UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

Creation of a *Danio rerio* mutant using CRISPR-Cas9 as a model system to study Primary Ciliary Dyskinesia (PCD)

Margarida Oliveira Almeida Rasteiro

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:
Doutora Susana Lopes
Professora Doutora Rita Zilhão

2017

1.1. Abstract

On the surface, vertebrates seem to have a bilateral symmetry. However, the disposition of the internal organs, such as the heart and liver says otherwise. The left-right axis differentiation is preceded by asymmetries in the gene expression pattern in the tissues around the left-right organiser (LRO).

The LRO (called Kupffer's vesicle in zebrafish) is a transient structure localized at the end of the notochord that is formed by cells contain a cilium protruding from its apical membrane. Most of these cilia are motile and their beating movement creates a fluid flow towards the left side of the LRO. In a way not yet fully understood, the cells surrounding the LRO sense the fluid directionality and trigger asymmetric expression of *nodal* and *cerl2*. The expression of *nodal* on the left side of the LRO triggers the Nodal-Lefty-Pitx2 pathway that leads to the typical positioning of the internal organs recognised as *situs solitus.*

Besides the laterality establishment, the respiratory epithelium, the ependymal cells lining the brain ventricles, the oviducts epithelium and the sperm cells also relay on motile cilia to function properly.

When motile cilia function is impaired, this sequence of events is not guaranteed and patients may suffer from heart congenital malformations bronchiectasis and infertility that characterize primary ciliary dyskinesia (PCD). PCD is caused by mutations in a vast number of genes that codify for cilia components. One of those genes is called *CCDC40* and codifies for a protein necessary for the assembly of the inner dynein arms (IDA) and the nexin-dynein regulatory complexes (N-DRC).

In this work, we generated a *ccdc40-/-* zebrafish mutant (116 aa) using a CRISPR-Cas9 approach.

Due to the long zebrafish maturation period, it was not possible to study the mutant phenotype. Instead, we studied the phenotype of embryos injected with a translation-blocking morpholino (MO). We noted that the organ *situs* was altered in 40 to 70% of the cases, (depending on the amount of MO-injected) and the developing of tail malformations (not related to the injected MO amount).

The *ccdc40-/-* zebrafish mutant was designed to replicate a human mutation and can be used to test gene editing approaches in an attempt to develop a PCD treatment.

Keywords: Motile cilia, Primary ciliary dyskinesia (PCD); CCDC40, CRISPR-Cas9, Zebrafish

1.2. Resumo alargado

Os vertebrados apresentam uma organização corporal complexa com três eixos: eixo antero-posterior, o eixo dorso-ventral e o eixo esquerda-direita. As assimetrias nos eixos antero-posterior e dorso-ventral são facilmente identificáveis exteriormente. No entanto, no que diz respeito ao eixo esquerda direita, as diferenças são notadas unicamente na disposição interna dos órgãos. No Homem, o coração está deslocado para o lado esquerdo, assim como o estômago e o baço, enquanto que o fígado e a vesícula biliar ficam do lado direito. Também os pulmões, devido à posição do coração, apresentam assimetria no número de lóbulos. Esta disposição (*situs solitus)* é determinada cedo durante o desenvolvimento embrionário e mantem-se conservada nos vertebrados. A diferenciação do eixo esquerda-direita é precedida por assimetrias na expressão de genes no organizador esquerda-direita (LRO, na sigla em inglês). O LRO é uma estrutura transiente localizada no final da notocorda, que no peixe-zebra é chamada de vesícula de Kupffer (KV).

Esta vesícula é formada por células, cada uma com um cílio diferenciado na sua membrana apical. A maioria destes cílios são móveis e com o seu movimento coordenado produzem um fluxo direcionado para o lado esquerdo do LRO. Esta assimetria no fluxo é depois transferida para uma assimetria na expressão de genes nos tecidos à volta do LRO.

Inicialmente há dois genes que são expressos de forma simétrica à volta do LRO: *nodal* e *cerl2*. A forma como o fluxo consegue controlar a expressão de *nodal* e *cerl2* ainda não é conhecida, mas foram propostas duas hipóteses que a tentam explicar. A primeira propõe que o fluxo gerado pelos cílios móveis transporta substâncias secretadas pelo LRO e as concentra do lado esquerdo, levando à assimetria na expressão *nodal* e *cerl2* verificada posteriormente. A outra hipótese propõe que os cílios móveis criam o fluxo de fluído que, no lado esquerdo, vai causar a deflação dos cílios imóveis, estimulando os seus mecanorecetores e induzindo uma via de ativação dependente de Ca^{2+} que culmina na expressão assimétrica de *nodal* e *cerl2*. No entanto, apesar dos esforços, ambas as hipóteses têm sido questionadas e o exacto mecanismo de transdução da informação permanece desconhecido.

Após a ação do fluxo, que é sentido pelas células que rodeiam o LRO, a expressão de *cerl2* é inibida no lado esquerdo. A inibição da expressão de *cerl2* (inibidor de *nodal*) permite uma maior expressão de *nodal* e consequentemente a ativação da cascata de sinalização Nodal-Lefty-Pitx2. Pitx2 é a proteína efetora da cascata iniciada por Nodal, sendo apontada como a responsável por induzir o fenótipo "esquerdo" nas células do lado esquerdo da mesoderme lateral (L-LPM).

Os cílios, que podem ser móveis ou imóveis (primários), estão presentes em quase todos os tipos células nos vertebrados e por isso, são determinantes quer no período de desenvolvimento embrionário quer na idade adulta. Quando o seu funcionamento fica comprometido pode desencadear várias doenças, como por exemplo a doença do rim poliquístico (PKD, na sigla em inglês) ou síndromes somáticas associadas a polidactilia, malformações neurológicas ou obesidade, no caso de alterações associadas aos cílios primários. Quando são os cílios móveis afetados, as alterações manifestam-se através de infertilidade, infeções respiratórias recorrentes e *situs inversus* que caracterizam a uma condição chamada discinesia ciliar primaria (PCD). A PCD afecta 1 em cada 10 000 nascimentos e é uma das causas de defeitos cardíacos congénitos. A presença de heterotaxia, condição na qual, pelo menos um órgão está no lado oposto ao que seria de esperar, aumenta 200 vezes a prevalência destes defeitos congénitos.

A PC pode ser causada por mutações que alterem ou inibam a produção de qualquer um dos componentes necessários à montagem das estruturas essenciais para o movimento dos cílios, como o par central, os raios, o complexo regulador da nexina, os braços interno e externo de dineínas ou outras proteínas citoplasmáticas que participam na montagem de componentes ciliares.

Este trabalho foca-se na proteína ccdc40, que é um dos componentes do braço interno de dineínas. Esta proteína é também responsável pela poliglutamilação dos microtúbulos contribuindo para a sua estabilização. Além disso, a proteína Ccdc40 forma um complexo com outra, a Ccdc39, e atua como uma régua molecular. Este complexo permite que os raios dos cílios apenas se liguem no local devido ao longo dos microtúbulos, isto é, a cada 96 nm. Na sua ausência, estes raios continuam a ligar-se aos microtúbulos, mas em locais onde isso não era suposto acontecer.

Em pacientes com PCD portadores de alterações neste gene, foram identificados locais da proteína onde as mutações são mais frequentemente encontradas: as mutações c.248delC e c.3129delC alteram a grelha de leitura enquanto que as outras (c.2440 C>T, c.961 C>T e c.1345C>T) produzem proteínas truncadas.

Neste trabalho, pretendeu-se criar um peixe-zebra mutante que produza a proteína Ccdc40 truncada e que possa servir como modelo de doença para estudar o impacto das mutações mais perto do terminal N da proteína (c.248delC e c.961 C>T). Estas mutações próximas da extremidade N representam cerca de 50% dos casos de mutações identificadas neste gene e produzem proteínas sem o domínio helicoidal (coiled-coil).

Para produzir este mutante usou-se o sistema CRISPR-Cas9. Este sistema foi adaptado a partir de um mecanismo de defesa descoberto em bactérias e permite direcionar a Cas9 (DNase) para um local específico do genoma, usando uma sequência guia de RNA complementar ao local escolhido. Uma vez clivado o DNA, a célula inicia o processo de reparação através da junção das extremidades não

homólogas (NHEJ), processo este que é propenso à inserção ou remoção de bases azotadas, levando à ocorrência de mutações que podem inativar a proteína.

Neste caso, o local escolhido como alvo da Cas9 foi a zona do genoma que codifica para o aminoácido 116 da proteína ccdc40 no peixe zebra. Esta região é homóloga da região do genoma humano onde ocorre a mutação c.961 C>T. Uma mutação nonesense ou uma alteração da grelha na leitura neste local impede totalmente a tradução do domínio em hélice, o que causa alterações na montagem da estrutura interna dos cílios e consequentemente na sua mobilidade.

De forma a poder estudar e futuramente validar o fenótipo resultante da injeção do morfolino bloqueador da tradução, foram analisados a posição do coração, fígado e pâncreas assim como defeitos na cauda e aparecimento de edema cardíaco, em embriões injetados com o morfolino. Os resultados mostraram que, independentemente da concentração injetada, os embriões apresentam defeitos relativos à posição do coração (coração à direita ou ao centro), fígado e pâncreas, assim como curvaturas na cauda e aparecimento de edema por volta dos 3dpf. Estes resultados, irão futuramente ser comparados com o fenótipo observado no mutante produzido, validando ou não este morfolino.

Além disso, este mutante permitirá estudar as alterações estruturais e funcionais dos cílios em que a proteína Ccdc40 apresenta uma mutação próxima da extremidade N e compará-las com a estrutura dos cílios de outa linha mutante já existente, *lok*, em que a proteína é mutada próximo da extremidade C e ainda possui parte do domínio helicoidal.

Futuramente, este mutante poderá ser útil para testar terapia génica direcionada à correção desta mutação que poderá mais tarde ser reproduzida em células humanas recolhidas de pacientes com PCD com o intuito de serem reimplantadas, possibilitando uma forma de tratamento que possa minimizar os efeitos da PCD no aparelho respiratório.

Palavras-chave: Cílios móveis, Discinesia ciliar primária, Ccdc40, CRISPR-Cas9, Peixe-zebra

1.3. Contents

1.4. List of Figures

1.5. List of Tables

1.6. Abbreviations list

aa	Amino acid
ATP	Adenosine triphosphate
Ca^{2+}	Calcium ion
CCDC39/40/65	Coiled-coil domain-containing protein 39/40/65
CHD	Chronic heart disease
CRISPR-Cas9	Clustered regularly interspaced palindromic repeats associated with caspase 9
Dand5/Cerl-2	DAN domain family member 5/Cerberus-like protein 2
DNAH5	Dynein axonemal heavy chain 5
DNAI	Dynein intermediate chain
DNAL	Dynein light chain
dNTP	Deoxynucleotide triphosphate
dpf	Days post fertilization
GRP	Gastrocoel roof plate
h	Hour(s)
hpf	Hours post fertilization
IDA	Inner dynein arm
iv	*Inversus viscerum*
Kif3A/B	Kinesin-Associated Protein 3A/B
KO	Knockout
KS	Kartagener syndrome
KV	Kupffer's vesicle
LB medium	Luria-Bertani medium
LPM	Lateral plate mesoderm
L-R	Left-right
lrd	Left-right dynein
LRO	Left-right organiser
min	Minute(s)
MO	Morpholino
MTOC	Microtubule organising centre
N-DRC	Nexin-dynein regulatory complex
O/N	Overnight
ODA	Outer dynein arm
PAM	Protospacer adjacent motif
PCD	Primary ciliary dyskinesia
PCR	Polymerase chain reaction
Pitx2	Pituitary homeobox 2
PKD	Polycystic kidney disease
PKD1L1	Polycystin 1 Like 1
PKD2	Polycystin 2
RA	Retinoic acid
rNTP	Ribonucleotide triphosphate
rpm	Rotations per min
RS	Radial spoke
RT	Room temperature
Shh	Sonic hedgehog
TGF	Transforming growth factor
V	Volts
WT	Wildtype

1. Introduction

1.1. Asymmetry axis formation during embryogenesis

Most of the living beings, specifically the vertebrates, have a complex body organization. The axis differentiation is crucial in the process of shaping multicellular organisms. Several existing pathways, spatially and temporally coordinated, cooperate to establish antero-posterior, dorso-ventral, and L-R (left-right) axis, breaking the embryo's initial radial symmetry.

The embryo's initial radial symmetry is only apparent, the ovule goes through several maturation steps that result in an asymmetrical environment inside this cell. The dorso-ventral axis formation is dependent on the establishment of animal and vegetal poles during oogenesis. The maternal dorsalizing factors are located in the vegetal pole and immediately after fertilization, together with zygotic factors, induce the dorso-ventral axis establishment by acting on Wnt /β-catenin, BMP (Bone morphogenic protein), Nodal, and FGF (fibroblast growing factor) pathways (reviewed by Langdon and Mullins 2011 [1]).

The formation of the anterior-posterior axis ultimately establishes the tail, trunk and head. This body segmentation results from the interaction of three signalling pathways: BMP, Nodal and Wnt [2]. The tail forms when all these three are activated, if only Nodal is on, that region will be the trunk, and the head develops at the end where all three pathways are off [3].

The third and last asymmetry axis formed is the one that establishes sidedness (L-R). This asymmetry is first noticed in gene expression in the left-right organiser (LRO) and then spreads to the LPM (lateral plate mesoderm). Ultimately, this information is translated into an asymmetric organ disposition [4].

1.2. L-R asymmetry: Nodal-Lefty-Pitx2 pathway

At first glance, most vertebrates show bilateral symmetry. However, there is an asymmetry within the body. This asymmetry, referred to as *situs solitus,* is highly conserved across species: the heart, stomach, pancreas and spleen lay on the left side of the longitudinal axis, while the liver and gallbladder are on the right. Additionally, the lungs also show asymmetries in the number of lobes. These morphological asymmetries are preceded by asymmetries in gene expression on the LRO during early embryogenesis [5].

The LRO is a transient embryonic organ (cavity) positioned at the end of the notochord and accordingly to the animal model it is given a different designation: in zebrafish is called Kupffer's vesicle, in mouse is referred to as the node, in chick as Hensen's node[a] and in Xenopus as gastrocoel roof plate [6]. Even though these structures are morphologically diverse, the mechanisms behind the L-R asymmetry establishment seem to be conserved across vertebrates [7].

The exact mechanisms that led to the onset of the L-R asymmetry axis are not fully understood, but it is known that a directional fluid flow on the LRO is essential in its establishment [8]. The cells surrounding the LRO have motile cilia protruding from their apical membranes that rotate generating a leftward flow [9]. This asymmetry in the flow is then transcribed into asymmetries in gene expression in the LRO surrounding cells.

[a] The Hensen's node does not have ciliated epithelium. Instead, in chick, the left-side expression of Nodal is achieved by the LRO displacement to the left side of the L-R axis during early gastrulation. The notochord forms on the right side of the LRO restricting Nodal expression only to the left side without requiring flow [113].

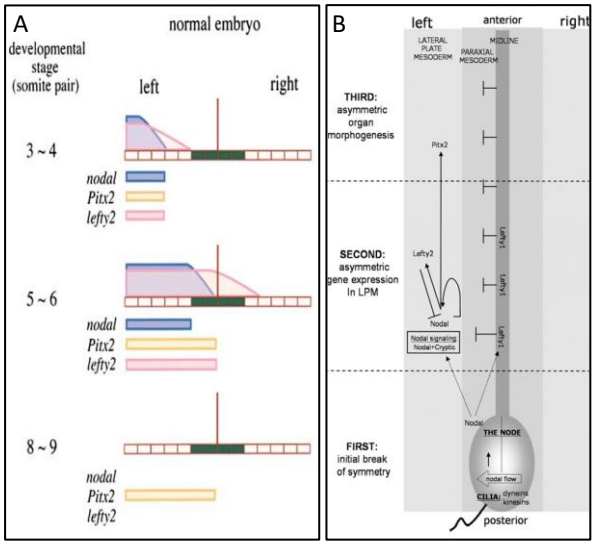


Figure 1.1 – Temporal and spatial distribution of Nodal, Pitx2 and Lefty2.
Lefty2 limits nodal diffusion by migrating faster and competing for the same receptors. Lefty1 is expressed in the midline (green) preventing Nodal from crossing to the right side. Pitx2 is expressed were Nodal is.

Horizontal aligned blocks stand for the LPM, the vertical brown line stands for the midline and blue, pink and yellow indicate the presence of Nodal, Lefty2 and Pitx2, respectively, along the LPM. Adapted from Meno *et al.* 2001 and Peeters *et al.* 2006 .

There are two specific genes, *nodal* and *dand5* (*nodal* inhibitor also known as *cerl2*), that are initially symmetrically expressed on the LRO. When the flow downregulates the expression of *dand5* on the left side of the LRO, it promotes the Nodal-signalling activation on this side [10]. As a consequence *nodal* expression becomes limited to the left side lateral plate mesoderm (Figure 1.1 A), the same side towards which the embryonic heart loops [5] while its antagonist, Dand5 is restricted to the right side of the LRO [11].

Nodal is a member of the TGFβ (transforming growth factor beta) family that, when lateralized on the LRO, spreads throughout the left-LPM where it induces its own expression as well as Lefty1, Lefty2 and Pitx2. Lefty1 and Lefty2, which are also members of the TGFβ family, compete with Nodal for the receptors preventing its action. Additionally, Nodal is only active as a dimmer while Lefty1 and Lefty2 are monomers, this lets them diffuse faster and farther, inhibiting the Nodal activity on the surrounding tissues. The *lefty1* expression is restricted to the left side of the ventral neural tube, where it acts as a midline barrier preventing the left-side establishing factors from crossing to the right side [12], while *lefty2* is expressed in the left LPM [13] (Figure 1.1 B).

Meno *et al*. demonstrated that, in *lefty2* mutants, Nodal expression lasts longer and can diffuse over to the right side [14]. Therefore, Lefty2 on itself does not decide which is going to be the left side, it functions as a regulator that restricts the duration of Nodal expression.

Finally, Pitx2 is the effector of Nodal signalling and is the most likely candidate for "informing" the cells on the left LPM that they are to adopt a leftward morphology [4]. However, there are some asymmetries such as the visceral looping that seem to be independent of Pitx2-signalling as observed in zebrafish [15].

1.3. How flow-receiving cells sense flow: two-cilia vs morphogen model

The asymmetric expression of Nodal is established when the motile cilia rotate and generate a fluid flow in the LRO [16]. The crown cells sense this leftward flow and trigger the Nodal-Lefty-Pitx2 pathway in a Ca^{2+} induced way [17]. This principle is supported by studies showing that mutants lacking motile cilia do not generate the required flow and thus have defects regarding gene expression and organ laterality [18]. Also, mutants having cilia with motility impairments display similar phenotype [19,20] making it clear that the defects are due to lack of movement and not because of the lack of cilia. The unequivocal connection between the cilia-generated flow in the LRO and the L-R asymmetry establishment was done by Nonaka *et al*. when they proved that the Nodal-Lefty-Pitx2 pathway asymmetry could be disturbed in WT mouse embryos or normalized in mutants showing laterality defects by inducing an artificially generated flow [16].

However, the exact mechanism on how the LRO senses the flow and transcribes it into gene expression is not yet clear. In an attempt to understand how this cilia-generated flow is translated into an asymmetry in molecular markers, two hypotheses were put forward.

In a summarized way, according to the mechanosensory hypothesis, cilia react to the flow's mechanical force. When the flow passes through the cilia, they bend and trigger the Nodal-Lefty-Pitx2 pathway [21]. On the other hand, the chemosensory hypothesis proposes that the Nodal-Lefty-Pitx2 pathway is triggered by a morphogen gradient created by the flow.

1.4. Morphogen gradient model

This model assumes that when beating, the motile cilia create a unidirectional flow towards the left side carrying secreted morphogens along with it. This assures that the left side of the organiser has a higher morphogen concentration than the right side and consequently the gene expression on the left side will diverge from that on the right [22].

However, these hypothetic free secreted morphogens capable of directly inducing the Nodal-Lefty-Pitx2 pathway have not yet been identified [9].

