CRISPR/Cas9

a Tool to Recapitulate Unresolved Human Cases in Mouse

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat)
submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by
Sinje Geuer
from Aachen

2016

Erster Gutachter: Prof. Dr. rer. med. Stefan Mundlos

Zweiter Gutachter: Prof. Dr. rer. med. Silke Rickert-Sperling

Disputation am 16.11.2016

Acknowledgment & Danksagung

An erster Stelle bedanke ich mich bei Prof. Dr. Stefan Mundlos, der mich als Doktorvater betreut und es mir ermöglicht hat an seinem Institut für Medizinische Genetik und Humangenetik der Charité (Universtitätsmedizin Berlin) und in seiner Arbeitsgruppe am Max Planck Institut für Molekulare Genetik zu promovieren. Bedanken möchte ich mich auch bei meinem Zweitgutachter Prof. Dr. Silke Ricke-Sperling für die Bewertung meiner Arbeit. Besonders dankbar bin ich meiner Betreuerin Dr. Eva Klopocki, die mich im ersten Jahr in die Gruppe und ihre Methoden eingeführt hat und mich auch nach ihrem Wegzug stets unterstützt und gefördert hat. I am very thankful that Dr. Dario Lupiáñez took over her task and helped me with advices - especially during the writing process of my thesis. In meinem Dank möchte ich besonders Dr. Malte Spielman hervorheben, der mich durch die vier Jahre meiner Promotion stets unterstützt hat, obwohl es nicht seine Verpflichtung war. Deswegen habe ich im letzten Jahr unsere Zusammenarbeit an gemeinsamen Projekten besonders genossen.

Many many thanks are devoted to the awesome research group that I was part of for so many years. It was always a lot of fun to work together knowing that you can find help everywhere. Apart of that, I enjoyed our non-work related events and the ability to find a party reason easily for instance the name day of a beloved cat!
During my time at the MPI, I had the pleasure to work in a cosy office with great people that were always open for scientific and less scientific discussions - many thanks for the time!

Besonders viel Dank geht an meine Freunde, die verhindert haben, dass ich zu sehr in der Forschung versinke: Iza, Aru, Anja und Julia mit Victor aus Berlin und meine Studienfreunde Karina und Niki, Steffi mit Philip und Christina. Vor allem meine Freunde aus der Schulzeit haben dafür gesorgt, dass ich die Welt außerhalb der Wissenschafft nicht vergesse: Karo und Kristin.
Ich will mich besonders bei Karina und Niki bedanken, die meine Arbeit korrigiert haben, bevor ich mich dem Urteil meiner Betreuer ausgesetzt habe. Many thanks to you, Aru, for improving my English and all the discussions regarding my thesis and future plans. I will never forget our exciting trip to India - thank you for showing me your country!

Der größte Dank geht an meine Familie, die es mir erst ermöglicht hat diesen Weg zu gehen. Ich danke meinen Eltern für die finanzielle aber auch emotionale Unterstützung, die mir den nötigen Mut gegeben hat, das Studium und die Promotion zu beginnen, und die Kraft, beides erfolgreich zu beenden. Ich habe mich stets von meinen Eltern, Reinald, und meinen Geschwistern Lars, Jan-Hendrik und Amelie geliebt, verstanden und unterstützt gefühlt. Dank euch konnte ich meine Ziele erreichen. Bei euch wurde ich aufgefangen, wenn ich Zweifel hatte und mir die Kraft fehlte. Meine Familie hat mir stets die Zuversicht gegeben, dass ich alles erreichen kann, und hat mich gleichzeitig gelehrt, mich und meine Ziele nicht zu wichtig zu nehmen. Durch euch weiß ich, was im Leben zählt!

Vielen Dank & Many Thanks!

Die vorliegende Arbeit wurde in der Zeit von Januar 2012 bis Mai 2016 am Max-Planck-Institut für Molekulare Genetik sowie am Institut für Medizinische und Humangenetik der Charité Universitätsmedizin unter der Leitung von Prof. Dr.Stefan Mundlos angefertigt.

Eidestatliche Erklärung

Hiermit erkläre ich an Eides statt, dass die vorliegende Dissertation in allen Teilen von mir selbständig angefertigt wurde und die benutzen Hilfsmittel und Literaturquellen vollständig angegeben worden sind. Weiter erkläre ich, dass ich nicht schon anderweitig eine Promotionsabsicht angemeldet oder ein Promotionseröffnungsverfahren beantragt habe.

Berlin, den _____

_____

Sinje Geuer

Contents

List of Figures

List of Tables

1 Introduction

For decades, geneticists have been investigating the human genome and struggle to understand the significance of the genetic code. By now, it is clear that genomic mutations can cause human disorders, some even claim that all kind of diseases are modulated by genomic predispositions [1]. The impact of genes on human health and thereby on human society evoked considerable attention and scientists supported international projects to decode the sequence and function of the human genome. Genes have been the main focus as their function is accomplished by encoded proteins, which are easy to study. But the Human Genome Project revealed that approximate 1% of the 3.2 billion nucleotides of the human genome are coding for genes [1]. The role of the large non-coding portion remained unclear for a long time and was therefore often thought to be 'junk'. By now it is clear that non-coding regions are important for gene regulation and chromosome organization [2, 3].

Considering that it was known quite early-on that genomic mutations have an impact on human health, it was surprising that each individual carries three million genomic variations including more than 70 *de-novo* mutations that are not inherited from the parents [1, 4, 5]. Apparently, most of these mutations are not harmful [6] raising the question of what is causing pathogenicity of certain mutations and how those mutations can be distinguished from the others. This is necessary to allow accurate diagnosis, risk estimation for patients and their children, as well as to determine the choice of efficient treatment [1]. Thus, genomic information and the understanding of variations and their impact is required for successful health care [7].

To improve diagnostic and the interpretation of variations, gene studies are aiming for the function of genes and their corresponding proteins. For this purpose, different animal models have been developed and the mouse proved to be especially qualified for functional studies. It resembles the main physiological human processes and encompasses high genomic similarities, while it has a short generation time and is easy and inexpensive to maintain in contrast to other animals like swine and primates [8]. Now, mice are the most widely used model for investigation of human disease [9]. The intense use of mouse models leaded to deep knowledge of its genomic sequence, improved gene annotation and the establishment of different genome editing technologies that enabled the investigation of gene function by genomic deletion or insertion.

In the recent years, it has become clear that non-coding mutations are also involved in human disorders and moved the focus towards the functional investigation of the non-coding genome and its regulative capability [10]. Nevertheless, our knowledge is still limited and does not allow us to predict the effect of non-coding variations. Therefore, we are still not able to distinguish between detrimental and non-detrimental mutations. Animal models are required to investigate the effect of single mutations on embryonic development and to elucidate their role in congenital disorders. Because of the broad knowledge, diverse technologies and similarity to human development, mouse is the animal model of choice to decipher whether genomic variations are causative for a certain disorders.

1.1 Variations in the Human Genome

In the 1000 Genomes Project (http://www.1000genomes.org) scientists from all over the world collaborated to define variations that can be found in the healthy genome. The resulting data gives an overview of non-pathogenic variations and represents therewith an important resource for medical studies and for estimation of pathogenic effects. Later it was shown that most of the 3 million variants found in each individual [1] are inherited from the parents, but each of us gains more than 70 *de-novo* mutations, of which the majority are point mutations [5]. *De-novo* mutations emerge in the germ line during meiosis but they can also arise in somatic cells in which they are manifested during replication [11, 12].
Although intense genetic studies elucidated the function of many genes, it remains often unclear if a coding mutation is deleterious [6]. Non-coding mutations are more difficult to evaluate since we are missing basic knowledge of its functionality. The international ENCODE project
(http://www.encodeproject.org) aims for a better understanding of DNA elements and revealed that many regulatory elements are located in "gene deserts". Although only 1% of the genome is coding for genes the ENCODE project uncovered that at least 80,4% encompass functionality. With the intense analysis of coding and non-coding regions, DNA modification (e.g. methylation), transcription factor binding sites and three dimensional inter- and intra chromosomal interactions ENCODE delivered immense insight into the regulative landscape [13]. This knowledge provides an important fundament to predict mutational effects.
Still clinicians and geneticists struggle to distinguish between deleterious and harmless genomic variations. The smallest alteration affects single nucleotides that can be deleted, inserted or transformed into another nucleotide species [11]. Structural variations affect longer DNA stretches which can be deleted, duplicated, inverted or translocated [14]. In general these mutations can be divided in two groups depending on their effect on either gene transcripts or gene regulation.

1.1.1 Single Nucleotide Variations

Single nucleotide variations or point mutations are alterations of single base pair that are classified in two groups: they either substitute existing single base pairs or change the number of nucleotides by deletion or insertion (see Fig.1). Point mutations can be caused by different intrinsic or extrinsic factors while the latter can be of chemical or physical nature [11]. Well known chemical factors are benzopyren (e.g. in tobacco smoke) and chemotherapeutic drugs [15]. The main physical influence is ionization radiation like UV-light which either directly damages DNA by inducing double strand breaks or creates radicals that interfere with DNA [11].
Intrinsic causes are by definition generated by cellular processes such as replication which can create nucleotide variations directly by allowing base pair mismatches. Indirect endogenous factors are for example reactive molecules and free radicals which are produced during metabolism and can interfere with DNA. Moreover, certain DNA-bases incorporate intrinsic instability. CpG islands (= stretches of Cytosin and Guanin) for instance harbor methylated Cytosine which are likely to undergo deamination and thereby transform to Uracil [11, 15].

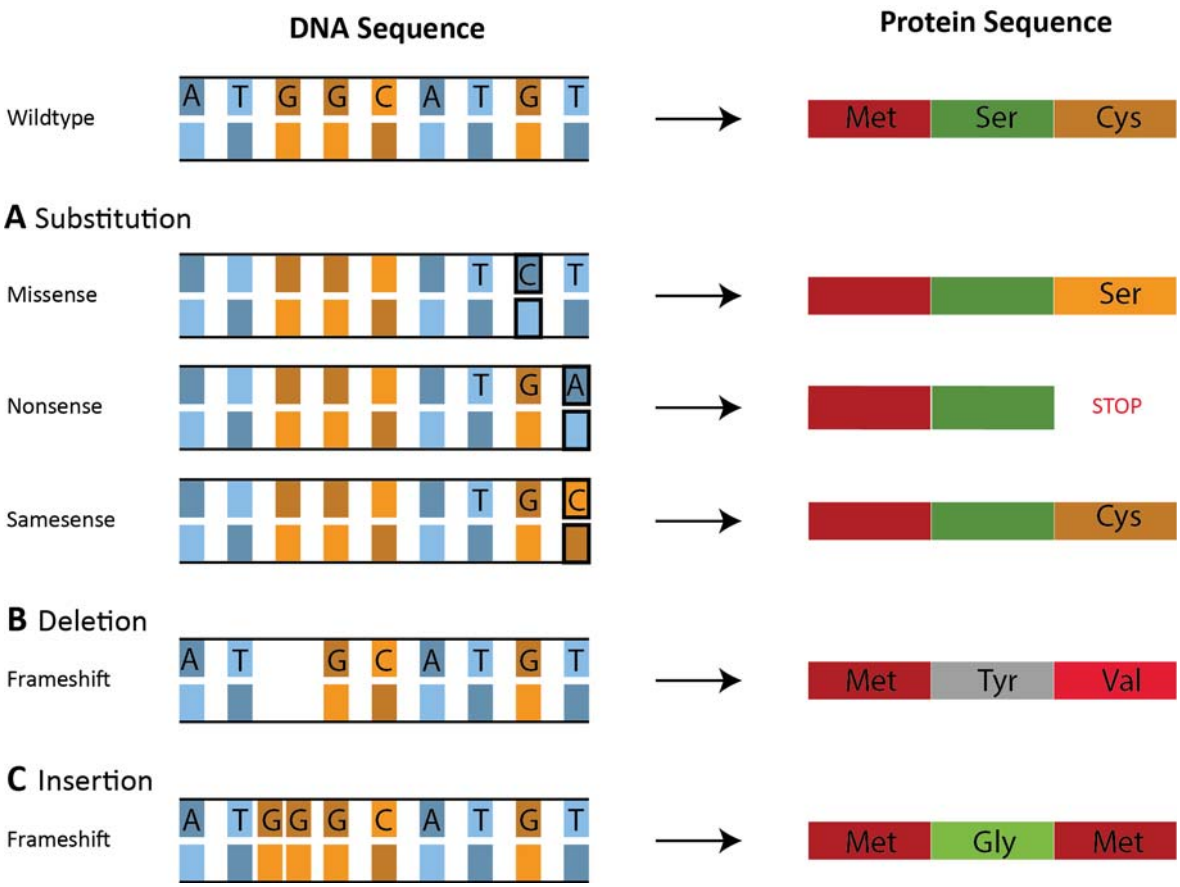

Figure 1: **Single nucleotide variations**
Single nucleotide variations are either (A) substitutions that cause missense, nonsense or samesense mutations on protein level. (B) Deletions and (C) insertions of single base pair lead to frameshift of codons and therewith change the sequence of the whole protein. Colored boxes are representing gene and its corresponding protein sequences; blue and orange boxes = different nucleotides; long boxes = different amino acids of a protein

The cellular machinery is able to repair all these mutations very efficiently but is not always accurate and might "oversee" mutations or even create some [15, 11].

Single nucleotide variations can be disease causing or without further effect depending on their localization in the genome. Variants that occur in human populations (or sub-populations) with a frequency of 1% or more are termed Single Nucleotide Polymorphisms (SNPs) and are likely to not be deleterious. Until now a huge number of SNPs have been collected and summarized in the dbSNP database of the National Center of Biothechnology Information (NCBI) [16, 17]. Point mutations can have an impact on the gene transcript by changing single codons and thereby protein sequences (see Fig.1A). These missense or non-synonymous mutations can change the functionality of a protein by changing amino acids that are crucial for biochemical function or by alteration of the protein structure. Nonetheless, missensene mutations might even be irrelevant if they do not occur at sensitive positions or if the new amino acid has similar biochemical characteristics [6]. The effect on protein level can be more extreme if a single base pair deletion results in a frame shift or if splice sites are modified [12]. While frameshift mutations result in alteration of the whole protein sequence (see Fig.1B and C) [12], splice site mutations can have different effects. They might cause transcription of

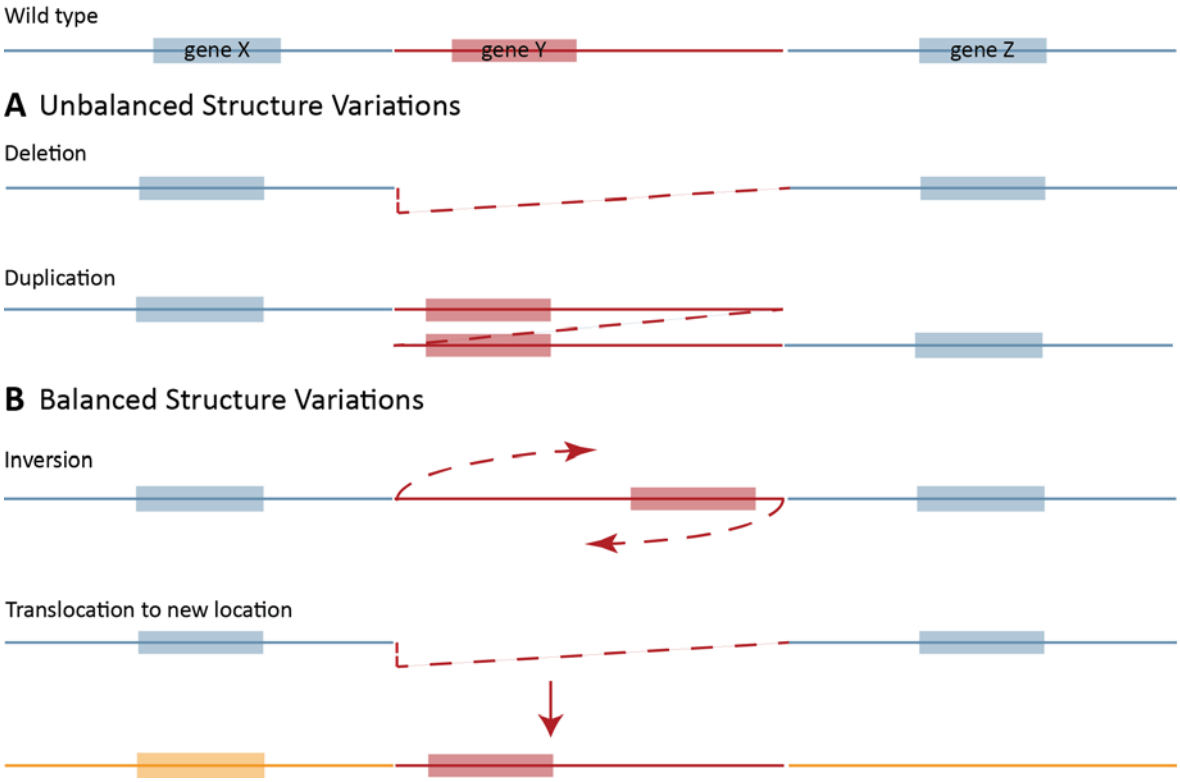

Figure 2: **Structure variations**
Structure variations like (A) deletions and duplications lead copy number changes of genomic regions while (B) inversions and translocations to other genomic regions alter the genomic order.

intron sequences, activation of new exons, down or upregulation of expression or transcript truncation due to exon skipping [18]. In general the effect of splice site mutations are hard to foresee, especially because the cellular splice machinery can switch to alternative splice sites. By now, different online tools are available to predict the effect of splice site mutations. Nevertheless, the outcome must be proved in expression experiments [18].
Finally, point mutations can cause knock-out phenotypes by creating nonsense mutations that generate additional STOP codons. Missense mutations can have a similar effect if essential amino acids are altered.
Nucleotide changes that take place in exons without alteration of the encoded amino acids - so called silent mutations - are typically thought to be not harmful. Nevertheless, they might effect genes on regulatory level if they destroy binding sites for regulative factors. Of course the same is true for all point mutations in non-coding parts of the genome [12].

1.1.2 Structure Variations

Structure variations (SVs) of the genome are rearrangements of DNA stretches greater than 1kb (and up to several megabases) that distinguish two individual genomes of one species. In general two different kinds of structure variations are known: Balanced and unbalanced SVs (see Fig.2). Unbalanced SVs (also CNVs = copy number variations) are alteration in the copy number of certain genomic regions that can be decreased by deletion or increased by duplication, triplication, insertion or translocation (see Fig.2A). Balanced SVs are changes in the order of sequences, that are either inverted or reciprocally translocated (see Fig.2B) [14].

Structure variations are induced by double strand breaks (DSBs) in the backbone of DNA [14] which can occur during different cellular processes or are caused by extrinsic factors like reactive oxygen species (ROS), ionized radiation or mechanically stress [19]. The cellular machinery allows SVs to arise by the repair of DSBs or via different mechanism like replication fork stalling (FoSTeS or MMBIR) and nonallelic homologue recombination (NAHR) (see explanation below) [20, 19, 21].

Nonallelic homologue recombination typically occurs during S- or G2-phase of the cell cycle [20] and requires long stretches of homology at different genomic positions [21]. These homologies are often low copy repeats (LCRs) [14] or retrotransposons (e.g. HERVs and LINEs) [21]. LCRs are by definition at least 1kb long, share more than 90% identity and occur in many copies within one genome. LCRs of at least 97% identity can interact via a distance of up to 10Mb and in this way cause misalignments of chromosomes or chromatids which can result in crossing over and recombination of unequal DNA fragments (=NAHR)[14]. The frequency of NAHR positively correlates with the length, proximity and identity of SVs [21] and the involved DNA fragment is flanked by LTRs [14]. Their orientation defines the type of CNV: Recombination between in the same direction orientated LTRs results in deletions or duplication while inverted orientation induces inversions. These recurrent rearrangements are characterized by two clear breakpoints mapping to the LCRs. In other cases, LCRs occur in complex mixtures with inverted and directed orientation potential leading to deletions, duplications and inversions [14].

Structure Variations can also be generated during DNA repair [14] via nonhomologous end-joining (NHEJ) [20]. The NHEJ machinery is activated to re-ligate free DNA endings if independent of homologue template and is characterized by extreme flexibility [14, 20]. It consists typically of four enzymes: First Ku binds open DNA endings and recruits further enzymes. Endonucleases then resect damaged DNA, polymerase *mu* or *lambda* fill-in DNA and ligase IV rejoins DNA strands. The flexibility of NHEJ is achieved in many ways. First, the machinery is independent from the order in which the different enzymes act and is secondly also working if one of the four enzymes is missing as long as the DNA-protein complex is stabilized by terminal microhomology of 2-6 base pairs. Third, ligase IV can be replaced by ligase I or III although it results in a lower efficiency. Last but not least both polymerases are able to use dNTP or rNTPs and are therefore independent of the metabolic stage of the cell [19]. Typically NHEJ creates small insertions and/or deletions (= indels, also called molecular scarring) at the fused endpoints [14]. The cause for that might be the order in which nucleases and polymerases act and the fact that polymerases can act as template independent [19]. In conclusion NHEJ is extremely efficient but error prone [14, 20].

Many complex structure variations are caused by errors during DNA replication called 'fork stalling and template switching' (FoSTeS) and 'microhomology mediated break induces replication' (MMBIR) [14, 21]. As the name suggests FoSTeS is characterized by stalling of the replication fork, whereby the lagging strand disconnects and anneals to another replication fork. MMBIR requires for the same process DSBs [21]. Depending on the new position of the disengaged strand deletions, duplications or inversions occur. FoSTeS is discussed to be the

major mechanism for the formation of CNVs [14], especially for unbalanced translocations [21]. It is the most common cause of nonrecurrent structure variations which typically show multiple break points and a mixture of different SV between them [14].

Structural Variations (SVs) are part of the healthy genome and occur even more often than point mutations [14, 22], nevertheless, they are also known for human diseases when they affect dosage sensitive genes or regulatory elements [14, 23] causing genomic disorders [14] or cancer [20]. The DECIPHER database gives an overview about published healthy and disease causing CNVs [23] and helps in this way to interpret new cases. The pathomechanism of SVs leading to health impairment is important for further treatment and consulting but is often not obvious and difficult to interpret [10, 14, 24].
Diagnostic needs to interpret structure variations in order to decipher potential pathogenicity. CNVs are easy to appraise if they change gene dosage: Deletions of genes can lead to knock-down or knock-out phenotypes while duplications might cause upregulation. Intragenic SVs are more complex, since they might have deletional effects but are also likely to create so called chimeric genes (e.g. fusion gens, truncated gens or rearrangement of exons) that might gain unexpected functionality [20]. In this case the effect of such a gain-of-function mutation needs to be validated in a model e.g. mouse.
Even if genes are not directly involved in the event of structural variations, they might be influenced by positional effects. This term summarizes the observation that genes disrupted from their initial genomic localization show different behavior leading to the conclusion that the noncoding part of the genome is influencing gene function. Until now, it is very well known that the landscape of genes regulates their spatial and temporal expression pattern. Different regulating elements are defined: enhancers promote gene expression while silencers inhibit it. Insulators influence the gene expression by preventing interactions between certain promotors and regulative elements [10]. The latter is very important since it was shown that enhancers can interact with target genes far away [25, 26]. Positional effects can trigger expression changes by separating genes from regulative elements and/or translocating them in the neighborhood of new regulative elements. Additional, long range effects has been proposed in which the chromatin status alters gene expression for instance by moving a gene from an euchromatin to a heterochromatin region which provokes silencing of the affected gene [27].
In the last years a new model has been introduced that focuses on the genomic 3D-structure that is thought to determine which element is allowed to act on a gene [10]. The basic idea is that the genome is divided into topological associated domains (TADs) which is a region of frequent interchromosomal interaction. These TADs have been detected by different technologies e.g. Chromosome Conformation Capture methods (3C, 4C, 5C) and High throughput 3C (HiC) [28, 29] showing that frequent interaction is restricted to a certain genomic region (= TAD) that can differ in size but is defined for each locus. It is thought that TADs arise from DNA-loops that are confined by border structures (also boundaries) that limit the activity of regulative elements (see Fig.3A) [10, 26]. TADs are highly conserved between species and different cell types [26] but can be changed by SV if boundaries are deleted, duplicated or

inverted. This can lead to pathogenic alterations by different mechanisms [30, 10]: Removal of boundaries allows ectopic activity of regulative elements called enhancer adoption (see Fig.3B). Tandem duplications across boundaries can lead to enhancer adoption because it enables interaction between the two copies(see Fig.3C). Inversion of boundaries causes reverse order of regulated elements which again can lead to enhancer adoption (see Fig.3D). In this way all structure variations might cause enhancer adoption of genes at breakpoints, which can trigger gain-of-function effects by misregulation of these genes [31]. On the other hand, alteration of the TAD structure can also remain without further effect. The reason for this is unclear but it is possible that the lack of required transcription factors that trigger enhancer-promotor interaction or too extreme spatial distances prevent enhancer adoption in certain cases [32].
The distance between enhancers and its target genes might also plays a role within a TAD. Duplications, deletions or inversions might move enhancer closer or further away from its target gene and thereby cause knock down or ectopic upregulation of genes [32].
Taken together it is clear that variations in non-coding regulative parts of the genome harbor the capability to induce genomic disorders. To elucidate what is the effect of a structure variation, information about the involved regulative elements and their restriction by borders or insulator are required. Available HiC-maps give an overview of intragenomic interacting regions and can help to find potential TADs [33, 34]. The VISTA enhancer browser lists expression pattern driven by certain enhancer [35] which shows in which tissue an enhancer might be active in vivo. Nevertheless, the outcome of TAD changes on transcriptionel level can be not predicted until now.

1.1.3 Detection of Genomic Variation in Human

The detection of genomic variation improved a lot in the last decades and many heritable diseases and their genetic causes are known. If mutations are known to cause a specific disorder the suspected locus is investigated with target specific methods before genome wide technologies are applied. Point mutations are sequenced with the Sanger approach while structure variations can be investigated via different molecular genetic methods like quantitative PCRs and multiplex ligation dependent probe amplification (MLPA) [14]. MLPA holds the advantages that up to 50 loci can be investigate simultaneously [36]. Nevertheless, translocation can only be detected with cytogenetic approaches for example fluorescence in situ hybridization (FISH) [14, 24]. Here fluorescence labeled DNA-probes are hybridized to metaphase or interphase chromosomes [21] and reveal the location of specific genomic regions with a resolution between 50kb-2Mb [36].
For genomic wide studies different techniques are available by now. Karyotyping and G-banding is the most historical method that allows detection of structural aberration of 5-10 Mb. Array comparative genome hybridization (aCGH or arrayCGH) revolutionized CNV screening. It allows genome wide comparison of the copy number of certain regions to a reference genome via hybridization to small DNA-oligos. The resolution depends on the genomic distance between these oligos [36]. While diagnostic arrayCGH focuses on known disease associated loci, research arrayCGHs are covering the whole genome with a resolution of up to

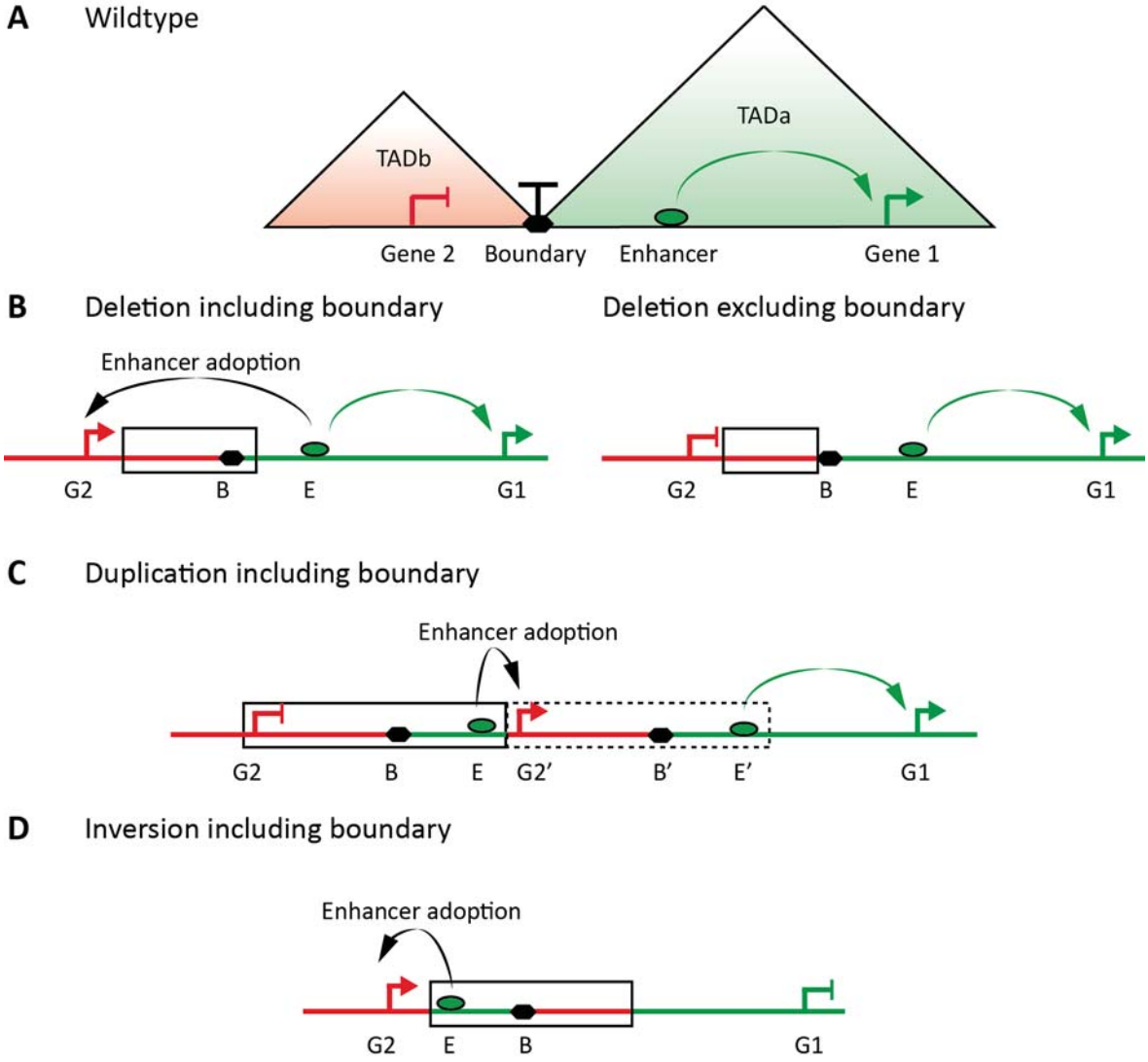

Figure 3: **The Model of Topological Associated Domains (TADs)**
(A) TADs are region of intrachromosomal interaction that are separated from each other by boundaries. Enhancer (green) and silencer (red) can exclusively act on genes within a TAD. (B) Deletion of boundaries facilitate pathogenic interaction between regulative elements and genes (=enhancer adoption). Thus similar deletions that do not involve boundaries are not causing misexpression. The same is true for (C) duplications and (D) inversions that cross boundaries. Graphic is adapted from Lupiáñez et al. (2015) [31].

1kb [14]. For a long time, genetics was lacking an appropriate method for genome wide detection of point mutations. The only approach had been single nucleotide arrays (SNParray) which are able to detect CNVs as well as specific SNPs [36]. Besides this SNParrays are also used in linkage analysis which is aiming for loci that segregate with the patient phenotype in consanguineous families [37, 38].

The development of next generation sequencing (NGS) finally enabled exome and genome wide detection of all kinds of genomic variants [14]. In general and especially for the detection of point mutations sufficient coverage of minimal 20 sequencing reads is demanded to exclude artefacts and to facilitate 95% coverage of coding regions [39]. Balanced structure variations are uncovered by pair-end sequencing that amplifies break point reads which map to two different loci in the control genome [14, 20, 36]. Unbalanced SVs are a greater challenge since many reads are needed to compare the read numbers between genomes especially because of

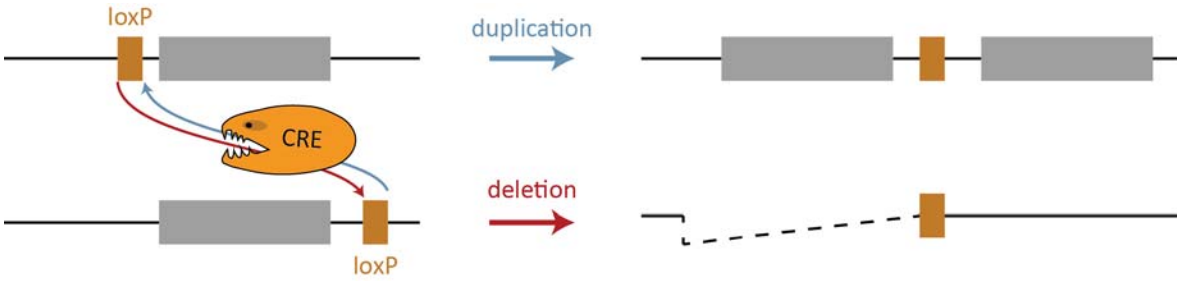

Figure 4: **Generation of in-tandem duplications with the loxP-Cre system**

sequence dependent amplification efficiencies [20].

Although we are capable of identifying genomic alterations the interpretation is often challenging. Since NGS enabled genome wide detection of all kind of mutations the need for precise and fast verification of pathogenicity is immense. The most accurate way to elucidate the impact on human health is the recapitulation of human mutations in mouse models, which require efficient genome editing technologies.

1.2 Genome Editing

Genomic modified organisms are important models for the investigation of gene function and the pathogenicity of human mutations [40]. Mice are favoured because of their short generation time and the low maintenance costs. Additionally, they resemble the main physiological human processes and development. Due to the high genomic similarities to humans and the good annotation of genes, mice are the most widely used model for the investigation of human disease [8, 9].
Genome editing has been strongly improved in the last decades and enabled intense analysis of cellular pathways and investigation of human mutations including determination of their pathogenicity. Most recently, genome editing was the key to the discovery of functional abilities in the non-coding parts of the genome [41, 42]. For these purposes deletion of genomic sequences or knock-out of genes and overexpression models has been proved to be a straight forward approach [41]. Nevertheless, investigation of genomic disorders or cancer requires often more specific models for point mutations and complex structure variations e.g. duplications, translocations and inversions [20, 14, 22].

The first genomic modification methods were imprecise and dependent on the treatment of animals with mutagen substances or radiation that induced random modification. This harbors the complication of identifying the mutation [43]. Targeted genome editing was facilitated in 1980s with the discovery of homologue recombination based mutagenesis which is since then the standard approach for the generation of mouse models [44]. Commonly, the desired mutation is generated in murine embryonic stem cells (ESC) which are injected into either blastocyst or zygotes in a process called aggregation. Later the embryos are transferred into foster animals which give birth to chimeric mutants [40, 45, 46].
Homologous recombination is facilitated by the cellular machinery that allows pairing of ho-

mologous sequences (of sister chromosomes or artificial constructs) during DNA replication. The proximity of homologous elements facilitates DNA polymerase to switch template and enables the integration of sequences beyond the homologue area [22]. This mechanism can be used to integrate constructs after cells transfection [43]. Recombination constructs typically consist of two 2-6kb long homologous segments (= homologue arms) that encase the to-be inserted DNA [47]. In a similar way deletions can be generated by using constructs with discontinuous homologue arms that are lacking the to be deleted region. Homologous recombination encompass several limitations. First, the size of deletion and insertion is restricted to a few kilobases. Secondly, inversions can be only created by integration and have thereby the same size limitations. Furthermore, it is impossible to create tandem duplications, because in this case the insert is homologous to the target region and therefore will form duplexes with the homologous genomic region resulting in replacement of the element instead of its duplication [48].
Another way for genomic integration is the flip-in System. Here the recombinase flipase (flp-recombinase from *Saccharomyces cervisiae*) is used to induce recombination at recognition sites called *frt*-sites. One *frt*-sites is required in the host genome and another one in the insert-construct. In the present of flipase recombination between both *frt*-sites leads to the insertion of the whole construct [49, 50]. Flipase is typically expressed from a second plasmid which is transfected together with the construct [51]. Usage of the flip-in system is limited since genome modification is initially needed to insert a *frt*-site in the host genome. This effort is reasonable if different kinds of insertions are needed. By now, a few murine embryonic cell lines are commercial available that harbor *frt*-sites at loci that are known to be unperturbed by insertions (e.g. *ColA1* and *Rosa26*) [50, 52]. With these pre-muted cells, gene dosage effects can be studied as long as the expression pattern is maintained, while positional effects of tandem duplication cannot be investigated.

To verify the pathogencity of human structure variation it is necessary to create tandem duplications and large SVs. Therefore another approach has been developed called TAMERE (trans-allelic targeted meiotic recombination) [53] or STRING (sequential targeted recombination-induced genomic) [54]. Both technologies use the viral recombinase Cre that recognizes a specific motif called *loxP* and induces recombination at these sites. If two *loxP* sites flanking a region Cre-mediated recombinases induces the deletion of this area. If the two *loxP* sites are located on different alleles at nonidentical positions the adjacent element will be duplicated due to translocation from one to the other allele. Consequently, the region is deleted on the other allele (see Fig.4).
The TAMERE and STRING approach is extremely time consuming since three different genomic insertions are required that has to be generated via homologue recombination: two *loxP* sites at different positions and alleles and one Cre-ORF (open reading frame). The different genotypes have to be united in one genome. Therefore, typically three transgenic animal lines are created which are then mated until the desired genotype is achieved [54, 53]. In result both, TAMERE and STRING technology, is time consuming and costly, because of its high cloning effort and the amount of required mating steps.

Another challenge for integration via homologue recombination is its low efficiency of 10^{-6} - 10^{-7} (prior to selection) [47]. The integration rate can be increased by introducing double strand breaks (DSB) [55], which are repaired by homology directed repair (HDR) [47] during G2- and S-phase [20]. The cellular HDR machinery depends on homologue DNA-strand pairing (e.g. sister chromatids) that stabilizes open DNA-endings and provides replication template bridging the gap. In most cases homologous regions are not in close proximity and DSBs are repaired by another mechanism called non-homologous end-joining (NHEJ) (see Fig.5) [19, 21, 22]. As described in 1.1.2 NHEJ is less accurate than homology directed repair (HDR) and might be used to create knock-out by frameshifts that occur by chance [42]. On the other hand NHEJ occurs more often than HDR since it is not restricted to a certain cell stage. Consequently both machineries are competing with each other and thereby reduce the efficiency of introducing DNA via DSB and HDR. Nonetheless, it is more efficient than homologous recombination without DSB [47].
DSBs can be introduced at the genomic target site with sequence specific endonucleases for instance transcription activator-like effector nucleases (TALEN) and Zinc-finger nucleases (ZFN) [43, 47]. ZFNs encompass sequence specificity for 3bp-sequences which are encoded in the key amino acids of the "finger chain". Adaption of sequence specificity is achieved by changes in the key amino acids and consequently requires complex design and cloning steps. Further pre-tests are necessary since sequence specificity is difficult to predict. The general low DNA-binding specificity and affinity is another limitation [43].
TALENs are artificial fusion proteins consisting of a DNA-binding domain originating from TALE (a unique transcription factor from the bacteria *Xanthomomas*) and the endonuclease domain of IIS-FokI. Sequence specificity is achieved by amino acid repeats of 33-35 residues that recognize single base pairs. TALEN proteins encompass a higher sequence specificity than ZNFs but require an immense cloning effort [43].
Recently a new endonuclease system has been developed that unites simple design and cloning, low cost and time effort with high efficiency and sequence specificity: CRISPR/Cas9 [43, 42, 56, 57, 58].

1.3 A New Generation of Genome Editing: The CRISPR/Cas9 System

1.3.1 The Origin of CRISPR/Cas9

The Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 System is part of the adaptive immune system of bacteria and archea and enables them to detect and eliminate foreign DNA. These two functions are clearly separated. While the endonuclease Cas9 degrades DNA, the crRNA guides the enzyme to foreign DNA by Watson-Crick base pairing to the so called protospacer sequence of extraneous sequences. Consequently alteration of the crRNA (also spacer) sequence leads to immune system adaption. Besides binding of the crRNA to the protospacer sequence Cas9 requires a protospacer adjacent motif (PAM) that ensures that only foreign DNA is targeted. The PAM sequence typically consists of three base pairs which are specific for different CRISPR/Cas9 systems and are therefore the only limitation [60].

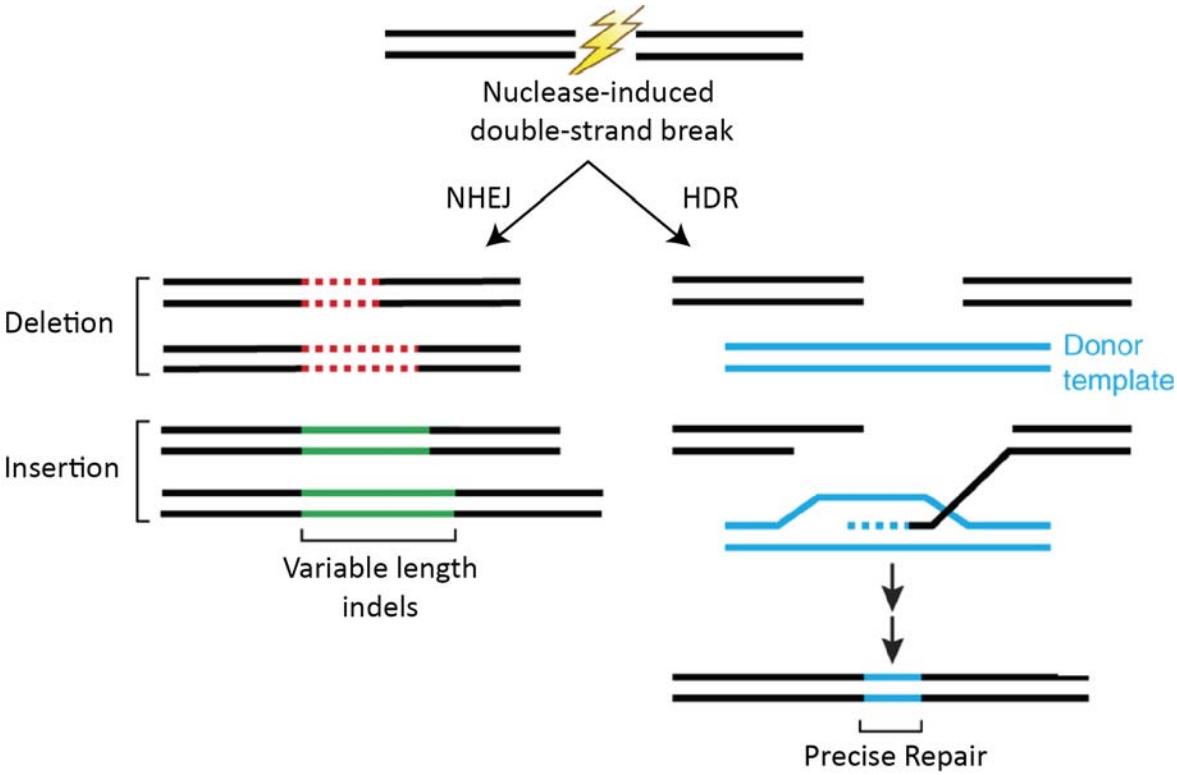

Figure 5: **Repair of double strand breaks by Nonhomologous end joining (NHEJ) and homology directed repair (HDR)**
NHEJ is an extreme efficient but error prone repair mechanism that creates indels at break point while HDR uses homologous template to replicate beyond a nick and therewith allows accurate filling of the gap. Figure is adapted from Sander et al. (2014) [59]

The CRISPR/Cas9 systems consist of three genes - one encodes endonuclease Cas9 and the others two RNA: the tranactivating crRNA (tacrRNA) and a precursor-crRNA (pre-crRNA) [56]. The maturation of the pre-crRNA is induced by a complex duplex structure with tacrRNA, that recruits RNAseII [61, 62]. By incorporation of the mature crRNA in Cas9 the assembling of the final effector complex is completed.

