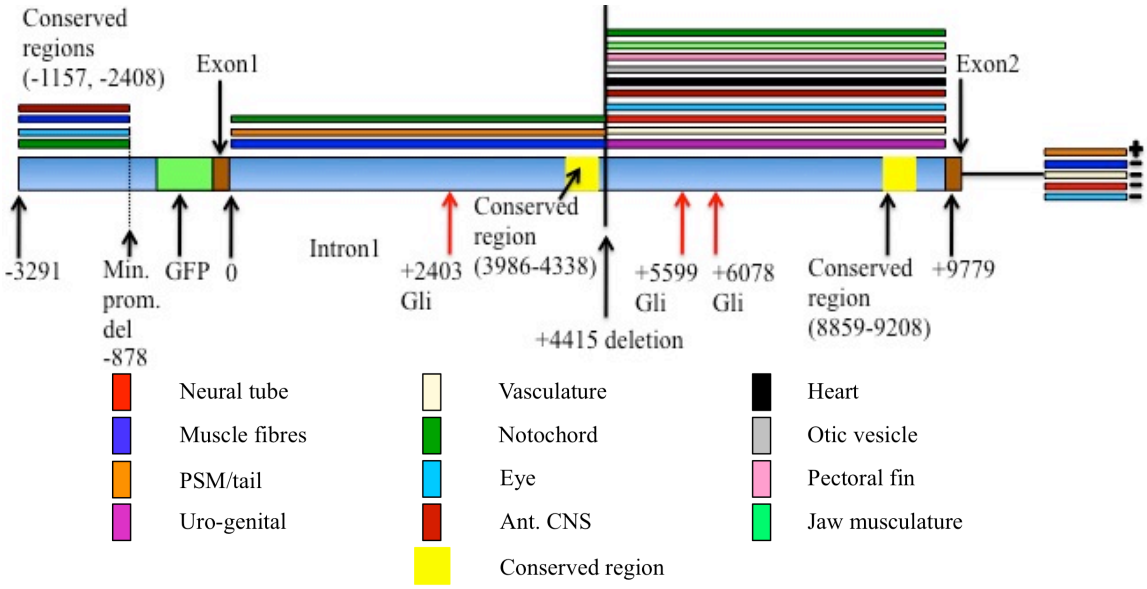

Figure 5.32: *A schematic representation of the enhancers controlling lama1 expression in the zebrafish.* Intron 1, in combination with -3291 bases upstream of the translational start site, is sufficient to drive the expression of a GFP reporter gene in a pattern that recapitulates that of the *lama1* BAC (FL). Tissues which are responsive to Hh signalling may have enhancer sites located around the Gli sites (red arrows). The conserved regions of DNA (yellow boxes) are also likely to regulate *lama1* expression. In addition to the regulatory elements located within intron 1 of *lama1*, other enhancer elements are likely to reside downstream of exon 2, which could maintain expression in the PSM (marked with a +), or repress expression in the muscle fibres, vasculature/pro-nephric tubules, neural tube and the eye at late developmental stages (marked with a -). Weak enhancers that regulate expression of *lama1* in the notochord, eye, anterior CNS, and the muscle fibres could also be located within 3291 bases upstream of the transcriptional start site. Tissue-specific enhancers are indicated with a colour-coded key, below the figure. Exons are shown as brown rectangles, GFP reporter gene as a light green rectangle.

The deletion analyses were performed using a transient transgenesis approach, which I have already shown causes a persistent GFP expression not present in the stable transenic line of *lama1:GFP* zebrafish. Based on the results I obtained using transient transgenesis, I predict that the regulatory elements downstream of exon 2 that are required for efficient expression of GFP in the PSM, will prove to be also necessary in a stable line of *lama1:GFP* zebrafish up to about 25hpf. After this developmental stage, *lama1* and GFP expression are down-regulated in the PSM, and also in the muscle fibres, neural tube, vasculature/pro-nephric tubules, and the uro-

genital region. One may speculate that this down-regulation is controlled by negative elements downstream of exon 2 (*Figure 5.32*). In this way, it is possible that transcriptional repressor proteins are up-regulated after 25hpf, which cause activation of the repressive enhancer elements. It is also possible that these inhibitive regulatory proteins are present all the time, although up to 25hpf they themselves are inactive due to the presence of other regulatory proteins.
Alternatively, repressive enhancer elements may not have a role in the regulation of *lama1*. Instead, down-regulation of *lama1* in zebrafish embryos could be the result of down-regulation or depletion of activating transcription factors.

