論 文 内 容 の 要 旨

Introduction

Sox (Sry-related HMG box) family transcription factors (TFs) play a major role as central coordinators of gene regulatory networks in the embryonic development processes such as cell differentiation and tissue development. Our group has been studying these Sox TF-mediated transcriptional regulation of early embryonic development using the zebrafish (*Danio rerio*) model[1]. Although, the major players of these signaling pathways are known, there are many crucial details yet to be revealed, in order to create a genome-wide picture of Sox TF-mediated gene regulatory network. For this purpose, chromatin immunoprecipitation followed by next-generation DNA sequencing (ChIP-seq) is a promising method, as it has already been used widely to map regulatory elements and analyze TF function throughout an entire genome. However, the reliability of the results obtained through ChIP-seq highly depends on the affinity and the specificity of the antibodies used to capture the protein complexes. Antibodies that show satisfying affinity and specificity during ChIP-seq are generally called "ChIP-grade" antibodies. Despite the popularity of ChIP-seq applications, the unavailability of ChIP-grade antibodies against each of the proteins of interest has hampered wider applications of ChIP-seq. Particularly, TF families like Sox proteins that share higher sequence similarities show antibody cross-reactivity and make it difficult to perform reliable ChIP-seq experiments.

Epitope tagging of the TFs and subsequent use of epitope tag specific antibodies for ChIP-seq is an interesting strategy to avoid this problem. Further, this strategy attracts much attention due to the ability to adapt Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 genome editing tool to epitope tag endogenous TFs and thus facilitates the expression of the tagged proteins at near-endogenous levels[2]. Amongst the candidate epitope tags that can be used for ChIP-seq such as FLAG, HA and V5, the FLAG peptide is the most widely reported epitope tag in ChIP-seq experiments. An important consideration when selecting suitable epitope tags for ChIP is to avoid lysine or histidine residues within the amino acid sequence due to the fact that lysine residues are the primary targets for formaldehyde cross-linking. Thus, epitope tags that contain lysines will be functionally destroyed (at least partially) by formaldehyde cross-linking. For this reason, the FLAG tag has now raised controversies of its usage in ChIP-seq experiments. Another consideration when selecting epitope tags for ChIP-seq is their ability to obtain substantial enrichment over the control samples[3]. On the other hand, the performance of an epitope tag in a ChIP-seq experiment depends not only on the amino acid sequence of the epitope used but also on the quality of the anti-epitope antibody. Therefore, the selection of optimal epitope tag/antibody combinations is a prerequisite for successful ChIP experiments and remains to be evaluated.

Objectives of the study

- To select optimal epitope tag/antibody combinations suitable for successful ChIP experiments
- To generate transgenic zebrafish lines with epitope tagged Sox proteins using CRISPR/Cas9 system

Part－1
Selection of optimal epitope tag/antibody combinations suitable for successful ChIP experiments

In this study, I selected five different epitope tags such as FLAG, HA, PA, Ty1 and V5, and monoclonal antibodies against each of these epitope tags based on their wide usage in research and the commercial availability.

There was a least of one candidate antibody clone to up-to four candidate antibody clones against each epitope tag. Thus, the most suitable antibody clone had to be selected for further experiments, and that required the understanding of the antibody performance in advance specially, under the ChIP assay conditions. Antibody performance can be evaluated by their specificity and/or affinity towards their cognate antigens. Here, I tried to develop a method to determine the antibody affinity or its quantitative determinant–dissociation constant (K_d), under IP conditions which is however challenging, mainly because the amount of precipitated protein during IP is often near or below the lowest limit of quantitative detection by Western blotting.

