Effect of cis-regulatory element CNE14 on tissue-specific GLI3 expression in developing mammals

Introduction

Hedgehog signalling is a pathway known to regulate development of limbs and central nervous system in all bilateral animals [1]. In vertebrates, gene expression changes in response to the signalling ligand Sonic Hedgehog (Shh) are mediated largely by zinc-finger transcription factors of the GLI family [2]. The main transcription factor implicated in digit patterning and limb development is GLI3, which is expressed in limbs and nervous tissue of CNS during development [2]. In the absence of Shh, Hedgehog signalling is off, and GLI3 is processed post-transcriptionally into a truncated transcription factor, repressing its target genes. When Shh is present in the tissue, intracellular signalling pathways prevent the processing of GLI3, leading to target gene activation by the full-length protein [1]. A gradient of Shh in the tissue leads to a corresponding gradient of GLI3 processing events, contributing to digit patterning in limbs and CNS [1]. Correct spatiotemporal expression of GLI3 during development (pre- and postnatally) is crucial for correct balance of GLI3 processing and successful formation of limbs, craniofacial structures and CNS [2].

GLI3 expression is regulated by several cis-regulatory elements. The intronic regions GLI3 contain several conserved non-coding elements (CNEs) that are implicated in gene regulation [2]. In 2015, Anwar *et al.* described two novel CNEs (CNE13 and CNE14) conserved in all tetrapods [2]. The individual CNEs seem to have distinct activities and tissue-specificity [2], suggesting they play differential roles in GLI3 regulation. The mechanisms of regulatory functions and tissue-specific activities remain unknown.

This project focuses on the regulatory actions of CNE14. The paper by Anwar *et al.* shows that CNE14 is capable of driving reporter gene expression in the developing pectoral fin of zebrafish, but does not induce gene expression in nervous tissue, contrary to other CNEs [2].

An *in vitro* assay in NIH3T3 cells (mouse embryonic fibroblasts) shows that CNE14 can induce reporter gene expression in a mammalian system [2], but does not address the tissue-specificity. This project sets out to investigate the regulating actions of CNE14 *in vivo* by addressing the levels and tissue-specificity of GLI3 expression in CNE14 knockout mouse models.

Relevance, potential impacts and importance

Hedgehog signalling is a main regulator of limb formation and CNS development in vertebrates [1]. Dysregulation of this pathway can lead to syndromes such as Greig Cephalopolydactyly syndrome and Pallister-Hall syndrome, displaying phenotypes such as polydactyly, deformity of limbs and craniofacial structures, and in some cases impaired brain development [3]. In order to characterize these syndromes more efficiently, it is of great importance to characterize and understand the expression of GLI3, as this particular transcription factor is an important modulator of the implicated pathway. The emergence of two novel intragenic cis-regulatory elements only last year [2] highlights the fact that our understanding of the spatiotemporal expression of GLI3 in development is still incomplete. Revealing details about this complex regulatory pattern and expanding our knowledge acquired in zebrafish to mammals will contribute to our understanding of how Hedgehog signalling regulates proper limb and brain development in humans and how it might go wrong.

Investigating the regulatory mechanisms of CNE14, especially concerning tissue-specificity, will contribute to our general knowledge on cis-regulatory elements and may provide important information that can be applied in parallel studies of other gene regulation patterns.

Hypothesis

My hypothesis is largely based on the two assays carried out by Anwar *et al*. in 2015 [2]. In the first assay, a zebrafish was transfected with a plasmid containing CNE14, a promotor and a reporter gene. The assay revealed that CNE14 is capable of promoting reporter gene expression in the developing limb buds of zebrafish [2], suggesting that CNE14 is important for limb-specific GLI3 expression.

The *in vitro* reporter assay in NIH3T3 cells revealed reporter gene expression driven by CNE14 [2], indicating that CNE14 is also an active regulatory element in mammals. This leads to the hypothesis that CNE14 is necessary for normal GLI3 expression patterns in limbs of developing mammals.

The conservation of CNE14 in all tetrapods supports the hypothesis that it has a central function in regulation. Additionally, comparative genomics studies predict several putative transcription factor binding sites within the region, further supporting a regulatory role [2].

Experimental plan and predictions

The experiment to test the hypothesis is a mouse knockout assay followed by gene expression analysis of a developing embryo. The knockout mouse will carry a deletion within intron 3 of GLI3, removing CNE14.

