Chapter 4

Analysis of transient and stable transgenic *lama1* GFP zebrafish embryos

4.1: Introduction

The identification of the DNA sequences required for gene expression, known as promoters or enhancers, can provide valuable information regarding the regulatory mechanisms that control that specific gene. These DNA sequences are often identified as functional by the activation of a reporter cassette in vivo, to which they are associated with (Fisher et al. 2006). DNA sequences capable of driving reporter gene expression are injected into zebrafish or mouse embryos at the one-cell stage, and can integrate into germ-line DNA. In this way, zebrafish or mouse transgenic lines are generated, as the integrated DNA sequence and reporter cassette will be inherited by the offspring (Bussmann and Schulte-Merker 2011).

The zebrafish is particularly amenable to transgenesis, due to the external fertilisation of hundreds of embryos each day, and their optical clarity. Multiple zebrafish can therefore be raised, with clear visualisation of a reporter gene. In contrast, mouse transgenesis is more expensive and more labour intensive, involving the transplantation of injected eggs to the oviducts of foster mothers. This also limits the number of transgenic mice which can be generated at one time, and unlike in the zebrafish, live-imaging of reporter gene expression throughout mouse embryonic development cannot be performed.

Analysis of reporter gene expression in transgenic zebrafish lines indicates which DNA sequences are required to completely recapitulate the endogenous gene mRNA expression. Absence of reporter gene expression from a given tissue suggests that critical enhancer DNA sequences are absent from the reporter cassette (Müller et al. 1999). Often, the regulatory elements controlling a gene are spread across a large distance of DNA (Bussmann and Schulte-Merker 2011), and so cannot be completely cloned into a reporter cassette. As a result, transgenic expression patterns may not represent the endogenous expression pattern of the respective gene (Zhang and Rodaway 2007; Bussmann and Schulte-Merker 2011). For these reasons, the bacterial artificial chromosome (BAC), which can contain up to 350kb of selected genomic DNA is emerging as a powerful tool to identify regulatory elements that control gene expression, and as a tool to create transgenic lines in both zebrafish and mice (Yang et al. 2006; Suster et al. 2009; Bussmann and Schulte-Merker 2011).

Homologous recombineering is the mechanism by which DNA is precisely exchanged between two molecules of identical sequence (Court et al. 2002; Sharan et al. 2009). It allows for DNA sequence insertion, modification, or deletion, and does not require the presence of restriction sites (Sharan et al. 2009). Therefore, homologous recombineering can be used to insert GFP-reporter cassettes into the transcriptional start site of a gene of interest within a BAC (Bussmann and Schulte-Merker 2011). Providing that the necessary enhancer elements are present within the BAC that would normally activate a given gene's expression, the GFP reporter gene will be activated instead.

Enhancer elements that control the expression of a gene (or reporter in the BAC) can be located upstream or downstream, or in the introns of the gene (Müller et al. 1999; Haeussler and Joly 2011; Kolovos et al. 2012). Some tissue-specific enhancers have even been identified within coding regions of DNA, such as the 5 'UTR in exon 1 of *Pax6* (Zheng et al. 2001) and *IGF-1* (McLellan et al. 2006), and within a translated coding exon in *Hoxa2* (Tumpel et al. 2008). It is important therefore when searching for the critical enhancer elements controlling a gene, that a BAC is selected which contains the entire relevant gene, including upstream and downstream DNA sequences.

To study the regulatory network controlling *lama1* expression in the zebrafish and to identify regulatory elements, I elected to perform BAC transgenesis in the zebrafish. I aimed to select a BAC containing the entire zebrafish *lama1* genomic sequence, including upstream and downstream DNA sequences. The goal was to insert a GFP cassette into the *lama1* gene start site by homologous recombination, to generate a reporter gene for the *lama1* gene containing BAC. I hypothesised that this *lama1* GFP reporter BAC will be expressed in a temporal and spatial pattern that recapitulates endogenous *lama1* expression. This hypothesis assumed the presence of all of the enhancers that regulate *lama1* transcription. Analysis of the *lama1* GFP expression pattern will reveal whether all the necessary enhancers and regulatory elements are present within the BAC.