5.3.2: Does Hh signalling directly regulate *lama1* expression?

The zebrafish intron 1 region, which drives *lama1* expression in the muscle fibres, PSM, and the notochord (bases 1 to 4415), contains a domain with >70% homology with *Fugu* fish (bases 3986 to 4339) and a single Gli binding site at position 2403 within intron 1 (*Figure 5.32*). One may hypothesise that this conserved DNA sequence and the Gli binding site are involved in the control of *lama1* expression. In agreement with this hypothesis, up-regulation of Hh signalling causes up-regulation of *lama1* expression in muscle fibres, anterior CNS, neural tube, vasculature, and the PSM, whilst down-regulation of Hh signalling causes a reduction of *lama1* expression in the PSM and uro-genital region. GFP expression in the muscle fibres and PSM is controlled by the intronic sequence that contains the conserved region and the Gli binding site described above (bases 1 to 4415). Thus, the Gli binding site in intron 1 could modulate *lama1* transcription by responding to increased levels of Gli activator or repressor protein caused by up-regulation or down-regulation of Hh signalling, respectively. Enhancer elements controlling *lama1* expression in the muscle fibres and PSM may therefore be located around this Gli site at position 2403 within intron 1 (*Figure 5.32*). The notochord element, which is not responsive to Hh signalling, could be located further upstream or downstream of the Gli binding sites, or in the conserved region (*Figure 5.32*).
There are also 2 Gli binding sites (positions 5599 and 6078) and a conserved region in the 3'half of intron 1, between bases 4416 and 9779 (*Figure 5.32*). This DNA sequence is required for *lama1* expression in the anterior CNS, eye, otic vesicles, jaw musculature, neural tube, heart, pectoral fins, vasculature/pro-nephric tubules and the uro-genital region. As Hh signalling also regulates the expression of *lama1* in the anterior CNS, neural tube, vasculature/pro-nephric tubules, uro-genital region (sections 3.2.5.1, 3.2.5.2), and GFP in the neural tube (section 4.2.8), it is possible that Hh signalling modulates *lama1* transcription in the above tissues through these Gli binding sites. Alternatively, critical enhancers controlling the expression of *lama1* in the anterior CNS, neural tube, vasculature/pro-nephric tubules, and the uro-genital region may be located between bases 4416 to 9779, but modification of *lama1* expression in response to alterations in the levels of Hh signalling could depend on the Gli binding site at position 2403.

There are also Gli binding sites within intron 2 (positions 922 and 1290) and intron 3 (position 1665) of zebrafish *lama1*. My results also indicate that a positive regulatory element controlling *lama1* expression in the PSM could be located downstream of exon 2 (*Figure 5.32*). Therefore, it is plausible that the Gli binding sites in introns 2 and 3 regulate *lama1* expression in the PSM. In support of this, I have shown that *lama1* expression in the PSM requires Hh signalling (section 3.2.5.1).

Similar to the zebrafish, intron 1 of mouse *Lama1* also contains three Gli binding site sequences (unpublished data, Kalin Narov, thesis) (Vokes et al. 2007). Overall, it could be suggested that Gli binding sites in intron 1 of zebrafish *lama1* function to modulate or fine-tune the expression of *lama1*. In mouse embryos, in which Hh signalling is critical for *Lama1* expression in the somites and neural tube, the presence of these three Gli binding sites could be essential for *Lama1* expression. However, there is no evidence as of yet that Hh signalling directly regulates *lama1* expression in the zebrafish or the mouse. The identification of three potential Gli binding sites within intron 1 of zebrafish *lama1*, combined with dnPKA data (sections 3.2.5.2, 4.2.8), brings us closer to the possibility that Hh signalling does in fact directly regulate *lama1* expression. Further deletion analyses are required to narrow down the regions required for *lama1* expression and its response to Hh signalling, and to test whether the conserved DNA sequences and Gli binding sites are capable of activating *lama1* expression. Sequences found to activate GFP expression could be analysed for the presence of Gli binding sites, and Gli binding to these sequences could be assessed by EMSA experiments. Potential functional sequences containing Gli binding sites could be modified by site directed mutagenesis to assess whether Gli binding is necessary for *lama1* reporter gene expression.