1.3.2 The Tool CRISPR/Cas9

CRISPR/Cas9 was discovered to be the new opportunity to target any sequence in any genome with a nuclease. While targeting with TALEN and ZFN is extremely complex and time consuming in design and production CRISPR/Cas9 is a simple tool that only needs the design of a primer like sequence [43, 41].
At the moment the CRISPR/Cas9 system from the human pathogen *Streptococcus pyogenes* is most widely used [43]. It was adapted to research needs by simplifying it to a two-component system in which only one RNA (singleguideRNA = sgRNA) is needed (see Fig.6). This single hybrid hairpin RNA already provides a tetraloop structure to load Cas9 with itself and thereby does not require RNAse processing as described above [64]. The sgRNA of 20bp can be easily designed with the only requirement that the PAM sequence is present at the 3-prime end of the protospacer [43]. The corresponding DNA-oligo is artificially produced

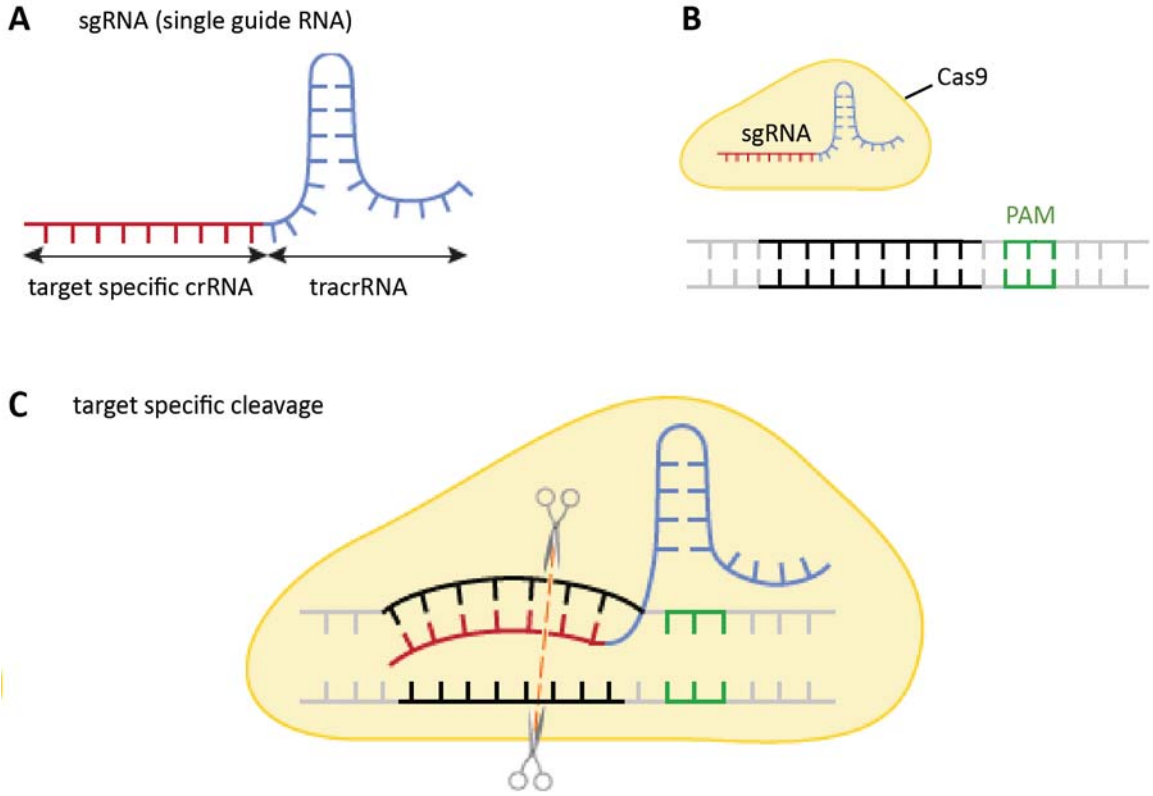

Figure 6: **The components of the CRIPSR/Cas9 System**
The CRIPSR/Cas9 System consist of two components - single guide (sgRNA) and Cas9 nuclease: (A) sgRNA consist of a target specific crRNA sequence and tracrRNA (B) that facilitate incorporation in Cas9. (C) Cas9 is guided by the sgRNA to a specific locus and cuts double stranded DNA when a PAM sequence is present (C). Graphic is adapted from www.clontech.com [63].

like any primer and cloned into an expression vector e.g. pX459.
CRISPR/Cas9 is already used in many different genomes from plants, animals and fungi [65]. The most common application are cell culture systems, where it can be easily introduced via transfection of a plasmid coding for Cas9 and the sgRNA [43]. Nevertheless, this expression vector can also be used for in vitro transcription of crRNA and Cas9, which can be injected into zygotes or blastocysts. This approach has been shown to be a straight forward way to generate transgenic animals e.g. in mouse (*Mus musculus*) [41], fly (*Drosophila melanogaster*) [65] and zebrafish (*Danio rerio*) [66].

As already discussed in 1.2 introducing of DSB allows genome editing via two different DNA repair mechanism - NHEJ and HDR. NHEJ introduces random indels at break points (see Fig.5) and can lead to frameshifts, missense or nonsense mutations. CRISPR/Cas9 provides us with a more accurate way. Since we can target almost any sequence we are able to design two guides flanking the gene of interest and remove it completely. It has been shown that the use of two guides is sufficient to introduce deletions of several kilobases and up to a few megabases. In addition, also inversions of the flanked region can be induced by the same approach [56]. The later could be especially interesting for the investigation of chromosomal rearrangements. Notably all these modifications are generated via the NHEJ pathway and therefore encompass indels at the breakpoints.

Aside from chromosomal rearrangement CRISPR/Cas9 also helps to introduce small DNA oligos of a few base paires (e.g. tags for protein pull down, recombination sites like *loxP* or defined point mutations) via HDR [43]. This has been proven in human cell cultures [64] as well as in mice stem cells and by blastocyst and zygote injection [41]. For *in vitro* and *in vivo* applications single stranded DNA-oligos of maximum 200bp are introduced that carry exogenous DNA-elements in the middle flanked by stretches of target-homology. This oligo serves as a template for the homology directed repair mechanism. Therefore, less indels are expected but also the efficiency is lower, since NHEJ is the favored repair mechanism in most cells [43]. A new approach to trigger the HDR pathway is to prevent NHEJ with the ligase IV and III-inhibitor Scr7. The application of Scr7 results in double insertion rates after CRIPSR/Cas9-transfection of human cell lines as well as after zygotic injections [67]

Independent of the desired modification it has to be taken in account that CRISPR/Cas9 is known for off-target effects, which are caused by a unspecificity of Cas9 regarding the binding ability of its guideRNA. Up to four nucleotide mismatches are tolerated at the 5-prime site of the sgRNA [68]. This might be a useful evolutionary tool for the adaption property of CRISPR/Cas9 [43], but is of course problematic for the use in research and beyond this. Right now different guideRNA design tools are available online that try to predict potential off-targets. Variable results of these tools suggest that the nature of CRISPR/Cas9 off-targets is poorly understood. Hence other solutions have been developed to overcome this challenge. For animal lines, the issue can be handled by mating with wildtype animals which will quickly separate the target locus from off-targets by segregation (if they are physically not to close). For cell systems or if no back crossing is needed, it is wise to generate controls by either using different guides for the same mutagenesis or investigating separate clones from the same experiment [43].

In conclusion CRISPR/Cas9 is a genome editing tool that allows rapid mutagenesis. Therewith it is a perfectly suited for recapitulation of human mutation to verify their pathogenicity and to elucidate the underlying pathomechanism.

1.4 The Human Limb as a Model for Mutation Verification

The skeleton consist of several hundred bones with distinct shape and size that are placed at a specific position [69]. This defined composition allows accurate measurement of alterations and comparison of malformations between human and mouse.
The skeleton fulfills three main function: It protects inner organs, is a reservoir for blood and minerals and allows muscle driven movements at joints, which are kept together by ligaments while muscle contraction is transmitted via tendons [69, 70]. Notwithstanding the high variability in vertebrate skeleton patterning, the overlaying structure and its development is extremely similar between species [71]. The main factors controlling skeletogenesis are well known regulators of either local (e.g. SHH, FGFs and WNTs) or systemic nature (e.g. growth and sex hormones) [69]. The focus of this work lies on bone development and pattern formation in limbs because it is easy to investigate and the define structure allows precise

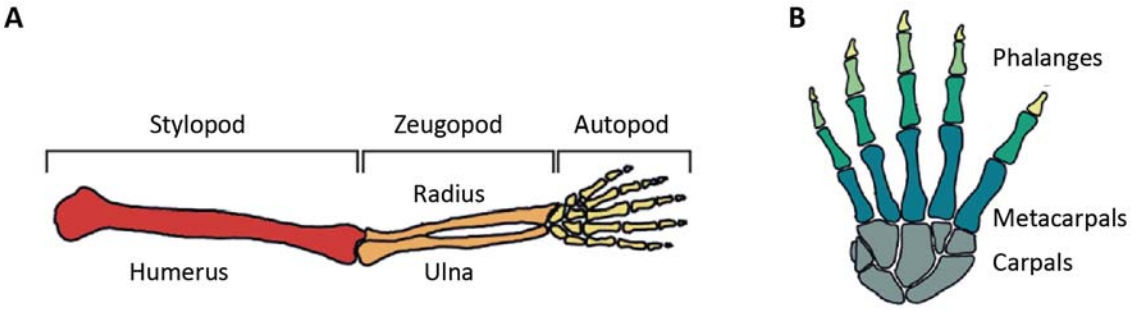

Figure 7: **Elements of the human limb skeleton**
The human limb skeleton is divided into three main elements: (A) Stylopod (humerus), Zeugopod (radius and ulna) and autopod (hand). (B) The hand consists of complex bone arrangements. The hand plate is formed by carpal and metacarpal bones. Finger identity is characterized by position, size and number of phalanges. Graphic is adapted from Hellmann et al. (2012) [74].

detection of alterations. Additionally, most limb malformation are compatible with life and thereby allows the generation of mouse models.

1.4.1 Skeletal Patterning

The patterning of the limb skeleton is defined for the three body axis: anterior-posterior, dorsal-ventral and proximal-distal. The proximal-distal patterning of the limb is characterized by three main elements: The stylopod (humerus/femur) is followed by zeugopod (ulna-radius/tibia-fibula) and autopod (hand/feet) (see Fig.7). On the anterior side radius, tibia and digit 1 (thumb) are located and posterior ulna, fibular and digit 5 (small finger) [71, 72, 73]. The dorsoventral axis is defined by the morphology of single bones and the presence of nails at the dorsal site [73].

In general limb patterning is determined by gradients of secreted morphogens that are expressed by organizers located in mesenchyme and ectoderm of the growing limb [71]. Limb development is initiated by the formation of the limb bud that consist of mesenchymal cells that grow out from the lateral plate mesoderm [75, 76] and induces the establishment of the apical ectodermal ridge (AER). The AER is located at the distal tip of the limb bud separating dorsal and ventral side (see Fig.8A) [73]. It drives further limb bud outgrowth [77] as well as proximal-distal patterning via different *Fibroblast Growth Factor* (FGF4, 8, 9, 17) [73, 75, 71]. It is thought that the proximodistal axis identity is already determined in mesoderm before outgrowth [78] and mesenchymal cells are kept in undifferentiated stage at the progress zone (PZ) by AER-FGF signaling [71]. Hyper-proliferation of cells causes outgrowth of the limb bud and increases the distance between proximal mesenchyme and AER. Thereby FGF-influence on proximal cells is reduced and allows differentiation of cell into chondrocyte fated ones marked by *SOX9* (*SRY-box9*) expression (see 1.4.2) [72].
The posterior-anterior patterning is guided by *Sonic Hedgehog* (*SHH*) expression in the zone of polarizing activity (ZPA) in the posterior limb bud mesenchyme (see Fig.8A) [71]. SHH is a morphogen with a high diffusion capability and creates a gradient that is fading towards the anterior site [79] which is tightly regulated by cholesterol and palmitoyl modification of

SHH [71]. The gradient is essential for the expression of anterior-posterior specific genes (e.g. 5-prime *HOXD* genes) and the formation of correct digit number and identity [71]. The importance of correct SHH localization was early shown in transplantation experiments in chicken, where anterior *Shh* expression induces additional digits. On the other hand knock out of *Shh* in mice leads to loss of anterior-posterior identity and zeugopod with only one single undefined bone digit [71, 75, 73].
The dorsal-ventral axis is determined in the mesenchyme before limb bud initiation and later transferred to the ectoderm by unknown molecular mechanisms [73]. In the limb bud the patterning is driven by three key genes. The dorsal ectoderm marker WNT7a (ectoderm) activates mesenchymal LMX1B (*LIM Homeobox Transcriptionfactor containing 1-beta*) expression but is inhibited on the ventral site by the morphogen BMP (Bone Morphogenetic Protein) driven *Engrailed* (*EN1*) expression [73, 75].

Although it is easier to look at axis patterning separately, a more precise picture is drawn by an integrative approach which gives insight into the regulation network across single signaling cascades (see Fig.8C). As already mentioned AER-produced FGF is crucial for growth and patterning along the proximal-distal axis. Beside this, it also promotes *SHH* expression in the posterior mesenchyme and in this way influences anterior-posterior patterning. Vice versa, SHH maintains *FGF* expression in the AER and is supported in this by BMP and EN1 from the ventral ectoderm. Lasty, dorsal WNT7a is another positive regulator of *SHH* expression [73].
The tight regulation between ZPA and AER signals is important for the onset and termination of limb development. This epithelial-mesenchymal (e-m) feedback loop (see Fig.8B) is initiated together with limb bud formation by *BMP* expression which is required for AER formation. BMP promotes the expression of its own antagonist *Gremlin1* (*GREM1*). In the following propagation phase GREM1 stimulates *FGF* expression in the AER through BMP inhibition. FGF drives upregulation of *SHH* which again leads to increased *GREM1* expression. Limb outgrowth is terminating this auto-enhancing loop: The distance between *SHH* expression cells and GREM1 positive cells is increasing and thereby decreasing SHH influence on *GREM1* expression. As a result *GREM1* is down- and BMP upregulated. This causes again inhibition of *FGF* expression in the AER that in consequence loses its functionality and limb bud growth is completed [71, 72].

1.4.2 From Mesenchyme to Bone

The skeleton of the limb is formed by a process called endochondral ossification that relays on a cartilaginous scaffold that is later replaced by bone (see Fig.9) [76, 80]. Mesenchymal cells expressing the transcription factor *SOX9* start to condense and differentiate into chondrocytes under the control of *SOX5* and *SOX6* [76] and form cartilage and respectively perichondrium at the outline [69]. Chondrocytes undergo three phases which are characterized by changes in morphology and composition of secreted extracellular matrix (ECM): proliferation, differentiation and apoptosis [81, 82]. At the distal site of the forming bone proliferating round chondrocyte give rise to elliptic forms of chondrocytes that proliferate fur-

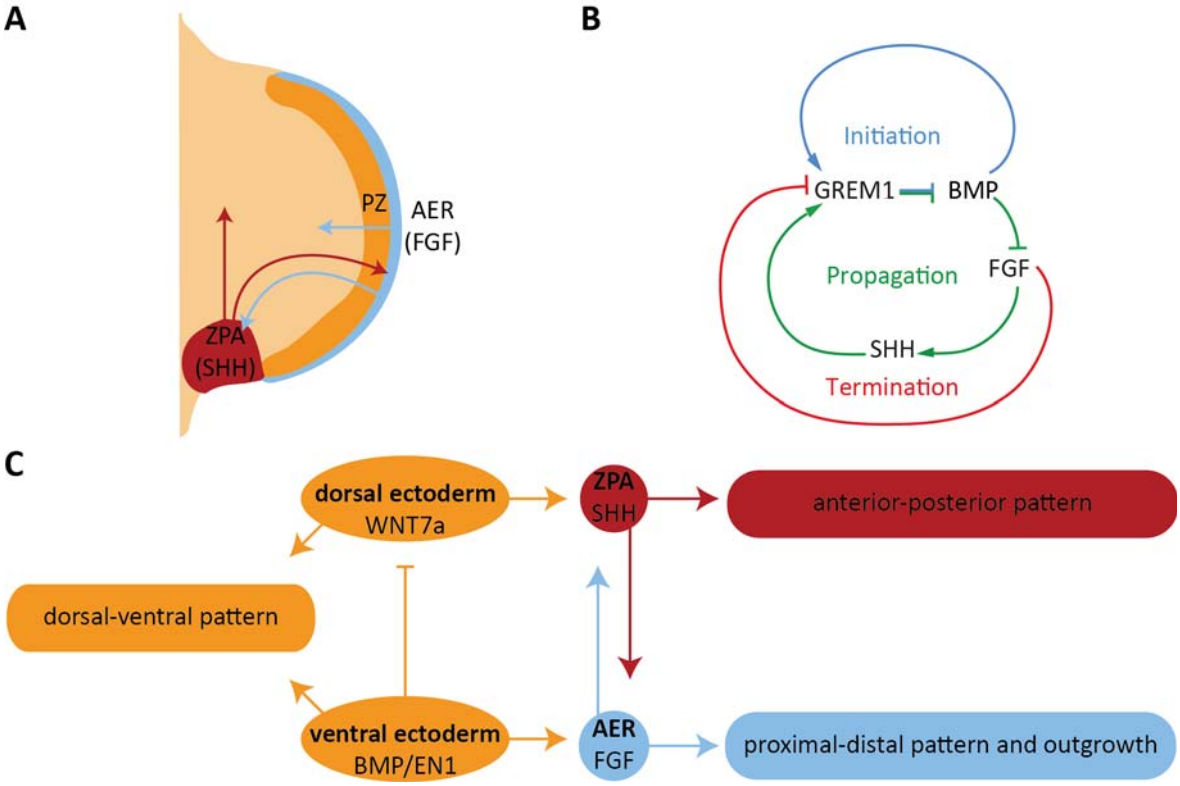

Figure 8: **Singling networks in the developing limb bud**
(A) Positioning of the two main organizer ZPA and AER and their major morphogens *SHH* and *FGFs*. (B) The epithelial-mesenchymal feedback loop consists of three phases: initiation, propagation and termination. Termination is caused by distance between *SHH* and *GREM1* expressing cells and leads to the breakdown of AER. (C) Signaling networks of the three spatial axis are interlinked. Graphic (A) and (C) are adapted from Niswander (2003) [73] and graphic (B) from Bénatzet and Zeller (2009) [71].

ther to form column like stacks. These stacks direct longitudinal growth [69]. In the center, the chondrocytes become hypertrophic by increasing their dry mass and fluid uptake [83]. This process is the main driver of embryonic bone growth besides chondrocyte proliferation [69, 83]. Hypertrophic chondrocytes guide mineralization in many ways. Adjacent perichondrial cells are promoted to become osteoblasts that produced collar matrix [69], which is the precursor of cortical bone [82]. Hypertrophic chondrocytes secrete the vessel growth factor VEGF that induces vessel invasion and enables thereby replacement of cartilage by trabecular bone [82]. With the blood chondroclast, osteoclast and osteoblast are migrating into the cartilage tissue and induce bone remodeling in which matrix is calcified (osteoblast) and cartilage and mature bone are degraded (chondroclast and osteoclast). Hereby hypertrophic chondrocytes fulfilled their task and undergo apoptosis. Bone remodeling is essential for bone formation because it allows adjustment of bone architecture to environmental changes during growth that causes various mechanical and metabolic challenges [69, 76].

The process of ossification takes place sequentially during outgrowth. When the appropriate size is reached chondrocytes stop proliferating and a second ossification center is formed at the epiphysis (bone endings). Consequently chondrogenic tissue is placed between two bony areas - epiphysis (at endings) and diaphysis (in center) - and form two growth plates at both sides. In this growth plate round chondrocytes are arrested [76] and serve as a pool for

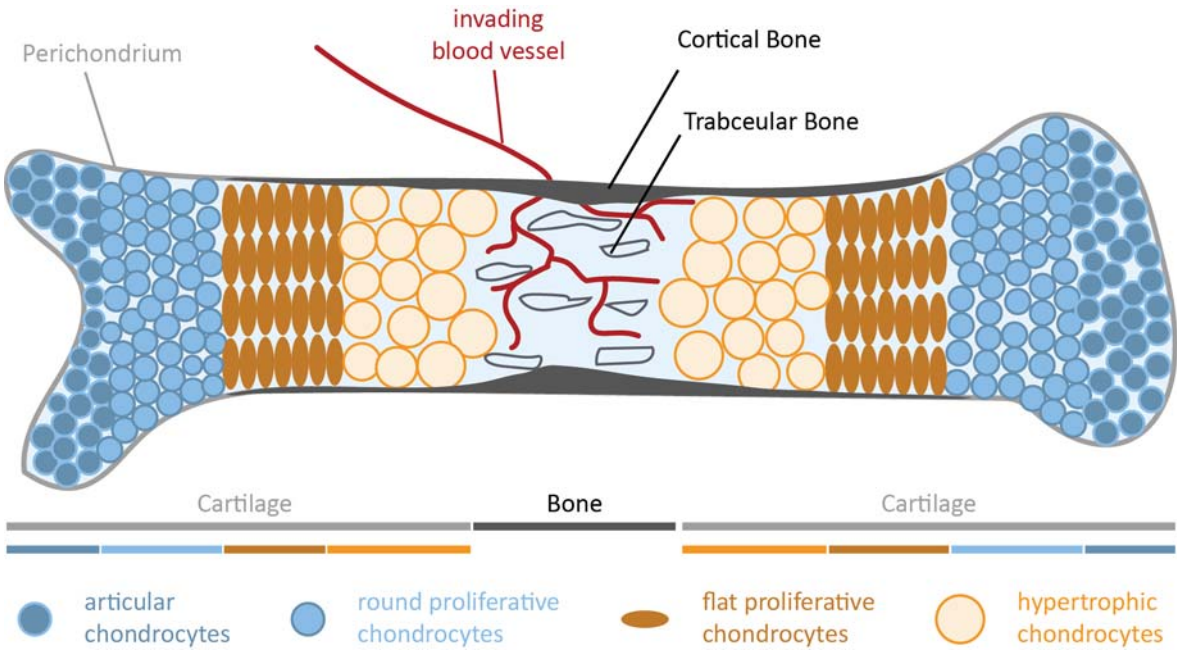

Figure 9: **Bone development**
Bone development is driven by complex regulation of round chondrocyte proliferation (blue) that gives rise to columnar chondrocytes (orange) that proliferate further until they become hypertrophic and induce formation of bone collar and vessel invasion before they undergo apoptosis. Osteoclast and chondroblast migrate via blood stream and cause bone remodeling and generation of trabecular bone. Graphic is adapted from Provot and Schipani (2005) [82].

proliferating columnar chondrocytes [69] which drives bone growth during childhood and is finally terminated with sexual maturation [81, 76].

1.4.3 T-Box Genes in Skeletal Limb Development

The *T-box* (*TBX*) gene family consists of several genes coding for transcription factors featuring the DNA-binding domain T-box [84]. This domain was first found in *Brachyury* (*T*) and later in many other genes, that were named after this feature. *TBX* genes derived from *Brachyury* by duplication and evolved different functions by structural or regulative mutations. Since they are known to be crucial for embryogenesis in metazoa it is not surprising that they are highly conserved between species [85]. Six *TBX* genes are known to be involved in limb development, because of their similar function they are often referred to in pairs: *TBX2/3*, *TBX4/5*, *TBX15/18* [85, 84].
TBX15 and *18* are expressed in the limb bud mesenchyme but also at the regions of skull and vertebrate formation. Homozygous knockout of murine TBX18 is lethal due to severe malformation of vertebral columns but no limb phenotype was observed [85]. The role of *TBX15* was revealed by Singh et al. (2005) [84]. *TBX15* is expressed at the posterior-distal axis from E9.0 to E16.5 in mouse embryos. *TBX15* is highly upregulated in the central part of the limb but excluded from the flanks where future digit I and V are initiated. Although the expression pattern implies a role in limb morphology, no interference with the known pattern marker (see 1.4.1) could be detected. Interestingly *TBX15* is expressed around condensating mesenchyme cells and at developing joints suggesting a function in chondrogenesis and osteogenesis. Indeed the phenotype of *TBX15*-deficient mice is supporting this interpretation.

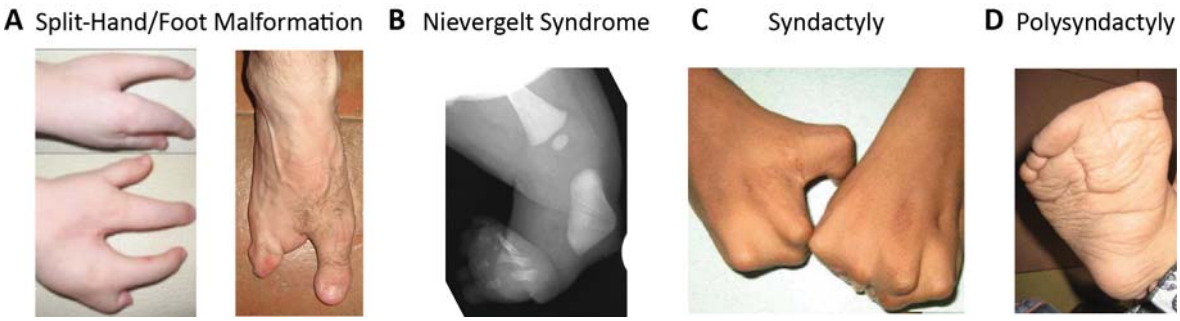

Figure 10: **Overview of limb phenotypes**
(A) Split-Hand/Foot Malformation (SHFM), (B) Nievergelt Syndrome (C) Syndactyly andd (D) Synpolydactyly with cup-shaped appearance

Knock-out-animals show a general retardation in size but also various alterations of bone morphology and delayed endochondral ossogenesis. This phenotype is caused by reduced proliferation of mesenchymal precursors and chondrocytes leading to a reduced number of pre-hypertrohic chondrocytes and thereby to general growth retardation and developmental delay. Absent bone matrix (for instance in scapula) can be explained by a reduced number of mesenchymal precursor cells that does not allow initiation of condensation and is also known to cause misshaped bones [84]. Homozygous *TBX15* loss-of-function was found to produce a similar phenotype in the human Cousin syndrome [86, 87]. Nevertheless, the limb phenotype deriving from each *TBX15* and *18* deletion is surprisingly mild in comparison to other *TBX* deficiencies. This might be explained by functional redundancy which is implied by identical expression pattern and 93% sequence homology, but has not be demonstrated yet [85, 84].

1.5 Limb Disorders

Genetic disorders of the skeleton are a heterogeneous group of malformations that differ in their clinical appearance, inheritance mode and genetic origin. They occur isolated or in syndromes and range from mild forms with growth retardation or shortening of fingers to deformations impairing mobility. Severe forms can even lead to neonatal lethality. Skeletal disorders are of significant importance because of their relatively high frequency although each type is rather rare. They are classified depending on the defect that affects either bone shape (dysostoses) or chondrogenic/osteogenic development (osteochondrodysplasia) [76].

1.5.1 Synpoldactyly

Synpoldactyly describes the manifestation of two distinguishable phenotypes: Polydactyly (additional digits) and syndactyly (fusion of digits) (see Fig.10C). Both are the most frequent congenital limb malformations [88, 89] with an incidence of 2:1.000 for polydactyly [88] and 3-10:10.000 for syndactyly [89]. Snydactyly is a very heterogeneous malformation in regards to it's clinic appearance, inheritance mode and genetic cause. Fusion of digits can be either cutaneous or osseous and is therefore likely to involve two different mechanisms. Osseous syndactyly is caused by irregular bone growth during early limb development. Cutaneous syndactyly reflects a fail of digit separation by interdigital apoptosis in late development, which is driven by the BMP pathway. Interestingly, inhibition of BMP signaling results in

Table 1: **Classification of split-hand/foot malformation (SHFM)**
SHFM is classified in seven groups regarding their associated genomic loci. Table is adapted from Sowinska et al. (2013) and extended with data from Tayebi et al. (2014) [95, 96]

SHFM Classification	Locus and Mutation	Inheritance Pattern
SHFM1	7q21.3-q22.1	autosomal dominant
	DLX5	autosomal recessive
	DYNC1l1 exon 15 and 17	autosomal dominant
SHFM2	Xq26	X-linked recessive
SHFM3	10q24 duplications	autosomal dominant
SHFM4	*TP63* mutation	autosomal dominant
SHFM5	2q31 deletions	autosomal recessive
SHFM6	*WNT10B*	autosomal recessive
SHFM/SHFLD	17p13.3 duplications	autosomal dominant

Polysyndactyly [89].
Polydactyly can arise either at anterior (preaxial polydactyly) or posterior side (postaxial polydactyly) of the limb [88]. In severe cases mirror-images have been reported, where all fingers/toes are duplicated [90]. More than 310 disorders with polydactyly are known of which most are syndromic (94%). All kinds of modes inheritance have been described but for many cases this still remains unresolved. This is due to the high number of disorders with unknown genetic causes or involvement of unknown genes. Nevertheless, around 100 genes have been identified that play a role in polydactyly [91].
Polydacytly is thought to result from defects in the anterior-posterior patterning [91]. Therefore, it is not surprising that *SHH* is one of the best known genes involved. It was earlier shown that ectopic anterior expression of *Shh* induces preaxial polydactyly in chicken [92], which is also true in mouse and human. This additional "anterior ZPA" disturbs the GLI3A/GLI3R gradient (see 1.4.1) and thereby transforms the thumb to triphalangeal morphology or produces additional triphalangeal digits [88]. In many polydactyly cases mutations in the *SHH*-regulating region ZRS had been found. This region lies in the intronic region of *LMBR1* and point-mutations in this area have been shown to cause misregulation of *SHH* and polydactyly [88]. Interestingly ectopic *Shh* expression can additionally cause syndactyly through its mitogenic properties. It was shown that prolonged interdigital *Shh* expression in chicken inhibits interdigital apoptosis and results in syndactyly via soft tissue-fusion [93]. The same mechanism has been proposed for human inversion of *SHH* [94] and duplication for ZRS [90]. Both results in severe synpolydactyly with fusion of all digits. This phenotype is often referred to as cup-shaped hands/feet (see 10D).

1.5.2 Split-Hand/Foot Malformation

Split-hand/foot malformation (SHFM) - also known as ectrodactyly - is a form of hypoplasia effecting predominantly the central rays of hands and/or feet (see 10A) [95, 97]. It occurs in one of 18.000 live borns and accounts for 8-17% of all limb malformation [95]. Clinical features encompass a high inter- and intrafamilial variability as well as different manifestations in each limb of a single patient. It can range from mild deformation as phalangeal hypoplasia and syndactyly, to aplasia of central digits or even monodactyly. The frequent occurrence of characteristic cleft like appearance hand and feet provokes the naming Split-hand/foot malformation [95]. SHFM is known as an inherited limb malformation but was also shown to arise from exogenous factors like retionic acid, cadmium, ethanol, caffeine and cocaine in rat [97]. Different mutations associated with SHFM lead to a classification in seven groups (see Tab.1), but the genetic pathomechanisms often remained unclear. Autosomal dominant inheritance with reduced penetrance is the most common inheritance pattern, while autosomal recessive mode is rare and X-linked SHFM was reported for just one case [95]. Noteworthy is that SHFM occurs isolated but is also found in more than 50 syndromes for instance Gollop-Wolfgang complex (GWC) [95, 97, 98].
SHFM arises in many species including mouse, in which human SHFM-mutations had been generated e.g. by homozygouse deltion of *Tp63* and *Lrp6* or *Dlx5/6*. Interestingly the *Dactylaplasia* mouse - which evolved spontaneously by mutation of *Dactylin* - shows incomplete penetrance as many human cases. Here *Mdc* was shown to be the modifier gene allowing the development of the SHFM phenotype in the presence of *Dactylin* mutations. Mouse models also revealed that developmental late disturbance of AER integrity is a possible molecular pathomechanism for SHFM and also occurs when pattern networks of other axis are involved, demonstrating the high interconnectivity of these (see also 1.4.2) [97].

A very rare subform of SHFM with additional hypoplasia of long bones (mainly tibia) is termed SHFM with long bone deficiency (SHFLD) and occurs with the incidence of 1:1.000.000 mostly with autosomal-dominant inheritance with reduces penetrance [95]. Four susceptibility loci are known for SHFLD (1q42.2-q43, 6q14.1, 2q14.2, 17p13.3) [99] of which duplications (and rarely triplication [100]) at 17p13.3 is the most common cause [95, 98, 101]. It accounts for 30% of all SHFLD cases and furthermore for 12% of all SHFM cases (with unaffected long bones) showing the close relationship but also variable expressivity of both malformations [95]. The only gene in the overlapping region of known duplications is *Basic Helix Loop Helix Protein A 9* (*BHLHA9*). The concrete function of *BHLHA9* in limb development is unclear [101, 102] but *BHLH* genes build a family of transcription factors that are known to influence cell proliferation and differentiation, and play a role in different developmental processes [103]. Since *Bhlha9* is expressed in the murine progress zone it is thought that increased BHLHA9 doses leads to disturbance of signaling between AER and PZ and causes thereby hypoplasia [101]. Interestingly, the opposite - *Bhlha9* knockdown - causes reduction of the pectoral fin in zebrafish [101], while in mouse homozygous *Bhlha9* deficiency produces the syndactyly phenotype because of reduced interdigital apoptosis [104]. Lately *BHLHA9* deletion was as well detected in a patient with syndactyly [104].

Table 2: **Non-skeletal anomalies in a case of Nievergelt-like disorder**
Data was summarized from Steichen-Gersdorf et al. (2008) [108]

Nievergelt-like Disorder - non skeltal anomalies
Neuological
Abnormal EEG
Myclonic jerks
Anpnoeic Spells
atrophy of cortical
atrophy of subcortical brain
dilateration ofsubarachnoid space
dilateration of ventricular system
Other Organs
Malrotation of colon
Enlarged clitorus and urogenital sinus
Horseshoe kidney

1.5.3 Radius Deficiency

Radius deficiency is a rare congenital limb malformation and is estimated to occur with an incidence of 1:9.000 - 1:55.000 [105]. The clinical manifestation ranges from hypoplasia of distal radius portions (type I) to severe forms with radius aplasia (type IV). In extreme cases the humerus is also affected as well (type V). Most of the time thumbs are also absent and carpal bones might be hypoplastic, too. Shortening and bowing of the ulna can occur in severe cases. Typically, radii hypoplasia/aplasia causes problems in the formation of adjacent joints leading to radial angulation of the wrist and absent or permanent elbow flexion [106]. Because of the limited number of cases the genetic pathogenicity is often unclear. Until now, mutations in *TBX5* (see also 1.4.3) and *SALL4* have been associated with syndromic radius hypoplasia/aplasia [76]. Recently Jahmsheer et al. (2013) linked a duplication on 5q35.2-5q35.3 to radius deficiency and hypothesized an involvement of *MSX2* and/or *FGFR4* misregulation [107]. The latter is especially interesting since shortening of radii had been found in SHFLD cases [99] suggesting a role of the AER-PZ network particularly FGF-signaling (see also 1.5.2).

1.5.4 Mesomelic Dysplasia and Nievergelt Syndrome

Mesomelic Dysplasia describes the shortening of limbs at the zeugopods in contrast to rhizomelic (stylopod) and acromelic (autopod) dysplasia [70]. The Nievergelt syndrome is a very rare type of mesomelic dysplasia and was characterized by the physician Nievergelt in 1944 as a deformation of bones in the lower leg to triangular shape (see 10B) [109]. The diagnosis is often thought to be quite easy because of the specific triangular deformation of long bones which is not found in other syndromes [110]. On the other hand until now less than 10 unrelated cases have been reported and all differ in their appearance regarding skeletal alterations as well as involvement of other organs, developmental delay and neurological dysfunctions, which were not described by Nievergelt [109, 111, 110, 112, 113, 114].

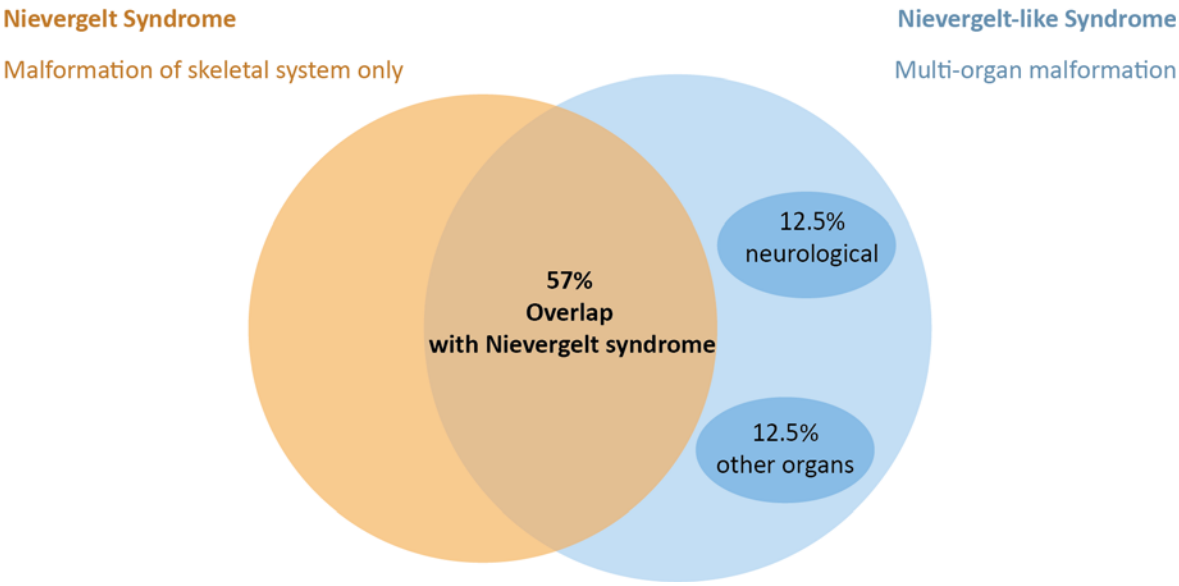

Figure 11: **Comparison of Nievergelt syndrome features and abnormalities in our patient**
The phenotype of the patient with Nievergelt-like disorder overlaps in 57% with the one of Nievergelt syndrome.

This issue was discussed in more recent cases and nicely summarized by Monga et al. (2003) who concluded, "it would be prudent to clump all these disorders as phenotypic variants of mesomelic dysplasia Nievergelt type [...]" [114]. Savarirayan et al.(2000) already reported a case of mesomelic dyplasia with triangular tibia that he preferred not to be classified as Nievergelt syndrome since his patient had no fibulae, which are supposed to be triangular in Nievergelt syndrome [112]. In the recent nosology of genetic skeletal disorders this case was classified as 'mesomelic dysplasia, Savarirayan type' [115] and is associated with duplications on chromosome 6p22.3 [116]. The genetic cause for the Nievergelt syndrome is still unknown [115]

In this work, a case of Nievergelt-like syndrome is studied, that is similar to Savarirayan type mesomelic dysplasia but includes further phenotypes (see Tab.2) like complex urogenitalia malformations e.g. intersex external genitalia. Internal genitalia as well as karyotyping revealed a female identity. Additionally, the patient suffered from severe neurological disfunctions that became apparent at the age of three weeks, when myoclonic jerks occur and changes in electroencephalogram (EEC) were detected. Later magnetic resonance imaging (MRI) revealed dilateration of the ventricular system and subarachnoid space and atrophy of the cortical and subcortical brain. By the age of three months the patient developed apnoeic spells with increasing frequency, which in the end lead to the death of the girl by four months of age [108]. Because of the phenotypical differences this case is not thought to be part of the Savarirayan type mesomelic dysplasia [112, 108]. Instead a "Nievergelt-like spectrum disorder[s]" was proposed since the patient phenotypes overlaps in 57% with the features of Nievergelt syndrome (see Fig.11) [108].
BAC-based arrayCGH was performed on the patient described above and a micro-deletion on chromosome 2q11.2 of around 500kb was detected (see Fig.12A) that encompasses only one single gene: *LAF4/AFF3*. The mutation has been proved to be *denovo* by SNP analysis.

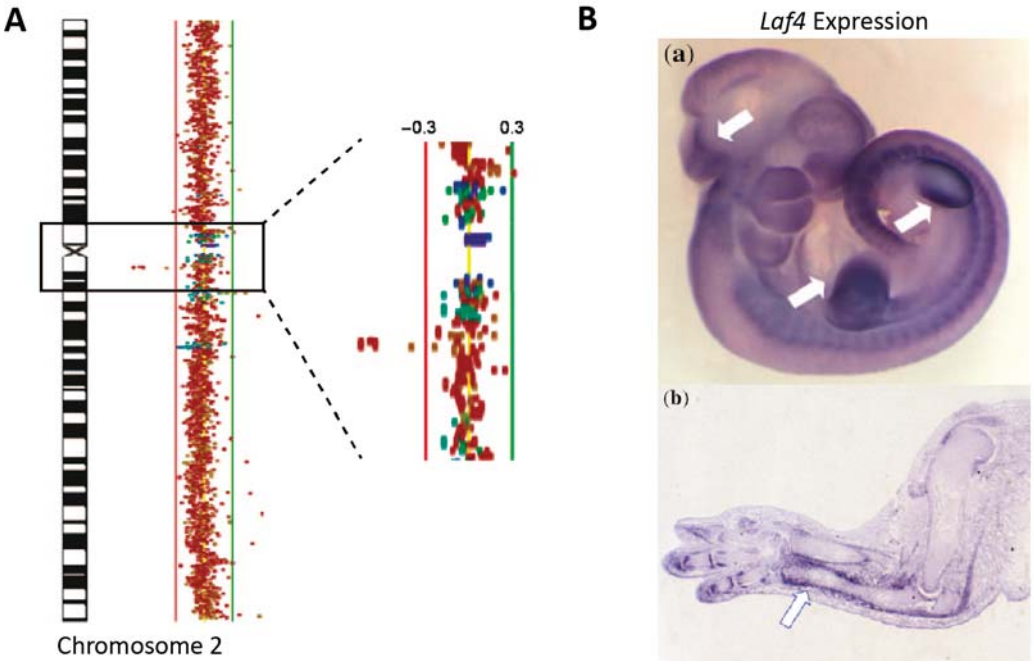

Figure 12: **Genotype in Nievergelt-like disorder**
(A)BAC-Array CGH revealed LAF4 deletion in our patient. (B-a) This gene is expressed in the hindbrain and limb bud mesenchyme in E11.5 mouse embryo and (B-b) in the perichondrium and joints of the developing limb at E15.5. Graphic is adapted from Steichen-Gersdorf et al. (2008) [108]

Since the second *LAF4* allele was unchanged, it was concluded that the complex phenotype arose from haploinsufficiency for *LAF4* with autosomal dominant inheritance mode [108].

LAF4/AFF3 is part of the *AFF* family that consists in total of four genes. Their common ancestor in *Drosophila melanogaster Lilliputian* plays a role in cell growth, differentiation, dorso-ventral patterning and wing development [117, 118], but the exact function of *LAF4* still remains unknown [119]. It is a putative transcription factor and was shown to bind direct or indirect to DNA and to encompass transcription activating activity *in vitro* [120]. Because of its expression during development in brain *LAF4* is expected to be an important developmental regulator. Beside this it is also detectable in lung, spleen, thymus and cell lines of the immune system [120] and in the developing limb (see Fig.12B) [108]. LAF4 locates subcellularly in nucleus speckle - a structure arising from splice machinery. Melko et al. (2011) showed that LAF4 and the other AFF proteins play a crucial role in alternative splicing [119]. They also detected *in silico* nuclear and nucleolar localization signals (NLS and NoLS) that are highly conserved between mouse and human as well as within *AFF* family [119].