Along with this, it has also been proposed that some of these potential morphogens (such as Shh, RA) could instead be encapsulated inside vesicles (nodal vesicular parcels) that once caught in the fluid flow are transported towards the left side where they eventually break down near the LRO wall setting the morphogens free (Figure 1.2A) [23].

However, unlike expected, slowing or accelerating the flow does not alter the situs [16]. Furthermore, mutants lacking motile cilia (*kif3A/B*) on the LRO show bilateral expression or absence of the Nodal-Lefty-Pitx2 pathway [18] while in mutants with motionless cilia (*lrd, iv*) the Nodal-Lefty-Pitx2 pathway can be expressed on the left side, on the right side, bilaterally or be totally absent [24].

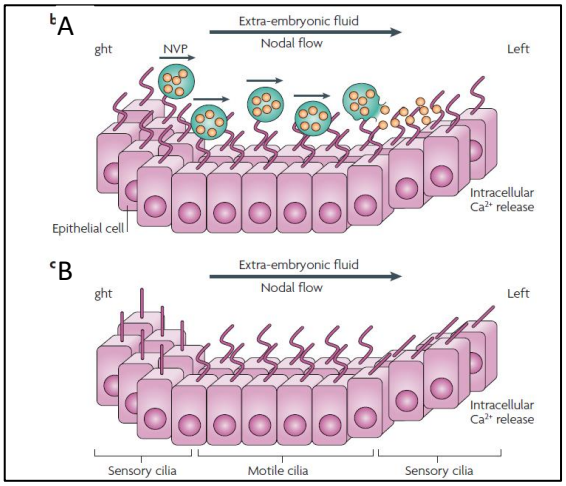


Figure 1.2 – Morphogen gradient model and two cilia model.
A: The morphogen model predicts that secreted morphogens, free or encapsulated in vesicles (NVP), are transported within the flux crating a concentration gradient that is sensed by the cilia in the left side triggering the Ca^{2+} release on the left side. B: The two -cilia model proposes that motile cilia create the flow while immotile cilia sense it. The flow detection occurs when the fluid pressure bends the sensory cilia triggering the L-R asymmetry breaking in a Ca^{2+} dependent way. Adapted from Fliegauf *et al*. 2007

According to the morphogen model, it was expected that both types of mutants presented the same phenotype, as they both have no fluid flow in the LRO [21]. Additionally, some other factors such as PKD2 [25] that are not involved in generating the flow, but seem to be important in sensing it, have been described to have a role in the gene asymmetric expression on the LRO.

1.5. Two-cilia model

In an attempt to justify the phenotype differences between mutants lacking motile cilia on the LRO (*Kif3A/B*) and mutants with cilia but immotile (*iv* and *Lrd*), it was demonstrated that the LRO has two types of cilia: motile cilia, those which generate the flow, and immotile cilia, likely those that may sense it (Mcgrath and Brueckner 2003).

This model assumes that immotile cilia on the left side of the LRO encounter a flow stronger than those on the right. The model advocates that the flow can bend the cilia stimulating the

mechanoreceptors to induce a Ca^{2+}-dependent signalling pathway (Figure *1.2*B) leading to the asymmetric release of signalling molecules [22].

Mutants with motionless cilia are incapable of generating the flow but preserve the sensing capacity, meaning that external stimuli can randomly activate the mechanosensors. On the other hand, mutants lacking motile cilia loose both abilities rendering them unable to generate or sense the flow [21]. This offers one explanation to the different phenotypes observed.

The explanation backing the mechanosensation connects PKD1 and PKD2 to the early L-R signalling. PKD1 and PKD2 were first identified in polycystic kidney disease. These are membrane proteins present in the kidney cell's primary cilia [26] that physically interact to change the conformation of PKD2, a Ca^{2+} channel, and allow it to open [27].

PKD1L1 (PKD1-related) was found to be the important protein, being restricted to the LRO cilia both in mouse [28] and in medaka[25].

Praetorius & Spring showed that bending the primary cilia in kidney culture cells led to an increase of Ca^{2+} into the cytoplasm that spreads to the nearby cells through gap-junctions [29] bringing up the possible role of the PKD1L1-PKD2 complex as a mechanosensor in the LRO cilia. Furthermore, both mutants in PKD2 and PKD1L1 do not succeed in establishing asymmetric gene expression at the LRO and LPM, displaying right lung isomerism regardless of the fact that both LRO morphology and cilia motility are apparently normal [28] [25]. PKD2 mice mutants also show several laterality defects [30] similar to those seen in mutants lacking motile cilia on the LRO [18].

These results connect the PKD1L1-PKD2 complex (located in all cilia) and Ca^{2+} levels to the L-R axis establishment [17][28]. In addition, Ca^{2+} signalling levels are typically higher on the left side of the LRO and become randomized in PKD2 mutants [30].

It has recently been described that PKD1L1 is not necessary to Nodal activation, instead, it acts as PKD2 regulator limiting Nodal expression to the left side [31]. Therefore, PKD1L1 acts as an intermediate between the flow and the Ca^{2+} signalling cascade activated by PKD2 [31].

In spite of Ca^{2+} asymmetries and flow response likely being connected, the exact mechanisms how they induce the signalling cascade that leads to gene expression asymmetries still needs to be clarified [32]. Furthermore, recently there has been some evidence against this hypothesis. The mechanosensation was recently questioned by measuring the Ca^{2+} influx in response to the flow in primary cilia *in vivo* [33]. The results showed that no Ca^{2+}influxes were observed in flow stimulated nodal cilia, suggesting that, if cilia are to act as mechanosensors, they must induce a non-Ca^{2+} related signalling pathway.

Nevertheless, the calcium indicator used by these researchers [33] was also questioned at subsequent commentaries and meetings and other alternatives and better calcium probes are being developed in labs from different countries. We shall have to wait for the new results to conclude if the mechanosensation is still feasible in the LRO.

1.6. Ciliary structure

The majority of cells only have one cilium, while others have several, such as the multiciliated respiratory epithelium [34].

Cilia are cellular hair-like structures projected from the plasma membrane that are built when cells get into a differentiated form and reabsorbed when they re-enter the cell cycle [35]. Cilia can be divided into sub-compartments consisting of basal body, transition zone, axoneme and ciliary tip.

Two similar proteins mainly form cilia: α-tubulin and β-tubulin [36]. These two proteins come together to form dimers that eventually aggregate into a polymer. This polymer rolls itself up to create a tube that serves as rails for the intraflagellar transport (IFT) [37]. The basal end of the microtubule is anchored to the microtubule organising centre (MTOC), which is derived from the centriole. The other end is let free so tubulin dimers can be added or removed accordingly to the cell's state.

In cells that have cilia or flagella, the MTOC[b] is called basal body. Cilia and flagella have similar structures, with the difference that cilia are shorter while flagella tend to be longer.

A cilium is not in continuity with the cytoplasm. All the building blocks necessary to its structure are produced in the cytoplasm and selectively transported through the transition zone (minus end) to the ciliary tip (plus end). Also, when the cilia need to be removed, all the components must be brought back to the cytoplasm.

Aside from building blocks, the motor proteins can also transport signalling molecules. This transport is done by kinesins (or KIFs) (anterograde) and cytoplasmic dyneins (retrograde) that shuttle in and out of the cilia respectively [38]. In humans, there are 46 kinesins [39] and 17 dyneins [40] identified with different motor domains handling different types of transport.

The distal tail domains of the motor proteins bind selectively to cargo while the motor domains interact with the microtubules. These motor domains have the ability to hydrolyze ATP using the generated energy to move forward along the microtubule [41].

In the absence of these motor proteins such as the KIF3 complex (anterograde transporter), the cilia are not correctly assembled. As a result, even though the basal bodies are present, *Kif3B* [18] and *Kif3A* [19] mice mutants have no cilia in the LRO leading to the randomization of the L-R asymmetry breaking.

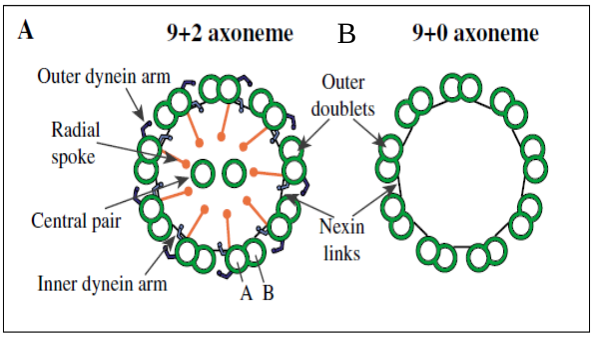


Figure 1.3 – Transversal views of the axoneme organisation.

A: The 9+2 structure is commonly found in motile cilia, showing 9 microtubule doublets, a central pair, radial spokes and dynein arms. B: The 9+0 structure is often associated with immotile cilia that have no dynein arms. From Dawe, et al.2006) [116].

1.7. Ciliary structure: motile vs immotile cilia

Cilia can be categorised as motile or immotile accordingly to their morphological structure. Immotile cilia (also called primary), present in almost every cell type in vertebrates, are usually short and motionless. Motile cilia are longer and have the ability to beat thanks to the presence of dynein arms in their ultrastructure.

Most motile cilia's axonemes exhibit a 9+2 arrangement (Figure 1.3A) with 9 doublets hold together by nexin fibres (N-DRC) and a central pair, as those generating the flow in zebrafish KV [42]. Each doublet is composed of two tubules, A (complete) and B (incomplete) where tubule A has two dynein arms attached: inner dynein arm (IDA) and outer dynein arm (ODA) [34]. Apart from holding the microtubules together, N-DRC also influences the axonemal bending by controlling the IDA's attachment to the A-tubule and mediating the communication between the central pair and the dynein arms through the RS [43].

Cilia with 9+0 arrangement (Figure 1.3 B) lack the central pair and dynein arms and are mostly found in immotile cilia. However, there are 9+0 cilia that have dynein arms and are capable of beating (such as those inducing nodal flow in mouse [44] or medaka [45]).

[b]The centrosome is also a type of MTOC and acts as an anchor to microtubules during mitosis.

IDA and ODA (Figure 1.4) are made up of axonemal dyneins (DNAH, DNAI, DNAL) that hydrolyse ATP and transform it into kinetic energy. The IDA and ODA touch the neighbour B-tubule and cause the two adjacent doublets to glide over each other, generating cilia beating pattern [46].

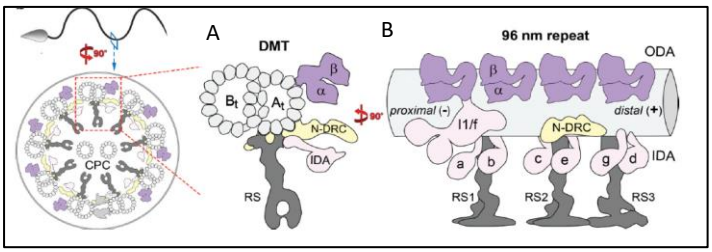


Figure 1.4 – Cilia ultrastructure and dynein placement in the axoneme in a transversal and longitudinal view.
A shows the relative positions of the duplet, IDA (pink), N-DRC (yellow), ODA (purple) and RS (grey).in a transversal view. B shows the relative position of the same components as A in a longitudinal view, evidencing the axonemal unit repeating at every 96 nm.
Adapted from J. Lin et al. 2014 [117].

In humans, ODA and IDA are made of 12 and 15 dyneins respectively (NCBI).When axonemal dyneins such as DNAH1 [48] and DNAH5 [49] are mutated or absent, the ciliary beating function is impaired and can give rise to primary ciliary dyskinesia symptoms.

1.8. Ciliopathies

Cilia can be found in almost every cell type [50]. Primary cilia have a major role in signalling pathways throughout development and in tissue homeostasis. They can act as antennae capturing signalling molecules (Shh, TGF), can react to light or odours in the eye or olfactory epithelium, react to fluid pressure in the kidney and sense pressure in the bone (reviewed by [34]).

Motile cilia also retain sensory skills, but while primary cilia are ubiquitous in almost all cell types (except red blood cells), in adults, motile cilia are localized in sperm, epithelial cells in the bronchi and oviducts, and ependymal cells lining the brain ventricles [51].

Due to this ubiquity, several human diseases and developmental disorders can arise when cilia are defective [52]. These disorders can range from PKD (polycystic kidney disease), pancreatic cancer, or somatic syndromes (characterised by loss of vision and audition, polydactyly, obesity and neurological malformations) when primary cilia are affected. When the problem lies in motile cilia, primary ciliary dyskinesia arises and may involve infertility, upper and lower respiratory impairment, hydrocephaly and *situs inversus*.

1.9. Primary ciliary dyskinesia (PCD) (MIM 244400)

A relationship between cilia and L-R axis establishment had been suspected since the discovery of Kartagener syndrome (KS) [53]. Kartagener described a condition causing *situs inversus*, otitis, sinusitis and bronchiectasis and later Afzelius group added the lack of dynein arms [54] and consequent male sterility [55] to the collection of indicators.

KS is a particular case of PCD. When a patient presents *situs inversus totalis*, they are said to suffer from KS, which happens in 50% of PCD cases [9].

While the perfect inversion of the *situs solitus* arrangement, called *situs inversus totalis*, carries little medical consequences [56], the problem lies in the heterotaxy spectrum cases [57] (Figure 1.5).

In humans, laterality defects can be manifested in distinct ways. All organs can be placed in their normal relative position (*situs solitus*) but in a mirrored image (*situs inversus*), what is estimated to occur in 1 out of 6 000 – 8 000 newborns [56]. Another situation, known as heterotaxy or *situs ambiguous*, occurs when at least one organ is misplaced along the L-R axis.

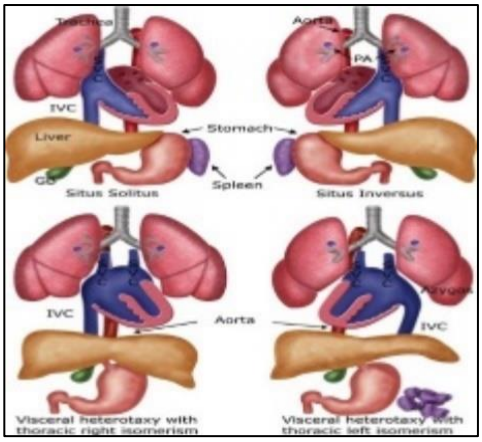


Figure 1.5 – Normal situs and situs abnormalities. GB, gallbladder; IVC: inferior vena cava; PA: pulmonary artery; SVC: superior vena cava
From: https://radiologykey.com/ultrasound-evaluation-of-the-fetal-hear

Heterotaxy syndrome is responsible for several congenital cardiac and gastrointestinal malformations [58] found in PCD patients [59] and it is thought to affect about 1 in 10 000 newborns or 1 in every 5 000 – 7 000 of live births with a chronic heart disease (CHD) [57].

According to Kennedy *et al*., 45% of patients diagnosed with PCD had *situs solitus*, 47.7% *situs inversus*, and 6.3% showed heterotaxy. Among this 6.3 %, the majority had cardiovascular malformations [60]. The same authors also reported that the prevalence of congenital heart diseases among patients diagnosed with heterotaxy is 200-fold higher in PCD-suffering patients than in the general population (1:50 vs.1:10 000) [60].

To complicate matters, heterotaxy often comes associated with isomerism. When isomerism is present, one organ that normally would be asymmetric presents itself symmetrically (Figure 1.5). Left isomerism is linked to polysplenia while right isomerism is associated with asplenia [61] but its prevalence in PCD patients is not known [60]. However, sometimes patients present anatomical uncommon arrangements, making it difficult to classify it as left or right isomerism [62].

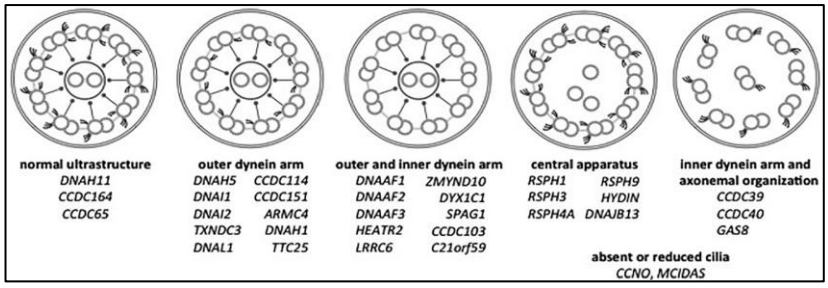


Figure 1.6 – Typical axoneme defects in different mutated genes in PCD.
It shows the list of PCD-related genes that when mutated result in normal axoneme structure, absence of ODA, IDA, or central pair are absent, or axoneme disorganisation. There are still two identified genes that when mutated impair the cilia formation. Adapted from Ferkol 2017 [63].

PCD can arise as the result of several inherited mutations in the cilia-motor machinery such as ODA [49] IDA, radial spokes or central pair components [59].

Mutations in DNAH1, DNAH5 and DNAH11 make up for 25% of all PCD cases [56], but so far more than 35 genes have been linked to PCD (with 70% of patients testing positive for biallelic mutations) [63]. However, a defect in any other one of the about 120 proteins that make up the cilia or in the cytoplasmic proteins [64] responsible for transporting or assembling cilia components can potentially affect the cilia's ultrastructure and cause PCD. Figure 1.6 shows the most common mutated genes and the consequent effect on cilia's axoneme structure.

1.10. Coiled-Coil Domain-Containing Protein 40 (CCDC40)

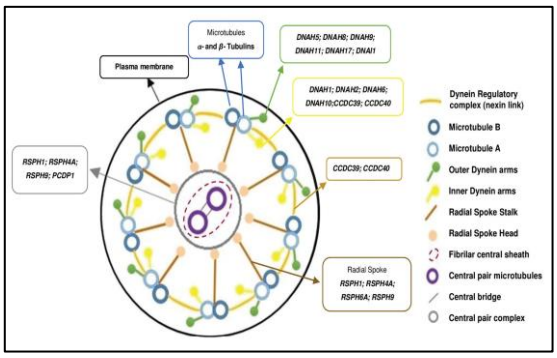


Figure 1.7 – Proteins known to be associated with each axoneme component in humans.
Ccdc40 encodes a protein necessary for the assembly of dynein regulatory complex (DRC) and inner dynein arm (IDA) complexes.
From Pereira *et al*. 2015 [118].

Among the genes responsible for PCD there is Coiled-Coil Domain-Containing Protein 40 (*CCDC40*) (HGNC: 26090) [65–67]. When mutated, *CCDC40* does not completely render the cilia immotile, however, they were found to beat in a fast, flickering way within a reduced amplitude [68].

CCDC40 stands for Coiled-Coil Domain-Containing Protein 40. These mentioned coiled-coil domains are common in proteins [69], they are made of two to five amphipathic α-helices that twist around each other to form a supercoil. Coiled-coil domains are implicated in homodimerization and are present in proteins involved in intracellular transport, molecular recognition, signal transduction and movement regulation [70].

CCDC40 (also known as FAP172, KIAA1640 or CILD15) is necessary for cilia motility, as it is needed for the correct assembly of the N-DRC and IDA [71,72] (Figure 1.7).

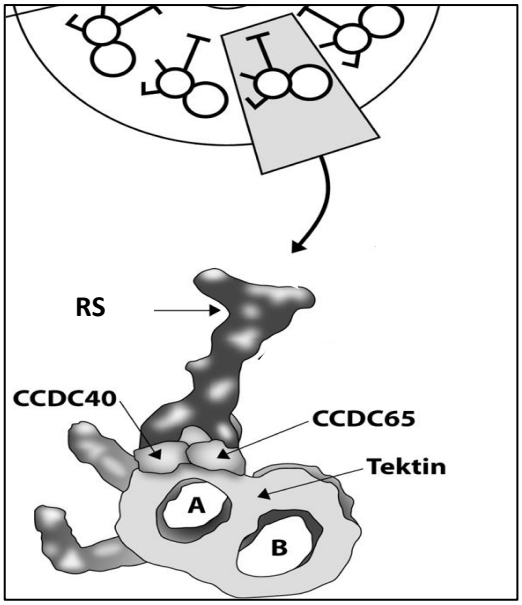


Figure 1.8 – Relative position of the CCDC40 protein within the motile cilia.
CCDC40 locates where the IDA, RS and N-DRC attach to the A-tubule.
Adapted from Werner-Peterson & Sloboda 2013 [73].

