Investigating the Functional Roles of GWAS Hits Common to Osteoarthritis and Cardiovascular Disease in the Development of the Heart and Skeletal System in Zebrafish Models

Nagham Dous

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield

Faculty of Medicine, Dentistry and Health

Department of Infection, Immunity and Cardiovascular Disease

March 2023

Declaration

I, Nagham Dous, confirm that the thesis is my own work. I am aware of the University's Guidance on the Use of Unfair Means (www.sheffield.ac.uk/ssid/unfair-means). This work has not previously been presented for an award at this, or any other, university.

Acknowledgements

"The true measure of success is not what you achieve, but who you become in the process." -Anonymous

Regrettably, this body of work bears only my name, when in reality, it stands on the shoulders of many individuals who played a crucial role in shaping its path. Therefore, I would like to take this opportunity to formally acknowledge and convey my profound gratitude to these individuals.

First and foremost, I want to express my heartfelt gratitude to my primary supervisor, **Dr. Heba Ismail**. Your guidance extended far beyond the realm of science, offering invaluable life lessons. I am eternally thankful for your continuous support, the myriad opportunities you paved for me, and your integrity which sets an outstanding example for all in the scientific community. To my secondary supervisor, **Professor Mark Wilkinson**, I extend my appreciation for consistently offering fresh perspectives to my project and your continuous support throughout my PhD journey. **Dr. Emily Noel**, my third supervisor, your boundless enthusiasm for science is truly inspiring. The spark in your eyes whenever you delve into scientific discussions never failed to ignite my own passion. I must also express my appreciation to my funders, the Healthy Lifespan Institute, for granting me this invaluable opportunity that I hold in the highest regard.

To my dearest loved ones, my parents, **Taghrid Hanna** and **Nady Dous**, my sister **Shaza Dous**, and my cherished canine companion, **Taxi**, I am eternally grateful for your constant support. Without your presence and your unconditional love, I would not have made it to where I am today. You have been my unwavering pillars of strength, and without your constant encouragement during those challenging moments, this thesis would not have come together.

A special thanks goes to my personal tutor and mentor throughout my PhD, **Dr. Richard Mead**. I will dearly miss the opportunity to seek your opinion and insightful advice, which have been invaluable. I also want to acknowledge the support of **Dr. Lynne Prince**, which has been a cornerstone of my progress.

To those special individuals who entered my life and changed its course, I offer my heartfelt thanks. **Dr. Ehab Ezzat**, your inspiring genetics class in year 9 was the catalyst for my journey, instilling in me a passion for genetics and inspiring me to find my purpose in making a positive impact through genetics. To **Dr. Mona Sarkis**, your recognition of my passion for genetics boosted my self-belief. **Dr. Nabil Elias**, your truly inspiring life and your patience in explaining basic algebra during my A-levels were indispensable.

I would be remiss not to acknowledge the individuals who equipped me with the skills necessary to present this work. Thanks to the Aquarium team, particularly **Michael Thomas** and **Susanne Surfleet**,

for their assistance with the fish work. My gratitude extends to **Dr. Nick van Hateren** for his guidance in the microscopy training and to **Dr. Sarah Baxendale** for her assistance in the behavioural analysis training. I would also like to thank **Rachel Tucker** and **Lisa Hateren** for their help with the ABI analyzer and ZEG. I extend my sincere gratitude to **Dr. Amnael Orozco** for his help with the micro-CT scans and making those lengthy analysis sessions less of a pain.

My heartfelt appreciation goes to my dear friends, who are, in truth, the most significant achievements in my life. A profound thanks to **Veronia Salama**, for always being ready to lend an understanding ear and offering unwavering encouragement. Your friendship has been a precious gift throughout my life. To my friend **Omar Mokhtar**, your friendship has been a source of strength through challenging times. To the true lifesavers whom I am extremely thankful for - **Arwa Abugable**, **Salma Srour**, **Youssef Maharem**, **Moataz Swidan**, **Ola Shehata**, **Nayera Ahmed**, and **Mohamed ElGhazaly**.

To my Firth Court companions who were there when I needed them most, **Salma**, **Mohamed**, and **Ola**, your support meant the world to me. To **Moataz** and **Ahmed ElSawaf**, who were always welcoming and got me through truly challenging times. To some friends who made my time in Sheffield extraordinarily special, **Nada Adham**, **Hend ElGhazaly**, **Jumana Osama**, **Nada Ghorab**, and **Ahmad ElBadry**, your friendship is a treasure. To **Robin Turner**, **Lloyd Harman**, and **Sahil Unkule**, who made my final year so much more bearable and supported me through my worst days, I am grateful.

Finally, I would like to thank my lab mates who added joy to each day, **Matina Christaki**, **Hamzah Huntul**, **Georgios Kalogiannis**, and **Nitchakarn Kaokhum**.

Table of Contents

List of Figures

List of Tables

Abbreviations

AA	Aortic Aneurysm
AD	Alzheimer's Disease
AE	Acrodermatitis Enteropathica
ALK1	Activin A Receptor Like Kinase 1
ANG	Angiogenin
BCIP	5-bromo-4-chloro-3’indolylphosphate
BMP	Bone Morphogenetic Proteins
BMPR1B	Bone Morphogenetic Protein Receptor Type-1B
CAMK2β	Calcium/Calmodulin-dependent Protein Kinase Type II Subunit Beta
CDMP	Cartilage-derived Morphogenetic Protein
CHD	Coronary Heart Disease
CHF	Chronic Heart Failure
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTSK	Cathepsin-K
CVD	Cardiovascular Disease
DMM	Destabilization of the Medial Meniscus
DPEP1	Dipeptidase 1
DPF	Day Post-fertilisation
ECM	Extracellular Matrix
ER	Endoplasmic Reticulum
ERK	Extracellular Signal-regulated Kinase
GAGs	Glycosaminoglycans
GDF5	Growth Differentiation Factor 5
GWAS	Genome Wide Association Studies
HDAC9	Histone Deacetylase 9

HHT	Hereditary Haemorrhagic Telangiectasia
IBD	Inflammatory Bowel Disease
IFN-γ	Interferon Gamma
JAK	Janus Kinase
JNK	c-Jun N-terminal kinase
KL	Kellgren and Lawrence Classification
LAP	Latency Associated Protein
M-6-P	Mannose-6-Phosphate
MAP	Microtubule Associated Protein
MAPK	Mitogen Activated Protein Kinase
MAPT	Microtubule Associated Protein Tau
MI	Myocardial Infarction
MII	Mucolipidosis II
MMPs	Matrix Metalloproteinases
mRNA	Messenger RNA
NFATc1	Nuclear Factor of Activated T-cells
NF-κB	Nuclear Factor Kappa-light-chain-enhancer of Activated B cells
NOG	Noggin
OA	Osteoarthritis
PAM	Protospacer Adjacent Motif
PBST	Phosphate-buffered Saline with Tween
PCR	Polymerase Chain Reaction
PD	Periodontal Disease
PFA	Paraformaldehyde
PI3K	Phosphoinositide 3-kinase
PIAS	Protein Inhibitors of Activated STAT

PTA	Phosphotungstic Acid
PTPs	Protein Tyrosine Phosphatases
PVDF	Polyvinylidene Fluoride
qPCR	Quantitative Polymerase Chain Reaction
RANK	Receptor Activator of Nuclear Factor Kappa-B
RANKL	Receptor Activator of Nuclear Factor Kappa-B Ligand
SAP	Shrimp Alkaline Phosphatase
SNP	Single Nucleotide Polymorphism
SOCS	Suppressors of Cytokine Signalling
STAT	Signal Transducer and Activator of Transcription
TAK1	Transforming Growth Factor-β (TGF-β)-activated Kinase 1
TB	Tuberculosis
TGFβ1	Transforming Growth Factor Beta-1
TNF-α	Tumour Necrosis Factor Alpha
VEGF	Vascular Endothelial Growth Factor

Abstract

Multimorbidity is the condition of living with two or more chronic diseases. A recent meta-analysis suggested a 24% increased risk of developing cardiovascular disease (CVD) in patients with osteoarthritis (OA). Therefore, a possible genetic link is hypothesised, underlying both diseases, and potentially other chronic diseases, causing multimorbidity. A bioinformatics analysis pipeline was designed to identify common risk genes between OA and CVD, as well as attempt to reveal common pathways underlying multimorbidity pathogenesis. Based on the conducted analysis, 37 genes were identified common between OA and CVD, using 133 OA and 2884 CVD Genome Wide Association Studies (GWAS). Cathepsin K (*CTSK*), Transforming Growth Factor Beta 1 (*TGFβ1*), and Growth Differentiation Factor 5 (*GDF5*) were selected for further functional genomic analysis based on data suggesting direct linkage between the three genes, and due to their involvement in additional chronic diseases.

CRISPR-Cas9 genome guide targets were designed to target each gene in zebrafish embryos, aiming to investigate their functional role in the musculoskeletal system and heart morphology. Additionally, the role of these genes in zebrafish locomotion was monitored to detect the targeting impact on the functionality of the skeletal system. Finally, micro-CT scans are utilised to examine the targeting effect on later skeletal development. Targeting of *ctsk* resulted in significant disruption to early cartilage and heart development, starting at 3 days post fertilisation (dpf). A similar phenotype was observed when *tgfβ1a* was targeted with milder impact on both structures, suggesting the early involvement of both genes in cartilage and heart development. By examining the efficiency of crispants (CRISPR edited zebrafish) locomotion, no significant alteration to the activity levels of crispants was detected at the larval stage, as well as in adult fish when compared to wildtype controls. However, 12-month-old adult *ctsk* crispants demonstrated a significant increase in bone surface area solely among male adults, confirming the resorptive function of Ctsk and implying a distinct gender-specific role. Conversely, targeting of *gdf5* did not display any alterations to the cartilage and heart morphology at 5dpf. However, by examining the bone structure of adult *gdf5* crispants, a significant increase in bone mineral density and surface area was detected in the absence of *gdf5*. The swimming abilities of *gdf5* crispants, however, did not exhibit any changes. Taken together, this suggests a role of *gdf5* in later bone maintenance through attenuation of bone tissue growth.

Following the bioinformatics analysis, a direct crosstalk between Ctsk and Tgfβ1 was hypothesised. To investigate the validity of this hypothesis, *ctsk* and *tgfβ1a* were co-targeted within the same embryo. Results indicate a possible recovery to the general morphology of co-targeted larvae when compared to separate gene targeted larvae. However, a differential impact of co-targeting was observed in regard to the heart morphology, highlighting the different gene function within different tissues. Using

quantitative PCR analysis, gene expression levels of *ctsk* and *tgfβ1a* were quantified in generated gdf5 F2 mutants. Data displayed a consistent reduction in the expression of both genes in gdf5 F2 mutants, suggesting a direct correlation between the three studied genes that function simultaneously through SMAD-dependent and SMAD-independent signalling to attain bone homeostasis. This allows disruptions in one gene to affect other correlated genes.

To summarise, *ctsk* and *tgfβ1a* were found essential in the early development of both the musculoskeletal and cardiovascular systems, while *gdf5* is not involved in the early development. Yet, its targeting affects the bone mineral density and surface area in adult zebrafish. Targeting one of the tested common risk genes may have a direct effect on the expression level of the other examined genes, emphasizing the intricacy of genetic pathogenesis contributing towards multimorbidity.

1. Introduction

1.1. Introduction to multimorbidity

Multimorbidity is the condition of living with two or more chronic diseases. It is estimated to affect 6 in 10 people between the ages of 65 and 84 years old. A dramatic increase is observed in older adults, affecting 8 in 10 people over the age of 85 (https://www.versusarthritis.org). Several factors have been identified in the pathogenesis of multimorbidity. The socio-economic status of an individual has been found to be one of the main factors contributing to the increase in multimorbidity (Kone et al., 2021). In fact, multimorbidity tends to start a decade earlier in underprivileged communities with poor living conditions and compromised quality of life (McPhail et al., 2016). Subsequently, multimorbidity becomes a direct contributor to the early death observed in less privileged communities. In the UK, 50% of people over the age of 65 suffer from multimorbidity. Treating each individual requires an average of £8,000 per year to cover healthcare costs (Sheridan et al., 2019), highlighting the financial burden of multimorbidity in addition to the health burden. A recent study of multimorbidity in Ontario, Canada, found a 29% increase in the prevalence of multimorbidity between 2003 and 2016, emphasising the need for a treatment to address the growing prevalence of this condition. The most common clusters of co-occurring chronic diseases are hypertension and osteoarthritis, followed by asthma and osteoarthritis (Kone at el., 2021).

Several biological mechanisms underlying multimorbidity have been linked to ageing and inflammation (Lopez-Otin et al., 2013). Some of the hallmarks of ageing associated with the progression of multimorbidity include epigenetic defects, telomere shortening, genomic mutations, inappropriate protein folding or processing, cellular senescence, stem cell exhaustion, alterations in intracellular interactions and mitochondrial dysfunction (Wetterling et al., 2021; Singer et al., 2019; Kennedy et al., 2014; Barnes et al., 2015).

Polypharmacy is the current approach to treating multimorbidity, which means that patients are given multiple drugs to treat each identified condition separately. The side effects of polypharmacy are numerous and unpleasant. In addition, polypharmacy may directly contribute to the development of multimorbidity, as some drugs administered to attenuate the pathogenic effects of one disease may induce the progression of another disease. This has already been observed with antipsychotic drugs, which have been found to increase the risk of diabetes (Nielson et al., 2010). In addition, anticholinergic drugs have been found to increase the risk of developing a cardiovascular disease or lead to cognitive impairment (Hanlon et al., 2020). Another example is the co-administration of non-steroidal anti-inflammatory drugs (NSAIDs) along with selective serotonin reuptake inhibitors (SSRIs), used to treat

depression, which leads to gastrointestinal bleeding (Dalton et al., 2003). This illustrates how an attempt to treat a disease chemically without considering its effect on other systems can lead to the development or progression of multimorbidity. Identifying a possible common cause linking these long-term conditions would therefore be the next step in making progress in this area.

In the UK, studies have found that 50% of patients with diabetes also have hypertension (Petrie et al., 2018) and a third of people with dementia suffer from depression (Muliyala and Varghese, 2010). Several studies have found an increased risk of developing cardiovascular disease development in people with osteoarthritis. Ong et al., 2013 observed a 1.53 OR (odds ratio) risk of cardiovascular disease occurrence in osteoarthritis patients when compared to healthy controls. A similar pattern was observed in a separate study of women with hip osteoarthritis, which showed a hazard ratio of 1.24, resulting in a 26.3% mortality from acquired cardiovascular disease among 7889 hip osteoarthritis patients (Barbour et al., 2015). Another study of 3099 sample size also showed a 1.22 hazard ratio in cardiovascular disease progression in osteoarthritis patients (Veronese et al., 2016). Hence, studying the affected key regulatory genes and their functions in relation to multiple systems may facilitate the identification of fundamental pathways underlying the development of multimorbidity. This project will specifically focus on the link between osteoarthritis and cardiovascular disease.

1.2. Introduction to osteoarthritis

Osteoarthritis (OA) is the main cause of disability worldwide and the most common musculoskeletal disease. It affects around 10 million people in the UK alone and it is estimated that 80% of the total population will start to show some developmental signs of OA as seen in radiography scans (Heidari, 2011). The number of OA patients has been increasing across the world, reaching 528 million patients in 2019, which signifies a 113% increase since 1990 (GBD, 2020). Knee osteoarthritis is the most common acquired form of osteoarthritis, followed by the hand and the hip (Long et al., 2022). The prevalence of the disease increases with age, with 73% of diagnosed osteoarthritis patients over the age of 55 (GBD, 2020). In addition to the devastating impact of OA on the individual quality of life, there is also a heavy financial burden, with treatment of OA costing more than $60 billion annually in the USA (Buckwalter et al., 2004).

1.2.1. Characterisation of OA

The disease is characterised by the degradation of cartilage and surrounding bone, as well as structural changes in the extracellular matrix (ECM) of the affected cartilage (Maldonado and Nam, 2013). By examining OA cartilage structures, some features have been observed in OA cases. These include thinning of the cartilage within the joints, which can lead to defects in joint function (DeGroot et al., 1999). Cartilage degeneration causes friction within the joint and damage to the meniscus, which can

lead to further thinning of the cartilage, creating a rougher joint surface (Pollard et al., 2008). This results in reduced mobility or limited range of motion in the joint, and in some cases, pain is experienced. Other phenotypes observed in the degrading cartilage of osteoarthritis include thickening of the joint capsule and degeneration of the underlying bones (Sharma et al., 1999). In some severe cases, inflammation of the synovium is observed, leading to thickening of the cartilage as part of the repair process, resulting in the formation of small projections known as osteophytes (Felson et al., 2005). **Figure 1.1** highlights the abnormalities observed in an osteoarthritic knee compared to a healthy knee.

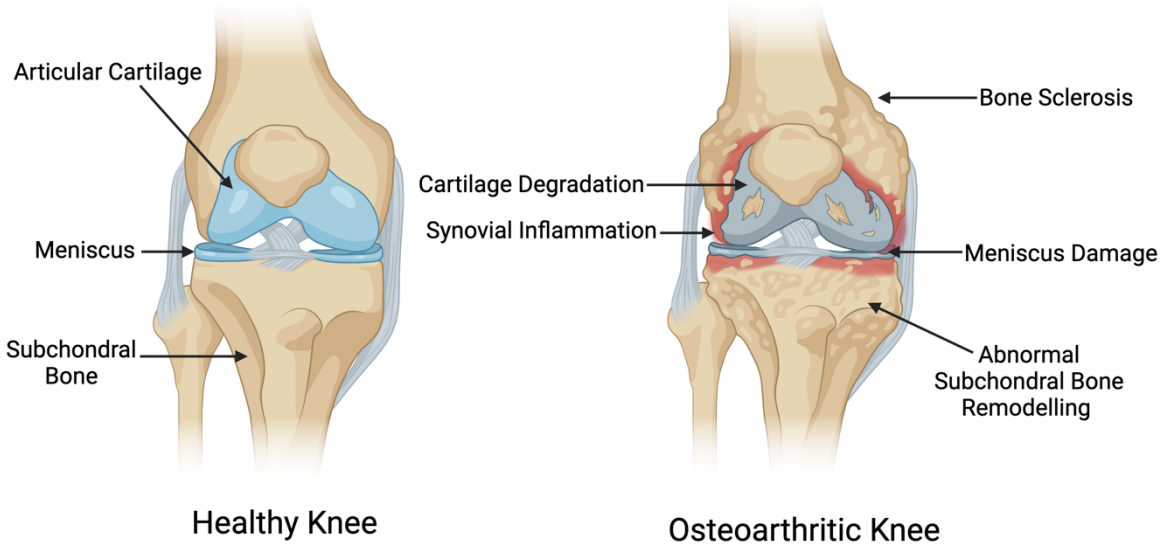


Figure 1.1: Knee morphology in osteoarthritis

Left side shows a representation of a healthy knee with intact articular cartilage, stable meniscus and subchondral bone. The right side shows an osteoarthritic knee with evidence of cartilage degradation, inflammation surrounding the cartilage within the synovial capsule, along with damage and thinning of the cushioning meniscal layer beneath the cartilage. The subchondral bone is often affected and shows abnormal remodelling along with general bone sclerosis. Figure generated using Biorender (https://www.biorender.com/).

1.2.2. Mechanical inducers of OA

OA was once considered a simple wear and tear condition, concluded from the late-onset occurrence and the increase in its prevalence in older adults, which is further exacerbated by the increase in life expectancy. Different factors contribute to the disease's pathogenesis such as weight gain, weakened muscles, and mechanical injuries induced during an exercise or sporadically (King et al., 2013; Alnahdi et al., 2012; Buckwalter et al., 2013). Due to the severe pain experienced by OA patients during movement, they usually acquire a sedentary lifestyle which leads to the development of other serious diseases, including hypertension, depression, and cardiovascular diseases (Skou et al., 2018).

1.2.3. Molecular mechanisms behind OA progression

Regulated bone remodelling is essential for the formation and maintenance of healthy bones. The process involves the breakdown of damaged bone tissue, and its replacement with new bone material. This process is controlled by a balance of activity between bone-building cells, known as Osteoblasts, and bone-resorbing cells, called Osteoclasts. **Figure 1.2** shows the different types of bone cells and their function in bone formation and maintenance. One of the main factors contributing to the progression of OA is an imbalance in the activity levels of these critical bone cells, leading to a disruption in the process of bone homeostasis. Research has provided clear evidence of a strong genetic basis, and in some cases even an inheritance pattern for the disease (Fernandez-Moreno et al., 2008; Aubourg et al., 2022). An important example is the TGF-β signalling pathway, which has been identified as a key regulator in the musculoskeletal system (Massague et al., 2000). Increased levels of TGFβ proteins expression have been found in OA patients (Kraan, 2018). In addition, mutations within *TGFβ* genes have been shown to exhibit pathogenic effects on bone homeostasis, leading to OA progression. This includes a T^{869}-C polymorphism mutation detected in *TGFβ1* that has been associated with spinal osteoarthritis (Yamada, 2000). Three other SNPs in *TGFβ1* (C^{1348}-T, T^{29}-C, and $T^{861\text{-}20}$-C) have been detected in southern Chinese women with symptoms of osteoarthritis (Lau et al., 2004). The protein has also been found implicated in osteophyte formation through its anabolic role in bone homeostasis (Davidson, 2007). In the TGF-β signalling pathway, the role of Suppressor of Mothers against Decapentaplegic (SMADs), a group of proteins involved in signal transduction, is significant. One important SMAD protein involved in transmitting key regulatory signals is SMAD3. Mutations within *SMAD3* were also found to induce OA progression by altering the TGF-β signalling, either by increasing or decreasing its expression (Muratovic et al., 2022; Yang et al., 2001). *Smad3* knockout mice show progressive degradation of articular cartilage, mimicking the phenotype observed in humans (Yang et al., 2001). However, *TGFβ* has a variable expression pattern in aged individuals, which has also been observed in osteoarthritic mouse models. In murine chondrocytes, TGFβ was found to signal via the ALK1-SMAD1/5/8 complex, instead of the ALK5-SMAD2/3 complex, as normally observed.

This alteration leads to an activation of catabolic functions resulting in an increased risk of OA progression in older adults and osteoarthritic mice (Davidson, 2007; van der Kraan et al., 2012).

Other key contributors in the progression of OA are certain types of fibroblast growth factors, including FGF2 and FGF18. FGF2 exhibits catabolic and anti-anabolic functions in mouse dental pulp cells (Li et al., 2012), which is balanced by the function of FGF18 as a cartilage growth factor that is highly involved in cartilage maturation as studied in human articular cartilage (Ellman et al., 2008). This suggests a critical role in cartilage homeostasis, highlighting the potential of these proteins as promising drug targets for osteoarthritis. Other identified molecular regulators are the HIF proteins. HIF1a, like FGF18, has an anabolic function (Pfander et al., 2004) and functions in a homeostatic manner with HIF2a, which is characterised as a catabolic protein (Saito et al., 2010). HIF2A also directs the expression of *Mmp13*, which targets collagen and induces degradation mainly in the connective tissues (Vincenti et al., 1998; Mengshol et al., 2001). Another important player in the regulation of the musculoskeletal system is GDF5. GDF5 plays a key role in the maturation of cartilage into bones through a process known as endochondral ossification. It is also involved in synovial joint formation (**Figure 1.1**). GDF5 has been implicated in anabolic homeostasis and reduction in its expression has been shown to contribute to osteoarthritis development (Mikic, 2004). *GDF5* gene will be further discussed in **Chapter 5**.