Method development–HiBiT-qIP: To determine antibody affinity under immunoprecipitation conditions

By employing a quantitative NanoLuc-based HiBiT detection system, a simple and relatively inexpensive approach for determining the antibody Kd under IP conditions was developed. This method was called HiBiT-qIP, which is short for "HiBiT-based quantitative immunoprecipitation". The HiBiT system is based on a split luciferase complementation of two NanoLuc subunits namely High BiT (HiBiT) and Large BiT (LgBiT). The high-affinity (K_d = 0.7 nM) binding of novel 1.3-kDa (11-amino acid) HiBiT peptide to 18-kDa LgBiT efficiently forms a stable complex that acts as the active NanoLuc luciferase, which enables HiBiT to serve as a quantitative luminescent peptide tag[4]. Thus, tagging a protein of interest with the HiBiT peptide facilitates sensitive quantification of the amount of HiBiT-tagged protein, which makes it possible to measure protein amounts of less than 1 amol (e.g., 0.05 pg of a 50-kDa protein). Furthermore, a simple add-mix-read assay protocol of the HiBiT detection system enabled to perform the IP-based equilibrium binding analysis more easily.

In order to determine the "apparent" Kd values of epitope tag/ antibody interactions, epitope tags mentioned above in their monomeric, dimeric (x2) or trimeric (x3) forms were fused to the GST protein and the HiBiT peptide. Varying amounts of the purified epitope-tagged GST protein were mixed with a fixed amount of cognate monoclonal antibody immobilized on anti-IgG magnetic beads in a stringent IP buffer, which contains similar components often used in standard ChIP. After overnight incubation, the unbound proteins were washed away and the amount of bound epitope-tagged GST protein was determined by adding excess amount of LgBiT protein to the diluted IP eluate and measuring the luminescence signal derived by the HiBiT/LgBiT complex. By data fitting to the binding model: $[L_b]/[L_{b\_max}] = [L_d]/(K_d+[L_d])$, a saturation curve of bound GST as a function of free GST was plotted and the "apparent" K_d values were determined. Interestingly, the HiBiT/LgBiT complex produced luminescence signal in a linear correlation with the amount of protein tagged with HiBiT peptide under the assay conditions used during HiBiT detection (0.001% final [SDS] and 0.1% [Triton X-100]) allowing the quantification of protein amounts as low as approximately 0.33 pg (0.01 fmol).

A considerable variation in K_d values could be observed for the antibodies tested against the monomeric form of epitope tags. However, each epitope tag tested had at least one high affinity antibody clone that can be used under ChIP buffer conditions. Interestingly, high affinity interactions were exhibited by Ty1, V5 and PA epitope tag/antibody combinations even though these epitope tags have been used less commonly in IP experiments than the FLAG and HA tags.

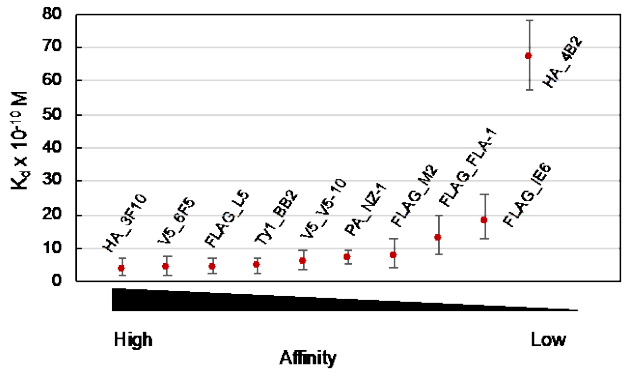

Figure 1. Affinity comparison of the antibody clones against the monomeric form of the epitope tags. Error bars depict the plus and minus confidence interval of the K_d value.

Moreover, a significant increase in affinity was observed with the use of epitope tags in dimeric or trimeric

form. The antibody clones that showed low affinity impressively recovered up to considerably high affinity while using the multimer forms. To examine the effects of multimer forms of epitope tag use on the efficiency of IP in real experimental conditions, I prepared crude cell lysates of Zebrafish embryos expressing Sox3 tagged with a monomeric or trimeric form of the FLAG tag and HiBiT at near-endogenous levels and performed IP. In accordance with the differences in affinity, a considerable improvement in IP recovery was clearly observed with the use of the trimeric form of the FLAG tag, and this effect was more pronounced if a limited amount of antibody was used. Through these experiments, I decided to use the trimeric form of epitope tags during the future experiments.