It localizes to the area where IDA and RS attach to the A-tubule [73] (Figure 1.8) and is also essential for tubulin polyglutamylation at the proximal extremity of the cilium [74], a post-translational modification that regulates the microtubule's stability [75].

CCDC40 does not act alone, it has been shown that mutations in *CCDC39* or *CCDC40* result in indistinguishable phenotypes [67]. Also, in a 2013 study, *CCDC40*, together with *CCDC3*9, when mutated were found to be the cause of PCD in 69% of the identified patients, especially among those previously diagnosed with "radial spoke defect" (loss of IDAs and axonemal disorganization) [76]. Furthermore, there are clues that the destabilization of CCDC40 (along with CCDC65) leads to the disassembly of the major structural components of the axoneme [73].

CCDC40, together with CCDC39, form a complex that was identified as a molecular ruler in cilia, responsible for forming a 96 nm-length gap between RS [77]. When the complex CCDC39/40 is absent, there is an inconsistent number of RS that bind along the A-tubule, meaning that the binding of the RS to the tubule does not depend on the CCDC39/40 complex. On the contrary, in normal cilia, it is thought that the complex blocks the radial spoke binding regions leaving available only the suitable ones at every 96 nm [77].

Antony *et al*. noted that RS components are detected in cilia from patients carrying *CCDC40/39* mutation however, there is no indication that they are correctly assembled or localized [76]. Therefore, the so-called "radial spoke defect" may not be due to the loss of RS, but instead due to its mislocalization or inability to attach to the microtubules.

Additionally, CCDC39/40 also regulates IDA and the N-DRC attachment to the A-tubule [77] making it a crucial element in the assembly of axonemal structures and spatial integrity.

Several studies seem to attest this notion by linking *CCDC40* mutations with IDA absence and axonemal disorganisation [65,67,76,78].

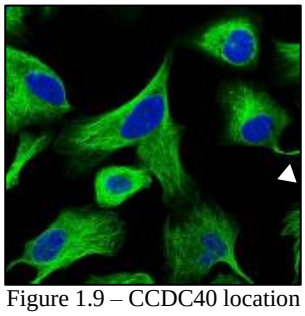


Figure 1.9 – CCDC40 location inside the cell.
CCDC40 is found in throughout the cytoplasm, were it co-localizes with microtubules, and in cilia. White triangle shows a cilium. Image available at: www.proteinatlas.org/ENSG00000141519-CCDC40/cell

CCDC40 has been identified in several organisms[c]. According to AmiGO2 database, it is located on cilia (respiratory epithelia, sperm flagellum, LRO) and also in the cytoplasm as shown in Figure 1.9.

In humans, it is located in the chromosome 17 (ENSG00000141519) and has 16 transcripts identified with at least 7 producing a protein.

It has two annotated domains: BRE1 (E3 ubiquitin ligase) [79] and SMC N terminal domain. The BRE1domain role is, in this case, not yet known [76], the SMC domain is larger and is believed to be involved in microtube transporting [72]. In zebrafish, ccdc40 is located on chromosome 6 (ENSDARG00000100584), has only two transcripts annotated with only one resulting in a protein.

In PCD patients, some of this gene's (CCDC40) locations seem to be more prone to mutations ("hot spots"). The most frequent mutations are c.248delC and c.3129delC, that modify the reading frame [71] and c.2440 C>T, c.961 C>T and c.1345C>T that produce truncated proteins [76].

1.11. CRISPR-Cas9 system as a tool for gene editing

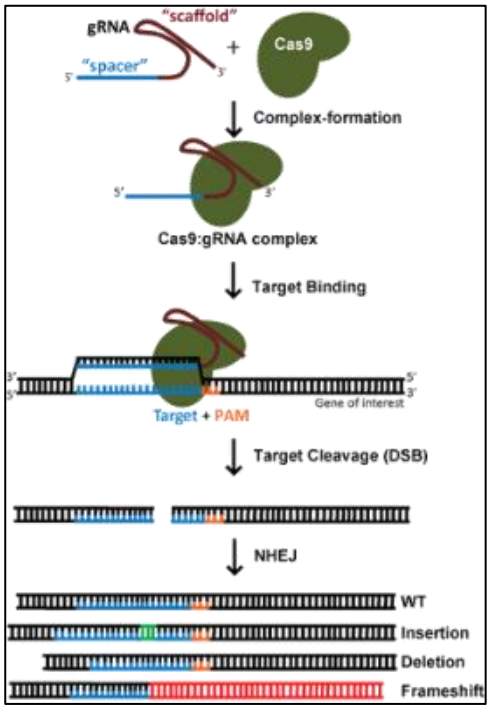


Figure 1.10 – Knockout generating process using CRISPR-Cas9 system.
After the Cas9 makes a DSB and the cell repairs the cut using the NHEJ pathway. This pathway is susceptible to insert or remove bases what can result in protein truncation or alterations in the reading frame. From https://www.addgene.org/crispr/guide/

The CRISPR/Cas9 (clustered regularly-interspaced palindromic repeats associated with Cas9) is based on a defence system evolved by bacteria to protect themselves against viral infections [80,81].

When a virus infects a bacterium, it stores a fragment of the viral DNA (spacer) in between PAM (Protospacer adjacent motif) sequences [80].

When the virus strikes again, the bacterium produces two types of RNA: crRNA (CRISPR RNA) and tracrRNA (trans-activating crRNA) [82]. The crRNA holds a sequence that is complementary to the spacer, while the tracrRNA facilitates the connection between crRNA and Cas9.

crRNA and tracrRNA form a complex with the Cas9 which acts as a helicase and nuclease. When the complex finds the matching sequence to the crRNA, the Cas9 makes a double-stranded break (DSB) in the DNA neutralising the virus.

There are at least three types of CRISPR/Cas systems [83] but type II is most used in genome editing [84]. This system was engineered to allow scientists to cut any DNA strand at a particularly chosen location by changing the

[c] http://amigo.geneontology.org/amigo/search/annotation?q=ccdc40

sequence in the guide RNA (merging of crRNA and tracrRNA) [84] as long as it is followed by a PAM sequence (Figure 1.10).

This PAM sequence consists of a 5'-NGG-3' sequence (where "N" can be any base) and is fairly common in the genome [85] therefore the CRISPR-Cas9 system can virtually cut any site along the genome.

This guide RNA (sgRNA) can be injected into a cell or embryo together with *cas9* mRNA or the Cas9 protein itself. Once inside the nucleus, the complex sgRNA-Cas9 anchors to the PAM, the Cas9 unzips the DNA and matches it to the target sequence in the sgRNA. If the match is successful, the Cas9 makes a DSB in the DNA. The cell reacts by trying to repair the cut using the non-homologous end joining (NHEJ)[d] pathway, however, this process is error-prone, leading to base insertions, deletions or frameshift mutations that can silence the gene [86].

The insertions or deletions produced are random, nevertheless, it is also possible to replace a mutated gene with a healthy copy. This can be achieved by injecting a DNA sequence carrying the correct sequence together with the sgRNA and Cas9. In this situation, after the Cas9 cuts the DNA, the cell can use the given DNA sequence as a template, when repairing the cut, following the homology-directed repair (HDR)[e] pathway [86].

CRISPR-Cas9 system can be used as knocking down tool allowing to study a gene's role in a cell or organism. Unlike previous methods, CRISPR-Cas9 can be used to target multiple genes at once [87], what makes it useful for studying processes or diseases regulated by several genes working together.

1.12. Zebrafish as animal model for PCD

An organism or cell culture that displays all or some of the physiological signs that can be observed in the actual disease can be considered a good disease model. It is important that it allows to understand the disease's physiological consequences, develop new diagnostic approaches or test new lines of treatment.

Zebrafish presents itself as o good model to study PCD. It has external embryonic development and the embryo is transparent enough to allow KV inspection and manipulation without killing the fish.

The presence of the KV is another reason favouring the use of zebrafish. It is vesicle full of liquid surrounded by ciliated cells. The cilia on these cells are one per cell, being 80% motile and 20% immotile but they all present a 9+2 structure [88].

Additionally, among all the genes identified in the human genome, around 70% have at least a homologous gene within the zebrafish genome [89]. In what concerns the CCDC40 gene, when aligning the human sequence with that of the zebrafish by using EMBOSS Needle alignment, they show a similarity of 43.9% (see 7.1). However, if we only consider the genomic sequence against which the sgRNAs were created, the similarity between zebrafish and human is 59.1%.

Importantly, this target region on zebrafish relates to the one that spans across the site where a common nonsense mutation was identified (961, C>T) in patients suffering from PCD [76].

[d] The NHEJ is a common mechanism used by cells to repair DSB. It is called this way because it does not relay on a template. Instead, the repairing enzymatic complex revises the ends by removing or adding bases and re-join the resulting extremities without checking for possible introduced errors [114].

[e] The HDR repair mechanism uses a similar undamaged sequence of DNA as a template. The repairing enzymatic complex interlace the damage and undamaged strands, get them to exchange sequences of nucleotides and fill in the missing gaps in an error-free way [115].
HDR is less common than NHEJ, as it can only occur in the presence of a template sequence. This template can only be accessed in diploid cells in S phase or when a donor sequence is artificially introduced into the cell.

1.13. Project goal

The main aim of this project was to generate a *ccdc40* zebrafish mutant using the CRISPR-Cas9 system to be studied as a disease model.

The reason behind this choice lays on the fact that mutations in this gene have been identified in patients diagnosed in our laboratory as well as in some published articles that associate it with PCD [67,90,91].

There is already a *ccdc40* zebrafish mutant line (*lok*) expressing a truncated protein. This mutant expresses an early terminated protein (non-sense mutation at 778 aa) containing almost the totality of the coiled-coil domain.

The mutant we generated expresses a truncated protein up to the 116 a.a. excluding completely the coiled-coil domain.

The CRISPR-Cas9 system appears as an alternative to the Morpholino (MO) technology, a knockdown method, that has been questioned lately [92,93]. The MO's have long been a popular method to study gene function, particularly in zebrafish. However, they can induce toxicity and have unpredictable off-targets. Kok *et al*. compared MO-induced and mutant phenotypes and concluded that, in 70% of the analysed genes, the results were not equivalent [93]. Therefore, when possible, MO-induced phenotypes should be verified by gene-editing methods such as CRISPR-Cas9.

To do so, the sgRNAs were designed to guide the CRISPR-Cas9 complex to the region of the zebrafish genome codifying for the orthologous human region containing one of these "hot spots" (Figure 1.11). After the designing, they were cloned into a pDR274 vector containing a gRNA scaffold that takes the role of tracrRNA and interacts with Cas9. The vector was replicated in competent bacteria and used as a template to produce RNA *in vitro*. This sgRNA produced included the gRNA (scaffold) and the complementary sequence to the target region.

In an attempt to cause a loss of function mutation, we injected sgRNA together with Cas9 mRNA into zebrafish embryos and tested some of those embryos to check the CRISPR-Cas9 system efficacy, while the majority was let to grow.

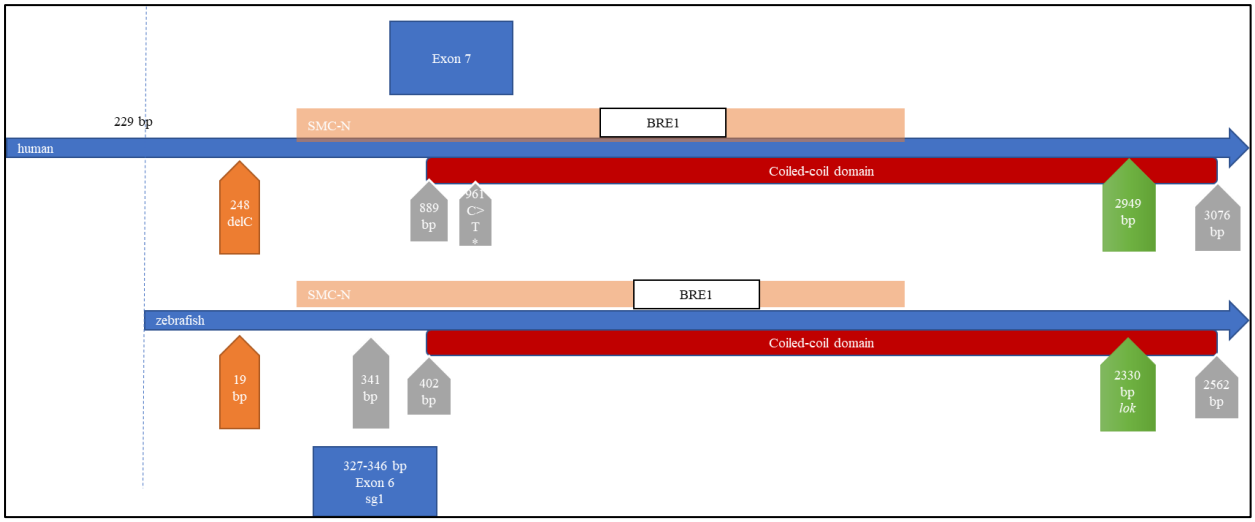


Figure 1.11 – Human and zebrafish ccdc40 cDNA sequence alignment.
Non-scaled diagram showing human and zebrafish ccdc40 cDNA sequence alignment EMBOSS Needle alignment algorithm. The CRISPR-Cas9 system was prepared to target the region around the position 341 bp in zebrafish, leading to a truncated protein without the coiled-coil domain. The *lok* mutation causes the production of a shortened Ccdc40 protein closer to the C-terminal.

We could identify crispant fish (founders) and now have their progeny (F1) growing until they are ready to be genotyped. However, as zebrafish is an organism that takes about three months to reach maturity, obtaining the homozygous mutant (*ccdc40-/-*) was not viable in this work's time span.

So, the second aim was to compare the CRISPR-Cas9 to the MO-induced phenotype. I, therefore, knocked down ccdc40 by using a transcription blocking MO. The MO was tested by injecting several concentrations (as explained in 2.12 in page 18). After the MO injection, parameters such as mortality, heart and gut situs, cardiac oedema and tail defects were evaluated to future comparison with *ccdc40-/-* mutant zebrafish.

2. Methods

2.1. Zebrafish maintenance

In this work were used wild-type *AB* lines (ZFIN ID: ZDB-GENO-960809-7), *Tg(sox17:GFP)* (ZFIN ID: ZDB-TGCONSTRCT-070117-57) *Tg(foxj1a:GFP)* (ZFIN ID: ZDB-TGCONSTRCT-131120-3) and *TgBac(cftr-GFP)* (ZFIN ID: ZFIN ID: ZDB-TGCONSTRCT-130423-1). These fish were kept in the CEDOC fish facility under a controlled environment approved by the Direção Geral de Alimentação e Veterinária (DGAV).

Adult fish were kept in fresh water (28º C, 700 µS, pH 7) under a 14:10 h light:dark cycle. Embryos were obtained by natural mating and incubated at 25ºC or 28ºC up to five days post-fertilization (dpf) after which they were transferred to a tank in the main system or sacrificed.

The natural mating was stimulated by placing a male and one or two females inside a mating box with a transparent divider allowing visual contact. The fish were removed from the main system and placed in the mentioned mating boxes in the evening and left separated, as described, during the night. In the morning, the dividers were removed and the water level lowered allowing the courtship behaviour between male and female and consequently the external fertilization of the eggs. This method allows for the synchronization of the embryonic development.

In each mating box, there was a sieve preventing the fish from eating the eggs. After the eggs were laid the adult fish were again placed in the same main system tank. The eggs were collected by filtering the water remaining in the mating box and then were placed in a Petri dish immersed in embryonic medium with methylene blue (5mM NaCl, 0.2mM KCl, 0.3mM $CaCl_2$, 0.3mM $MgSO_4$ and distilled water H_2O, pH 6.5). The embryonic medium renewal and the removal of dead embryos were done daily to prevent bacterial and fungal infections.

The embryonic staging was done according to what is described in Kimmel *et al.* [94]

2.2. Zebrafish euthanasia

Larvae up to 5dpf were sacrificed by being placed in a bleach solution. Larvae older than 5dpf and adult fish were euthanized using an overdose of tricaine (MS-222) solution in a 250 mg/L concentration [95,96].

2.3. CRISPR-Cas9 target sequences - Design of sgRNA

To study the functionality of ccdc40 gene (ENSDARG00000100584) in zebrafish, the CRISPR-Cas9 technique was used to induce DNA breaks. The CRISPR-Cas9 target sequences inside the ccdc40 gene were determined using the crisprscan tool available online (crisprscan.org). This tool scans the gene-of-interest coding sequence (ENSDART00000169752) and gives back a list of potential targets sites ranked accordingly to its predicted efficacy. All the 20 nucleotide-long sequences presented are followed by an NGG motif also called PAM (protospacer adjacent motif) that is required for the Cas9

recognition. The designed sgRNA oligonucleotides are expected to induce the cutting at any sequence following the formula 5′-GG- N18 -NGG-3'. The GG-N18 (underlined) is included in the transcribed sgRNA while the NGG integrates the genomic DNA only (not present in the sgRNA). If the sequence has a mismatch in one of the first two GG's it is called non-canonical but the Cas9 can still work properly [97].

The provided list shows information concerning the score assessed by the tool, the sequence's position in the chromosome, its canonical state and the number of off-targets.

From the referred list, we chose two oligonucleotides that along with having the highest score (higher scores mean higher predicted efficacy), were closer to the initial ATG (to create a mutation as early as possible), these had no off-targets (those with no off targets minimize induced cuts in other genome sites) and were considered canonical (no mismatches in the first GG providing a stronger binding to the target site). The capital letters shown indicate the sequence that must be present in the sgRNA and transcribed in vitro.

To allow the sgRNA cloning into a vector, for each target site were designed two complementary oligonucleotides exhibiting single-stranded overhanging ends (bold) (ordered from Stab Vida; Oeiras, Portugal) (Table 2.1).

Table 2.1 – sgRNA sequences

ccdc40	sgRNA #1 oligonucleotides - 5'>3'	sgRNA #2 oligonucleotides - 5'>3'
Forward (Fw)	**TAGG**TAGTAGTTGAGAAGGCAG	**TAGG**CCGAGGACCACCTGTATG
Reverse (Rv)	**AAAC**CTGCCTTCTCAACTACTA	**AAAC**CATACAGGTGGTCCTCGG

2.4. sgRNA annealing and cloning into pDR274 using BsaI restriction site

Both pairs of (forward and reversed) designed oligonucleotides were dissolved in miliQ H_2O (100 µM) and annealed. To perform the annealing reaction, 1.5 µl of the forward oligonucleotide (100 µM) and 1.5 µl of the reverse oligonucleotide (100 µM) were added to 23 µl of miliQ H_2O and 24 µl of annealing buffer (10 mM Tris, pH7.5-8.0, 50mM NaCl, 1m EDTA). This solution was then heated to 95ºC for 5 min and let cool down to RT.

Simultaneously, pDR274 plasmid vector (Addgene #42250) (see 7.2) was incubated with BsaI (NEB #R0535) and brought to a linear conformation. The linearization was accomplished by mixing 5 µg of the circular plasmid, 2 µl of the enzyme, 5 µl of NEBuffer™ 3.1 (NEB # B7203S) and 38 µl of miliQ water and incubating the resulting solution O/N at 37ºC. To check the linearization success, 1µl of the reaction product was loaded into a 0.8% TBE-agarose gel and compared with the "non-insert" control ligation reaction. The remaining volume was purified using a DNA clean & concentrator kit (Zymo research #D4003) and quantified by NanoDrop™ 2000.

The pDR274 plasmid contains three useful features: a T7 promoter found upstream of a partial gRNA sequence and a domain conferring kanamycin resistance. The partial gRNA sequence allows the sgRNA sequence and Cas9 interaction. The annealed oligonucleotides were inserted between the T7 promoter and the gRNA scaffold where the BsaI restriction site is found. To ensure this, the oligonucleotides (10 µM) and the linearized plasmid (100 ng) were incubated O/N at 16ºC with 0.5 µl T4 ligase (Sigma-Aldrich #9015-85-4), 5 uL T4 buffer (2x) (Sigma-Aldrich #KEM0046B) and 2.17 µl miliQ water.