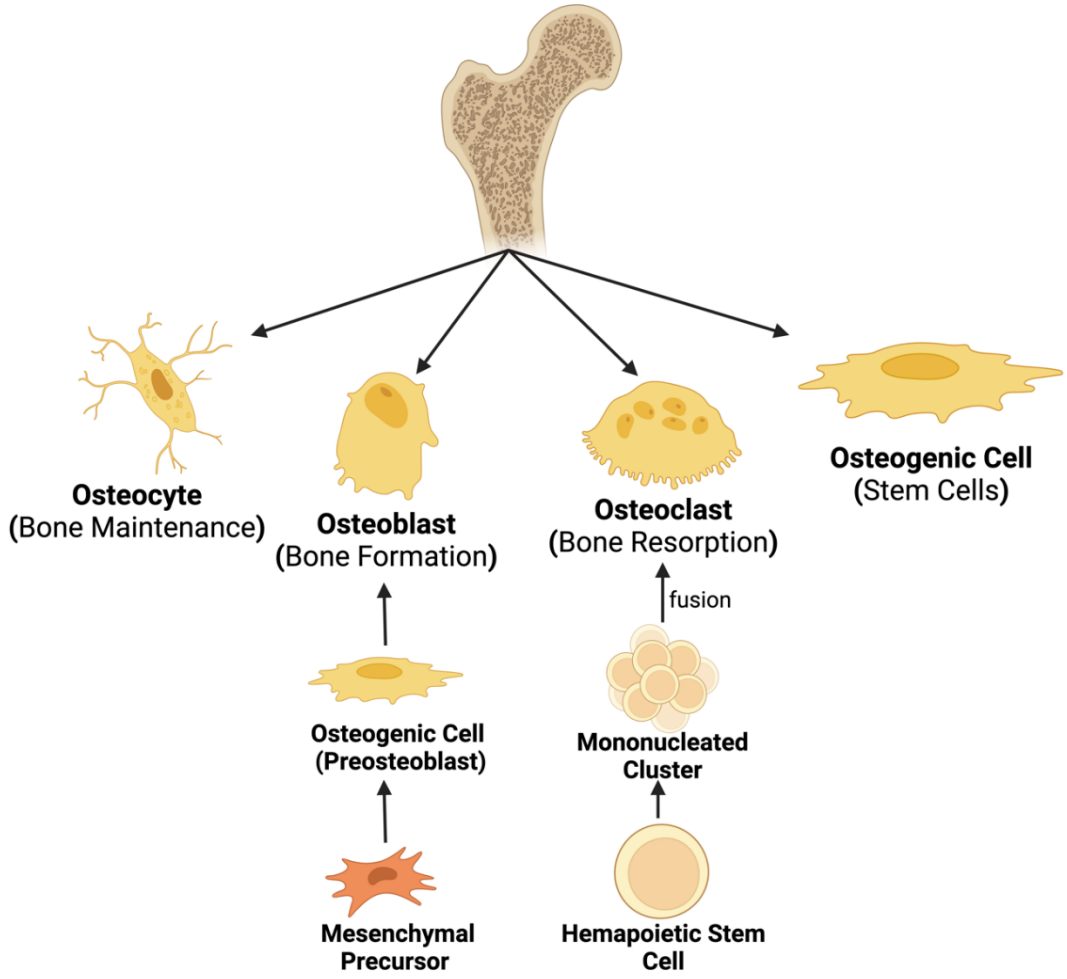


Figure 1.2: Types of bone cells

Bones are made up of four types of cells. Osteocytes are present in all mineralised bone and are responsible for maintaining bone function and health. Osteoblasts are the building blocks of bone tissue and are formed as single nucleated cells from mesenchymal precursor cells. Osteoclasts, on the other hand, are the breaking blocks in bone tissues, essential for bone remodelling and resorption. They are derived from haemopoietic stem cells in a mononucleated form. Several mononucleated pre-osteoclast cells then coalesce to form a mature multinucleated osteoclast. Osteogenic stem cells, also known as, osteoprogenitor cells, are the undifferentiated cells with the potential to develop into different types of bone-related cells. Figure generated using Biorender (https://www.biorender.com/).

1.2.4. Available OA treatments

There is currently no cure for OA, so patients rely mainly on painkillers on a regular basis, along with gentle exercise, supportive devices and, in severe cases, arthroplasty (joint replacement surgery). The drugs administrated are NSAIDs and paracetamol, which work by attenuating excessive inflammation (Costa et al., 2017). Importantly, excessive inflammation over a prolonged period leads to further degradation of cartilage tissue (Sokolove et al., 2013). Current treatments used to attenuate the progression of osteoarthritis include steroid injections (Arroll et al., 2004), electrical nerve stimulation (Osiri et al., 2000), and hyaluronic acid injections (Migliore et al., 2015), which have been shown to prevent further cartilage degradation by stabilising the affected cartilage. Alternative treatments that may help improve the symptoms of OA include non-strenuous aerobic exercise (Hunter et al., 2008), acupuncture (Selfe et al., 2008), glucosamine (Reginster et al., 2001) and chondroitin sulphate (Henrotin et al., 2010).

1.3. Introduction to cardiovascular disease

Cardiovascular disease (CVD) includes any condition that affects the heart or blood vessels, such as aortic disease, congestive heart failure, or coronary heart disease. According to the World Health Organisation (WHO), CVD affects one third of the population and is the leading cause of death worldwide (WHO, 2021). It is considered a complex disease, involving many genes and other environmental factors such as obesity, diabetes, hypertension, physical inactivity and smoking. A large cardiovascular study known as the INTERHEART trial, showed that increasing the intake of vegetables and fruit, exercising and avoiding smoking can reduce the risk of CVD by 80% (Yusuf et al. 2004). However, as with OA, a genetic basis has also been found, accounting for between 10.3% and 39.4% of CVD pathogenesis (Osztovits et al., 2011).

1.3.1. Molecular mechanisms behind CVD progression

Identified mutations associated with cardiovascular disease have been linked to ECM regulation. One example is illustrated by the overexpression of TGFβ1, which leads to the progression of coronary artery disease (CAD) characterised by plaque build-up in the arterial wall due to dysregulation of ECM degradation (Tahsiro et al., 2002). Atherosclerosis is another CVD that has been identified as a major contributor to cardiovascular mortality worldwide (Rafieian-Kopaei, 2014). Similar to CAD, recent studies have demonstrated a clear association between ECM dysregulation and the progression of atherosclerosis (Gialeli et al., 2021). Moreover, increased levels of TGFβ1 in serum were observed among patients suffering from atherosclerosis (Gomez-Bernal et al., 2023).

Other identified CVD biomarkers genes include *NOTCH1*, which causes aortic valve disease due to a point mutation, resulting in a premature stop codon (Garg et al., 2005). *PCSK9,* identified in familial hypercholesterolaemia with two identified non-synonymous mutations (Abifadel et al., 2003), and a point mutation in the *MYH7* gene, causing hypertrophic cardiomyopathy (Rose et al., 2020).

1.4. Introduction to Genome Wide Association Studies

Genome Wide Association Study (GWAS) is an observational study in which millions of individuals are tested for their genetic variants by undertaking various types of genome sequencing analysis using microarrays and next-generation sequencing to detect Single Nucleotide Polymorphisms (SNPs), indels, and Copy Number Variations (CNVs). GWAS aim to identify genetic biomarkers associated with the diseases being studied. First published in the year 2005, more than 50,000 significant associations between genetic variants and human traits and diseases have been found (Visscher et al., 2012; Visscher et al., 2017). While GWAS primarily focus on SNPs, it can also be applied to other genomic variants such as CNVs or sequence variations (Siminovitch et al., 2004). These associations have shed light on the biology of certain diseases, their likelihood of occurrence, as well as possible treatments. One example would be the discovery of the IL-12/IL-23 pathway as a contributor to the onset of Crohn's disease (Wang et al., 2009), which has led to several clinical trials of potential drugs targeting this pathway (Moschen et al., 2019).

1.4.1. Experimental workflow

There are several steps required to conduct GWAS, including data collection, genotyping, quality control, association testing, and post-association analysis. This section provides an overview of the end-to-end process of the GWAS experimental workflow.

The first process in a GWAS is data collection, where genetic and phenotypic information is collected from study cohorts or from biobanks and repositories. Large human sample sizes are required in order to ensure the reproducibility of the associations found. Considerable care is required as the characteristics of the study cohort must be carefully selected based on the aim of the study. Using the data collected, genotyping of blood samples or tissue biopsies is performed using microarrays, targeted SNPs, or next-generation sequencing. Microarrays are most commonly used due to their cost-effectiveness. However, it is likely that Whole Genome Sequencing (WGS) will become the preferred method as the cost of WGS technologies decrease over the years.

Quality control is then performed to remove ambiguous positions and individuals, for example if the genotyped sex differs from the known sex. It is also used to remove bias when the individuals being tested have a high degree of relatedness. The importance of quality control is to reduce noise in the data, as well as to reduce the false positive discovery rate (Anderson et al., 2010; Laurie et al. 2010; Marees et al., 2018). Using this high-quality dataset, association testing is performed using biometric techniques to detect genotype-phenotype associations. The most commonly used techniques are linear regression models for continuous data or logistic regression models for binary data (presence or absence of disease) (Pirinen et al., 2012). Different p-value thresholds are set for different populations to minimize the risk of false positives (Bakker et al., 2005).

Finally, other post-association tests are performed, such as replication and meta-analysis. Replication is the process of repeating the performed tests in multiple independent cohorts in order to verify the results, while meta-analysis is the process of using standardized statistical pipelines to combine results from multiple cohorts of smaller size.

Over the past 15 years, genomic risk loci have been identified for multiple diseases and traits including anorexia nervosa (Duncan et al., 2017), major depressive disorder (Hyde et al., 2016), subtypes of cancer (Milne et al., 2017; Sud et al., 2017), type 2 diabetes (Zhao et al., 2017), coronary artery disease (Nikpay et al., 2015), schizophrenia (Li et al., 2017), inflammatory bowel disease (de Lange et al., 2017), insomnia (Jansen et al., 2019), and educational attainment (Lee et al., 2018), among others. This abundance of genotype-phenotype associations was unprecedented before the advent GWAS (Lohmueller et al., 2003).

GWAS risk loci often include multiple genes whose function or relevance was previously unknown (Hirschhorn, 2009). Experimental investigation of these loci may lead to the discovery of new biological mechanisms underlying a particular disease. One prominent example is the discovery of the role of autophagy in Crohn's disease after studying the SNPs associated with the genomic risk loci in GWAS, including the genes *ATG16L1* (rs2241880) ch2:233,210,051-233,295,674, and *IRGM* (rs1000113) chr5:150,846,521-150,902,402 (Hampe et al, 2007).

1.5. Possible links between osteoarthritis and cardiovascular disease

Several studies on OA and CVD, have identified many factors that are common to both conditions and may lead to a general link underlying both pathological mechanisms. Some of these studies suggest that CVD is particularly high in patients with late-stage hip or knee OA (Nuesch et al. 2011; Hawker et al. 2014). This correlation may happen as patients with late-stage hip or knee OA are usually older adults, and old patients have a higher risk of CVD. However, diverse findings have been reported from various studies. As a result, a comprehensive meta-analysis was conducted in 2016 to consolidate all the data

from these studies. Out of 19 studies, 11 suggested that OA patients had a significant increased risk of CVD, five others showed no significant difference in the development of CVD, and the remaining three studies suggested that OA patients had a lower risk of CVD. Taking into account various factors, including age association, the meta-analysis results proved that the risk of CVD is usually 24% higher in OA patients compared to non-OA patients (Wang et al., 2016).

Common factors found to be associated with both conditions include: First, physical inactivity as a result of the severe pain that OA patients experience with simple movement, which limit the activity levels of OA patients, resulting in increased risk of CVD. Second, the regular use of non-steroidal anti-inflammatory drugs (NSAIDs) to suppress joint pain, as there are currently no drugs available for OA. NSAIDs are known to have adverse effects on the cardiovascular system, which can lead to myocardial infarction and hypertension. Thirdly, muscle weakness has been observed in patients with osteoarthritis (OA), which can limit their activity and potentially impact the heart as a muscle-dominant organ. Fourthly, the observed tissue perfusion or ischemia, which may affect the bones by reducing the amount of nutrients supplied to the cartilage, contributing to the progression of OA. And fifth, inflammation, which is associated with both diseases (Rahman et al., 2013).

Another suggested link between OA and CVD is the presence of a metabolic syndrome, as it has been observed in 60% of OA patients compared to only 23% in the rest of the population (Puenpatom and Victor, 2009; Zhuo et al., 2012). Signs of metabolic syndrome include elevated blood pressure, blood glucose level, blood cholesterol and waist circumference. The syndrome is often associated with obesity and an increase in adipokines and pro-inflammatory cytokines, which activate cartilage cells or chondrocytes. When chondrocytes are over-activated, they produce matrix metalloproteinases (MMPs) and pro-inflammatory cytokines, which can dramatically increase the likelihood or severity of both OA and CVD (Rocher et al., 2019). Another metabolic factor that can affect chondrocytes and increase MMPs production is leptin. Leptin is a hormone produced by fat cells that plays a role in appetite and energy regulation. It has been shown to promote collagen degradation leading to cartilage loss (Ding et al., 2007). Upregulated levels of leptin have also been linked to myocardial infarction (Kain et al., 2018). Finally, leptin could be implicated as a common link between metabolic syndrome, OA, CVD, and obesity. Adiponectin is another hormone that is abundantly present in adipose tissue and is found reduced in metabolic syndrome (Padamalayam and Suto, 2013). However, insufficient information is currently available to understand the exact role of adiponectin, as it displays both pro-inflammatory and anti-inflammatory functions. One explanation is the existence of different adiponectin isoforms with different properties, such as molecular weight and globularity. Although the exact function is unclear, the general data suggest an involvement in the meta-inflammation that affects OA and CVD (Berenbaum et al., 2013).

Recognising that general inflammation and adipokines caused by obesity are major factor in several chronic diseases, and that obesity is an easily modifiable factor, a study was conducted to determine the effect of weight loss on the level of inflammation and general improvement in wellbeing. The results showed a 28% improvement in physical wellbeing following a 10% reduction in body weight. A significant decrease in low-grade systemic inflammation was also detected (Woldbaek et al., 2005). A recent study on the effect of weight loss on osteoarthritis, confirmed a significant >25% improvement in knee osteoarthritis pain, and an overall improvement in health-related quality of life following a 20% weight loss (Messier et al., 2018). However, the joint structure should be examined to determine whether any changes occur with weight loss.

1.6. Zebrafish as a model organism for functional analysis of GWAS hits

Zebrafish are a great tool for genetic studies, with a number of advantages over in-vitro cellular work and mouse models. First, zebrafish have a genome that is 70% similar to that of humans (Howe et al., 2013). In addition, 84% of the genes found to be associated with human disease have been found in zebrafish (Kettleborough et al., 2013). The small size of zebrafish, not exceeding 4 cm, and the large number of eggs laid by females, about 100-300 embryos per week, make them easy to maintain in designed aquariums. The embryos are fertilised and develop externally which permits simple examination as they are physically transparent. Zebrafish possess a fast rate of development, which makes it possible to monitor the effect of alterations over several generations. **Figure 1.3** illustrates the early developmental stages of zebrafish embryos, which reach adulthood at 3 months. Genetic manipulation can be easily performed using various methods, including CRISPR-Cas9, and zebrafish show high tolerance to such modifications (Varshney et al., 2016). Finally, the availability of fluorescent transgenic lines that label different structures allow for clear visualisation of structures of interest, providing an excellent tool for developmental studies.

In the case of examining the functional roles of GWAS hits, zebrafish displays an array of advantages that make it an exceptionally good model. Zebrafish provides a sufficient in-vivo system with a substantial number of organisms that can be tested at a low cost when compared to mouse models. The availability of a sequenced genome and al large number of CRISPR targets, together with high tolerance to genetic manipulations, further facilitates the genetic manipulation process. Multiple genes can be targeted within the same organism, and the generational effect can be easily followed due to the rapid rate of development. In addition, any induced mutation can be selected to produce homozygous and

heterozygous mutants. The effect of a mutation can be examined across different body systems rather than in a single cell type.

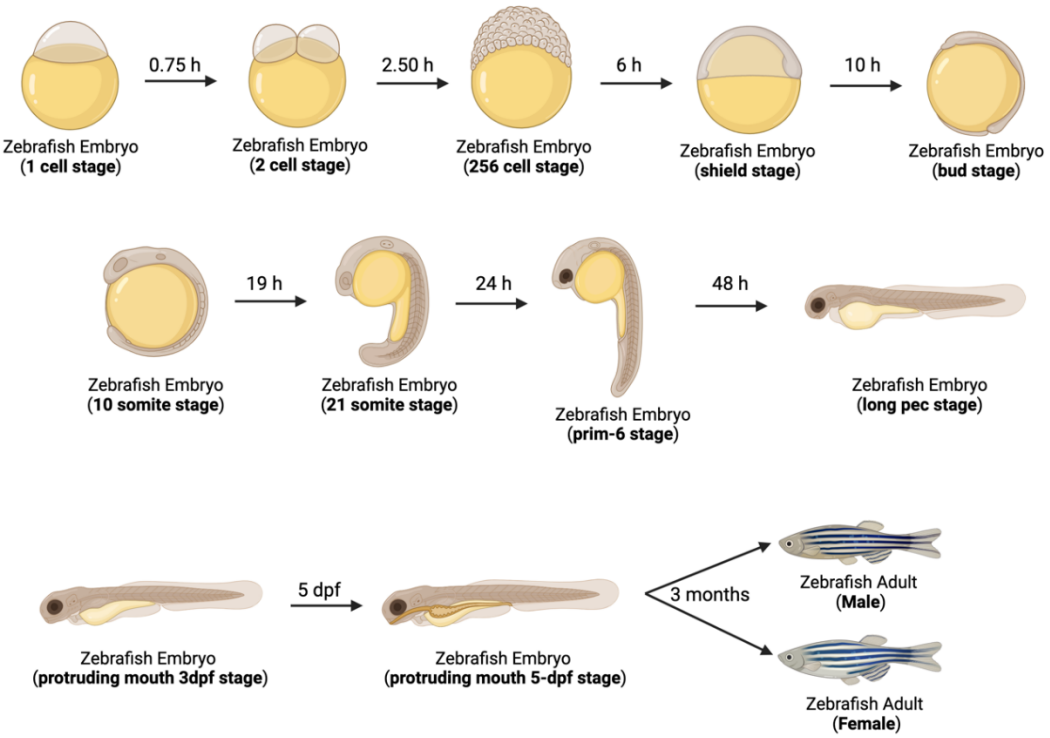


Figure 1.3: Zebrafish developmental stages

Zebrafish development from a single cell embryo at 0 hours to adulthood, normally reached at 3 months of age. Arrows represent the progress to each stage and the time taken to reach each stage is highlighted in hours above the indicated arrow. Figure generated using Biorender (https://www.biorender.com/).

1.6.1. Zebrafish as a model organism for studying the musculoskeletal system

The use of zebrafish in musculoskeletal system studies is a growing field (Mackay et al., 2013) due to the availability bone-related transgenic lines that allow us to monitor early skeletal system development (skeletogenesis) in detail in a live organism. In terms of bone composition, zebrafish and humans share many similarities as both bone structures are composed of collagen fibres and hydroxyapatite crystals. Zebrafish undergo bone mineralisation (Apschner et al., 2011) and bone tissue formation and resorption similar to humans (Witten et al., 2001). Cartilage and bone can be visualised separately using different fluorescent colours, allowing for easier imaging. In musculoskeletal studies, the *Tg(Col2a:mCherry)* transgenic line can be used to visualise the cartilage in red (Hammond et al., 2012), while a bone-specific transgenic line such as *Tg(sp7:EGFP)b1212* is used to highlight the bones in green (DeLaurier et al., 2010). At 3dpf, a synovial joint start forming, allowing the jaw area to mimic OA sites such as

the knee in humans. One study examined six genes known to be associated with OA that were selected using GWAS and the ensembl database. Functional studies were performed in zebrafish and the expression patterns of all six genes were successfully identified (Mitchell et al., 2013). This highlights the remarkable potential of using zebrafish models in functional studies to confirm the role of biomarkers identified by GWAS. Furthermore, the use of zebrafish to study OA allows us to eliminate the mechanical strain and obesity factors in OA development, due to the weightlessness of the water environment. This allows a more comprehensive focus on the genetic link, while restricting the impact of environmental factors.

1.6.2. Zebrafish as a model organism for studying heart structural development

Zebrafish heart shows high similarity to the heart structure of humans. It consists of a myocardium, endocardium, epicardium and fibroblasts, which are highly conserved between the two species. Furthermore, critical structures found in the developed human heart including both main chambers, the atrioventricular valve, the aortic valve, the pacemaker, and the coronary vasculature, are also found in zebrafish (Staudt and Stainier, 2012). **Figure 1.4** displays early stages of zebrafish heart development up to 5dpf.

On the other hand, there are some structural differences between the human and zebrafish heart. Zebrafish have a single atrium and ventricle, forming a two-chambered heart with a single circulatory system, rather than the dual circulation found in humans. The adult zebrafish heart has a different structural composition with excessive tuberculation (a process that involves maturation of the heart, segmentation of structures to form the different chambers, and looping of structures to achieve correct alignment within the organism) when compared to the human heart (Jensen et al., 2016). Nevertheless, zebrafish provide an excellent tool for studying structural development in the musculoskeletal and cardiovascular systems. Available transgenic lines were used to induce a CRISPR-Cas9 mutation targeting the gene of interest. The effect of gene targeting on the musculoskeletal and heart structural development was monitored using *Tg(Col2a1aBAC:mCherry)* (Hammond et al., 2012) and *Tg(myl7:LifeactGFP)* (Reischauer et al., 2014). Cartilage is shown in red and the heart in green.

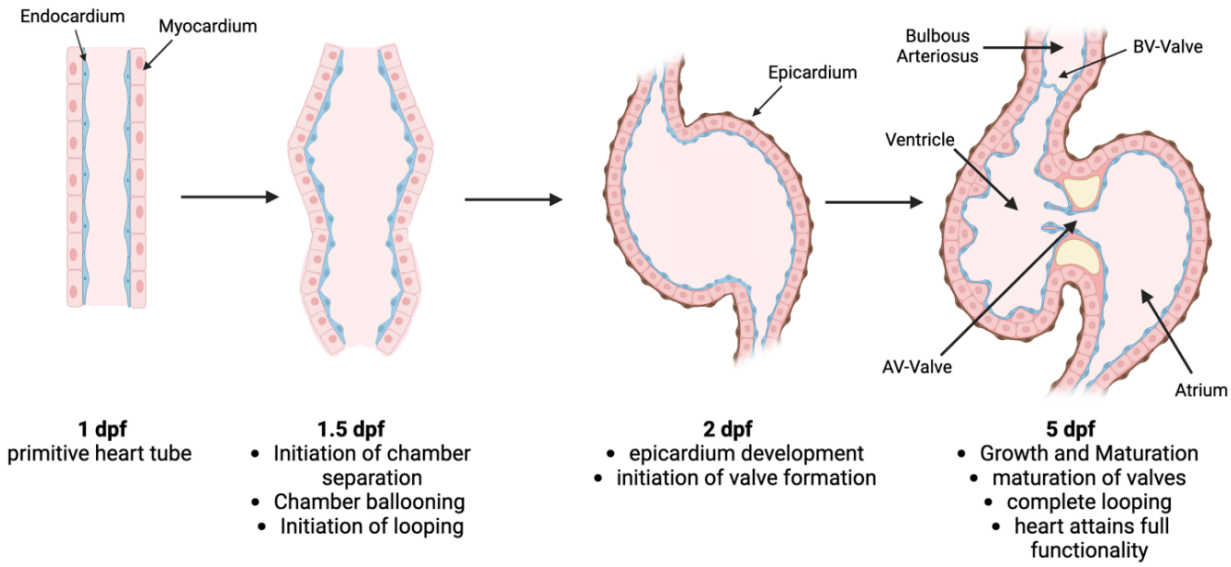


Figure 1.4: Zebrafish heart development and looping

The arrows from left to right represent the early development of the zebrafish heart, starting with the primitive heart tube present at 1-day post-fertilization (1dpf) with a basic inner endocardium wall and outer myocardium. At 1.5dpf, the tube starts ballooning, initiating the formation of the two heart chambers and their separation. At 2dpf, a clear separation of the chambers becomes visible, with the heart structure looping to align itself in its correct orientation within the embryo. At this stage, an outermost layer of epicardium starts forming, along with the valves between the chambers. By 5dpf, clear separation of the two chambers is observed, with complete looping of the two chambers and functional valves separating the Bulbous arteriosus from the ventricle, and another AV-valve between the atrium and ventricle. Figure generated using Biorender (https://www.biorender.com/).