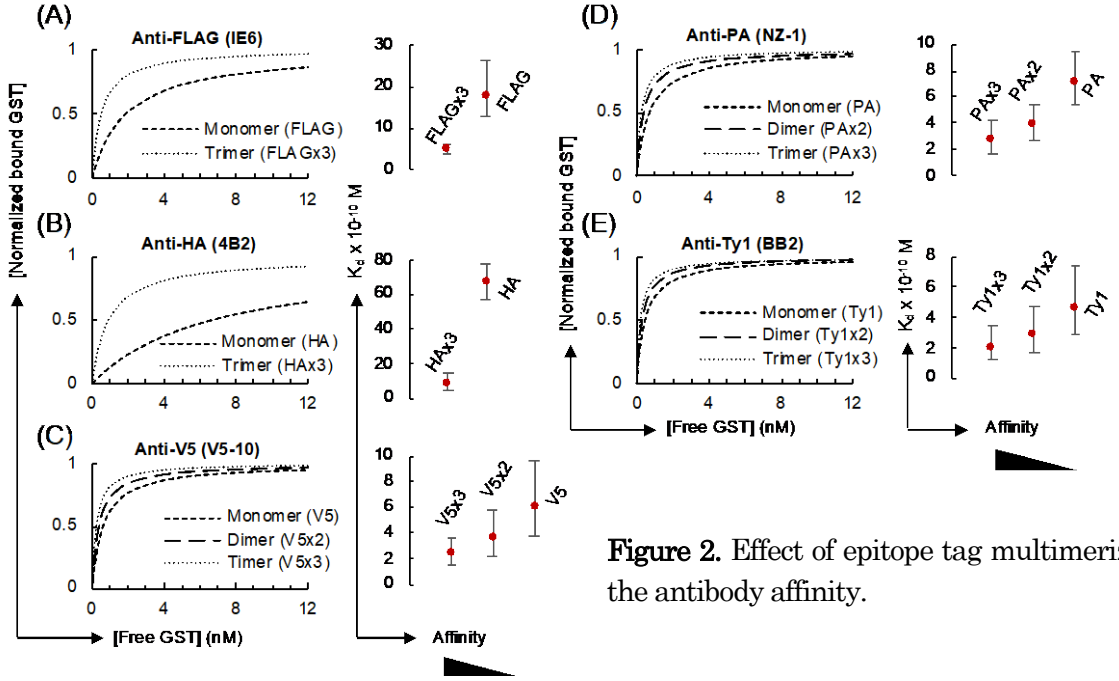

Figure 2. Effect of epitope tag multimerization on the antibody affinity.

Part－2

Generation of transgenic zebrafish lines with epitope tagged Sox3 gene using CRISPR/Cas9-mediated knock-in approach

As I could select optimal epitope tag/antibody combinations that performs well under ChIP assay conditions, next, I tried to tag zebrafish *Sox* genes endogenously with those epitope tags using CRISPR/Cas9-mediated knock-in (KI) approach. At the first instance, to tag the *Sox3* gene with the trimeric form of FLAG or PA epitope tag at the 3' end of the coding DNA sequence (CDS) was attempted. To achieve that, at first, I prepared a composite of epitope tags containing an epitope tag FLAG or PA in trimeric form, followed by a TEV protease cleavage site, a biotin acceptor domain (Bio tag) and, most C-terminally, the HiBiT peptide. High-affinity Bio-tag was included to be able to perform cell-type specific ChIP-experiments in the future studies by expressing BirA (Biotin ligase) selectively. Also, to pull-down biotinylated protein complexes by streptavidin and to identify those proteins that partner with target TFs using mass-spectroscopy.

To knock-in the composite of epitope tags to the 3' end of the Sox3 CDS, two candidate gRNAs were selected in-silico based on the predicted high on-target cleavage efficiency and low off-target cleavage potential (Fig. 3A). Next, cleavage efficiency of those gRNAs were evaluated *in vivo* by injecting candidate gRNAs and Cas9 protein (RNP complex) into the cytoplasm of 1-cell stage zebrafish embryos followed by heteroduplex mobility assay (HMA) (Fig. 3B). Based on the intense smear and the light single band compared to the uninjected control, crRNA number 1 considered as the most efficient in inducing mutations at the target site.