In the forward strand, BsaI cuts immediately after the recognition sequence 5'-GGTCTCN-3' whereas in the reverse strand it cuts only four nucleotides ahead (3'-CCAGAGN$(N)_4$-5') creating non-compatible sticky ends that match the overhanging ends added to the designed oligonucleotides.

This directional restriction cut by BsaI not only forces the correct insertion of the annealed oligonucleotides into the plasmid backbone but it also prevents the plasmid from self-ligating.

2.5. *E. coli* (DHα5) transformation and identification of positive colonies

The ligation product was used to transform competent *E. coli* bacteria. In this case, DHα5 was the elected strain, rendering a transformation efficiency >$1x10^6$ colony forming units for every mg (Thermo Fisher #18265017).

The bacteria, stored at -80ºC, were thawed on ice for 20 min. Then, a 50 µl aliquot of DHα5 together with 10 µl of the ligation product were subjected to a thermal shock (30 min on ice, 1 min at 42ºC, 2 min on ice) to boost the plasmid uptake. Next, followed an incubation at 37ºC and 250 rpm for 1h 30 min maximizing air-liquid interface. The incubation step was intended to let the bacteria recover from the thermal shock and assure the plasmid stability inside the cells. Subsequently, the culture was centrifuged (5000x g, 10min, RT), to increase cell density. These transformed bacteria were streaked in a LB-agar plate treated with kanamycin (30µg/µl) and incubated once again at 37ºC O/N with the agar layer facing upwards.

Because the cloning plasmid has a domain conferring resistance to kanamycin, only the bacteria that up-took it were able to subsist. However, some bacteria could have assimilated the plasmid without the oligonucleotides inserted. This oligonucleotide chain is too small to be resolved on an agarose gel, so to avoid this predicament, colonies were screened by PCR (colony PCR). This technique relied on the use of M13R primer (5'- TGTAAAACGACGGCCAGT-3'), that anneals with the plasmid backbone, and the forward designed oligonucleotide for the corresponding Cas9 decided-target. When amplification occurs, it meant the colony-inducing bacterium integrated the plasmid with the insert. However, when no amplification was detected it could be because the cloning process was unsuccessful or the PCR failed.

To detect the true positive colonies, we picked around six colonies for each sgRNA-dedicated agar plate were directly used as template in a colony PCR reaction and streaked in a new agarose plate. Each of these colony PCR reactions was prepared by combining 19.55µl of miliQ water, 2.5µl of buffer (10x, provided with the enzyme), 1.25µl of $MgCl_2$ (50mM, provided with the enzyme), 0.5µl of dNTPs (10mM),0.2µl of NZYTaq DNA polymerase (Nzytech MB00101) and 0.5µl of each primer: M13R (10 µM) and forward sgRNA oligonucleotide (10 µM), adding up to a total volume of 25µl. The colony PCR followed the steps from a common PCR i.e. initial denaturation (95ºC, 10min), denaturation (95ºC, 1min), annealing (51ºC for sgRNA#1 and #2, 1min), elongation (72ºC, 1min) and final elongation (72ºC, 10min) for 30 cycles. The M13R primer is tolerant to a wide range of temperatures, therefore the annealing conditions were adjusted to the designed oligonucleotides structure.

The positive colonies were inoculated in primary liquid cultures (LB medium, kanamycin 30µg/µl) and grown O/N at 37ºC and 250 rpm. These cultures sustain high bacteria densities, essential to a good plasmid yield.

The plasmids were isolated using a commercial kit from Zymo Research (ZR Plasmid Miniprep™ #D4015) and then Sanger sequenced by Stab Vida using M13R primer.

All the described bacteria manipulations, including agar plate preparation, were done in a sterile environment.

2.6. *In vitro* transcription of sgRNAs using T7 promoter

The plasmids (for sgRNA #1 and sgRNA #2) isolated from the positive colonies and confirmed by Sanger sequencing were linearized downstream of the cloning site using a "unique cutter" enzyme. The linearization was achieved by incubating together 1.25 µl of HindIII (NEB # R0104), 5 µl of NEBuffer ™ 2.1 (NEB # B7202S), 5µg of plasmidic DNA and miliQ water (up to 50 µl) for 10h at 37ºC followed by storage at 4ºC. The then linear DNA was purified using a Zymo Research commercial kit (Zymo Research #D4003) and since it contained a T7 promoter, it was used as a template to produce RNA.

Besides the purified DNA template with a promoter, the RNA *in vitro* transcription reaction required triphosphate ribonucleotides (rNTP) (mix containing 10 mM of ATP, UTP, CTP and GTP), transcription buffer 5x ($MgCl_2$, 1 M; NaCl, 5 M; Tris-HCl, 1 M pH8), dithiothreitol (DTT) (50 mM), T7 RNA polymerase (NEB #M0251S) and miliQ water adding up to a total volume of 50 µl.

DTT is a redox reagent that stabilizes the sulfhydryl groups within the T7 protein and without it, the enzyme efficiency drops considerably.

First, the water, 10µl of transcription buffer, 5 µl of DTT and 5 µl of the rNTP mix were added together and incubated for 5 min. Secondly, the DNA template was added (1.5 µg) and after a 1-minute incubation, 1 µl of RNA inhibitor was joint to the solution, followed by another 1-minute incubation. Ultimately, T7 RNA polymerase was joined. First, 2 µl preceded a 2h-incubation and then another 1 µl was followed by a 1h-incubation.

Lastly, 1 µl of DNase (provided with RNA clean & concentrator kit by Zymo Research #R1013) was incubated for 30 min with the solution to remove the template DNA initially added and make sure just RNA remained. All the incubation periods were performed at 37ºC.

The generated RNA was purified with an RNA clean & concentrator kit by Zymo Research (#R1013) and stored at -20ºC. The produced RNA was not 5′-capped or 3′-polyadenylated as it was not to be translated. This protocol makes the RNA more prone to degradation, however as it is to be promptly used, the quality decay is minimal.

This resulting RNA was injected together with Cas9 mRNA into one-cell stage zebrafish embryos.

2.7. Cas9 mRNA production

For Cas9 mRNA synthesis, the plasmid pCS2-ncas9n (Addgene, #47929) was linearized using NotI, a "unique cutter) enzyme.

The *in vitro* transcription reaction done here was very similar to the one done to synthesize the sgRNAs with the exception that here the reaction also included 5ul of G(5')ppp(5')G RNA Cap Structure Analog (20mM)(NEB #S1407S). Therefore, as expected from a mRNA, the Cas9 mRNA is 5′-capped. This feature is crucial as the Cas9 protein needs to be produced by the cell after co-injection with sgRNA for the CRISPR-Cas9 system to work.

Raquel Jacinto, a lab's PhD student performed the synthesis of this mRNA.

2.8. Micro co-injection of Cas9 mRNA and sgRNA into one-cell stage zebrafish embryos

After the collection, the fertilized eggs were lined up against a glass slide inside a petri dish, without embryonic medium and injected using a thin glass needle attached to a Narishigi pico-pump injector. The needle was calibrated using a graticule (10mm/0.1mm Graticule Ltd., Tonbridge, Kent) in a way that each pulse delivered 1.4 nL into the yolk.

The solution filling the needle contained Cas9 mRNA and sgRNA. Two different proportions of these components were injected: 1) 100ng/µl Cas9 mRNA and 50ng/µl sgRNA, 2) 75ng/µl Cas9 mRNA and 37.5ng/µl sgRNA.

Around 200 one-cell stage embryos were injected with each corresponding Cas9 mRNA/sgRNA proportion and allowed to develop at 28ºC. At 24 hpf, each petri dish was sampled by picking 10 embryos used to extract genomic DNA. This DNA was amplified and a 15% polyacrylamide-agarose gel electrophoresis (PAGE) was performed. Each gel contained 6.17 ml of Milli-Q water, 1.2 ml of TBE buffer (10x) (Invitrogen™ 15581044), 4.5 ml of acrylamide/bis-acrylamide (29:1 Nzytech MB04501), 120µl of ammonium persulfate (10%), 9.6µl of TEMED and was let to polymerize between two glass plates with 1.5mM spacer. After 3 h of electrophoresis at 150 V, the polyacrylamide gel was stained with GreenSafe Premium (Nzytech MB13201) (2ul) diluted in TBE buffer (1x) (50ml) for 10 min, washed twice in Milli-Q water and then visualised using ChemiDoc™ XRS+ System (Bio-Rad).

The aim was to detect heteroduplexes that migrate at a different rate due to the existence of mismatches between the WT and the mutated sequence.

Whenever the experiment identified heteroduplexes, the fish were put in a separate tank and transferred to the main system to grow, if not they were euthanized before 5 dpf.

2.9. Primers' design and optimization

The CRISPR-Cas9 system targeted two specific sites in the ccdc40 gene leading to base insertions or deletions around the cutting site. The amplification of the region around each target site by PCR made possible scanning the fish for DNA editions. The primers used in this PCR are listed in Table 2 and were designed using the NCBI primer blast tool (ncbi.nlm.nih.gov/tools/primer-blast/) by restricting the PCR product size between 100 and 200 bp, the primer size between 21 and 25 bases and the T_m from 63 to 67ºC.

The pair 1 flanked a 143 bp fragment including the sgRNA#1 target site. The pair 2 amplified a 183 bp fragment around the sgRNA#2 target site.

To assess the best PCR product different combinations of reagents were used in an optimization protocol, each primer pair underwent an optimization essay by being tested for different final concentrations of $MgCl_2$ (2 mM, 3mM, 4 mM), dNTP's (0.25 mM, 0.5 mM) and primers (0.2 µM, 0.5 µM). Pair 1 worked best with 3 mM of $MgCl_2$, 0.25 mM of dNTP's, 0.2 µM of primers and 0.2 µl of NZYTaq DNA polymerase (MB00101) under the follow conditions: 95ºC for 10 min (initial denaturation); 95ºC for 20 s (denaturation), 60ºC for 30 s (annealing), 72ºC for 1 min (extension) for 35 cycles; 72ºC for 10 min (final extension) followed by storing at 12ºC. Pair 2 worked best with 2.5 mM of $MgCl_2$, 0.5 mM of dNTP's, 0.2 µM of primers and 0.2 µl of NZYTaq DNA polymerase (MB00101) under the same conditions as pair 1 except for the annealing step: 61ºC for 30 s.

Table 2 - Primers flanking the CRISPR - Cas9 target sites.

	Forward	Reverse
Pair 1: ccdc40 sgRNA#1	5'-GCCCTGCTGCAGGTAGTAGTTGAG-3'	5'-GCTGCAGCTTCGCTATTGGTCTCT-3'
Pair 2: ccdc40 sgRNA#2	5'-GGCGCCATGACCTTGCTG-3'	5'-CGTTGTACATTTGTTTGACTAGCAG-3'

2.10. Genomic DNA extraction

Genomic DNA was extracted from 24 hpf embryos. In each case, we collected 3 batches of 10 embryos that were dechorionated, placed in a 1.5 ml Eppendorf tube with 20µl of NaOH (5mM) and incubated for 20 min at 96ºC. The embryos were then smashed with a pipette tip and kept at 4ºC for another 20 min. Lastly, the remaining extract was centrifuged at 12 500g for 10 min with 2 ul of Tris-HCl (1 M, pH8). The supernatant was transferred to a new Eppendorf tube and frozen at -20ºC or used straight away as a template.

2.11. Heteroduplexes detection using PAGE analysis

When two single-stranded DNA or RNA molecules deriving from different sources anneal together they form a heteroduplex. In this case, the different sources are the WT embryos and the crispant embryos where the CRISPR-Cas9 system made a double-stranded cut. The repair (NHEJ) of this cut is prone to errors and therefore may insert or remove several base pairs in the vicinity of the cut. When amplifying this region by PCR using DNA from a pool of embryos as a template, WT (non-altered) fragments can hybridize with repaired fragments (with base insertions or deletions), forming heteroduplexes. Due to the alterations carried out by the CRISPR-Cas9 system, the annealing with WT fragments is not perfect. The existence of mismatches alters the DNA double-stranded conformation and as a result, it changes its migration rate.

On a PAGE experiment (polyacrylamide gel electrophoresis), heteroduplexes generally migrate slower than homoduplexes [98] making its visualization a good method to screen for crispants (potentially modified embryos).

The search for heteroduplexes was done at two different time-points. First, right after the co-injection of Cas9 mRNA and sgRNA and then again when the injected fish reached maturity.

The analysis at the first time-point included randomly picking ten injected embryos from each petri dish to guarantee that the Cas9 was cutting at the targeted genome region. The analysis at the second time-point involved screening each individual adult fish to find crispants. At this point, the potential crispant fish were mated with wild-type fish and their progeny was analysed to access the mutation presence in the germline. In this case, the analysis included three batches of ten embryos from each couple (potential crispant + WT).

In both cases, the selected embryos were dechorionated and processed to extract genomic DNA. This DNA was then amplified by PCR using the primers flanking the corresponding Cas9-targeting region.

PCR followed the conditions specified in “Primers design and optimization”.

The annealed PCR products were then loaded and resolved by electrophoresis on a non-denaturing PAGE.

In the lanes with positive heteroduplexes, the bands were isolated and the DNA extracted by incubating the acrylamide with 20ul of TE (supplied with DNA clean & concentrator kit, Zymo research #D4003) at 37ºC O/N.

The isolated DNA was further amplified using Xpert HighFidelity DNA Polymerase (GRiSP #GE07.0250) accordingly to the manufacturer’s instructions and a T_a= 60.5ºC. After the PCR product purification (DNA clean & concentrator kit, Zymo Research #D4003). Then, the DNA was sequenced (Stab Vida) for confirmation. Overlapping sequences in the chromatograms downstream to the PAM region were interpreted as a sign of heterozygosity and are expected in the heteroduplexes’ samples.

2.12. Microinjection of zebrafish embryos with Morpholino

A morpholino is a synthetic molecule with a structure similar to the natural nucleic acids. It acts as an antisense agent by binding to the RNA and preventing its translation.

The translation-blocking morpholino (ccdc40 MO) was purchased from Gene Tools LLC (Philomath, OR, USA) and received in a lyophilized form. Upon arriving, it was suspended in Milli-Q water, achieving a final concentration of 1mM, heated up to 65ºC for 10 min and then aliquoted.

This ccdc40 MO was designed against the 5'UTR region of the ccdc40 gene, excluding the start codon (to allow future rescue experiences): 5'-TTTTAATTCACAGTCCTTTAGCGGA- 3'.

The ccdc40 MO was injected into the yolk in each one cell-stage embryo to maximize the morpholino uptake by the cells. To do so, right after the eggs were laid, they were lined up against a glass slide and injected with a previously calibrated borosilicate glass needle loaded with a ccdc40 MO solution. This needle, attached to the air injector, which allows the standardization of the volume injected in each pulse 1.4nl. The solutions loaded into the needle held several MO concentrations: 4 ng, 4.5 ng, 5ng, 5.5 ng, 6 ng, 6.5 ng, 7 ng, and 7.5 ng.

In order to characterize the MO effects, were taken into consideration, heart laterality, liver and pancreas relative position and defects such as oedema and tail crooks.

Heart laterality was screened at 30hpf, liver and pancreas relative position at 50hpf and oedema and tail crooks at 5dpf in fish incubated at 28ºC.

2.13. Heart and gut screening

The heart and gut situs were screened to assess the impact of the morpholino injection in the zebrafish left-right development.

In wild-type embryos, the heart can be seen ventrally on the right side of the body under the right-sided eye. The heart tube starts to gain its form around 21hpf and at 36hpf it becomes possible to distinguish the atrium from the ventricle. The S-shape looping occurs between 26 and 48hpf [99], therefore the heart situs screening was done at 30 hph. In wildtype embryos, this looping characteristically occurs towards the left side but it can also be formed towards the right side or the loop can be absent (central heart). The screening of non-injected and MO-injected embryos was done in vivo under a stereoscope using a 60x magnification. The three different phenotypes were placed in separate Petri dishes and later fixed at 50 hpf for further gut *situs* analysis.

The pancreas is formed as a result of the fusion of the dorsal and ventral bud that happens at 44hpf [100]. The liver morphogenesis starts at 28 hpf with the thickening of the endoderm rod under the first somite and is completed around 50 hpf when the hepatic duct is formed [101]. To assess the positions of these two organs we used the 50 hpf *tg(sox17:GFP)* larvae and fixed them in 4% PFA-PBS O/N. These larvae and then evaluated under the Zeiss Lumar V12 stereoscope's fluorescence after the yolk removal by careful dissection.

3. Results

3.1. CRISPR-Cas9 preparation

CRISPR-Cas9 is a useful method to create mutants. For it to work, first, it is necessary to produce *cas9* mRNA (or the protein itself) and the sgRNA holding a complementary sequence to the genomic region we want to disrupt. The cas9 mRNA was already available in the laboratory, however, the specific sgRNAs had to be produced. Besides that, in order to screen the region of interest for CRISPR-Cas9 induced 'indels', we had to design and optimise a pair of primers for each target region.

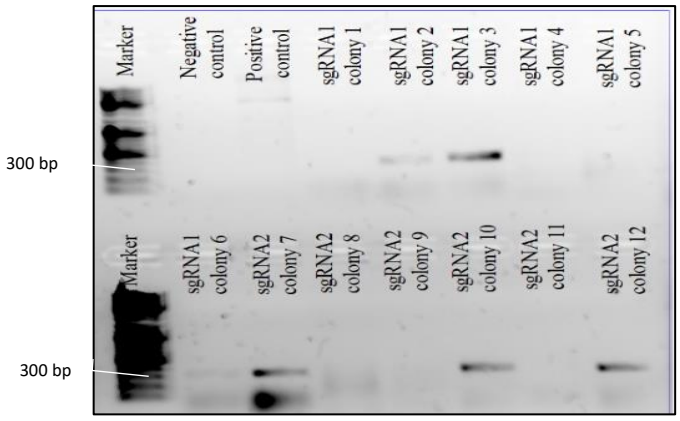


Figure 3.1 – Colony PCR results.
Colonies 2, 3, 7, 10 and 12 were positive showing amplification around 300 bp.
1% TBE-agarose gel, 100 V, 30 min

3.2. SgRNA cloning into pDR274 and sgRNA synthesis

The sgRNAs sequences were inserted into a pDR274 plasmid using the restriction enzyme BsaI. This plasmid containing the sgRNAs template was replicated in *E. coli* (DHα5) incubated in a solid medium with kanamycin.

To verify if the bacteria had, in fact, the plasmid with the insert we performed a colony PCR (cPCR) and run the products on a 1% TBE-agarose gel. Colonies 2, 3, 7, 10 and 12 showed amplification (~300 bp) (Figure 1.1) and therefore, were sent to sequence.

The Sanger sequencing results showed that only colony 10 (sgRNA#2) had the plasmid with the insert (Figure 3.3 A). To save time and reagents, colonies 13, 14 and 15 (for sgRNA#1) were promptly sequenced without undergoing a cPCR beforehand. The results showed that bacteria in colony 14 had integrated the plasmid with the sgRNA#1 template (Figure 3.3 B).

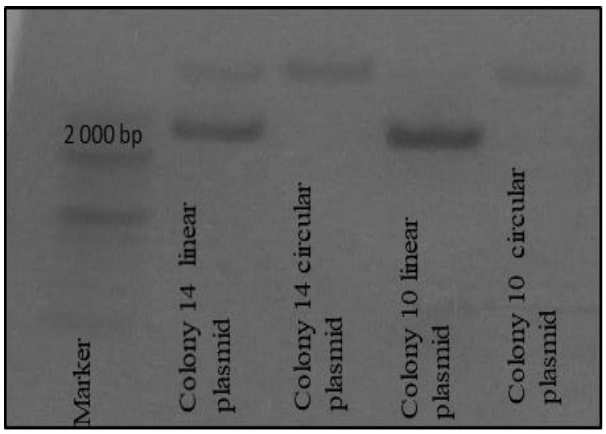


Figure 3.2 – pDR274 linearization with HindIII.
The plasmid has approximately 2100 bp, but when in a circular conformation, it coils slowing its migration rate.
1% TAE-agarose gel, 100 V, 30 min

Colonies 10 and 14 were inoculated into a primary culture. When they reached confluence, we isolated the plasmids and linearized them using the HindIII restriction enzyme. The linearization was tested by running a sample on an agarose gel (Figure 3.2). The linearized plasmids were used as a template to produce the sgRNAs. Table 3.1 shows the sgRNAs quantification.