LAF4 was discovered to play a role in other diseases. For instance it forms fusion proteins with MLL (Mixed-Lineage-Leukemia) causing acute lymphoblastic leukemia (ALL). In most of the ten fusion versions of MLL the transcription regulating domain is replaced by the fusion partner and is thought to cause thereby misregulation of various genes [119]. Furthermore *LAF4* has been identified as a susceptible locus for rheumatic arthritis [121]. Because of its close relation to *FRM2*, which plays a role in intellectual disability [122], and its expression during early cortical development in brain [123], it is likely that *LAF4* mutations can also

cause neurological dysfunction as in our patient [108].

1.6 Aim of this Study

Genomic variations arise in every generation and are therefore extremely common and variable. They are primarily classified in two groups: structural rearrangements and single nucleotide polymorphisms. Both types have the potential to cause hereditary disease or cancer but can just as well be completely inconsequential. Human geneticists struggle to distinguish between both the outcomes. The effect of variations within genes are often easier to predict as long as the affected gene and its functionality is known. Non-coding parts, on the other hand, are crucial for gene regulation but they continue to be poorly understood. This lack of knowledge makes the interpretation of variations in this part of the genome almost impossible. Structure variations are a more serious challenge, since they can also effect gene expression in adjacent areas. Therefore, animal models are required to elucidate the pathogenic potential of certain mutations. However, but the traditional genome editing technologies do not allow introduction of human mutations with required precision, efficiency and speed.

The skeleton and in particular the limb is a heterogeneous structure with distinct shape, size and position of each bone. Therewith, it is an ideal research model to investigate pathogenic mechanism of genomic variations since the morphological affects are easily recognizable. Since most limb malformations are not lethal it is possible to create a model organism that verifies the pathogenicity. The mouse is suitable model organism due to its high genomic and skeletal similarities that allows precise comparison of human and murine phenotypes. In this study the new genome editing technology CRISPR/Cas9 was used to recapitulate human variation in mouse which are thought to cause limb malformations because it allows rapid generation of mutations.

The aim of this work is to investigate if CRISPR/Cas9 allows to recapitulate specific human structural (SVs) and single nucleotide variations (SNVs) in murine ES cells and if these ES cells can be used to create mice with those mutations. It should be elucidate how efficient the generation of SVs and SNVs is. The ability to phenocopy human congenital disease in mouse will be illuminated to clarify whether the mouse is a suitable model to verify pathogenicity of mutations in human.
To answer these key questions, five mutations that were found in human patient cases of different skeletal limb malformations were recapitulated in mouse:

1. Nievergelt-like syndrome with deletion at *LAF4* locus
2. SHFM/SHFLD with duplication of *BHLHA9* locus
3. Synpolydactyly with duplication of *SHH* locus
4. Radial Aplasia with duplication at *TBX15* locus
5. SHFM with a pointmutation within *FNDC3A*

2 Material

2.1 Instruments

Table 3: Centrifuges

Name	Supplier
Chilling Centrifuge 5417R	Eppendorf
Chilling Centrifuge Sorvall RC-5	TermoElectron
Microtiter Centrifuge 5415 D	Eppendorf
Microtiter plate centrifuge 5461	Eppendorf

Table 4: Microscops

Name	Supplier
Camera AxioCam HRc	Zeiss
Camera AxioCam MRm	Zeiss
Light source KL1500 LCD	Leica
LSM700	Zeiss
Microscope Axiovert 200M	Zeiss
Microscope DMR	Leica
Stereo microscope MZ6	Leica
Stereo microscope MZ7-5	Leica

Table 5: Thermo Cycler

Name	Supplier
ABIPrism HT 7900 Real-time Cycler	Applied Biosystems
GeneAmp PCR System 2700	Applied Biosystems
GeneAmp PCR System 2720	Applied Biosystems
GeneAmp PCR System 9700	Applied Biosystems

Table 6: Other Instruments

Name	Supplier
BioRobot M48 workstation	Qiagen
Bioruptor UCD-300	Diagenode
Imaging system Curix 60	Agfa
Microtome 2050 Supercut	Reichert-Jung
Microtome Cool Cut HM355S	Microm
Microtome RM2255	Leica
Nanodrop	Thermo Scientific
Semi-dry blotting chamber	BioRad
LAS 4000 Imaging System	Fuji

2.2 Chemicals

Table 7: Chemicals

Chemical	Supplier
2-Methoxyethylacetet (MEA)	Polysciences
Alcian Blue	Sigma
Alizarin Red	Sigma
BMP Purple AP Substrate	Roche
Boehringer Blocking Reagent(BBR)	Roche
Blocking Reagent	Perkin Elmer
CDPStar	Roche
Enhanced Chemo Luminescence	Rotilumin, Car Roth
Fluoromount G	Affymetrix
Formamide p.a.	Merck
NP 40 Nonident P40	Fluka 74385
Proteinase K	Boehringer 1000144
Protein Ladder, pre-stained	Fermentas
Silver Nitrate	Roth
Protease Inhibitor Cocktail Complete	Roche
QuickExctract DNA Extraction Solution	Quiagen
RNAse type III from bakers yeast	Sigma R 6750
Tetramisole-Hydrocloride	Sigma L9756

2.3 Buffers and Solutions

2.3.1 Biochemical Methods

Table 8: Buffers for Southern Blot

Buffers	Ingredients
10xDIG	in ddH_2O: 1M maleic acid, 1,5M NaCl (pH7,5)
1M NaPi	in ddH_2O: 1MNa_2HPO_4 (pH7,2; autoclaved)
10xSSC	in ddH_2O: 50% 20xSSC
50mM	in ddH_2O: 5% 1M NaPi
Church-Hybridisation buffer	in ddH_2O: 0,5M NaPi, 7% SDS, 4µM EDTA
Church-Wasching buffer	in ddH_2O: 40mM NaPi 1% SDS
Denaturation solution	in ddH_2O: 0,5M NaOH, 1,5M NaCl
DIG2	in DIG1 10% BBR (autoclaved)
DIG1	in ddH_2O: 10% 10xDIG1
DIG3	in ddH_2O: 0,1M Tris (pH9,5), 0,1M NaCl
Neutralization solution	0,5M Tris-HCl (pH7,5) 1,4M NaCl

Table 9: Buffers for Western Blot and Protein Extraction

Buffers	Ingredients
1x Transfer Buffer	in ddH_2O: 25mM Tris, 192mM Glycine, 20% Metahnol
4x Protein loading buffer	in 600 ml ddH_2O: 2ml 1M Tris, 4ml Glycerine, 2ml 20% SDS, 400µl 1% Bromphenol Blue, 1ml Beta-Mercaptoethanol; pH7.5
5x Running Buffer	in ddH_2O: 25mM Tris, 250 mM Glycine, 0,1% SDS
Blocking Solution	in TBST: 5% milk powder
Membrane lysis buffer	in ddH_2O: 50mM NaF, 30 mM Nappi, 5mM EDTA, 1% Triton-X 100,1xTBS

2.3.2 Histology

Table 10: Buffers for Immunohistochemistry

Buffers	Ingredients
Blocking Solution	in PBS: 3% horse serum, 0,3% BSA, 0,1% TritonX-100
PBSX	in PBS: 0,02% TritonX-100
TSA	in PBS: 10% horse serum, 0.5% Blocking Reagent, 0,1% Triton-100

Table 11: MMA Embedding

Buffers	Ingredients
Infiltration Solution	in MMA: 10% polyethylene glycerol, 0,33% benzoyl peroxide
Polymerization Solution	in MMA: 10% polyethylene glycerol, 0,55% benzoyl peroxide, 0,05% N,N-dimethyl-p-toluidine

Table 12: Buffer for Whole Mount RNA In-Situ

Buffers and Solutions	Ingredients
Alkaline Phosphatase Buffer	12ml 5M NaCl, 30ml 1M $MgCl_2$, 6ml 10% TWEEN 20, 60ml 1M Tris (pH9,5), 300mg Tetramisole, ad 600ml bidest. water
Citric Acid	1M Citric Acid Monohydrate in DEPC-H_2O
Final Fixation Buffer	50ml 4% PFA/PBS, 400µl Glutaraldehyde(25%), 500µl 0,5M EDTA(pH8,0)
First Fixation Buffer	50ml 4% PFA/PBS, 400µl Glutaraldehyde(25%), 500µl 10% TWEEN20
Heparin	100mg/ml 4xSSX-DEPC
Hybe-Buffer	25ml Formamide, 12,5ml 20xSSC-DEPC(pH4,5), 25µl Heparin, 100µl 10% TWEEN 20, ad 50ml DEPC-H_2O
5xMABT	100ml 1M Maleic Acid (pH7,5), 30ml 5M NaCl, 10ml 10% TWEEN 20, ad 200ml with bidest. water
Maleic Acid(pH7,5)	1M Maleic Acid in bidest water, adjust pH with NaOH, autoclaved
PBST-DEPC	in DEPC-H_2O: 10% 10xPBS, 1% 10%-TWEEN
PBSTGlycine	2mg Glycine add 1ml PBST-DEPC
PBSTTetramisole	500mg Tetramisole-Hydrochloride in 1l PBST-DEPC
RIPA	2,5ml SDS(10%), 15ml 5M NaCl-DEPC, 5ml NP40, 2,5g DEoxycholate, 1ml 0,5M EDTA-DEPC, 25ml 1M Tris(pH8,0), ad 500ml DEPC-H_2O
RNase A	50ml RNase solution, 500µl RNase stock (10mg/ml)
RNase solution	5ml 5M NaCl, 500µl 1M Tris(pH7,5), 500µl 10 % TWEEN 20, ad 50ml bidest water
tRNA	10mg RNA type III add 1ml DEPC-H_2O
20xSSC-DEPC(pH7,0)	175,3g NaCl, 88,2g Sodium Citrate add 1000ml DEPC-H_2O, adjust pH with HCl
20xSSC-DEPC(pH4,5)	20xSSC-DEPC(pH7,0), adjust pH with 1M Citric Acid
SSCFAT	100ml 20xSSC, 500ml Foramide, 10ml 10% TWEEN 20, ad 1l bidest water

2.3.3 Other Buffers

Table 13: Other Buffers

Buffers	Ingredients
10xPBS	80g NaCl, 2g KCl, 14,4g Na_2HPO_4, 2,4g KH_2PO_4 ad 1l H_2O
4% PFA/PBS	40gPFA ad 1l 1xPBS
4% PFA/PBS-DEPC	40gPFA ad 1l 1xPBS-DEPC
Alcian Blue Solution	150mg/l Alcian Blue, 80% EtOH, 20% acetic acid
Alizarin Red Solution	50mg/l Alizarin Red in 1%KOH
Cell-Lysisbuffer	10mM TrisHCl, 10mM EDTA, 0,2%SDS, 100mM NaCL, 10µl/ml ProteinaseK
DEPC-H_2O	0,1% DEPC in bidest water
DNA Loading Buffer	6%Sucrose, 0,7% Orange G, dissolved in Bidest
GC-PCR Buffer	1,7ml 1M SO_4, 6ml Tris/HCl (pH8,8), 0,2ml $MgCl_2$, 70µl Beta-Mercapto-Ethanol, 2, 03 ml bidest. water
GC-PCR Mix	250µl 5mM dNTPs, 250µl DMSO,
Proteinase K	20mg in 1ml Proteinase K Buffer
Proteinase K Buffer	1ml 1M Tris (pH7.0), 0,1ml 0,5M EDTA, add 50ml DEPC-H_2O
TissueLysisbuffer	17mM Tris(pH7.5), 17mM EDTA, 170mM NaCl, 0,85% SDS, 3µl/ml ProteinaseK

2.4 Kits

Table 14: Kits

Name	Supplier
BigDye Terminator v3.1 Sequencing Kit	Applied Biosystems
BCA Protein Assay Kit	Pierce
Myoalert Detection Kit	Lonza
NucleoSpin Plasmid	Macherey-Nagel
Nucleobond PC100	Macherey-Nagel
NucleoSpin Gel and PCR Clean-up	Macherey-Nagel
PCR DIG Probe Labeling Kit	Roche
RNAeasy-Kit	Qiagen
SYBR Green qPCR Master Mix	Qiagen
TaqMan Reverse Transcription Reagents	Roche/Applied Biosystems

2.5 DNA-Oligos

All DNA-Oligos are listed in 5'prime to 3'prime orientation. Primers and CRISPR guides were purchased as HPSF purified products from Eurofins MWG (Ebersberg).

2.5.1 Genotyping-Primer

Table 15: Genotyping Primer

Line	Sequence Forward	Sequence Reverse	Specificity
Bhlha9 5/3	caaaggcactcccaacttgt	tccaccctatccagctgttc	5'BR
Bhlha9 5/3	ctgtgtacagcccggatagg	gttgctctccaccgtctctc	3'BR
Bhlha9 5n1	ccgggagcagagttctagg	gtccagctatgcctgttggt	5'BR
Bhlha9 5n2	gaacaggcaaggagcatagc	aatgagctcctggctgtagg	5'BR
Bhlha9 5n1/2	gttcggacctccacacagtt	agagtcaagccactcccaga	3'BR
Bhlha9-ColA1	ggtgggaaggaatttctggt	gatggtttagcagggggatt	transgene
Bhlha9-ColA1	ggggttcttggactgttgaa	ggtcctgtcctttctggtga	wild type
Fndc3a	tccaaccccaaaccaaaata	ctcaggcgagctacaggttc	*Fndc3a* exon 26
Hprt-Cre	tgctgtttcactggttgtgcggcg	tgccttctctacacctgcggtgct	*Cre* Recombinase
KO-Laf4	aaagtggctgggaaagaggga	ctgcatgtggtatcgttgcac	knockout
KO-Laf4	aggcagaggtgagttcattgt	aagtgaagtccatacggcctt	wildtype
Laf4-trunc	cccaagcttgttgactccat	gacctggaacagacccttga	5'BR
Laf4-trunc	ggacgaaccttgttctgagg	cagacgagggaaggaaacag	3'BR
prx:Bhlha-ColA19	gatgtaggagggcgtggata	gatgtaggagggcgtggata	Hygromyicin resistence
Shh	ctgcttgaggttcagggtgt	gcttgtagccagcagctttc	5'BR
Shh	caccacacggagagtgctaa	ggcagcaagcagaaagtacc	3'BR
Tbx15	ggccccataatccttttcat	tcttttccaaatgggctctg	5'BR
Tbx15	atagccctcctgctcaacct	cattaggtcccctgaagctg	3'BR

Table 16: *FNDC3A* specific primer to analyze the sequence of the human gene

Name	Sequence Forward	Sequence Reverse	Specificity
FNDC3A_ex1	gaagcccgagagggtgtg	ccctcagtcccagcctaac	exon 1
FNDC3A_ex2	catcttgaatgtacaatgttttactg	aaaacacgaattaacaggaactcac	exon 2
FNDC3A_ex3	tgtaatccccattttccttcc	tgcttttgtgtgctggtagtg	exon 3
FNDC3A_ex4	ctgttggcttgagtgtttagatg	cacaacacttggtaagcctcag	exon 4
FNDC3A_ex5	gctaccgtgcctaccctactc	tggcatctattatatccaagcg	exon 5
FNDC3A_ex6	gtttatggagacacaaaggttgc	caaactgtcgaaagcagcag	exon 6
FNDC3A_ex7	agggagaagaagggattgtatc	gcttgtcacagaatacagtatgatcc	exon 7
FNDC3A_ex8	gggcaggtttatatttgtgagc	gtgtaccaagctacatttggacc	exon 8
FNDC3A_ex9	tggaatctgaatacagttacagttgg	gcccagcccacaaagtc	exon 9
FNDC3A_ex10	aaaggtgacactattgaaaacagg	accaggggaaacaaagaaac	exon 10
FNDC3A_ex11	atgaacattttcgcaggacc	tcaatccactcatgcaccac	exon 11
FNDC3A_ex12	ctgtcaagaaggaattgatgc	gcttttgctagactggtaacacatc	exon 12
FNDC3A_ex13	ttacatctgtcacaaacttgttc	aaatgccatcgtctcagc	exon 13
FNDC3A_ex14	ttttcacatattcattattgccc	aaccccaccacaccagtaag	exon 14
FNDC3A_ex15	tggttgaagctgtttgtctcc	ttttggactggcttgagataac	exon 15
FNDC3A_ex16	caacagcatggagtaattggg	atccaggacaacccaaagtg	exon 16
FNDC3A_ex17	tgataccagaacgttaaacattacc	gtgatcctaccccatttggc	exon 17
FNDC3A_ex18-19	ttcaaattatcagtgcaatatttgtg	ggagtcatatgttaaagccaaatg	exon 18-19
FNDC3A_ex20	tgcaatcagtggacatttaagg	ggcttactctgtataatataggcgg	exon 20
FNDC3A_ex21	tggccaaaagctatttaatcaac	cagagcctccgagttaacaaag	exon 21
FNDC3A_ex22	attgcagggtccaggtaaag	ctgaaaggaccagctccaag	exon 22
FNDC3A_ex23	aacagtgggaaaggttacaagc	tcttgcactcctgggctc	exon 23
FNDC3A_ex24	ttttggggtaacttatcactcc	gggcttgcctatacgattttg	exon 24
FNDC3A_ex25	tgattttaacctctttggatgc	gtgtagttcatactgaaaagatttgg	exon 25
FNDC3A_ex26	aacacattttctgtccagcc	caatgccagtgctaaatgcc	exon26
FnIII_3UTR_1	gcatttagcactggcattga	ggccaagtaaactgtctgttcttt	3'UTR
FnIII_3UTR_2	tgcaagccacaaaaatatcaa	gcccctccagttactcatca	3'UTR
FnIII_3UTR_3	tgttaggtagaggctggcact	tttgtttctaagcccctcca	3'UTR
FNIII_3pUTR1	cgtcatccttgtgctgtttg	gcccctccagttactcatca	3'UTR
FNIII_3pUTR2	tggaggggcttagaaacaaa	tgccccaaacaaaaaggtta	3'UTR
FNIII_3pUTR3	gctgcttctcccataactgc	tgttgtttctggggaaggtt	3'UTR
FNIII_3pUTR4	catttgagcctcactgcaaa	aaggcaaggagaggttaggg	3'UTR
FNIII_3pUTR5	tgtcaaatgatctgtgctgtagaa	attggaagaggggagaccat	3'UTR

Table 17: Other Primers

Name	Sequence Forward	Sequence Reverse	Application
laf4del_cDNA	accatggacagcttcgactt	tctctgagtcgctggaggag	to detect truncated *Laf4* transcript
cLaf4	tggacagcttcgacttagcc	tcctgagagagccctgttgt	to proof expression of *Laf4* 002 isoform
cLaf4_002	tttcagtcatcagccagcag	gtttgggaactccaacgaga	to proof expression of *Laf4* 002 isoform
Laf4_check	ttacggaggaaagagcgaga	gggggaaagttctgaacaca	to proof of *Laf4* KO in Laf4-KO line
probeNeo	ccacagtcgatgaatccagaaaag	ttttgtcaagaccgacctgtcc	to generate southern blot probe against Neomycin-Resistence gene
FNDC3a_24/26	tgtaatgaagctggggaagg	aaaagttccaaccccaaacc	to proof splicing of *FNDC3a* clone in mouse
RCBTB1_c9_1	gctgggaactggcaataaaa	ggtactggagaaaggcacga	to check splicing of *RCBTB1*
RCBTB1_c9_2	agcacacatcatggtggaga	ggtactggagaaaggcacga	to check splicing of *RCBTB1*

2.5.2 Sequencing Primers

Table 18: Sequencing Primers

Name	Sequence	Target
LINC00441_ex1seq	gtgggtccaggcgcct	LINC00441 Exon1 PCR from human gDNA
DLEU7_ex1seq	gaccacacacgctgtgagac	DLEU7 Exon7 PCR from human gDNA
pX459_F	ttagtacaaaatacgtgacgtagaaagt	CrispR guide in pX459
pX459_R	ttagtacaaaatacgtgacgtagaaagt	CrispR guide in pX459

2.5.3 qPCR-Primers

Table 19: Primers for Genomic Copy Number Analysis at the Human *Laf4* Locus

Name	Sequence Forward	Sequence Reverse
Laf4_5'out	tggagggaaaatgagtcgag	aatttgcctctcatgggttg
Laf4_5'max_1	aggcaggacgactcagacat	ttcgaccaccaaaagtctcc
Laf4_5'max_2	gcaatttagcatgtgcaagg	cccagttcaaaggttgcagt
Laf4_5'max_3	gacgtgcacataggtcagga	ctgccactttctctgaagca
Laf4_min_1	ggcatttcagaaggaagcag	tgctagtcctgggactttgg
Laf4_min_2	ctctgcaccactgagtccaa	tttactgttccggtcccttg
Laf4_3'max_1	caaggttcagaaaccagcaa	ttgtcgccctacttcctgag
Laf4_3'max_2	cgtgaccacaccaatcttttt	gtggtcgccgagtgtttatt
Laf4_3'max_3	aaaatctgtaatagcttgaacacaa	tgtcacagcgagactccttc
Laf4_3'max_4	cttatcttactacgggcctgacc	tctgtactcctggtagccgtga
Laf4_3'max_5	gccttgtgcctctgcatatc	tcagcacccaacactgtagc
Laf4_3out	catacgccttgttggaaatg	tgggaatgctgtctcaatgt
Laf4dup_1	gagtcccttccagcactctc	cagggtttacatttctggca
Laf4dup_5out	cgcaagttcagcgaaataaa	gtccgggtggaagaggtt
Laf4dup_1	agccgaagagtcatcctctc	ctgagcattctcaaatccga
Laf4dup_3	gctgccagaaatgtaaaccc	gaggatgactcttcggcttg
Laf4dup_4	tcatcgtggctcctaaagagt	cttccggttctaggccaact
Laf4dup_5	gactctcctgacctcttggc	ggagtgcgaggtttggtt
Laf4dup_6	tccacagttggcctagaacc	ggtgagaatccggagctgt
Laf4_dup7	gggattccccagaattccta	cttgcacccagggtttacat
Laf4_break	attagctaggcgtggtggtg	ctcatgcggtgagtgtagga
Laf4_3dup1	gccacagactaaagttggca	ggcttagccacatagggtct
Laf4_3dup2	aacgactgatttgcattcca	ccacagacatgcttctgtca
Laf4_3dup3	cctgggaatacctgcttcat	ttggaatttgttattgatccaag
Laf4_3dup4	cagcaaattgtggttccaat	tgtactggctgccctctatg
Laf4_3dup5	caacccagatgagagttcca	tcattgcaggcctcatacat
Laf4_3dup6	accagacgtgtgagtggaag	ttgctcaagatgtcatgggt
3dup3.1	ccacaggatctctgactgga	gggggtaaattctgtgacca
3dup3.2	gcagtaatacctggcacacg	tgctttgatttgaacccatc
3dup3.3	tgccttctgactgtaagcga	gccttgacagtcatgtcacc
3dup4.1	gcaaattgtggttccaatcc	tgccctctatgtcagcaatg
3dup4.2	tcacacccatcctgtgctta	cctccgtgctaaaccagtgt
3dup4.3	gttttgcctgacctctcagc	gggacttccttgggctactc

Table 20: Primers for Genomic Copy Number Analysis in Mouse

Primer Set	Sequence Forward	Sequence Reverse
mAlbumin	ctgcaatcctgaaccgtgt	ttccaccagggatccactac
mF8_EX8	aaattcgctcggttgctaaa	tgaaggtgcatagtcccagtc
qB9_5p1	tatctctgtgagccgtgagg	tcctttactgggatggaagg
qB9_5p2	ggctgtgatgacaggagcta	cctcacggctcacagagata
qB9_3p1	agccaggagctcattccat	ctttgagcctcagtttgtgc
qB9_3p2	caccgagccatctcagttc	tcttggtcaggtgtgatggt
qB9_I_1	agccaggagctcattccat	ctttgagcctcagtttgtgc
qB9_I_2	ccctgacagatacgggaaat	gcaggacacaggacgagtta
qB9_I_3	cacagtcaactgcagggaag	acccataggtttggcaagag
qB9_I_4	ttctccagcccgttagattt	ccaacctcgtctacaaacca
qSHH_5p1	gcgaggatggtgacagacta	tctccttgcagaaggaacag
qSHH_5p2	cctagtctacccaccacagcta	ggctgggtatagggactgg
qSHH_3p1	tgctggtactttctgcttgc	tcctgttgagagctccgtaa
qSHH_I_1	ccaaatgagaaggctgaaca	gccagatgacccttatgtcc
qSHH_I_2	tgctgcataatggagaggtt	ggtgcacctatactgcctga
qSHH_I_3	ctgcagcattctgtggaaa	ccaaggaactcttggcactt
qSHH_I_4	ccaagggctgggaacata	ctcttctgttaggcatggca
qTBX_5p1	catgacccaggcctttctat	agacattgatcagcccacag
qTBX_5p2	ctgtgggctgatcaatgtct	agcaagccaccctctctaaa
qTBX_3p1	gaggtccaacctgatgcttt	tgataaacacggaatggacaa
qTBX_3p2	agcttaggcactcaaagagaca	aaagcatcaggttggacctc
qTBX_I_1	cagcactgcagataatgcct	atgcacatgacttcagcaca
qTBX_I_2	tgaaaggattcctgttggg	cttgggcagtttacacagaca
qTBX_I_3	agcagaactgggccaataac	tttctcaagaaccagaacattca
qTBX_I_4	accttggctgagggtagttg	ccttacaagagctgcttgagata

Table 21: Primer for RNA Expression Analysis in Mouse

Target	Sequence Forward	Sequence Reverse
Bhlha9	agcgcatcctggattacaac	ggaccaaggatagagcagtga
Bhlha9	gattacaacgaggcgttcaa	acgcaggaccaaggatagag
Bhlha9	gggaattccagcttcagtgt	tgtaggcggcactaaggg
Gadph	ctgcaccaccaactgcttag	ggatgcagggatgatgttct
Rbm33	aaagcatccaagggattcac	accactcgaaggttctgctt
Rbm33	cacatgtcagctccaaggtc	ctggccttctgtgtgagaaa
Rbm33	atccggagaacaggaatctg	gcatcatggccttcatactg
Tbx15	ctggcagaaacagaactgga	gtcctcaaaggtcagtgtgc
Tbx15	ctctttgggaccctgcttag	agaatggaactcccacaagg
Tbx15	ccttgtgggagttccattct	ctgtcgtccatgactcgttt
Shh	tgtacagcgacttcctcacc	agcgtctcgatcacgtagaa
Shh	ggacagctcacaagtcctca	aggaagcaaggatcaccaga

2.5.4 Oligos for CRISPR Experiments

For introducing point mutations in exon 26 of the murine gene *Fndc3a* PAGE purified single stranded DNA (ssODN = dFnIII_3F) with the following sequence was ordered at IDT DNA:

tcagctgaaggctggaataccggaatgaagtatcaggacccttgc
aaatctgttaataaaatattagaaaaagaatggagtacaaagtct

Table 22: CRISPR Guides

Name	Sequence	Target Locus	Score	Genomic Score	Exonic Score
sgB9_5	aggcacagaccatggtcgac	*BHLHA9*	89	1.3	0.2
sgB9_3	tcaagtgcgtaagccacgat	*BHLHA9*	91	1.1	0.0
sgB9_5n1	gttccgtgcccccatggcgg	*BHLHA9*	93	0.8	0.8
sgB9_5n2	aggcacagaccatggtcgac	*BHLHA9*	89	1.3	0.2
sgB9_3n	agcccggataggaccctctc	*BHLHA9*	87	1.7	0.6
sgLaf4_5'1	gaatggtcctgcgttttcgc	*Laf4*	93	0.8	0.3
sgLaf4_3'2	gtcccataggggaaaacgcg	*Laf4*	92	0.9	0.6
sgSHH_5'1	cagtacaaatgtgggtgcgt	*SHH*	82	0.8	0.8
sgSHH_3'	catgagtctaccccgcttgt	*SHH*	91	1.0	0.3
sgTBX15_5'	cagactcttgttaccgtgta	*TBX15*	90	0.9	0.2
sgTBX15_3'	gccttgctatgggtcgaaga	*TBX15*	91	0.8	0.2
sgFnIII_3F	gggcccttgtaaatctgttg	*FNDC3a*	75	1.7	0.7

2.6 Enzymes and Antibodies

Restriction enzymes, ligase, polymerases and other DNA modifying enzymes were purchased from MBI Thermo Scientific.
FNDC3A antibody was ordered from Sigma Aldrich (HPA008927) and HSP60 antibody from Abcam.

2.7 Bacteria and Plasmids

Cloning steps of plasmids (see table 23) were performed in *Escherichia coli* TOP10 cells.

Table 23: Plasmids

Name	Application	Supplier
pGKFrtATG-PNN	targeting of ColA locus in C2 cells	Open Biosystem, adapted by Martin Franke
pTAgfp	sub cloning of PCR products and RNA in-situ probes	Dr. Jochen Hecht (MPIMG, Berlin)
pX459	expression of sgRNA and Cas9	
10252	targeted Knockout of *Bhlha9* in ES-cells	KOMP

2.8 Media for Cell Culture

Table 24: Ingredients

Name	Supplier
100xglutamin	Lonza BE17-605E
1xPBS	Lonza BE17-512F
100xpenicillin/streptomycin	Lonza DE17-603
Beta-Mercaptoethanol	Sigma M-7522
Bicarbonate free	Gibco 52100
DMEM	Lonza BE12-733
DMSO	Sigma Aldrich D-2650
FCS for feeder	Biochrome
FCS for ES cells	PAN Biotech P122011
Fugene	Promega TM-238
Gelatin	Sigma G-1393
KO-DMEM	Gibco 10829-018
LIF	Chemicon ESG1107
Mytomycin C	Sigma M-4287
Non-essential Amino acids	Gibco 11140-35
Nucleosides	Chemicon ES-008D
OptiMEM	Gibco
Trypsin-EDTA	Gibco 25300-054
Water	Lonza BE 17-724Q

Table 25: Media Composition

Media	Application	Composition
2x ESC Freezing	Freezing ESCs	in KO-DMEM: 20%FCS, 20%DMSO
2x Feeder Freezing	Freezing Feeders	in DMEM: 20%FCS, 20%DMSO
2x HAF Freezing	Freezing HAFs	in DMEM: 20%FCS, 20%DMSO
2x Plate Freezing	Freezing ESC-plates	in Bicarbonate-free DMEM: 20%FCS, 10%DMSO
20% FCS Feeder	Freezing Feeders	in DMEM: 20%FCS
20% FCS ESC	Freezing ESCs	in KO-DMEM: 20%FCS
20% FCS HAF	Freezing HAFs	in DMEM: 20%FCS
ESC Media	cultivating ESC	in KO-DMEM: 15%FCS, 2mM glutamine, 0,05U/ml penicillin/streptomycin, 1x non-essential amino acids, 0,1mM Beta- Mercaptoethanol, 1x nucleosides, 1000U/ml LIF
HAF Media	Cultivating HAFs	in DMEM: 10%FCS, 1% glutamine, 1% penicillin/streptomycin
Feeder Media	Cultivating Feeders	in DMEM: 10%FCS,2mM glutamine, 0,05U/ml penicillin/streptomycin
Gelatine	Coating dishes	in water: 0,1% gelatin

2.9 Software

Table 26: Software

Name	Application
Adobe Illustrator	Image processing
ApE	Sequence analysis
Axio Vision	Digital photography
DNA Star Seqman	Sequence analysis
Microsoft Office	Data analysis, text processing
Qbase software package	qPCR analysis
SDS	qPCR analysis
ZEN	Digital photography

2.10 Internet Resources

Table 27: Internet Resources

Resource	Address
BioGPS	http://biogps.org
CRISPR Design	http://crispr.mit.edu/
EMBL-EBI	http://www.ebi.ac.uk/
Ensembl Genome Browser	http://www.ensembl.org/index.html
Human Protein Atlas	http://www.proteinatlas.org/
Mouse Atlas	http://www.emouseatlas.org/emap/home.html
NCBI	http://www.ncbi.nlm.nih.gov/
Primer3	http://primer3.ut.ee
UCSC Genome Browser	http://genome.ucsc.edu/
Vista Enhancer Browser	http://enhancer.lbl.gov/
YUElab	http://promoter.bx.psu.edu/hi-c/view.php

3 Methods

3.1 Biomolecular Methods

If not described different all classical molecular biological or microbiological experiments were performed according to the handbook
"Molecular Cloning: A Laboratory Manual" (Sambrook et al., 2012).

3.1.1 DNA-Isolation

For genotyping mice, ear and tail biopsies are provided by the animal house. To isolate DNA 50µl QuickExtract is added to each sample. After 20 minutes incubation at 65°C, the reaction is stopped by two minutes incubation at 96°C. After cooling down the extraction solution containing DNA is separated from the sample. Sample and DNA are stored at -20°C.

For embryo genotyping amnions are lysated in 500µl Tissue-Lysation Buffer by overnight incubation at 55°C. Next day 0,25ml 5MNaCl are added. After 10 minutes shaking at RT and another 10 minutes on ice the samples are centrifuged for 20 minutes at 9.000rpm and 4°C. Then 500µ µl of the supernatant are added to 1ml 100% Ethanol p.a (-20°C). After 30 minutes of centrifugation at 13.000rpm and 4°C the DNA-Pellet is washed in 70% Ethanol and centrifuged for another 15 minutes at 13.000rpm at RT. After 30 minutes drying the DNA is dissolved in 100µl bidest. water.

To isolate DNA from cultivated cells, cells are lysed by overnight incubation in 50µl Cell-Lysisbuffer per 96-plate well at 55°C. At the next day the DNA is precipitated in 1ml isopropanol containing 100µl 8M LiCl. After 20 minutes centrifugation at 10.000rpm and 4°C the DNA-Pellet is washed twice in 70%Ethanol and centrifuged at 10.000rpm. After 30 minute drying the DNA is dissolved in the appropriate amount of bidest. water.

Plasmid-Isolation from *E.coli* culture were performed with the Mini-Kit (Machery-Nagel) for 5ml bacteria culture and with the MIDI-Endotoxinfree-Kit (Machery-Nagel) for 50ml bacteria culture according to the specification of the manufacturer.

3.1.2 RNA-Isolation

RNA-Isolation were performed with the RNA-Extraction Kit from Macherey-Nagel according to the specification of the manufacturer.

3.1.3 Generation of cDNA

Generation of cDNA was performed with the TaqmanKit from Roche by using hexameres and following the specification of the manufacturer.

Standard Genotyping		Laf4-KO Genotyping	
Ingredients	End concentration	Ingredients	End concentration
10xTaq-Buffer	1x	10xTaq-Buffer(MgSo)	1x
10mM dNTPs	160µM	10mM dNTPs	160µM
Primer F(10pm/µl)	400nM	Primer F(10pm/µl)	500nM
Primer R(10pm/µl)	400nM	Primer R(10pm/µl)	500nM
Taq-DNA Polymerase	0,002U/µl	Taq-DNA Polymerase	0,002U/µl
Template	1ng/µl	Template	1ng/µl
Water	ad to 50µl	Water	ad to 25µl
Cre Genotyping		**Colony PCR**	
Ingredients	End concentration	Ingredients	End concentration
10xTaq-Buffer	1x	10xTaq-Buffer	1x
10mM dNTPs	50µM	10mM dNTPs	16µM
Primer F(10pm/µl)	500nM	Primer F(10pm/µl)	400nM
Primer R(10pm/µl)	500nM	Primer R(10pm/µl)	400nM
Taq-DNA Polymerase	0,002U/µl	Taq-DNA Polymerase	0,002U/µl
Template	1ng/µl	Template	1ng/µl
Water	ad to 50µl	Water	ad to 50µl

Figure 14: Protocols for Genotyping PCRs

3.1.4 Polymerase Chain Reaction (PCR)

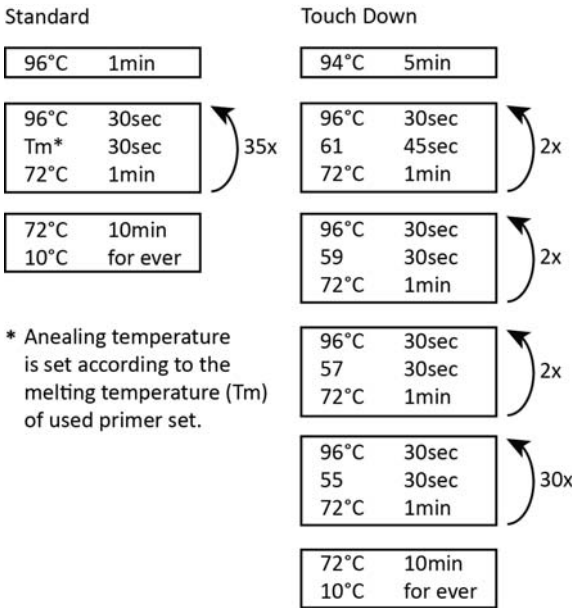

Figure 13: Cycler Programs for Genotpying PCRs

Genotyping PCR of different mouse lines is performed with primers listed in table 15. Protocols and cycler programs for each genotyping PCR are listed in figure 14 and figure 13. The settings and the expected product sizes for each genotyping is summarized in table 28. The line FNDC3a was genotyped by Sanger sequencing of the wild-type PCR that was run under standard conditions at 60°C resulting in a product of 802bp.

Table 28: Genotyping PCRs

Line	Protocol	Program Tm[℃]	product size[bp]			
Conventional Lines			Wildtype		Mutant	
CMV-Cre	Cre	Standard 55	-		650	
Bhlha9-ColA1	Standard	Standard 60	1644		1972	
Laf4-KO	Laf4-KO	Standard 55	100		200	
prx:Bhlha9-ColA1	Standard	Touch Down	987		780	
CRISPR Lines			5' BP	3' BP	dup	del
Bhlha9 5/3	Standard	Standard 59	818	963	1072	749
Bhlha9 5.1	Standard	Standard 59	541	605	568	715
Bhlha9 5.2	Standard	Standard 59	541	605	504	601
Laf4-trunc	Standard	Standard 61	422	460	504	601
Shh	Standard	Standard 58	769	571	675	665
Tbx15	Standard	Standard 58	761	571	850	482

3.1.5 Sanger-Sequencing

For Sanger Sequencing of PCR products or plasmid BigDyev3.1 (Applied Biosystem) is used according to the specification of the manufacture. The reactions are precipitated and transferred to the capillary electrophoresis (ABI 3700) by Mohsen Karabsyan at Charité - Universitätsmedizin (Institute for Molecular Genetics).

3.1.6 Quantitative Genomic and Real-Time PCR

In order to measure the copy number of genes and their expression quantitative and real time PCRs are performed. For one reaction 4ng genomic DNA or 1.5ng cDNA are used in a reaction of 12µl consisting of 0,25µm primer mix (forward and reverse) and 6µl Sybergreen (Qiagen). To measure the efficiency of primers a standard curve is performed by running the reaction with five different concentrations:

Genomic DNA: 2; 1; 0.5; 0.25; 0.125 and 0 ng/µl

cDNA: 3, 1.5; 0.75;0.375, 0.1875 and 0 ng/µl

For genomic DNA the copy number of F8 and Albumin and for real time PCR the expression level of the house keeping gene Gapdh was used for normalization.

3.1.7 Southern Blot

Southern blot is a method to detect defined DNA-Sequences on Membrane. The DNA is first enzymatic digested and fragments are separated by gel electrophoresis. DNA is blotted on a membrane where it can be detected by hybridization with label DNA probes. In this work Southern blot was used to detect C2 clones with the right integration of the construct

prx:BHLHA9 (see 3.4.4).

Day1-Blotting In the first step genomic DNA of C2-clones are enzymatic digested with EcoRV resulting in fragment size that differ between wildtype (26kb) and mutant clones (3,7kb). These fragments are later detected with a probe that is specific for the *ColA1* locus. Samples are loaded on a 1% agarose gel, that run for 6 hours at 80V. The gel is denaturized for 30 minutes in denaturation solution which is stopped by two 20 minutes step in neutralization solution. Now the DNA is blotted on a nylon membrane overnight: A block is wrapped in Whatman paper and put in a tray with 10xSSC. Another two layers of Whatman papers are added on the block. Everything is soaked in 20xSSC. The gel is put on the papers an air is removed with a glass rod. The membrane is cut to the size of the gel, dipped in 10xSSC and added on the gel. Again, air bubbles need to be removed before the next two layers of wet Whatman paper are added. Everything is sealed with cling film to prevent evaporation. A window is cut at the position of the gel and a pack of paper towel is fixed with a metal blot on top. The paper tower soaks up the SSC in the tray and facilitates in this way the transfer of the DNA from the gel to the membrane.

Day2 - Hybridization On the next day the blot can be checked under UV light that visualize the DNA on the membrane. If this is the case neutralization is performed in 50mM NaPi before DNA is fixed on the membrane by 2h incubation at 80°C. Hybridization is performed in a glass falcon with the DNA side facing inside. First the membrane is pre-hybridized for 30 min. in Church-Hybridization buffer at 65°C on a rotating device (10rpm). Meanwhile 4-12µl probe is denatured in 25ml Church-Hybridization buffer at 100°C for 10 minutes. For hybridization the membrane is incubated in the probe containing buffer at 65°C overnight.

Day3 - Developing At the next morning two 10 minutes washing steps are performed in Church-Washing Buffer at 65°C and one at RT. After five minutes incubation in 1xDIG1/0.3% TWEEN the membrane is sealed in a plastic bag with blocking solution DIG2 and mixed for 30 minutes. For antiDIG-antibody binding the solution is replaced by DIG1 for 20 minutes. The plastic bag is removed and the membrane is washed for 20 min. in 1xDIG1/0,3%TWEEN and for 5 min. in DIG3. The membrane is placed in a developing cassette and covered with CDPstar solution for 5 minutes. X-ray films are exposed for 5-20 minutes depending on the efficiency of the probe and developed with Imaging system Curix 60.

3.2 Cloning Methods

For cloning linear plasmids are dephosphorylated with alkaline phosphatase (Fermentas) and the to be inserted DNA fragment is phosphorylated with T4-Phosphorkinase (Fermentas) according to the manufacture protocol. Ligation and Transformation is performed in *E.coli* TOP10 as described in Sammbrock et al. (2002).

3.2.1 CrispR Cas9 Guide Design

Before starting the CrispR Design it has to be considered at which position breakpoints should be induced in order to genocopy the patient. For this the corresponding position in the mouse genome has to be determined as well as how close the break point needs to be to this position. This is highly depending on the conservation between human and mouse genome and gene density at the target locus. The guide RNA is the most important factor of the CrispR experiment and has a high impact on the efficiency and specificity. The online software CrispR Design Tool (http://crispr.mit.edu) [124] is used to find potential guide sequences in region of interest. Therefore, the sequence type 'other region' and the target genome 'mouse' need to be selected. As a result, the software will provide different guide sequences in a rank order list based on predicted specificity. This main score should be as high as possible but never below 70. The other two scores reflect the likelihood of unspecific bindings (off-targets) in the whole genome and the exome, which is influenced by the number and position of guide mismatches. A mismatch between guide and target sequence still allows binding of the guide when it is located more distal to the PAM sequence. Because of that off-targets with mismatches close to the PAM sequence are less troublesome since it is relative unlikely that Cas9 will cut. A genomic or exonic score of zero means that there is no predicted off-target; therefor, the score should be as low as possible. Furthermore, CrispR Design Tool quotes the exact position of different off-targets and genes that are affected by them. In this way off-targets can and should be avoided if they contain vital genes or genes that play a role in the patient syndrome. After the guide sequence had been chosen overhangs are added to enable cloning into pX459. For the sense sequence it is 'CACCG' at the 5-prime side and for the antisense sequence 'AAAC' at the 5-prime and 'C' at the 3-prime end. The guides (sgOligos) are ordered as primer via MWG.

3.2.2 Cloning of CrispR sgOligos

The Cas9 expression vector pX459 is linearized with BbsI and dephosphorylated with alkaline phosphatase. In order to clone the guide DNA in the Cas9 expression vector pX459 the ordered single stranded sgOligos are diluted to 100pmol/μl as recommended by the company. To anneal the sgOligos 10µl of each stock solution are mixed with 10µl 10x-ligation buffer and 70µl bidest. water and placed in a preheated 95℃ metal heating block that is switched off. After 15 minutes this block is taken out of the heater and cooled down for another 45min at RT. The annealed DNAs are phosphorylated in T4-PNK buffer A as described above. Ligation and Transformation are performed as written in Sammbrock et al. (2002). Since most of the clones will carry the vector with the right insert no prescreening is performed. Instead Minis are inoculated with four clones, cleaned up and sequenced with the Primer pX459_F and pX459_R.