1.7. Hypothesis and aims

Several studies have identified an increase in the risk of CVD development in individuals with OA (Ong et al., 2013; Veronese et al., 2016). Furthermore, as the number of GWA-studies increases, the same biomarker hits started emerging in both OA and CVD studies separately (Tachmazidou et al., 2019; Zhang et al., 2021; Pare et al., 2010). Finally, a recent meta-analysis indicated a 24% rise in the risk of developing CVD in OA patients by pooling data from 19 distinct studies, looking at the association between OA and CVD (Wang et al., 2016). Therefore, a hypothesis can be made regarding a common genetic link underlying the pathogenesis of both OA and CVD, alongside other chronic diseases.

The following list outlines the primary objectives of this study, in the greater aim of identifying common links between chronic diseases with a specific focus on OA and CVD. The identification of a common genetic link could inform the selection of appropriate drug targets, to avoid polypharmacy, or investigate the repurposing of pre-existing drugs.

- Identifying common risk genes between OA and CVD by combining data from the GWAS Catalog and GWAS Central databases from available OA and CVD GWA-studies.
- Utilizing bioinformatic tools to gain further insights into the functions of the identified genes and their associated pathways.
- Targeting of the identified common risk genes using CRISPR-Cas9 genome editing tool in single cell zebrafish embryo, followed by monitoring the impact of the targeting on the skeletal and heart development morphology in the *Tg(Col2a:mCherry)* cartilage line, and the *Tg(myl7:lifeActGFP)* heart transgenic line. This objective aims to identify the intrinsic role of the gene in early development.
- Examining the impact of targeting on the function of the skeletal system by monitoring the swimming abilities of fish during their larval and adult stages.
- Investigating the effect of targeting on subsequent bone development by examining bone structural development in adult zebrafish using micro-CT scans.
- Attempting to investigate the impact of a single gene targeting on the expression of other studied genes.

Taken together, this postulation may provide an explanation for multimorbidity's nature and the concurrent incidence of several chronic diseases within the same individual. Exploring common genes and mechanism underlying chronic diseases may reveal fundamental pathways in multimorbidity pathogenesis and help identify less detrimental treatment approaches to polypharmacy.

1.8. Functional genomic analysis pipeline

Following the identification of common risk genes between OA and CVD, we designed and executed a functional genomic analysis pipeline to explore gene depletion effects and examine our primary study objectives. To understand the role of a gene, the consequences of its absence must be examined, which will reveal its intended function in the different systems. CRISPR-Cas9 genome editing tool will be used to target the identified common risk genes in zebrafish embryos. The effect of specific gene targeting will be examined in relation to exhibited phenotype within the context of cartilage and heart development. The study will employ the specific fluorescent transgenic lines, *Tg(Col2a:mCherry)* for morphological observation of the cartilage in red and the *Tg(myl7:lifeActGFP)* line for visualising the heart in green. Genotyping will be conducted on produced crispants to match with the corresponding resultant phenotype. Moreover, the functionality of the skeletal system will be assessed by monitoring the impact of mutations on the swimming abilities of zebrafish at the larval and adult stages, using the Viewpoint Zebrafish Box system. Finally, to investigate the effect of the gene targeting on later bone development or maturation, micro-CT scans will be performed on the adult crispants. This pipeline facilitates screening for the functional role of different genes across the two examined structures.

2. Materials and Methods

2.1. Materials

Common laboratory reagents were purchased from Sigma Aldrich, Roche Diagnostics, or ThermoFisher Scientific, unless otherwise stated. DNA oligonucleotides were synthesised and purchased from Integrated DNA Technologies (Leuven, Belgium).

2.2. General zebrafish techniques

2.2.1. Zebrafish husbandry and maintenance

All zebrafish work was carried out under the regulations of The Home Office Guidance on the Operation of Animals (Scientific Procedures) Act 1986 (ASPA), project license: PP9182752 and personal License: I38989005. The fish were housed at the Centre for Developmental and Biomedical Genetics (CDBG) at the University of Sheffield. They are maintained in a semi-closed environment, where fresh water gets inflexed by approximately 10% every 24 hours. The temperature is kept between 26.5 and 28°C and the pH is nearly neutral. A 14-hour light/ 10-hour dark cycle is implemented. High-water quality is essential for successful nurturing and breeding. To maintain the good quality, water in the aquarium is passed through various filters. Starting with a carbon filter to dechlorinate the water, it then passes through an ultraviolet steriliser to mutate any micro-organisms. The removal of toxic waste, particularly ammonia, is another essential step, and a pressurised sand filter and a biological filter are used to reduce the amount of ammonia, then any ammonia that remains is converted to nitrates, which is a much less harmful product. To ensure the consistency of healthy breeding, the system water is regularly checked to preserve the following mineral concentrations including 0.01-0.02 mg/L ammonia, <25 ml/L nitrate, 0.5-1.0 mg/L nitrite and 0.00 chlorine.

Young larvae are fed rotifers and dry food until they are 11 days old, when they are switched to Artemia as their primary food source. When the fish are old enough, they are transferred from petri dishes to glass tanks containing 10 litres of system water, with a maximum of 40 fish per tank. To breed the fish to obtain embryos, two methods can be used depending on the experiment to be performed (Audeeh et al., 2012).

2.2.2. Embryo collection

The first method is commonly known as marbling. This involves placing a sieve container with different coloured marbles fitted in a closed deep container underneath. Fish are attracted to the marbles, so they

gather to swim over them, and the females lay their eggs and subsequently males spray their sperm over to fertilise the eggs. The eggs fall through the sieve so the fish do not have access to them, otherwise they would be eaten by the adults. Fertilised eggs can be easily obtained by draining the tank through another finer sieve.

Alternatively, fish pairing can also be used to obtain the embryos. This method is generally less desirable as it involves a lot of handling of the fish to place a male and a female in a small transparent container with a clear divider between them. The container has a sieve and is enclosed in another transparent container so that the eggs can fall without the fish having access to them. The set-up presented in **Figure 2.1** is prepared the day before the embryos are required. Fish must be left overnight to allow for sufficient time for them to adapt to their new environment and more importantly for the females to be ready for egg laying. Otherwise, the males will be ready prior to the females. The next morning, dividers are removed, and the fish are allowed to mate. Embryos are collected immediately after fertilisation. This is essential for age-sensitive experiments such as the microinjections performed in this study, which must be done at the single cell stage.

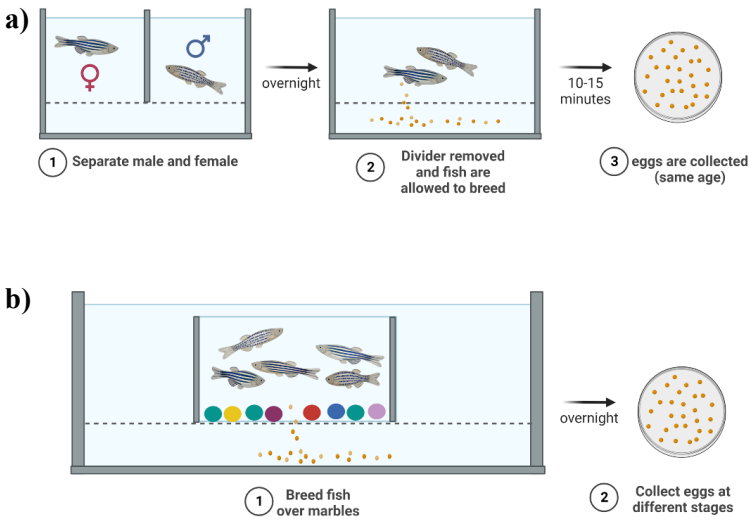


Figure 2.1: Zebrafish embryo collection methods

(a) One male and one female are placed in a single smaller tank, separated by a transparent divider. The following day, the divider is removed, and the pair is allowed to breed so that the eggs can be collected from the lower container. This procedure provides access to embryos that are born at the same time and will therefore develop simultaneously for time-critical experiments. **(b)** A sieve-bottomed container containing marbles is placed in the tank to attract fish to mate. Eggs can be collected from the bottom tank at different stages. Once the embryos have been collected, they are maintained in E3 medium

(0.17mM KCl, 0.33mM $CaCl_2$, 5mM NaCl, 0.33mM $MgSO_4$) with 0.001% methylene blue in a 28.5°C incubator with a regulated light cycle.

2.2.3. Zebrafish anaesthesia

Embryos (<5.2 days post fertilisation) are anaesthetised with 0.4% w/v tricaine in E3 (PharmaQ, Hampshire). Upon completion of handling, embryos are transferred to fresh E3 medium, and adults are monitored while recovering in system water.

2.2.4. Transgenic lines

Different transgenic lines were used for different purposes. For the visualisation of the cartilage, the transgenic line *Tg*(*Col2a1aBAC:mCherry*)^{hu5900} was used, where cartilage is visible in fluorescent red. To study the heart structure the *Tg(myl7:lifeActGFP)* line was chosen, highlighting the heart in green. The widely available wildtype AB fish ZDB-GENO-960809-7, as well as wildtype transparent line Nacre were also part of this study.

Zebrafish Line	Structure	Reference
Wildtype AB (ZDB-GENO-960809-7)		
Wildtype Nacre		
Tg(Col2a1aBAC:mCherry)^{hu5900}	Cartilage in Red	Hammond et al., 2012
Tg(myl7:lifeActGFP)	Heart in Green	Reischauer et. al. 2014

Table 2.1: Zebrafish transgenic lines included in this study

2.3. CRISPR-Cas9 genome editing technique

2.3.1. Single-guide RNA and primer design

To design the CRISPR sgRNA, the UCSC genome browser (UCSC Genome Browser Home, 2019) was used to select the best targets for the guide RNA. The genomic size was approximately 20

nucleotides, based on targets with the lowest off-target score and high overall score assigned by UCSC, preferably above 900 and located in an exon at the beginning of the gene. The selected sequence was then obtained with the addition of 150 bp flanking nucleotides on each side. For guides to be used efficiently in CRISPR modulations, suitable primers must be designed around the mutated region for later confirmation of induced mutation through genotyping. To confirm the availability of such primers, the entire target region with a total of 300 bp of flanking nucleotides was entered into Primer3Plus, which automatically selects the most appropriate forward (F) and reverse (R) primers, **Table 2.3**. All our primers were selected within the 50-60% CG ratio to ensure accuracy.

For each candidate gene, two target-specific sgRNA were designed. These are listed in **Table 2.2**. Two guide RNAs were used to increase the efficiency of the knockout, and to potentially remove a larger DNA sequence for easier detection. The use of a single guide RNA designed by Synthego, rather than combining two separate crRNA and tracrRNA has been shown to increase the efficiency of the knockout and was therefore chosen as the first step in the pipeline. To 1.5nmole of lyophilised target-specific sgRNA, 20µl of nuclease-free water was added for reconstitution. A working solution was then prepared using 30% of the stock solution and further diluted with nuclease-free water.

Target Gene	Target Exons	sgRNA
tgfß1a	1 and 2	Guide 1= GAATCCGGAGCGGACGACGA Guide 2= GGAACTGTATCGCGGAGTGG
gdf5	1 and 2	Guide 1= GAGATCTTCGCGTTATCG Guide 2= GGTTTTTTCCCACGCGGCAA
ctsk	4 and 5	Guide 1= GATAGATACCGTTCCAACGT Guide 2= GAGGGCACTGACCGCCGCCG

Table 2.2: CRISPR-Cas9 single-guide RNAs

Two guide sequences utilised to target specified exons. Sequences selected using the Zebrafish CRISPR target tool UCSC (Genome Res. 2002 Jun;12(6):996-1006).

2.3.2. CRISPR guides and primers validation

To validate the designed guides (**Table 2.2**) and primers (**Table 2.3**), embryos were injected at single cell stage with the designed guides, along with Cas9 protein. DNA was extracted from injected and control embryos, and a simple PCR was performed to check for clear bands on a 1% agarose gel (ThermoFisher Scientific,17850) at the expected size on the DNA ladder. After confirming that the primers are specific to our targets, a fluorescent PCR was performed to detect the efficiency of the designed guides in modulating the gene under investigation, indicated by a disruption of the wildtype peak observed in control samples.

Target Gene	CRISPR Primer Sequence
tgfβ1a	F= TGTAAAACGACGGCCAGTGGAGGTGGTGAGGAAGAAGC R= GTGTCTTGTTTGCTTGGTGCTGACCG
gdf5	F= TGTAAAACGACGGCCAGTACTTTCTCTGAGGGCATGGC R= GTGTCTTCTCAGAGGAAGATCAGAGGATGA
ctsk	F= TGTAAAACGACGGCCAGTATGTACCGCGATCCAGCAAA R= GTGTCTTTCATTGTTAGGCTGTGAAATCGT

Table 2.3: CRISPR primer sequences

Designed using Primer3Plus. An M13 tail sequence was added to the forward primer (F), while a pigtail end was added to the reverse primer (R), allowing the primers to be used for CRISPR-STAT (fluorescence PCR) analysis.

2.3.3. Microinjection of single cell zebrafish embryos

In preparation for the CRISPR microinjection, borosilicate glass needles are pulled using a P-1000 flaming/brown micropipette puller to obtain fine needle tips. Agarose moulds are made using 1.5% agarose in E3 medium, and embryos are placed in the mould for precise positioning. The injection mixture consists of 0.5µl Cas9 protein (Sigma Aldrich, CAS9PROT-50UG), 0.5µl of each 30% working solution sgRNA (Synthego), 1µl Dilution buffer, and 1µl Phenol red (Sigma Aldrich, P0290). The mixture is prepared and allowed to stand on ice before being used for the injection. The mixture is loaded into needles using microloader tips (Eppendorf, Cat no. 5242956003) and fixed to a PV820

Pneumatic PicoPump (World Precision Instruments). The fine tip of the needle is gently broken using forceps and air is pumped through the picopump injector into mineral oil (Sigma, M5904) on a glass graticule to calibrate the injection bolus. In this study we injected two consistent volumes of 0.5nl and 2nl to obtain the final optimal concentrations of each guide RNA (100pmol and 400pmol respectively).

2.4. Molecular biology techniques

2.4.1. Zebrafish larval DNA extraction

For embryonic/larval tissue DNA extraction a Sigma Aldrich Extraction-N-Amp kit (Sigma Aldrich, XNAT2-1KT) was utilised to yield high concentration DNA. For each sample, a mixture of 25µl extraction buffer and 7µl tissue preparation solution was added to separate wells of a 96-well plate. Embryos were incubated at room temperature (RT) for 10 minutes and then boiled at 95°C for 5 minutes to dissolve the tissue. 25µl of neutralisation buffer was then added to each well and vortexed briefly to mix. A sample of the extracted DNA was then diluted 1:10 in nuclease free water to achieve the appropriate DNA concentration required for the PCR reaction.

2.4.2. DNA extraction from adult fin clips

For DNA extraction from adult fin tissue a NaOH/Tris method was used. For each fin clip 25µl of 50mM NaOH (VWR Chemicals, 1310-73-2) was added for full immersion and boiled in a thermocycler at 95°C for 5 minutes, followed by a cooling step at 4°C for 10 minutes. 6µl of 500mM Tris-HCL (pH 8.0) (VWR Chemicals, J831-500ML) was added to each sample and centrifuged for 1,500xg, 5 minutes at RT. For PCR reactions 1.5µl of stock DNA was used per 20µl reaction volume.

2.4.3. RNA extraction

1ml of TRIzol reagent (Ambion, 15596018) was added to a 1.5ml tube containing 20-50 zebrafish larvae (<5.2dpf). Using a pellet pestle attached to a homogeniser, samples were homogenised for 1 minute to disrupt tissue and release RNA. Samples were then centrifuged at 4°C for 10 minutes at 13,000 rpm. The supernatant was transferred to a fresh 1.5ml tube, and 200µl of 1-bromo-3-chloropropane (BCIP) (Sigma Aldrich, B9673) was added per 1ml of supernatant. Tubes were vortexed briefly, and allowed to stand for 10 minutes at RT before being further centrifuged for 15 minutes at 4°C. At this stage the samples separated to form three distinct layers, phenol chloroform, DNA, and an upper RNA layer. The RNA layer was pipetted into a fresh 1.5ml tube and an equal volume of 70% ethanol was added. The RNA Clean & Concentrator Kit (Zymo Research, R1019) was used to purify and concentrate the extracted RNA. 30µl of nuclease-free water (Zymo Research, W1001) was used to elute the RNA and its concentration was measured using a nanodrop. The extracted RNA can be stored at -80°C.

2.4.4. Gel electrophoresis

A 1% agarose gel was prepared by mixing 1g of agarose powder (ThermoFisher Scientific,17850) with 100ml 1x TAE buffer warmed up in a microwave. The mixture was allowed to cool, and 10µl PAGE Gel Red (Biotrend, 41008-T) was added to stain the DNA before being poured into a gel tray with a well comb inserted. After allowing the gel to set, then 3µl of the PCR product was mixed with 0.5µl of 6x DNA loading dye (ThermoFischer, R0611) and loaded into the designated well. 3µl of DNA ladder (New England Biolabs, N0550S) was loaded onto the first lane to act as a reference for band sizes. Gels were run at 120V for 45 minutes and then imaged in a BioDoc-It 220 Imaging System Transilluminator (UVP, 97-0182-02).

2.4.5. Genotyping using CRISPR-STAT (Fluorescence PCR)

To initially confirm the presence of a mutation induced by CRISPR targeting, a fluorescence-based PCR method known as CRISPR-STAT was our technique of choice. The readouts are peaks at specific locations and heights that appear to be disrupted or missing in modulated samples. The key elements in this reaction are primers with a fluorescently labelled M13 pigtail ends. This technique was performed according to the protocol of Varnshey et al., 2016. The PCR master mix was prepared using 7.94ml 10x PCR Buffer II, 795µl 100mM dNTP mix (25mM each), 7.94ml 25mM MgCl2 solution, and 635µl AmpliTaq Gold on ice (Applied Biosystems, N8080241). A primer mix was prepared using 485µl TE (pH 8.0) (Invitrogen, AM9849), 5µl M13F-tailed forward primer (100µM), 5µl pigtail reverse primer (100µM), and 5µl M13-FAM primer (100µM). Primer sequences are listed in **Table 2.3**.

To each 100µl of PCR mix, 6µl of M13-FAM primer mix was added and 5µl of this mixture was aliquoted into separate wells of a 96-well plate. To each well, 1.5µl of 1:10 diluted DNA was added, and samples were left to run in a BioRad thermocycler following the programme in **Table 2.4**.

Cycle	Temperature and Duration
1	94°C for 12 mins
2-41	94°C for 30s, 57°C for 30s, 72°C for 30s
42	72°C for 10 mins
43	4°C hold for infinity

Table 2.4: PCR thermocycler program

To confirm the success of the PCR, 2.5µl of PCR product was loaded onto a 1% agarose gel to perform gel electrophoresis. When clear bands were observed on the gel, 2.5µl of the remaining PCR product was loaded into a 96-well Axygen plate (Axygen, AXP440). A mixture of 1:50 dilution of GeneScan 400HD ROX (Thermofisher, 402985) in HiDi formamide (Thermofisher, 4401457) was prepared and 10µl volumes added to the PCR product in separate wells. The plate was sealed and heated to 95°C for 5 minutes, before loading the plate into an ABI 3130xl or 3730xl genetic analyser using pre-set CRISPR-STAT module parameters.

2.4.6. Sequencing

Sequencing was done via four different providers: The Sheffield Microarray/Genomic Core Facility (Sheffield, England), Eurofins (Ebersberg, Germany), Genewiz (Leipzig, Germany) and Source BioScience (Nottingham, England).

Samples were prepared by mixing 5µl FirePol Master Mix (Solis Biodyne, 04-11-00125), 0.5µl forward primer, 0.5µl reverse primer, 17µl nuclease-free water and 1µl of stock DNA to be tested in PCR tubes. Samples were left to run in a thermocycler following the protocol in Table 2.5. below. 3µl of PCR product was run on a 1% agarose gel to confirm the success of the PCR. Samples were then purified using 0.05µl of Exonuclease I enzyme and 1µl of Shrimp Alkaline Phosphatase (SAP) (ThermoFisher, 783901000UN) to every 5µl of PCR product and incubated at 37°C for 15 minutes, followed by 80°C for a further 15 minutes to deactivate the enzymes. The DNA concentration was then measured using a Nanodrop ND-Spectophotometer to send the appropriate concentration according to the sequencing company along with the designed primers.

Step	Temperature and Duration
Step 1	95°C for 3 minutes
Step 2	95°C for 30 seconds
Step 3	58°C for 30 seconds
Step 4	72°C for 1 minutes
Step 5	Repeat step 2-4 30 times
Step 6	72°C for 5 minutes
Step 7	Hold at 4°C for infinity

Table 2.5: PCR thermocycler conditions for sequencing samples

2.5. Microscopy and imaging

2.5.1. Stereomicroscope for general morphology imaging

All brightfield imaging was performed using a ZEISS Axio Zoom V16 microscope (ZEISS, 495010-0004-000) with an attached ZEISS Axiocam 503 mono (ZEISS, 426560-9040-000). 2D fluorescence images were aquired using the same equipment along with an attached HXP 200c illuminator (ZEISS, 435716-0000-000) and the ZEISS ZEN Pro imaging software.

2.5.2. Lightsheet microscopy

Lightsheet microscopy (ZEISS Lightsheet Z.1) is an excellent tool for imaging zebrafish, with a 360° angle to access a clear view of all fluorescent structures in different transgenic lines, and therefore can be used to generate 3D constructs for visualisation purposes. As shown in **Figure 2.2**, fish are mounted in 1-1.5% low-melting point agarose (Sigma Aldrich, A9414) in glass capillaries. Up to 5 fish can be mounted per capillary, and black capillaries were used for mounting fish up to 5dpf. The acquisition chamber was filled with MilliQ water for cartilage/bone imaging, while filled with E3 plus 8.4% tricaine for heart imaging. Imaging the heart also requires the chamber temperature to be set to 10°C and the

fish to be anaesthetised on ice to prevent the heart from beating and moving while the sample is imaged as z-stacks.

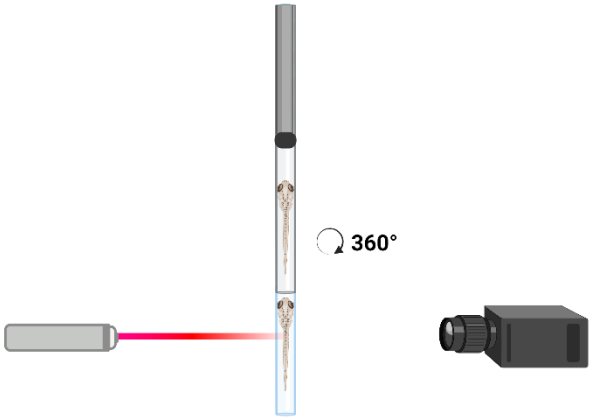


Figure 2.2: Zebrafish larvae setup for lightsheet imaging

Fish are mounted in low melting point agarose in glass capillaries. The capillaries are immersed in the MilliQ or E3 filled acquisition chamber, and samples are pushed through the liquid to allow for direct laser penetration. The capillary has a 360° field of rotation, allowing complete imaging of the structure of interest. Figure generated using Biorender (https://www.biorender.com/).

For cartilage imaging of the *Tg(Col2a:mCherry)* transgenic line, 25% laser power was passed through a 405/488/561 filter, and reflected onto a 560nm mirror line feeding to a 585 camera, visualising cartilage in red. Images are acquired in 16-bit depth from a single side laser angle chosen on the basis of preview quality. The exposure time is adjusted for each sample. For heart imaging in the *Tg(myl7:lifeActGFP)* line the same parameters were used but the 488nm laser was activated to show the heart in green. We also used a higher magnification of 1x on a 10x objective and single sided laser projection. All Lightsheet imaging was performed using the ZEISS ZEN Pro imaging software. After imaging, the samples were collected for DNA extraction and genotyping.