Further deletion analysis within intron 1 will also establish whether individual enhancers are sufficient to drive tissue-specific expression of *lama1*, or whether multiple enhancers co-operate together for the efficient expression of *lama1* in a specific tissue, as is the case for *Shh* expression in the zebrafish floor plate (Ertzer et al. 2007).

5.3.3: Enhancers that regulate *lama1* expression in the pectoral fin, heart, and jaw musculature are distinct from those controlling *lama1* expression in the trunk musculature

I have shown that *lama1* (and GFP in the *lama1:GFP* stable line) are expressed throughout the pectoral fins, at 49hpf and 72hpf, including in the pectoral fin musculature. Construct Δmin.prom*lama1*GFPΔ5k, which removes the region -3291 to -878 upstream of the *lama1* transcription start site, and the region 4416 to 9779 within intron 1, shows normal GFP expression within the somitic muscle fibres but a loss of GFP expression within the pectoral fins, suggesting an uncoupling of the enhancers. This demonstrates that although both muscle progenitor cells for somitic and fin muscles derive from the somite (Neyt et al. 2000), *lama1* activation in these cells is controlled by different mechanisms. GFP is expressed throughout the paraxial mesoderm including somites 2 and 4, which are the origins of the myogenic migratory cells that form the pectoral fin musculature (Neyt et al. 2000). Like *myoD*, it is possible that *lama1* expression is initiated in the somite, and is then down-regulated after somitogenesis. Following migration to the fin bud, pectoral fin myogenic progenitor cells initiate *myoD* (Neyt et al. 2000) and *lama1* expression, suggesting different mechanisms for the regulation of *myoD* and *lama1* occur between the fin bud and the somite. There is a precedent for this in tetrapods. For instance, *Myf5* expression in epaxial and limb muscle progenitor cells is controlled by distinct enhancers and different signalling mechanisms (Borycki et al. 1999; Buchberger et al. 2007). Results indicate therefore that Δmin.prom*lama1*GFPΔ5k is sufficient for *lama1* expression in the somites, but not for initiation of *lama1* in the pectoral fin. This means that the enhancers regulating *lama1* expression in the pectoral fin are probably located between bases 4416 and 9779 within intron 1, whereas the somitic enhancer is located between bases 1 and 4415 of intron 1 (*Figure 5.32*). Therefore, it is possible that mutation in the regulatory elements controlling *lama1* expression in the somitic muscle fibres could cause a specific loss of *lama1* within the somite. In this case, expression of *lama1* in the eye, anterior CNS, neural tube, notochord and the pectoral fin would be normal, and their development should be unaffected. Somitic muscle fibres lacking Laminin α1 however could show signs of dystrophy and muscle detachment from the myotendinous junctions, as observed by Sztal et al. (2012) in zebrafish lacking *lama1*.

Likewise, my data raise the possibility that mutations within the pectoral fin enhancer in intron 1 of *lama1* could result in lack of *lama1* expression specifically within the muscle fibres of the pectoral fin. Although no role for Laminin α1 in the development of the pectoral fins has been reported, Laminins are required for pectoral fin development (Webb et al. 2007), and limb development in chick (Godfrey et al. 1988) and mouse embryos (Godfrey and Gradall 1998). Basal lamina components are concentrated in the limb buds just before the differentiation process of myogenic cells begins (Godfrey and Gradall 1998), suggesting that, similar to in vitro experiments (Foster et al. 1987), the extra-cellular matrix promotes myogenic differentiation. Zebrafish carrying mutations in *lama5* have defective epidermal fin fold morphogenesis due to

disruptions in the integrity of basement membranes in the apical ectodermal ridge (AER). In the absence of an apical fold, pectoral fin outgrowth is stunted (Webb et al. 2007). Accordingly, loss of other components of the AER basement membrane cause similar defects. Limb abnormalities are observed in mice carrying mutation in both *Nidogen1* and *Nidogen2* (Bose et al. 2006). Loss of the AER in this mutant causes abnormal patterning and morphogenesis of the limb bud, in part due to altered secretion of FGFs from the AER, which function to regulate the proliferation of the underlying mesechymal cells (Martin 1998). Defects in mesenchymal cell proliferation result in stunted limb growth. The AER is also required for the secretion of BMPs, which play a crucial role in digit patterning and digit separation (Dahn and Fallon 2000; Bose et al. 2006). Likewise, loss of Laminin α5 also results in syndactyly, caused by abnormal AER development and digit separation (Miner et al. 1998).