3.3 Cell Culture Method

3.3.1 Splitting Cells

To split cells, they are washed twice in 1xPBS and trypsinized in the appropriate volume Trypsin-EDTA for 5 minutes at 37℃ in a cell incubator. Cells are then separated by up- and down-pipetting with a 1000µl pipette. The reaction is stopped by adding the double amount of media to the cells. This suspension is then added to 5ml media and centrifuged for 5 minutes at 1000rpm. After discarding the supernatant, the cell pellet is re-suspended in the appropriate media volume and seeded on the needed area.

3.3.2 Thawing Cells

Cell-vials are thawed in 37 ℃ water bath, then 1ml Media is added to the cells and the cell suspension is transferred to a 15ml falcon with 5ml Media. After centrifugation at 1000rpm for 5minutes the supernatant is discarded and the cell pellet is re-suspended in the appropriate media volume and seeded on the needed area.

3.3.3 Cryoconservation of Cells

For cryoconservation in vials cell are washed twice in 1xPBS and trpysinized as described in section 3.3.1 on page 50. After centrifugation the cell pellet is re-suspended in 1ml 20%FCS-Media per vial. To freeze one vial 1ml of this cell suspension is added to 1ml freezing media. Vials are frozen in a cryo-box with Isopropanol at −80℃.

3.4 Embryonic Stem Cell Culture

3.4.1 Culturing Embryonic Stem Cells

Embryonic Stem Cells (ES cells or ESC) are in general cultured in gelatinized dishes on a layer of mitosis inactivated murine embryonic fibroblasts (MEFs), also referred to as feeder cells. MEFs helps to maintain pluripotency of ESC by secreting Leukemia Inhibitory Factor (LIF).
For this study feeder cells are derived from either wildtype CD1 mice or transgenic DR4-mice carrying resistance against Hygromyicin, Neomycin, Puromycin and 6-Thioguanin. If not described differently ES cells are cultured on CD1-feeder cells. A day before seeding ES-cells dishes are coated with 0,1% gelatin and inactivated feeder cells are plated. After 12 hours feeder cells are attached to the bottom and can be checked for 80% confluence. Then ES cells can be plated. ES cells grow rapidly and divide every 18-24 hours. For this reason, medium has to be changed every day. If ESC colonies are too big or too close to each other cells are differentiating. This can be observed by irregular shape and structure of the colonies. To prevent this bad performance cells have to be splitted in advance – typically every second day. To reduce stress medium has to be changed two hours before trypsinization.

3.4.2 Preparation of Murine Embryonic Fibroblasts

For preparation of MEFs form E13,5-14,5 mouse embryos are dissected from the uterus and extra-embryonic membranes and placed into 1xPBS. The head and soft tissue are discarded and the remaining carcasses are washed in 1xPBS and minced into cubes about 2-3mm with a pair of fine scissors. The cubes are transferred into a 50ml falcon with 5-20ml Trypsin-EDTA depending on the number of embryos. The cubes are further minced by up and down pipetting and then incubated for 10 minutes at 37°C. The tissue cubes are again minced by up and down pipetting. Inactivation of trypsinzisation is achieved by adding double volume of feeder medium. To get a single cell suspension the cells are separated by up and down pipetting. The cells are plated on a 15cm dish (2-3 embryos per dish) and incubated overnight. At the next day an aliquot of medium is taken for Mycoplasma-test that is performed with Myoalert Detection Kit according to the specification of the manufacture. The cells are frozen as described in section 3.3.3 on page 50.

When MEFs reach confluency, proliferation is inhibited by mitosis inactivation with Mytomycin treatment. Mytomycin C inhibits DNA synthesis by CpG cross-linking and in this way prevents cell division without further effects on protein synthesis and secretion. After three hours Mytomycin is removed, cells are washed thrice with PBS and cultivated for another day. Then inactivated MEFs are cryoconserved as described in section 3.3.3.

3.4.3 Cell Transfection with the CrispR/Cas9-System

ES-Cells (wt-line G4) are thawed from vials and seeded on two 6cm-dishes in two different concentrations: 70% and 30%. After two days the dish with better performance (see 3.4.1) is trypsinized, the cell number is determined with a haemocytometer and 400.000 cells are plated on one 6-well-plate well. After one day recovering cells are fed with 1,75ml ES-medium without Penicillin and Streptomycin and 30 minutes before transfection all reagents are warmed up at RT.
For transfection 125µl DNA-solution (125µl Optimem containing 5–8µg pX458) and 125µl FuGENE-solution (100µl Optimem, 25µl FuGene) are combined. This transfection mixture is incubated at RT to enable the formation of particles. After 15 minutes the whole transfection mixture is added drop-wise to the 6-well plate. After 12 hours the transfection is stopped by feeding with normal ES medium and the cells are allowed to recover for 24 hours.
After recovering ES cells are splitted to four 6cm-dishes with DR4-feeder layer. Selection is started directly by adding Puromycin to ES medium (final Puromycin concentration: 2µg/ml). After 48 hours selection is stopped by switching to normal ES medium and cells are allow to recover for 4–6 days, depending on the occurrence of colonies.

3.4.4 Cell Transfection of C2-ES cells with the flip-in System

C2-ESC derive from the mouse line C2 that carries a frt side at the *ColA1* loci. This side allows integration via the flip-in system that consists of the flipase coding plasmid pCAGGS-Flpe and a plasmid with a second frt side (pGKFrtATG). The flipase binds both frt sides and

induces recombination into the locus. This System was used to integrate the pGKprx:Bhlha9 construct that drives overexpression of *Bhlha9* in the limb bud.
Before transfection 300.000 C2 ES cells are seeded on a 6cm dish and cultivated for one day. On the next day the media of ESC is exchanged to let them recover for two hours. The DNA solution is prepared by mixing 5µg of the construct and 1µg of pCAGGS-Flpe with 125µl OptiMEM. 125µg of the DNA solution is combined with Lipofectamin Solution (25µl Lipofectamin in 110µl OptiMEM) and incubated for 15 minutes at RT. The transfection mix is added to 1,25ml ESC media without penicillin or streptomycin and used to replace the media on the C2-ESC. After 3-5 hours the transfected cells are split on three 6-cm dishes in three different ratios (1/3,1/6 and 3/6). At the next day selection is started with ESC media containing Hygromycin (150µg/ml). Clones are expected to be seen after seven to ten days of selection.

3.4.5 Picking of Clones

A day before picking ESC clones 96-well-F-bottom-plates with feeder cells need to be prepared. Colonies can be picked when they are visible with a stereo microscope. At least three hours before picking medium of the growing ES cells has to be exchanged. Dishes with ES colonies are washed twice with 1xPBS, then 3ml 1xPBS are added. Colonies are picked with the 10µl pipette under a microscope and single colonies are transferred to 96-well U-bottom plate containing 30µl Trypsin-EDTA. After colonies have been picked (but not more than 10 minutes after picking) U-bottom plates are incubates for 10–15 minutes in the incubator. Then 60µl ES medium is added to each well. Colonies are disaggregated by pipetting up and down and transferred to the 96-well-plates with the prepared feeder layer.

3.4.6 Split and Freeze in 96-Well-Plates

The clones that had been picked need to be cryconserved and genotyped. For this purpose the 96-well-plates containing the clones are splitted in three parts after two or three days depending on their size. Two parts are frozen in U-bottom plates while one part is kept in ES medium without LIF. The cells in this so called DNA-plate are allowed to differentiate since they are used for DNA isolation and genotyping.
Before starting splitting the cells 96-well-U-bottom plates with 50µl 2x-plate freezing medium are prepared. The ES cells in 96-well-plates are trypsinized as described in 3.3.1, but trypsinization is stopped by adding 100µl bicarb-free medium containing 20% FCS to each well. After disaggregating them by pipetting up and down 50µ µl of the cell suspension is added to the previously prepared U-bottom plates with 2x-plate freezing medium. These freezing-plates are sealed with autoclave tape, packed in polystyrene boxes and kept at −80°C. After two days freezing plates can be stored at −80°C without polystyrene box.
The leftovers in the initial clone plate are filled up with 200µl ES medium without LIF. The medium of these cells is exchanged every second day until they are harvested at 90% confluence for DNA extraction (3.1.1).

3.4.7 Thawing and Expansion of Cells from 96-Well-Plates

ES Cells in 96-well-U-bottom plates are thawed by adding 100µl pre-warmed ES medium to single wells. The cell suspension is then transferred to a 15ml falcon with 1ml medium and centrifuged at 1000rpm for 5 minutes. After discarding the supernatant, the cell pellet is re-suspended in 200µl medium and seeded on a 96well-plate. After approximate two days cells can be transferred to 24well-plate and after another two days to 6well-plate. These cells can be frozen in three vials after two days. 100µl of the cell suspension is used to culture ES cells feeder cell free for reconfirming the genotype.

3.4.8 Diploid Aggregation

All mouse line of this work had been generated by the animal facility (Max Planck Institute for Molecular Genetics,Berlin) via diploid aggregation as described by Artus and Hadjantonakis (2011). In this method zygotes are isolated from pseudo pregnant mice and cultured until they reach morula stage. ES cells are injected into the morula and retransferred into foster animal that give birth to chimeric mice carrying wildtype and mutant cells [46].

3.5 Biochemical Methods

3.5.1 Protein Extraction from Fibroblasts

Cultured fibroblasts are lysed on the 10cm culture dish by adding 1ml membrane lysis buffer that contains inhibitor cocktail (Complete Roche). After 10 minutes incubation at RT cell are disrupted by up- and down pipetting. Lysate is transferred to a 1.5ml test tube and sonicated with the Bioruptor in a cooled water bath. Four Series of impulses with low power for 30 sec. and 30 sec. rest are performed. Sonicated lysate is centrifuged for 5 minutes with maximum speed at 4℃. The protein containing supernatant is transferred to a new tube and stored at -20℃.
Protein concentration was carried out using the BCA Protein Assay Reagent Kit according the manufacturer protocol. Extinction was measured at 563nm with a titer-plate reader and concentration was determined with the help of a BSA standard curve.

3.5.2 SDS PAGE and Western Blot

To compare expression level and protein size of FNDC3A between patient and healthy controls SDS PAGE and Western Blot was performed with in samples from human fibroblast. For this purpose, a 3% polyacrylamid gel is prepared according to Sambrock et al. (1998). Protein samples are obtained from cell culture as described in 3.5.1 and boiled for 10 minutes after adding 4x protein loading buffer. Pre-stained and probes are loaded together to the gel that is run for 30min at 80V and then at 100V until bromphenol blued reaches the lower edge of the gel.
Proteins are transferred on a Polyvinyl fluoride (PVDF) membrane that needs to be activated with methanol before. After equilibration in transfer buffer proteins a gel-membrane sandwich is assembled that is surrounded by Whatman paper and placed in a semi-dry blotting chamber. Proteins are transferred at 22V for 30 minutes.

Before protein detection the membrane is blocked in 5% milk powder in TBST for one hour. The primary antibody is incubated overnight at 4℃ in the dilution 1:100 for FNDC3a and 1:1000 for HSP60. The next day membrane is washed three times in TBST for 5min and incubated in HRP-coupled secondary antibody in the dilution 1:1000 for 1 hour. After three time washing in TBST visualization is carried out with enhanced chemo luminescence (ECL) with the help of LAS 4000 Imaging System.

3.6 Histological Methods

3.6.1 Immunohistochemistry on Cultured Cells

For immunohistochemical stainings fibroblast are cultured on glass cover slips to 70% confluence. After three times washing in PBS cells are fixated for 5 minutes in 4%PFA (in PBS) at RT. Blocking is performed for at least one hour in blocking solution at RT. Primary antibody is diluted in blocking solution overnight at 4℃ in a humidified dark chamber. The next day primary antibody is removed with three washings step (PBX) and cells are incubated with secondary antibody and DAPI (1:1000 dilution) for 1h at RT in a humidified dark chamber. After 3 washes in PBS cover slips are mounted on slides in Fluoromount G.

3.6.2 Methylmethacrylat (MMA) Embedding and Sectioning

For histological stainings mouse limbs are embedded in MMA and then sectioned with the Microtome. For this limbs are fixed in 4% PFA overnight at 4℃. Then samples are subsequently washed thrice in 1xPBS and dehydrated in 50%, 70% and 100% EtOH for at least one hour per step and then twice in 100% EtOH for 24 hours, twice in UltraClear for two hours. The tissue is infiltrated with MMA-Infiltration solution (10% Polyethelene Glycerol and 0.33% benzoyl peroxide in MMA) for 24 hours at 4℃. The limbs are embedded in MMA polymerization solution (10 % Polyethylene Glycerol, 0.55% Benzoyl Peroxide, 0.5% N,N-dimethyl-p-toluidine in MMA) which needs to polymerize at 60℃ for at least one day. 5µm sections are achieved with a Microtome (LeicaRM2255).

3.6.3 Immunohistochemistry

3.6.4 Histological Stainings

Histological stainings were carried out on MMA embedded sections of 5µm. MMA is removed by three subsequent washes in MMA-removing solution (2-Methoxyethyl-acetat) for 10 minutes and subsequent rehydration twice in Xylol for 3min, twice in 100%EtOH, once in 90%, 70%, 50% EtOH for 1min and 5min in VE water.
DIC microscopy is performed on dehydrated unstained MMA sections. Histological stainings are performed as follows:

VanCossa/Nuclear Fast Red staining Rehydrated MMA-sections are kept for 3min in 1% $AgNO_3$, unbounded silver is removed by washing under floating VE water for at least 3min. The sections are incubated for 2min in 10%Formamid/5%$NOCO_3$, which is removed by washing under floating tape water for at least 10min. The staining is finished with 5

minutes incubation in 5%$NO_2S_2O_3$ which is again removed by washing under floating tape water for at least 10min. The sections are counterstained for 5 minutes in Nuclear Fast Red-Aluminum Sulfate solution. Unbound dye is removed by short dipping in VE water. Ready stained sections are air dried for several hours and mounted in Entellan.

Masson-Goldner Trichrome staining Rehydrated MMA sections are kept for 2min in Weigert's Iron Hematoxylene, which is removed by washing in floating tape water for 10 minutes. Acid Fuchsin – Ponceau Red is added for 5 minutes and inactivated by dipping in 1% acetic acid. The staining is differentiated for 10 minutes in Orange G-Phosphormolybden and again stopped in 1% acetic acid. Incubation in Fast Green is performed for another 10 minutes and finally stopped by dipping in 1% acetic acid. Acid is briefly removed in VE water. Slides are air dried for several hours before they are mounted in Entellan.

3.6.5 Skeletal Staining

Skeletal stainings are used to reveal the structure of every bone in the complete skeleton and to allow to distinguish bone from cartilage. For this purpose, sacrificed mice from the embryonic stage E14.5 are treated witch Alcian Blue (cartilage staining) and Alizarin Red (bone staining).
First the skin is disrupted by 1 min incubation in a 65°C water bath. Then the skin can be easily removed with forceps as well as any inner organs apart from the brain. For E18.5 or older stages also muscle are partly removed. The carcass is fixed overnight at 4°C in 100% EtOH, which is replaced with Alcian Blue solution (see 13). After 12 hours skeletons are washed in 100% EtOH for 12 hours. The pH of the tissue is increased by a 6-hour treatment with 1% KOH (in ddH_2O). Now, Alizarin Red solution (see 13) is added for another 6 hours. To make the skeleton more visible tissue is removed by KOH digestion. Depending on the appearance of the skeleton different concentrations are used:

3% KOH: adult mice
1% KOH: postnatal mice
0,5% KOH: prenatal mice

In general, the KOH steps have to be adapted to the amount of tissue left. KOH digestion is stopped when the tissue is transparent. Then skeletons are dehydrated in 20%, 40% and 80% gelatin for at least 12 hours per step. They are stored in 80% gelatin at RT in the dark.

3.6.6 RNA-*in situ* Hybridization

The Method of Whole Mount RNA *in situ* Hybridization is used to make the expression pattern of single genes visible in whole embryos. Because of the variability of the staining intensity it is essential to include a wildtype control in each RNA-*in situ* which is handled the same as the mutant embryos.
For RNA-protection it is necessary to dissect embryos in PBS-DEPC and to fix them overnight

in 4%PFA-DEPC on 4℃. At the next day embryos are dehydrated on ice through a series of Methanol-PBST dilutions: 25%, 50%, 75% and two times 100% Methanol. Each step should last at least 10 minutes. From this time point on embryos can be stored in 100% Methanol on −20℃. Whole Mount *in situ* hybridization takes four to five days depending on the binding strengths of the probe and the expression level of the RNA.

Day 1 – Hybridization Embryos are rehydrated on ice through a series of Methanol-PBST dilutions: 75%, 50%, 25% and PBST-DEPC. The solutions should be exchanged when the embryos are sunken down. The embryos are washed for another 10 minutes in PBST-DEPC and are bleached in 6% hydrogen peroxide (in PBST) for one hour on ice. After three 10 minutes steps of washing in PBST embryos are digested with Proteinase K in Proteinase K buffer on a shaker at RT. The concentration of Proteinase K and the digestion time is depending on the embryonic stage. In this work only E11.5 mouse embryos have been used. This stage is digested in 10µg/ml PK for 3 minutes.
After digestion embryos are washed twice for five minutes in PBST/glycine, then twice for five minutes in PBST, followed by three five minute incubations in RIPA buffer. After three times washing in PBST for five minutes, embryos are fixed for 20 minutes in 4%PFA-DEPC with 0,2% Glutaraldehyd in PBS/0.1% Tween20. Now embryos are washed thrice in PBST for five minutes. Then embryos are incubated in Hybe-Buffer/PBST (1:1) until they settled down. If so, they are incubated for 10 minutes in Hybe-Buffer. Now, embryos are sorted for different probes in cryovials with Hyb-Buffer and kept at 65℃ for five minutes. Meanwhile the hybridization solution is prepared consisting of Hybe-Buffer containing 100µg/ml tRNA and DIG-labeled probe in a dilution of 1:1000. The hybridization solution is also heated up to 65℃. Then Hyb-Buffer is removed from the embryos and 1ml hybridization solution are added.The probe hybridizes overnight at 65℃. Hyb-Buffer and SSC/FA/T are also heated up to 65℃ overnight.

Day 2 – Removing of Unbound Probe After hybridization embryos are transferred to 12-well-plates with nets and washed twice in hot Hyb-Buffer at 65℃. Then embryos are kept at RT to cool down and are washed for five minutes in 1:1 Hybe-Buffer/RNAse Solution. Now RNA is removed at 37℃ by two 30 minute incubations in RNAse Solution containing 100µg/ml RNAseA. The embryos are then washed in 1:1 Solution of RNAse Solution and SSC/FA/T for five minutes at RT and heated up to 65℃ in SSC/FA/T. From now on embryos are washed in preheated SSC/FA/T – twice for five minutes, then thrice for ten minutes and six times for 20 minutes. When embryos are cooled down to RT they are washed for ten minutes in 1:1 solution of SSC/FA/T and 1xMABT and then incubated twice for ten minutes in 1xMABT. Meanwhile 10%BBR Blocking Solution is prepared by dissolving 5g BBR in 50ml 5xMABT at 50℃. During one hour of blocking preabsorbtion of the DIG antibody is performed by diluting it 1:5000 in 1%BBR-1xMABT and let mixing it on a rotating wheel for one hour at 4℃. Embryos are incubated in this antibody solution overnight at 4℃ on a shaker.

Day 3 – Removing of Unbound Antibodies Unbound antibodies are removed by several washing steps with PBST/Levamisole: The embryos are washed in this solution three times for five minutes, then eight times for one hour and are finally kept in PBST/Levamisole at 4℃ overnight.

Day 4 – Staining The embryos are washed thrice in alkaline phosphatase buffer for 20 minutes before they are transferred to a fresh 12well-microtiter plate with 1.5ml BM Purple AP Substrate to induce the staining-reaction. The staining reaction can take several hours depending on the probe and the expression level of the target gene. Sometimes it if necessary to prolong the incubation to the next day. For this purpose, embryos should be kept in alkaline phosphatase buffer overnight (4℃).

4 Results

4.1 CRISPR/Cas9 - a Tool for Rapid Genome Editing

Mouse models have been used for a long time to verify the pathogenicity of human mutations. Nevertheless, the generation of specific mutations by genome editing was until now time-consuming and cost-intensive. The CRISPR/Cas9 system allows efficient targeting of specific sequences with the nuclease Cas9 without any cloning effort [43, 42, 56, 57, 58]. Therefore, this technology was chosen to test if specific human mutations can be generated in a shorter time than usual. For this purpose, five human mutations that are associated with limb disorders has been selected to be recapitulated in mouse using CRISPR/Cas9 system (see Fig.16):

1. Nievergelt-like syndrome with deletion at *LAF4* locus
2. SHFM/SHFLD with duplication at *BHLHA9* locus
3. Synpolydactyly with duplication at *SHH* locus
4. Radial Aplasia with duplication at *TBX15* locus
5. SHFM with a point mutation within *FNDC3A*

To generate mutant mice murine ESC were targeted with CRISPR/Cas9. A protocol was established that allows to obtain ES cell with the desired mutation within seven weeks (see Fig.15).
The seven-weeks protocol includes one week for the design of sgRNA and genotyping Primer (for qualitative and quantitative PCR - see below) and one week for the cloning process including sequencing of the construct. Within further two weeks of cell culture ES cells are transfected with the CRISPR/Cas9 system and clones are picked and cryoconserved. In the fifth week genotyping of the obtained clones is performed - first with PCR and later with qPCR. In the following week positive clones are expanded. At week seven, clones can be aggregated. Mutant chimeric mice are born three weeks later can be either investigated for rapid verification of pathogenicity or they are used for matings to set up a mutant mouse line.

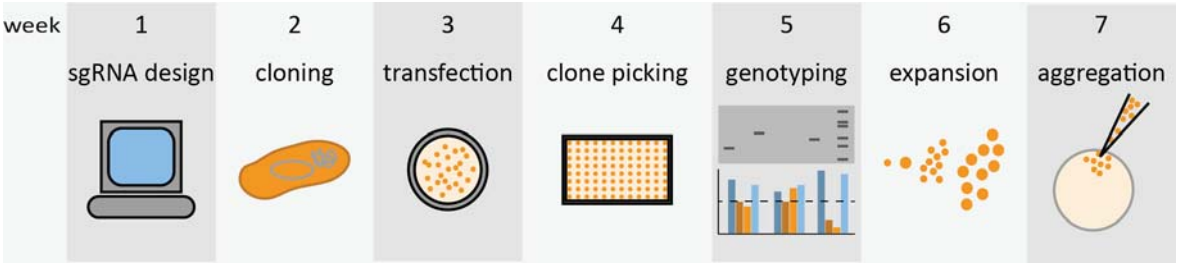

Figure 15: **Seven-weeks time frame for CRISPR/Cas9 application in murine ES-cells**

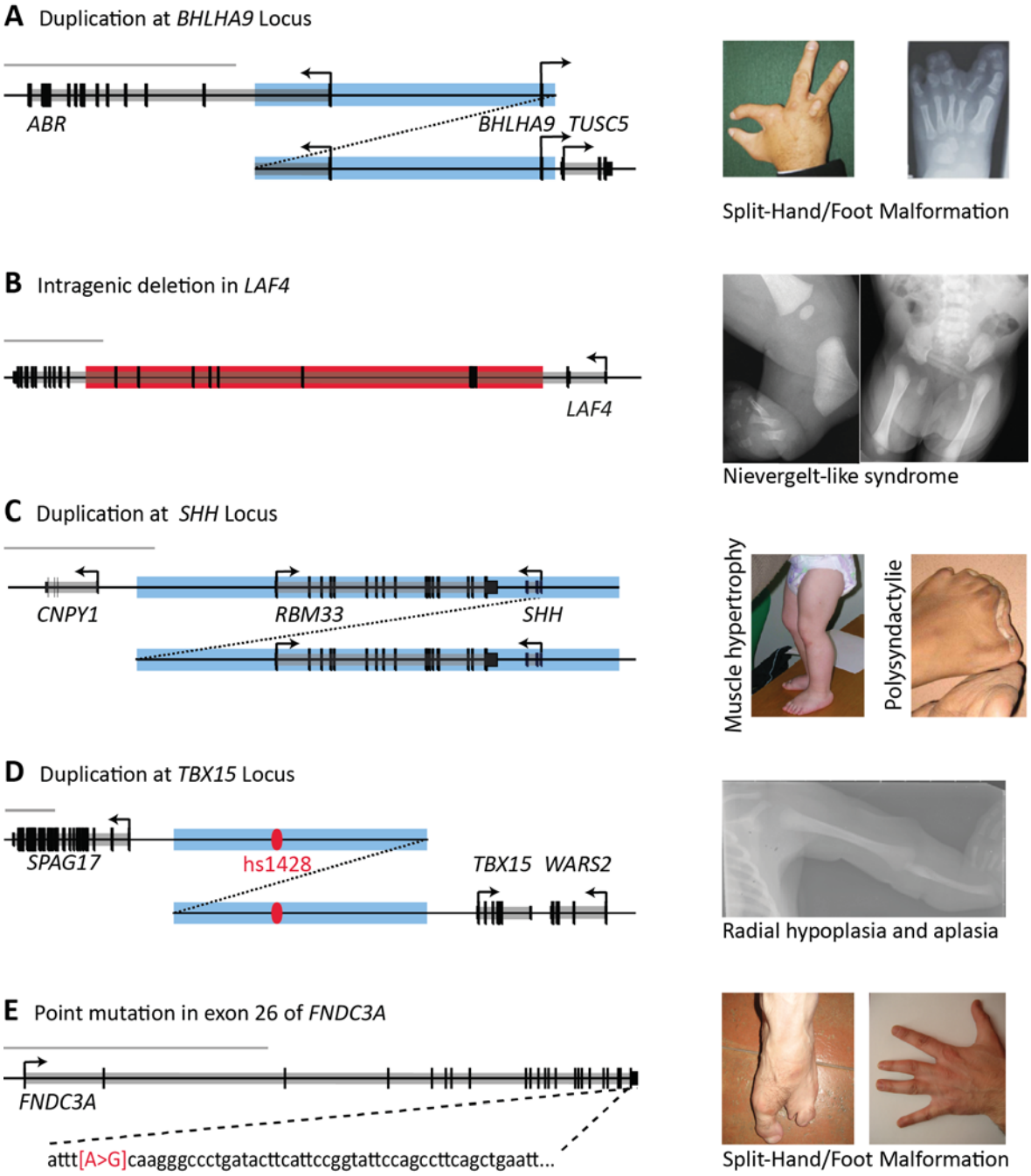

Figure 16: **Human cases for CRISPR/Cas9 application**
Five different human mutation had been genocopied in mouse: (A) A Duplication of *BHLHA9*, which is often found in SHFM and SHFLD, (B) an intragenic truncation of *LAF4* (see 4.2.3) in Nievergelt-like syndrome, (C) a duplication at *SHH* locus in a single case of synpolydactyly and a single case of muscle hypertrophy, and (D) a duplication of the limb enhancer hs1428 in a single case of radial aplasia and hypoplasia. (E) Point mutation in *FNDC3A* in a single case of SHFM. Photo of muscle hypertrophy patient is adapted from Kroeldrup et al. (2012) [125]. (bar = 100kb)

4.1.1 CRISPR/Cas9 System Induces Genomic Rearrangement in Murine Embryonic Stem Cells

For the generation of structural variations in ES cells two sgRNA were used, that guide Cas9 to the desired breakpoints (see Fig.17A). To genotype the obtained ESC clones, four break point specific primers were designed, that allow the detection of wildtype but also rearranged sequences. For each type of rearrangement a different combination of primers allows the amplification of a specific PCR product. Inversions create two distinguishable sequences at the 5-and 3-prime end respectively (see Fig.17A).
Genotyping of clones revealed that different kind of rearrangements can be achieved: deletions, duplications and inversions (see Fig.17B). Sanger sequencing of wildtype and rearrangements specific PCR products showed that clones with the same genotype hold minor differences at breakpoints in form of single base pair deletion or insertion (see Fig.17C). Interestingly, some clones are positive for more than two genotypes (e.g. deletion, inversion and duplication) or positive for just one inversion breakpoint (BP) (see Inv/Dup and Inv/Del in fig.17B). This indicates that not all rearrangements are complete. Therefore, genomic qPCRs were performed on positive clones. Figure 17D demonstrates the position of different primer sets within (red) and outside (blue) of the rearranged region. The latter are supposed to have a stable copy number while the rearranged region shows less copies for deletion or more for pure duplication. Noteworthy, clones with pure inversion or duplication with deletion do not encompass any copy number changes. The occurrence of pure duplications shows that additional copies can be gained. The qPCR for one of these duplication is shown in figure 17D. Here region 1 is present in four copies while region 3 has only three copies indicating that at least one of the duplications is not complete.
In conclusion, CRISPR/Cas9 enables induction of structure variations like deletions, duplications and inversions. Accurate detection requires rearrangement specific PCRs and qPCRs as well as sequencing of the breakpoints.

4.1.2 Efficiency of CRISPR/Cas9 Induced Rearrangements

In this work, four loci were targeted with two CRISPR/Cas9 sgRNAs to introduce rearrangements. The size of these rearrangement ranges from 154kb to 370 kb. To elucidate the effect of different sgRNA positioning in the same locus the *Bhlha9* locus was targeted with three different construct combination in three parallel experiments: 5/3 (sg5 and sg3), 5.1 (sg5n1 and sg3n) and 5.2 (sg5n2 and sg3n). All designs differ marginally in the genomic position of sgRNAs and define a similar region of around 160kb.
Table 29 summarizes the efficiency of all detectable rearrangements. The number of pure *wildtype* clones reflects the general efficiency of rearrangements. *Laf4* is the locus with the highest general efficiency, followed by *Shh* locus and *Bhlha9* 5.2 design. *Tbx15* and *Bhlha9* 5.1 design have a relative low targeting efficiency of around 20%. Targeting with the Bhlha9 5/3 guides was not successful and did not result in mutant ESC. This leads to the conclusion that efficiency is not or not exclusively depending on region size.
Of all listed genotypes only a few are of general interest, these are duplications, deletions and inversions with two mapped breakpoints. Detection of only one inversion break point hints

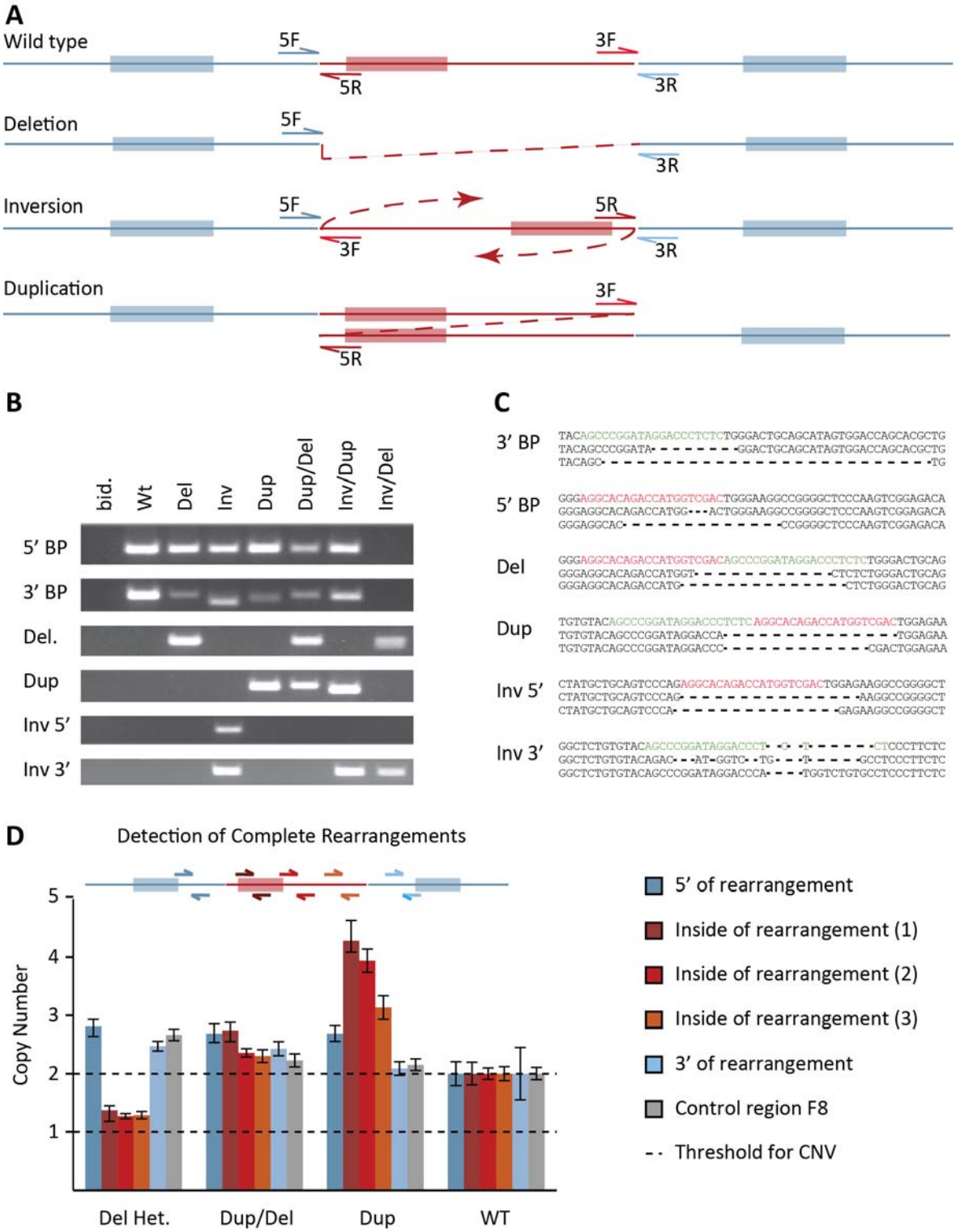

Figure 17: **Genotyping of CRISPR/Cas9 induced genomic rearrangements**
(A) The combination of two primer sets that map to the breakpoints enable the detection of deletions, inversions and duplications by PCR. (B) Genotyping of targeted ESC clones reveals different combination of rearrangement that all arise from one single experiment. (C) Sanger sequencing of wildtype PCR products as well as rearrangement specific amplification show different indels at the breakpoint region close to the guide sequence (5-prime sgRNA = red, 3-prime sgRNS = green). (D) Genomic qPCR inside and outside of targeted region is performed to distinguish complete from incomplete rearrangements. Gel pictures in (B) were provided by Guillaume Andrey (MPI for Molecular Genetics, Berlin).

Table 29: CRISPR/Cas9 targeting efficiencies [%]

Target Locus	Shh	Tbx15	Bhlha9 5/3	Bhlha9 5n1	Bhlha9 5n2	Laf4
Re-arranged region size [kb]	267	370	154	157	162	353
Deleted	10,4	5,2	0,0	2,1	7,2	13,2
Inverted						
2 breakpoints mapped	8,3	0,0	0,0	0,0	8,3	14,2
1 breakpoints mapped	19,7	1,0	0,0	1,0	5,2	6,9
Duplicated	3,0	0,0	0,0	0,0	3,1	28,1
Deleted/inverted						
2 breakpoints mapped	2,1	0,0	0,0	0,0	1,0	3,8
1 breakpoints mapped	5,2	0,0	0,0	0,0	2,1	3,1
Deleted/duplicated	2,1	11,0	0,0	0,0	1,0	6,2
wildtype	35,4	80,2	100	85,4	53,1	13,1
Targeting efficiency	64,6	19,8	0	14,4	46,1	86,9

towards incomplete rearrangements. The efficiency for different genotypes differ but a trend towards more deletion and inversion is seen while incomplete inversions are more common than complete ones.

Possibly locus dependent efficiency arises from the DNA accessibility and might be influence by heterochromatin. DNase hypersensitive sites indicate genomic position that can be targeted by nucleases and might be a marker for efficient CRISPR sgRNAs. To verify this hypothesis three different designed were applied: Bhlha9 5/3 (sg5 and sg3), Bhlha9 5n1 (sg5n1 and sg3n) and Bhlha9 5n2 (sg5n2 and sg3n). The guide location were chosen depending on DNaseI hypersensitive sides for ES cells that are listed in the UCSC track UW DNaseI HS (see Fig.18). Just the guides sg5n2 and sg3n are located within DNaseI hypersensitive sides. The guide sg3 is binding proximal 500bp upstream of sg3n. While the distance between sg5n2 and sg5 (3kb) or sg5n2 (8kb) is much higher. Interestingly, the Bhlha9-5n2 design leaded to the highest efficiency of around 50% while Bhlha9-5n1 achieved only 20% and Bhlha9 5/3 was not performing at all (see Tab.29). This suggest that positioning of sgRNAs in DNaseI hypersensitive sides might enhance successful CRISPR/Cas9 targeting.

4.1.3 CRISPR/Cas9 Enables Rapid Generation of Point Mutation

In a family case of SHFM a homozygous point mutation in exon 26 of FNDC3A was associated with SHFM (see 4.6). To verify if the point mutation is causing SHFM it was recapitulated in mice by targeting murine ESC with CRISPR/Cas9. To introduce the point mutation Cas9-generated double strand breaks were used to induces homology directed repair (HDR). The experimental procedure for introducing of point mutations with CRISPR/Cas9 is the same as for rearrangements (see 4.1.1) with the only different that only one sgRNA is needed

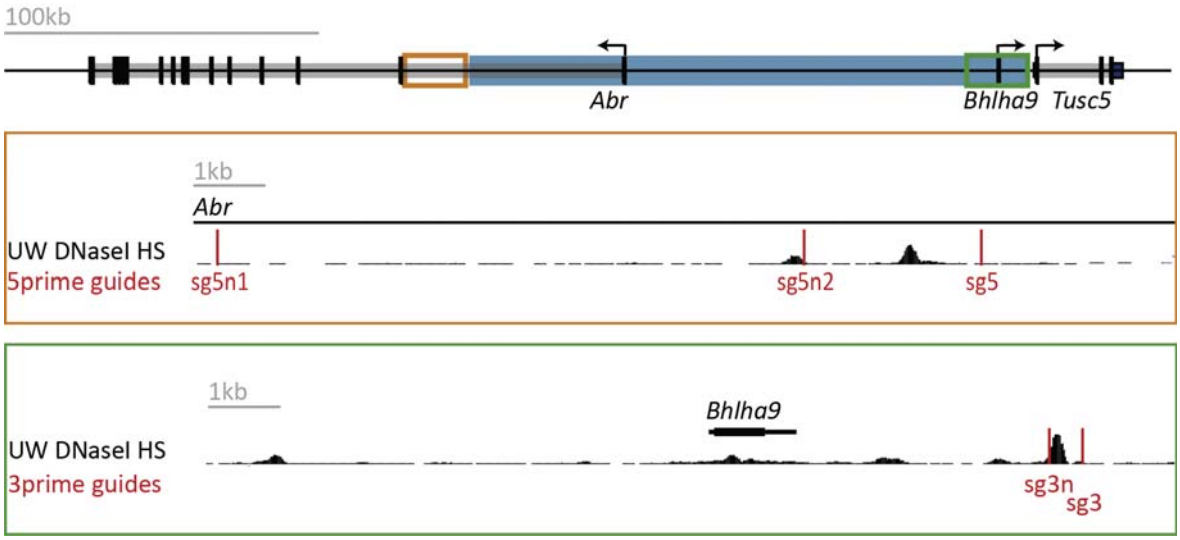

Figure 18: **Design of *Bhlha9* targeting in respect of DNase hypersensitive sites**
DNase hypersensitive sites at *Bhlha9* duplication (blue bar) and different guide sequences: For the duplication of *Bhlha9* three guides had been designed for the 5prime break point (5,5n1,5n2) and two for the 3-prime break point (3 and 3n1). A close up of the 5-prime region (red box) and the 3-prime region (green box) with the UW DNaseI HS track from UCSC browser shows that just the guide 3n and 5n2 are located at a DNase hypersensitive site

and that a template is co-transfected with the Cas9-plasmid (see Fig.19A). This template is a single stranded DNA oligo (ssODN) of 90bp that is reverse-complement to the guide and carries three mutations: the patient mutation, a PAM mutation to prevent Cas9 from cutting and a silent mutation to allow specific genotyping by PCR. The PAM is located close to the splice site of exon 26 and is therefore highly conserved. That is why three base pairs were replaced to generate another sequence that is also conserved between species (e.g. in rat). After incorporation of the ssODN via HDR the mutations are manifested by replication (see Fig.19A). With the use of the NHEJ inhibitor Scr7 one mutant clone was obtain out of 256. Without Scr7 no positive clones were found. In both cases more than 90% of the clones carry small indels at the break point.
Sequencing of the mutant clones showed that all three mutations are incorporated in the correct manner, but additionally a two base pairs deletion is located at the splice site (see Fig.19B). Sequencing of *Fndc3a* mRNA from the mutant clone proves correct transcription of wildtype and mutant allele with expected sequence at the exon junction (see Fig.19C). Therefore, it can be concluded that CRISPR/Cas9 allows rapid introducing of point mutations in ES cells.

4.2 Creating Mouse Model with Nievergelt Syndrome

4.2.1 *Laf4* Deficiency does not Interfere with Limb Development in Mouse

In 1.5.4 a case of Nievergelt-like syndrome is described which is characterized by complex malformation affecting skeleton, genitalia and neurological function. ArrayCGH analysis of the patient revealed a deletion of the unknown gene *LAF4*. Since *Laf4* is expressed in the developing limb bud of mice, it was assumed that the anomalies of the patient are caused by *Laf4* deficiency [108]. To verify the effect of *Laf4* deletion Chan Wing Lee (Charité -

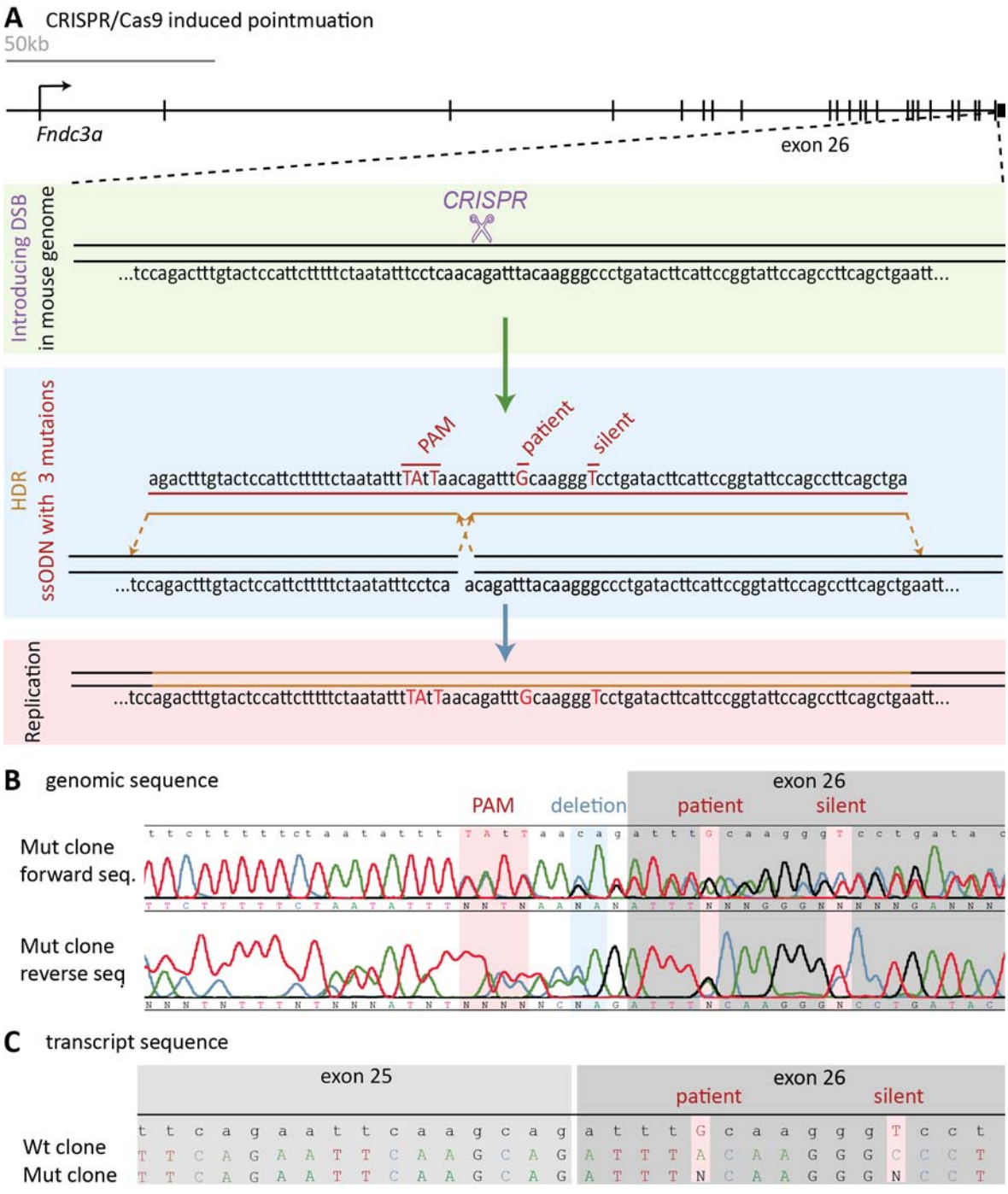

Figure 19: **CRISPR/Cas9 induced point mutation in *Fndc3a* does not interfere with splicing machinery**
(A) Together with the CRISPR/Cas9 plasmid a single stranded DNA oligo (ssODN) is transfected in ES cells. Cas9 is guided to the position of the patient point mutation, where it induces a double strand break. The ssODN serves as a template for homology directed repair (HDR) and thereby allows the incorporation of the patient mutation, the PAM mutation and a third silent mutation. Replication is necessary for manifestation on both strands. (B) Sequencing of the observed clones proves the correct localization of all three mutations (red bars) but also a 2bp deletion at the splice site (blue). (C) Sequencing of Fndc3a mRNA reveals that wildtype and mutant allele is transcribed with correct exon junction sequence.