2.6. Zebrafish gene expression analysis

2.6.1. Quantitative Real Time PCR (qRT-PCR)

Using 1-2µg of extracted RNA (described in **section 2.4.3)**, cDNA can be prepared by reverse transcription. A master mix consists of 2µl 10x RT buffer, 0.8µl 25x dNTP, 2µl 10x RT random primers, 1µl reverse transcriptase, 1µl RNase inhibitor, 1-2µg RNA (variable amount per sample), and **x** µl nuclease-free water to make up a final volume of 20µl reaction. All reagents above are obtained from High-Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific, 4368814). The samples are

then left to run in a thermocycler according to the programme described in **Table 2.6**. cDNA is then diluted 10-fold by adding 180µl of nuclease-free water to each tube.

Step	Temperature and Duration
Step 1	25°C for 10 minutes
Step 2	37°C for 120 minutes
Step 3	85°C for 5 minutes
Step 4	Hold at 4°C for infinity

Table 2.6: Thermocycler conditions for reverse transcription.

For qRT-PCR, a SYBR Green based master mix is prepared using 63.5µl Fast SYBR Green (ThermoFisher Scientific, 4385612), 4.22µl 10µM forward primer, 4.22µl 10µM reverse primer, and 2.11µl nuclease-free water. The primers used are listed in **Table 2.7**. This master mix is sufficient for 10 samples. To an appropriate qPCR 96-well plate, 7µl master mix is aliquoted along with 5µl of diluted cDNA. The plate is briefly vortexed and run on a Bio-Rad CFX96 C1000 Touch Real Time PCR instrument according to the protocol described in **Table 2.8**.

Target Gene	CRISPR Primer Sequence
tgfß1a	F= GGAGGTGGTGAGGAAGAAGC R= ATCTTCTGAACCCTGCAGCC
gdf5	F= ACTTTCTCTGAGGGCATGGC R= AGCGGATGGGAAAGTCACAG
ctsk	F= ATGTACCGCGATCCAGCAAA R= ACTTCTTGCCTCTCGGTGTG

Table 2.7: qRT-PCR (Whole-Coding) primer sequences

The above primers were used to detect DNA expression levels in qPCR reactions. The same primers were also used to generate in situ hybridization probes for specific localised RNA expression detection. The primers were designed using Primer3Plus.

Step	Temperature and Duration
Step 1	50°C for 2 minutes
Step 2	95°C for 10 minutes
Step 3	95°C for 15 seconds
Step 4	60°C for 30 seconds
Step 5	72°C for 30 seconds
Step 6	Repeat step 3-5 39x
Step 7	65°C for 5 seconds
Step 8	95°C for 50 seconds

Table 2.8: Conditions for qRT-PCR.

2.6.2. In situ hybridization probe design

To design suitable primers, the cDNA sequence of the candidate gene was obtained from Ensembl (https://www.ensembl.org), taking into account the presence of multiple transcripts, and their variability in exon number and location. The cDNA sequence was exported to Primer3Plus to generate suitable forward and reverse primers that encompass the exons of interest within the sequence and have a good CG ratio. Using the designed primers, a PCR reaction was performed using 5µl FirePol Master Mix (Solis Biodyne, 04-11-00125), 0.5µl 10µM forward primer, 0.5µl 10µM reverse primer, 1µl stock DNA, and 3µl MilliQ water to give a final volume of 10µl per tube. Mixture was left to run in a thermocycler using the protocol listed in **Table 2.5**.

3µl of PCR product was run on 1% agarose gel in 1xTAE buffer to confirm the success of the PCR, as indicated by strong, clear, single bands at the expected position along the DNA ladder.

2.6.3. TOPO cloning protocol

To amplify the DNA fragment obtained, a TOPO cloning kit was used, which contains a pCRII vector (**Figure 2.3**) with an attached polyT tail overhang. This overhang polymerises with the polyA tail added during the PCR reaction, allowing transformation into E. coli bacteria. All reagents were obtained from the TOPO cloning kit (Invitrogen, 45-0640) and used in the following amounts: 0.5µl vector, 0.5µl salt solution, 1µl of PCR product and 1µl nuclease-free water. The mixture was left at RT for 5 minutes and then transferred to ice.

Cells were allowed to thaw on ice before adding 1µl of reaction mixture. Cells were incubated on ice for 30 minutes, then subjected to a short heat shock at 42°C for 30 seconds, followed by a recovery incubation on ice for 2-minutes. 250µl RT SOC bacterial growth medium (Sigma Aldrich, S1797) was added to the cells, and incubated horizontally at 37°C for 1 hour at 200-350rpm. Bacteria were spun down at 3000rpm for 5 minutes at RT, 30µl of supernatant was discarded and the remaining liquid was pipetted back with the cell pellet to increase the concentration and plated onto ampicillin agar plate (Sigma A0166. 100µg/mL). Plates were pre-warmed at 37°C for 30 minutes, then 40µl of Xgal (ThermoFisher, R0404) and 40µl of IPTG (Merck, I6758) were added to the plate, which was allowed to dry for a further 30 minutes before the bacteria were spread over it. The plates were then incubated overnight at 37°C to allow for bacterial growth. Successful recombinant cells will appear as white colonies. 4-6 white colonies were selected and resuspended separately in 10µl MilliQ water. A colony PCR reaction was performed on each resuspended colony to determine if transformation of the desired fragment had occurred. For this, 5µl of colony suspension was added to 0.4µl of 10µM forward primer, 0.4µl of 10µM reverse primer (**Table 2.7**), 10µl FirePol, and 4.2µl MilliQ water. The reaction was then run on a thermocycler using the programme in **Table 2.9**.

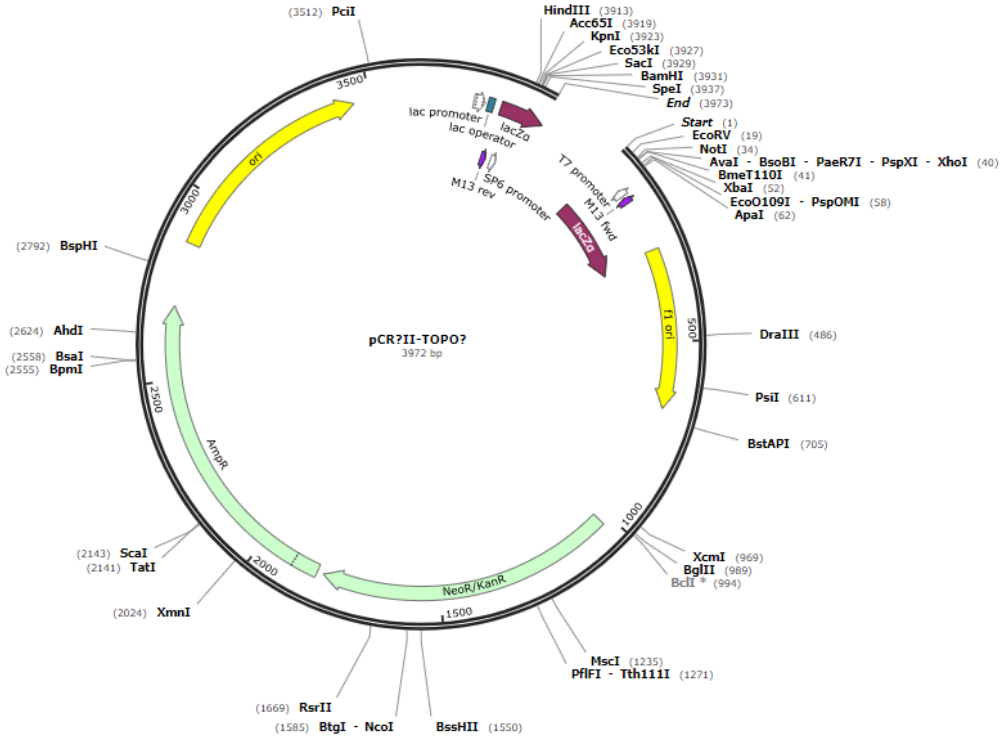


Figure 2.3: pCR-II TOPO cloning vector

Figure presents vector map used for cloning fragment of interest, including restriction sites for specific enzymatic restriction digestion and linearisation. Construct map obtained from Snapgene.

Step	Temperature and Duration
Step 1	94°C for 2 minutes
Step 2	94°C for 20 seconds
Step 3	60°C for 20 seconds
Step 4	Drop 1°C per cycle
Step 5	72°C for 45 seconds
Step 6	Repeat steps 2-5, 9 times

Step 7	94°C for 20 seconds
Step 8	50°C for 20 seconds
Step 9	72°C for 45 seconds
Step 10	Repeat steps 7-9, 14 times
Step 11	72°C for 3 minutes
Step 12	10°C for 5 minutes

Table 2.9: Colony PCR reaction program

3µl of PCR product was run on 1xTAE agarose gel for 45 minutes at 110 volts. If positive, the remainder of the colony/milliQ mixture was used to grow transformed colonies, which were processed in a plasmid miniprep kit (Zymo Research, D4200-B) to extract the desired plasmid. Plasmids were sent for sequencing using an M13R primer to check the orientation of the insert within the plasmid and to select a restriction enzyme to cut at the 5' end to generate an antisense probe.

Once a suitable restriction enzyme has been selected, a digestion is performed by mixing 10µl DNA, 5µl Cutsmart buffer (New England Biolabs, Catalogue No. B7204), 0.5µl restriction enzyme and 34.5µl MilliQ water. The mixture is incubated overnight at 37°C and purified the next day using the Qiagen PCR Purification Kit (Qiagen, 28104) or the DNA Clean and Concentrator Kit (Zymo Research, D4033). At this stage the plasmid should be fully linearised. To check the state of linearization, 1µl of the presumed linearised DNA is mixed with 4µl MilliQ water and 1µl loading dye and run on a 1% agarose gel for 2 minutes at 100 volts. As a negative control, an additional sample is prepared with 0.5µl of the original non-linearised plasmid, 4.5µl MilliQ water and 1µl loading dye is added to the gel. If the plasmid has been linearized, a single band will appear on the gel, whereas if the plasmid failed to linearise, it would appear as an intact structure in the form of multiple bands or a thick smeared band.

The linearised plasmid can then be used to transcribe the probe. A digoxygenin (DIG) RNA labelling mix (Roche,11175033910) was used to label the RNA probes. 2µl DIG RNA Labelling Mix was added to 2µl transcription buffer, 1µl RNAse out, 2µl RNA polymerase, 1µg of linearised DNA and a variable amount of MilliQ (depending on DNA concentration) to give a final volume of 20µl. The reaction was

incubated at 37°C for 2 hours. 1µl of DNAse was then added to the tube, briefly vortexed and incubated for a further 15 minutes at 37°C. The labelled RNA probe was then purified using the RNAeasy Kit (Qiagen, 74004) before being used for in situ hybridization staining.

2.6.4. Whole mount in situ hybridization

In situ hybridization is a widely utilised technique in zebrafish research as it allows the expression levels of specific RNAs of interest to be clearly visualised in a localised manner within the organism. The technique relies on the binding specificity of RNA probes designed as described in **section 2.6.2**. Zebrafish larvae are fixed in 4% PFA overnight at 4°C with gentle agitation. Embryos were then washed in 1xPBST for 5 minutes at RT. This step was repeated three times to ensure complete removal of PFA before dehydration in a series of 30%, 50%, 70% and 100% methanol dilutions to preserve the structure of the fish. Each wash was performed on a rocker for 5 minutes at RT and the final 100% methanol wash was repeated 3 times before the embryos were stored at -20°C until required for staining.

To initiate the staining process, fish was rehydrated by reversing the dehydration process with 70-30% methanol/PBST washes before a final step of 4x 5-minute 1xPBST washes. For larger RNA probes to penetrate, and to reduce non-specific probe binding during hybridization, proteinase K (proK) was used at different concentrations and time durations depending on the embryo/larval stage at which the samples were fixed. Young embryos at 24 hours were incubated in 10µg/ml proK for 5 minutes at RT, while for 5dpf samples, embryos were incubated in 20µg/ml proK for 20 minutes at RT on a rocker. Proteinase K was then be replaced with 4%PFA-PBS for 20 minutes at RT for fixation. Embryos were then washed 5 washes with 1x PBST.

Hybridization buffer (Hyb-) was prepared as described in **Table 2.10.** For prehybridization 50µg/ml heparin and 500µg/ml tRNA were added to 1ml hybridization buffer per reaction tube. Embryos were prehybridized for at least 1 hour at 70°C before hybridization was started by incubating samples in hybridization buffer containing probe (1:100 dilution) at 70°C overnight.

The probe was removed, and the embryos washed briefly in 100% Hyb- at 70°C. A series of 75%, 50% and 25% Hyb- washes were performed for 15-minute at 70°C, followed by 2x SSCT, 15-minute wash, then 0.2xSSCT for 30 minutes, repeated twice, all at 70°C. All reagents up to this stage require preheating prior to addition to the samples. All other steps from this point were performed at RT. A series of 75%, 50% and 25% 0.2xSSCT washes for 10 minutes were performed before a final wash for a further 10 minutes in 1xPBST. Embryos were then incubated in blocking buffer for at least 2 hours at RT before incubation in 1:50 Anti-Digoxigenin-AP Fab fragments overnight at 4°C with agitation.

The next day, antibodies are discarded, and the embryos are washed 8 times in 1xPBST for 15 minutes each at RT. To prepare for staining, embryos are first washed 3 times for 5 minutes in staining buffer. Staining solution is then added (1:50 NBT/BCIP in staining buffer) is then added and embryos are incubated at RT in the dark with agitation. Depending on the probe, staining may take a variable amount of time to develop and should be monitored frequently under a dissecting microscope. Once staining appears, the reaction can be stopped by washing off the staining solution with 1xPBST 3 times for 5 minutes. To fix the stain, embryos were incubated overnight in 4%PFA-PBS at 4°C. Finally, PFA was removed by 2 washes in 1xPBST for 10 minutes each and the embryos were dehydrated again in a series of 30%, 50%, 70% and 100% methanol washes. Stained embryos can be stored in 100% methanol at 4°C until imaging. Embryos were imaged in Murray's BBA solution (2:1 benzylbenzoate: benzylalcohol) to reduce background and obtain clearer images.

Reagent	Contents	Final Volume	Source
DEPC- treated water (autoclave)	1 ml DEPC 1 L Milli Q water	1 L	(Sigma Aldrich, D5758-5ML)
1x PBS	10 PBS tablets 1 L DEPC water	1 L	(Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, D5758-5ML)
10x PBS	10 tablets 100 ml DEPC water	100 ml	(Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, D5758-5ML)
20x SSC	70.12 g NaCl 35.28 g Sodium Citrate 400 ml DEPC water	400 ml	(Sigma Aldrich, S9888-25G) (Fisher Scientific, 6132-04-3, 68-04-2) (Sigma Aldrich, D5758-5ML)
2x SSCT	5 ml 20xSSC 45 ml 1xPBS 250 µl 20% tween 20	50 ml	(ThermoFisher Scientific, 15557044) (Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
20% Tween 20	10 ml tween 20 40 ml DEPC water	50 ml	(Sigma Aldrich, P1379-100ML) (Sigma Aldrich, D5758-5ML)
100% methanol in PBST	50 ml Methanol 250 µl 20% tween 20	50 ml	(Sigma Aldrich, P1379-100ML)
75% methanol in PBST	37.5 ml Methanol 12.5 ml 1xPBS 250 µl 20% tween 20	50 ml	(Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
50% methanol in PBST	25 ml Methanol 25 ml 1xPBS 250 µl 20% tween 20	50 ml	(Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
25% methanol in PBST	12.5ml Methanol 37.5 ml 1x PBS 250 µl 20% tween 20	50 ml	(Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
4% PFA-PBST	10 ml 16% PFA 30 ml 1x PBS	40 ml	(Thermo Scientific, 28906) (Sigma Aldrich, P4417-100TAB)
PBST	5 ml 10xPBS 45 ml DEPC water 250 µl 20% tween 20	50 ml	(Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, D5758-5ML) (Sigma Aldrich, P1379-100ML)
Proteinase K	1 µl Proteinase K stock 1 ml PBS 250 µl 20% tween 20	1 ml	(Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
1M Citric Acid	14.705 g trisodium citrate dihydrate 50 ml DEPC water	50 ml	(Sigma Aldrich, D5758-5ML)
Hybridization buffer (Hyb-)	25 ml Formamide 12.5 ml 20xSSC 250 µl 20% tween 20 460 µl 1M citric acid 11.24 ml DEPC water	50 ml	(Honeywell, 616-052-00-8) (ThermoFisher Scientific, 15557044) (Sigma Aldrich, P1379-100ML)
75% Hyb	37.5 ml wash hyb 12.5 ml 2xSSC 250 µl 20% tween 20	50 ml	(ThermoFisher Scientific, 15557044) (Sigma Aldrich, P1379-100ML)
50% Hyb	25 ml wash hyb 25 ml 2xSSC 250 µl 20% tween 20	50 ml	(ThermoFisher Scientific, 15557044) (Sigma Aldrich, P1379-100ML)
25% Hyb	12.5 ml wash hyb 37.5 ml 2xSSC	50 ml	(ThermoFisher Scientific, 15557044)

Reagent	Contents	Final Volume	Source
	250 µl 20% tween 20		(Sigma Aldrich, P1379-100ML)
0.2x SSCT	0.5 ml 20xSSC 49.5 ml 1xPBS 250µl 20% tween 20	50 ml	(ThermoFisher Scientific, 15557044) (Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
75% 0.2x SSC	37.5 ml 0.2xSSC 12.5 ml 1xPBS 250 µl 20% tween 20	50 ml	(ThermoFisher Scientific, 15557044) (Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
50% 0.2x SSC	25 ml 0.2xSSC 25 ml 1xPBS 250 µl 20% tween 20	50 ml	(ThermoFisher Scientific, 15557044) (Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
25% 0.2x SSC	12.5 ml 0.2xSSC 37.5 ml 1xPBS 250 µl 20% tween 20	50 ml	(ThermoFisher Scientific, 15557044) (Sigma Aldrich, P4417-100TAB) (Sigma Aldrich, P1379-100ML)
Blocking Buffer	20 mg/0.02g BSA 200 µl sheep serum 9.8 ml PBST	10 ml	(Sigma Aldrich, A3059-50G) (Sigma Aldrich, S3772)
DIG-AP antibody in blocking buffer	0.2 µl Anti-Digoxigenin-AP Fab fragment 1 ml blocking solution	1 ml	(Roche Diagnostics, 11093274910)
Staining Buffer	5 ml 1M tris-HCl pH9.5 5 ml 500mM MgCl2 5 ml 1M NaCl 34.75 ml DEPC water 250 µl 20% tween 20	50 ml (make fresh)	(Alfa Aesar, J62084) (Sigma Aldrich, S9888-25G) (Sigma Aldrich, D5758-5ML) (Sigma Aldrich, P1379-100ML)
Staining Solution	22.5 µl 100mg/ml NBT 35 µl 50mg/ml BCIP 9.9425 ml staining wash	10 ml	(Roche Diagnostics, 11383213001) (Roche Diagnostics, 11383221001)

Table 2.10: In situ hybridization reagents

2.7. Zebrafish behavioural analysis

2.7.1. Larvae behavioural analysis

The Viewpoint software (https://www.viewpoint.fr/) was used to detect any changes in the behaviour of crispant or mutant larvae up to 5dpf. Using a square based 96-well plate in the Viewpoint Zebrafish box machine, a 10-minute analysis was performed to look at the behaviour of fish in separate wells under two different conditions of light and dark throughout the 10-minute cycle. Fish were placed in 300µl of E3 medium, which was consistent across all wells. From the obtained Excel data sheet, the inadur (inactivity duration) and the lardur (hyperactivity duration) were our chosen parameters, these were averaged for each fish throughout the cycle and this value plotted in Graphpad prism to analyse using a student non-parametric t-test comparing the average time embryos spent in each category of motion for wildtype compared to crispants or mutants. Larvae should be handled minimally and carefully using a wide Pasteur pipette to minimise disturbance and stress to the larvae. In addition, the age of the larvae tested, and the time spent in the Viewpoint box should be carefully considered to avoid surpassing the unregulated stage without taking appropriate actions.

2.7.2. Adult behavioural analysis

The Viewpoint Zebrafish Adult system was used to test the effect of candidate gene modulation on adult fish locomotion. Up to 10 fish were placed in the system chamber to be monitored, as shown in **Figure 2.4.** Each fish was placed in a separate tank containing 600ml of system water. The tanks were stacked, and the fish were monitored by the system for 8 hours under two different conditions of light and dark throughout the whole cycle. The first hour was excluded from the analysis as fish require an adaptation period when transferred to the different environment of the chamber. Inadur (duration of inactivity) and Lardur (duration of hyperactivity) were the two parameters of interest as they indicate the time spent by each fish in an inactive state or in a hyperactive state of rapid movement. The data can be interpreted in several ways, such as a single average speed of the fish throughout the experiment, or a consistent monitoring of speed changes over time. Fish should be handled and transferred with care. When fish are transferred from the aquarium to the behavioural analysis room, they should be transferred in a black box with minimal movement and sufficient amount of system water. Experiments should be scheduled according to the light cycle and feeding times of the aquarium.

Figure 2.4: Adult behavioural analysis setup

Fish are housed separately in tanks filled with system water. Up to 10 tanks can be stacked per run. Tanks are monitored to detect any individual changes in behaviour.

The time spent by each fish in each state was averaged and compared with wildtypes using a student non-parametric t-test between the two groups. Alternatively, the behaviour of each fish over the 7-hour monitoring period was also examined to detect any possible patterns suggesting rapid exhaustion as indicated by a sudden decrease in speed or rather an increase in movement over time in crispants/mutants compared to wildtype controls of the same age.

2.8. Tailfin injury-regeneration model

Inducing an injury to the zebrafish embryo tailfin is a common tool used to study the regenerative capacity of the organism under different conditions. Prior to performing the injury, 3-5dpf larvae were anaesthetised in 1ml of 0.4% Triciane in 50ml of E3 medium. A blunt cut was made using a sterile 10a scalpel (Swann-Morton, 0302) as shown in **Figure 2.5.** Fish were then placed in fresh E3 and left for 24 hours. The next day, the fish were imaged using a ZEISS Axio Zoom V16 microscope in the brightfield display and the images were analysed in ImageJ by measuring the growth length of the regenerated notochord.

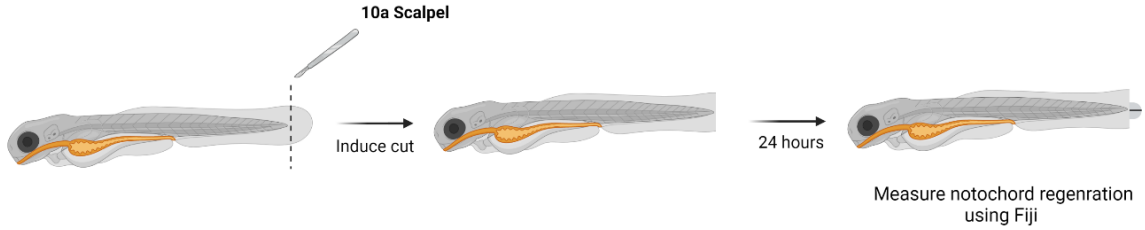


Figure 2.5: Tailfin injury-regeneration model

Anaesthetised fish are placed horizontally under a light microscope. A minor incision is induced using a 10a scalpel just below the notochord. Fish are allowed 24 hours to recover and regenerate tissue at the incision site. Images of fish were taken using a stereomicroscope in brightfield, and the length of the regenerated notochord is measured using Fiji (ImageJ) to measure the effect of experimental fish regenerative abilities versus wildtypes.

2.9. Microcomputed tomography scans

2.9.1. Phosphotungstic acid staining

This method is used to stain the soft tissue to visualise different organs from micro-CT scans along with the bones. Fish were fixed in 4% PFA overnight at 4°C. The next day, the PFA was removed with a 1xPBS wash for 10 minutes, repeated 3 times. Fish were then transferred to 35% ethanol to be washed on a rocker at RT for 20 minutes, followed by a further wash of 50% ethanol for 20 minutes under the same conditions. Finally, 0.3% phosphotungstic acid stain (PTA) (Sigma Aldrich, 12501-23-4) prepared in 100% ethanol was added to the samples and left at RT for 24 hours. The following day, a series of ethanol dilution washes (70%, 90%, 95% and 100%) were performed for 30 minutes at room temperature with gentle rocking. The samples were stored in 100% ethanol at 4°C until scanning (Lin et al., 2018).