Model system

M. musculus is an ideal model system for this project, as the Hedgehog signaling pathway is conserved between mice and humans. The mouse genome and general gene regulation mechanisms resembles the human system to such an extent that it is highly probable that results obtained in mice apply to humans as well. The relatively short development time makes it possible to create a knockout mouse for research purposes within a reasonable time frame [4].

Creating a knockout mouse

Mouse embryonic stem cells (ES cells) are engineered to carry a mutated cope of the GLI3 gene with a deletion in intron 3 disrupting CNE14. This is achieved by taking advantage of the cells' pathway for homologous recombination. A replacement vector is introduced into the ES cells, and will be incorporated into the genome by a homologous recombination event, replacing the wild type GLI3 gene.

The construct will carry two sequences homologous to regions flanking GLI3, inducing recombination. It will also include the mutated GLI3 gene as well as selection markers. A gene encoding neomycin phosphotransferase (Neo^r) provides resistance against the drug Neomycin that

interferes with protein synthesis [4]. Including Neo^r in the plasmid allows for positive selection by treating the transfected cells with Neomycin – only successful integrants will resist the toxicity [4]. A negative selection marker is included to address the issue of random integration (by recombination with other sequences than the homology regions). Random integration may result in cells acquiring a copy of Neo^r, without incorporating the desired GLI3 mutation. To negatively select for these, a gene encoding HSV thymidine kinase (HSV-tk) is placed adjacent to the region flanked by homology sequences (see figure 1). Successful integrants will not incorporate HSV-tk, but random integrants containing Neo^r will likely contain the HSV-tk gene, as the two genes are not far apart. HSV-tk is able to phosphorylate and activate a compound called Ganciclovir, which inhibits DNA synthesis [4]. Treating the cells with Ganciclovir will negatively select for random integrants containing HSV-tk, without affecting the successful integrants.

Figure 1 shows the relative positions of the elements of the plasmid construct.

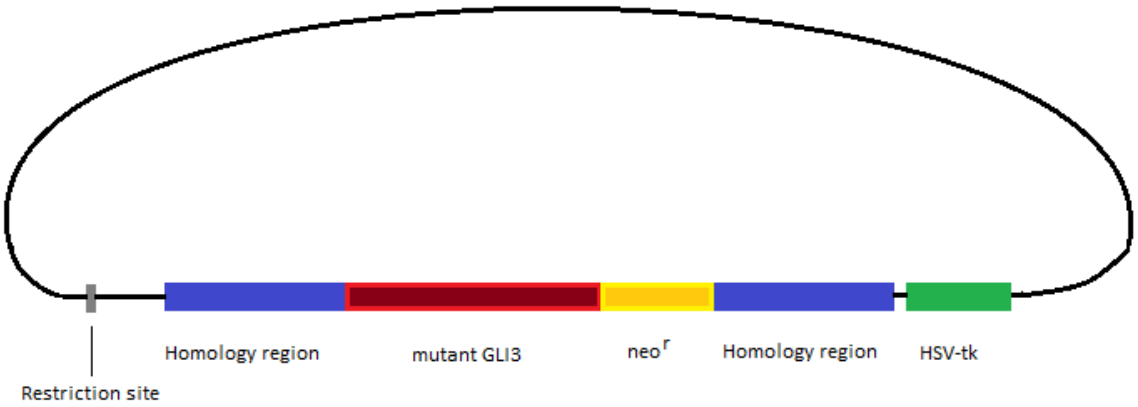

Figure 1 -schematic representation of plasmid vector

The cells are transfected with the construct by electroporation, and the construct is linearized by treatment with restriction enzyme to allow homologous recombination. Positive and negative selection is carried out by treating cells with Neomycin and Ganciclovir to enrich for successful recombination.

This knockout could be carried out with a higher efficiency by genomic editing with CRISPR/Cas9. However, I have chosen to include the other protocol, as this is the procedure of most labs that do not have access to CRISPR/Cas9 technology.

The engineered ES cells are injected into mouse blastocysts that are implanted in the uterus of a foster mother. The developed mice will be mosaics – only part of their cells will carry the mutation. The chimeric mice are mated and their offspring will be either wild type or heterozygous for the mutation, depending on whether the mutations was present in the germ cells. The offspring is genotyped to check the presence of the mutation [4].