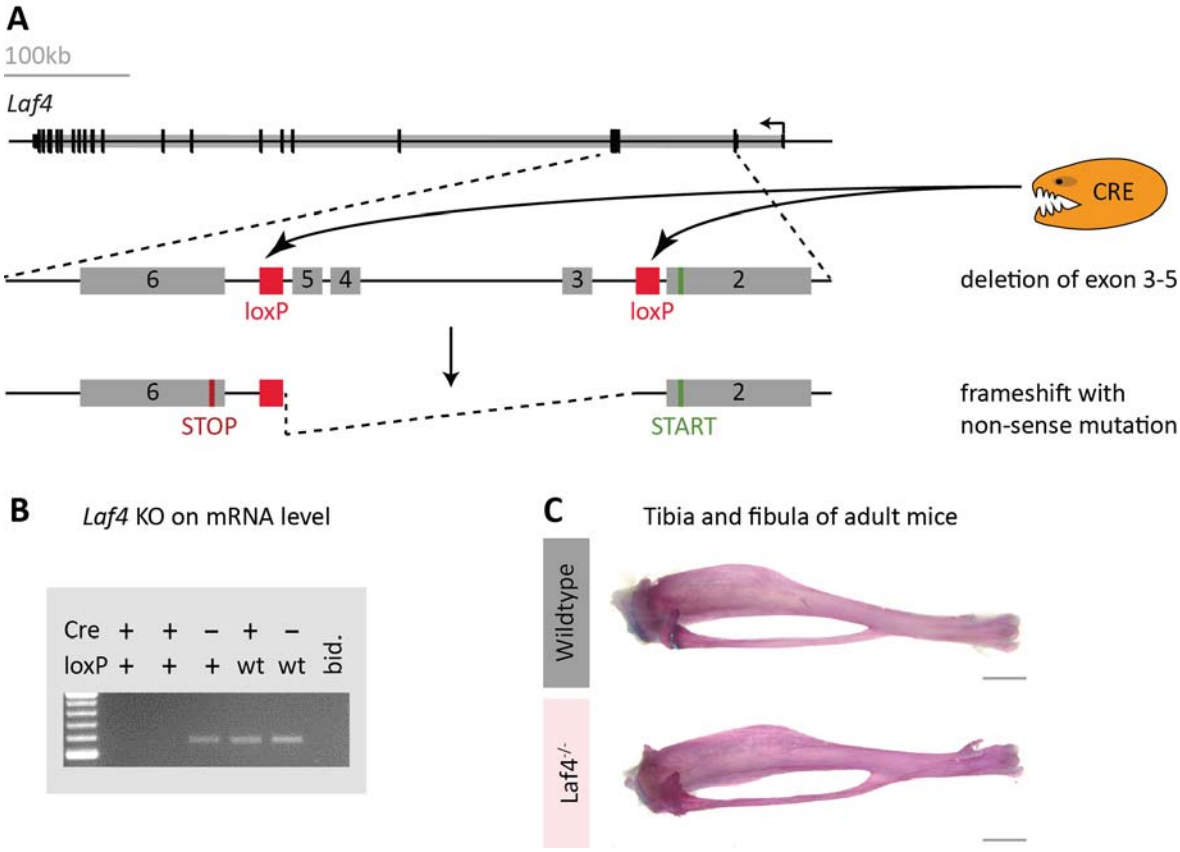

Figure 20: ***Laf4*-KO mice without phenotype**
(A) *Laf4*-KO mice were generated by introducing *loxP* sites and breeding with Cre-expressing mice. This caused deletion of exon 3-5 and a frameshift mutation that gives rise to a STOP codon in exon 6. (B) *Laf4*-specific PCR on cDNA reveal lack of *Laf4* mRNA. (C) Alizarin Red staining of homozygous mice shows no morphological changes of tibia and fibula (bar = 2mm). KO-Mice were designed and cloned by Chan Wing Lee (Charité - Universitätsmedizin Berlin).

Universitätmedizin Berlin) generated a *Laf4* Knock-out (KO) mouse line. For this purpose mice with *loxP* sites that encase exon three to five (Laf4-floxed) were generated and bred with a CMV:Cre line. In CMV:Cre mice expression of Cre-Recombinase is regulated by the ubiquitous active CMV promoter. Therefore, in double mutants (Laf4-floxed;CMV:Cre) Cre recombinase is deleting exon three to five of *Laf4* which causes a frameshift mutation that gives rise to a STOP codon in exon 6 (see Fig.20A). To confirm the knock-out *Laf4*-transcript specific PCR were performed. Only homozygous *Laf4*-floxed mice that are transgenic for *Cre* were negative for this PCR while a product could be obtained from wildtype animal, *Cre*-transgenic mice and *Laf4*-floxed without Cre-ORF (see Fig.20B). This proves absence of *Laf4* transcripts on RNA level in homozygous *Laf4*-KO mice. Skeletal preparation of these mice revealed no skeletal changes, tibia and fibula morphology was normal (see Fig.20C). Therewith, it can be concluded that *Laf4*-deficiency is not causing Niervergelt-like phenotype or any other obvious anomalies in mouse.

4.2.2 *LAF4* is Truncated in Patient with Nievergelt-like Syndrome

Since *Laf4*-KO mice do not develop Nievergelt syndrome, the genotype of the index patient was investigated more closely. *LAF4* deletion was original detected with BAC-arrayCGH.

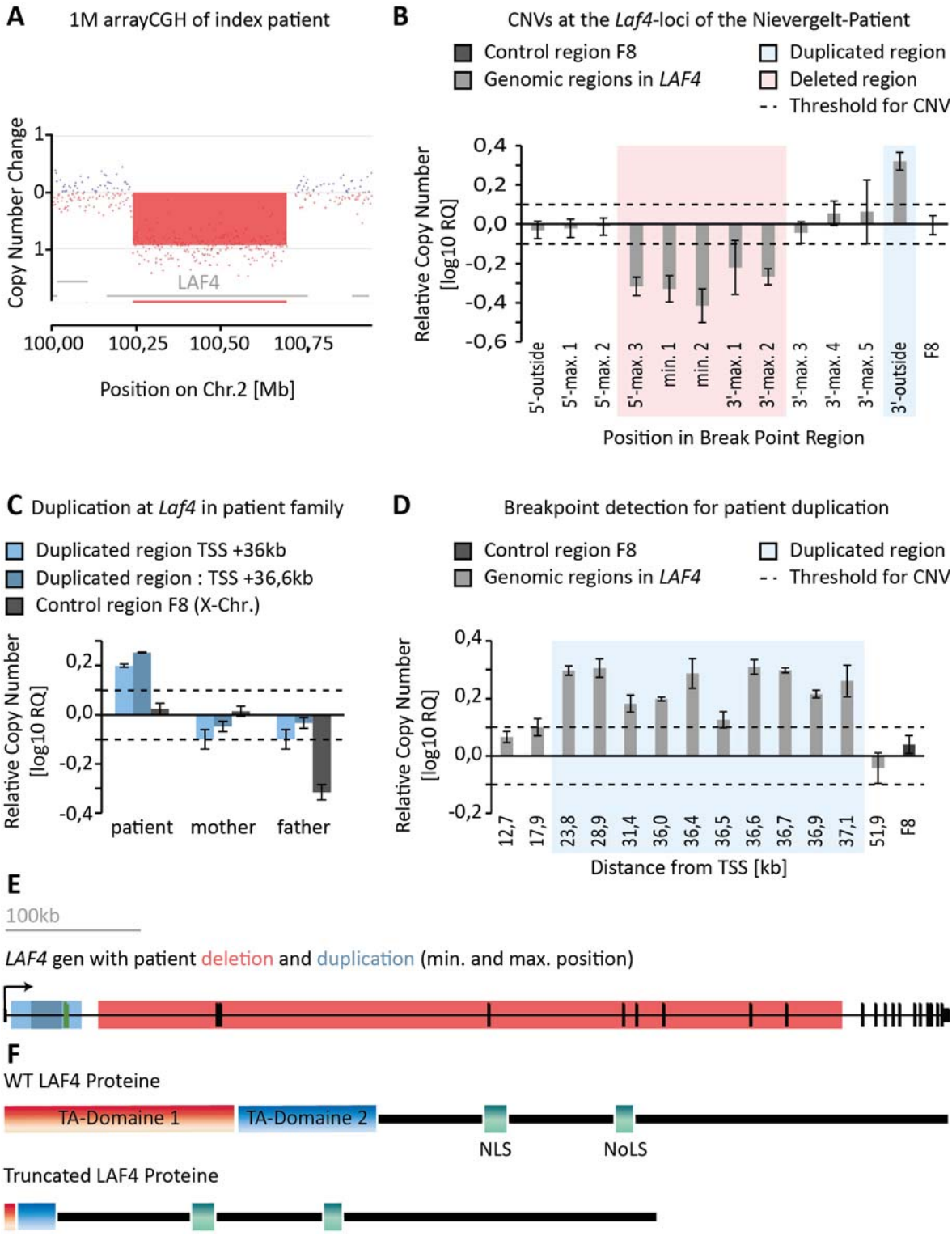

Figure 21: **Nievergelt-like disorder - patient genotype**
(A) 1M arrayCGH reveals deletion within *LAF4* but does not involve the whole gene. (B) Genomic qPCR of the patient recapitulates arrayCGH findings and detects additionally duplication. (C) Comparison of genomic qPCR of the patient and her parents shows that the duplication is *de novo*. (D) The breakpoints of the duplication was narrowed down by genomic qPCR and lays upstream of transcriptional start side (TSS) with minimal position TSS+23.8-37.1kb and maximal position TSS+17.9-51kb. (E) Mutated *LAF4* allele carries two mutations: a large deletion within the ORF (red) and a duplication (blue). Maximal duplication (bright blue) size includes the translation start (green). Minimal duplicated region (dark blue) is located between exon 1 and 2 without touching translation start. (F) The intragenic deletion is in frame and possibly give rise to a truncated LAF4 protein in which transcription activating domains are almost completely lost. Nucleus and nucleoli location signals (NLS and NoLS) are not affected. 1M-array was performed in the Array-group at Charité - Universitätmedizin Berlin and analyzed by Malte Spielman (Charité - Universitätmedizin Berlin).

The more precise 1M-array elucidated a much smaller deletion that is located within *LAF4* but does not involve the whole gene (see Fig.21A). To define the array result genomic qPCRs were performed (see Fig.21B) with primer sets located within the minimal (min.1 and min.2) and the maximal deleted region (5'/3'-max1-4) and outside of the detected deletion (5'- and 3'-outside). In this way it was elucidated that the deletion is spanning exon three to eleven and that the breakpoints are located within intron 2 and intron 11, respectively. The translation start site is not affected by the deletion (see Fig.21D). Since the codon order is not affected the translation of a truncated protein is possible. The transcription activating domains of this mutant protein are almost deleted and the left overs are fused. The nucleus and nucleoli localization sides (NLS and NoLS) are unaffected (see Fig.21F).
Additionally, to the deletion a *de-novo* duplication was detected at the start of *LAF4* (see Fig.21C). The duplicated region is located upstream of the transcriptional start side (TSS) and was narrowed down by qPCR (see Fig.21C). The minimal duplicated region is located within the first intron starting 23,8kb behind TSS and ending 73,1kb downstream of TSS. The maximal duplicated region includes the translational start side and ranges from TSS+17,9kb to TSS+51,5kb.
Thus, it can be concluded that the patient with Nievergelt-like disorder is carrying a smaller deletion than expected that encompasses 500kb and includes exon 3-11 of *LAF4* which possibly give rise to a truncated LAF4 protein (see Fig.21E).

4.2.3 Truncation of *Laf4* Results in Nievergelt Syndrome in Mouse

To verify if the intragenic deletion in *LAF4* causes Nievergelt-like disorder the same mutation was recapitulate in mouse by guided targeting of murine ESC with CRISPR/Cas9. Mutant mice heterozygous and homozygous for the human truncation of *Laf4* were obtained from aggregation after CRIPSR/Cas9 induced deletion (see above and Fig.22A). The expression of the mutant *Laf4* gene was detected in mutant ES cells on mRNA level by amplification of a product spanning the sequence from exon 1 to exon 12. The PCR product can not be amplified from wildtype templates or genomic DNA from mutant ES cells (see Fig.22B).
$Laf4^{trunc/wt}$ mice were born with low chimerism and unable to give birth to heterozygous mice suggesting that the mutation is lethal. As a result, all data presented arises from chimeric animals.

RNA whole mount insitu hybridization (WISH) of *Laf4* in E11.5 showed identical expression pattern in wildtype and $Laf4^{trunc/trunc}$ mice (see Fig.23). Therefore, it can be concluded that the truncated *Laf4* gene is expressed in the same pattern as the wildtype version.
Heterozygous and homozygous mice were obtained and analyzed in E18.5. In this stage all homozygous mice had reduced body size, club feet at hind limbs and mesomelic dysplasia. Polydactyly of hind limbs and intestinal prolapse were frequent. Heterozygous mice encompassed higher variability. Noteworthy, more severe affected heterozygous mice had the same phenotype as homozygous mice while in less severe cases animals had normal body size and never polydactyly. Interestingly, $Laf4^{trunc/trunc}$ mice did not response to muscle stimulation while wildtype animals showed muscle contraction, suggesting a neurological phenotype sim-

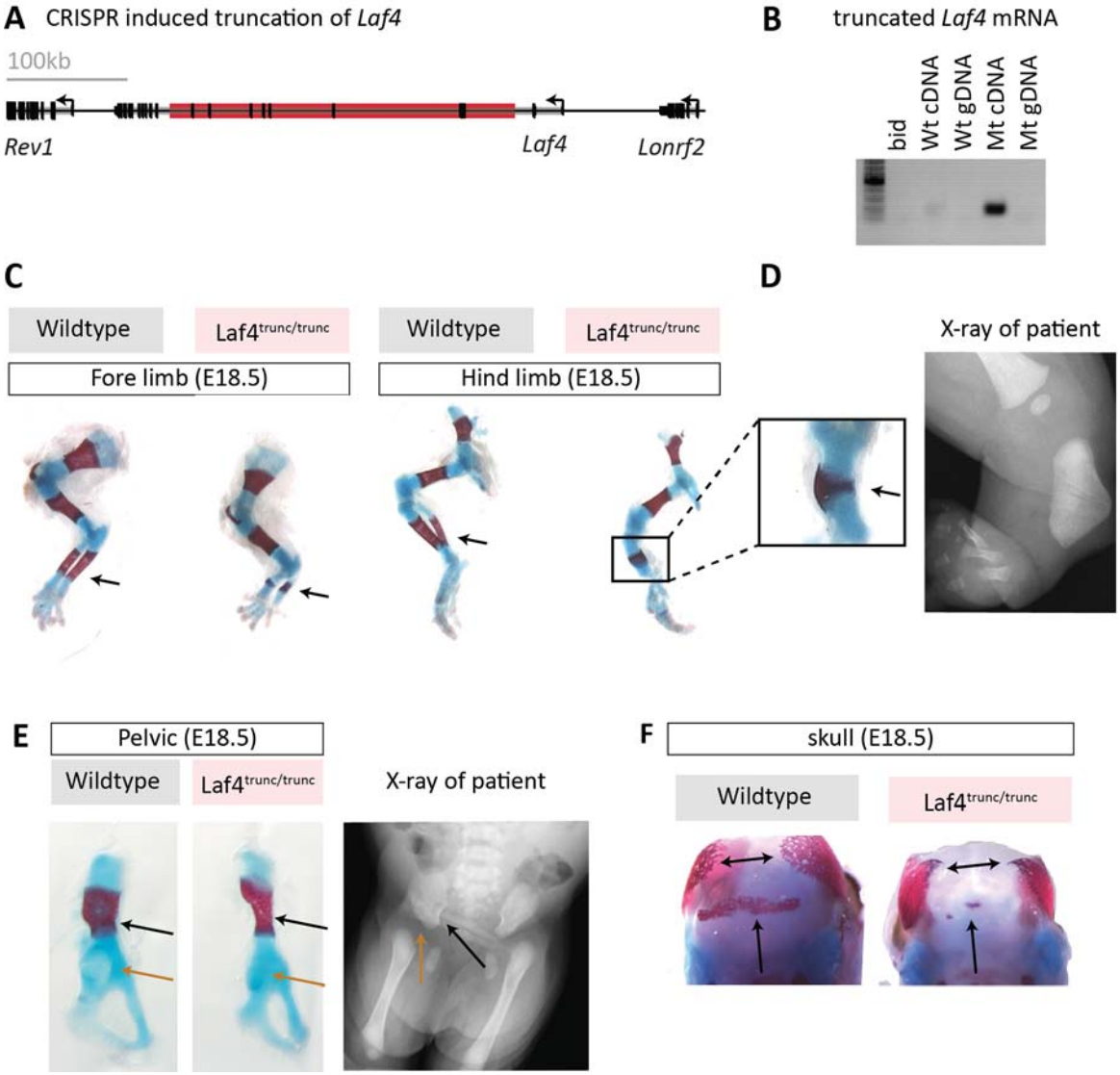

Figure 22: ***Laf4*-truncation induces Nievergelt syndrome in mouse**
(A) With the CRISPR/Cas9 system the patient deletion was generated in ESC (red bar). (B)Expression of truncated *Laf4* transcript was detected on mRNA-level by PCR. (C) Skeletal staining of E18.5 embryos shows general size reduction of fore and hind limb and uncover triangulate ossification center in radius and tibia (arrows) (D) which were also found in x-rays of the patient. (E) As in the patient the pelvic of mutant mice is hypoplastic, especially noticeable in iliac wing (black arrow) and acetabulum (orange arrow). (F) Delayed ossification of flat bones in the skull

ilar to the patient.

Skeletal preparation at E18.5 revealed severe limb malformation in heterozygous and homozygous mice (see Fig.22C and D). The stylopod was reduced in length and radius and tibia were marked by triangular ossification center (see Fig.22C) and recapitulated thereby the patient phenotype (see X-ray in fig.22D). Fibula was dys- or aplastic in homozygous animals, while femur, humerus and tibia were broader. Ankle joints of the hind limb, elbow and knee were malformed. Interestingly, the right hind limb was always more sever affected while triangular shape was sometimes not visible on the left side. The pelvis of mutants was hypoplastic as in the patient (see Fig.22E). This was especially prominent for the acetabulum (orange arrow) and iliac wing (black arrow). Additional to the limb phenotype bones of the skull were delayed in ossification (see Fig.22F).

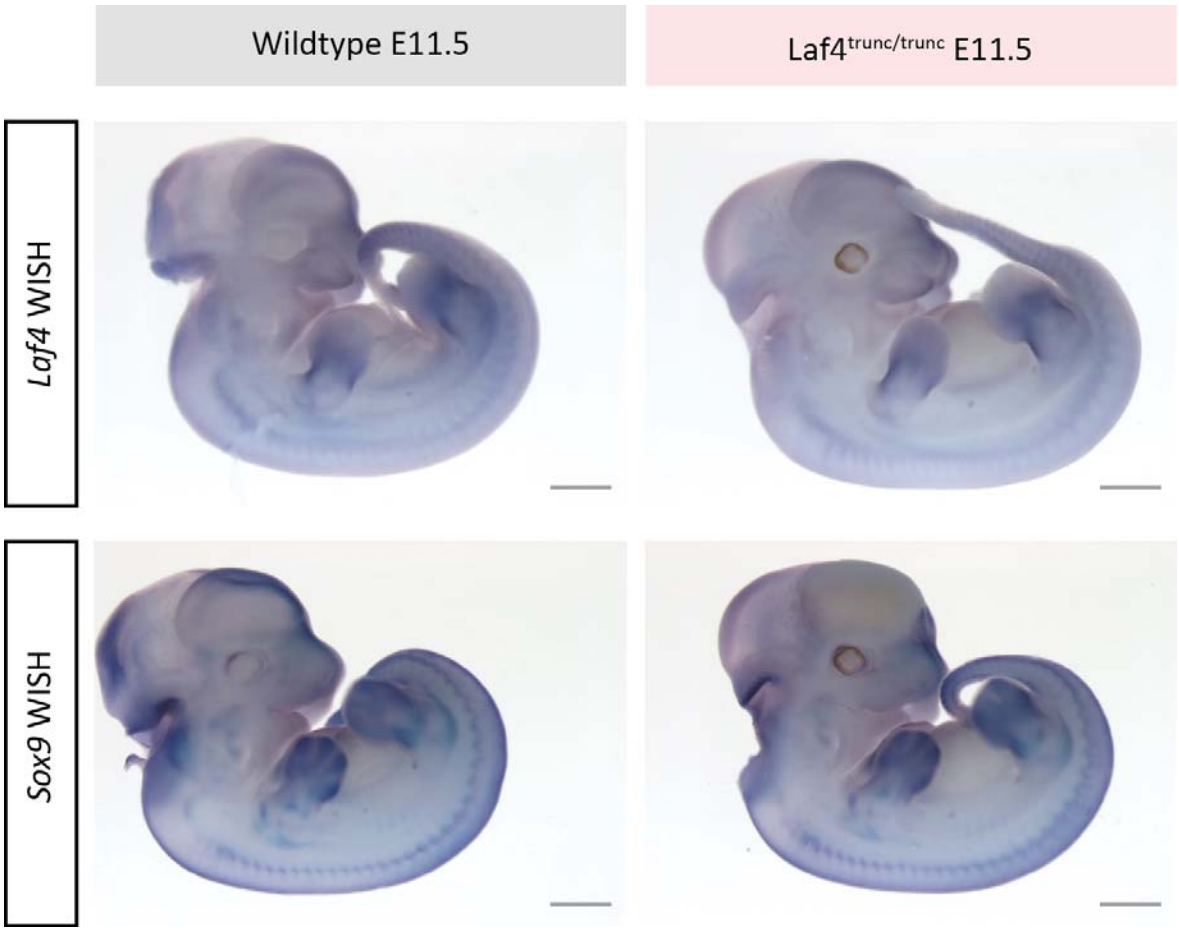

Figure 23: **Laf4-truncation does not change expression pattern of Laf4 and Sox9**
RNA Whole Mount *insitu* shows no expression pattern change of *Laf4* and *Sox9* in $Laf4^{trunc/trunc}$ E11.5 embryos in comparison to wildtype. (bar = 1mm)

It can be concluded that *Laf4*-truncation leads in mouse to the same Nievergelt-like phenotype as in the patient, which is even extended by triangular shaped radius. Furthermore, hints for neurological dysfunction can be observed but were not further investigated.

To elucidate which developmental process is involved in Nievergelt-like phenotype expression pattern of *Sox9* was checked in E11.5 embryos and histological bone structure were analyzed in E16.5. RNA-WISH of *Sox9* shows no changes in expression pattern between wildtype and $Laf4^{trunc/trunc}$ suggesting correct establishment of bone anlage (see Fig.23).
To visualize the histology of the tibia Masson-Goldner/Trichrome staining was performed on hind limb sections (see Fig.24A). $Laf4^{trunc/trunc}$ tibia was triangular shaped, wider and shorter. The smaller ossification center (orange) and less mineralization (green) suggests general delay in bone development in the mutant. Interestingly, the concave side of the triangular tibia showed less mineralization than the convex side which leaded to incomplete or disrupted appearance of the cortical bone. To measure the positioning of the growth plate the orientation of the growth plate was compared with the axis that is connecting the distal and proximal epiphysis. Indeed, the ankle between the epiphyseal axis and the growth plate border is reduced to 68° in $Laf4^{trunc/trunc}$ mice in comparison to 81° in wildtype animals. DIC (Differential Interference Contrast) imaging was perform to visualize cell morphology

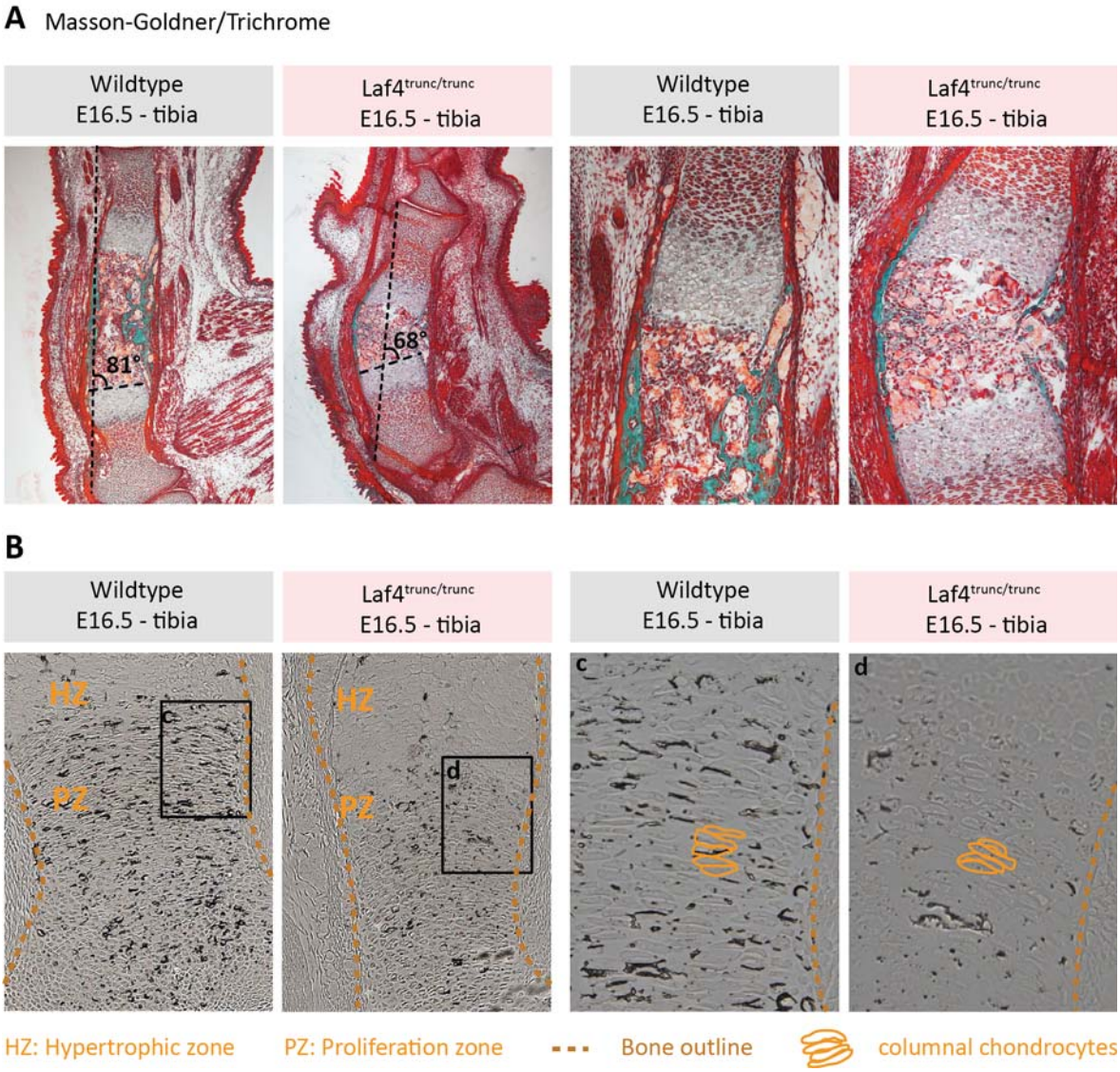

Figure 24: **Histology of *Laf4*$^{trunc/trunc}$ reveals dislocation of growth plate**
Histology was performed of E16.5 limb sections. (A) Masson Goldner / Trichrome staining shows calcified bone (green), hypertrophic zone (light grey), followed by columnar chondrocytes (dark red) and round proliferating chondrocytes (red-orange). *Laf4*$^{trunc/trunc}$ tibia show a general delay in bone development, triangular shape, broadening and longitudinal shortening. The black line marks the orientation of epiphyses and growth plate, respective. The ankle between these axes changes from 81° in wildtype to 68° in *Laf4*$^{trunc/trunc}$ tibia. The cortical bone at the concave side of *Laf4*$^{trunc/trunc}$ tibia is disruption (arrow). (B) DIC microscopy of unstained tibia sections shows normal morphology and orientation of columnar stack.

in tibia section, which is unchanged in mutant mice (see Fig.24B). In the proliferation zone (PZ) stacks of columnar chondrocytes were formed correctly.

In conclusion *Laf4* truncation leads to the same Nievergelt-like disorder in mouse as in human by affecting the orientation of the growth plate of the tibia without obvious effect on the development of bone anlage and bone tissue.

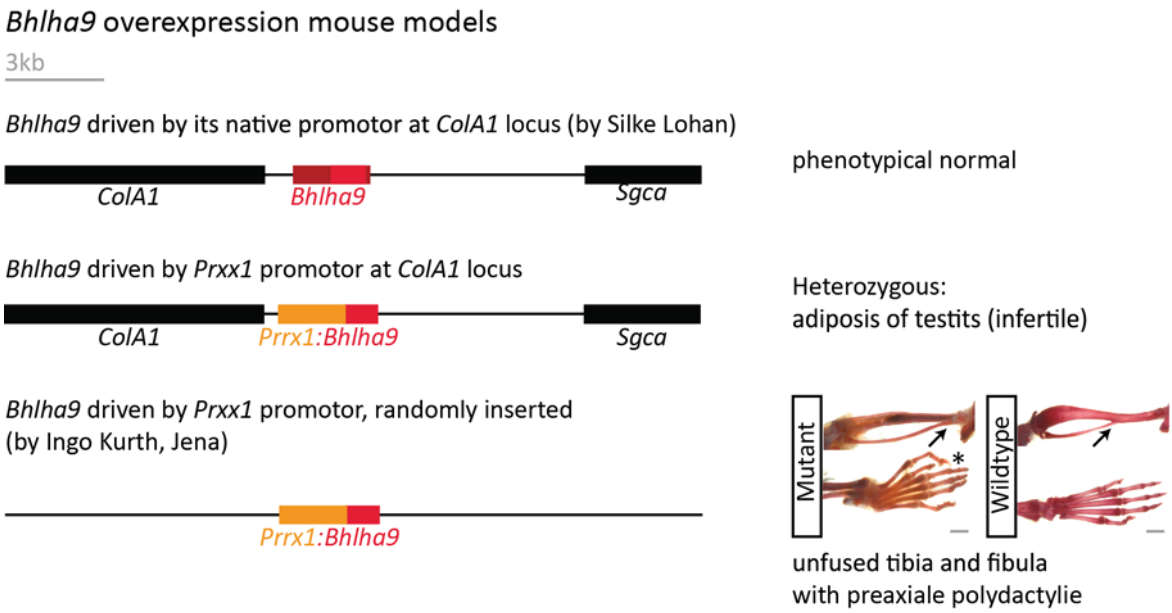

Figure 25: ***Bhlha9* overexpression mouse models do not develop SHFM**
Additional *Bhlha9* gens had been integrated into the mouse genome to induce Bhlha9 overexpression. Via the flip-in system (see 1.2) constructs were inserted at the *Cola1* locus, which drive Bhlha9 expression either with the native promoter (by Silke Lohan, Charité Universitätsmedizin Berlin) or *Prrx1* promoter (*Prrx1:Bhlha9*). Both mice do not show any skeletal phenotype. *Prrx1:Bhlha9* mutant male are infertile due to adiposis of testis. Therefore, only heterozygous chimeras had been investigated. Random insertion of *Prrx1:Bhlha9* (generated by Ingo Kurth, Jena) causes preaxial polydactyly with additional triphalangeal digit (asterisk) and unfused tibia and fibula (arrow).

4.3 Limb Development is Sensitive to *Bhlha9* Deficiency but Duplication is not Causing SHFM in Mouse

BHLHA9 duplications are associated with SHFM and SHFLD (see 1.5.2). To elucidate the effect of high *BHLHA9* dosage three different mouse models were generated before the CRISPR/Cas9 system was available (see Fig.25). The lines have in common that they carrier an additional *Bhlha9* gene outside of its native locus. Silke Lohan (Charité Universitätsmedizin Berlin) introduced a construct in the *ColA1*-locus (flip-in system, see 1.2), in which *Bhlha9* expression is driven by its potential native promoter. These mice did not show any abnormal phenotype. To investigate if increased *Bhlha9* overexpression can cause SHFM, a construct was inserted at the same position which allows to drive *Bhlha9* expression via the strong mesenchymal promoter of *Prrx1*. However, this did not lead to any skeletal phenotype in heterozygous chimeric mice. Homozygous or non-chimeric mice could not be obtained because of infertile males with adipose testis. In the third mouse model (generated by Ingo Kurt, University of Jena) the same construct was inserted at a random unknown position. Most of these mice had preaxiale polydactyly at the hind limb with an additional triphalangeal digit and unfused tibia and fibula. In the fore limb synostosis of carpal bones was present (data not shown). Genomic copy analysis of these mice suggests that they carried in total four *Bhlha9* copies at different loci (data no shown). Unfortunately, the insertion points could not be mapped. In summary overexpression of *Bhlha9* with insertion outside of its native locus is not able to produce SHFM in these mice.

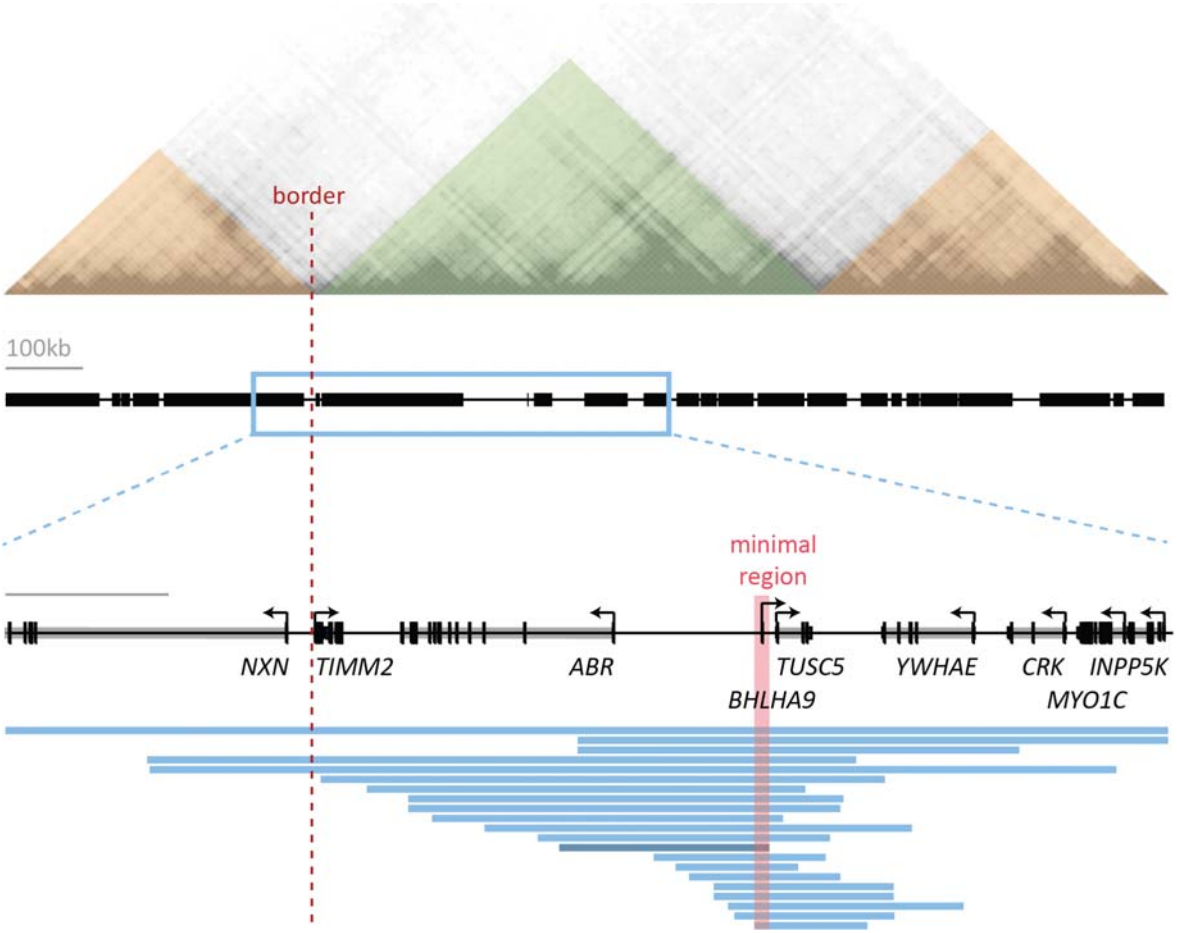

Figure 26: **Overview of the human *BHLHA9* locus (hg19)**
Different duplications were found in SHFM/SHFLD patients in DECIPHER data base (blue bars). The red line marks the region of minimal overlaps that includes *BHLHA9*. The TAD structure (colored triangles; Hi-C data from [33]) of the overview of this locus reveals a boundary 5-prime of *BHLHA9* that is crossed by some duplications. The duplication that has been genocopied in this work is marked dark blue.

Figure 26 gives an overview of the human *BHLHA9* locus (hg19) including TADs that are marked with colored triangles. The Hi-C data were obtained from the Hi-C browser from Ren Lab [33]. Positions of duplications in SHFM/SHFLD cases (blue bars) were extracted from DECIPHER. Some of these duplications include a TAD border at the 5'prime end of *BHLHA9* locus while most duplication are located within the *BHLHA9*-TAD (in green). The minimal region of all duplication is marked in red and includes just one gene - *BHLHA9*. To rule out that a positional effect is causing SHFM in human the smallest duplication (Hg19 Chr.17:1.050.815-1.179.926) with sequenced breakpoints (marked in dark blue) was selected for recapitulation in mouse with CRISPR/Cas9.
CRISPR/Cas9 targeting of the murine *Bhlha9* locus gave raise to ES cells that carried a duplication on one allele while the same region was deleted on the second one. Mice obtained from these cells were bred with wildtype animals to separate the deletion from the duplication. In the following a deletion and a duplication line was established and bred to homozygosity. Heterozygous animals for both genotypes are always normal, while homozygous deletion leads to syndactyly with incomplete penetrance and variable expressivity (see Fig.27A) as it was already shown for *Bhlha9* deletions that do not involve adjacent regions

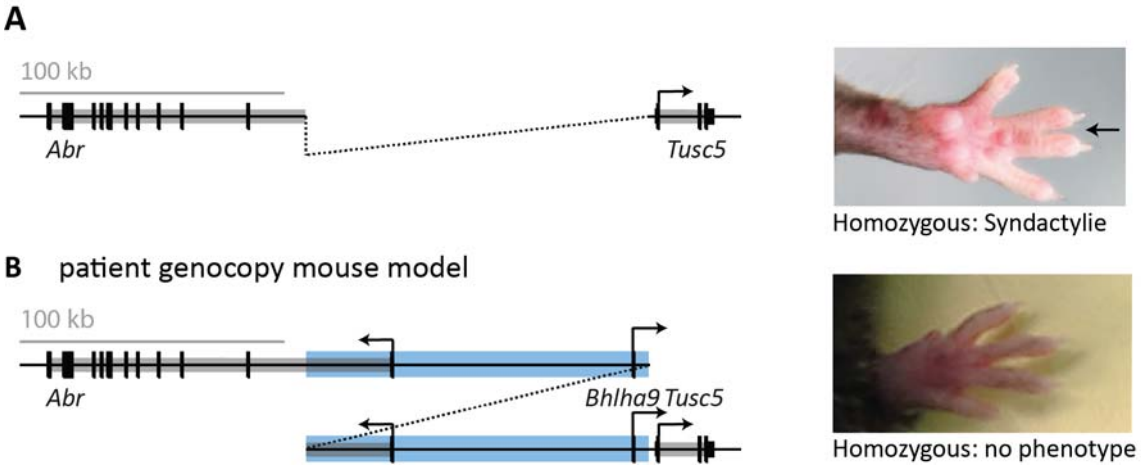

Figure 27: **Phenotypes of *Bhlha9* duplication and deletion**
Targeting murine Bhlha9 locus with CRISPR/Cas9 resulted in a duplication and deletion at one allele, respectively. (A) Homozygous deletion leads to syndactyly with incomplete penetrance. (B) Homozygous duplication results in no phenotype.

[104]. Duplication of *Bhlha9* did not induce any abnormalities in mouse (see Fig.27B) and it must be concluded that recapitulation of the patient duplication did not result in SHFM phenotype.