2.9.2. Micro-CT fish scanning

For adult fish bone visualisation, Skyscan 1172 desktop x-ray high-resolution microtomography was used. Fish were fixed in 10% neutral buffered formalin overnight at 4°C with rocking. When ready for scanning, the fish were wrapped in clingfilm to ensure a secure fit in the plastic moulds. Skyscan 1172 software was used to perform scans using the 0.5mm AI filter, medium camera binning 2000x1048, and

6µm pixels as constants for all scans. Fish skull scans take 15 minutes and full body scans take about 2 hours.

2.9.3. Micro-CT analysis and 3D construct generation

The scans were first processed using the software NRecon (Micro Photonics Inc.) to reconstruct cross-sectional slices from scanned angle projections. The brightness histogram was set to 0-0.16 and the ring artefact to 10. The reconstructed images were then uploaded to the analysis software CTAn (CT-Analyser, Bruker) and the region of interest was manually selected every 5-10 slices across the scans to limit the area analysed and remain as close as possible to the region of interest. This is essential due to the variation in the region of interest through the fish body. The tail has a much smaller region of interest compared to the belly. Eliminating irrelevant background, increases the quality and accuracy of the software measurements. A binary threshold was also manually selected to limit background noise while preserving the structure of interest. A task list of parameters was set; filtering of Gaussian blur in 3D space with a radius of 0.65, threshold based on the selected binary threshold as a low value and 255 as a standardised high threshold, despeckle to remove white speckles in 2D space for an area of less than 10 pixels and saving bitmaps as TIFF files for further analysis. Finally, 3D analysis, which includes all the baseline measurements to obtain total bone volume and bone mineral density. The software generates an Excel sheet with the specified measurements, and 3D constructs were then created using the same reconstructed images in CTvox (Bruker).

2.10. Bioinformatics analysis pipeline

To identify common risk genes between the two studied diseases, we combined available GWAS data acquired from both GWAS Catalogue (https://www.ebi.ac.uk/gwas/) and GWAS Central (https://www.gwascentral.org/). A bioinformatics analysis pipeline was designed to extend our knowledge of the identified genes by determining any known function (KEGG) (https://www.genome.jp/kegg/), pathway involvement (WebGestalt) (https://www.webgestalt.org/) and protein-protein interactions (STRING) (https://string-db.org/). Common SNPs were also detected using the same GWAS data analysed with Venny 2 (https://bioinfogp.cnb.csic.es/tools/venny/). Genes associated with the identified SNPs were determined using the NCBI 1000 genome browser (https://www.internationalgenome.org/). The pipeline is discussed in detail in **Chapter 3**.

2.11. Statistics and reproducibility

All statistics were generated using GraphPad Prism Software (Version 9). For single comparison graphs, a student t-test was performed, while a one-way ANOVA test was used for multiple group comparisons to account for the false discovery rate (FDR). Details of each statistical test are given in the figure legend with the exact calculated p-value. For accurate statistical comparison, each experiment was repeated three times, unless otherwise stated in the figure legend.

3. Common Risk Genes and Pathways underlying the Pathogenesis of OA and CVD

3.1. Common risk genes identified between OA and CVD

In the aim of unravelling the common risk genes between OA and CVD, and in order to understand more about the pathways in which they are involved and their indicated functions, a bioinformatic analysis pipeline was designed, shown in **Figure 3.1**. Data were retrieved from 133 osteoarthritis and 2884 cardiovascular disease GWA-studies documented in the GWAS Catalog (https://www.ebi.ac.uk/gwas) and the GWAS Central (https://www.gwascentral.org) databases as Excel sheets. The data were filtered using OpenRefine software (https://openrefine.org) to remove duplicates or errors. Common genes and SNPs between all data sets were determined using Venny 2.0.1 (https://bioinfogp.cnb.csic.es/tools/venny). 37 genes were identified as common between the two diseases and are shown in **Figure 3.4**, along with 3 common SNPs (**Figure 3.5)** For the identified common genes, KEGG (https://www.genome.jp/kegg/pathway.html) and Reactome (https://reactome.org) were used to assign their function, combined with literature searches to identify additional functions. **Figure 3.2** shows the identified common pathways and molecular functions, including TGF-β signalling, osteoclast differentiation, cytokine-cytokine receptor interactions, which have an impact on elastic fibres formation and extracellular matrix organisation. WebGestalt (http://www.webgestalt.org) was used to link candidate genes with associated diseases and pathways. Identified common risk genes were found to be highly involved in inflammatory bowel disease (IBD), tuberculosis (TB), leukaemia and pancreatic cancer, as well as several infection and inflammatory response pathways. To determine the baseline expression levels of candidate genes in different tissues, we used the FUMA GWAS online tool (https://fuma.ctglab.nl). **Figure 3.3** highlights the baseline expression pattern of all identified genes in different tissues. Purple rectangles indicate tissues of particular interest due to their abundant presence in the skeletal or cardiac structures. This analysis revealed a positive correlation between the expression pattern of *CTSK* and *TGFβ1* across almost all tissues. Identifying potential protein interactions between the common genes was another approach taken to decipher the link between these pathways. Therefore, protein-protein interactions between the identified common genes were explored using STRING (https://string-db.org). In this analysis, only evidence-based interactions sourced from experiments, databases and text mining with the highest confidence (0.900) were considered. The same protein-protein analysis was performed twice, on human and zebrafish data to observe any difference in protein interactions between the two organisms, but no difference in interactions was observed and the network model is shown in **Figure 3.4,** showcasing the 37 identified common risk genes.

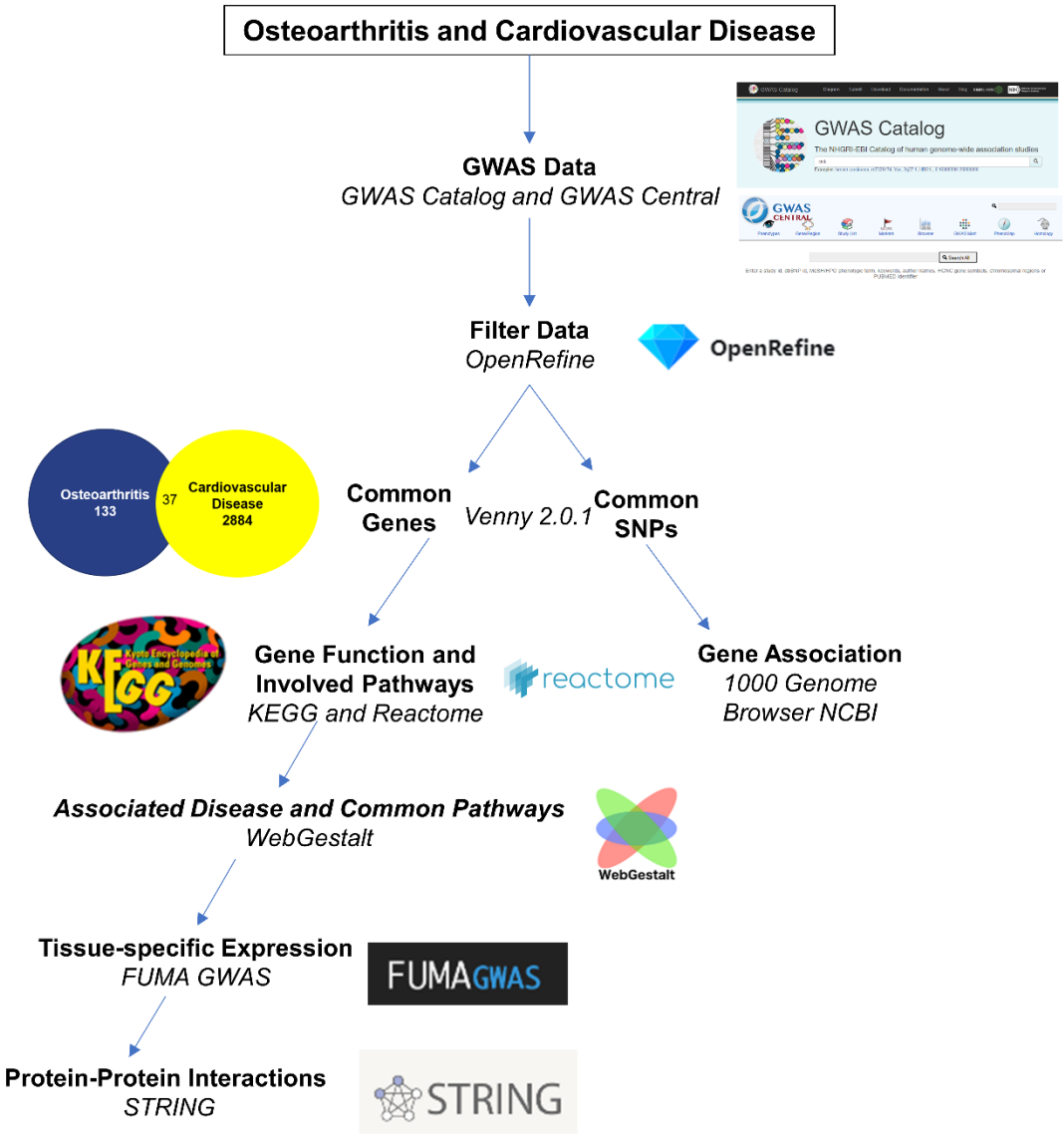


Figure 3.1: Bioinformatics workflow for the identification of common risk genes in OA and CVD retrieved from GWAS data

This pipeline is designed to identify common risk genes and understand more about their properties and the pathways in which they are involved. It starts with GWAS data obtained from the GWAS Catalog and GWAS Central. Venny 2.0.1 was used to identify the common genes and common SNPs. Two separate routes are followed for the genes and SNPs. For gene analysis, KEGG and Reactome were used to assign the function and identify common pathways. WebGestalt was used to link candidate genes with identified associated diseases and involved pathways. FUMA GWAS was used to determine baseline expression levels in different tissues. Protein-protein interactions are also examined using STRING to reveal other possible links. The other route follows the common SNPs and maps them to their candidate genes using the 1000 Genome Browser on NCBI.

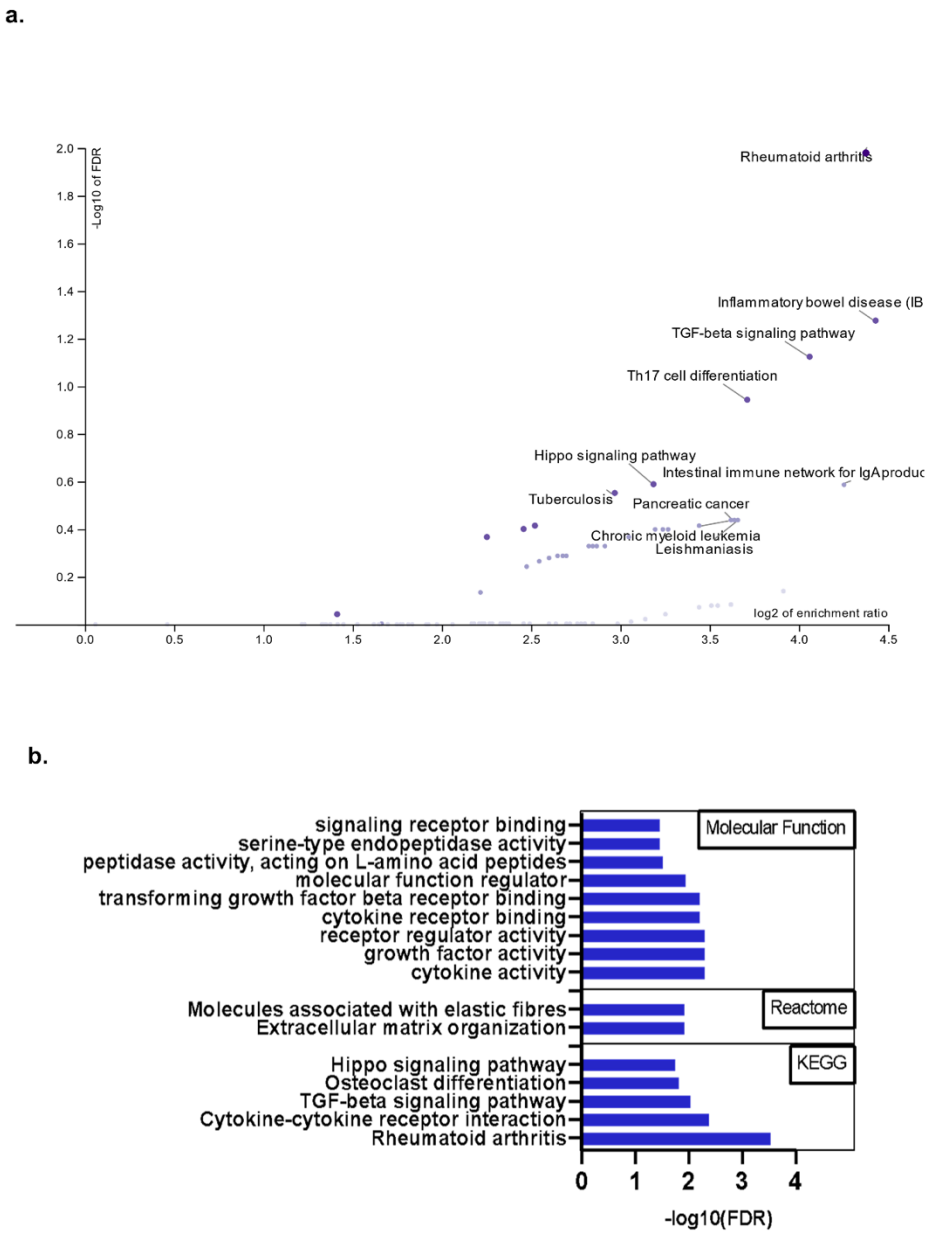


Figure 3.2: Associated pathways and molecular functions

(a) WebGestalt analysis of the identified common risk genes showing high involvement with the presented diseases, as shown by their high enrichment ratio (x-axis), and a high -Log10 of False Discovery Rate (FDR) (y-axis). This highlights a low likelihood of false discovery and an accurate association with the diseases, becoming more significant up the y-axis. Several identified common diseases are associated with inflammation and cancer (http://www.webgestalt.org). **(b)** Analysis performed using KEGG (https://www.genome.jp/kegg/pathway.html) and Reactome (https://reactome.org), showing high involvement of common risk genes in several receptor/cytokine binding pathways, TGF-β signalling and ECM organisation. Many are associated with rheumatoid arthritis and osteoclast differentiation.

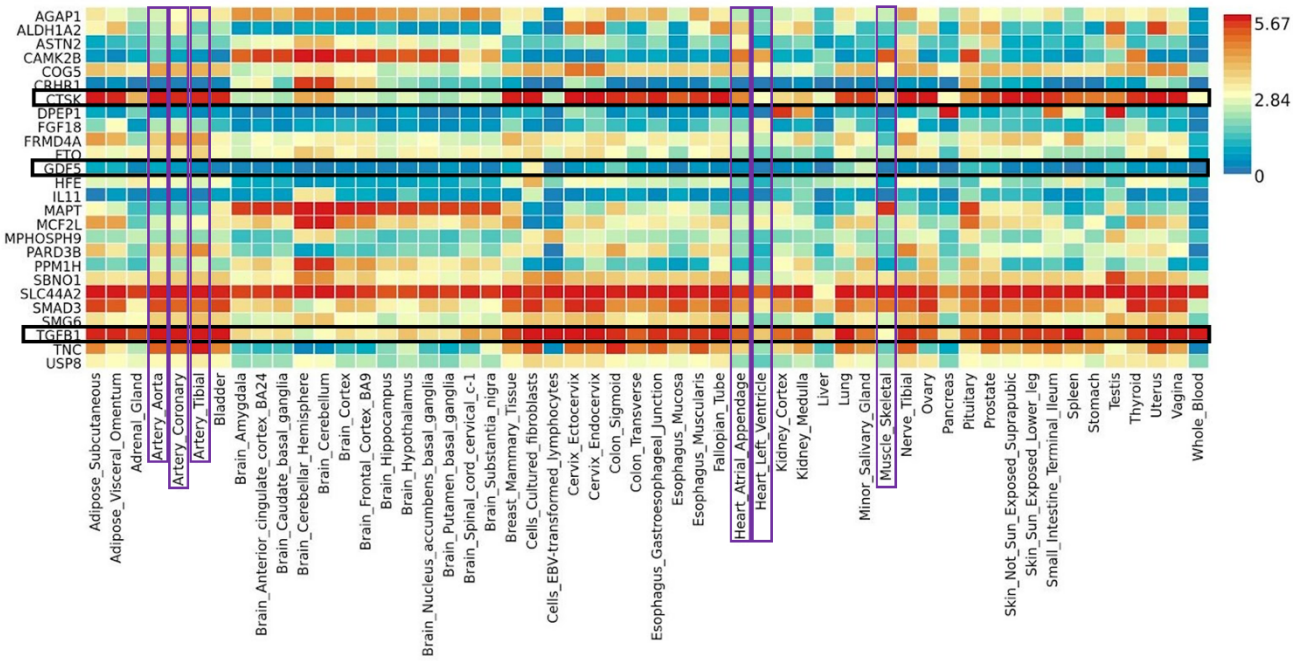


Figure 3.3: Baseline gene expression levels in different tissues

The figure shows the baseline expression levels of identified common risk genes in different tissues. Highly expressed genes appear in the red part of the scale, while low expression is shown in a blue gradient. Black rectangles highlight genes that have been studied at a functional level in this project. Purple rectangles highlight tissues of interest due to their association with the cardiovascular or musculoskeletal systems. *CTSK* and *TGFβ1* genes show a positive correlation in their baseline expression pattern across different tissues. Figure generated using FUMA GWAS (https://fuma.ctglab.nl/).

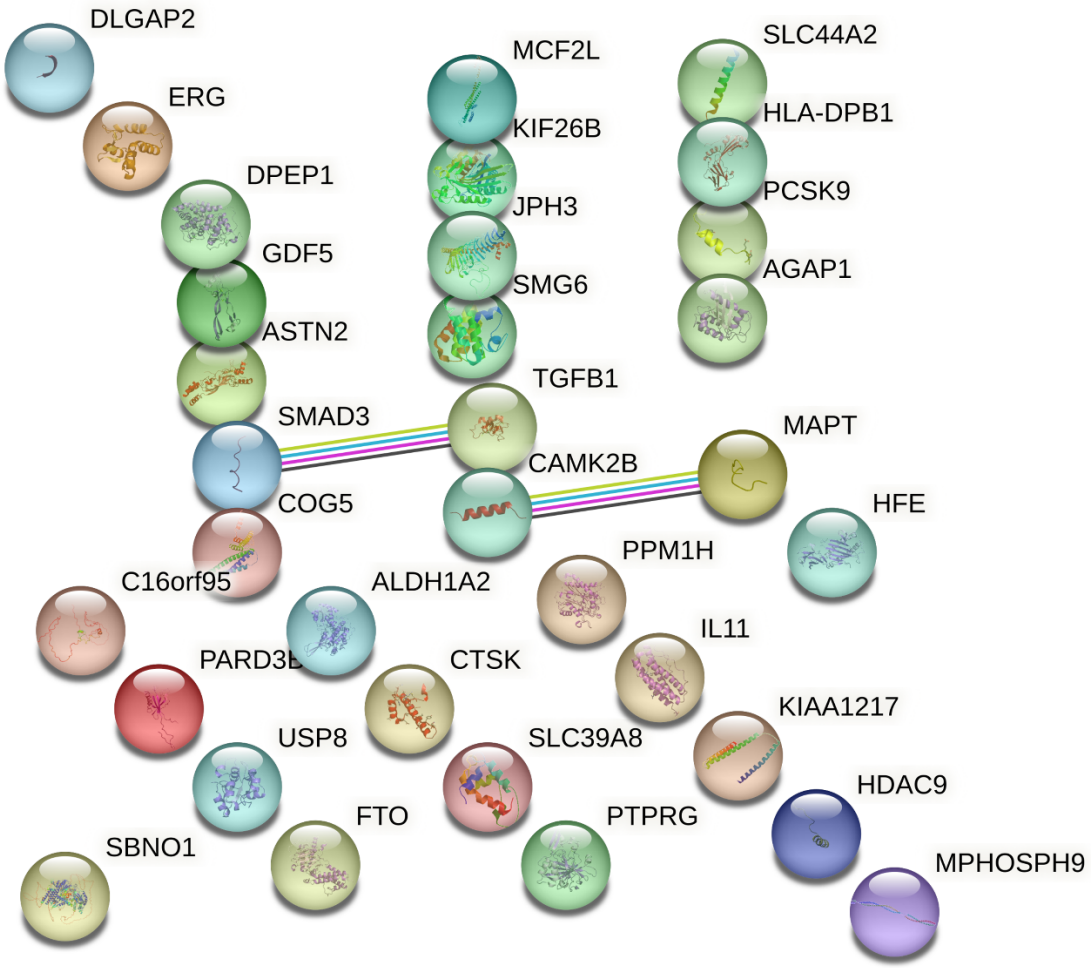


Figure 3.4: Protein-Protein interactions between encoded proteins of the 37 identified common risk genes

Figure highlights direct protein-protein interactions between all common risk genes identified between OA and CVD. Data were filtered to exclusively show evidence-based interactions with the highest confidence (0.9000) to rule out any unsupported data. Interactions between proteins are highlighted by coloured lines connecting their respective spheres. The length of the connecting lines does not correlate to the directness of the interaction. Each colour of the lines indicates different resources from which the interaction was detected. Figure generated using STRING (https://string-db.org/).

The most prominent interaction found is between TGFβ1 and SMAD3 via TGFβ receptors, which can phosphorylate SMAD3. This interaction is discussed further in **Chapter 6**. Other important interactions include a link between MAPT and CAMK2B. Microtubule-associated protein tau (*MAPT*) is the gene that encodes for the TAU protein. TAU is one of the two major pathogenic proteins involved in the pathogenesis of Parkinson's disease and Alzheimer's disease, **section 6.1.4**. The main functional role of TAU is the formation and stability of the microtubule cytoskeleton, particularly in neurons. Tau acts as a linker between axonal microtubules and neural plasma membrane components increasing the stability of the neuronal cytoskeleton (Cleveland et al., 1977). When found in an abnormal state, the protein can cause irregular clumps known as neurofibrillary tangles, which are accumulated misfolded proteins that localise within neurons or the brain, leading to the aforementioned associated diseases (Bloom., 2014; Zhang et al., 2018). In addition, MAPT has been found to be highly expressed in cardiomyocytes, smooth muscle cells and fibroblasts, suggesting a strong involvement in structural regulation that is not limited to the nervous system. Calcium/calmodulin-dependent protein kinase type II subunit beta (CAMK2B) is a kinase that is abundantly expressed in the brain and regulates several signalling pathways. It requires the binding of calcium to calmodulin, leading to autophosphorylation of the subunit. CAMK2B has been shown to regulate the formation of dendritic spines and synapses between neurons. It also regulates the organisation of the actin cytoskeleton, making it an essential element in neural plasticity (Koller et al., 2020). The protein also plays a role in skeletal muscle through its function in regulating the sarcoplasmic reticulum, a specific type of ER specialised in calcium regulation for muscle movement (Abisambra et al., 2013). CAMK2B and MAPT show a significant co-expression pattern as CAMK2B was discovered to phosphorylate TAU in Alzheimer's patients (Panda et al., 2003).

In terms of SNPs analysis, three SNPs, rs1126464 (Tachmazidou et al., 2019; Zhang et al., 2021), *rs13107325* (Tachmazidou et al., 2019; Pare et al., 2010) and rs11591147 (Nelson et al., 2019; Qiu et al., 2017) were found to be shared between OA and CVD. The 1000 Genome Browser (NCBI) is an online tool that generates libraries of human genetic variants based on sequencing data from thousands of people. It can therefore locate the SNPs that have been studied to their associated genes. Each of the three common SNPs identified was mapped to its candidate gene and found to be located on the *DPEP1*, *SLC39A8* and *PCSK9* respectively (**Figure 3.5**).