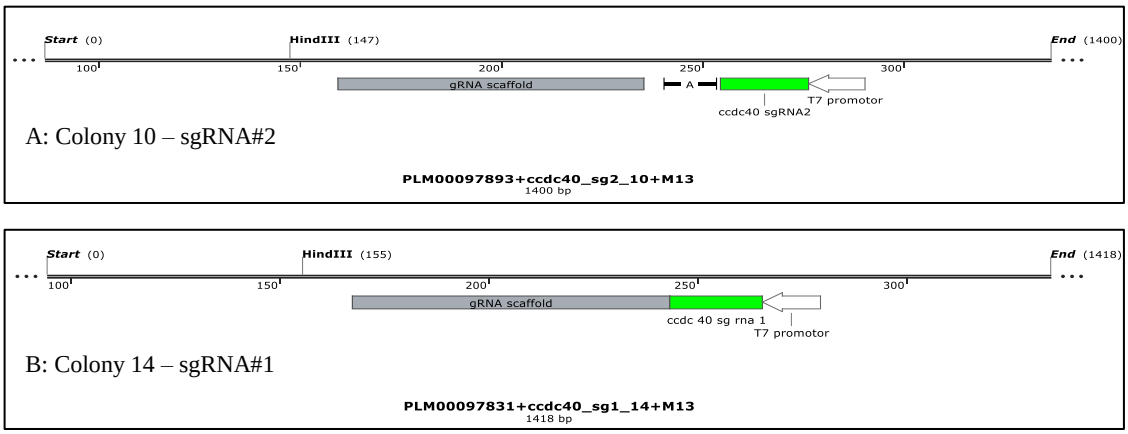


Figure 3.3 – Sequencing results from colony 10 (sgRNA#2) and colony 14 (sgRNA#1).
Both sgRNA templates were inserted between the T7 promoter and the gRNA scaffold. In colony 10 the sgRNA template seems to have been partially duplicated (feature A). For detailed sequence see Figure 7.2 (colony 10) and Figure 7.3 (colony 14).

Table 3.1 – sgRNA#1 and#2 nanodrop quantification

Sample	Mean of RNA concentration (ng/µl)	Total volume (µl)	Amount of sgRNA (ng)
sgRNA#1 (colony 14)	873	9	7857
sgRNA#2 (colony 10)	829	9	7461

3.3. Primers optimization

To test CRISPR-Cas9 efficacy after injection and later screen for crispants, we designed two primer pairs, each flanking sgRNA#1 and sgRNA#2 target-sites. The optimization was done as explained in 2.9.

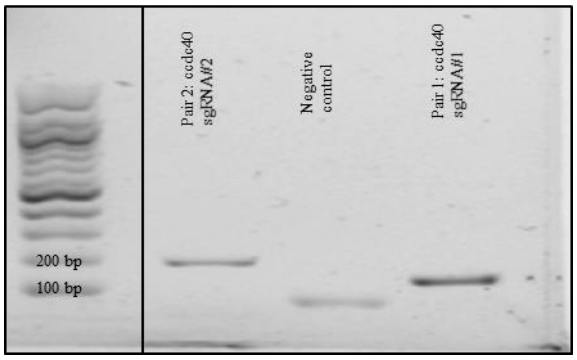


Figure 3.4 – Primer-pairs optimization for zebrafish genomic DNA.
Primer pair 1 and 2 show amplified a band with 143 and 184 bp respectively.
0.8 % TBE-agarose gel, 120 V, 45 min

Primer pair 1 and 2 specifically amplified only one band with 143 bp and 184 bp respectively (Figure 3.4). At this point, primers were ready to be used in testing CRISPR-Cas9 efficacy after injection and later to screen for crispant fish.

3.4. Co-injection of cas9 mRNA and sgRNA and CRISPR-Cas9 effectiveness analysis by PAGE assay

In our first attempt, we injected *cas9* mRNA, in a concentration of 100 ng/µl, together with each sgRNA, in a concentration of 50 ng/µl in one-cell stage zebrafish embryos. Most of the injected embryos developed into larvae with malformations such as curly-up, curly-down, tail atrophy, cyclopia, and cardiac oedema (Figure 3.5 A). Those with severe deformities were not viable and were not picked when testing for cas9-induced mutations. When analysing normal-looking 24 hpf-embryos, no heteroduplexes were found, meaning that the CRISPR-Cas9 system was not working in those embryos and that could be the reason for their normal appearance.

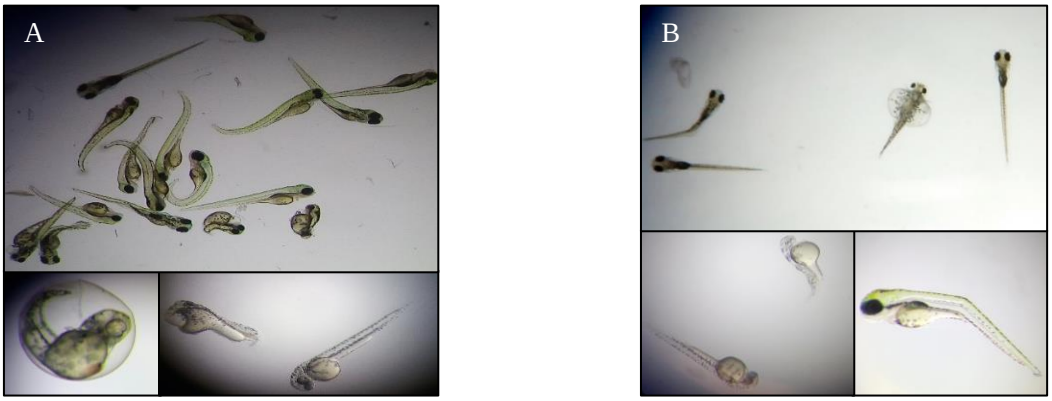


Figure 3.5 – Deformities found in zebrafish larvae (3dpf) after co-injection with Cas9 mRNA and sgRNA.
A: Embryos injected with 100 ng/µl of Cas9 mRNA and 50 ng/µl of sgRNAs. The malformations included curly-up and curly-down tails, cyclopia, cardiac oedema and tail atrophy.
B: Embryos injected with 75 ng/µl of Cas9 mRNA and 37.5 ng/µl of sgRNAs. The malformations included cardiac oedema and tail kinks.

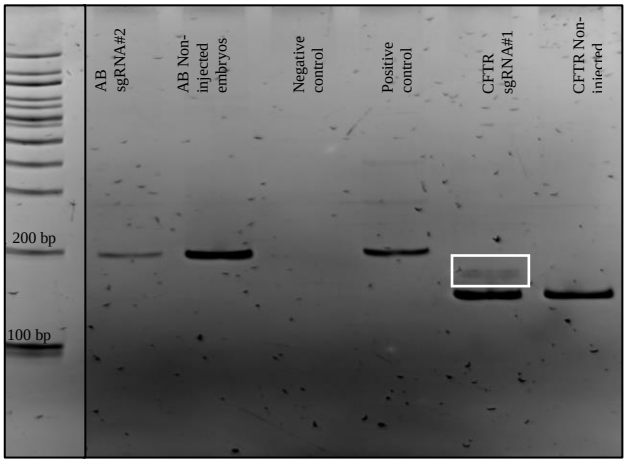


Figure 3.6 – Heteroduplexes detected in PAGE-based analysis of CRISPR-Cas9 system efficacy in *TgBac(cftr-GFP*) embryos. Here we can see heteroduplexes in the sample extracted from *TgBac(cftr-GFP)* embryos injected with Cas9 mRNA (75 ng/µl) and sgRNA#1 (37.5 ng/µl). (white box).

The deformities could be linked to the toxicity of the injected components, therefore, the amounts of Cas9 mRNA and sgRNA were lowered to 75 ng/µl and 37.5 ng/µl respectively.

This time around, the embryos injected with lower concentrations of *cas9* mRNA and sgRNA developed into larvae with fewer deformities (Figure 3.5 B). The PAGE-based analyses showed heteroduplexes in the sample from the *TgBac(cftr-GFP)* embryos injected with sgRNA#1 (Figure 3.6). Since only sgRNA#1 yield positive results, the following attempts focused on it rather than sgRNA#2. sgRNA#2 presents itself partially duplicated (Figure 3.3 A) what can one reason for its ineffectiveness.

To verify if the CRISPR-Cas9 system had successfully cut the target region in defective larvae, a group of them was tested independently. As expected, samples extracted from pools of defective embryos showed a larger collection of heteroduplexes (Figure 3.7). Opportunely, samples coming from apparently non-defective embryos also presented several heteroduplexes (white boxes in Figure 3.7).

These results confirmed that CRISPR-Cas9 system was working, thus the remaining batch of injected fish (possible founders) was transferred to the fish facility main system to grow until reaching maturity.

3.5. Search for founders: zebrafish screening and genotyping

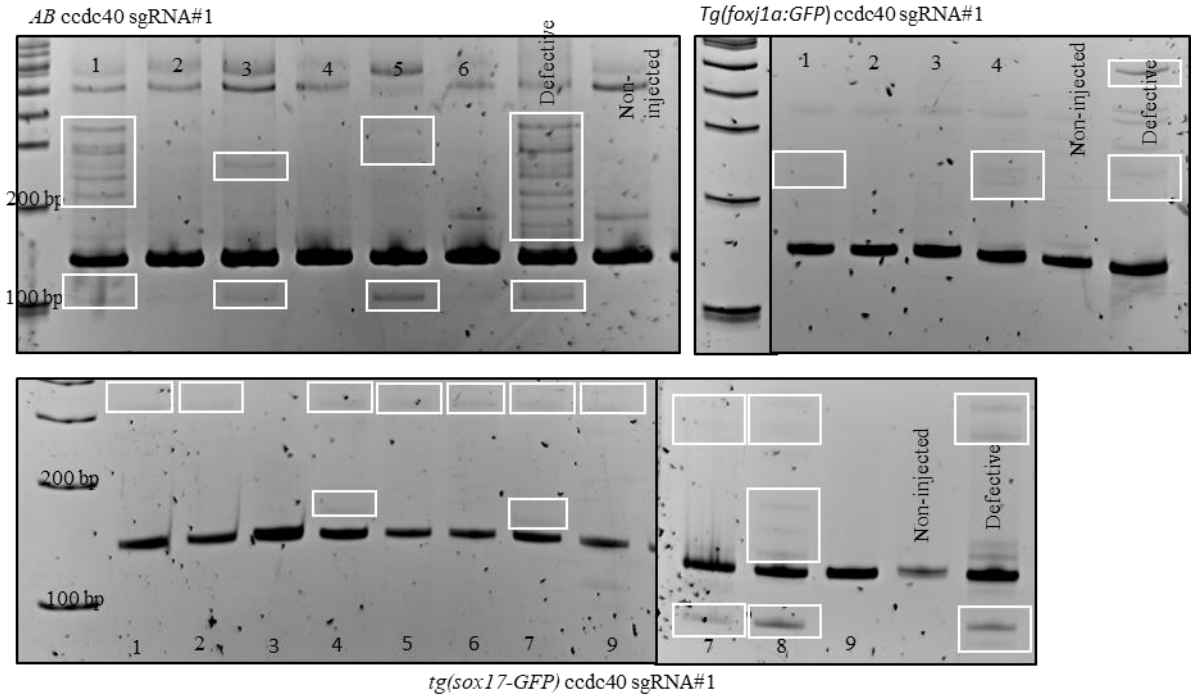


Figure 3.7 – Heteroduplexes detected in PAGE-based assay for testing CRISPR-Cas9 efficacy in *AB, Tg(foxj1a:GFP*) and *Tg(sox17-GFP*) embryos injected with Cas9 mRNA (75 ng/µl) and sgRNA#1 (37.5 ng/µl).
Heteroduplexes are highlighted in white boxes.

This is the most time-consuming task when generating mutants using CRISPR-Cas9 method [102]. It is not only necessary to assure that a certain fish carries a mutation, but, because these can present mosaicism, it is also necessary to guarantee that that mutation is present in the germline cells. This can be done by directly analysing the gametes or by screening the progeny of each individual founder fish.

In this case, the fish, that developed from the CRISPR-Cas9 injected embryos, were screened through the analysis of their progeny in a PAGE-based assay.

The first to be screened were the *TgBAC(cftr-GFP*) zebrafish injected with ccdc40 sgRNA#1. This aquarium only had four females and none of them was a crispant. From the remaining fish (30 males), male 23 analysis (Figure 3.8) showed a similar heteroduplex pattern as before in all three pools of ten embryos(see Figure 3.6).

The next aquarium to be screened was the one with *Tg(foxj1a:GFP)* zebrafish injected with ccdc40 sgRNA#1. From the ten fish (all male) screened in this aquarium, only male 3 produced positive results. In replica A, heteroduplexes appeared slightly above the 300 bp marker (not coincident with the observations in Figure 3.6), while in replica B, a band was detected around the 200 bp mark, consistent with the pattern observed in Figure 3.6.

Nine of the AB fish injected with sgRNA#1 were also screened, but none yield positive results. The remaining fish (*tg(sox17-GFP)*) had not yet reached full maturity at the end of this project's time limit.

The DNA from the revealed heteroduplex bands, as well as from the WT bands, was isolated and sequenced. The Sanger sequencing results showed a chromatogram overlapping in heteroduplex-isolated DNA from male 23 Tg*Bac(cftr-GFP)* (all three polls of embryos) (Figure 3.9) (for complete results, see appendix 7.4). In male 3 *Tg(foxj1a:GFP)* samples, the overlapping was only detected in replica B. The bands in replica A, around 300 bp, were probably due to some unspecific PCR amplification or contamination. The band slightly above 200 bp in replica A, already had appeared in the PAGE-based assay when testing CRISPR-Cas9 efficacy (Figure 3.7).

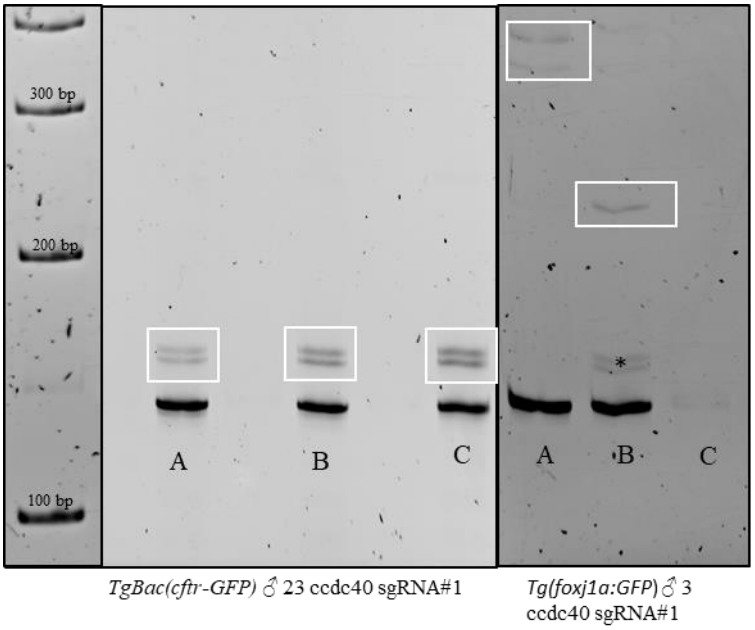


Figure 3.8 – Heteroduplexes detected in a PAGE-based assay used to test *TgBac(cftr-GFP)* ♂ 23 and Fox:GFP ♂ 3 offspring.
The analysis was run in triplicate (A, B, C): in ♂ 23 the results were uniform while in ♂3 were detected two different sets of heteroduplexes (lane C showed a very faint WT band with no heteroduplexes detectable). The bands marked with a * probably resulted from a contamination with ♂ 23 samples.

The chromatogram overlapping is a tell-tale sign of heterozygosity [103], therefore these results allowed the identification of male 23 *TgBAC(cftr-GFP)* and male 3 *Tg(foxj1a:GFP)* as founders (F0).

The analysis of the founders through their offspring is an effective way of assuring that the CRISPR-Cas9 induced indels are passed on through the germline, however, as the sequenced DNA derives from a batch of ten embryos, at this point was not yet possible to determine the exact induced mutation. Among this batch of embryos, it is possible to find a maximum of three different alleles (and therefore peaks). Each embryo can carry the WT allele, inherited from the AB mother, and one of the two paternal alleles, that, in theory, can both carry induced indels. The exact induced mutation can only be determined when analysing a single fish. Aiming for that, the offspring resulting from the mating between male 23 *TgBAC(cftr-GFP)* and male 3 *Tg(foxj1a:GFP)* with AB females (F1), were kept in the fish facility main system to be screened when reaching maturity.

Assuming that only one paternal allele was mutated, 50% of these fish should carry the mutation. However, if both paternal alleles were mutated, the DSBs were resolved differently in each allele, resulting in two different mutations. Therefore, we will have 50% of the fish carrying one mutation and

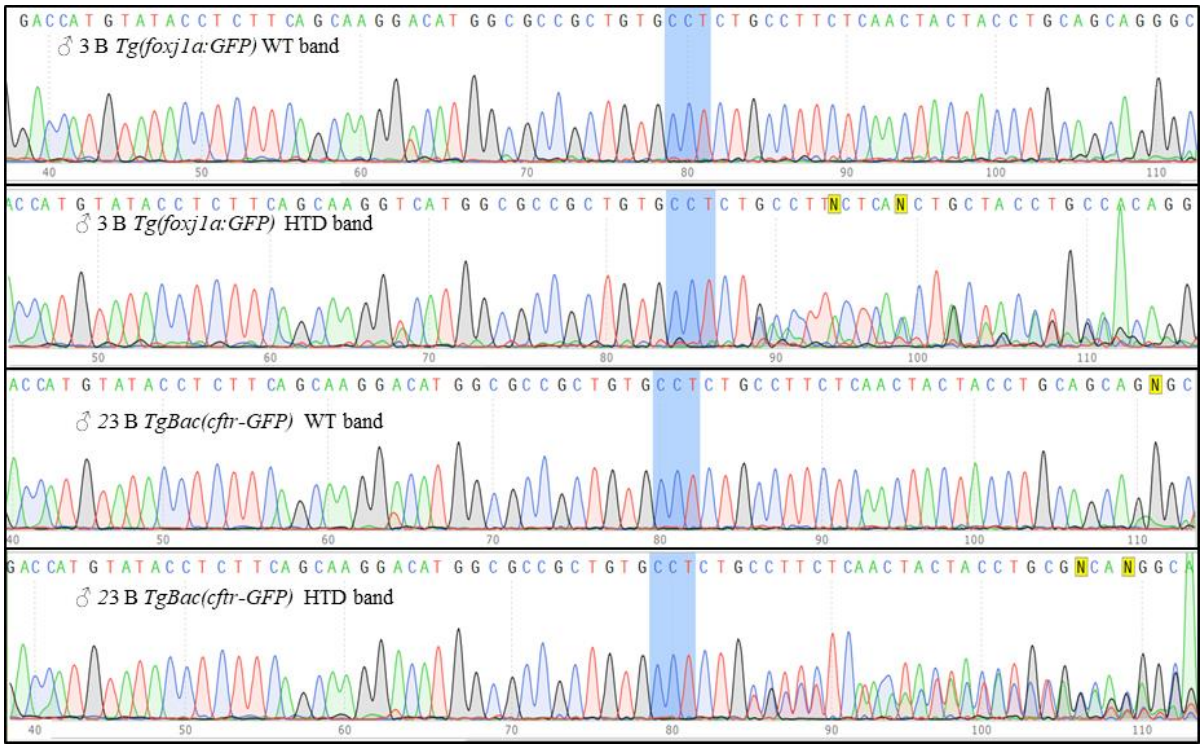


Figure 3.9 – Sanger sequencing results for wildtype (WT) and heteroduplex (HTD) band isolated from ♂ 23 B *TgBAC(cftr-GFP)* and ♂ 3 B *Tg(foxj1a:GFP).*
The referred bands were isolated from the gel shown in Figure 3.8 .
The WT sequences showed individualized peaks, while HTD sequences exhibited overlapping in the chromatogram immediately after the PAM sequence 3' edge.
This sequencing was performed using a reverse primer, therefore the PAM sequence (CCT) highlight in blue is the reverse complement sequence of the canonical PAM sequence 5'-NGG-3'.