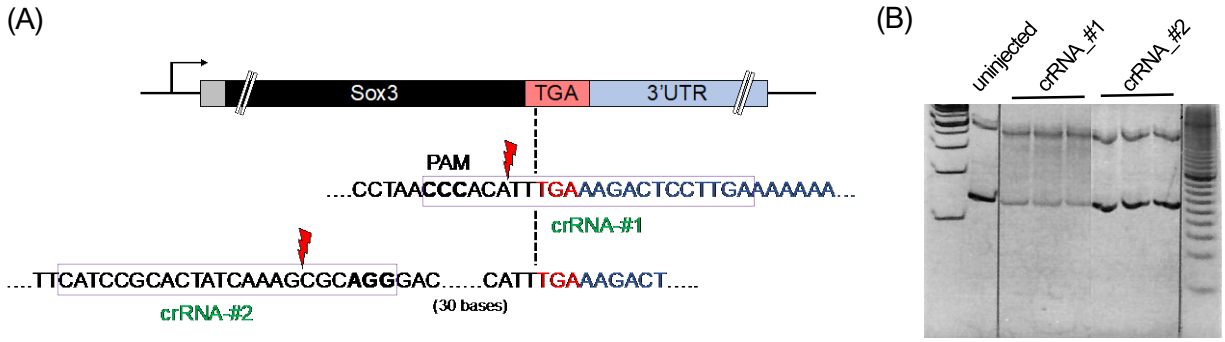

Figure 3. (A) Candidate crRNA location and sequence (B) Heteroduplex mobility assay to evaluate the cleavage efficiency of crRNAs. Three replicates of embryo lysates per crRNA

As the KI DNA donor template, which will assist homology directed repair (HDR) of the double strand break (DSB) caused by RNP complex, a single-stranded DNA (ssDNA) template that consists of the composite of epitope tags flanked by homology arms of Sox3 coding DNA sequence (CDS) and the Sox3 3'UTR regions was used. Due to the reported low KI efficiency in zebrafish and the limited reports about successful long sequence insertions to the zebrafish genome, several factors such as the orientation of the ssDNA, homology arm length, and symmetry or asymmetry of homology arms that would yield higher KI efficiency were examined.

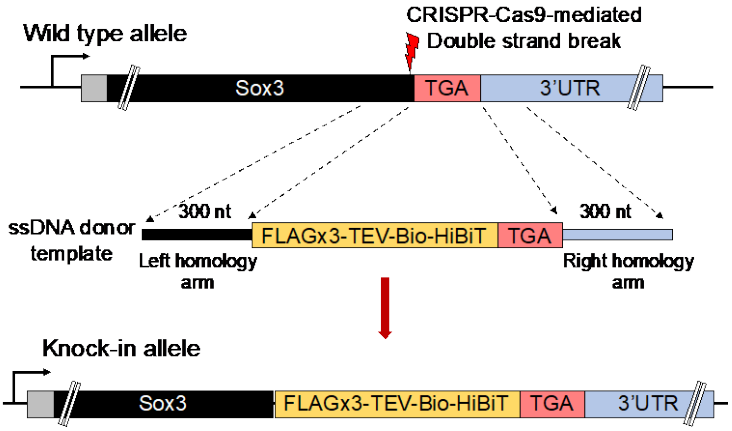

Figure 4. ssDNA donor template design

Interestingly, only the 'target' ssDNA (the strand bound by the gRNA) injected embryos produced expected amplicons for the KI-allele specific PCR. However, PCR with primers outside of the insertion (wild-type PCR) did not produce amplicons derived from the KI-allele but, only the wild-type allele, suggesting the low copy number of the KI allele (Fig. 5A). Further, 19 out of 20 injected embryos showed the expected PCR amplicon for the KI-allele specific PCR indicating the success of the CRISPR experiment and the higher KI efficiency (Fig. 5B).