The heterozygous mice are mated, and statistically one fourth of the resulting embryos are homozygous. Upon extraction, the embryos are sequenced to establish their genotype.

Gene expression analysis

Analysis of GLI3 gene expression in homozygous CNE14 mutants and wild type mice is done by **whole mount in situ hybridization** followed by **histochemical detection** of GLI3:

The embryos are extracted at E15, when they are sufficiently developed to display signs of limb deformity, and robust enough to allow intact extraction. They are treated (see [5]) and incubated overnight in pre-hybridization solution with a GLI3 mRNA probe, created by PCR. The embryos are washed with blocking solution to prevent unspecific antibody binding and subsequently incubated with probe-specific antibodies. After washing to remove excess antibody, the embryos are incubated with dye that stains the GLI3 mRNA bound antibodies so the tissue-specific expression patterns are visualized [5]. This procedure is done in parallel in wild type embryos for comparison.

Predictions

If the hypothesis is correct, the knockout embryos will show decreased GLI3 expression in developing limbs compared to wild type embryos. The CNS will not display altered expression levels, as CNE14 is thought to be a limb-specific regulator [2].

Possible results

Provided are three sets of possible results and the conclusions and inferences drawn from each.

Observations: Knockout embryos show decreased histochemical detection of GLI3 mRNA in limbs compared to wild type. GLI3 mRNA signals in CNS are similar in wild type and knockout. The embryo displays phenotypic polydactyly and general deformity of limbs.

Conclusions: CNE14 is necessary for normal GLI3 expression patterns in developing limbs of mice. Deletion of CNE14 is sufficient to induce the observed phenotype in developing mouse embryos.

Inferences: We can infer a causal relationship between the mutation and the observed phenotype: Deletion of CNE14 results in dysregulation of GLI3 expression, leading to decreased GLI3 mRNA levels in the limbs, causing the observed phenotype of polydactyly and limb deformity in mouse embryos. The biological resemblance of mice and humans allows us to infer that this is also the case in humans.

These results support the hypothesis by showing that CNE14 is necessary for normal GLI3 expression in developing limbs of mice. The current literature describes GLI3 regulation as a complex pattern involving many components that have not all fully been characterized [1]–[3]. This complexity is the reason that the underlying causes of GLI3 related phenotypes have not all been clarified [3]. Investigating the mechanisms of CNE14 further would contribute to revealing these mechanisms. CNE14 contains many putative binding sites for known proteins, mostly transcription factors [2]. Addressing whether these proteins bind to CNE14 *in vitro* and later *in vivo* would shed light on the underlying molecular mechanism of GLI3 regulation.

Observations: Decreased or non-existing GLI3 mRNA levels are detected in both limbs and CNS of the developing embryo. The embryo displays a severe phenotype including polydactyly, deformity of limbs and cranio-facial structures, as well as cerebral anomalies (impaired brain development).

Conclusions: CNE14 is necessary for wild type GLI3 expression patterns in developing mouse embryos, and is necessary for normal limb and brain development.

Inferences: Deletion of CNE14 impairs expression of GLI3 during development, leading to the observed phenotype in mice. Due to biological similarities, this is also the case in humans.

These results would be rather surprising. The hypothesis states that CNE14 is important in the tissue-specificity of GLI3 expression, not necessary for the general expression in all tissues. However, the results support that CNE14 has an important regulatory role, so further investigation of the mechanism could reveal important information about how the cell regulates GLI3 expression in development. As previously mentioned, comparative genomics have revealed

putative transcription factor binding sites within CNE14 [2]. This potential binding is an interesting target of investigation, as it might explain why CNE14 is necessary for GLI3 expression – perhaps it recruits some essential protein components.

Observation: There is no detectable difference in levels or distribution of GLI3 mRNA in knockout embryos compared to wild type embryos.

Conclusion: CNE14 is not necessary for wild type GLI3 expression patterns in mouse embryos.

Inferences: CNE14 is dispensable in regulating GLI3 expression in mice as well as humans. It may not have any regulatory function in GLI3 expression, and the conservation between species is due to other factors. If CNE14 is in fact involved in GLI3 expression (other experiments must be conducted to confirm this), then other mechanisms exist to restore normal GLI3 expression after the loss of CNE14.