Together, the expression of *lama1* in the pectoral fin bud, controlled by enhancer elements located between 4416-9779bp within intron 1, is likely to play a crucial role in AER basement membrane stability, thereby regulating the patterning and outgrowth of the fin bud. The presence of Laminin α1 in the ECM of the fin bud may also have an important role in the differentiation of the myogenic cells.

The region of intron 1 which regulates *lama1* expression in the pectoral fin contains 2 Gli binding sites and a conserved region of DNA. If *shh* (expressed in the posterior of the pectoral fin, as in other species (Mercader 2007)) controls the expression of *lama1* in the pectoral fin, then it is possible that the pectoral fin enhancer overlaps with the Gli binding sites. If Hh has no role for *lama1* expression in the fin, then the enhancers could instead be located upstream or downstream of the Gli binding sites, possibly within the region of DNA which is conserved with *Fugu*. Analysis of *lama1* expression in *smu* zebrafish at 48hpf would help clarify the role of Hh signalling in the expression of *lama1* in the pectoral fin. Further investigations using a stable line generated from Δmin.prom*lama1*GFPΔ5k-injected embryos will establish whether GFP expression is lost throughout the pectoral fin, or in specific cell types. If in the stable line GFP was only lost from the posterior of the pectoral fin, then this would support the idea that Hh signalling regulates the expression of *lama1*. However, *shh* expression in the posterior of the pectoral fin could be responsible for expression of *lama1* throughout the entire fin, and so a loss of *lama1* in a specific region would not be observed.

Myocardial progenitor cells arise from the lateral plate mesoderm (Yelon 2001), and so are distinct in origin from the somitic muscle fibres. Rather than a dependence on Hh and FGF8 signalling as is the case for somitic muscle fibres in the zebrafish, Wnt and BMP signalling appear crucial for the specification of cardiac progenitor cells (Yelon 2001; Liu and Stainier 2012). Unlike in the fast muscle compartment of the somite which requires retinoic acid signalling for its formation (Hamade et al. 2006), retinoic acid signalling restricts the number of cardiac progenitor cells that form (Keegan et al. 2005). A role for Notch signalling in cardiac

differentiation has also been described (Klaus et al. 2012). It can be suggested that *lama1* expression in the heart and somitic muscle is controlled by separate regulatory mechanisms which operate through different enhancers. Although the role of Laminin in cardiac development has not been studied in depth, an association between *lama2* mutation and cardiac abnormalities has recently been reported (Carboni et al. 2011).

Muscles of the jaw also arise from a distinct origin to the somitic mesoderm (Schilling and Kimmel 1994; Shih et al. 2008). At 12hpf, a ventral mesoderm layer is generated in the zebrafish head, which contributes to the formation of the pharyngeal and branchial arches. The jaw musculature is derived from these arches (Schilling and Kimmel 1994). In line with the fact that head and somitic muscle have distinct origins, the specification cues are also different (Shih et al. 2008). Pitx2, Tbx1 and Tcf21 initiate myogenesis in the head through regulation of *MyoD* and *Myf5* in the mouse (Shih et al. 2008), whilst in the zebrafish a role for Ret tyrosine kinase signalling has been identified to specifically regulate myogenesis of opercular muscles (Knight et al. 2011), which express *lama1*. The absence of GFP expression in jaw musculature in Δmin.prom*lama1*GFPΔ5k-injected embryos suggests that the enhancers controlling *lama1* expression in the head mesoderm have been lost in this construct, suggesting that, as in the pectoral fins and the heart, the enhancers controlling *lama1* expression in the head mesoderm are likely located between bases 4416 to 9779 bases within intron 1 (*Figure 5.32*).

Overall, my deletion analyses have lead to the identification of a distinct region of intron 1 which controls *lama1* expression in the somites, PSM, and notochord, and a region which controls *lama1* expression in the anterior CNS, neural tube, eye, otic vesicle, jaw musculature, heart, vasculature/pro-nephric tubules, uro-genital region, and the pectoral fin. I have also determined that other enhancer elements in intron 1 or downstream of exon 2 contribute to the maintenance of *lama1*, during zebrafish embryonic development. In combination with my deletion analyses, bioinformatic analysis has allowed the prediction of the mechanisms responsible for tissue-specific expression driven by these enhancer elements.