50% carrying the other. Fish carrying a certain mutation are to be in crossed among each other, and 25% of the resulting offspring (F3) should be the sought-after *ccdc40* homozygous mutants.

3.6. Morpholino antisense technology - phenotype characterization

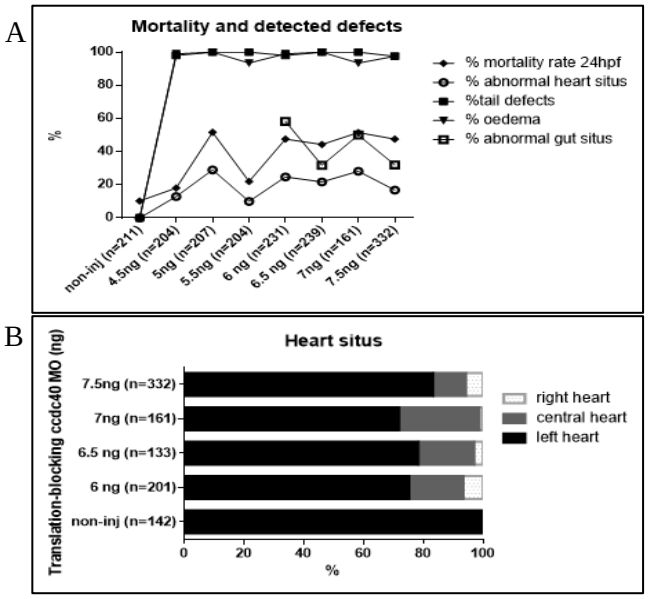


Figure 3.10 – Morphological defects observed in ccdc40 translation-blocking MO injected embryos.
A: Percentage of mortality, tail defects, oedema and abnormal heart and gut situs accordingly to the injected amount of MO.
B: Heart laterality phenotype amid MO injected embryos. Black, grey and white bars represent the percentage of embryos with heart positioned on the left side, in the middle and on the right side respectively.

To study the *ccdc40* loss-of-function phenotype and later compare it with the *ccdc40* -/- mutant, and also to access MO toxicity and effects on laterality axis establishment, we injected several

Figure 3.11.-. Phenotype of lavae (3dpf) injected with ccdc40 translation-blocking MO (6 ng) showing cardiac oedema and tail crooks.

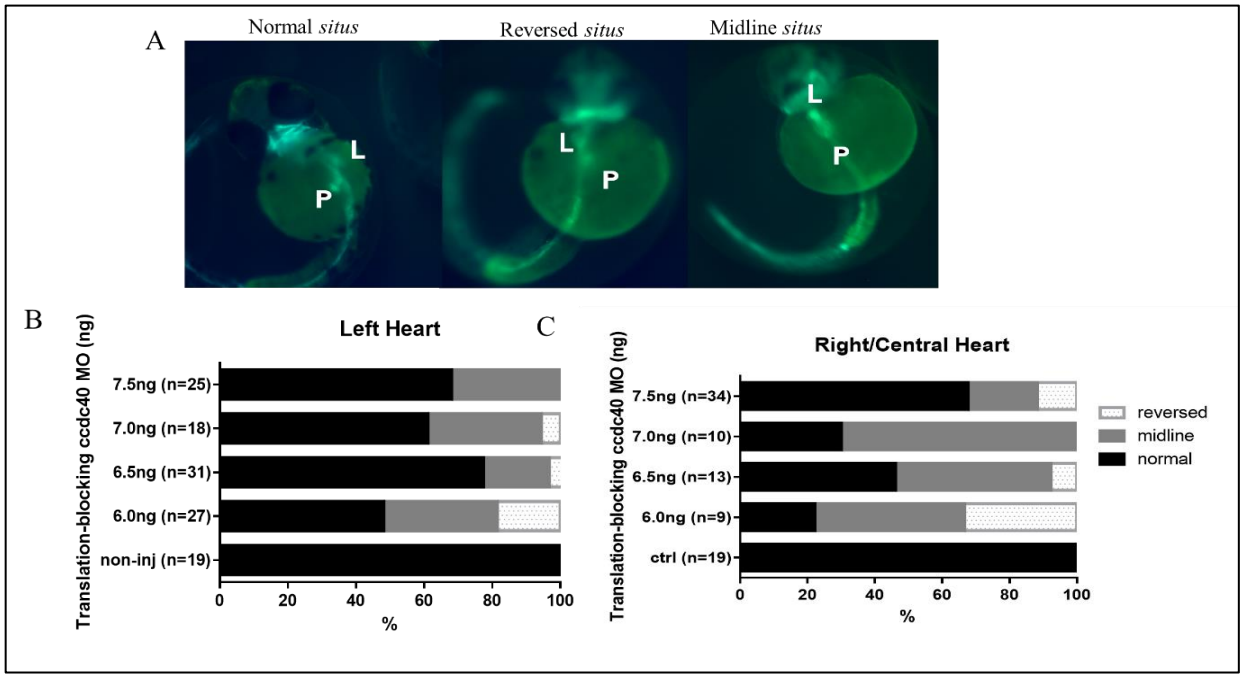


Figure 3.12 – Gut situs from *tg(sox17-GFP)* embryos injected with translation-blocking ccdc40 MO
A: Pictures showing the gut situs of MO-injected *tg(sox17-GFP)* embryos (50 hpf) taken under the fluorescence steroscope.
B and C: The bar graphs represent the gut phenotype of *tg(sox17-GFP)* embryos injected with the translation-blocking ccdc40 MO. The gut situs was evaluated separately in embryos with the heart on the left (B) and embryos with right or central heart (C).
L-liver; P-pancreas
Normal-liver on the left and pancreas on the right; Midline-both liver and pancreas centred; Reversed-liver on the right and pancreas on the left.

quantities of *ccdc40* translation-blocking MO into one-cell stage embryos. We evaluated mortality rates (at 24 hpf), external morphological defects such as cardiac oedema and tail crooks (4 dpf), and internal morphological features such as heart (at 30 hpf), liver and pancreas (at 50 hpf) placement along the L-R axis.
All the evaluated parameters seemed to stabilize from 6 ng onwards, suggesting that these defects were not a direct cause of toxicity. Furthermore, the development of cardiac oedema, tail crooks and lethargy were consistent throughout all the MO amounts injected, being present in almost all of the embryos (Figure 3.11).
Particularly, heart *situs* abnormalities presented themselves in about 20% (6 ng) to 30% (7 ng) of all the injected embryos analysed, in contrast with non-injected ones. (Figure 3.10 B). Among those with heart misplacement, it was noticeable a higher proportion of central hearts, with 70% (6 ng) to 95% (7 ng) of the embryos having a heart with no loop.

Besides the heart *situs*, the injected embryos also were screened for the liver and pancreas positioning (Figure 3.12A). This evaluation was done separately for embryos with a left heart and for embryos with right or central heart (Figure 3.12 B-C) and in both cases, the gut *situs* was affected.

Embryos with a left-sided heart presented between 23% (6.5 ng) and 50% (6 ng) of gut positioning defects, while in embryos with a central or right-sided heart the gut *situs* defects were present in 32% (7 ng) to 77% (6 ng). The fact that these numbers do not correlate directly to the amount of MO-injected can be due to the randomization of the *situs* establishment.

Also, in the case of central and right-hearted embryos, the number considered was low, what, together with the morpholino variability, adds to this discrepancy. Interestingly, when the gut was mispositioned the prevalent phenotype was the midline *situs*. This happened despite the heart position but was more evident in embryos with a central or right-sided heart.

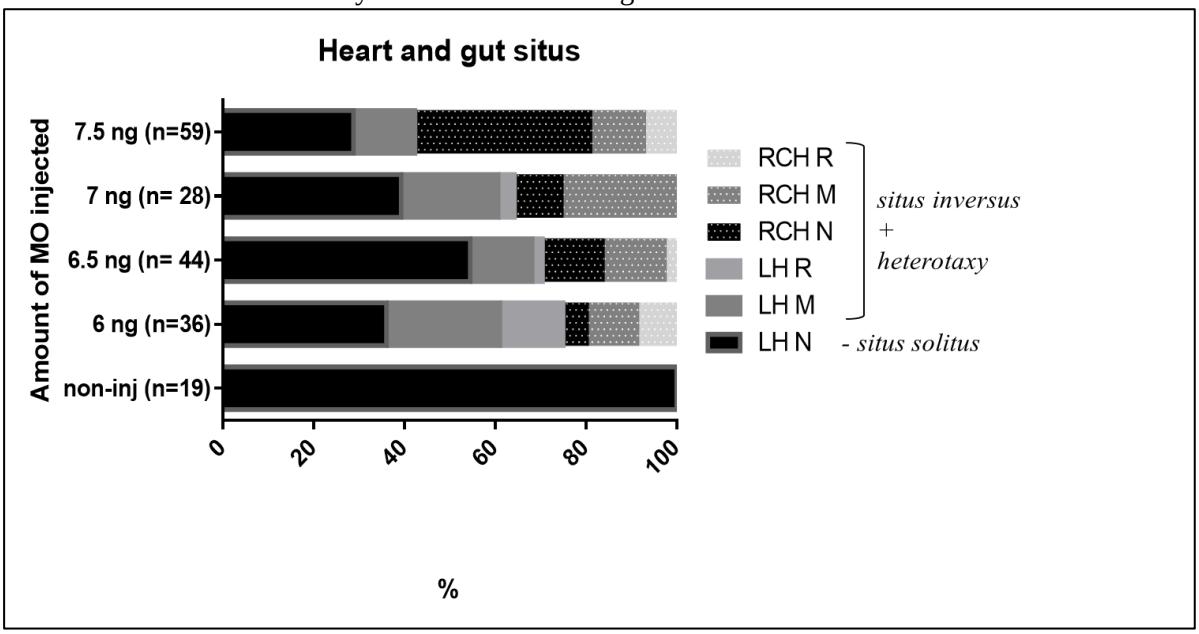


Figure 3.13 – Heart and gut situs combined.
The gut situs assessment for embryos with central and right hearts was done together. Therefore, cases of situs inversus and heterotaxy cannot be separated.
RCH: central or right heart, LH: left heart, N: normal gut (left liver and right pancreas), R: reversed gut.

Figure 3.13 shows the merge of heart and gut *situs* results. Since the embryos with central and right-side hearts were not separated, it is not possible to distinguish *situs inversus* from heterotaxy. Still, we can say that in 40% (6.5 ng) to 70% (7.5 ng) of the cases presented *situs* defects (Figure 3.13) as well as tail crooks, oedema and lethargy.

4. Discussion

In this project, our goal was to create a *ccdc40 -/-* zebrafish mutant line (position 116 aa). As the chosen animal model takes three months to reach maturity and the desired mutant can only be obtained after three generations, we studied MO-injected embryos instead.

Our results in MO-injected embryos showed *situs* defects in up to 70% (7.5 ng) of the cases as well as tail crooks, oedema and lethargy.

In studies with *lok* mutants, known to have a recessive mutation in *ccdc40*, embryos are shown to develop a similar phenotype with curly down tails, organ *situs* problems in 50% of the cases (*situs inversus* and heterotaxy) [90], lethargy and cardiac oedema [104].

In both cases, it makes sense that the organ *situs* is randomly established because when cilia are immotile or beat in an abnormal way, *nodal* expression does not become asymmetric to the left on the LPM and instead it can become either absent, bilateral, expressed on the right or on the left side of the LPM.

Although both phenotypes represent a Ccdc40 knockdown, the MO-induced phenotype appeared stronger. MO-induced *situs* defects seemed accordant with *lok* phenotype, ranging from 40% to 70%, however, the cardiac oedema is more severe and the tail curvatures more convoluted *Figure 3.11*. in the morphants than in the *lok* homozygous mutants.

At 3 dpf WT embryos break out of the chorion and straighten their tails. In *lok* homozygous mutants, as well as in our MO-injected embryos this does not occur. They kept their curvatures even after hatching. *lok* mutants have their tails bent ventrally and in MO-injected embryos, the tail can bend downwards, upwards or to the sides and present more than one curvature. *lok* mutants [71,104], as well as the *ccdc40* MO-injected embryos in this study, died at 7 dpf.

In zebrafish, *ccdc40* gene has two transcripts annotated. One encodes for the CCDC40 protein and the other is a non-coding retained intron transcript[f].

In humans, there are six transcripts producing protein and other six transcripts with premature stop codons[g]. Based on the literature, around half of the homozygous patients with a CCDC40 mutation carry the frameshift mutation c.248delC [76,90] This mutation inserts a stop codon upstream the region codifying the coiled-coil domain. The other half of the mutations take place much closer to the end of the codifying region (such as c.3129delC and c.2824-2825insTGT) allowing the integral transcription of smaller CCDC40 isoforms, containing a truncated coiled-coil domain, and non-coding transcripts.

The axonemal ultrastructures shown by Antony *et al*. (2013) [76] and by Becker-Heck *et al*. (2011) [90] are from nasal cilia collected from patients carrying early termination mutations and are similar to the ciliary ultrastructure reported for the *lok* homozygous mutants. We were not able to find reported TEM images from cilia collected from PCD patients with early mutations such as c.248delC (closer to the N terminal) to compare the phenotypes. Both events, mutations closer to the N extremity or mutations closer to the C terminal of the protein, cause PCD, meaning that the complete CCDC40 variant is necessary for the correct IDA assembly in the ciliary axoneme. The other protein variants may play a regulatory role. It was described that transcripts with premature stop codons can have a role in gene regulation [105,106]. In the *lok* mutants, only the protein-coding transcript is affected, with the non-coding transcript still being produced normally. In the MO-injected embryos, both transcripts were

[f] http://www.ensembl.org/Danio_rerio/Gene/Summary?db=core;g=ENSDARG00000100584;r=6:21779819-21821089

[g] http://www.ensembl.org/Homo_sapiens/Transcript/ProteinSummary?db=core;g=ENSG00000141519;r=17:80036632-80100613;t=ENST00000269318

knocked down, offering an explanation to the more severe defects, that in this situation can result from both IDA assembly problems and absence of regulatory elements.

Apart from being structurally necessary for correct assembly of IDA, Ccdc40 is also needed for the recruitment of tubulin polyglutamilases, [74] that promote tubulin stabilization. This was found in *Chlamydomonas*, that only expresses the protein-coding transcript[h], not clarifying if proteins with truncated coiled-coil domains can still do the recruitment of the tubulin polyglutamilases. However, in case they still preserve that ability, the truncated Ccdc40 protein can be able to promote tubulin stabilization, even if with less efficiency in both *lok* zebrafish mutants and PCD patients carrying early termination mutations near the C terminal.

To verify this, the ciliary structure organization of the CCDC40 -/- mutant line obtained by CRISPR-Cas9 must be analyzed using an antibody for glutamylated tubulin and compared with that of *lok* mutants.

It would also be interesting to analyse the cilia beat frequency of both mutants to verify if higher polyglutamylation levels are reflected in increased the cilia motility control as suggested in [107]

The analysis of cilia beat frequency in *lok* mutants revealed that cilia beat within a reduced amplitude and with a more rigid movement [108]. This stiffer movement can be due to the portion of the coiled-coil domain that is present and still allows IDA assembly (but in lower numbers as found in Becker-heck *et al*. 2011 [90]) that together with an intact ODA results in the presence of some residual movement.

In the case of MO-injected embryos, the ccdc40 protein is not present at all and the IDA assembly should not be possible. We did not measure the cilia beat frequency in MO-injected embryos but, if this assumption is right, no movement is expected in MO-injected or *ccdc40* -/- mutants because in both cases the coiled-coil domain necessary for movent regulation [70] is not present.

According to Grimes *et al*. the spinal curvatures in *lok* mutants appeared as a result of abnormalities in CSF flow and vertebrae morphogenesis [109]. The resulting phenotype is not viable, however, if *lok* embryos are injected with WT *ccdc40* mRNA at one cell stage, they survive to develop spinal curvatures that resemble human idiopathic scoliosis [109]. The mRNA injected only counteracts the absence of the protein-coding transcript and the non-coding transcript is not compensated for. If the co-injection of mRNA for both coding and non-coding transcripts leads to a complete phenotype rescue resulting in fish with not spinal curvatures, then the non-coding transcript may play an important role in osteogenesis regulation.

One possible explanation for the spinal defects detected in our knockdown of *ccdc40* is in the evidence connecting cilia and osteogenesis in a PKD1-dependent signalling pathway. PKD1 (present in the cilia membrane of the KV cilia) is also expressed in cartilage and somites and Pkd1 mutant mice develop problems in bones and cartilage, exhibiting spinal curvatures and vertebra deformities [110]. We know from previously unpublished work in the lab that the KV cells will later integrate the tail tissue [111]. In particular, they will become part of the notochord, hypochord and somites. So, it is possible that defects in Pkd1 could lead to tail deformities. Nevertheless, a link between the lack of Ccdc40 and abnormal expression of Pkd1 would need to be first established.

[h] http://plants.ensembl.org/Chlamydomonas_reinhardtii/Transcript/Domains?db=core;g=CHLREDRAFT_170513;r=DS496120:1834038-1840048;t=EDP04735

5. Conclusion and Prospects

Based on the MO phenotype, CCDC40 is essential for correct organ *situs*. It seems that the phenotype obtained when injecting translation-blocking MO is more severe than that obtained in *lok*

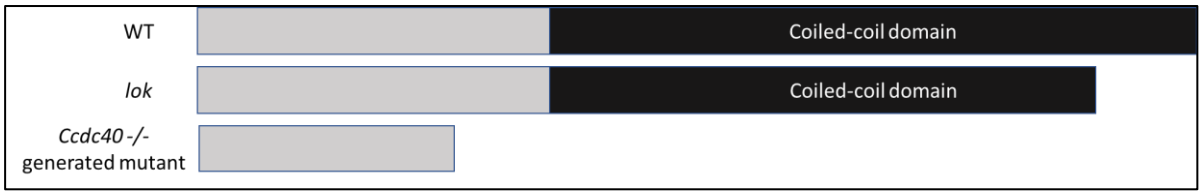


Figure 5.1 – Diagram showing the predicted zebrafish Ccdc40 protein domains in WT, *lok* mutants and Ccdc40 generated mutant using CRISPR-Cas9 system.
In WT zebrafish, the coiled-coil domain ranges from the 134 aa to the 854 aa. In *lok* mutants, this domain is truncated at the 778 aa position. The mutant line generated in this work (*Cccdc40* -/-) synthesize a protein truncated at the 116 aa, with no coiled-coil domain.

mutants.

lok mutants assemble a ccdc40 protein that still has part of the coiled-coil domain, while the homozygous *ccdc40* -/- mutant produced expresses only the first 346 of the total 2826 bp, which excludes the all the extension of the coiled-coil domain (Figure 5.1). Similar mutations causing PCD have been identified in patients and up to date, no one investigated if a mutation truncating a part of the protein C-terminal (*lok*-like) and a mutation closer to the N-terminal, excluding the total extension of the coiled-coil domain, result in similar phenotypes such as ultrastructural defects and tubulin post-translational modification patterners. The comparative analysis of both phenotypes will help to clarify these questions.

Since this MO was designed to exclude the initial ATG codon to make possible rescue experiences, we attempted to clone the ccdc40 gene into a PCS2+ *mcherry* vector but, with the tested pair of primers (appendix *7.5*), it was not possible to obtain the desired amplification. However, with the production of mRNA for the protein coding sequence and for the non-coding transcript, it will be possible to study the effected of each one alone in the cilia motility and development of spinal cord in zebrafish.

To conclude this project, in the immediate future, it is necessary to incross the heterozygotes (F1) to achieve a homozygous mutant and identify the exact mutation introduced. When getting this *ccdc40* -/- mutant, its phenotype can be compared with the MO phenotype validating it or not.