These results do not support the hypothesis. CNE14 is not essential for GLI3 regulation in this developmental stage. However, the conservation of CNE14 suggests an important function. To establish if CNE14 plays a role in GLI3 expression later in development, a similar experiment could be conducted, extracting the embryos at a different stage.

Considering the observations of reporter gene expression by Anwar *et al.* [2], a plausible hypothesis could be that CNE14 regulates a different target gene than GLI3. Identification of such a target gene could be attempted by a large-scale gene expression analysis of CNE14 knockouts. Detection of alterations in mRNA levels could indicate potential targets of CNE14 regulation.

Summary – CNE14 function in GLI3 expression

This is Sonic Hedgehog
He is not a cartoon - he is a signalling molecule!

He makes sure we're all born with a proper brain and the right amount of fingers and toes.
But how?

I'm GLI3 - I decide how many fingers you get!

He regulates **GLI3**

GLI3 is found on two forms:
When Sonic Hedgehog is not present, GLI3 is truncated - shorter than usual.
This means GLI3 acts as a repressor of genes, preventing protein expression.
When Sonic signals to the cell, GLI3 changes to the full-length form, acting as an activator of protein expression!

So why do we care if GLI3 is late or gets lost?
We need her in the right place at the right time, otherwise we might experience impaired brain development or polydactyly (too many fingers or toes).

Researchers are even talking about GLI3 being involved in cancer. We better keep an eye on her!

Now GLI3 might show up in different tissues, she might not show up at all, or she could be too early or fashionably late - or maybe nothing changes

GLI3 is a protein - and proteins come from genes
Inside the GLI3 gene, researchers found something called **CNEs**

CNEs regulate the expression of GLI3 in different tissues at different times in development - they basically tell GLI3 where to be and when to be there
(This is important because Sonic expects her to be certain places at specific times)

SIDEBAR - CNE
Conserved (everyone has them, must mean they are important)
Non-coding (means that they don't code for a protein)
Element (stretch of DNA)

Wait - Non-coding? But they're inside a gene coding for GLI3 protein!?!
Well, proteins are encoded in fragments called exons, and in between the exons are introns. These can contain CNEs.

COOL!
Tell me more!

All the CNEs work through different mechanisms, so we need to study them one at a time. We are particularly interested in **CNE14**
The hypothesis is that CNE14 acts to promote GLI3 expression in nervous tissue and developing limbs. This is quite conveinient - remember that GLI3 is essential for development of limbs and the nervous system!

We **delete CNE14** and observe what happens to **GLI3** expression

How do you test that?

Another cool thing:
CNEs are actually found in many different genes - not just GLI3.
The more we find out about how GLI3 CNEs work, the more we will possibly know about the regulation of other important genes!

References

[1] J. Briscoe, “The mechanisms of Hedgehog signalling and its roles in development and disease,” *Nature*, vol. 14, no. July, pp. 416–429, 2013.

[2] S. Anwar, R. Minhas, S. Ali, N. Lambert, Y. Kawakami, G. Elgar, S. S. Azam, and A. A. Abbasi, “Identification and functional characterization of novel transcriptional enhancers involved in regulating human GLI3 expression during early development,” *Dev. Growth Differ.*, no. 57, pp. 570–580, 2015.

[3] F. Demurger, A. Ichkou, S. Mougou-zerelli, M. Le Merrer, A. Delezoide, L. Faivre, C. Baumann, S. Nampoothiri, L. Pasquier, B. Isidor, D. Lacombe, M. Delrue, S. Mercier, N. Philip, E. Schaefer, M. Holder, A. Krause, F. Laffargue, M. Sinico, D. Amram, A. Liquier, M. Rossi, J. Amiel, F. Giuliano, G. Andre, O. Boute, A. Dieux-coeslier, M. Jacquemont, A. Afenjar, L. Van Maldergem, A. Munnich, O. Raoul, S. Romana, L. Devisme, and D. Genevie, “New insights into genotype – phenotype correlation for GLI3 mutations,” *Eur. J. Hum. Genet.*, vol. 33, no. October 2013, pp. 92–102, 2015.

[4] B. Hall, A. Limaye, and A. Kulkarni, “Generation of Gene Knockout Mice,” *Curr Protoc Cell Biol*, pp. 1–23, 2009.

[5] E. Mcglinn and J. H. Mansfield, “Detection of Gene Expression in Mouse Embryos and Tissue Sections,” in *Vertebrate Embryogenesis*, 2011, pp. 259–292.