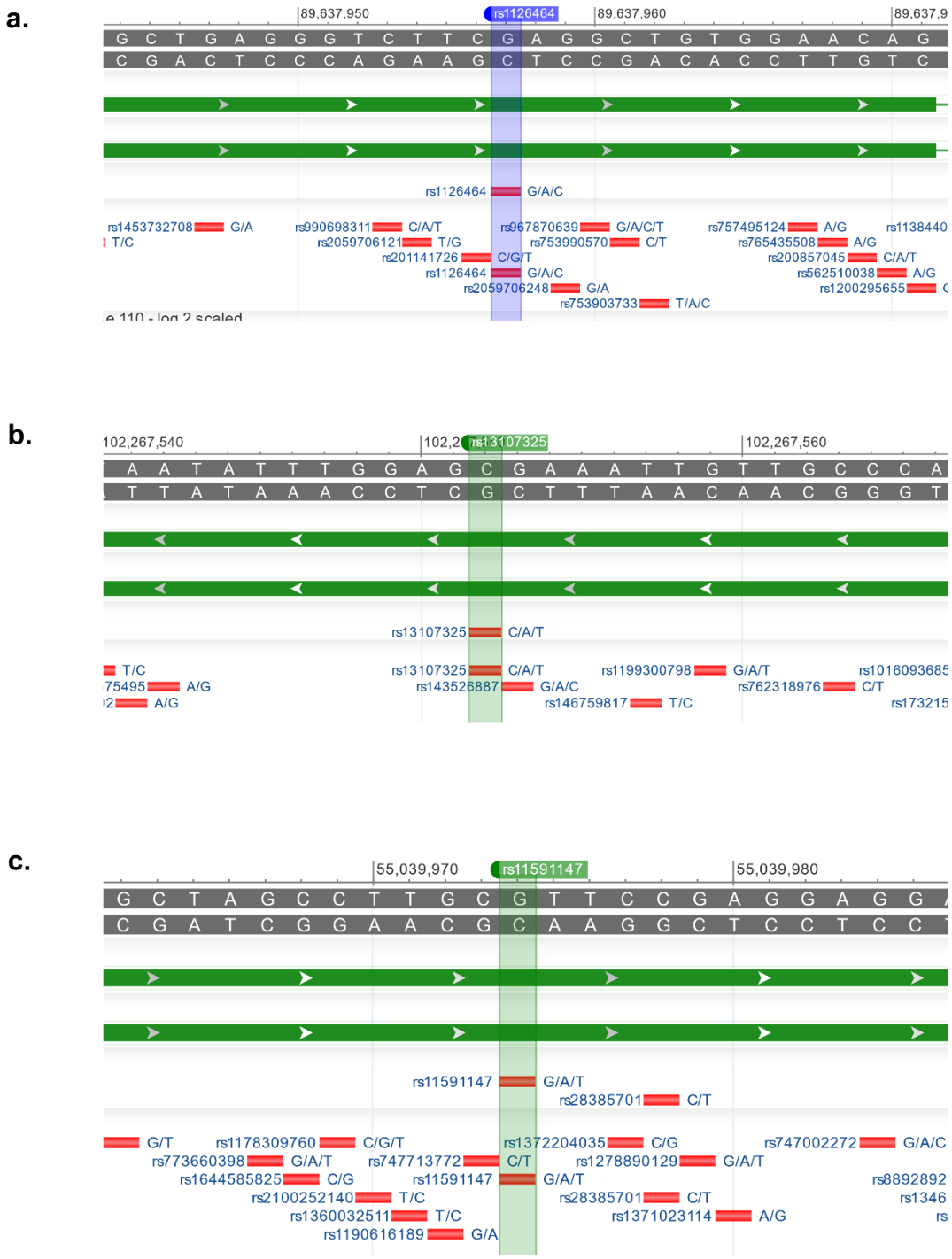


Figure 3.5: Two out of the 3 identified common SNPs are linked to zinc transportation

Following the initial step of identifying common risk genes and SNPs using OA and CVD GWAS data, 3 SNPs were found to be common between the two diseases and are presented in this figure. **(a)** rs1126464 located on the *DPEP1* gene. **(b)** rs13107325 located on the *SLC39A8* gene. **(c)** rs11591147 located on the *PCSK9* gene in humans. Associated genes were mapped using the NCBI 1000 Genome Browser (https://www.ncbi.nlm.nih.gov/genome/gdv/).

Furthermore, using murine cartilage microarray data generated in our laboratory from mouse hip injury samples, the fold change in expression level of the identified common genes was examined, 4 hours post injury. This information allows us to detect genes that are highly involved in the inflammatory response following injury or rather associated with tissue repair. **Figure 3.6** presents the results obtained. *Il11*, *Gdf5*, *Smad5*, and *Fgf18* show the highest increase in expression 4 hours post hip injury. Among the three genes studied in this project, to be described in detail in the following chapters, *Gdf5* was found to be highly upregulated post-injury with a 4.50-fold change. *Tgfβ1* showed a similar pattern with a 1.32-fold increase in its expression level after injury. In contrast, *Ctsk* showed a significant reduction in its expression after injury with a -1.78-fold change.

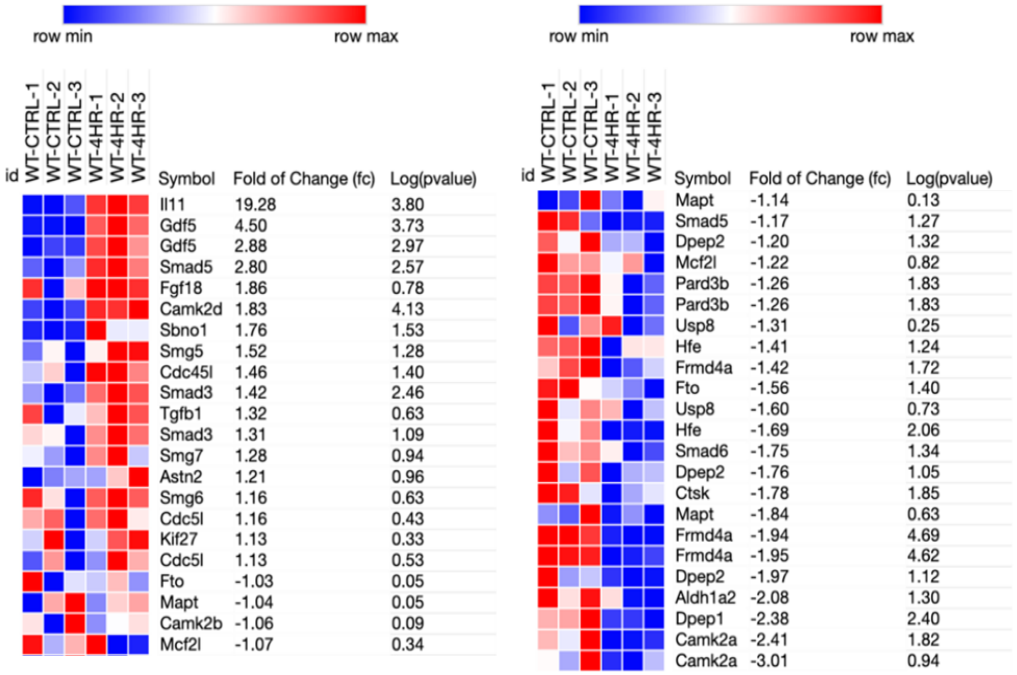


Figure 3.6: Heatmap of gene expression profile of candidate genes in murine cartilage before and after injury

The figure summarises the change in expression levels of identified common risk genes in healthy wildtype murine cartilage samples (WT-CTRL1-3) compared to injured murine cartilage (WT-4H1-3), 4 hours post-injury induction. Data obtained by interrogating the transcriptomic data from the microarray and highlighting gene overexpression through a gradient that goes up to bright red, while blue indicates a reduction in gene expression levels. *Il11*, *Gdf5*, *Smad5*, and *Fgf18* show the highest increase in expression levels 4 hours post injury. Among the genes studied in detail in this project, *Gdf5* shows a significant increase of 4.50-fold post-injury, along with *Tgfβ1*, which also shows an increase in expression of 1.32-fold. Conversely, *Ctsk* was downregulated by -1.78-fold at 4 hours post-injury. Figure was generated using Morpheus (https://software.broadinstitute.org/morpheus/).

3.2. Possible link to zinc transportation

Interestingly, two out of the three common SNPs found between OA and CVD were located on the genes *DPEP1* (rs1126464) and *SLC39A8 (rs13107325)*, **Figure 3.5**. Both genes display a role in zinc transportation. *DPEP1* encodes the enzyme dipeptidase 1, which among other similar proteins, is essential for the metabolism of glutathione via dipeptide hydrolysis, releasing amino acids that can be used for energy production or further metabolism (Eisenach et al., 2013; Cui et al., 2019). This protein requires zinc as a cofactor for its transport and activity (Kawasaki et al., 1993). Currently, there is limited information on the role of DPEP1, but it was found to be highly expressed in colorectal cancer samples when compared to healthy mucosa (Eisenach et al., 2013). This study established the elevation of DPEP1 as a reliable marker for colorectal cancer, but the exact mechanism of action remains elusive. Recently, a role in kidney disease has been identified, revealing an additional aspect of DPEP1's cellular functions. A variant of a single genetic locus, rs164748, was found to be associated with the development of kidney disease through ferroptosis due to a defect in iron trafficking and the locus was mapped to *DPEP1* (Guan et al., 2021). This highlights an additional role in iron transport and modulation.

SLC39A8, also known as ZIP8, is a member of the solute carrier family 39 (SLC39). It is an integral membrane protein recognised as a zinc transporter that regulates the influx of zinc from the extracellular matrix into the cytoplasm (Begum et al., 2002). Defects in this protein can be the major cause of a number of disorders due to the essential role of the zinc in numerous molecular processes. One example is acrodermatitis enteropathica (AE), a disease characterised by inefficient zinc absorption due to a mutation in the *SLC39A8* gene, resulting in a weakened immune system, skin rashes and hair loss (Lehneret et al., 2006). *SLC39A8* has also been implicated in several neurological disorders, including autism, but the mechanism of pathogenesis remains undetermined (Wahlberg et al., 2018). Schizophrenia has recently been linked to a mutant variant of *SLC39A8,* known as rs13107325, which was found to regulate zinc concentration, which affects dendritic spine density, and is commonly found in schizophrenia (Li et al., 2022). Another variant, A391T, which causes a significant decrease in serum manganese levels, has been reported to increase the risk of cardiovascular disease (Waterworth et al., 2010) and Crohn's disease (Li et al., 2016), while surprisingly having the opposite effect on Parkinson's disease by decreasing the risk of prognosis (Pickrell et al., 2016).

This observation sheds light on the importance of zinc and its precise regulation. It also adds a layer to our understanding of the underlying prognosis of chronic diseases caused by an alteration in the mineral concentration in the body. Several indicators and previous studies hinted to the presence of a link between zinc transportation defects and the studied chronic diseases (Huang et al., 2018; Li et al., 2021). Firstly, zinc is abundant in connective tissues, which makes up a large part of the joints and the heart.

Thus, this metal ion is essential for structural development as well as for later function. One study generated *Slc39a13* knockout mice, which is another zinc transporter belonging to the same SLC39 family. It has 67% similarity to the protein sequence of SLC39A8. The knockout mice showed defective development of chondrocytes, osteoblasts and fibroblasts in several connective tissues. This study also confirmed the link between zinc metabolism and the BMP/TGF-β signalling pathway (Fukada et al., 2008). In addition, another study looked at SLC30A3 and found it to be essential for MAPK signalling. All mentioned pathways have been identified as key regulators of the two diseases studied, as well as other chronic diseases (Zhang et al., 2021), and will be discussed further in **section 3.3**.

3.3. Common pathways between OA and CVD

Based on the performed bioinformatic analysis, in addition to literature searches, the following signalling pathways were found to be involved in OA and CVD. There are four common pathways that showed the highest correlation, which include the JAK/STAT (Mascareno et al., 2001; Malemud, 2018), MAPK/JNK (Muslin, 2008; Loeser et al., 2008; Sadoshima et al., 2002; Ge et al., 2017), PI3K/AKT/mTOR (Aoyagi and Matsui, 2011; Chen et al., 2012), and NF-κB (van der Heiden et al., 2010; Rigoglou and Papavassiliou, 2013) signalling pathways. In the following section, the consequences of the dysregulation of these pathways will be discussed, thus highlighting the potential theoretical molecular mechanisms underlying OA and CVD.

3.3.1. JAK/STAT pathway

The JAK/STAT signalling pathway is responsible for transducing signals from a wide range of cytokines and growth factors that control essential cellular processes such as proliferation, differentiation, migration and apoptosis (Rawlings et al., 2004). Primary activation is initiated by the binding of a ligand to JAK receptors (JAK1, JAK2, JAK3 and TYK2), which brings the receptor subunits into close proximity, forming either a homodimer if the ligand is erythropoietin or a growth factor (Renuald et al., 2003), or a heterodimer if interferons or interleukins have been bound to the receptor (O'Shea et al., 2012). Trans-phosphorylation occurs, allowing subsequent phosphorylation of proteins known as STATs, signal transducer and activator of transcription (Ivashkiv et al., 2004).

The pathway can be positively regulated by three mechanisms, providing a direct single pathway that translates external signals into a transcriptional response. Dysregulation of the JAK/STAT signalling pathway can lead to several diseases including cutaneous leishmaniasis when downregulated (Fernandez-Figueroa et al., 2016), and several types of leukaemia when present in an upregulated state (James et al., 2005; Ripoll et al., 2016). In osteoarthritis, as in rheumatoid arthritis, JAK signalling is hyperactivated, leading to an increase in the expression levels of *MMPs*, a decrease in apoptosis of inflamed synovial tissue and an increase in the apoptosis of chondrocytes (Malemud, 2018). Several

cardiac diseases have been found to have a genetic pathogenic basis linked to the JAK/STAT signalling pathway, such as myocardial ischemia (MI). Upregulation of the *ANG* gene caused by an increase in STAT signalling, contributes to the development of MI (Mascareno et al., 2001).

Regarding the suppression mechanism of the JAK/STAT pathway, three other factors have been found to negatively regulate the JAK signalling pathway. Suppressor protein groups are protein inhibitors of activated STATs (PIAS), which were found to inhibit the JAK/STAT pathway by inhibiting the phosphorylation of STATs (Niu et al., 2018), similar to PTPs, protein tyrosine phosphatases (Bohmer et al., 2014), and suppressors of cytokine signalling (SOCS) (Liau et al., 2018).

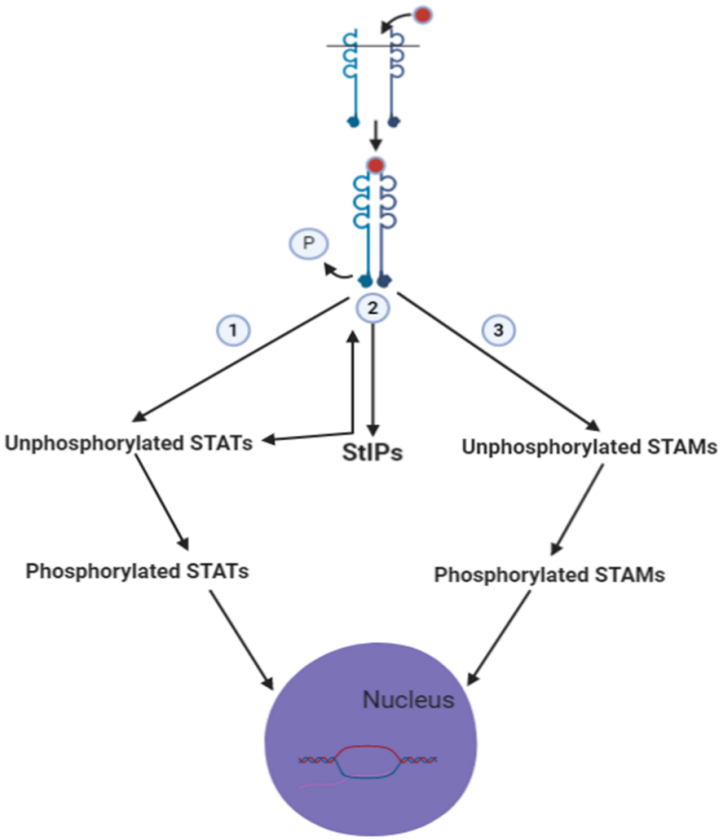


Figure 3.7: JAK/STAT regulatory pathways

Ligand binds to JAK receptor subunits, which form a dimer and release a phosphate. **1)** STAT proteins are phosphorylated at a tyrosine residue near their C-terminus. Phosphorylated STATs can now enter the nucleus**. 2)** StIPs are protein scaffolds that can bind to JAKs or unphosphorylated STATs and help in the phosphorylation process to facilitate the entry of STATs into the nucleus. **3)** The final step of the pathway relies specifically on JAK1 and JAK3 to phosphorylate STAMs, the phosphotyrosine proteins containing ubiquitin-interacting and SH3 domains. In the nucleus, STATs and STAMs can activate or repress target genes. Figure generated using Biorender (https://www.biorender.com/).

3.3.2. MAPK pathway

Mitogen-activated protein kinases (MAPKs) comprise of three families of protein kinases: extracellular signal-regulated kinase (ERK), p38 MAPK and c-Jun kinase (JNK). They regulate various cellular functions, such as cell cycle regulation, proliferation, differentiation, transformation and apoptosis (Zhang et al., 2002; Wada et al., 2004). Endothelial growth factors such as VEGF can activate the ERK signalling pathway and thereby regulate the cell cycle (Pedram et al., 1997). p38 MAP kinases transduce signals into intracellular responses via serine-threonine kinases that regulate gene expression, and subsequently many processes in the cell. It plays a role in the differentiation and survival of adipocytes, chondroblasts and cardiomyocytes, which has a direct effect on the progression of OA and CVD (Nebreda and Porras, 2000).

JNK is another member of the MAPK family (Li et al., 2015). Some cytokines such as TNF-a or IL-1β, along with other growth factors bind to extracellular receptors, which get phosphorylated and activate several proteins, leading to the activation of JNK 1 or JNK 2. Upon their activation, JNKs can phosphorylate the c-Jun transcription factor at two serine residues near the N-terminus (Weston and Davis, 2007). NF-KB transcription factors act as scaffolding proteins to assist in the activation of JNK proteins (Zhang et al., 2015). Our group has identified a direct link between the JNK pathway and aggrecan degradation in osteoarthritis. The signalling pathway RAF-6/TAK-1/MKK-4/JNK-2 pathway, which is induced by IL-1, increasing aggrecanase levels in the matrix and JNK2 knockout mice showed delayed progression of osteoarthritis in murine experimental models (Ismail et al., 2016).

The MAPK pathway was also found to be highly involved in the development and maintenance of the cardiovascular system. ERK1/2 and JNK are essential members of the MAPK family, which were identified for their role in cardiovascular development (Srivastava, 2006; Wagner et al., 2007). ERK1/2 have been shown to play a role in morphogenesis during early heart growth. Several FGFs and VEGFs signal through ERK1/2 and contribute to cardiomyocyte formation and valve assembly (Lavine et al., 2005; Rajasingh et al., 2007).

Jnk has also been implicated in early cardiac development through the non-canonical Wnt ligand, Wnt11 (Eisenberg et al., 1999; Pandur et al., 2002). *Jnk* has been shown to be activated downstream of Wnt11 to promote early cardiac cell differentiation and orientation for morphogenesis (Zhou et al., 2008; Gerits et al., 2007). In the context of cardiovascular disease, the overexpression of MEK1, which is an upstream activator of ERK1/2 signalling, has been shown to increase the expression of Ras in transgenic mouse models, contributing to the increase in cardiomyocyte hypertrophy (Bueno et al., 2000).

3.3.3. PI3K/AKT/mTOR pathway

The third identified common signalling pathway is PI3K/AKT/mTOR, which has an impact on two crucial processes that occur during the development of OA; ECM degradation and chondrocyte depletion in synovial joints (Li et al., 2013; Huang et al., 2021). This is due to the overactivation of the signalling pathway, leading to an increase in reactive oxygen species and pro-inflammatory cytokines leading to chondrocyte apoptosis (Facchini et al., 2011). PI3K/AKT overactivation can also increase the activity of matrix metalloproteinases (MMPs), particularly MMP-13, MMP-3 and MT1-MMP (Prasadam et al., 2013).

MMPs are known for their role in ECM degradation, and some are involved in the P3K/AKT/mTOR pathway through several cofactors, including Runx3, which is required for MMP-13 production (Mengsgol, 2001), IL-1β by inducing MMP-3 (Shakibaei et al., 2007), the IL-6 cytokine family (including IL11), which increases both MMP-1 and MMP13 (Litherland et al., 2008). Additionally, leptin, a hormone over-represented in obesity, which has been shown to increase the expressions of several *MMPs* (Hui et al., 2012). Similar activation of signalling pathways has been observed in some types of cardiovascular disease, such as cardiac fibrosis, where IL-1β, IL-6 and TNF-a protein levels are increased, thereby increasing the activity of PI3K/AKT/mTOR signalling pathway (Atefi et al., 2011; Song et al., 2010; Comstock et al., 1998; Yamauchi-Takihara et al., 1995).

3.3.4. NF-kB pathway

The NF-kB pathway is involved in several processes including proliferation, differentiation, apoptosis, immune regulation and ageing (Guttridge et al., 1999; Khandewal et al., 2011; Li et al., 2002; Salminen et al., 2008). It is found in an over-activated state in some cases of OA and CVD (Goldring and Marcu, 2009; Tilstra et al., 2011). NF-kB is normally repressed by IkB proteins, which bind to the NF-kB members and inhibit their effect. When IkB kinases (IKK) are activated, they cause the subsequent phosphorylation of IkB proteins. Following phosphorylation, IkB proteins get degraded and are therefore unable to bind NF-kB to inhibit their migration to the nucleus. Once the NF-kB complexes are free, they enter the nucleus to begin regulating target genes that express inflammatory-related molecules (Yan et al., 2008; Wan et al., 2010; Snow et al., 2016). Induced genes include *MMPs* (Wu et al., 2011), *TNF-a* (Tian et al., 2005), *IL-1β* (Greten et al., 2007) and other cytokines that spread to further activate other inflammatory signalling pathways. This suggests that NF-kB may affect both diseases through its structural regulation or inflammatory regulatory properties.

3.4. Selected genes for functional genomic analysis

Of the 37 identified common risk genes, 3 genes were selected for further functional genomic analysis by targeting them in zebrafish embryos using CRISPR-Cas9. *CTSK, GDF5 and TGFβ1* were specifically chosen due to their involvement in other common chronic diseases in addition to OA and CVD. The three genes have all been found involved in Alzheimer's disease (Dauth et al., 2011; Wu et al., 2021; Bernhardi et al., 2015), Parkinson's disease (McGlinchey et al., 2020; Goulding et al., 2022; Comino et al., 2022) and various types of cancer, including prostate cancer (Wu et al., 2022; Conti et al., 2021; Shiota et al., 2021). Based on the FUMA GWAS analysis, *CTSK and TGFβ1* showed a positive correlation pattern across most tissues, suggesting a direct linkage between the two genes. Additionally, *GDF5 and TGFβ1* are both members of the TGFβ superfamily. They are known to function simultaneously within the BMP and TGFβ signalling pathways, which indicates the presence of a link between the two genes. Studying these genes would therefore cover a range of common chronic diseases in humans, allowing us to consider their function in a broad sense as key regulators of multimorbidity. It will also allow us to detect common patterns and common pathways behind the pathogenesis of multimorbidity, which would allow us to explore the potential of drug targeting of key regulators and limiting polypharmacy.