The mutation induced in *ccdc40* -/- mutant lies in a region homologous to the sequence in the human genome where the mutation c.961 C>T has been identified. This mutation appears to be in a region regarded as a “hot spot” where at least other two mutations have been identified [76]. Therefore, with the exact mutation identified, the *ccdc40* -/- mutant line can be used to test gene therapy approaches to try to correct it back to normal. These same approaches can then be applied in patient-derived cells, leading to a way of alleviating the respiratory symptoms present in PCD in a similar way to what has been done to other diseases [112].

6. Bibliography

1. Langdon YG, Mullins MC. 2011 Maternal and Zygotic Control of Zebrafish Dorsoventral Axial Patterning. *Annu. Rev. Genet.* **45**, 357–377. (doi:10.1146/annurev-genet-110410-132517)

2. Piccolo S, Agius E, Leyns L, Bhattacharyya S, Grunz H, Bouwmeester T, De Robertis EM. 1999 The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature* **397**, 707–710. (doi:10.1038/17820)

3. Agathon A, Thisse C, Thisse B. 2003 The molecular nature of the zebrafish tail organizer. *Nature* **424**, 448–452.

(doi:10.1038/nature01822)

4. Hamada H, Meno C, Watanabe D, Saijoh Y. 2002 Establishment of vertebrate left–right asymmetry. *Nat. Rev. Genet.* **3**, 103–113. (doi:10.1038/nrg732)

5. Collignon J, Varlet I, Robertson EJ. 1996 Relationship between asymmetric nodal expression and the direction of embryonic turning. *Nature* **381**, 155–158. (doi:10.1038/381155a0)

6. Nakamura T, Hamada H. 2012 Left-right patterning: conserved and divergent mechanisms. *Development* **139**, 3257–3262. (doi:10.1242/dev.061606)

7. Blum M, Weber T, Beyer T, Vick P. 2009 Evolution of leftward flow. *Semin. Cell Dev. Biol.* **20**, 464–471. (doi:10.1016/j.semcdb.2008.11.005)

8. Levin M, Palmer AR. 2007 Left-right patterning from the inside out: Widespread evidence for intracellular control. *BioEssays* **29**, 271–287. (doi:10.1002/bies.20545)

9. Dasgupta A, Amack JD. 2016 Cilia in vertebrate left-right patterning. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **371**, 20150410. (doi:10.1098/rstb.2015.0410)

10. Nakamura T *et al.* 2012 Fluid flow and interlinked feedback loops establish left-right asymmetric decay of Cerl2 mRNA. *Nat. Commun.* **3**, 1322. (doi:10.1038/ncomms2319)

11. Marques S, Borges AC, Silva AC, Freitas S, Cordenonsi M. 2004 The activity of the Nodal antagonist. *Genes Dev.* , 2342–2347. (doi:10.1101/gad.306504.2342)

12. Tsukui T *et al.* 1999 Multiple left-right asymmetry defects in Shh-/- mutant mice unveil a convergence of the Shh and retinoic acid pathways in the control of Lefty-1. *Proc. Natl. Acad. Sci.* **96**, 11376–11381. (doi:10.1073/pnas.96.20.11376)

13. Meno C, Ito Y, Saijoh Y, Matsuda Y, Tashiro K. 1997 Two closely-related left-right asymmetrically expressed genes, and domains, chromosomal linkage and direct neuralizing lefty-2 : their distinct expression activity in lefty-1 Xenopus embryos. *Genes to Cells* **2**, 513–524. (doi:10.1046/j.1365-2443.1997.1400338.x)

14. Meno C *et al.* 2001 Diffusion of Nodal Signaling Activity in the Absence of the Feedback Inhibitor Lefty2. *Dev. Cell* **1**, 127–138. (doi:10.1016/S1534-5807(01)00006-5)

15. Ji Y, Buel SM, Amack JD. 2016 Mutations in zebrafish pitx2 model congenital malformations in Axenfeld-Rieger syndrome but do not disrupt left-right placement of visceral organs. *Dev. Biol.* **416**, 69–81. (doi:10.1016/j.ydbio.2016.06.010)

16. Nonaka S, Shiratori H, Saijoh Y, Hamada H. 2002 Determination of left–right patterning of the mouse embryo by artificial nodal flow. *Nature* **418**, 96–99. (doi:10.1038/nature00849)

17. McGrath J, Somlo S, Makova S, Tian X, Brueckner M. 2003 Two populations of node monocilia initiate left-right asymmetry in the mouse. *Cell* **114**, 61–73. (doi:10.1016/S0092-8674(03)00511-7)

18. Nonaka S, Tanaka Y, Okada Y, Takeda S, Harada A, Kanai Y, Kido M, Hirokawa N. 1998 Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell* **95**, 829–837. (doi:10.1016/S0092-8674(00)81705-5)

19. Takeda S, Yonekawa Y, Tanaka Y, Okada Y, Nonaka S, Hirokawa N. 1999 Left-right asymmetry and kinesin superfamily protein KIF3a: New insights in determination of laterality and mesoderm induction by KIF3A(-/-) mice analysis. *J. Cell Biol.* **145**, 825–836. (doi:10.1083/jcb.145.4.825)

20. Supp DM, Brueckner M, Kuehn MR, Witte DP, Lowe LA, McGrath J, Corrales J, Potter SS. 1999 Targeted deletion of the ATP binding domain of left-right dynein confirms its role in specifying development of left-right asymmetries. *Development* **23**, 83–88. (doi:10.1016/j.jdiacomp.2008.01.002.Postural)

21. Tabin CJ, Vogan KJ. 2003 A two-cilia model for vertebrate left-right axis specification. *Genes Dev.* **17**, 1–6. (doi:10.1101/gad.1053803)

22. Okada Y, Nonaka S, Tanaka Y, Saijoh Y, Hamada H, Hirokawa N. 1999 Abnormal Nodal Flow Precedes Situs Inversus in iv and inv mice. *Mol. Cell* **4**, 459–468. (doi:10.1016/S1097-2765(00)80197-5)

23. Tanaka Y, Okada Y, Hirokawa N. 2005 FGF-induced vesicular release of Sonic hedgehog and retinoic acid in leftward nodal flow is critical for left–right determination. *Nature* **435**, 172–177. (doi:10.1038/nature03494)

24. Lowe LA, Supp DM, Sampath K, Yokoyama T, Wright CVE, Potter SS, Overbeek P, Kuehn MR. 1996 Conserved left–right asymmetry of nodal expression and alterations in murine situs inversus. *Nature* **381**, 158–161. (doi:10.1038/381158a0)

25. Kamura K, Kobayashi D, Uehara Y, Koshida S, Iijima N, Kudo A, Yokoyama T, Takeda H. 2011 Pkd1l1 complexes with Pkd2 on motile cilia and functions to establish the left-right axis. *Development* **138**, 1121–1129. (doi:10.1242/dev.058271)

26. Pazour GJ, San Agustin JT, Follit JA, Rosenbaum JL, Witman GB. 2002 Polycystin-2 localizes to kidney cilia and the ciliary level is elevated in orpk mice with polycystic kidney disease. *Curr. Biol.* **12**, R378–R380. (doi:10.1016/S0960-9822(02)00877-1)

27. Hanaoka K, Qian F, Boletta a, Bhunia a K, Piontek K, Tsiokas L, Sukhatme VP, Guggino WB, Germino GG. 2000 Co-assembly of polycystin-1 and -2 produces unique cation-permeable currents. *Nature* **408**, 990–994. (doi:10.1038/35050128)

28. Field S *et al.* 2011 Pkd1l1 establishes left-right asymmetry and physically interacts with Pkd2. *Development* **138**, 1131–1142. (doi:10.1242/dev.058149)

29. Praetorius HA, Spring KR. 2003 Removal of the MDCK cell primary cilium abolishes flow sensing. *J. Membr. Biol.* **191**, 69–76. (doi:10.1007/s00232-002-1042-4)

30. Pennekamp P, Karcher C, Fischer A, Schweickert A, Skryabin B, Blum M, Dworniczak B. 2002 The Ion Channel Polycystin-2 Is Required for Left-Right Axis Determination in Mice. **12**, 938–943.

31. Grimes DT *et al.* 2016 Genetic Analysis Reveals a Hierarchy of Interactions between Polycystin-Encoding Genes and Genes Controlling Cilia Function during Left-Right Determination. *PLoS Genet.* **12**, 1–25. (doi:10.1371/journal.pgen.1006070)

32. Grimes DT, Burdine RD. 2017 Left–Right Patterning: Breaking Symmetry to Asymmetric Morphogenesis. *Trends Genet.* **33**, 616–628. (doi:10.1016/j.tig.2017.06.004)

33. Delling M *et al.* 2016 Primary cilia are not calcium-responsive mechanosensors. *Nature* **531**, 656–660. (doi:10.1038/nature17426)

34. Fliegauf M, Benzing T, Omran H. 2007 When cilia go bad: cilia defects and ciliopathies. *Nat. Rev. Mol. Cell Biol.* **8**, 880–893. (doi:10.1038/nrm2278)

35. Quarmby LM, Parker JDK. 2005 Cilia and the cell cycle?: Figure 1. *J. Cell Biol.* **169**, 707–710. (doi:10.1083/jcb.200503053)

36. Horani A, Brody SL, Ferkol TW. 2014 Picking up speed: advances in the genetics of primary ciliary dyskinesia. *Pediatr Res* **75**, 158–164. (doi:10.1038/pr.2013.200)

37. Rosenbaum JL, Witman GB. 2002 Intraflagellar transport. **3**. (doi:10.1038/nrm952)

38. Hirokawa N. 1998 Superfamily and Dynein Kinesin of and the Mechanism Proteins Organelle Transport. *Science (80-.).* **279**, 519–526. (doi:10.1126/science.279.5350.519)

39. Miki H, Setou M, Kaneshiro K, Hirokawa N. 2001 All kinesin superfamily protein, KIF, genes in mouse and human. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 7004–11. (doi:10.1073/pnas.111145398)

40. Gray KA, Yates B, Seal RL, Wright MW, Bruford EA. 2015 Genenames . org : the HGNC resources in 2015. **43**, 1079–1085. (doi:10.1093/nar/gku1071)

41. Roossien DH, Miller KE, Gallo G. 2015 Ciliobrevins as tools for studying dynein motor function. *Front. Cell. Neurosci.* **9**, 252. (doi:10.3389/fncel.2015.00252)

42. Kramer-Zucker AG. 2005 Cilia-driven fluid flow in the zebrafish pronephros, brain and Kupffer's vesicle is required for normal organogenesis. *Development* **132**, 1907–1921. (doi:10.1242/dev.01772)

43. Lin J, Tritschler D, Song K, Barber CF, Cobb JS, Porter ME, Nicastro D. 2011 Building blocks of the nexin-dynein regulatory complex in chlamydomonas flagella. *J. Biol. Chem.* **286**, 29175–29191. (doi:10.1074/jbc.M111.241760)

44. Lee JD, Anderson K V. 2008 Morphogenesis of the node and notochord: the cellular basis for the establishment and maintenance of left-right asymmetry in the mouse. *Dev. Biol.* **237**, 3464–3476. (doi:10.1002/dvdy.21598.Morphogenesis)

45. Okada Y, Takeda S, Tanaka Y, Belmonte JCI, Hirokawa N. 2005 Mechanism of nodal flow: A conserved symmetry breaking event in left-right axis determination. *Cell* **121**, 633–644. (doi:10.1016/j.cell.2005.04.008)

46. Porter ME, Sale WS. 2000 The 9 + 2 axoneme anchors multiple inner arm dyneins and a network of kinases and phosphatases that control motility. *J. Cell Biol.* **151**, 37–42. (doi:10.1083/jcb.151.5.F37)

47. NCBI. In press. (outer dynein arm) AND "Homo sapiens"[porgn:__txid9606] - Gene - NCBI. See https://www.ncbi.nlm.nih.gov/gene/?term=(outer+dynein+arm)+AND+%22Homo+sapiens%22%5Bporgn%3A__txi

d9606%5D (accessed on 10 August 2017).

48. Ben Khelifa M *et al.* 2014 Mutations in DNAH1, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella. *Am. J. Hum. Genet.* **94**, 95–104. (doi:10.1016/j.ajhg.2013.11.017)

49. Olbrich H *et al.* 2002 Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left–right asymmetry. *Nat. Genet.* **30**, 143–144. (doi:10.1038/ng817)

50. Wheatley DN, Wang AM, Strugnell GE. 1996 Expression of primary cilia in mammalian cells. *Cell Biol. Int.* **20**, 73–81. (doi:10.1006/cbir.1996.0011)

51. Davy BE, Robinson ML. 2003 Congenital hydrocephalus in hy3 mice is caused by a frameshift mutation in Hydin, a large novel gene. *Hum. Mol. Genet.* **12**, 1163–1170. (doi:10.1093/hmg/ddg122)

52. Badano JL, Mitsuma N, Beales PL, Katsanis N. 2006 The Ciliopathies: An Emerging Class of Human Genetic Disorders. *Annu. Rev. Genomics Hum. Genet.* **7**, 125–148. (doi:10.1146/annurev.genom.7.080505.115610)

53. Afzelius B a. 1976 A human syndrome caused by immotile cilia. *Science (80-.).* **193**, 317–319. (doi:10.1126/science.1084576)

54. Camner P, Mossberg B, Afzelius BA. 1975 Evidence of congenitally nonfunctioning cilia in the tracheobronchial tract in two subjects. *Am. Rev. Respir. Dis.* **112**, 807–809. (doi:10.1164/arrd.1975.112.6.807)

55. Eliasson R, Mossberg B, Camner P, Afzelius BA. 1977 The Immotile-Cilia Syndrome. *N. Engl. J. Med.* **297**, 1–6. (doi:10.1056/NEJM197707072970101)

56. Peeters H, Devriendt K. 2006 Human laterality disorders. *Eur. J. Med. Genet.* **49**, 349–362. (doi:10.1016/j.ejmg.2005.12.003)

57. Shiraishi I, Ichikawa H. 2012 Human Heterotaxy Syndrome. *Circ. J.* **76**, 2066–2075. (doi:10.1253/circj.CJ-12-0957)

58. Sutherland MJ, Ware SM. 2009 Disorders of left-right asymmetry: Heterotaxy and situs inversus. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **151**, 307–317. (doi:10.1002/ajmg.c.30228)

59. Shapiro AJ *et al.* 2014 Laterality Defects Other Than Situs Inversus Totalis in Primary Ciliary Dyskinesia. *Chest* **146**, 1176–1186. (doi:10.1378/chest.13-1704)

60. Kennedy MP *et al.* 2007 Congenital heart disease and other heterotaxic defects in a large cohort of patients with primary ciliary dyskinesia. *Circulation* **115**, 2814–2821. (doi:10.1161/CIRCULATIONAHA.106.649038)

61. Brueckner M. 2007 Heterotaxia, congenital heart disease, and primary ciliary dyskinesia. *Circulation* **115**, 2793–2795. (doi:10.1161/CIRCULATIONAHA.107.699256)

62. Bartram U, Wirbelauer J, Speer CP. 2005 Heterotaxy syndrome - Asplenia and polysplenia as indicators of visceral malposition and complex congenital heart disease. *Biol. Neonate* **88**, 278–290. (doi:10.1159/000087625)

63. Ferkol T. 2017 Movement. *Paediatr. Respir. Rev.* , 16–17. (doi:10.1016/j.prrv.2017.06.005)

64. Omran H *et al.* 2008 Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins. *Nature* **456**, 611–616. (doi:10.1038/nature07471)

65. Sui W *et al.* 2015 CCDC40 mutation as a cause of primary ciliary dyskinesia: a case report and review of literature. *Clin. Respir. J.* , n/a-n/a. (doi:10.1111/crj.12268)

66. Warner WA, Sanchez R, Dawoodian A, Li E, Momand J. 2013 NIH Public Access. **80**, 631–637. (doi:10.1111/j.1747-0285.2012.01428.x.Identification)

67. Blanchon S *et al.* 2012 Delineation of CCDC39/CCDC40 mutation spectrum and associated phenotypes in primary ciliary dyskinesia. *J. Med. Genet.* **49**, 410–416. (doi:10.1136/jmedgenet-2012-100867)

68. Casey JP, McGettigan PA, Healy F, Hogg C, Reynolds A, Kennedy BN, Ennis S, Slattery D, Lynch SA. 2015 Unexpected genetic heterogeneity for primary ciliary dyskinesia in the Irish Traveller population. *Eur. J. Hum. Genet.* **23**, 210–217. (doi:10.1038/ejhg.2014.79)

69. Reichlin S. 1969 Handbook of Experimental Pharmacology. *Am. J. Med. Sci.* **258**, 366. (doi:10.1097/00000441-196911000-00008)

70. Burkhard P, Stetefeld J, Strelkov S V. 2001 Coiled coils: A highly versatile protein folding motif. *Trends Cell Biol.* **11**, 82–88. (doi:10.1016/S0962-8924(00)01898-5)

71. Becker-heck A *et al.* 2011 The coiled-coik domain containing protein CCDC40 is essentioal for morile cilia function and left-right axis formation. **43**, 79–84. (doi:10.1038/ng.727.The)

72. Merveille A-C *et al.* 2011 CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. *Nat. Genet.* **43**, 72–8. (doi:10.1038/ng.726)

73. Werner-Peterson R, Sloboda RD. 2013 Methylation of structural components of the axoneme occurs during flagellar disassembly. *Biochemistry* **52**, 8501–8509. (doi:10.1021/bi4011623)

74. Lin H, Zhang Z, Guo S, Chen F, Kessler JM, Wang YM, Dutcher SK. 2015 A NIMA-Related Kinase Suppresses the Flagellar Instability Associated with the Loss of Multiple Axonemal Structures. *PLoS Genet.* **11**, 1–25. (doi:10.1371/journal.pgen.1005508)

75. Kubo T, Yagi T, Kamiya R. 2012 Tubulin polyglutamylation regulates flagellar motility by controlling a specific inner-arm dynein that interacts with the dynein regulatory complex. *Cytoskeleton* **69**, 1059–1068. (doi:10.1002/cm.21075)

76. Antony D *et al.* 2013 Mutations in CCDC39 and CCDC40 are the Major Cause of Primary Ciliary Dyskinesia with Axonemal Disorganization and Absent Inner Dynein Arms. *Hum. Mutat.* **34**, 462–472. (doi:10.1002/humu.22261)

77. Oda T;, Yanagisawa H, Kamiya R, Kikkawa M. 2014 A molecular ruler determines the repeat length in eukaryotic cilia and flagella. *Science (80-.).* **346**, 857–860. (doi:10.1126/science.1260214)

78. Lobo LJ, Zariwala MA, Noone PG. 2014 Primary ciliary dyskinesia. *Qjm* **107**, 691–699. (doi:10.1093/qjmed/hcu063)

79. Kim J, Hake SB, Roeder RG. 2005 The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol. Cell* **20**, 759–770. (doi:10.1016/j.molcel.2005.11.012)

80. Bolotin A, Quinquis B, Sorokin A, Dusko Ehrlich S. 2005 Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–2561. (doi:10.1099/mic.0.28048-0)

81. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E. 2005 Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182. (doi:10.1007/s00239-004-0046-3)

82. Deltcheva E, Chylinski K, Sharma CM, Gonzales K. 2011 CRISPR RNA maturation by trans -encoded small RNA and host factor RNase III. *Nature* **471**, 602–607. (doi:10.1038/nature09886.CRISPR)

83. Makarova KS, Brouns SJJ, Horvath P, Sas DF, Wolf YI. 2012 Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. ...* **9**, 467–477. (doi:10.1038/nrmicro2577.Evolution)

84. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012 A Programmable Dual-RNA – Guided. **337**, 816–822. (doi:10.1126/science.1225829)

85. Tanentzapf G, Devenport D, Godt D, Brown NH. 2007 Integrin-Dependent Anchoring of a Stem Cell Niche. *Nat Cell Biol* **9**, 1413–1418. (doi:10.1038/nature13579.Structural)

86. Iliakis G *et al.* 2004 Mechanisms of DNA double strand break repair and chromosome aberration formation. *Cytogenet. Genome Res.* **104**, 14–20. (doi:10.1159/000077461)