4.4 Duplication at the *SHH* Locus is Causing Hypertrophy

Two identical 299kb duplication at the *SHH* locus (Hg19 Chr.7: 155.351.052-155.650.275) were found in two different patients (see Fig.28A). One patient (analyzed by the human genetics apartment of Charité Universtätsmedizin Berlin) had synpolydactyly with cup shaped hands and feet, the other patient was published by Kroedrup et al. (2012) and had muscle hypertrophy, which was most severe at the stylopod. The duplication of the patient with muscle hypertrophy is in-tandem [125] while the duplication of the synpolydactyly case could not be mapped since no patient tissue was available. The duplication includes two genes: *SHH* and *RBM33* (*RNA Binding Protein 33*). Analysis of Hi-C data (see Fig.28A) [33] shows that the TADs of these genes (red and green triangle) are separated by a boundary element (red dash line) that is included in the duplication. Recapitulation of the duplication in mice with CRIPSR/Cas9 resulted in chimeric animals with duplication on one allele that arose from a deletion on the other allele (= Shh-dup/del). Therefore, mice heterozygous for the deletion (Shh-del) were used as control together with wildtype animals.
Shh-dup/del mice had a broader appearance in comparison to Shh-del mice (see Fig.28B). Further phenotypes are marked by incomplete penetrance and were only apparent in three out of nine Shh-dup/del mutants of the same litter. These effected mutants had enlarged brown fat tissue between shoulder blades with partly darkened color (see arrow in Fig.28B and C). All animal with these fat tissue anomalies had additional broader upper fore limb (see Fig.28D). Skeletal staining of the fore limb showed no difference between mutants and wildtype (see Fig.28E). This suggests that the observed enlargement is due to hypertrophy of soft tissue for instance muscles.
RNA WISH for *Shh* and *Rbm33* showed no difference in expression pattern in E11.5 embryos

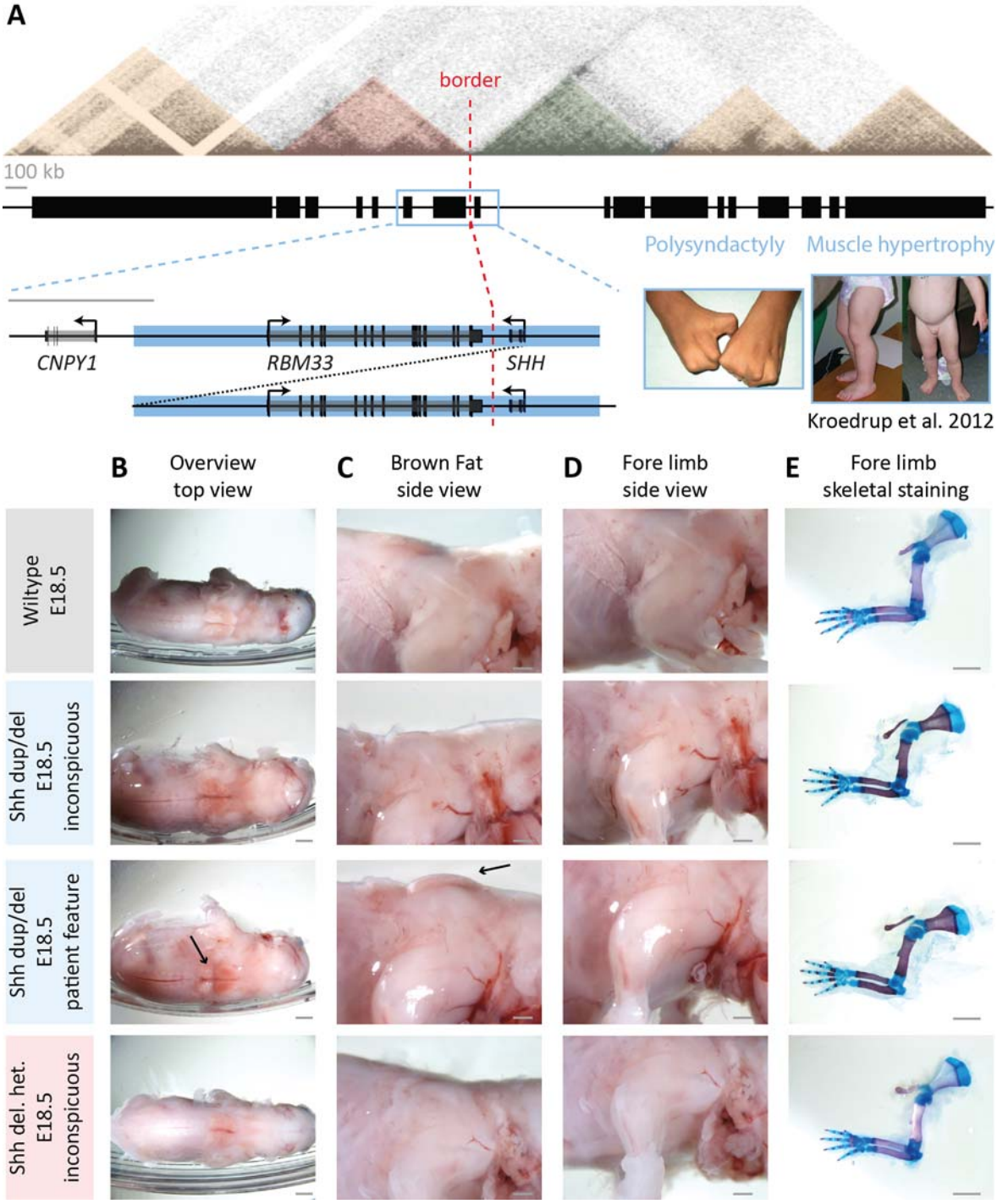

Figure 28: **Duplication at the *SHH* locus including *SHH* in human muscle hypertrophy and polysyndactyly**
(A) Polysyndacytly with cup shaped like appearance and muscle hypertrophy (latter was published by Kroeldrup et al. 2012 [125]). Overview of the TAD structure show a border between the TAD of *SHH* (green triangle) and the one of the neighboring gene *RBM33* (red triangle). This border (red dash line) is included in the duplication (blue bars) (Hi-C data from [33]). *Shh* duplication mice that are chimeric for heterozygous duplication and deletion (dup/del) have (B) a heavy stature that distinguish them from wildtype and mice with heterozygous deletion. Some Shh dup/del mice have enlarged brown fat (arrow) (bar = 2mm), which results in a bent like posture in side view (C) (bar = 1mm). (D) These animal also show broader upper fore limb (bar = 1mm). (E) Skeletal stainings of fore limbs show no difference in bone morphology. (bar = 2mm)

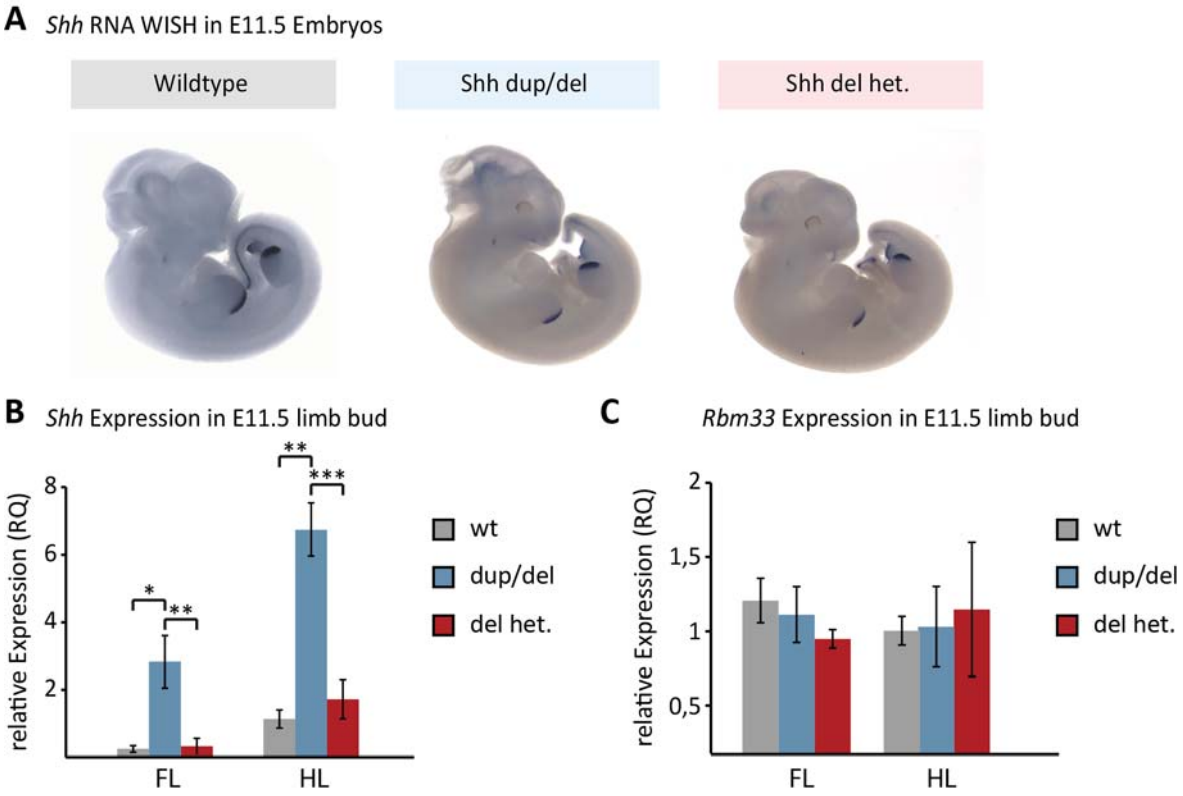

Figure 29: ***Shh* is upregulated but not misexpressed in Shh dup/del mice**
Shh and *Rbm33* expression in chimeric Shh dup/del and heterozygous Shh del mice (stage E11.5) show (A) no pattern different in the RNA WISH of *Shh* at *Rbm33* but (B) significant upregulation of *Shh* in fore and hind limb bud of Shh dup/del animals. (C) *Rbm33* expression level in unchanged.

(see Fig.29A) but qPCR detected *Shh* upregulation. Its expression level was significant sevenfold increased while *Rbm33* expression was unchanged (see Fig.29B).
In conclusion tandem duplication of *Shh* results in its upregulation with unchanged expression pattern and most likely causes muscle hypertrophy and not polysyndactyly.

4.5 Duplication at the *TBX15* Loci

A 522kb duplication at the locus of *TBX15* (Hg19 Chr.1: 118.806.296-119.328.333) was found in a patient with radial hypoplasia and aplasia (analyzed by the human genetics apartment of Charité Universtätsmedizin Berlin) (see X-ray in Fig.30B). Analysis of Hi-C data [33] showed that the duplication is located within the *TBX15* TAD and does not involve any boundary element (see Fig.30A). *TBX15* itself is not affected by the duplication but the enhancer hs1428. Data from the Vista Enhancer Browser were used to analyze the activity of this enhancer. Indeed, the data show that the enhancer is able to drive limb bud mesenchyme specific LacZ expression (see Fig.30C).
To reveal the pathogenicity of the duplication chimeric mice with the corresponding duplication were generated with CRISPR. This duplication arose from a translocation from the other allele, which in results carries a deletion. Therefore, the mutant hs1428 dup/del mice are compared to mice with heterozygous deletion (hs1428 del) of the same area. For both mutants - hs1428 dup/del and hs1428 del - no skeletal or other abnormalities were obtained.

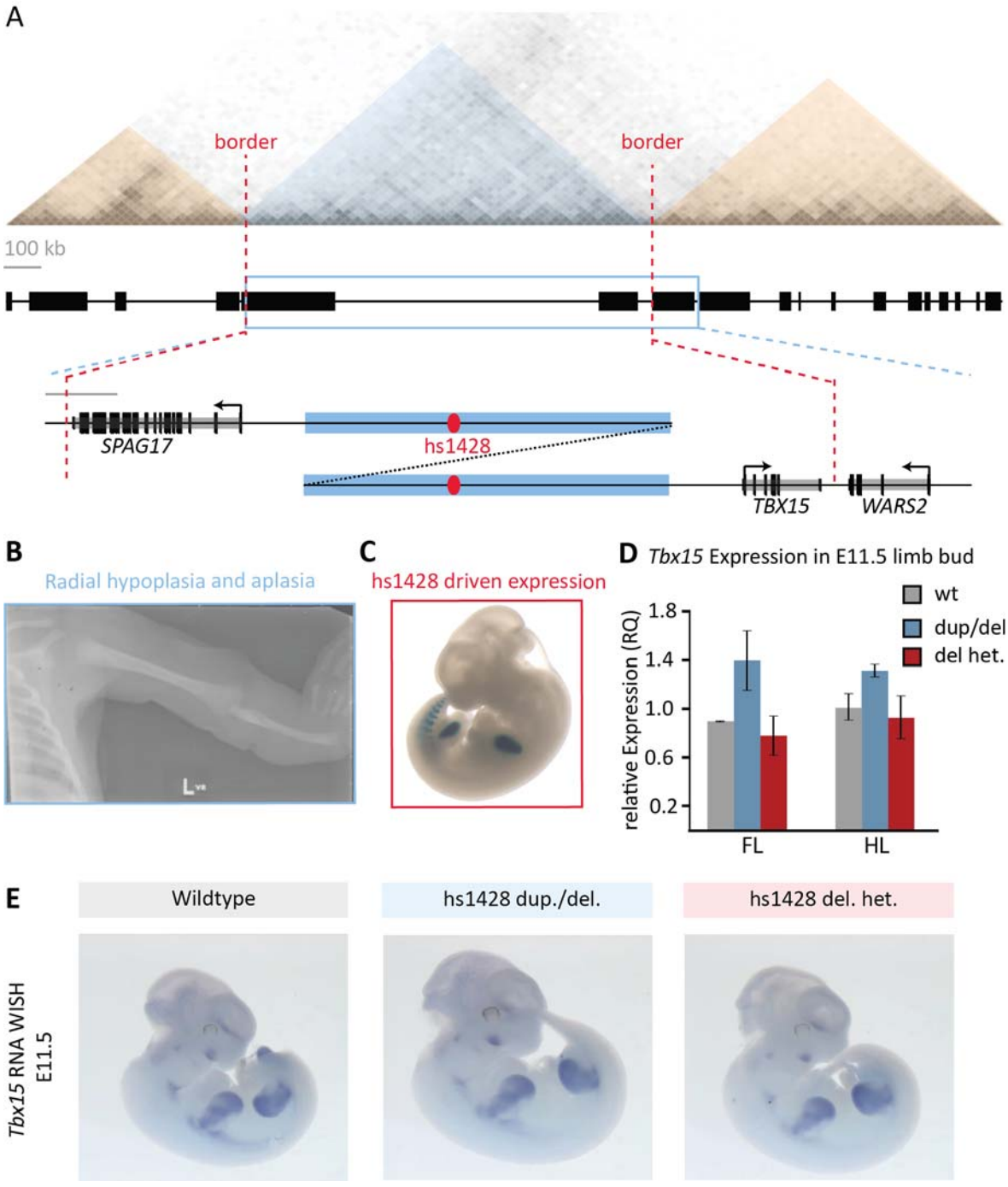

Figure 30: **Heterozygous duplication of hs1428 in a case of radial aplasia**
Duplication at the *TBX15* locus of a patient with radial hypoplasia and aplasia: (A) The duplication is located upstream of *TBX15*. Hi-C data ([33]) shows the duplication is located within the TAD of *TBX15* (blue triangle) and does not involve any boundary. (B) X-ray of the left arm of the patient with radius aplasia. (C) The duplication includes the enhancer hs1428 that drives limb specific expression of a Lac-Z reporter construct (staining from Vista Enhancer Browser). Chimeric mice which are heterozygous for both duplication and deletion (dup/del) or only for deletion (del het.) had been investigated together with wildtype (D) *Tbx15* expression level in E11.5 limb buds of dup/del and del het. animals is unchanged as well as (E) expression pattern.

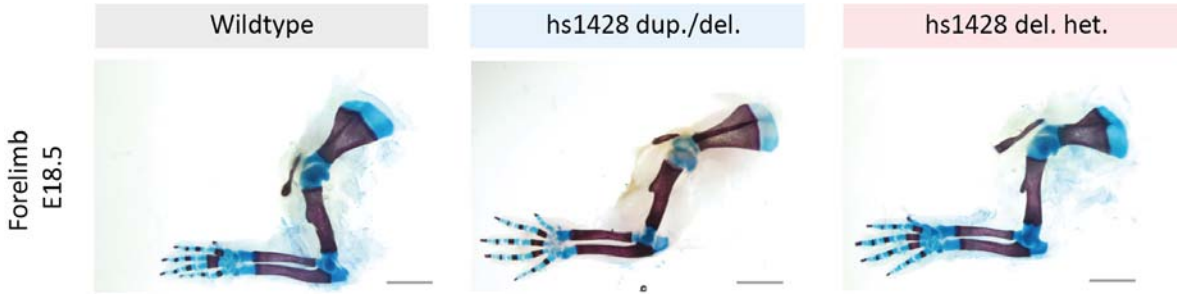

Figure 31: **Heterozygous duplication of hs1428 does not induce radial aplasia in mouse**
Skeletal staining of mice with duplication of hs1428 show no phenotype in chimeric animals. (bar = 2mm)

Skeletal stainings of fore limbs showed normal sized radii for all mutants (see 30E). Expression pattern analysis with RNA WISH (see Fig.30A) as well as expression analysis with qPCR (see Fig.30D) revealed no changes. Therefore, it must be concluded that the patient duplication of the enhancer hs1428 is not able to produce the human phenotype in chimeric mice. Still, it has to be investigated whether non-chimeric or homozygous mice will have radial hypoplasia.

4.6 Missense Mutation in FNDC3A is associated with SHFM

In a huge family from Syria SHFM was obtained in two generations with autosomal recessive inheritance pattern (see Fig.32B). Affected individuals have oligodactyly of feet with cleft like appearance while hands show mild syndactyly (see index patient in Fig.32A). Standard diagnostic procedures were performed by the genetic laboratory of Charité - Universitätsmedizin in Berlin but no mutation could be observed.

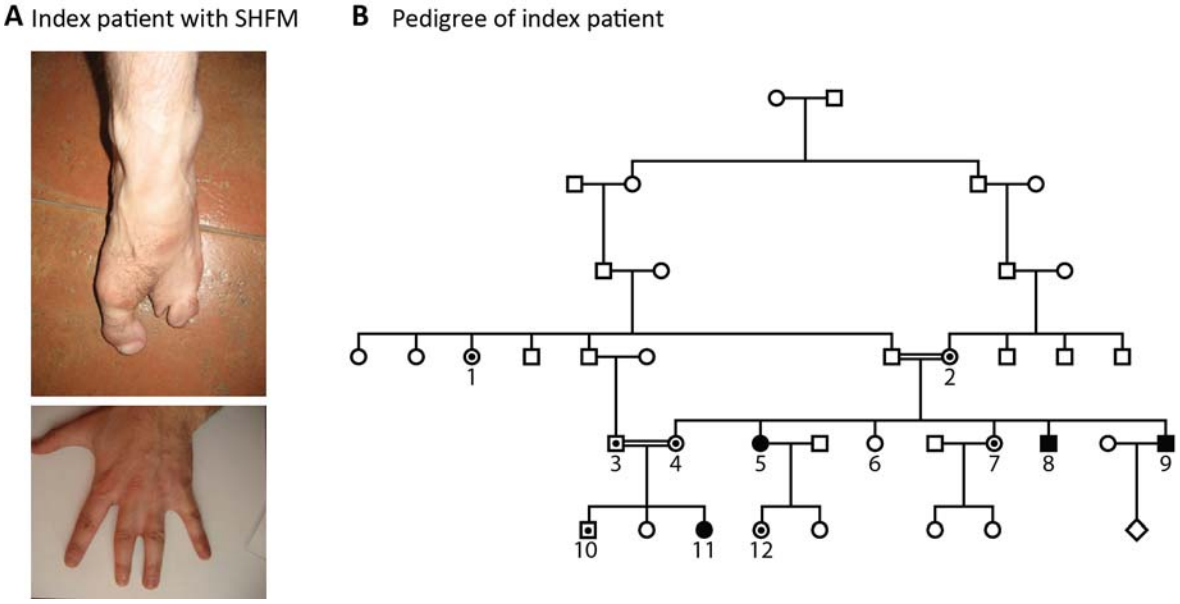

Figure 32: **The FNDC3A mutation Tyr1096Cys segregates with recessive SHFM in a consanguineous family**
The single family case with recessive SHFM is characterized by (A) mild syndactyly of hands and oligodactyly of feet with cleft like appearance. (B) Pedigree of this family with two consanguineous couples show four affected family members (#5, 8, 9 and 11). Sanger sequencing show that affected individuals are homozygous for the mutation Tyr1096Cys and that healthy parents are carrier (#3,4 and 7). Other healthy family members carry only the wildtype allele (#6) or the heterozygous mutation (#1,7,10 and 12)

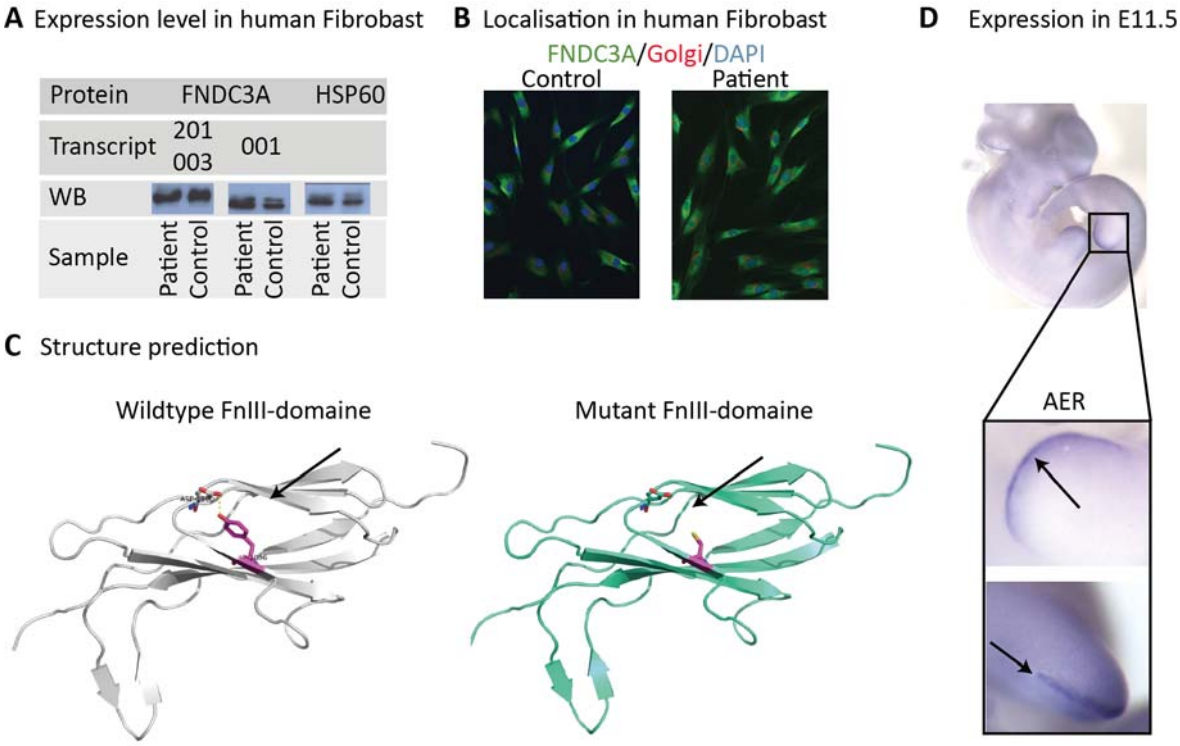

Figure 33: **FNDC3A Tyr1096Cys might interfere with protein structure but not dosage and localization**
Investigation in patient and control fibroblast shows that the mutated FNDC3A is (A) not down regulated and has (B) the same subcellular localization. (C) Structure prediction based on the known crystal structure of the 6th Fibronectin type-III domain in FNDC3C revealed a potential hydrogen bond between Tyr1096 and Asp1100 that get lost in the mutant version Tyr1096Cys. (D) RNA WISH for Fndc3a in mouse E11.5 embryos shows expression in AER of limb buds. Structure prediction was performed by Catherine Sargent (former Worth) (Yaspo group at Max-Planck Institute for Molecular Genetics, Berlin).

SNParray based linkage analysis was performed commercially with a 250k SNP array from Affymetrix and revealed a linkage interval at chromosome 13 between 48 and 52Mb. Exome sequencing of three affected family members (# 5, 8 and 9 in Fig.32B) (performed by NGS facility of Charité - Universitätsmedizin in Berlin) detected two homozygous point mutation: A missense mutation in *FNDC3A* (Tyr1096Cys) and splice site mutation in *RCBTB*. Sequencing of the *RCBTB* transcript from patient fibroblast proved normal splicing (data not shown), therefore the missense mutation in *FNDC3A* was investigated further. Sanger sequencing of 12 family members showed that affected family members are always homozygous for the mutation while healthy family members are heterozygous or carry the wildtype variant only (see Fig.32B). Parents of patients are always heterozygous. Thus, the missense mutation Tyr1096Cys in *FNDC3A* segregates with the SHFM phenotype and fits to autosomal recessive inheritance mode with complete penetrance.

4.6.1 Effect of the Tyr1096Cys Missense Mutation on Human FNDC3A Protein

Fibroblasts obtained from patient skin (obtained by Aleksander Jamsheer from University of Poznan, Poland) were used to investigate the effect of *Fndc3a* missense mutation on protein level. Western Blot analysis detected normal expression level of FNDC3A (see Fig.33A) and immunostainings showed that cytosolic localization is unchanged (see Fig.33B). Structure

prediction was performed by Catherine Sargent (former Worth; Yaspo group at Max-Planck Institute for Molecular Genetics, Berlin) with the help of a homology model for the 9th Fibronectin type-III domain of FNDC3A that was obtained from Protein Model Portal. The homology model bases on the known crystal structure of the 6th Fibronectin type-III domain in FNDC3C and predicts a hydrogen bond between Tyr1096 and Asp1100. The substitution Tyr1096Cys leads to the loss of this hydrogen bond and is thereby destabilizing the structure and allows opening of the tunnel structure (see Fig.33C). In summary these data do not clearly show what the effect of Tyr1096Cys substitution in FNDC3A is, but structure prediction suggests destabilization of the protein, which might cause pathogenic consequences.

4.6.2 Fndc3a Role in Murine Limb Bud Development

To elucidate if *FNDC3A* is expressed during limb development, its expression pattern was analyzed with RNA-WISH in E11.5 mouse embryos. Indeed, *Fndc3a* is expressed in the AER (see Fig.33D) implying that *Fndc3a* plays a role in limb development.
Fnd3a-knockout mice were kindly supplied by Grant McGregor (University of California, Irvine). These mice did not have any obvious limb anomalies and skeletal preparation of homozygous KO animals showed normal bone morphology (data not shown) suggesting that *Fndc3a* is not required for normal limb development in mouse.
To clarify if the point mutation Tyr1096Cys in *FNDC3A* could drive the SHFM phenotype in human the same mutation was successful recapitulated in murine ES cells with the use of CRISPR/Cas9 (see 4.1.3). Out of these ES cell mice were generated that are heterozygous for the mutation. As in human the heterozygous mutation does not cause any limb malformation. Therefore, mice need to be bred to homozygosity to clarify if the point mutation Tyr1096Cys in *Fndc3a* causes SHFM.

5 Discussion

5.1 CRISPR/Cas9 Enables Rapid Generation of Rearrangement in ES Cells

CRISPR/Cas9 is an efficient technology for specific genome editing via double strand breaks [41, 60]. Here two sgRNA were successfully used to create deletions, duplications and inversions at different loci in murine ES cells (see 4.1.1). These ES cells were used to create mouse models that recapitulate certain human rearrangements (see 4.1).
In the past, genomic rearrangements were difficult to create and required complex design and cloning of targeting constructs which usually took several months or even a year. Additionally, intensive screening for the right modification had to be performed with Southern Blot. Homologue recombination is the traditional mechanism for the generation of small deletions and insertions [43]. In-tandem duplications and inversions could be generated so far with the *loxP-Cre* system, which requires the generation of three lines for duplication and two for deletion. These lines have to be mated to combine the three genotypes in one animal [53, 54]. This process takes at least a year while CRISPR/Cas9 allows the generation of deletions, duplications and inversions of several hundred kb in mice within 10 weeks. A similar approach has been performed in zebrafish where TALEN and ZNF were used to generate larger inversions and deletions of more than one 1Mb [126]. However, the design and cloning of specific TALEN and ZNF proteins is complex and time consuming [41, 43] and both technologies were therefore not widely used.

5.1.1 Translocation

This work did not aim for translocations, nevertheless, this is the only rearrangement type that is missing in mouse models but is frequently found in humans [21, 14]. The mechanism for a translocation should be the same as for duplication since in both cases genomic material is transferred from one chromosome to another. Close proximity of both chromosomes is important for translocation. The fixed position of chromosomes in the nucleus (chromosome territories) [127, 128] results in a close neighborhood of certain chromosome while others are never in proximity which prevents interchromosomal interaction and translocation.
A recent study shows that CRISPR/Cas9 is able to induce translocations in *C.elegans* and human cell lines (e.g. HEK cells). In both cases different chromosomes were targeted with two guides to fuse the endings of them [129, 130]. Contrarily, in most human translocations intra-chromosomal parts are inserted into another chromosome and require three double strand breaks: One in the target chromosome and two to excise the to be translocated piece. These DSBs could be induced with CRISPR/Cas9 but it still needs to be tested if controlled translocation can be achieved.
Although translocations could be generated in human cells, it is unclear if this can also be done in ES cells. A general experience is that human cell lines are much easier to target with CRISPR/Cas9 than murine ESCs and insertions are achieved with a much higher efficiency. The reason for this is still unknown and might be caused by higher transfection rates in cell lines, which could be checked using GFP expressing vectors. It is also possible that ES

cells encompass a higher internal protection against genomic modifications. They are the source of all somatic cells and in the adult organism stem cells form important cell pools of precursors. From an evolutionary point of view, it is therefore reasonable to equip these cells with strong protection against mutations, which will be not carried out in all cells because of its energetic expense. Indeed, it was shown that germline cells have a reduced mutation rates in comparison to somatic cells [131]. Translocations were successful induced in HEK cells, which is derived from kidney tissue, and are likely to be less protected against mutations than stem cells.

5.2 CRISPR/Cas9 Requires Precise Genotyping

Specific break points allow to detect and distinguish different structure variations with PCRs. This revealed that we can gain duplications, deletions and inversions with CRISPR/Cas9. Surprisingly, for some ESC clones more than two rearrangements were detected. Additionally, sometimes only one of two inversion break point could be amplified. Both results suggest, that generated SVs were not complete. Therefore, precise genotyping of CRISPR-targeted cells was established resulting in a three-steps protocol. First clones are screened with qualitative PCRs that detect rearrangement specific breakpoints. Then specificity is proved by Sanger sequencing and quantitative PCR verifies the genomic copy number. This intense investigation is mandatory because of unprecise performing of CRISPR/Cas9 due to the error prone DNA repair machinery.

5.2.1 Indels Occur at Break Points of Rearrangement

PCR based screening of clones gave rise to different sized amplicons suggesting that clones with the same rearrangement still differ in their genotypes. Sanger sequencing revealed small deletions and insertions of proximal 20bp at the predicted break points. Sometimes they reach sizes of 100bp. Indels (also called molecular scarring) are known to be caused by the cellular DNA repair machinery and are typical for NHEJ [14]. NHEJ is the common mechanism for DSB repair because it is extremely stable due to its variability regarding enzyme specificity and order in which enzymes appear. Polymerases act at DSB without a template and fill the gap therefore with random base pairs causing small insertions. The ligase that is supposed to ligate open endings can also act prior to DNA synthesis which results in small deletions [19]. Since indels are also found for TALEN and ZNF applications [126] it can be concluded that molecular scarring is caused by DSB induced NHEJ activity.

The fact that indels occur at the break points has to be considered in the design of CRISPR-experiments and exons should not be targeted for the generations of rearrangements since frameshifts will occur. However, the frequent formation of indels can be used to create frameshifts on purpose by targeting the investigated gene with one sgRNA.
Although indels are typically rather small they can also span more than 100bp and genotyping primers have to be located further away from breakpoints. A distance of minimum 200bp allowed reliable detection of rearrangement in this work.
Indels are also found in clones without SVs with a frequency of more than 90% and are thereby

a good intrinsic control for Cas9 activity especially if the desired rearrangement could not be achieved. Indeed, the majority of wildtype specific break point PCRs show indels. If that is not the case it is likely that either the position of the guide in the vector is incorrect or that the region can not be targeted. In that case new guides need to be designed.

5.2.2 Duplications without Deletions Suggests Gain of Copies.

Surprisingly, in rare cases duplication could be generated without deleting the second allele (see 4.1.1). In consequence, more than two copies of the rearranged region are present in these clones, which raise the question where the additional copy comes from. During S-phase of the cell cycle the whole genome is replicated to allow its equal distribution during mitosis. Thus, between S-phase and Telophase each locus is present in four copies [132]. If Cas9 creates duplications in this time frame it is possible that the corresponding deletion segregates to the other daughter cell after cell division. This hypothesis could be verified by blocking proliferation during CRISPR/Cas9 application which should not engender pure duplications.

5.2.3 Incomplete Rearrangements

For some clones only one inversion break point could be detected while the second one was missing. Thus, inversions do not always include the whole target region. Quantitative PCRs revealed that this is also true for other rearrangements, for instance a duplication clone was found with copy numbers between three and four depending on the detected position. It is likely that parts of the targeted region are lost due to the relative big size of rearrangements that increase the likelihood for further DSB. The exact composition of incomplete rearrangements is difficult to elucidate and therefore those clones are excluded from aggregation.

5.3 Efficiency of CRISPR Induced Rearrangements

CRISPR/Cas9 is a powerful tool to generate rearrangements of different kinds in murine ES cells. Nevertheless, most of the clones encompass indels at the break point instead of structural variations (SVs), which was already reported for mammalian cell lines [56]. Deletions and inversions occur with a much higher frequency than duplications because duplications require a simultaneous targeting of both alleles and close proximity of them.
It was reported previously that deletions appear more often than inversions [56]. This could not be confirmed in this work for all loci, although a general trend towards higher deletion rate could be obtained for three out of four experiments. Possibly, inversions and duplications require stabilization and ligation of the excised fragment and rely therefore much more on the cell machinery than deletion that are only facilitated by Cas9 nuclease activity.
The power of CRISPR/Cas9 to induce SVs is reduced by the occurrence of incomplete rearrangements, which was not reported before. Incomplete inversions can be easily distinguished from the true ones due to specific sequences on both endings. Just a third of all inversions are complete, thus the efficiency is reduced by 66%.

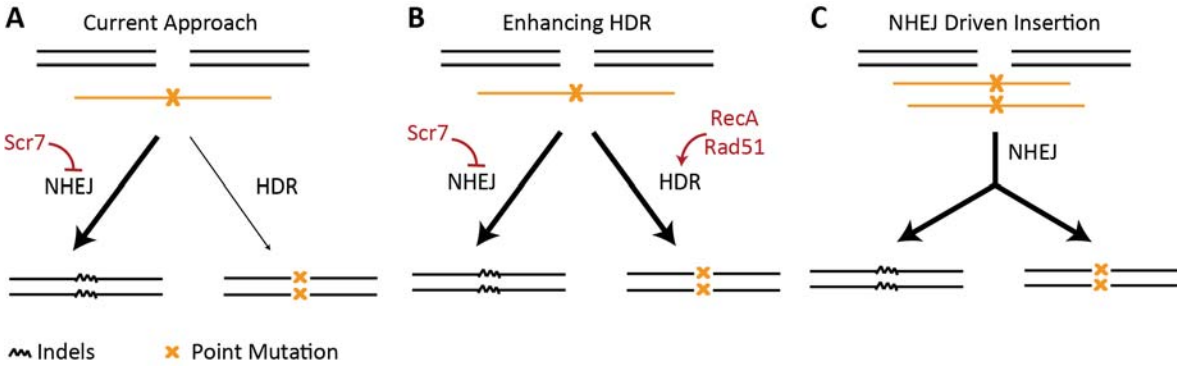

Figure 34: **Approaches for generation of specific point mutation**
(A) Current approach, that was used in this work, (B) optimization of current approach by increasing HDR efficiency (not established) (C) using NHEJ to mediate integration of double stranded oligos.

5.3.1 Locus Specific Efficiencies

Locus specific efficiency rates were previously reported for deletions [57] and have been linked to the size of target regions [56]. This correlation could not be confirmed in this study, maybe because the SVs were ten times larger than those of published experiments. The efficiency for deletions (and other rearrangement) was much higher for 353kb (*Laf4* locus) than for 160kb (*Bhlha9* locus), while a similar sized targeting of 360kb at the *TBX15* locus was less efficient which suggests that performance depends not completely on the size of rearrangements but on other locus specific features. These features could be DNA modifications, transcriptional activity or the accessibility of genomic region.
Accessibility of a locus is mandatory for Cas9 activity and was therefore the factor chosen to be investigated to increase the efficiency. The accessibility of DNA can be affected by chromatin condensation and transcriptional activity. Although many chromatin markers are known we are not aware of the composition that marks regions accessible for Cas9. To investigate this further different guides were designed for the *Bhlha9* locus which are located inside, outside or close to DNaseI hypersensitive sides. These sites indicate open genomic positions that can be targeted by DNase [133] and might be more accessible for Cas9. Indeed, guides within hypersensitive sides were able to facilitate targeting while guides that were just close but not at DNaseI hypersensitive sides (HS) did not work. In contrast genome wide studies in human iPS cells did not show any correlation between DNaseI HS and NHEJ (or HDR) efficiency [57]. It is possible that DNaseI HS sometimes come together with other factors (e.g. chromatin and histon markers) that are the real origin of CRISPR variability or combination of multiple markers are needed to predict efficiency. This has to be investigated in further genome wide studies with different cell types.

5.4 CRIPSR/Cas9 Induced Point Mutation

CRISPR/Cas9 facilitates insertion of short sequences of around 10bp in iPS cells and after zygote injection [58, 57]. In contrast to the generation of rearrangements insertion of oligos requires homology directed repair (HDR). Here this approach was successfully used in murine stem cells to generate mice with a certain point mutation in *Fndc3a*. While Yang et al. (2013) reported an efficiency of 2% for iPS cells [57], the application in ESC was less efficient. No

positive clones were obtained with the protocol of Yang et al. In a second experiment inhibition of the NHEJ pathway with Scr7 gave rise to one positive clone in 300. Ligase IV inhibitor Scr7 is reported to increases HDR frequency by preventing NHEJ completely (see Fig.34A) [67]. In contrast, all negative clones of the *Fndc3a* experiment encompass indels showing that NHEJ was not completely inhibited. On the other hand, it demonstrates correct targeting of *Fndc3a* by Cas9. Therefore, the focus for enhancing insertion capability should be on promoting HDR and blocking NHEJ.

Inhibition of NHEJ is challenging because of its intrinsic stability caused by redundancy of required enzymes (see 1.1.2). HDR efficiency is in general low, thus it might be easier to increase HDR activity instead. A crucial step in HDR is the homology search that is performed by RecA/RAD51 filament complex. Single stranded templates serve as a scaffold for assembly of this complex [134]. This implies that the homology search is not performed at the position of DNA damage but by potential templates. Thus, a possible way to increase the insertion rate is enhancing homology search by overexpression of RecA and RAD51 from a plasmid that is transfected together with the CRISPR/Cas9 system (see Fig.34B). A recent study shows that the RAD51 activator RS-1 increases integration rate ten times better than Scr7 [135].
A totally different approach for integration is to use the highly efficient NHEJ machinery. It was shown that ligation-mediated insertions of double stranded oligos can be facilitate by sticky ends that were generated with zinc-finger nucleases [47, 43]. Indeed, the CRISPR system can also induce single strand breaks when Cas9 is replace by nickases (see Fig.34C) and was already successfully used for precise insertion of a 164bp template [136]. It needs to be investigated if this method is working in ES cells and if it allows a higher insertion efficiency than the approach used in this work. (QUELLE FOR ESC?)

5.5 Comparison of Direct and Indirect Mouse Generation

Generation of mice with CRISPR/Cas9 is often done directly by injection of sgRNA and Cas9 proteins in zygotes or blastocysts, mainly because ES cell culture and aggregation technology are not required. Injection takes just eight weeks in comparison to 10 weeks for aggregation. Certainly, both methods are characterized by low time requirement in comparison to alternative methods. However, the application in ES cell holds numerous advantages that are summarized in table 30.
The ESC approach allows screening of many clones. For the most experiments 300 clones had been investigated, but it is possible to obtain more than 1000 clones if necessary. This allows generation of genotypes with low efficiency rate. Because of financial and space restriction much less mice are allowed to be born after injection. Additionally, many of these mice will not show the desired genotype due to variable efficient rates and incompleteness of rearrangement, as it was found in ESC. There is only one study that list efficiencies for the generation of different rearrangements at one locus, which is for deletions 1-35%, for inversions 1-20% and for duplication 1-5% [137]. Unfortunately, no qPCR data from injection experiments are

Table 30: Comparison of direct (zygote injection) and indirect (aggregation) mouse generation and valuation of characteristics. Rearrangement sizes for injection were summarized from [137], for aggregation from [138].

	Injection	Aggregation
ESC culture	not required	required
time	8 weeks	10 week
chimerism	always	minor contribution (<1%) with tetroploid aggregation
genotypes		
deletion	up to 900 kb	up to 1.6Mb
inversion	up to 900 kb	up to 1.6Mb
duplication	up to 900 kb	up to 1.6Mb
point mutations	easy to obtain	obtained
mosaicism	can not be excluded	excluded by prescreening
screening		
sample number	few (approx.100)	high (approx.1000)
exclusion of unwanted genotypes	not possible	possible
re-generation	unpredictable	possible
re-targeting	time consuming	simple

published and therefore it can not be said how many of the detected rearrangements involve the whole target region.

Inefficient mouse generation is problematic in regards to financial but also ethical aspects. Furthermore, chimeric mice that are obtained from injection experiments are likely to carry different kinds of mutations since CRISPR/Cas9 is applied in a multi-cell stage and will result in F1 offsprings with different genotypes. Since the genotype of ES cells can be detected prior to aggregation mosaicism can be excluded.

Aggregation of ESC allows re-generation of mice with the same genotype while injection does not allow any prediction of the outcome. This is necessary when a mutation turns out to be lethal or to interfere with fertility. In such a case generation of mice from the same ESC clone enables investigation of the developmental background of a disorder by obtaining chimeric mice in different stages as it was done in the *Laf4*-project.

Finally, CRISPR/Cas9 application in ES cells allows re-targeting of a clone to create multiple mutations without the consuming establishment of a line for the first mutation. Thereby, it enables the investigation of the effect of multiple mutations.

5.6 Nievergelt-like Syndrome is Caused by Intragenic Truncation of *LAF4*

LAF4 was the first gene discovered to be causative for the Nievergelt-like phenotype, which if deleted in humans results in triangular tibia, fibula aplasia, brain dysfunction and numerous

organ malformations [108]. In contrast, *Laf4* knock-out does not induce any skeletal abnormalities (see 4.2.1). Re-investigation of the patient with more precise arrayCGH revealed that *LAF4* is not deleted but affected by intragenic in-frame truncation. Additionally, a smaller duplication was detected in the first intron.
The truncation causes almost total loss of both transcription activating domains while the nucleus and nucleoli localization signals are still intact. A putative mutant protein could causes gain- or loss-of-function phenotype. The latter was already excluded to be pathogenic by *Laf4*-KO mouse model. Therefore, CRISPR/Cas9 was used to generate the truncation in mouse. Indeed, the mouse features the typical Nievergelt Syndrome characteristics like mesomelic dysplasia and triangular tibia. In conclusion the *Laf4* truncation causes Nievergelt-like syndrome in mouse via a gain-of-function effect.
Additionally, the mouse recapitulates patient specific phenotypes like fibula aplasia, neuronal dysfunction and lethality. In contrast, kidney, genitalia and colon malformations were only found in the patient. Inter-species differences might cause the missing of these features (see 5.11). Since *Laf4* is not expressed in the embryonic gastro-intestinal truck it is possible that the truncated *Laf4* version does not affects its development. It can be hypothesized that the duplication at the promoter region affects transcriptional regulation of *LAF4*, drives ectopic expression in these organs and thereby causes the extended phenotype of the patient. This could be investigated with a reporter assay in which LacZ expression is either driven by the normal or duplicated promotor region of *Laf4*. For this approach the duplicated area has to be narrowed down and breakpoints should be sequenced.
$Laf4^{trunc}$ mice show a general delay in skeletal development in form of reduced ossification of long and flat bones. This implies that LAF4 plays a role in bone development which is important for both endochondral and dermal ossification. Flat bones are formed by dermal ossification during which in contrast to endochondral ossification no cartilage scaffold is formed but mesenchymal cells directly differentiate into osteoblasts. Osteoblasts are required for bone mineralization in flat and long bones [139]. Therefore, it is likely that *Laf4* truncation interferes either with osteoblast differentiation or their functionality.