3.5. Discussion

Using 133 osteoarthritis and 2884 cardiovascular disease GWAS obtained data, 37 risk genes were identified to be common between the two diseases. Through further research following the generated bioinformatic analysis pipeline and pathway analysis, candidate genes were found to exhibit two main functional roles. First, a structural role in tissue maintenance and extracellular matrix was identified based on the high involvement of the analysed genes in the TGF-β and Hippo signalling pathways, both of which are essential for cell proliferation, differentiation and apoptosis (Moustakas et al., 2002; Harvey et al., 2003; Jia et al., 2003). Additionally, genes have been found to be critical for osteoclast differentiation, elastic fibre formation and extracellular matrix formation and degradation. The second major function identified is in inflammation and the regulation of the inflammatory response. Candidate genes exhibit essential functions in regulating cytokine-cytokine interactions during an inflammatory response. They were also found to be highly associated with rheumatoid arthritis, a disease recognised for its pathogenic basis that leads to inflammatory and structural disruptions (Silman and Pearson, 2022; Kurko et al., 2013; Dedman, 2020).

The analysis pipeline also revealed a positive correlation between the expression levels of *TGFβ1* and *CTSK*, genes selected for functional genomic analysis. Based on the FUMA GWAS, which shows the baseline expression levels in different tissues, both genes showed a similar pattern in most tissues. They were highly expressed in heart tissues and moderately expressed in skeletal muscle cells suggesting a critical role in both systems, which was concluded by our experimental work on these two genes and will be discussed in detail in the following chapters. Additional interactions were also revealed by STRING analysis of direct protein-protein interactions. The analysis revealed links between TGFβ1 and SMAD3, which have been studied and discussed extensively. It also revealed interactions between MAPT and CAMK2B, which have been previously implicated in Alzheimer's disease (Panda et al., 2003). However, both genes encode for proteins that are highly functional in the musculoskeletal cells (Abisambra et al., 2013) and the vascular smooth muscles (Zhang et al., 2018), making them good candidates for further investigation in the context of OA and CVD, as well as ageing multimorbidity. It should be noted that this interaction does not currently show a direct effect on bone maintenance to provide a direct link to OA, but MAPT is known for its role in microtubule stability within muscle fibres (Goedert et al., 1989). Therefore, a possible link may be mediated by muscle-bone crosstalk. The involvement of CAMK2B in both diseases is an observation that deserves attention, especially as CAMK2B is a calcium-dependant protein kinase and therefore its pathogenic effect could be a result of mineral dysregulation in calcium signalling.

The following obtained result also flagged the importance of cellular metal ion regulation, which when dysregulated can manifest in an array of diseases including OA and CVD. When the identified single

nucleotide polymorphisms (SNPs) from the GWAS data of both diseases were examined, 3 SNPs were found to be common, two of which are associated with zinc transportation. The SNPs were mapped to the *DPEP1* and *SLC39A8* genes, which either utilise zinc as a cofactor, or regulate its transport from the extracellular matrix to the cytoplasm. Recently, DPEP1 was also found to regulate iron trafficking in kidney cells (Guan et al., 2021). Taken together, the separately obtained data highlight the importance of trace metal ion regulation through concentration or transportation on several systems, which may underly the increase in multiple chronic diseases within the same individual, multimorbidity.

Using microarray data generated in our laboratory, the change in gene expression levels of the identified common risk genes was determined. *IL11*, *GDF5*, *SMAD5*, and *FGF18* showed the highest increase in expression levels 4 hours after hip injury in mice. This may highlight their role in tissue repair post injury, or a role in the inflammatory response, or both.

Of the 37 identified common risk genes, *CTSK, GDF5* and *TGFβ1* were selected for further functional genomic analysis using CRISPR-Cas9 gene-targeting approaches in zebrafish embryos, as described in **section 1.8**. These genes were specifically selected based on the correlation in their expression patterns detected in the FUMA GWAS analysis, specifically *CTSK* and *TGFβ1.* In addition, their pathogenic involvement in other prevalent chronic diseases, along with OA and CVD, renders them remarkable candidates to study for the investigation of common pathways underlying multimorbidity. Moreover, the varying significant changes in their expression levels observed in the murine hip injury microarray data, suggest a possible role of these genes in tissue repair, or the inflammatory response. Therefore, targeting those genes as part of a functional genomic analysis could potentially reveal some of their cellular functions. Finally, the designed bioinformatic analysis pipeline revealed a high correlation of four pathways between the two diseases, from the identified common risk genes. The four pathways include the JAK/STAT (Tang et al., 2017; Godoi et al., 2023), MAPK/JNK (Lee et al., 2018; Thouverey and Caverzasio, 2015; Zhang et al., 2019; Zaidi et al., 2010; Tian et al., 2019), PI3K/AKT/mTOR (Chen et al., 2021; Hinoi et al., 2015; Zhang et al., 2013), and NF-kB (Krum et al., 2011; Bitzer et al., 2000; Freudlsperger et al., 2013) signalling pathways. An association has been found between the identified pathways and the selected genes, which justifies performing the functional genomic analysis on these genes, in the aim of unravelling the common underlying mechanisms behind multimorbidity. The following chapters will explore the roles of *ctsk*, *gdf5*, and *tgfβ1a* genes in the musculoskeletal and heart structural development in zebrafish.

4. Cathepsin K is Essential for Early Skeletogenesis and Heart Development

4.1. Introduction

Cathepsin K (CTSK), also known as CTSO2, PYCD and PKND, is a member of the papain family of lysosomal cysteine proteases, which consists of more than 15 different proteases, mainly found in the endosomal and lysosomal regions (Shi et al., 2000; Yadati et al., 2000). The family exhibits a range of molecular functions including ECM regulation (Buck et al., 1992; Everts et al., 1996), antigen presentation (Shi et al., 2000), in addition to other regulatory functions in the immune system (Asagiri et al., 2008). The encoded enzymes are synthesised in an inactive form due to an endogenous inhibitor present in the pro-domain that regulates the catalytic domain within the protein. The protein is activated following the cleavage of the catalytic peptide (Turk et al., 1997). Similarly, Cathepsin K is activated, exhibiting critical roles in bone remodelling, resorption and osteoclast function through the digestion of essential bone matrix proteins, such as collagen, osteonectin and osteopontin (Troen., 2004; Bonnet et al., 2017; Yoshida et al., 2018). By studying its complex mechanism of action, CTSK has been identified for its involvement in various processes including protein metabolism (Invest et al., 2003) and homeostatic regulation of different body systems (Zaidi et al., 2018). CTSK has also been discovered to play a role in the development of various systems, such as the respiratory system (Buhling et al., 2004), the nervous system (Dauth et al., 2020), the cardiovascular system (Hua et al., 2015; Guo et al., 2018), and the musculoskeletal system (Wilson et al., 2009).

4.1.1. Activation and regulation of CTSK

CTSK precursors are synthesised in the endoplasmic reticulum and then are transported to the Golgi apparatus where they bind to a high-mannose oligosaccharide chain, mannose-6-phosphate (M-6-P). M-6-P is considered an essential lysosomal enzyme signal for CTSK transport to lysosomes via clathrin-coated vesicles (Alimena et al., 1988). CTSK is mainly secreted by osteoclasts in their active state (Kirschke et al. 1995). However, subsequent research revealed that osteoblasts and osteocytes also secrete CTSK (Mandelin et al., 2006). Once secreted, it is activated either through the activity of other proteases or autocatalytically by cleaving its own propeptide, activating its enzymatic function (McGlinchey et al., 2015).

The *CTSK* gene can be activated through RANK ligand (RANKL) binding to the RANK receptor via the NF-kB pathway (Troen, 2006). This pathway activates the pro-osteoclastogenic factor nuclear factor of activated T-cells (NFATc1), which translocate to the nucleus within osteoclasts, activating *Ctsk*

(Balkan, 2009). Other stimulatory factors can also activate the same pathway, including some interleukins, TNFα, vitamin D and parathyroid hormone (Troen, 2006). TGF-β signalling is another pathway that has a direct effect on the regulation of *CTSK* expression. It has been shown to be essential for the processing of pro-CTSK into active CTSK through the activation of C4-S GAGs (a collagen IV-derived glycosaminoglycans that play an important role in the development and remodelling of tissue membranes), which cleaves the endogenous inhibitor region of CTSK, thereby increasing the activity of CTSK (Flanagan-Steet et al., 2018).

Other conditions can affect the activation of CTSK, such as the pH levels, which must remain acidic (Christensen and Shastri, 2015). To achieve this acidic state, an acidified milieu layer is formed around osteoclasts to allow for CTSK secretion and activation (Takito et al., 2018). When stored in lysosomes, the pH must remain below 5 for optimal regulation (Turk et al., 2012). Mechanical loading has also been found to stimulate the expression of osteoblastic *Ctsk*, as well as *Ctsk* in osteocytes (Bonnet et al., 2018).

Region-specific expression patterns of *CTSK* have been discovered where inflammatory cytokines such as IFN-Y, IL6 and TGF-a can induce the expression of *CTSK* from macrophages (Kamolmatyakul et al., 2001; Li et al., 2019; Singh et al., 2021). Meanwhile, TGF-β was found to inhibit *CTSK* in fibroblasts (van der Brule et al., 2005).

4.1.2. Cathepsin K cellular function

Cathepsin K was first identified as a collagen I degrading protease (Garnero et al., 1998). Collagen I is an abundantly present protein that makes up about 90% of the bone matrix. However, CTSK was later found to also degrade collagen II (Kafieneh et al., 1998), emphasising its essential role in bone resorption. CTSK can also cleave other matrix-related proteins to optimise bone resorption, including the matrix metalloprotease 9, MMP9 (Christensen and Shastri, 2015). It has therefore been identified as a catabolic factor in bone homeostasis (Costa et al., 2011). However, further studies revealed the role of cathepsins as anabolic factors in some cases, depending on the cell type in which they are expressed (Durdan et al., 2022). Wenqian Fang et al., 2019 showed that targeting *Ctsk* in mice may lead to an increase in cardiac fibrosis and cause cell death post myocardial infarction, highlighting a protective role of *Ctsk* as when impaired or inhibited, an increase in cardiomyocyte death was observed.

4.1.3. Skeletal diseases associated with Cathepsin K

Mutations in *CTSK* can lead to a heritable bone disorder known as Pycnodysostosis. First identified by Gelb in 1996, it was classified as an autosomal recessive disorder caused by a *CTSK* mutation that results in a loss of function of the enzyme, which prevents the protein from degrading collagen I. This

leads to abnormal musculoskeletal development with an overall increase in bone mineral density, leading to bone hardening and increased susceptibility to fractures, facial and skeletal deformities, and dwarfism (Gelb et al., 1996). Carriers usually suffer from a variety of progressive symptoms including dental problems and cavities, scoliosis due to the deformed vertebral column, and obstructive sleep apnoea due to disruptions in the upper airway. This disease was investigated in mice, where knocking out *Ctsk* resulted in phenotypes similar to those observed in patients with pycnodyspstosis. Knockout mice developed osteopetrosis due to a reduction in bone matrix resorption and showed an overall increase in bone mass (Saftig et al., 1998). Conversely, overexpression of *Ctsk* showed an increase in bone turnover rate (Kiviranta et al., 1992). More recently, a study revealed a mechanism of articular cartilage protection in *Ctsk* homozygous knockout mouse models. The mutants had delayed progression of OA due to a reduction in bone turnover in the already degrading cartilage (Sokki et al., 2018).

Changes affecting CTSK expression or activation leading to a disruption to its finely tuned homeostatic role can lead to osteopetrosis or osteoporosis. Osteopetrosis is characterised by a remarkable increase in bone mineral density due to the reduction of resorption in the absence or reduction of CTSK activity, leading to larger more brittle bones (Lotinun et al., 2013). Conversely, the abundant presence of CTSK in its active state can lead to osteoporosis, which is a condition characterised by a reduction in total bone mass and density due to the increased resorptive activity of CTSK during bone remodelling (Logar et al., 2007).

Regarding its disease-specific expression, OA patients show increased *CTSK* expression levels in cartilage tissue, synovial fluid and serum when compared to healthy individuals (Logar et al., 2007). The increase in *CTSK* expression shows a direct correlation with the severity of OA (Konttinen et al., 2002). Furthermore, an upregulation of *CTSK* mRNA expressions has also been observed in osteoarthritic cartilage of mice, suggesting an enzymatic role in the pathogenesis of osteoarthritis (Morko et al., 2004). On the other hand, microarray data obtained in our laboratory from murine cartilage, 4-hours post a scalpel-induced minor hip injury, suggest a significant decrease in *Ctsk* expression of -1.78-fold change.

A function of CTSK was discovered in patients with periodontal disease (PD). Periodontitis is a localised chronic gum infection that results in serious damage to the gums and bone surrounding the teeth. Using microarray data, qPCR and immunohistochemistry, *CTSK* was found to be highly expressed in PD patients suggesting an essential role of the protease in osteoclast differentiation through a paracrine signalling response (Heo et al., 2021). Although *CTSK* has been discovered and extensively studied for decades now, there remains several gaps in our knowledge regarding its mechanism of action. Several upstream factors that activate *CTSK* expression have been identified, yet the mechanism behind *CTSK* downregulation or inhibition to attain bone homeostasis remains unknown. Limited

knowledge is currently available regarding the interactions between CTSK and other cathepsins, as well as, the post-translational modifications undertaken following transcription, which could provide clues to the negative regulation of the gene. In addition, most research on CTSK has focused on its role in bone development and regulation, with little reference to its impact on non-skeletal tissues. Several studies have pointed to a notable role of *CTSK* in the respiratory and cardiovascular systems suggesting a direct involvement of the gene, but the mechanism remains elusive.

4.1.4. Cardiovascular diseases associated with Cathepsin K

Cathepsin K has been found to be highly expressed in several cardiovascular diseases. In patients with chronic heart failure (CHF), high levels of CTSK were found in serum samples (Zhao et al., 2015). High levels of CTSK were also observed in the plasma of patients with coronary heart disease (CHD) and myocardial infarction (MI) (Fang et al., 2019), atherosclerosis (Dubland and Francis, 2015), hypertension (Cheng et al., 2006) and cardiac hypertrophy patients (Hua et al., 2013).

Atherosclerosis is a cardiovascular disease characterised by the accumulation of plaque within the arteries. High levels of CTSK were detected in atherosclerosis patient samples, suggesting a significant contribution to the development or progression of the disease. Such elevation result in an increase in the degradation of collagen and other ECM related proteins, leading to plaque formation and accumulation in the arteries (Dubland and Francis, 2015).

Furthermore, aortic aneurysm (AA) has also been found to be associated with elevated CTSK levels (Zhao et al., 2015). In this condition, the aorta becomes enlarged and more fragile, which may lead to fatal tears within the aorta. Studies in *Ctsk* knockout mice demonstrated the role of CTSK in AA progression by inducing smooth muscle cell apoptosis, elastin degradation and T-cell proliferation, suggesting the use of a CTSK inhibitor as a promising treatment (Sun et al., 2011). CTSK protein levels are found to be elevated in patients with cardiovascular disease. Therefore, Hua et al., 2015 investigated the correlation between ageing and the development of cardiovascular disease in mice. Mice exhibited the same elevated levels of CTSK, along with several cardiac abnormalities, such as changes in heart morphology, contractility, and calcium levels, which could lead to heart failure. In this study, *Ctsk* knockout mice were generated and showed a significant attenuation to heart failure progression when compared to wildtype aged mice. *Ctsk* knockout also displayed a cardioprotective role in diabetic mice, and significantly reduced cardiac oxidative stress (Guo et al., 2017).

The role and effect of *CTSK* is not only limited to its expression levels in tissues; precise localisation of the protein is also essential, and impairments in this process can lead to a variety of diseases. One example is congenital heart defects (CHD), where CTSK mislocalisation alters TGF-β signalling,

leading to myocardial and valvular deformities. These adverse effects were attenuated through the inhibition of CTSK, restoring the heart function of CHD mouse models (Lu et al., 2020).

To date, elevated levels of CTSK have been detected in several types of cardiovascular disease, but the molecular mechanism behind this observation remains to be determined. Available studies that generated *Ctsk* knockout mice models to study the heart examined the knockout effect in relation to a cardiovascular disease rather than its intrinsic role in early development. Therefore, studying the effect of targeting *CTSK* on the structural development of the heart may provide additional information about the nature of the function displayed by the encoded protein.

4.1.5. Cathepsin K expression pattern in zebrafish

In zebrafish, *ctsk* gene expression starts as early as 5.25hpf and is essential for subsequent early developmental stages such as segmentation, pharyngula and hatching (Thisse et al., 2004). Using In situ hybridization, Ctsk has been found present in tissues associated with various systems including the cardiovascular, digestive, nervous, muscular, respiratory and sensory systems. In musculoskeletal tissues, it is highly expressed in the bone, cartilage and skeletal muscle, particularly in osteoclasts, and at a lower level in chondrocytes (Thisse et al., 2004; Petrey et al., 2012). Thus, zebrafish provide a suitable model to further investigate the functional role of *ctsk* in the musculoskeletal system and heart development in zebrafish.

4.2. Hypothesis and aims

GWAS have identified *CTSK* as a potential marker for osteoarthritis (Tachmazidou et al., 2019; Boer et al., 2021) and cardiovascular disease (LeBlanc et al., 2016). The gene has also been identified as a biomarker for other chronic diseases, including prostate cancer (Wu et al., 2022) and type-2 diabetes (Saxena et al., 2007). Therefore, I hypothesise that targeting *ctsk* would have a remarkable impact on the skeletal and heart structural development in zebrafish and may provide insights regarding the gene function in other systems, and potentially multimorbidity.

In this chapter, I aim to:

- Investigate the impact of targeting *ctsk* using CRISPR-Cas9 mutagenesis on early skeletogenesis in zebrafish. The focus will be on the jaw cartilage structure, as there are observed similarities between zebrafish jaw cartilage and common osteoarthritis joints. Also, the synovial like joint present within the zebrafish jaw is the only joint structure in zebrafish that experiences mild mechanical loading as part of the feeding process. Mechanical loading is a known contributing factor to the progression of osteoarthritis in human joints, therefore, studying the jaw specifically would make the study more comparable to the OA observed in humans.

- Examine the effect of targeting *ctsk* on the heart development in the *Tg(myl7:lifeActGFP)* transgenic line, highlighting the heart in green.

- Test the swimming efficacy of *ctsk* targeted larvae and adult fish, to determine any alterations in the skeletal system's functionality caused by the induced mutations.

- Assess the ongoing necessity for *ctsk* in bone maintenance and later skeletal development by conducting micro-CT scans of adults crispants.

4.3. Results

4.3.1. Ctsk protein domain and selected CRISPR guide targets in zebrafish

CTSK is located on chromosome 1: 150,796,208-150,807,434 in humans and chromosome 16: 29,492,749-29,502,787 in zebrafish. In both species it consists of 8 exons. The gene consists of 329 amino acids (333aa in zebrafish) which are translated to form the enzyme with a V-shaped catalytic diad cysteine-histidine active site (McGrath et al., 1997). The protein alignment performed using the Ensembl genome browser shows 64.5% similarity between the human and zebrafish Ctsk encoded proteins. The most highly conserved exons with available PAM sites for CRISPR-Cas9 targeting, are exons 4 and 5. Therefore, two CRISPR RNA guides were designed to target both exons for mutagenesis and are shown in **Figure 4.1**.

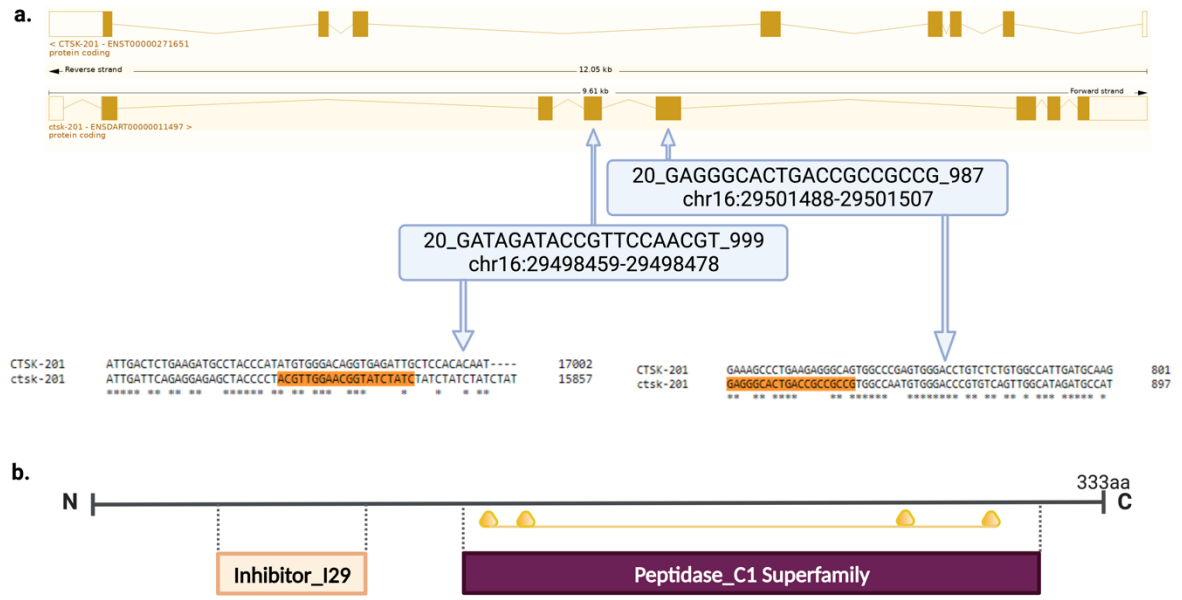


Figure 4.1: Cathepsin K gene composition and protein domain

(a) *CTSK* is composed of 8 exons in both human (top) and zebrafish (bottom). The blue rectangles highlight the location of selected CRISPR target guides in zebrafish *ctsk* with their exact position and the targeting score guide 1= 999 and guide 2= 987 (blue rectangle, left) based on the USCS genome browser tool (UCSC, 2019, https://genome.ucsc.edu/). **(b)** Zebrafish *ctsk* encodes a 333 amino acid protein with two conserved subunits, Inhibitor_I29, which prevents substrate from binding to the active site and can only allow protein activation through the interaction with a secondary peptidase or via an autocatalytic cleavage. The second subunit is a Peptidase C1 unit, which encodes for the lysosomal enzyme activated following the cleavage of inhibitor_I29. Yellow triangles represent the four active sites present in the encoded protein. Figure generated using Biorender (https://www.biorender.com/).

4.3.2. Abnormal cartilage development in *ctsk* targeted zebrafish larvae

To investigate the effect of targeting *ctsk* on early skeletogenesis, single cell zebrafish embryos were injected with two CRISPR-Cas9 guides (**Figure 4.1)** targeting exons 4 and 5. Both exons are conserved across the two species and have pre-identified high-scoring CRISPR targets based on UCSC data and were therefore selected for injection into single cell embryos. 400pmol of *ctsk* guide RNAs were injected into each embryo along with Cas9 protein. The efficiency of the designed guides was tested in a trial injection followed by genotyping using CRISPR-STAT to confirm the disruption of the wildtype allele peak, with the expected size of 306.4. Loss of this peak indicates successful targeting of the *ctsk* gene. Larvae were allowed to grow and imaged at 3dpf using a stereomicroscope to examine the general morphology of crispants compared to wildtype controls. Lightsheet imaging was also performed to visualise the jaw cartilage structure using the *Tg(Col2a:mCherry)* transgenic line, which labels collagen II with red fluorescence (**Figure 4.2**). Our analysis began at 3dpf, the stage at which the jaw cartilage structure becomes visible within the *Tg(Col2a:mCherry)* transgenic line.