87. Smith et al. 2006. 2014 NIH Public Access. **153**, 910–918. (doi:10.1016/j.cell.2013.04.025.One-Step)

88. Tavares B *et al.* 2017 Notch/Her12 signalling modulates, motile/immotile cilia ratio downstream of *Foxj1a* in zebrafish left-right organizer. *Elife* **6**. (doi:10.7554/eLife.25165)

89. Howe K *et al.* 2013 The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498–503. (doi:10.1038/nature12111)

90. Becker-Heck A *et al.* 2011 The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. *Nat. Genet.* **43**, 79–84. (doi:10.1038/ng.727)

91. Sugrue KF, Zohn IE. 2017 Mechanism for generation of left isomerism in Ccdc40 mutant embryos. , 1–17. (doi:10.1371/journal.pone.0171180)

92. Place ES, Smith JC. 2017 Zebrafish atoh8 mutants do not recapitulate morpholino phenotypes. *PLoS One* **12**, 1–12. (doi:10.1371/journal.pone.0171143)

93. Kok FO *et al.* 2016 Reverse genetic screening reveals poor correlation between Morpholino-induced and mutant phenotypes in zebrafish F. *Dev Cell.* **32**, 97–108. (doi:10.1016/j.devcel.2014.11.018.Reverse)

94. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. 1995 Stages of embryonic development of the zebrafish. *Dev. Dyn. an Off. public* **203**, 253–310. (doi:10.1002/aja.1002030302)

95. Matthews M, Varga ZM. 2012 Anesthesia and euthanasia in zebrafish. *ILAR J.* **53**, 192–204.

(doi:10.1093/ilar.53.2.192)

96. Leary S, Underwood W, Anthony R, Cartner S. 2013 *AVMA Guidelines for the Euthanasia of Animals: 2013 Edition*. (doi:10.1016/B978-012088449-0.50009-1)

97. Zhang Y, Ge X, Yang F, Zhang L, Zheng J, Tan X, Jin Z, Qu J, Gu F. 2015 Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells. *Sci. Rep.* **4**, 5405. (doi:10.1038/srep05405)

98. Zhu X *et al.* 2014 An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. *Sci Rep* **4**, 6420. (doi:10.1038/srep06420)

99. Keßler M, Just S, Rottbauer W. 2012 Ion flux dependent and independent functions of ion channels in the vertebrate heart: Lessons learned from zebrafish. *Stem Cells Int.* **2012**, 1–9. (doi:10.1155/2012/462161)

100. Nandy D, Mukhopadhyay D. 2011 Growth Factor Mediated Signaling in Pancreatic Pathogenesis. (doi:10.3390/cancers3010841)

101. Field HA, Ober EA, Roeser T, Stainier DY. 2003 Formation of the digestive system in zebrafish. Liver morphogenesis. *Dev. Biol.* **253**, 279–290. (doi:10.1016/S0012-1606(02)00017-9)

102. Brocal I, White RJ, Dooley CM, Carruthers SN, Clark R, Hall A, Busch-Nentwich EM, Stemple DL, Kettleborough RNW. 2016 Efficient identification of CRISPR/Cas9-induced insertions/deletions by direct germline screening in zebrafish. *BMC Genomics* **17**, 259. (doi:10.1186/s12864-016-2563-z)

103. Gu J-Y, Xu J-H, Yu H, Yang Y-Q. 2012 Novel GATA5 loss-of-function mutations underlie familial atrial fibrillation. *Clinics* **67**, 1393–9. (doi:10.6061/clinics/2012(12)08)

104. Zhao C, Malicki J. 2007 Genetic defects of pronephric cilia in zebrafish. *Mech. Dev.* **124**, 605–616. (doi:10.1016/j.mod.2007.04.004)

105. Lewis BP, Green RE, Brenner SE. 2003 Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 189–92. (doi:10.1073/pnas.0136770100)

106. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE. 2007 Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* **446**, 926–929. (doi:10.1038/nature05676)

107. Suryavanshi S *et al.* 2014 NIH Public Access. **20**, 1–20. (doi:10.1158/2326-6066.CIR-13-0034.PD-L1)

108. Sullivan-Brown J, Schottenfeld J, Okabe N, Hostetter CL, Serluca FC, Thiberge SY, Burdine RD. 2008 Zebrafish mutations affecting cilia motility share similar cystic phenotypes and suggest a mechanism of cyst formation that differs from pkd2 morphants. *Dev. Biol.* **314**, 261–275. (doi:10.1016/j.ydbio.2007.11.025)

109. Grimes DT, Boswell CW, Morante NFC, Henkelman RM, Burdine RD, Ciruna B. 2016 Zebrafish models of idiopathic scoliosis link cerebrospinal fluid flow defects to spine curvature. *Science (80-.).* **352**, 1341–1344.

110. Boulter C, Mulroy S, Webb S, Fleming S, Brindle K, Sandford R. 2001 Cardiovascular, skeletal, and renal defects in mice with a targeted disruption of the Pkd1 gene. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12174–12179. (doi:10.1073/pnas.211191098)

111. Bordalo DM. 2015 Fate map of zebrafish left-right organizer. See https://run.unl.pt/handle/10362/15790 (accessed on 27 September 2017).

112. Schwank G *et al.* 2013 Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* **13**, 653–658. (doi:10.1016/j.stem.2013.11.002)

113. Blum M, Feistel K, Thumberger T, Schweickert A. 2014 The evolution and conservation of left-right patterning mechanisms. *Development* **141**, 1603–13. (doi:10.1242/dev.100560)

114. Lieber MR. 2011 The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End Joining Pathway. *Annu. Rev. Biochem.* **79**, 181–211. (doi:10.1146/annurev.biochem.052308.093131.The)

115. Brandsma I, Gent DC. 2012 Pathway choice in DNA double strand break repair: observations of a balancing act. *Genome Integr.* **3**, 9. (doi:10.1186/2041-9414-3-9)

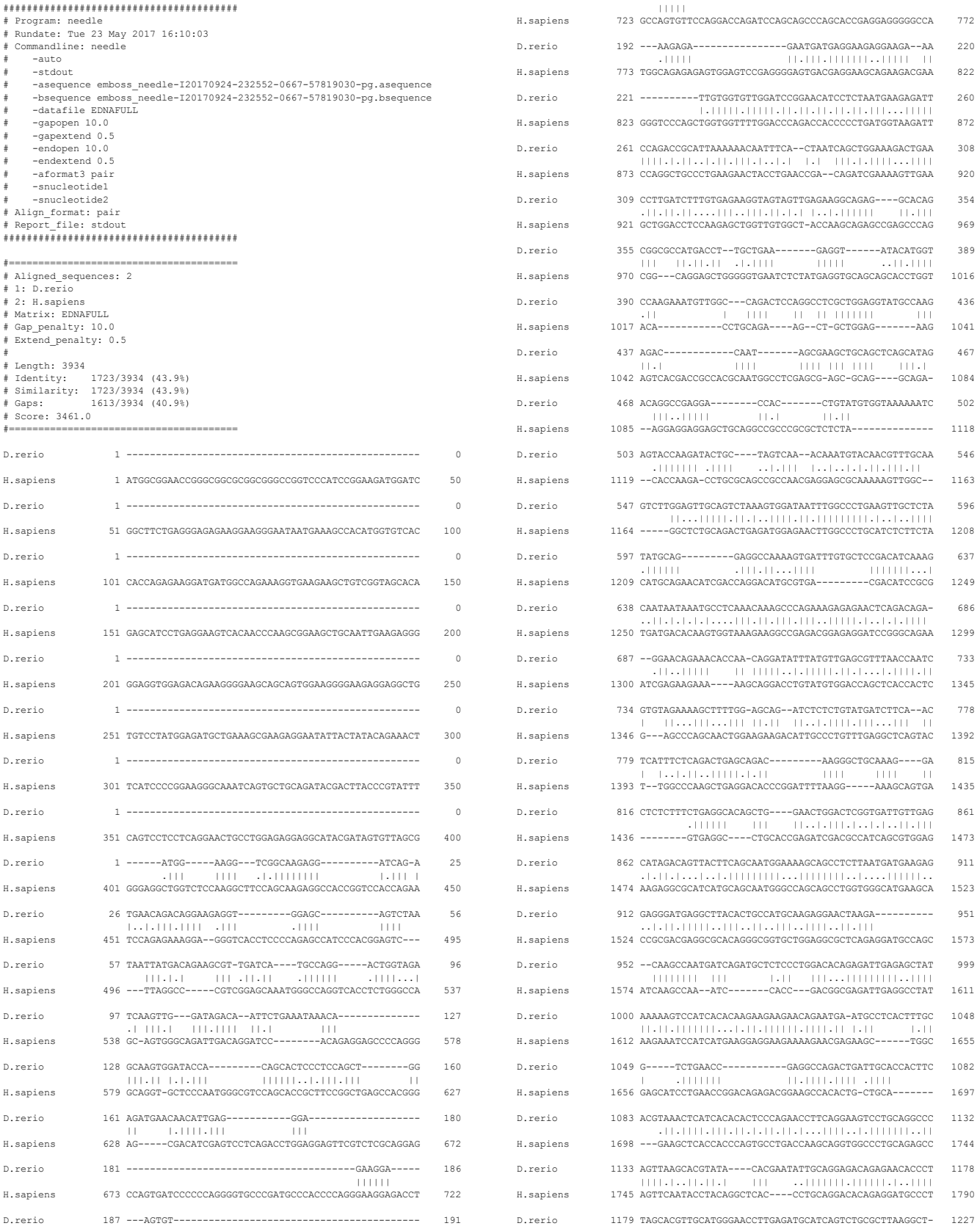
116. Dawe HR, Farr H, Gull K. 2006 Centriole/basal body morphogenesis and migration during ciliogenesis in animal cells. *J. Cell Sci.* **120**, 7–15. (doi:10.1242/jcs.03305)

117. Lin J, Okada K, Raytchev M, Smith MC, Nicastro D. 2014 Structural mechanism of the dynein powerstroke. *Pmc* **16**, 479–485. (doi:10.1038/ncb2939.Structural)

118. Pereira R, Sá R, Barros A, Sousa M. 2015 Major regulatory mechanisms involved in sperm motility. *Asian J. Androl.* **19**, 5–14. (doi:10.4103/1008-682X.167716)

7. Appendices

7.1. Human- zebrafish ccdc40 alignment



7.2. pDR274 structure

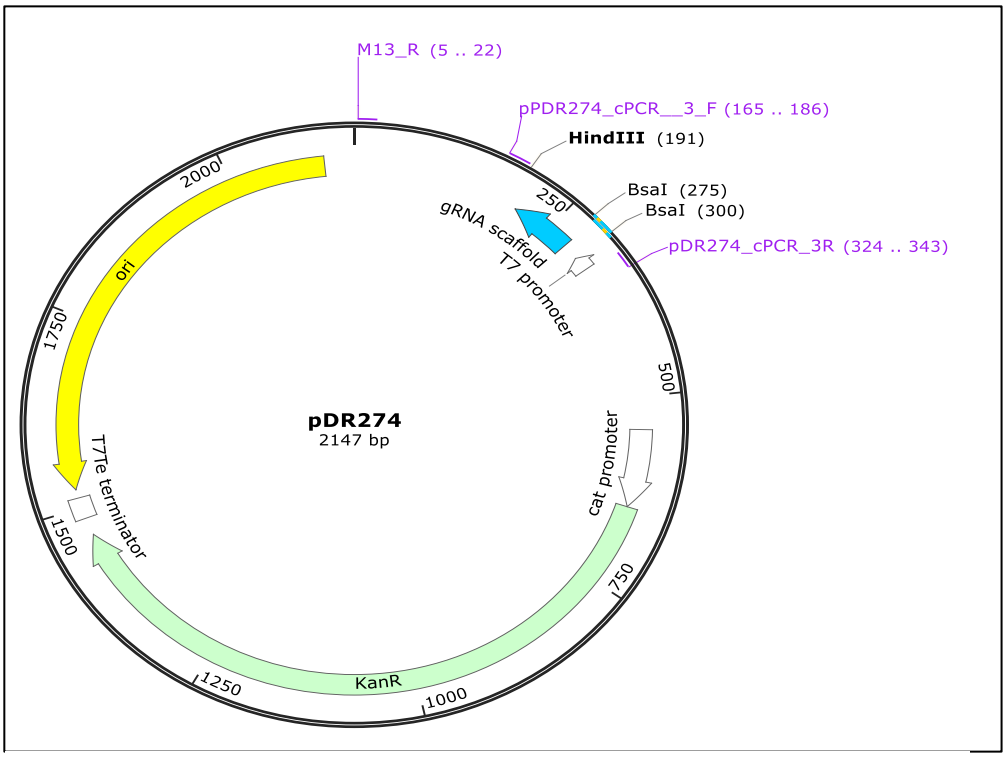


Figure 7.1 – pDR274 plasmid structure
pDR274 includes a replication origin (ori), a domain conferring kanamycin resistance (KanR), a T7 promoter and a gRNA scaffold. The gRNA scaffold is immediately downstream of the BsaI restriction cloning site and allows the interaction with Cas9.
It also shows a HindIII restriction site used to linearize the plasmid with the sgRNA template.
The primer M13_R was used to confirm the sgRNA template insertion in the plasmid by sanger sequencing. pDR274 _cPCR_3F and pDR274 _cPCR_3R were used as controls.

7.3. Plasmid sequencing results showing the detailed sequence

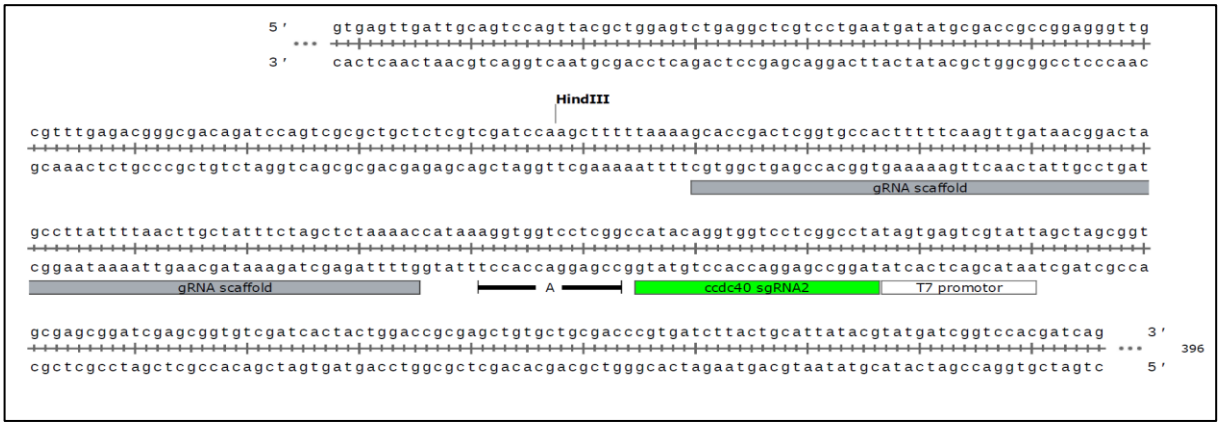


Figure 7.2– Sequencing results for colony 10 (sgRNA2)
The sgRNA2 template was inserted between the T7 promoter and the gRNA scaffold.
The sgRNA2 template appears to be partially duplicated as indicated by feature A.

Figure 7.3 – Sequencing results for colony 14 (sgRNA2)
The sgRNA1 template was inserted between the T7 promoter and the gRNA scaffold.

7.4. Crispants sequencing

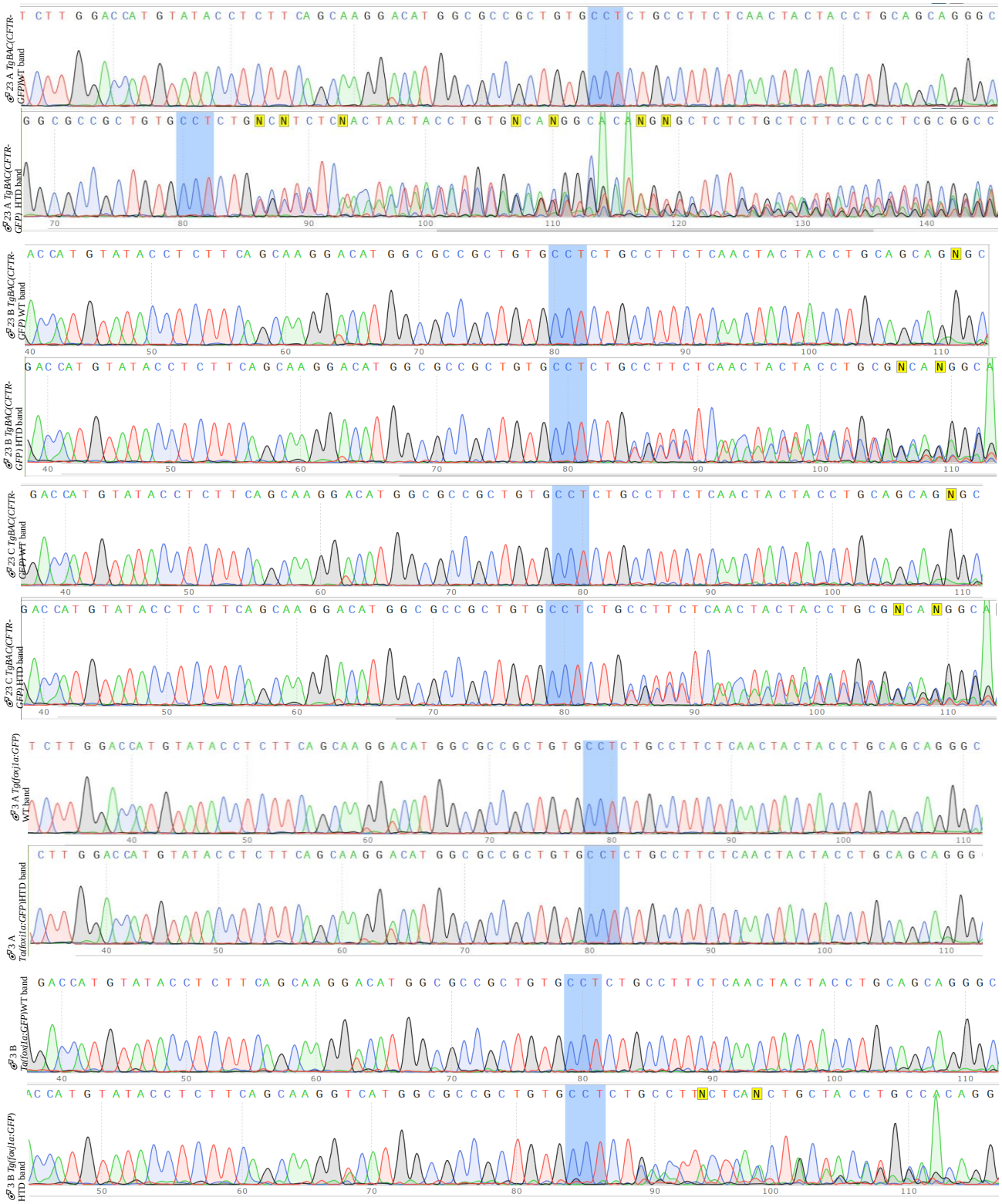


Figure 7.4 – Completed Sanger sequencing results for wildtype (WT) and heteroduplex (HTD) bands isolated from ♂ 23 *TgBAC(CFTR-GFP)* and ♂ 3 *Tg(foxj1a:GFP).*
The bands were isolated from the gel shown in Figure 3.8 .
This sequencing was done using a reverse primer, therefore the PAM sequence - 5'-CCT-3', highlight in blue is the reverse complement sequence of the canonical PAM sequence 5'-NGG-3'. All three HTD replicas from ♂ 23 *TgBAC(CFTR-GFP)* and replica B from and ♂ 3 *Tg(foxj1a:GFP).* showed overlapping in the chromatogram immediately after the PAM sequence 3' edge. The WT sequences and replica A from and ♂ 3 *Tg(foxj1a:GFP).* showed individualized peaks.

7.5. ccdc40 cloning primers

Table 7.1 – Primers tested to amplify zebrafish ccdc40 cDNA

	Sequence (5'->3')
Forward primer	GGGAAAGAGCTCATGGAAGGTCGGCAAGAG
Reverse primer	AGTACCCGGGCTGGCTGTTTGGACATTC