5.6.1 Triangular Tibia Results from Uneven Positioning of Growth Plate

Misshaped bones like triangular tibia result either from misplaced bone anlage or dysregulation of longitudinal bone growth, which mainly takes place in the growth plate [140]. Bone anlage defects are also causative for bone aplasia. Since both were found in $Laf4^{trunc}$-mice determination of bone anlage was checked by comparison of *Sox9* expression patterns, but were confirmed to be normal. Another reason for frequent fibula aplasia in the mouse model can be reduced number of condesing cells. If not enough cells are condensing bone can not be formed [84]. Possibly the amount of fibula forming cells are reduced and allow just sometimes formation of rudimentary fibula due to stochastic effects.
Histological investigations of chimeric mice clarified that the bone tissue including growth plate developed in the right manner. In other mouse models alteration in bone morphology - especially reduction of bone length - was found to be caused by polarization defects in proliferating chondrocytes that becomes obvious in the failure of chondrocytes to form

columnar stacks [141, 142]. Additionally, the perichondrium, which restricts chondrocyte spreading [143], is also lost in mice with chondrocyte polarization defects [141]. Anyhow, this is not the case in *Laf4*^{trunc}-mice. The only detectable change is dysorientation of the whole growth plate, that is not described in literature so far and it is unclear which pathway can be responsible for this change. It could be caused be minor anlage defects that could not be detected with WISH since the resolution of this method is too low.

It is still unclear how positioning of the growth plate is established and the *Laf4*^{trunc} mouse offers the first opportunity to study this process. Unfortunately, postnatal lethality of the mice prevents intense investigation of the phenotype and the underlying pathomechanism. This could be overcome with the generation of a line with inducible truncation, for example with the loxP-Cre system. If *loxP* sites are inserted into intron 2 and 10 of *Laf4* truncation can be induced limb specific by Prxx1-promotor driven Cre-expression that should prevent postnatal lethality, if it is due to organ failure. If these animals are not viable a *Laf4*^{loxP} line can be established that is crossed to Prxx1::Cre mice which allows the investigation of embryo stages. A big advantage of *Laf4*^{loxP} mice would be that it can be crossed to different Cre-expressing lines which allows induction of the truncation at specific time points. In the Cre line CRE-ERt2 Tamioxifen is used to induce Cre expression at certain time point and can be also applied perinatal during embryo development [144]. It could be used to elucidate at which developmental stage the malformation is induced e.g. during anlage formation or bone maturation.

5.6.2 Neuronal Dysfunction

Neuronal dysfunctions became apparent in *Laf4*^{trunc}-mice because of the lack of muscular reflects in E18.5 embryos. Although this can be caused by muscular defects it is more likely that *Laf4* affects brain function, since *Laf4* is expressed in the developing brain. It reaches its highest level in the time of increased neurogenesis [119] suggesting an important role in brain development. Interestingly, CGG trinucleotide repeat expansion mutation in the promotor region of *LAF4* is decreasing its expression and has been associated with intellectual disability (ID) [145]. Consistent with this *Laf4* is downregulated in the cortical brain of *Pax6*-KO mice that develop brain structure defects [123]. Taken all these data together *LAF4* is likely to play a major role in brain development. Contrarily, no neurological phenotype could be observed in *Laf4*-KO mice, but it was also not specifically investigated for that. Thus, it is possible that a mild brain phenotype was overseen. If *Laf4* deficiency has an impact on brain development, it is unlikely that the same mechanism is taking place in *Laf4*^{trunc} mutant since the phenotype is much more severe in the latter. No brain histology was performed on *Laf4*^{trunc}-mice but it is likely that neuronal dysfunction can be explained with structural defects of the brain. Thus, histology of both - *Laf4*^{trunc} and *Laf4*-KO brains - has to be performed to reveal the role of *Laf4* in brain development and the pathomechanism of *Laf4* mutations.

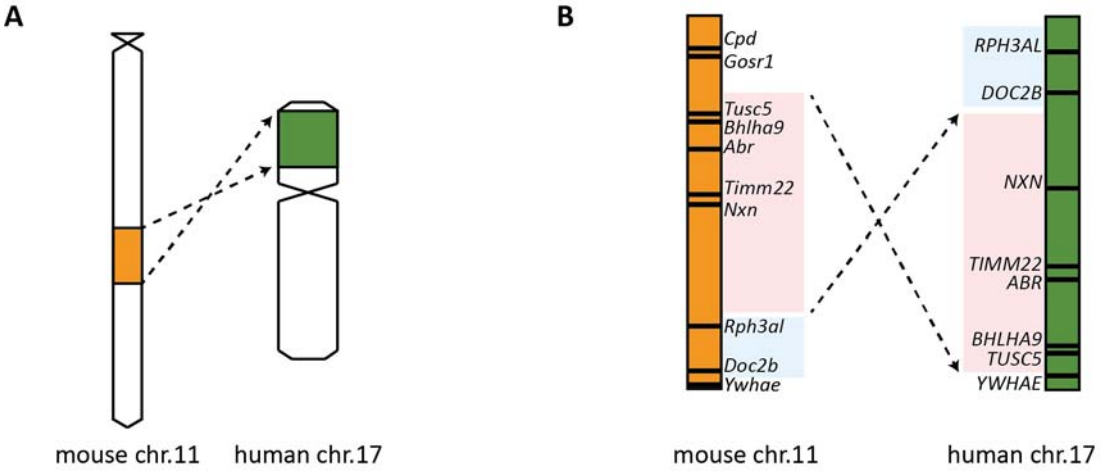

Figure 35: Comparison of the chromosmal structure at the BHLHA9 locus of mouse and human. Adapted from Yingling et al. (2003) [146]

5.7 *BHLHA9*-duplication does not Result in SHFM Phenotype in Mouse

BHLHA9 duplications are strongly associated with SHFM and SHFLD due to frequent occurrence in those patients. However, the role of *BHLHA9* doses in SHFM and the underlying molecular pathomechanism is unknown since additional *Bhlha9* copies do not affect limb development in mouse. That is why a specific duplication of a SHFM patient was recapitulated in mouse with CRISPR/Cas9. Again these mice do not develop SHFM and suggest a striking difference between human and mouse limb development. Noteworthy, homozygous deletion of the same region leads to syndactyly in mouse as it was shown for the pure *Bhlha9* deletion. Indeed, (heterozygous) *BHLHA9* deletion was also found in a human syndactyly case [104].

5.7.1 What can we *learn* from *Bhlha9* Induced Syndactyly

The fact that *Bhlha9* deletion causes syndactyly in human and mouse shows that limb development is sensitive regarding BHLHA9 doses. Interestingly, BHLHA9 seems to be a positive regulator of apoptosis and deficiency causes reduced interdigital apoptosis - mainly at central rays [104]. Thus it can be expected, that overexpression leads to increased apoptosis at central rays and thereby explains emerging of SHFM/SHFLD.
Noteworthy, *BHLHA9* driven SHFM and *Bhlha9* induced syndactyly features the similar characteristic in mouse and man: incomplete penetrance and variable expressivity. This clearly implies that in mouse and human the same modifier is involved. Since it is likely that SHFM and syndactyly is caused by a misregulation of apoptosis the modifier is possibly a apoptosis regulator. Since most people with *BHLHA9* duplication suffer from SHFM it is likely that a common variation can be linked to the phenotype. It is possible that in mouse only the variant that prevents the occurrence of SHFM is present.

5.7.2 Differences in Murine and Human BHLHA9 Network

In contrast to BHLHA9 deficiency overdoses does not cause the same phenotype in mouse and human. This finding suggests a similar function but differences in the involved network.

The most obvious difference between human and mouse is the inversion of the genomic region upstream of *Bhlha9* (see Fig.35). Due to the different size and location of the *BHLHA9* duplication in humans it is unlikely that SHFM is caused by ectopic interaction within the region. Nevertheless, it is possible that the altered genomic order allows better regulation of the additional copy in mouse and prevents overexpression. Maybe, this could be tested by re-inverting this region in ESC with *Bhlha9* duplication. However, it is possible that the involved regulative elements are different between mouse and human. Due to evolutionary distance between both species it is unlikely that regulative elements that were separated from *Bhlha9* by inversion evolved in the same way after.
The missing phenotype could also be caused by alterations in BHLHA9 itself. BHLH-proteins are known to interact with different proteins. In BHLHA9 these interactions could be mediated via the proline rich C-terminus. Interestingly in human the frequency of these prolines is 21,7% while it is just 15,7% in the mouse orthologue [104]. This implies that the interaction network of BHLHA9 differs between mouse and human, for example in the type or chemical properties of interaction partners. If that is the case, it might be feasible to overcome this differences by overexpression of the human *BHLHA9* in the murine limb bud, which should induce SHFM. Still it can not be excluded that the interspecies differences are more complex like it is discussed in 5.11.1 and 5.11.2.

5.8 *SHH* Duplication Causes Muscle and Brown Fat Hypertrophy and not Synpolydactyly

Duplication of *SHH* was found in two patients with different developmental disorders. Kroedrup et al. (2012) reported a patient with muscle hypertrophy that was most prominent at the legs [125]. In contrast the second patient suffers from severe synpolydactyly with cup-shaped appearance of hands and feet (see 4.4). Surprisingly, both patients carry almost the same duplication that includes *SHH* and the neighboring gene *RBM33* (RNA-binding motif protein 33) and crosses the 5-prime border of the *SHH*-TAD. While SHH is a known morphogen and crucial for numerous developmental processes, little is known about *RBM33*. Data from BioGPS, EMBL-EMBI, and Human Protein Atlas suggest an almost ubiquitous expression. Therefore, it can not be excluded that *RBM33* duplication affects muscle and/or skeletal development. Nevertheless, due to the ubiquitous expression more complex malformations are expected.
To elucidate whether the duplication is responsible for one of the disorders or if it is even irrelevant for development, the patient duplication was recapitulated in mouse with the CRISPR/Cas9 approach. The phenotype of (chimeric) mice is marked by incomplete penetrance with heavy body stature, enlarged brown fat tissue between shoulder blades and broad fore limbs. The skeletal system is not affected in any way, suggesting that the duplication of *SHH* does not interfere with skeletal development and does not cause synpolydactyly.
In summery CRISPR/Cas9 technology allowed a rapid verification of the pathogenicity of *SHH* duplication and lead to a clear association of the duplication with muscle hypertrophy. However, it is unclear why the second patient has normal muscle mass which suggests either incomplete penetrance or nonidentical genotypes. The mouse model carries the *Shh* dupli-

cation in tandem as it was the case for the patient with muscle hypertrophy [125] while the location of the additional *SHH* copy was not mapped in the synpolydactyly patient. Therewith a possible explanation for the different phenotype is a translocation of *SHH* to a region where ectopic expression is driven by other enhancers.

5.8.1 *SHH* Duplication does not Interfere with Skeletal Development

Synpolydactyly is linked to ectopic *SHH* expression in the anterior mesenchyme of the early limb bud [88, 73]. The patient duplication includes a potential boundary and thereby possibly fuses parts of the regulative landscape of *SHH* and *RBM33*. This might enable enhancer from the *RBM33*-TAD to act on *SHH* and induce its ectopic expression. Although ectopic interaction between the *RBM33*- and *SHH*-TAD and enhancer adoption was not experimentally confirmed (e.g. by 4C) it is to be expected to emerge as it was shown for similar mutations at other loci [31]. This mechanism would explain the development of synpolydactyly, but the mouse model with the corresponding mutation does not show any skeletal abnormalities. This is in line with the unchanged expression pattern of *Shh* in limb buds of mutants (see Fig.29).
Since it is not clear if the *SHH* duplication of the patient with synpolydactyly is in-tandem or if it is translocated to another locus it is possible that the mouse model does not exactly recapitulate the patient genotype. A translocation to another locus could again explain misregulation of *SHH* due to ectopic interaction with enhancer of that locus.

5.8.2 *SHH* Role in Limb Muscle Development and Muscle Hypertrophy

Skeletal muscles derive from the dorsal part of somites called dermomyotome [147]. In mouse muscle precursor cells (MPC) from the hypaxial side migrate into the limb bud [148] and differentiate to muscle fibers at distinct positions where muscles are required (see Fig.36). Myogenesis is regulated by *Myf5* expression which triggers the onset of differentiation into myoblasts that fuse to myotubes later [148, 149]. *Myf5* is essential for limb muscle development as it was shown in knock-out mice which are lacking limb muscles [147].

ZPA-derived SHH is important for different steps of myogenesis. It promotes distal migration of MPCs and maintains indirectly MPC proliferation and survival [148]. Grafting experiments in chicken showed that overexpression of *Shh* in the limb bud results in strong muscle hypertrophy caused by hypertrophic fibers (and not increased fiber numbers). Increased myoblast proliferation was suggested as a possible mechanism due to *in vitro* experiments [92]. Since *Shh* is upregulated in the limbs of dup/del mice a similar mechanism can be assumed in mouse and the patient.
SHH initiates myogenesis directly by activating *Myf5* expression. Knockout of *Shh* leads to decreased number of *Myf5* expressing MPCs and delays myogenesis [148]. In the contrary scenario *Shh* upregulation could increase the number of *Myf5* expressing cells, resulting in increased number of cells that undergo differentiation towards myoblasts which leads to enlarged muscle fibers and muscle hypertrophy.

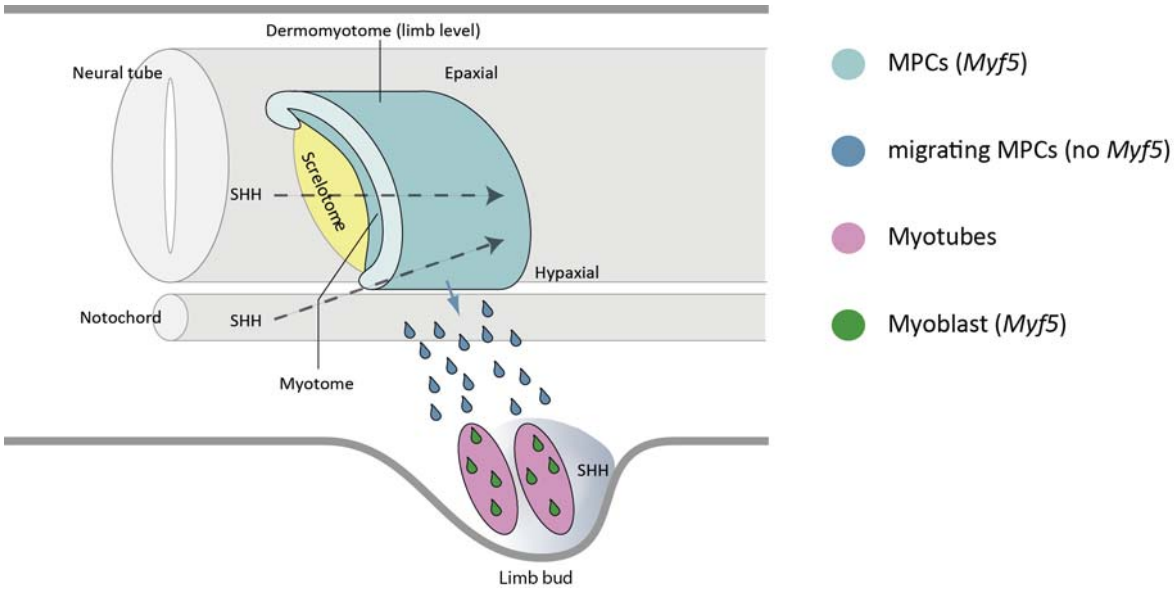

Figure 36: **Muscle and brown fat tissue originate from the same cell lineage**
The Muscle Precursor Cells (MPCs) derive from the dermomyotome and are marked by *Myf5* expression. They give rise to skeletal muscle but also brown fat. Proliferation of MPCs is enhanced by SHH signaling. Scheme is adapted from Endo (2015) [149]

Although *SHH* overexpression is the most likely pathomechanism for muscle hypertrophy in the patient, increased gene dosage - meaning genomic copy number - is not the main factor. Indeed, the patient has three instead of two *Shh* copies, nonetheless, the mutant mouse with muscle hypertrophy features normal copy number due to a deletion of the second allele. Thus, *Shh* overexpression results from changes in the regulative environment. This is in line with a recently published case of corpus callosum agenesis (= missing connection between cerebral hemispheres) with a similar in-tandem duplication of *SHH* with a breakpoint within *RBM33* [150] and not downstream of it as in the other two patients. As a result, enhancers that potential drive the muscle hypertrophy phenotype by ectopic interaction are not included and thereby explains the absent of muscle hypertrophy in the third case.

5.8.3 Link between Brown Fat and Muscle Hypertrophy

The duplication of *Shh* resulted not only in the expected phenotype of muscle hypertrophy but caused additional extreme enlargement of inter scapula brown fat deposits. Skeletal muscles originate from a dermomyotomic cell (DMC) pool which is also the main source for brown fat precursors [151]. Both, muscles and brown fat, arise from a common ancestor that expresses *Myf5* [152, 153] (see Fig.36). This link between muscle and fat development is supported by the finding that two mouse mutants deficient for the muscle markers *Myogenin* or *Myf5* and *MyoD* show changes in muscle and fat formation. In contrast to our mouse in those mutants fat hypertrophy is accompanied by muscle hypotrophy [154, 155].
Shh is a key player for dermomytomic differentiation and proliferation. It is expressed by the notochord and the floor plate of the neural tube and is penetrating the somites in a gradient manner [156]. *Shh* level was not checked in the neural tube of the mouse mutants but over-

expression could explain the fat and muscle phenotype. As a promoter of proliferation, *Shh* upregulation is likely to increase the pool of DMCs that allows both enhanced muscle and fat formation. This can also explain the general massive stature of mutants by additional hypertrophy of trunk musculature, which can not be caused by ZPA-derived SHH. To test this hypothesis *Shh* expression level has to be investigated in the neural tube. In a next step the number of DMCs and their proliferative potential should be determined for example with the help of BrdU staining.

The dual phenotype in *Shh*-dup/del mice can be explained by increased SHH level in somites. Stronger effect on the limb muscles in comparison to other skeletal areas is potentially caused by additional effects of limb bud specific SHH signaling as it was described above. However, it is surprising that no alteration of brown fat tissue was reported for the patient [125]. Brown fat is an important tissue in all mammalian especially in small animal like mice where its thermogenic capacity is required to maintain body temperature. In human only infants are in need of larger amounts of brown fat tissue, which is replaced early by white fat [152]. Hence, brown fat hypertrophy could not be recognized in the patient because the effect was not visible from the outside.

5.9 Duplication of hs1428 does not Cause Radial Dysplasia

In a case of radial dysplasia and aplasia a duplication of the human enhancer 1428 (hs1428) was found, which is located in the *TBX15*-TAD. LacZ-stainings from the Vista-Enhancer browser show that hs1428 is able to drive limb bud specific expression and overlaps with mesenchymal *Tbx15* expression pattern in mouse. Therefore, it is likely that this enhancer is contributing to *TBX15* expression and mutation of this enhancer might cause limb phenotypes.
TBX15 is a known regulator of skeletal developmental and plays a crucial role in the onset of chondrogenesis [84]. Loss-of function mutations lead to the rare Cousin-Syndrome [87, 86] that features the same characteristics as the murine knock-out of *Tbx15* mice with general body size reduction and aplasia of certain bones [84]. Thus dysplastic radii in our patient could be linked to loss of *TBX15* expression in the limb, but the duplication of hs1428 suggests rather upregulation or misexpression of *TBX15*. Anyhow, the mutation was considered to be possibly causative because of the important function of *TBX15* in skeletal development. To determine if this is true and to elucidate the pathogenic mechanism the corresponding duplication was generated in mouse with the help of CRISPR/Cas9. Unfortunately, heterozygous mice did not show any abnormalities and the expression pattern and level was unchanged in chimeric animals implying that duplication of hs1428 is not pathogenic. To conclude this with certainty homozygous mutant need to be investigated. On the other hand, it is possible that the effect of the duplication is rescued in mouse by *Tbx18* [84, 85] but not in human due to inter-species differences that are discussed in 5.11.1 and 5.11.2.

5.10 *FNDC3a* Missense Mutation in a case of Split-Hand/Foot Malformation

A homozygous missense mutation in *FNDC3A* (Tyr1096Cys) was found to segregate in a consanguine family with Split-Hand/Foot Malformation (SHFM). *FNDC3A* is poorly investigated and its function is still unknown. It is expressed during dental development and spermatogenesis and was proposed to play a role in cell adhesion and production of matrix components like collagen type I [157, 158]. Expression in adult organs (e.g. kidney and trachea) [158] suggests additional functions beyond developmental processes. Detection of *Fndc3a* mRNA in mouse limb bud AER implies a role in embryonic limb development.
Extracellular matrix is an important scaffold for inter-cellular signaling [159]. For the limb, the adhesiveness of mesenchymal cells undergoes spatiotemporal changes during cartilage formation [160] and alterations in cell adhesion have been shown to lead to limb malformations such as distal truncations [161]. Mutation of *FNDC3A* might interfere with the integrity of the extracellular matrix of the AER and in this way disrupt signaling to the PZ but also within the AER leading to central growth retardation and limb truncation.
Surprisingly, *Fndc3a* deficient mice do not show any limb phenotype but defects in spermatogenesis [157] showing that *Fndc3a* loss-of-function does affect developmental pathways but that the underlying pathomechanism must be different to the one resulting in SHFM. This is in line with the observation that *FNDC3A* is not down regulated in patient fibroblasts. Furthermore intracellular localization is also unchanged and excludes ectopic *FNDC3A* activity.

FNDC3A consists of eight fibronectin type III (FnIII) domains, which can be found in 2% of all mammalian proteins and are important for protein-protein interaction [162, 163]. FnIII domains form tonne-like structures that are thought to be in a constant process of folding and unfolding to allow interactions partners to bind. Aromatic amino acids like Tyrosine and Tryptophan are highly conserved within FnIII domains and are required for the primary and tertiary structure [164]. Indeed, studies showed that mutation of Tyrosine and Tryptophan residues lead to reduced thermostability [165]. The Tyrosine that is mutated in SHFM patients is predicted to stabilize the structure by forming a hydrogen bond to Asparagin1100. Since the substitution with Cysteine does not support the tertiary structure (e.g. by forming sulfide bonds) the mutation is predicted to destabilize the protein and can thereby result in a malfunction of the protein for instance by allowing ectopic interaction. First destabilization of *FNDC3A* should be investigated with native protein gel electrophoresis. Then interaction partners can be detected by immunoprecipitation (Co-IP) followed by mass-spectrometry.
Noteworthy, although Tyrosine residues are important for structure maintenance the effect of single amino acids on stability is too low to explain the high conservation. Therefore, it has been proposed that Tyrosine is an internal marker for protein structure aberrations. In the native conformation of FnIII domains Tyrosines are facing the interior and are not visible for the cell machinery. But in the unfolded stage it is recognizable for the cell and can act as a signal for miss-folded proteins and induce either refolding via chaperons or elimination by lysosomes or proteinases [164]. Indeed, other studies reported recognition of protein folding substrates based on aromatic amino acids. In the light of this hypothesis Tyr1096Cys

substitution could lead to a defect in recognizing damaged FNDC3A proteins which in result will accumulate in the cell. Protein accumulation is known to disturb cellular processes and signaling that is important for correct limb development. Therefore, accumulation of damaged FNDC3A is a likely pathomechanism for the mutation Tyr1096Cys. In summary the mutation is likely to destabilize FNDC3A and additionally prevents repair and/or elimination of miss-folded FNDC3A and increases thereby the pathogenic defect.
Apparently, this effect does not interfere with any other developmental or non-developmental function of FNDC3A. Developmental processes take place in a restricted time frame and are known to be extremely sensitive towards signal disturbance. Some mutations are tolerated better in matured organs while they can be deleterious in embryogenesis. This might be also the case for FNDC3A. Furthermore, the present of co-factors and interaction partners is likely to play an important role in FNDC3A-driven pathomechanism, since the presence of eight FnIII domains implies a high interaction in protein networks. The fact, that Fndc3a loss-of-function affects spermatogenesis extremely but not all the other organs in which *Fndc3a* is expressed, shows that functionality and interaction partners differ depending on cell types.

To elucidate whether the substitution Tyr1096Cys in *FNDC3A* leads to SHFM, a mouse model was generated with the help of CRIPSR/Cas9 induced HDR. Heterozygous mice do not show any skeletal abnormality as it is also the case in humans. Unfortunately, breeding to homozygosity could not be finished yet.

5.11 CRISPR/Cas9 as a Powerful Tool for Rapid Verification of Pathogenicity of Human Mutations

This work demonstrates that CRISPR/Cas9 allows to verify pathogenicity of human mutation within ten weeks by investigation of chimeric mice. In some cases, establishment of a line was necessary to determine the effect of a mutation in homozygous animal (e.g. for *BHLHA9* duplication and deletion, hs1428 duplication and *FNDC3A* missense mutation). Recapitulation of a human disorder in mouse is an approved approach to link a certain mutation to a distinct phenotype. Furthermore, it is the last step in verification of pathogenicity. Since mouse models have been used for a long time, enormous knowledge regarding their physiology and genomic editing has been collected which is missing for other animals. Animal models that might fit more for the investigation of certain disorders due to similarities to human are expensive in maintenance and are therefore not widely used. If the mouse model does not recapitulate the human phenotype, the only way to find correlation between a human mutation and a (congenital) disorder is to find more than one (unrelated) patient with the same or at least similar mutation as it is the case for *BHLHA9* duplications that are causing SHFM.

5.11.1 Missing Phenotype in Mouse Models can be Explained by Inter-Species Differences

In two cases - *BHLHA9* and hs1428 duplication - genocopying of the human mutation did not lead to the expected phenotype in mouse. While for hs1428 it has to be considered that

Table 31: Overview of recapitulated human mutations and their verification capability

Locus	Mutation	Verification	Potential Pathomechnism
LAF4	intragenic deletion	+	gain-of-function
BHLHA9	duplication	−	upregulation
BHLHA9	deletion	+	loss-of-function
SHH	duplication	+	upregulation
TBX15	duplication of hs1428	−	misregulation

the mutation might not be causative, *BHLHA9* duplications are expected to be causative for SHFLD and SHFM due to the high number of unrelated cases with different sized duplications [95, 98, 100, 102, 99]. Thus other aspects have to be taken in account:

1. Interspecies variations of the developmental environment
2. Functional bias between orthologue genes

Mouse models are used in 59% of all studies related to human health and body because of their genetic, system physiologic and organ similarities. On the other hand, it is clear that human differ at lot from mice in morphometry, physiology, life span and more [9]. Many studies have been published on this issue [166, 167, 168] and it is clear that most mouse models only partly resemble human diseases including congenital disorders of the skeleton [169]. A recent published comparison of postnatal development and aging between man and mice revealed striking differences. Depending on the feature that is investigated one humane year corresponds to 2 - 57 mice days [170]. It must be concluded that postnatal maturation takes place in different systemic environments in mouse and man due to different timing of organ development. It is likely that this is also true for prenatal development and is responsible for variable pathogenicity and sensitiveness towards certain mutations.
Orthologue genes are considered to have the same function between species, but even highly conserved developmental genes e.g. from the *Wnt*-family are partly differential expressed for instance in mouse and chicken [171]. This variability might account for different morphology but can also arise by different functionality in the developmental network. Murine *Wnt5a* expression pattern is mimicked by *Wnt5b* in chicken orthologue [171] suggesting reverse function. Indeed, comparison of embryonic expression pattern revealed striking differences in spatial- and temporal regulation [172].

5.11.2 Differences in Phenocopying Ability of Functional and Regulative Mutations

It is striking that both mutations that interfere directly with protein functionality (*BHLHA9* deletion and *LAF4* truncation) are able to recapitulate the human phenotype. The two disorders that could not be reproduced in mice are predicted to be caused by misregulation of certain gens. Just one out of three mutations with regulative effect induced the patient phenotype (*SHH* duplication). This suggests that functional defects are more likely to affect murine and human development in the same manner than regulative defects (see Tab.31).

Deletions cause total loss of a gene product that forbid any kind of expression adaption of this gene while a duplication can be compensated by regulative feed-back loops that force a certain expression level independent on the genomic copy number. Regulation of gene expression is important for the organism to adapt to a changing environment. This is why regulative networks can differ a lot between species not only due to different proteins within the network but also because of different thresholds that decide on the sensitiveness and preciseness of regulation [173]. Therefore, it is possible that duplications causing overexpression do affect human but not mice but because of more restrict regulation of gene expression. This might be the case for *BHLHA9* duplications. It is clear that limb development is sensitive for *BHLHA9* deletions that causes syndactyly in mouse and human. As discussed before it is possible that overexpression causes the SHFM (see 5.7), which might be prevented in mouse by more restricted expression regulation. This could be investigated with RiboSeq which enriches for translated mRNA and reflects thereby the amount of produced gene products.
In another scenario it is possible that the duplicated gene is upregulated but does not affect the development in mouse due to less sensitivity towards upregulation maintained by regulative loops within the interaction network, that are less efficient in human. The fact that also in human *BHLHA9* duplication does not always cause SHFM suggests that mechanisms are present that allow to overcome the effect of *BHLHA9* overexpression but can be disturbed by a modifier or maybe even by stochastic effect.
It is known that pathogenic effects of deleted genes are often prevented by other genes with redundant functions. This is not happening by chance but is tightly regulated and redundant genes are often overexpressed if the other one is not functional anymore [173]. If two (or more) genes have the same function it is necessary that their expression is tightly controlled or that the network is not sensitive to enhanced activity at all. In both scenarios redundancy results in increased stability towards gain and loss of copy numbers. This mechanism is likely to take place in the regulation of *BHLHA9* and *TBX15* since both are members of huge gene families which enables redundancy [172].

In summery reduced recapitulation ability of human disorders that are caused by regulative mutations can be caused by differences in three mechanisms:

1. Regulation of gene expression on transcriptional and translational level
2. Regulation of protein activity
3. Sensitivity regarding protein dosage alterations

All these effects are more likely to occur if redundant proteins are present. They also apply for pure regulative mutations that do not affect the copy number of any genes but cause changes in expression level or ectopic expression. The pathomechanism of those mutations are exclusively dependent on the cellular expression machinery and its regulation. As discussed above regulation networks differ a lot between species and therefore it is possibly that the pathomechanism leading to a certain disorder is only present in human but not in mouse.

6 Summary

For decades geneticist have been investigating the human genome and revealed that mutations can cause congenital disorders. The current challenge is to distinguish between pathogenic and nonpathogenic mutations to enable precise diagnostic and risk estimation as well as to determine the choice of efficient treatment. Nevertheless, the effect of coding and noncoding mutations is often not predictable. Therefore, mouse models are needed to verify their pathogenicity. For many years it was extremely time consuming to recapitulate human mutations in mice due to enormous cloning efforts. Additionally, structural variations (SVs) of several hundred kbs could not be generated. Now, CRISPR/Cas9 allows rapid genome editing by sequence specific induction of double strand breaks that facilitate structural rearrangements of more than one Mb and also the integration of point mutations.

This work shows that CRISPR/Cas9 can be used in murine ESC for rapid recapitulation of human mutations and generation of viable chimeric mice with duplications, deletions and inversions of several hundred kbs or single nucleotide substitutions. Both, introducing of SVs and point mutations are marked by variable efficiency that seem to be caused by locus specific factors.

CRISPR/Cas9 was used to generated six mouse models that helped to verify the pathogenicity of *Laf4* truncation, *Bhlha9* deletion and *Shh* duplication while the known pathogenicity of *Bhlha9* duplication could not been proved. The effect of the duplication of the enhancer hs1428 at the *Tbx15* locus and of the Tyr1096Cys mutation in FNDC3A still remains unclear.
Intragenic inframe truncation of *LAF4* was found in a patients presenting Nievergelt-like syndrome that is characterized by triangular shaped tibia which was recapitulated in mouse by *Laf4* truncation while the *Laf4*-KO mouse was normal suggesting a dominant gain-of-function effect of intragenic *Laf4* truncation.
BHLHA9 duplications are found in many patient with SHFM or SHFLD. Nevertheless, overexpression of *Bhlha9* does not result in SHFM in mouse. Even the recapitulation of a specific patient duplication did not cause any limb abnormalities. Thus, it is assumed that BHLHA9 and/or its interaction network differs between human and mouse. Contrarily, the deletion of *Bhlha9* leads to syndactyly as it was already published for mouse and human.
SHH duplication was found in a patient with muscle hypertrophy and in another patient with synpolydactyly. The generation of the corresponding duplication in mouse clearly shows that the tandem duplication leads to soft tissue hypertrophy and not synpolydactyly.
A duplication of the enhancer hs1428 at the *TBX15* locus was found in a case of radial aplasia. In mouse, the heterozygous recapitulation of the mutation did not affect skeletal development. Now, homozygous mice are needed to clarify if the mutation is pathogenic.
In a consanguineous family the homozygous Tyr1096Cys mutation in FNDC3A was found to segregate with recessive SHFM. Since functional investigation could not detect any changes, a mouse models was produced that will help to reveal the effect on limb development.

The unsuccessful generation of SHFM in mouse by *Bhlha9* duplication shows that mouse models do not always recapitulate human physiology and developmental process. Since only disorders with potential regulative pathomechanism could not be reproduced it is likely that inter species difference on the regulative level prevents the emergence of the expected phenotype. Consequently, mouse model can prove the pathogenicity of a mutation but never its insignificance.

7 Zusammenfassung

Durch Jahrzehntelange Forschung am menschlichen Genom ist inzwischen bekannt, dass Mutationen potentiell Krankheiten verursachen. Für die humangenetische Diagnostik ist es essentiell pathogene von nicht-pathogenen Varianten zu differenzieren. Daher ist es problematisch, dass der Effekt von codierenden und nicht codierenden Mutationen nicht vorhersehbar ist. Um die Pathogenität zu verifizieren werden humane Mutationen in der Maus reproduziert. Dies war bisher jedoch sehr aufwendig und dadurch limitiert, dass Struktur Veränderungen (SVs) von mehreren kbs nicht generiert werden konnten. Das kürzlich entwickelte CRISPR/Cas9 System erlaubt nun durch sequenzspezifische Doppelstrangbrüche die exakte Generierung von großen SVs und Punktmutationen.

Diese Arbeit zeigt, dass CRISPR/Cas9 die Generierung von humanen Mutationen wie Deletionen, Duplikationen, Inversionen bzw. Punktmutationen in murinen ESCs ermöglicht, aus denen chimäre Mäuse erzeugt werden. Die Effizienz der Erzeugung von SVs und Punktmutationen hängt jedoch stark von Lokus spezifischen Faktoren ab.

Mit CRISPR/Cas9 wurden sechs unterschiedliche Mauslinien erzeugt und die Pathogenität der intragenen *LAF4* Deletion, der *BHLHA9* Deletion und der *SHH* Duplikation nachgewiesen. Die bereits bekannte Pathogenität der *BHLHA9* Duplikation konnte das Mausmodel nicht bestätigen. Der Effekt der Tyr1096Cys Mutation in *FNDC3a* und der Duplikation des Enhancers hs1428 am *TBX15* Locus ist immer noch unklar.
Die intragene inframe-Deletion in *LAF4* wurde in einem Patienten mit Nievergelt-ähnlichem Syndrom gefunden. Sowohl der Patient als auch das Mausmodel zeigen die typische dreieckige Verformung der Tibia. Da *Laf4*-KO keine Extremitätenveränderung verursacht, wird von einem gain-of-function Mechanismus ausgegangen.
Aufgrund von zahlreichen klinischen Fällen sind *BHLHA9* Duplikationen dafür bekannt SHFM bzw. SHFLD auszulösen. In der Maus haben jedoch weder die Überexpression noch die Duplikation von *Bhlha9* einen Effekt auf die Extremitätenentwicklung. Daher wird vermutet, dass die Funktionalität des menschlichen *BHLHA9* sich vom murinen Ortholog signifikant unterscheidet.
Die Duplikation von *SHH* wurde in einem Patienten mit Synpolydactyly gefunden wie auch in einem Fall mit Muskelhypertrophy. Das entsprechende Mausmodel zeigt jedoch klar, dass die in-tandem Duplikation zu Hypertrophy von weichem Gewebe und nicht zu Synpolydactyly führt.
Die Duplikation des Enhancer hs1428 am *TBX15* Locus wurde in einem Patienten mit radialer Aplasie gefunden. Die gleiche heterozygote Mutation hat in der Maus keinen Effekt. Für eine endgültige Beurteilung werden jedoch homozygote Tiere benötigt.
Die Missense Mutation Tyr1096Cys in *FNDC3A* segregiert innerhalb einer konsanguinen Familien mit rezessivem SHFM. Da funktionelle Untersuchungen nicht ausreichten, um die Pathogenität der Mutation zu beurteilen, wurde sie im Mausmodel nachgestellt. Es lässt sich jedoch noch keine abschließende Aussage treffen, weil die Tiere noch zur Homozygotie gezüchtet werden müssen.

Obwohl bekannt ist, dass *BHLHA9* Duplikationen SHFM erzeugen, konnte dies in der Maus nicht nachgewiesen werden. Dies verdeutlicht, dass die Maus die humane Physiologie und Entwicklung nicht perfekt wiederspiegelt. Da vor allem Mutationen mit einem potentiell regulativen Pathomechanismus in der Maus nicht das menschliche Krankheitsbild reproduzierten, wird angenommen, dass dies vor allem durch Spezies bedingte Unterschiede in der Genregulation verursacht wurde. Folglich ist die Maus ein gutes Model um die Pathogenität von Mutationen nachzuweisen, jedoch nicht um selbige auszuschließen.

References

[1] J. J. McCarthy, H. L. McLeod, and G. S. Ginsburg, "Genomic medicine: A decade of successes, challenges, and opportunities", *Science Translational Medicine*, vol. 5, no. 189, 189sr4–189sr4, Jun. 12, 2013.

[2] M. Pheasant and J. S. Mattick, "Raising the estimate of functional human sequences", *Genome Research*, vol. 17, no. 9, pp. 1245–1253, Sep. 2007, bibtex: pheasant_raising_2007.

[3] ENCODE Project Consortium, "The ENCODE (ENCyclopedia of DNA elements) project", *Science (New York, N.Y.)*, vol. 306, no. 5696, pp. 636–640, Oct. 22, 2004.

[4] A. Eyre-Walker and P. D. Keightley, "The distribution of fitness effects of new mutations", *Nature Reviews. Genetics*, vol. 8, no. 8, pp. 610–618, Aug. 2007.

[5] J. A. Veltman and H. G. Brunner, "De novo mutations in human genetic disease", *Nature Reviews Genetics*, vol. 13, no. 8, pp. 565–575, Aug. 2012.

[6] A. Eyre-Walker, M. Woolfit, and T. Phelps, "The distribution of fitness effects of new deleterious amino acid mutations in humans", *Genetics*, vol. 173, no. 2, pp. 891–900, Jun. 2006.

[7] B. J. Wilson and S. G. Nicholls, "The human genome project, and recent advances in personalized genomics", *Risk Management and Healthcare Policy*, vol. 8, pp. 9–20, Feb. 16, 2015.

[8] S. E. Lewis, "A consideration of the advantages and potential difficulties of the use of transgenic mice for the study of germinal mutations", *Mutation Research*, vol. 307, no. 2, pp. 509–515, Jun. 1, 1994.

[9] L. Demetrius, "Of mice and men", *EMBO reports*, vol. 6, S39–S44, S1 Jul. 1, 2005.

[10] M. Spielmann and S. Mundlos, "Structural variations, the regulatory landscape of the genome and their alteration in human disease", *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, vol. 35, no. 6, pp. 533–543, Jun. 2013.

[11] J. S. Bertram, "The molecular biology of cancer", *Molecular Aspects of Medicine*, vol. 21, no. 6, pp. 167–223, Dec. 2000.

[12] S. B. Dowton and R. A. Slaugh, "Diagnosis of human heritable diseases–laboratory approaches and outcomes", *Clinical Chemistry*, vol. 41, no. 5, 785–794; discussion 813–816, May 1995.

[13] H. Qu and X. Fang, "A brief review on the human encyclopedia of DNA elements (ENCODE) project", *Genomics, Proteomics & Bioinformatics*, vol. 11, no. 3, pp. 135–141, Jun. 2013.

[14] P. Stankiewicz and J. R. Lupski, "Structural variation in the human genome and its role in disease", *Annual Review of Medicine*, vol. 61, pp. 437–455, 2010.

[15] J. J. Wyrick and S. A. Roberts, "Genomic approaches to DNA repair and mutagenesis", *DNA repair*, Sep. 15, 2015.

[16] L. Zhang and W.-H. Li, "Human SNPs reveal no evidence of frequent positive selection", *Molecular Biology and Evolution*, vol. 22, no. 12, pp. 2504–2507, Dec. 2005.

[17] D. L. Wheeler, T. Barrett, D. A. Benson, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. DiCuccio, R. Edgar, S. Federhen, L. Y. Geer, Y. Kapustin, O. Khovayko, D. Landsman, D. J. Lipman, T. L. Madden, D. R. Maglott, J. Ostell, V. Miller, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, K. Sirotkin, A. Souvorov, G. Starchenko, R. L. Tatusov, T. A. Tatusova, L. Wagner, and E. Yaschenko, "Database resources of the national center for biotechnology information", *Nucleic Acids Research*, vol. 35, pp. D5–12, Database issue Jan. 2007.

[18] M. A. Lewandowska, "The missing puzzle piece: Splicing mutations", *International Journal of Clinical and Experimental Pathology*, vol. 6, no. 12, pp. 2675–2682, 2013.

[19] M. R. Lieber, "The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway", *Annual Review of Biochemistry*, vol. 79, pp. 181–211, 2010.

[20] K. Inaki and E. T. Liu, "Structural mutations in cancer: Mechanistic and functional insights", *Trends in genetics: TIG*, vol. 28, no. 11, pp. 550–559, Nov. 2012.

[21] B. Weckselblatt and M. K. Rudd, "Human structural variation: Mechanisms of chromosome rearrangements", *Trends in genetics: TIG*, Jul. 21, 2015.

[22] P. J. Hastings, J. R. Lupski, S. M. Rosenberg, and G. Ira, "Mechanisms of change in gene copy number", *Nature Reviews. Genetics*, vol. 10, no. 8, pp. 551–564, Aug. 2009.

[23] G. J. Swaminathan, E. Bragin, E. A. Chatzimichali, M. Corpas, A. P. Bevan, C. F. Wright, N. P. Carter, M. E. Hurles, and H. V. Firth, "Decipher: Web-based, community resource for clinical interpretation of rare variants in developmental disorders", *Human Molecular Genetics*, vol. 21, R37–44, R1 Oct. 15, 2012.

[24] C. L. Usher and S. A. McCarroll, "Complex and multi-allelic copy number variation in human disease", *Briefings in Functional Genomics*, Jul. 9, 2015.

[25] E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker, "Comprehensive mapping of long-range interactions reveals folding principles of the human genome", *Science (New York, N.Y.)*, vol. 326, no. 5950, pp. 289–293, Oct. 9, 2009.

[26] W. de Laat and D. Duboule, "Topology of mammalian developmental enhancers and their regulatory landscapes", *Nature*, vol. 502, no. 7472, pp. 499–506, Oct. 24, 2013.

[27] B. Fernandez, J Siegel-Bartelt, J.-A. Herbrick, I Teshima, and S. Scherer, "Holoprosencephaly and cleidocranial dysplasia in a patient due to two position-effect mutations: Case report and review of the literature", *Clinical Genetics*, vol. 68, no. 4, pp. 349–359, Oct. 1, 2005.

[28] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren, "Topological domains in mammalian genomes identified by analysis of chromatin interactions", *Nature*, vol. 485, no. 7398, pp. 376–380, May 17, 2012.

[29] E. P. Nora, B. R. Lajoie, E. G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N. L. van Berkum, J. Meisig, J. Sedat, J. Gribnau, E. Barillot, N. Blüthgen, J. Dekker, and E. Heard, "Spatial partitioning of the regulatory landscape of the x-inactivation centre", *Nature*, vol. 485, no. 7398, pp. 381–385, May 17, 2012.

[30] R. R. Haraksingh and M. P. Snyder, "Impacts of variation in the human genome on gene regulation", *Journal of Molecular Biology*, vol. 425, no. 21, pp. 3970–3977, Nov. 1, 2013.