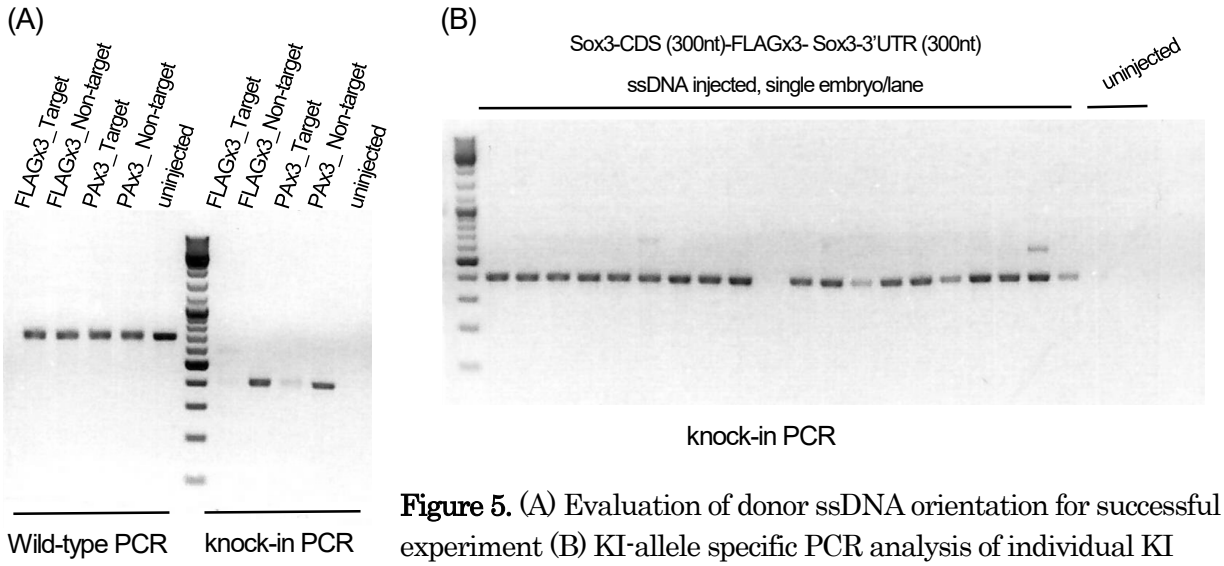

Figure 5. (A) Evaluation of donor ssDNA orientation for successful KI experiment (B) KI-allele specific PCR analysis of individual KI

To examine the efficacy of germline transmission of the KI insert, at the age of four months, the potential knock-in founders were mated to each other and genomic DNA was extracted from clutches of F1 embryos. For KI-allele specific PCR of each pooled clutch, 3 out of 10 pairs showed expected amplification, suggesting the germline transmission (Fig. 6).

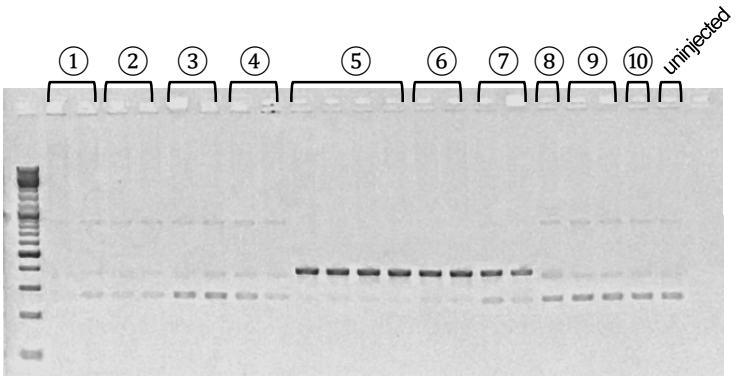

Figure 6. KI screening of extracts from embryo clutches produced by potential founders.

References

[1] Okuda, Y., Ogura, E., Kondoh, H., Kamachi, Y. B1 SOX Coordinate Cell Specification with Patterning and Morphogenesis in the Early Zebrafish Embryo. *PLos Genet*, 6(5):e1000936 (2010).
[2] Savic et. al. CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins, *Genome Res.*, 25:1-9 (2015).
[3] Brinkman, A.B. & Stunnenberg, H.G. Strategies for Epigenome Analysis in *Epigenomics* 3-18 (Springer Netherlands, 2009).
[4] Schwinn, M. K. et al. CRISPR-mediated tagging of endogenous proteins with a luminescent peptide. *ACS Chem. Biol.* 13, 467–474 (2018)