[31] D. G. Lupiáñez, K. Kraft, V. Heinrich, P. Krawitz, F. Brancati, E. Klopocki, D. Horn, H. Kayserili, J. M. Opitz, R. Laxova, F. Santos-Simarro, B. Gilbert-Dussardier, L. Wittler, M. Borschiwer, S. A. Haas, M. Osterwalder, M. Franke, B. Timmermann, J. Hecht, M. Spielmann, A. Visel, and S. Mundlos, "Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions", *Cell*, vol. 161, no. 5, pp. 1012–1025, May 21, 2015.

[32] T. Montavon, L. Thevenet, and D. Duboule, "Impact of copy number variations (CNVs) on long-range gene regulation at the HoxD locus", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 50, pp. 20 204–20 211, Dec. 11, 2012.

[33] F. Jin, Y. Li, J. R. Dixon, S. Selvaraj, Z. Ye, A. Y. Lee, C.-A. Yen, A. D. Schmitt, C. A. Espinoza, and B. Ren, "A high-resolution map of the three-dimensional chromatin interactome in human cells", *Nature*, vol. 503, no. 7475, pp. 290–294, Nov. 14, 2013.

[34] S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden, "A 3d map of the human genome at kilobase resolution reveals principles of chromatin looping", *Cell*, vol. 159, no. 7, pp. 1665–1680, Dec. 18, 2014.

[35] A. Visel, S. Minovitsky, I. Dubchak, and L. A. Pennacchio, "VISTA enhancer browser–a database of tissue-specific human enhancers", *Nucleic Acids Research*, vol. 35, pp. D88–92, Database issue Jan. 2007.

[36] A. C. J. Gijsbers and C. A. L. Ruivenkamp, "Molecular karyotyping: From microscope to SNP arrays", *Hormone Research in Pædiatrics*, vol. 76, no. 3, pp. 208–213, 2011.

[37] J. D. Wall and J. K. Pritchard, "Haplotype blocks and linkage disequilibrium in the human genome", *Nature Reviews. Genetics*, vol. 4, no. 8, pp. 587–597, Aug. 2003.

[38] R. Mehrian-Shai and J. K. V. Reichardt, "A renaissance of "biochemical genetics"? SNPs, haplotypes, function, and complex diseases", *Molecular Genetics and Metabolism*, vol. 83, no. 1, pp. 47–50, Oct. 2004.

[39] S. H. Lelieveld, M. Spielmann, S. Mundlos, J. A. Veltman, and C. Gilissen, "Comparison of exome and genome sequencing technologies for the complete capture of protein-coding regions", *Human Mutation*, vol. 36, no. 8, pp. 815–822, Aug. 2015.

[40] J. T. Fujii and G. R. Martin, "Developmental potential of teratocarcinoma stem cells in utero following aggregation with cleavage-stage mouse embryos", *Development*, vol. 74, no. 1, pp. 79–96, Apr. 1, 1983.

[41] H. Wang, H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang, and R. Jaenisch, "One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering", *Cell*, vol. 153, no. 4, pp. 910–918, May 9, 2013.

[42] A. R. Bassett and J.-L. Liu, "CRISPR/cas9 and genome editing in drosophila", *Journal of Genetics and Genomics*, vol. 41, no. 1, pp. 7–19, Jan. 20, 2014.

[43] Y. Peng, K. J. Clark, J. M. Campbell, M. R. Panetta, Y. Guo, and S. C. Ekker, "Making designer mutants in model organisms", *Development*, vol. 141, no. 21, pp. 4042–4054, Jan. 11, 2014.

[44] C. Mussolino, T. Mlambo, and T. Cathomen, "Proven and novel strategies for efficient editing of the human genome", *Current Opinion in Pharmacology*, Anti-infectives • New technologies, vol. 24, pp. 105–112, Oct. 2015.

[45] M. Tanaka, A.-K. Hadjantonakis, K. Vintersten, and A. Nagy, "Aggregation chimeras: Combining ES cells, diploid, and tetraploid embryos", *Methods in molecular biology (Clifton, N.J.)*, vol. 530, pp. 287–309, 2009.

[46] J. Artus and A.-K. Hadjantonakis, "Generation of chimeras by aggregation of embryonic stem cells with diploid or tetraploid mouse embryos", *Methods in Molecular Biology (Clifton, N.J.)*, vol. 693, pp. 37–56, 2011.

[47] S. J. Orlando, Y. Santiago, R. C. DeKelver, Y. Freyvert, E. A. Boydston, E. A. Moehle, V. M. Choi, S. M. Gopalan, J. F. Lou, J. Li, J. C. Miller, M. C. Holmes, P. D. Gregory, F. D. Urnov, and G. J. Cost, "Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology", *Nucleic Acids Research*, vol. 38, no. 15, e152, Aug. 2010.

[48] L. Zhang, R. Jia, N. J. Palange, A. C. Satheka, J. Togo, Y. An, M. Humphrey, L. Ban, Y. Ji, H. Jin, X. Feng, and Y. Zheng, "Large genomic fragment deletions and insertions in mouse using CRISPR/cas9", *PLOS ONE*, vol. 10, no. 3, e0120396, Mar. 24, 2015.

[49] J. R. Broach, V. R. Guarascio, and M. Jayaram, "Recombination within the yeast plasmid 2 circle is site-specific", *Cell*, vol. 29, no. 1, pp. 227–234, May 1982.

[50] C. S. Branda and S. M. Dymecki, "Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice", *Developmental Cell*, vol. 6, no. 1, pp. 7–28, Jan. 2004.

[51] K. S Koch, T Aoki, Y Wang, A. E Atkinson, A. S Gleiberman, O. K Glebov, and H. L Leffert, "Site-specific integration of targeted DNA into animal cell genomes", *Gene*, vol. 249, no. 1, pp. 135–144, May 16, 2000.

[52] H. Yu, X. Wang, L. Zhu, Z. He, G. Liu, X. Xu, J. Chen, and G. Cheng, "Establishment of a rapid and scalable gene expression system in livestock by site-specific integration", *Gene*, vol. 515, no. 2, pp. 367–371, Feb. 25, 2013.

[53] Y. Hérault, M. Rassoulzadegan, F. Cuzin, and D. Duboule, "Engineering chromosomes in mice through targeted meiotic recombination (TAMERE)", *Nature Genetics*, vol. 20, no. 4, pp. 381–384, Dec. 1998.

[54] F. Spitz, C. Herkenne, M. A. Morris, and D. Duboule, "Inversion-induced disruption of the hoxd cluster leads to the partition of regulatory landscapes", *Nature Genetics*, vol. 37, no. 8, pp. 889–893, Aug. 2005.

[55] F Smih, P Rouet, P. J. Romanienko, and M Jasin, "Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells.", *Nucleic Acids Research*, vol. 23, no. 24, pp. 5012–5019, Dec. 25, 1995.

[56] M. C. Canver, D. E. Bauer, A. Dass, Y. Y. Yien, J. Chung, T. Masuda, T. Maeda, B. H. Paw, and S. H. Orkin, "Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/cas9 nuclease system in mammalian cells", *Journal of Biological Chemistry*, vol. 289, no. 31, pp. 21 312–21 324, Jan. 8, 2014.

[57] L. Yang, M. Guell, S. Byrne, J. L. Yang, A. De Los Angeles, P. Mali, J. Aach, C. Kim-Kiselak, A. W. Briggs, X. Rios, P.-Y. Huang, G. Daley, and G. Church, "Optimization of scarless human stem cell genome editing", *Nucleic Acids Research*, vol. 41, no. 19, pp. 9049–9061, Oct. 2013.

[58] H. Yang, H. Wang, and R. Jaenisch, "Generating genetically modified mice using CRISPR/cas-mediated genome engineering", *Nature Protocols*, vol. 9, no. 8, pp. 1956–1968, Aug. 2014.

[59] J. D. Sander and J. K. Joung, "CRISPR-cas systems for editing, regulating and targeting genomes", *Nature Biotechnology*, vol. 32, no. 4, pp. 347–355, Apr. 2014.

[60] B. Wiedenheft, S. H. Sternberg, and J. A. Doudna, "RNA-guided genetic silencing systems in bacteria and archaea", *Nature*, vol. 482, no. 7385, pp. 331–338, Feb. 16, 2012.

[61] M. P. Terns and R. M. Terns, "CRISPR-based adaptive immune systems", *Current opinion in microbiology*, vol. 14, no. 3, pp. 321–327, Jun. 2011.

[62] E. Deltcheva, K. Chylinski, C. M. Sharma, K. Gonzales, Y. Chao, Z. A. Pirzada, M. R. Eckert, J. Vogel, and E. Charpentier, "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III", *Nature*, vol. 471, no. 7340, pp. 602–607, Mar. 31, 2011.

[63] (). CRISPR/cas9 overview, [Online]. Available: `http://www.clontech.com/US/Products/Genome_Editing/CRISPR_Cas9/Resources/About_CRISPR_Cas9` (visited on 10/19/2015).

[64] L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, and F. Zhang, "Multiplex genome engineering using CRISPR/cas systems", *Science (New York, N.Y.)*, vol. 339, no. 6121, pp. 819–823, Feb. 15, 2013.

[65] S. J. Gratz, C. D. Rubinstein, M. M. Harrison, J. Wildonger, and K. M. O'Connor-Giles, "CRISPR-cas9 genome editing in drosophila", *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]*, vol. 111, pp. 31.2.1–31.2.20, 2015.

[66] A. P. W. Gonzales and J.-R. J. Yeh, "Cas9-based genome editing in zebrafish", *Methods in Enzymology*, vol. 546, pp. 377–413, 2014.

[67] T. Maruyama, S. K. Dougan, M. C. Truttmann, A. M. Bilate, J. R. Ingram, and H. L. Ploegh, "Increasing the efficiency of precise genome editing with CRISPR-cas9 by inhibition of nonhomologous end joining", *Nature Biotechnology*, vol. 33, no. 5, pp. 538–542, May 2015.

[68] Y. Fu, J. D. Sander, D. Reyon, V. M. Cascio, and J. K. Joung, "Improving CRISPR-cas nuclease specificity using truncated guide RNAs", *Nature Biotechnology*, vol. 32, no. 3, pp. 279–284, Mar. 2014.

[69] H. M. Kronenberg, "Developmental regulation of the growth plate", *Nature*, vol. 423, no. 6937, pp. 332–336, May 15, 2003.

[70] R. Savarirayan and D. L. Rimoin, "The skeletal dysplasias", *Best Practice & Research Clinical Endocrinology & Metabolism*, vol. 16, no. 3, pp. 547–560, Sep. 2002.

[71] J.-D. Bénazet and R. Zeller, "Vertebrate limb development: Moving from classical morphogen gradients to an integrated 4-dimensional patterning system", *Cold Spring Harbor Perspectives in Biology*, vol. 1, no. 4, a001339, Oct. 2009.

[72] R. Zeller, J. López-Ríos, and A. Zuniga, "Vertebrate limb bud development: Moving towards integrative analysis of organogenesis", *Nature Reviews. Genetics*, vol. 10, no. 12, pp. 845–858, Dec. 2009.

[73] L. Niswander, "Pattern formation: Old models out on a limb", *Nature Reviews. Genetics*, vol. 4, no. 2, pp. 133–143, Feb. 2003.

[74] T. V. Hellmann and J. N. a. T. D. Mueller, "Missense mutations in GDF-5 signaling: Molecular mechanisms behind skeletal malformation", Oct. 12, 2012.

[75] B. R. Olsen, A. M. Reginato, and W. Wang, "Bone development", *Annual Review of Cell and Developmental Biology*, vol. 16, no. 1, pp. 191–220, 2000.

[76] U. Kornak and S. Mundlos, "Genetic disorders of the skeleton: A developmental approach", *American Journal of Human Genetics*, vol. 73, no. 3, pp. 447–474, Sep. 2003.

[77] J. K. Ng, K. Tamura, D. Büscher, and J. C. Izpisúa-Belmonte, "Molecular and cellular basis of pattern formation during vertebrate limb development", *Current Topics in Developmental Biology*, vol. 41, pp. 37–66, 1999.

[78] A. T. Dudley, M. A. Ros, and C. J. Tabin, "A re-examination of proximodistal patterning during vertebrate limb development", *Nature*, vol. 418, no. 6897, pp. 539–544, Aug. 1, 2002.

[79] R. D. Riddle, R. L. Johnson, E. Laufer, and C. Tabin, "Sonic hedgehog mediates the polarizing activity of the ZPA", *Cell*, vol. 75, no. 7, pp. 1401–1416, Dec. 31, 1993.

[80] E. Blitz, S. Viukov, A. Sharir, Y. Shwartz, J. L. Galloway, B. A. Pryce, R. L. Johnson, C. J. Tabin, R. Schweitzer, and E. Zelzer, "Bone ridge patterning during musculoskeletal assembly is mediated through SCX regulation of bmp4 at the tendon-skeleton junction", *Developmental Cell*, vol. 17, no. 6, pp. 861–873, Dec. 2009.

[81] R. Cancedda, P. Castagnola, F. D. Cancedda, B. Dozin, and R. Quarto, "Developmental control of chondrogenesis and osteogenesis", *The International Journal of Developmental Biology*, vol. 44, no. 6, pp. 707–714, 2000.

[82] S. Provot and E. Schipani, "Molecular mechanisms of endochondral bone development", *Biochemical and Biophysical Research Communications*, vol. 328, no. 3, pp. 658–665, Mar. 18, 2005.

[83] K. L. Cooper, S. Oh, Y. Sung, R. R. Dasari, M. W. Kirschner, and C. J. Tabin, "Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions", *Nature*, vol. 495, no. 7441, pp. 375–378, Mar. 21, 2013.

[84] M. K. Singh, M. Petry, B. Haenig, B. Lescher, M. Leitges, and A. Kispert, "The t-box transcription factor tbx15 is required for skeletal development", *Mechanisms of Development*, vol. 122, no. 2, pp. 131–144, Feb. 2005.

[85] M. King, J. S. Arnold, A. Shanske, and B. E. Morrow, "T-genes and limb bud development", *American Journal of Medical Genetics Part A*, vol. 140A, no. 13, pp. 1407–1413, Jul. 1, 2006.

[86] E. Lausch, P. Hermanns, H. F. Farin, Y. Alanay, S. Unger, S. Nikkel, C. Steinwender, G. Scherer, J. Spranger, B. Zabel, A. Kispert, and A. Superti-Furga, "TBX15 mutations cause craniofacial dysmorphism, hypoplasia of scapula and pelvis, and short stature in cousin syndrome", *The American Journal of Human Genetics*, vol. 83, no. 5, pp. 649–655, Nov. 17, 2008.

[87] E. Dikoglu, P. O. Simsek-Kiper, G. E. Utine, B. Campos-Xavier, K. Boduroglu, L. Bonafé, A. Superti-Furga, and S. Unger, "Homozygosity for a novel truncating mutation confirms TBX15 deficiency as the cause of cousin syndrome", *American Journal of Medical Genetics Part A*, vol. 161, no. 12, pp. 3161–3165, Dec. 1, 2013.

[88] E. Anderson, S. Peluso, L. A. Lettice, and R. E. Hill, "Human limb abnormalities caused by disruption of hedgehog signaling", *Trends in Genetics*, vol. 28, no. 8, pp. 364–373, Aug. 2012.

[89] S. Malik, "Syndactyly: Phenotypes, genetics and current classification", *European Journal of Human Genetics*, vol. 20, no. 8, pp. 817–824, Aug. 2012.

[90] S. Lohan, M. Spielmann, S. C. Doelken, R. Flöttmann, F. Muhammad, S. M. Baig, M. Wajid, W. Hülsemann, R. Habenicht, K. W. Kjaer, S. J. Patil, K. M. Girisha, H. H. Abarca-Barriga, S. Mundlos, and E. Klopocki, "Microduplications encompassing the sonic hedgehog limb enhancer ZRS are associated with haas-type polysyndactyly and laurin-sandrow syndrome", *Clinical Genetics*, vol. 86, no. 4, pp. 318–325, Oct. 2014.

[91] L. G. Biesecker, "Polydactyly: How many disorders and how many genes? 2010 update", *Developmental Dynamics*, vol. 240, no. 5, pp. 931–942, May 1, 2011.

[92] D. Duprez, C. Fournier-Thibault, and N. L. Douarin, "Sonic hedgehog induces proliferation of committed skeletal muscle cells in the chick limb", *Development*, vol. 125, no. 3, pp. 495–505, Feb. 1, 1998.

[93] J. J. Sanz-Ezquerro and C. Tickle, "Autoregulation of shh expression and shh induction of cell death suggest a mechanism for modulating polarising activity during chick limb development", *Development*, vol. 127, no. 22, pp. 4811–4823, Nov. 15, 2000.

[94] L. A. Lettice, S. Daniels, E. Sweeney, S. Venkataraman, P. S. Devenney, P. Gautier, H. Morrison, J. Fantes, R. E. Hill, and D. R. FitzPatrick, "Enhancer-adoption as a mechanism of human developmental disease", *Human Mutation*, vol. 32, no. 12, pp. 1492–1499, Dec. 1, 2011.

[95] A. Sowińska-Seidler, M. Socha, and A. Jamsheer, "Split-hand/foot malformation - molecular cause and implications in genetic counseling", *Journal of Applied Genetics*, vol. 55, no. 1, pp. 105–115, 2014.

[96] N. Tayebi, A. Jamsheer, R. Flöttmann, A. Sowinska-Seidler, S. C. Doelken, B. Oehl-Jaschkowitz, W. Hülsemann, R. Habenicht, E. Klopocki, S. Mundlos, and M. Spielmann, "Deletions of exons with regulatory activity at the DYNC1i1 locus are associated with split-hand/split-foot malformation: Array CGH screening of 134 unrelated families", *Orphanet Journal of Rare Diseases*, vol. 9, no. 1, p. 108, Jul. 29, 2014.

[97] P. H. G. Duijf, H. v. Bokhoven, and H. G. Brunner, "Pathogenesis of split-hand/split-foot malformation", *Human Molecular Genetics*, vol. 12, R51–R60, suppl 1 Feb. 4, 2003.

[98] K. Lezirovitz, S. R. P. Maestrelli, N. H. Cotrim, P. A. Otto, P. L. Pearson, and R. C. Mingroni-Netto, "A novel locus for split-hand/foot malformation associated with tibial hemimelia (SHFLD syndrome) maps to chromosome region 17p13.1–17p13.3", *Human Genetics*, vol. 123, no. 6, pp. 625–631, May 21, 2008.

[99] F. Petit, J. Andrieux, B. Demeer, L.-M. Collet, H. Copin, E. Boudry-Labis, F. Escande, S. Manouvrier-Hanu, and M. Mathieu-Dramard, "Split-hand/foot malformation with long-bone deficiency and BHLHA9 duplication: Two cases and expansion of the phenotype to radial agenesis", *European journal of medical genetics*, vol. 56, no. 2, pp. 88–92, Feb. 2013.

[100] E. Nagata, N. Haga, Y. Fujisawa, M. Fukami, G. Nishimura, and T. Ogata, "Femoral-tibial-digital malformations in a boy with the japanese founder triplication of BHLHA9", *American Journal of Medical Genetics Part A*, n/a–n/a, Sep. 1, 2015.

[101] E. Klopocki, S. Lohan, S. C. Doelken, S. Stricker, C. W. Ockeloen, R. S. T. d. Aguiar, K. Lezirovitz, R. C. M. Netto, A. Jamsheer, H. Shah, I. Kurth, R. Habenicht, M. Warman, K. Devriendt, U. Kordaß, M. Hempel, A. Rajab, O. Mäkitie, M. Naveed, U. Radhakrishna, S. E. Antonarakis, D. Horn, and S. Mundlos, "Duplications of BHLHA9 are associated with ectrodactyly and tibia hemimelia inherited in non-mendelian fashion", *Journal of Medical Genetics*, vol. 49, no. 2, pp. 119–125, Jan. 2, 2012.

[102] F Petit, A.-S. Jourdain, J Andrieux, G Baujat, C Baumann, C Beneteau, A David, L Faivre, D Gaillard, B Gilbert-Dussardier, P.-S. Jouk, C Le Caignec, P Loget, L Pasquier, N Porchet, M Holder-Espinasse, S Manouvrier-Hanu, and F Escande, "Split hand/foot malformation with long-bone deficiency and BHLHA9 duplication: Report of 13 new families", *Clinical genetics*, Jun. 21, 2013.

[103] S. Jones, "An overview of the basic helix-loop-helix proteins", *Genome Biology*, vol. 5, no. 6, p. 226, May 28, 2004.

[104] O. Schatz and E. L. \. N. Ben-Arie, "Gene dosage of the transcription factor fingerin (bHLHA9) affects digit development and links syndactyly to ectrodactyly", *Human Molecular Genetics*, ddu257, May 22, 2014.

[105] A. Afshar, "Combined congenital radial and ulnar longitudinal deficiencies: A case report", *Journal of Hand and Microsurgery*, vol. 7, no. 1, pp. 191–193, Jun. 2015.

[106] K. C. Oberg, J. M. Feenstra, P. R. Manske, and M. A. Tonkin, "Developmental biology and classification of congenital anomalies of the hand and upper extremity", *The Journal of Hand Surgery*, vol. 35, no. 12, pp. 2066–2076, Dec. 2010.

[107] A. Jamsheer, A. Sowińska, D. Simon, M. Jamsheer-Bratkowska, T. Trzeciak, and A. Latos-Bieleńska, "Bilateral radial agenesis with absent thumbs, complex heart defect, short stature, and facial dysmorphism in a patient with pure distal microduplication of 5q35.2-5q35.3", *BMC Medical Genetics*, vol. 14, p. 13, Jan. 24, 2013.

[108] E Steichen-Gersdorf, I Gassner, A Superti-Furga, R Ullmann, S Stricker, E Klopocki, and S Mundlos, "Triangular tibia with fibular aplasia associated with a microdeletion on 2q11.2 encompassing LAF4", *Clinical genetics*, vol. 74, no. 6, pp. 560–565, Dec. 2008.

[109] M. Urban and S. Krüger, "Alice vance ("das bärenweib"): A historical case of nievergelt syndrome", *American Journal of Medical Genetics*, vol. 76, no. 2, pp. 145–149, Mar. 5, 1998.

[110] M. Vasil, A. Baxova, and K. Kozlowski, "Severe limb abnormalities: Nievergelt or new syndrome?", *American Journal of Medical Genetics*, vol. 70, no. 1, pp. 48–51, May 2, 1997.

[111] R. Petrella, M. D. Ludman, J. G. Rabinowitz, F. Gilbert, and K. Hirschhorn, "Mesomelic dysplasia with absence of fibulae and hexadactyly: Nievergelt syndrome or new syndrome?", *American Journal of Medical Genetics*, vol. 37, no. 1, pp. 10–14, Sep. 1, 1990.

[112] R. Savarirayan, V. Cormier-Daire, C. J. Curry, M. B. Nashelsky, V. Rappaport, D. L. Rimoin, and R. S. Lachman, "New mesomelic dysplasia with absent fibulae and triangular tibiae", *American Journal of Medical Genetics*, vol. 94, no. 1, pp. 59–63, Sep. 4, 2000.

[113] B. Tüysüz, C. Zeybek, G. Zorer, z. Sipahi, and S. Üngür, "Patient with the mesomelic dysplasia, nievergelt syndrome, and cerebellovermian agenesis and cataracts", *American Journal of Medical Genetics*, vol. 109, no. 3, pp. 206–210, May 1, 2002.

[114] P. Monga, M. Swamy, and K. Rao, "Rhomboid shaped tibia and hypoplastic fibula: A variant of nievergelt syndrome", *American Journal of Medical Genetics Part A*, vol. 118A, no. 4, pp. 394–397, May 1, 2003.

[115] L. Bonafe, V. Cormier-Daire, C. Hall, R. Lachman, G. Mortier, S. Mundlos, G. Nishimura, L. Sangiorgi, R. Savarirayan, D. Sillence, J. Spranger, A. Superti-Furga, M. Warman, and S. Unger, "Nosology and classification of genetic skeletal disorders: 2015 revision", *American Journal of Medical Genetics. Part A*, Sep. 23, 2015.

[116] R. Flöttmann, J. Wagner, K. Kobus, C. J. Curry, R. Savarirayan, G. Nishimura, N. Yasui, J. Spranger, H. Van Esch, M. J. Lyons, B. R. DuPont, A. Dwivedi, E. Klopocki, D. Horn, S. Mundlos, and M. Spielmann, "Microdeletions on 6p22.3 are associated with mesomelic dysplasia savarirayan type", *Journal of Medical Genetics*, vol. 52, no. 7, pp. 476–483, Jul. 2015.

[117] F. Wittwer, A. van der Straten, K. Keleman, B. J. Dickson, and E. Hafen, "Lilliputian: An AF4/FMR2-related protein that controls cell identity and cell growth", *Development (Cambridge, England)*, vol. 128, no. 5, pp. 791–800, Mar. 2001.

[118] M. A. Su, R. G. Wisotzkey, and S. J. Newfeld, "A screen for modifiers of decapentaplegic mutant phenotypes identifies lilliputian, the only member of the fragile-x/burkitt's lymphoma family of transcription factors in drosophila melanogaster", *Genetics*, vol. 157, no. 2, pp. 717–725, Feb. 2001.

[119] M. Melko, D. Douguet, M. Bensaid, S. Zongaro, C. Verheggen, J. Gecz, and B. Bardoni, "Functional characterization of the AFF (AF4/FMR2) family of RNA-binding proteins: Insights into the molecular pathology of FRAXE intellectual disability", *Human molecular genetics*, vol. 20, no. 10, pp. 1873–1885, May 15, 2011.

[120] C. Ma and L. M. Staudt, "LAF-4 encodes a lymphoid nuclear protein with transactivation potential that is homologous to AF-4, the gene fused to MLL in t(4;11) leukemias", *Blood*, vol. 87, no. 2, pp. 734–745, Jan. 15, 1996.

[121] A. Barton, S. Eyre, X. Ke, A. Hinks, J. Bowes, E. Flynn, P. Martin, YEAR Consortium, BIRAC Consortium, A. G. Wilson, A. W. Morgan, P. Emery, S. Steer, L. J. Hocking, D. M. Reid, P. Harrison, P. Wordsworth, W. Thomson, and J. Worthington, "Identification of AF4/FMR2 family, member 3 (AFF3) as a novel rheumatoid arthritis susceptibility locus and confirmation of two further pan-autoimmune susceptibility genes", *Human Molecular Genetics*, vol. 18, no. 13, pp. 2518–2522, Jul. 1, 2009.

[122] F. Correia, C. Café, J. Almeida, S. Mouga, and G. Oliveira, "Autism spectrum disorder: FRAXE mutation, a rare etiology", *Journal of Autism and Developmental Disorders*, vol. 45, no. 3, pp. 888–892, Mar. 2015.

[123] O. Britanova, S. Lukyanov, P. Gruss, and V. Tarabykin, "The mouse laf4 gene: Exon/intron organization, cDNA sequence, alternative splicing, and expression during central nervous system development", *Genomics*, vol. 80, no. 1, pp. 31–37, Jul. 2002.

[124] P. D. Hsu, D. A. Scott, J. A. Weinstein, F. A. Ran, S. Konermann, V. Agarwala, Y. Li, E. J. Fine, X. Wu, O. Shalem, T. J. Cradick, L. A. Marraffini, G. Bao, and F. Zhang, "DNA targeting specificity of RNA-guided cas9 nucleases", *Nature biotechnology*, vol. 31, no. 9, pp. 827–832, Sep. 2013.

[125] L. Kroeldrup, S. Kjaergaard, M. Kirchhoff, K. Kock, C. Brasch-Andersen, M. Kibaek, and L. B. Ousager, "Duplication of 7q36.3 encompassing the sonic hedgehog (SHH) gene is associated with congenital muscular hypertrophy", *European Journal of Medical Genetics*, vol. 55, no. 10, pp. 557–560, Oct. 2012.

[126] A. Gupta, V. L. Hall, F. O. Kok, M. Shin, J. C. McNulty, N. D. Lawson, and S. A. Wolfe, "Targeted chromosomal deletions and inversions in zebrafish", *Genome Research*, vol. 23, no. 6, pp. 1008–1017, Jun. 2013.

[127] T. Cremer and M. Cremer, "Chromosome territories", *Cold Spring Harbor Perspectives in Biology*, vol. 2, no. 3, a003889, Mar. 2010.

[128] A. J. Fritz, A. R. Barutcu, L. Martin-Buley, A. J. van Wijnen, S. K. Zaidi, A. N. Imbalzano, J. B. Lian, J. L. Stein, and G. S. Stein, "Chromosomes at work: Organization of chromosome territories in the interphase nucleus", *Journal of Cellular Biochemistry*, vol. 117, no. 1, pp. 9–19, Jan. 1, 2016.

[129] X. Chen, M. Li, X. Feng, and S. Guang, "Targeted chromosomal translocations and essential gene knockout using CRISPR/cas9 technology in caenorhabditis elegans", *Genetics*, genetics.115.181883, Oct. 19, 2015.

[130] R. Torres, M. C. Martin, A. Garcia, J. C. Cigudosa, J. C. Ramirez, and S. Rodriguez-Perales, "Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR–cas9 system", *Nature Communications*, vol. 5, p. 3964, Jun. 3, 2014.

[131] Y. Luo and J. C. Schimenti, "MCM9 deficiency delays primordial germ cell proliferation independent of the ATM pathway", *Genesis (New York, N.Y.: 2000)*, Sep. 21, 2015.

[132] A. R. Leman and E. Noguchi, "Linking chromosome duplication and segregation via sister chromatid cohesion", *Methods in molecular biology (Clifton, N.J.)*, vol. 1170, pp. 75–98, 2014.

[133] M. A. Keene, V. Corces, K. Lowenhaupt, and S. C. Elgin, "DNase i hypersensitive sites in drosophila chromatin occur at the 5' ends of regions of transcription", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 78, no. 1, pp. 143–146, Jan. 1981.

[134] S. C. Kowalczykowski, "An overview of the molecular mechanisms of recombinational DNA repair", *Cold Spring Harbor Perspectives in Biology*, vol. 7, no. 11, 2015, bibtex: zotero-null-547.

[135] J. Song, D. Yang, J. Xu, T. Zhu, Y. E. Chen, and J. Zhang, "RS-1 enhances CRISPR/cas9- and TALEN-mediated knock-in efficiency", *Nature Communications*, vol. 7, p. 10 548, Jan. 28, 2016.

[136] F. A. Ran, P. Hsu, C.-Y. Lin, J. Gootenberg, S. Konermann, A. E. Trevino, D. Scott, A. Inoue, S. Matoba, Y. Zhang, and F. Zhang, "Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity", *Cell*, vol. 154, no. 6, pp. 1380–1389, Sep. 12, 2013.

[137] J. Li, J. Shou, Y. Guo, Y. Tang, Y. Wu, Z. Jia, Y. Zhai, Z. Chen, Q. Xu, and Q. Wu, "Efficient inversions and duplications of mammalian regulatory DNA elements and gene clusters by CRISPR/cas9", *Journal of Molecular Cell Biology*, vol. 7, no. 4, pp. 284–298, Jan. 8, 2015.

[138] K. Kraft, S. Geuer, A. J. Will, W. L. Chan, C. Paliou, M. Borschiwer, I. Harabula, L. Wittler, M. Franke, D. M. Ibrahim, B. K. Kragesteen, M. Spielmann, S. Mundlos, D. G. Lupiáñez, and G. Andrey, "Deletions, inversions, duplications: Engineering of structural variants using CRISPR/cas in mice", *Cell Reports*, Feb. 4, 2015.

[139] A. D. Berendsen and B. R. Olsen, "Bone development", *Bone*, vol. 80, pp. 14–18, Nov. 2015.

[140] B. Alman, "Skeletal dysplasias and the growth plate", *Clinical Genetics*, vol. 73, no. 1, pp. 24–30, Jan. 1, 2008.

[141] P. Kuss, K. Kraft, J. Stumm, D. Ibrahim, P. Vallecillo-Garcia, S. Mundlos, and S. Stricker, "Regulation of cell polarity in the cartilage growth plate and perichondrium of metacarpal elements by HOXD13 and WNT5a", *Developmental Biology*, vol. 385, no. 1, pp. 83–93, Jan. 1, 2014.

[142] M.-G. Ascenzi, C. Blanco, I. Drayer, H. Kim, R. Wilson, K. N. Retting, K. M. Lyons, and G. Mohler, "Effect of localization, length and orientation of chondrocytic primary cilium on murine growth plate organization", *Journal of Theoretical Biology*, vol. 285, no. 1, pp. 147–155, Sep. 21, 2011.

[143] Y. Li and A. T. Dudley, "Noncanonical frizzled signaling regulates cell polarity of growth plate chondrocytes", *Development (Cambridge, England)*, vol. 136, no. 7, pp. 1083–1092, Apr. 1, 2009.

[144] R. Feil, J. Wagner, D. Metzger, and P. Chambon, "Regulation of cre recombinase activity by mutated estrogen receptor ligand-binding domains", *Biochemical and Biophysical Research Communications*, vol. 237, no. 3, pp. 752–757, Aug. 28, 1997.

[145] S. Metsu, L. Rooms, J. Rainger, M. S. Taylor, H. Bengani, D. I. Wilson, C. S. R. Chilamakuri, H. Morrison, G. Vandeweyer, E. Reyniers, E. Douglas, G. Thompson, E. Haan, J. Gecz, D. R. FitzPatrick, and R. F. Kooy, "FRA2a is a CGG repeat expansion associated with silencing of AFF3", *PLOS Genet*, vol. 10, no. 4, e1004242, Apr. 24, 2014.

[146] J. Yingling, K. Toyo-Oka, and A. Wynshaw-Boris, "Miller-dieker syndrome: Analysis of a human contiguous gene syndrome in the mouse", *American Journal of Human Genetics*, vol. 73, no. 3, pp. 475–488, Sep. 2003.

[147] M. Buckingham, L. Bajard, T. Chang, P. Daubas, J. Hadchouel, S. Meilhac, D. Montarras, D. Rocancourt, and F. Relaix, "The formation of skeletal muscle: From somite to limb", *Journal of Anatomy*, vol. 202, no. 1, pp. 59–68, Jan. 2003.

[148] C. Anderson, V. Williams, B. Moyon, P. Daubas, S. Tajbakhsh, M. Buckingham, T. Shiroishi, S. Hughes, and A.-G. Boryck, "Sonic hedgehog acts cell-autonomously on muscle precursor cells to generate limb muscle diversity", *Genes and Development*, vol. 26, no. 18, pp. 2103–2117, 2012.

[149] T. Endo, "Molecular mechanisms of skeletal muscle development, regeneration, and osteogenic conversion", *Bone*, Muscle Bone Interactions, vol. 80, pp. 2–13, Nov. 2015.

[150] K. Wong, R. Moldrich, M. Hunter, M. Edwards, D. Finlay, S. O'Donnell, T. MacDougall, N. Bain, and B. Kamien, "A familial 7q36.3 duplication associated with agenesis of the corpus callosum", *American Journal of Medical Genetics Part A*, vol. 167, no. 9, pp. 2201–2208, Sep. 1, 2015.

[151] R. Atit, S. K. Sgaier, O. A. Mohamed, M. M. Taketo, D. Dufort, A. L. Joyner, L. Niswander, and R. A. Conlon, "-catenin activation is necessary and sufficient to specify the dorsal dermal fate in the mouse", *Developmental Biology*, vol. 296, no. 1, pp. 164–176, Aug. 1, 2006.

[152] P. Seale, S. Kajimura, W. Yang, S. Chin, L. M. Rohas, M. Uldry, G. Tavernier, D. Langin, and B. M. Spiegelman, "Transcriptional control of brown fat determination by PRDM16", *Cell Metabolism*, vol. 6, no. 1, pp. 38–54, Jul. 11, 2007.

[153] P. Seale, B. Bjork, W. Yang, S. Kajimura, S. Chin, S. Kuang, A. Scimè, S. Devarakonda, H. M. Conroe, H. Erdjument-Bromage, P. Tempst, M. A. Rudnicki, D. R. Beier, and B. M. Spiegelman, "PRDM16 controls a brown fat/skeletal muscle switch", *Nature*, vol. 454, no. 7207, pp. 961–967, Aug. 21, 2008.

[154] P. Hasty, A. Bradley, J. H. Morris, D. G. Edmondson, J. M. Venuti, E. N. Olson, and W. H. Klein, "Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene", *Nature*, vol. 364, no. 6437, pp. 501–506, Aug. 5, 1993.

[155] B. Kablar, K. Krastel, S. Tajbakhsh, and M. A. Rudnicki, "Myf5 and MyoD activation define independent myogenic compartments during embryonic development", *Developmental Biology*, vol. 258, no. 2, pp. 307–318, Jun. 15, 2003.

[156] J. K.-H. Hu, E. McGlinn, B. D. Harfe, G. Kardon, and C. J. Tabin, "Autonomous and nonautonomous roles of hedgehog signaling in regulating limb muscle formation", *Genes & Development*, vol. 26, no. 18, pp. 2088–2102, Sep. 15, 2012.

[157] K. L. Obholz, A. Akopyan, K. G. Waymire, and G. R. MacGregor, "FNDC3a is required for adhesion between spermatids and sertoli cells", *Developmental Biology*, vol. 298, no. 2, pp. 498–513, Oct. 15, 2006.

[158] F. Carrouel, M.-L. Couble, C. Vanbelle, M.-J. Staquet, H. Magloire, and F. Bleicher, "HUGO (FNDC3a): A new gene overexpressed in human odontoblasts", *Journal of Dental Research*, vol. 87, no. 2, pp. 131–136, Feb. 2008.

[159] R. Xu, A. Boudreau, and M. J. Bissell, "Tissue architecture and function: Dynamic reciprocity via extra- and intra-cellular matrices", *Cancer metastasis reviews*, vol. 28, no. 1, pp. 167–176, Jun. 2009.

[160] N. Wada, "Spatiotemporal changes in cell adhesiveness during vertebrate limb morphogenesis", *Developmental Dynamics*, vol. 240, no. 5, pp. 969–978, May 1, 2011.

[161] T. P. Yamaguchi, A. Bradley, A. P. McMahon, and S. Jones, "A wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo", *Development*, vol. 126, no. 6, pp. 1211–1223, Mar. 15, 1999.

[162] D. Craig, M. Gao, K. Schulten, and V. Vogel, "Tuning the mechanical stability of fibronectin type III modules through sequence variations", *Structure*, vol. 12, no. 1, pp. 21–30, Mar. 16, 2004.

[163] R. Pankov and K. M. Yamada, "Fibronectin at a glance", *Journal of Cell Science*, vol. 115, no. 20, pp. 3861–3863, Oct. 15, 2002.

[164] E. Hoxha and S. R. Campion, "Structure-critical distribution of aromatic residues in the fibronectin type III protein family", *The protein journal*, Feb. 22, 2014.

[165] R. N. Gilbreth, B. M. Chacko, L. Grinberg, J. S. Swers, and M. Baca, "Stabilization of the third fibronectin type III domain of human tenascin-c through minimal mutation and rational design", *Protein engineering, design & selection: PEDS*, vol. 27, no. 10, pp. 411–418, Oct. 2014.

[166] L. Demetrius, "Aging in mouse and human systems", *Annals of the New York Academy of Sciences*, vol. 1067, no. 1, pp. 66–82, May 1, 2006.

[167] R. Chamanza and J. A. Wright, "A review of the comparative anatomy, histology, physiology and pathology of the nasal cavity of rats, mice, dogs and non-human primates. relevance to inhalation toxicology and human health risk assessment", *Journal of Comparative Pathology*,

[168] J. Paton-Hough, A. D. Chantry, and M. A. Lawson, "A review of current murine models of multiple myeloma used to assess the efficacy of therapeutic agents on tumour growth and bone disease", *Bone*, vol. 77, pp. 57–68, Aug. 2015, bibtex: zotero-null-638.

[169] K. A. Maupin, C. J. Droscha, and B. O. Williams, "A comprehensive overview of skeletal phenotypes associated with alterations in wnt/-catenin signaling in humans and mice", *Bone Research*, vol. 1, no. 1, pp. 27–71, Mar. 29, 2013.

[170] S. Dutta and P. Sengupta, "Men and mice: Relating their ages", *Life Sciences*,

[171] A. Martin, S. Maher, K. Summerhurst, D. Davidson, and P. Murphy, "Differential deployment of paralogous wnt genes in the mouse and chick embryo during development", *Evolution & Development*, vol. 14, no. 2, pp. 178–195, Apr. 2012.

[172] F. Fougerousse, P. Bullen, M. Herasse, S. Lindsay, I. Richard, D. Wilson, L. Suel, M. Durand, S. Robson, M. Abitbol, J. S. Beckmann, and T. Strachan, "Human-mouse differences in the embryonic expression patterns of developmental control genes and disease genes", *Human Molecular Genetics*, vol. 9, no. 2, pp. 165–173, Jan. 22, 2000.

[173] I. Barbaric, G. Miller, and T. N. Dear, "Appearances can be deceiving: Phenotypes of knockout mice", *Briefings in Functional Genomics & Proteomics*, vol. 6, no. 2, pp. 91–103, Jan. 6, 2007.

8 Scientific Publications and Awards

Article

K. Kraft* & **S. Geuer***, A. J. Will, W. L. Chan, C. Paliou, M. Borschiwer, I. Harabula, L. Wittler, M. Franke, D. M. Ibrahim, B. K. Kragesteen, M. Spielmann, S. Mundlos, D. G. Lupiáñez, and G. Andrey, "Deletions, inversions, duplications: Engineering of structural variants using CRISPRcas in mice", Cell Reports, Feb. 4, 2015.

*Corresponding first authors

Scientific Congresses and Courses

2015

Annual Conference of German Society for Human Genetics (GfH)
Oral presentation: Intragenic Truncation of Laf4 causes Nievergelt-like Syndrome

Pro Retina Potsdam Meeting 2015

2014

ESHG - 3rd Course in Next Generation Sequencing
ESHG Fellowship for the 3rd Course in Next Generation Sequencing
Poster prize: Missense Mutation in FNDC3A causes Split-Hand/Foot Malformation

2013

ESHG - 26th Course in Medical Genetics
ESHG Fellowship for the 26th Course in Medical Genetics
Poster prize: Missense Mutation in FNDC3A causes Split-Hand/Foot Malformation

Annual Conference of German Society for Human Genetics (GfH)
Oral presentation: Missense Mutation in FNDC3A causes Split-Hand/Foot Malformation

2012

European Human Genetics Conference 2015 (ESHG)

A Abbreviations

3C	Chromosome Conformation Capture
AER	Apical Ectodermal Region
arrayCGH	Array Comparative Genome Hybridization
Chr.	Chromosome
CNV	Copy Number Variation
CRISRP	Clustered Regularly Interspaced Palindromic Repeats
del	deletion
DIC	Differential Interference Contrast
dup	duplication
ECM	Extracellular Matrix
ESC = ES cell	Embryonic Stem Cell
FnIII domain	Fibronectin type III domain
HDR	Homology Directed Repair
Hg19	Human Genome Version 19
Hi-C	High throughput 3C
indel	small region of few bp insertion and deletion (often caused by NHEJ)
KO	Knock Out
NAHR	Nonallelic Homologue Recombination
NHEJ	Nonhomologue End Joining
NLS	Nucleus Localization Signal
NoLS	Nucleoli Localization Signal
ORF	Open Reading Frame
PAM	Protospacer Adjacent Motif
sgRNA	single guide RNA
SHFM	Split-Hand/Foot Malformation
SHFLD	Split-Hand/Foot Malformation with Long Bone Deficiency
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variation
ssODN	single-stranded Oligo Donor Nucleotide
SV	Structure Variation
TAD	Topological Associated Domain
TALEN	Transcription Activator-Like Effector Nucleases
TSS	Transcription Start Site
WB	Western Blot
WISH	Whole Mount *in situ* Hybridization
ZFN	Zink Finger Nuclease
ZPA	Zone of Polarizing Activity