The *ctsk* targeted larvae exhibit a distorted overall appearance (**Figure 4.2 C&D**), manifested by a curved and stunted body length compared to the control group (**Figure 4.2 A&B**). Additionally, the cranial structure appears to be disproportionally smaller with reduced eye size and a retracted or underdeveloped jaw in comparison to the control larvae. While these were the most common abnormalities found, extreme variability in phenotype was also observed, as demonstrated in (**Figure 4.2 C&D**) and (**Figure 4.3 C&D**). Starting from 3dpf, there were clear deformities observed in the jaw cartilage of *ctsk* crispants. These deformities included misaligned smaller Meckel's cartilage, non-developed or partially developed synovial joint (**green arrow, c&d**), irregular paraquadrate cartilage (**yellow arrow**) and an array of abnormalities to ceratohyal cartilage orientation. The latter appeared less arched and bent (**c&d**). Due to variations in the acquired phenotype, conducting comparable measurements in the larvae posed a challenge. Therefore, we opted to classify the observed phenotypes based on the severity of their deformities in general. At 5dpf, the phenotype persisted, and additional abnormalities were apparent due to the continuous rapid development of the musculoskeletal system. When examining the general morphology of the larvae, spinal curvature remained noticeable, and the swim bladder failed to develop in 66.6% of the targeted larvae (**Figure 4.3 C&D**). The same cartilage structures exhibited comparable abnormalities to those of targeted larvae at 3dpf. Additionally, at this stage, the basihyal structure (**yellow circle**) developed and was found to be smaller in size with a deformed shape in 30% of the analysed crispants (**Figure 4.3 c&d**). This underscores the fundamental role of *ctsk* in early skeletogenesis in the form of cartilage development starting from 3dpf.

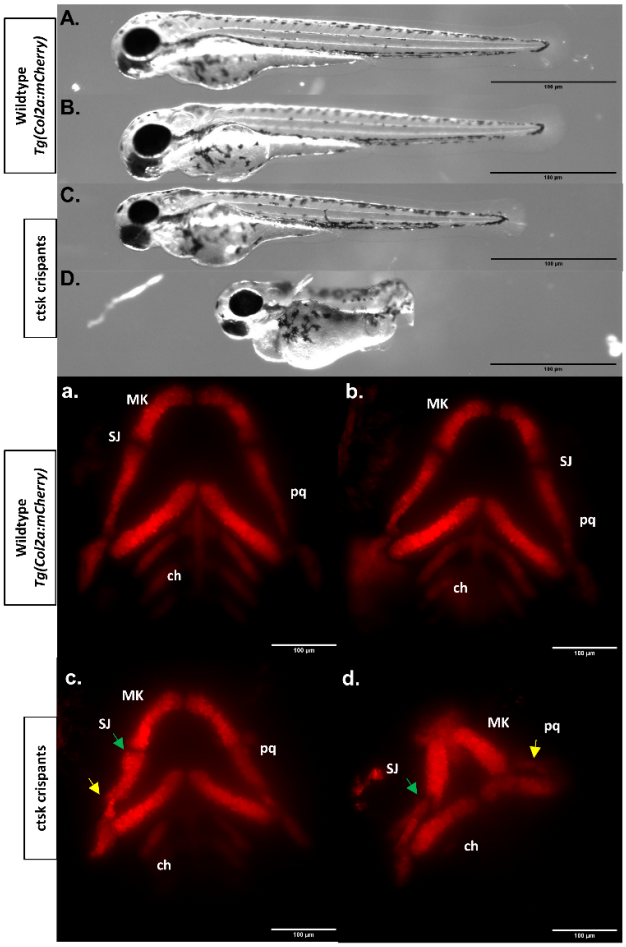


Figure 4.2: Cathepsin K is essential for early skeletogenesis at 3dpf

The upper panel illustrates the overall morphology, and the lower panel presents the corresponding jaw cartilage structure of the larvae under examination. Panels **A** and **B** show the brightfield ventral view of the wildtype controls (n=32), while panels **C** and **D** exhibit two distinct phenotypes when targeting *ctsk* (n=32). **C** displays a mild phenotype manifested by a retracted smaller jaw (n=15), whereas **D** exhibits a severe monstrous phenotype characterized by a significantly shorter and disrupted skeletal structure and a distinct retracted jaw (n=9); however, some larvae did not exhibit any phenotype in their general structure (n=8). Regarding the control group, un-injected larvae obtained from *Tg(Col2a:mCherry)* line display the expected cartilage structure in red, as observed from a lateral view (n=20). 400pmol ctsk sgRNA targeted larvae (n=26) showed an abnormality in their jaw cartilage where the structures of Meckel's (MK) and Palatoquadrate (pq-yellow arrow) were bent and exhibited non-uniform overall shape as demonstrated in **c** (n=12). Meanwhile, more severe deformities were observed in **d** (n=5), whereas some subjects did not exhibit any abnormal phenotype (n=9). The modulation seemingly affects the synovial joint (SJ) (indicated by the green arrow), while the ceratohyal (ch) cartilage shows less impact at this stage. The data was obtained from three different experimental replicates. The scale bar is measured at 100µm.

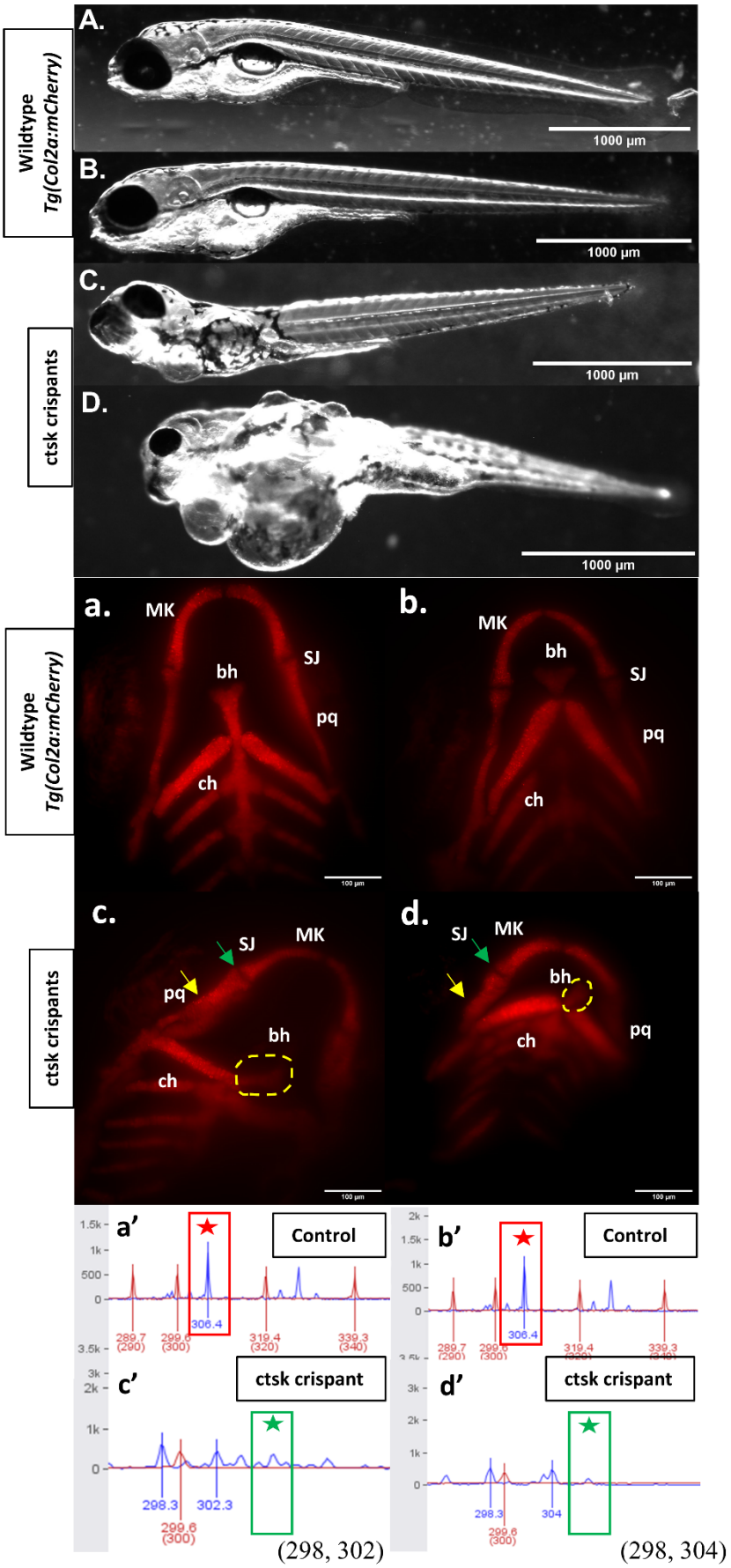

Wildtype
Tg(Col2a:mCherry)
ctsk crispants
A.
B.
C.
D.
1000 µm
a.
b.
c.
d.
MK
SJ
bh
pq
ch
100 µm
a'
b'
c'
d'
Control
ctsk crispant
306.4
289.7 (290)
299.6 (300)
319.4 (320)
339.3 (340)
298.3
302.3
304
(298, 302)
(298, 304)

Figure 4.3: Persistence of abnormal skeletogenesis in *ctsk* crispants at 5dpf

The top panel compares the morphology of *ctsk* crispants (ventral view) (**C&D**) (n=23) to that of wildtype controls (**A&B**) (n=24). The bottom panel displays the corresponding jaw cartilage structure in the larvae presented and are matched to their genotype at the bottom (**a'-d'**). Control larvae (**a&b**) obtained from the *Tg(Col2a:mCherry)* line show the expected cartilage structure in red lateral view at 5dpf (n=24). (**c&d**) 400pmol *ctsk* targeted larvae exhibited abnormal jaw cartilage, which affected the Meckel's (MK), Basihyal (bh-yellow circle), Palatoquadrate (pq-yellow arrow), synovial joint (SJ-green arrow) and ceratohyal (ch) structures (n=30). Of those, (n=14) developed mild abnormalities, illustrated in **c**, while (n=9) developed more severe defects, illustrated in **d**. Furthermore, (n=7) showed no phenotype, and larvae in this group appeared to develop jaw cartilage in a similar manner to wildtype controls. (**a'&b'**) The CRISPR-STAT analysis demonstrates the anticipated wildtype allele peak of the intact *ctsk* at 306.4 (red star), whereas (**c'&d'**) affirm the loss of the wildtype allele (indicated by the green star) in targeted larvae. The data was obtained from three distinct experimental replicates. The top panel scale bar is 1000µm, and the bottom panel scale bar is 100µm.

4.3.3. Abnormal heart structure and volume in *ctsk* targeted embryos

Although *CTSK* has been extensively researched in relation to bone homeostasis, its effect on heart development has not yet been identified, despite emerging evidence of its involvement in several cardiovascular diseases, as described in **section 4.1.4**. Therefore, to ascertain the impact of *ctsk* mutagenesis on heart development, I injected the same sgRNAs targeting *ctsk* described previously into the *Tg(myl7:lifeActGFP*) transgenic line to visualise the heart structure through fluorescently tagged cardiomyocytes. For this experiment, two different concentrations of *ctsk* single-guide RNAs were injected to assess the possibility that a lower concentration would result in a partial knockout, avoiding the detrimental effect observed with the higher concentration, and enable crispants to be raised to adulthood. Embryos were injected with 100pmol (0.5nl) and 400pmol (2nl) of two CRISPR single-guide RNAs. As negative controls, embryos were injected with 400pmol of tyrosinase targeting guides to confirm the success of the injection. Tyrosinase plays a crucial role in the production of pigmentation in zebrafish, and when it is targeted, larvae appear transparent. This provides a great screening tool to visualise the efficiency of CRISPR targeting and to ensure that the obtained phenotype is not due to the invasiveness of the needle during the injection process. At 3dpf, 100pmol injected larvae showed a mild effect on general structure of the larvae in 30.8% of 26 targeted larvae, while 400pmol targeted larvae displayed much significant deformities in the general structure as described in **section 4.3.2**, leading to a mild phenotype in 47.4% of the 19 embryos and a severe phenotype in 21% of the injected fish (**Figure 4.4** and **Figure 4.6).** Furthermore, the heart also showed different phenotypes in the two concentrations. In the lower concentration, hearts develop a similar morphology to wildtype, with the exception of some shape irregularities and a smaller ventricle chamber in 37% of 100pmol targeted larvae (**Figure 4.5 e&f**). However, higher concentration hearts show a disruption in the looping process that is essential for correct orientation of the heart within the organism, resulting in heart misorientation or complete inversion in 80% of larvae at 3dpf (**Figure 4.5 g&h**).

When larvae reached the 5dpf stage, 43% of the targeted group failed to develop a swim bladder when injected with 100pmol (**Figure 4.4 E&F**), while 62.5% of 400pmol targeted larvae failed to develop the swim bladder (**Figure 4.6 G&H**). Defective heart looping was also observed even in the lower concentration of targeted larvae (**Figure 4.5 e&f**), as well as the higher concentration (**Figure 4.5 g&h**). Additionally, both the atrium and ventricle chambers appear smaller in 400pmol (**Figure 4.5 g&h**). This was further confirmed by the heart volumetric measurements performed using Imaris Software and shown in **Figure 4.8** that determined a significant *(p-value= 0.0185)* decrease in the heart volume of 400pmol targeted larvae surprisingly when compared to 100pmol targeted larvae, but not significant when compared to wildtype controls *(p-value= 0.4639).*

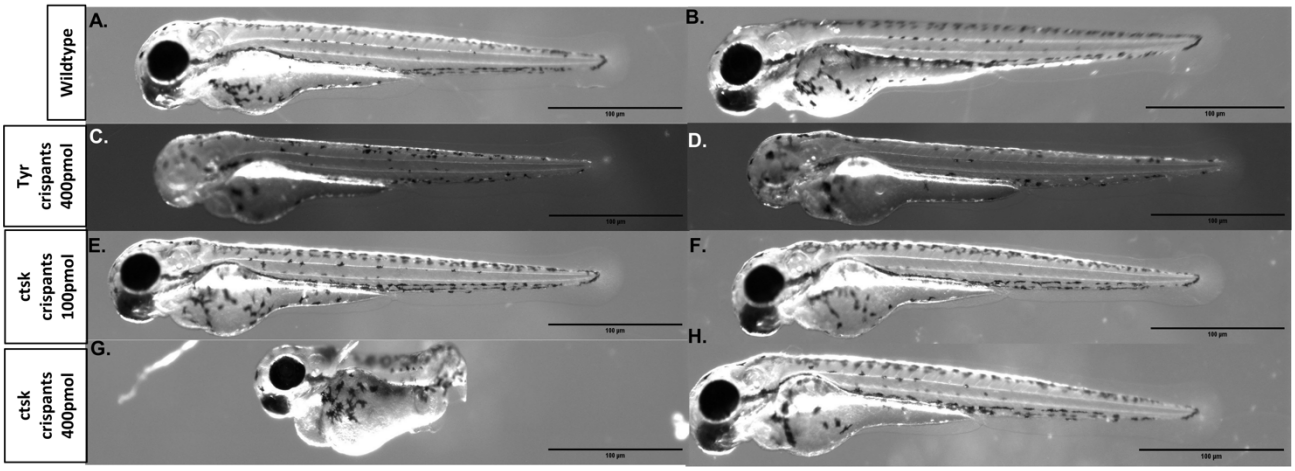


Figure 4.4: Clear deformities in whole body structure of 400pmol *ctsk* crispants at 3dpf

(A&B) The general morphology of wildtype controls was imaged in brightfield ventral view (n=24). **(C&D)** tyrosinase targeted larvae as negative controls showing loss of pigmentation that indicated successful targeting and displayed no abnormalities in the whole-body structure. This confirms that observed phenotypes were not a result of needle invasiveness during the injection (n= 8). **(E&F)** 100pmol *ctsk* targeted larvae (n= 18) showing unaffected structure that appears to develop in a manner similar to the wildtype controls. **F** shows slight curvature of the spine, observed in n=3. **(G&H)** 400pmol *ctsk* targeted larvae (n= 14), some display monstrous severe phenotype as shown in **G** (n= 5), while others display mild abnormalities such as a retracted jaw and slight curvature as shown in **H** (n=8), (n=1) do not show a phenotype. Scale bar= 100µm.

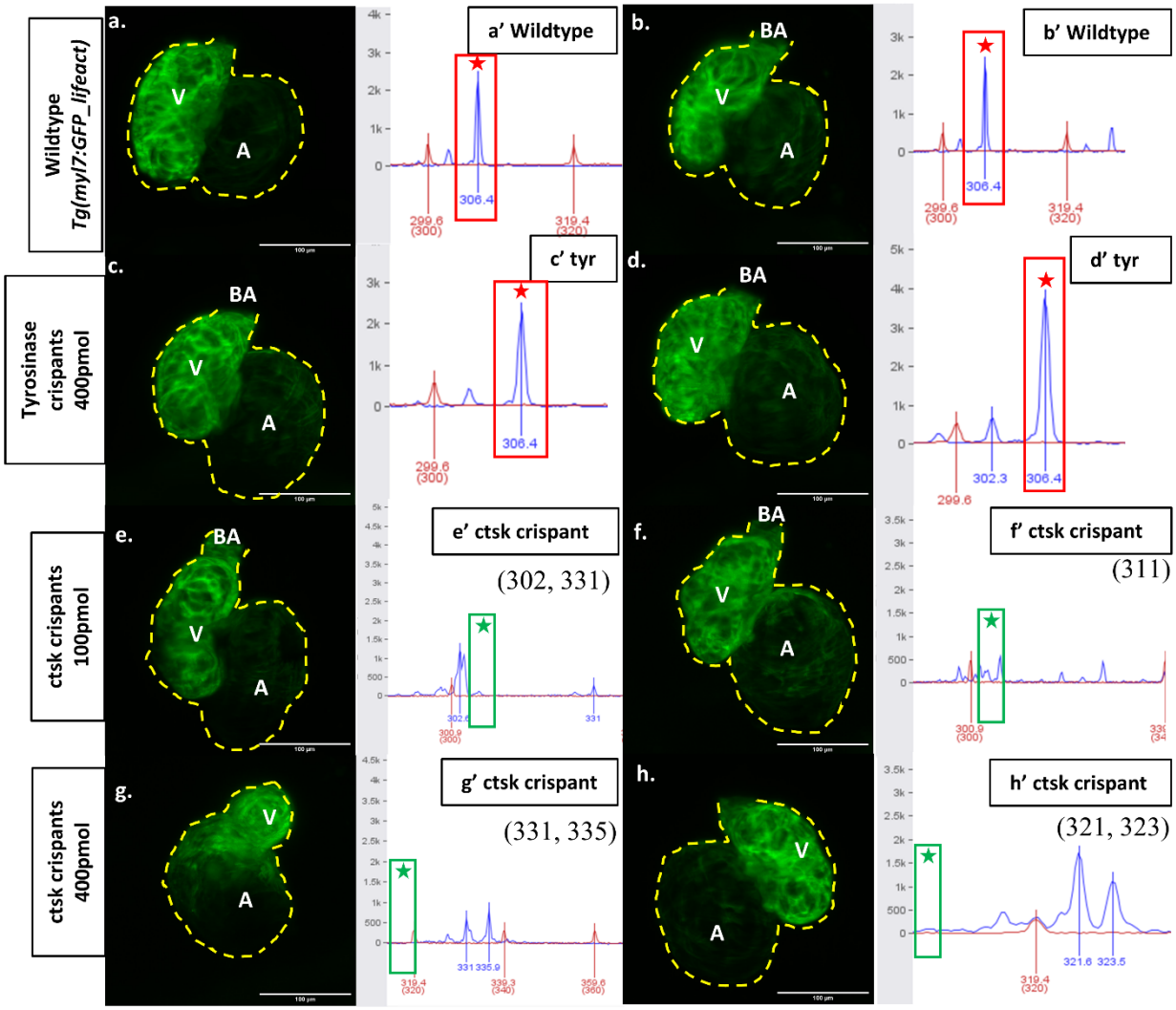


Figure 4.5: Abnormal heart development in *ctsk* crispants at 3dpf

(a&b) Wildtype un-injected larvae (n=16) representing the expected heart structure (outlined in yellow) at the 3dpf stage in the *Tg(myl7:lifeActGFP)* transgenic line to visualise the heart in green (lateral view). **(a'&b'**) CRISPR-STAT genotyping shows *ctsk* wildtype peak at 306.4 (red star) **(c&d)** 400pmol tyrosinase CRISPR targeted larvae (n=8) as negative controls and their corresponding genotype **(c'&d'). (e&f)** 100pmol *ctsk* targeted larvae (n=8) appears to have smaller irregularly shaped ventricles in some generated crispants (n=3), while the rest of targeted larvae display no phenotype (n=5), and confirmation of the successful targeting is shown in the corresponding CRISPR-STAT **(e'&f').** A green star indicates the expected peak position when absent. **(g&h)** 400pmol *ctsk* targeted larvae (n= 14) showing a highly disrupted heart morphology in 5 of the examined crispants, a mild phenotype similar to 100pmol crispants was observed in 8 of crispants, while only one larva exhibited no phenotype. Both chambers appear smaller than controls, but has not been measured separately, and the modulation of *ctsk* was confirmed **(g'&h').** Additionally, misorientation of the heart is observed causing inversion of the heart. V= Ventricle, A= Atrium. BA= Bulbus Arteriosus. Scale bar= 100µm.

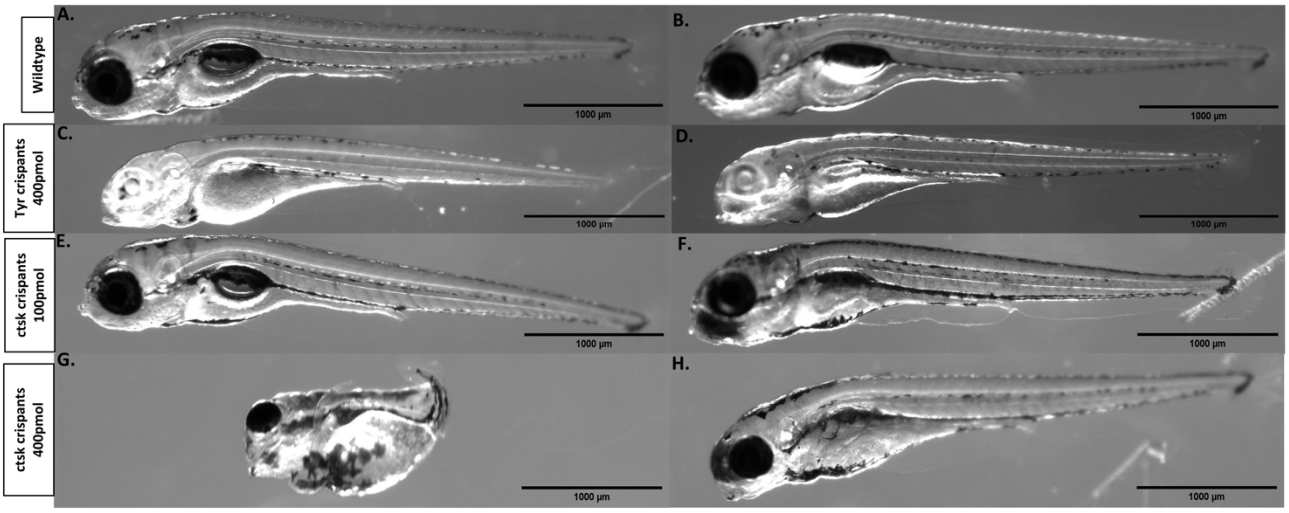


Figure 4.6: Clear deformities in whole body structure of 400pmol *ctsk* crispants at 5dpf

(A&B) The general morphology of the wildtype controls was imaged in a brightfield ventral view (n=22). **(C&D)** tyrosinase targeted larvae were used as negative controls showing loss of pigmentation indicating successful targeting with no signs of abnormalities in the whole-body structure (n= 8). **(E&F)** 100pmol *ctsk* targeted larvae (n= 8) showing mildly affected structure with body curvature and skull abnormalities in 3 of examined larvae, while the remaining 5 did not display any skeletal defects. 43% failed to develop an inflated swim bladder. **(G&H)** 400pmol *ctsk* targeted larvae (n= 20), some display monstrous severe phenotype as depicted in **G** (n= 10), while others demonstrated mild abnormalities such as a retracted jaw, smaller skull and slight curvature as shown in **H** (n=9), (n=1) do not show a phenotype. Out of the 20, 400pmol *ctsk* targeted larvae, 62% failed to develop an inflated swim bladder. Scale bar= 1000µm.

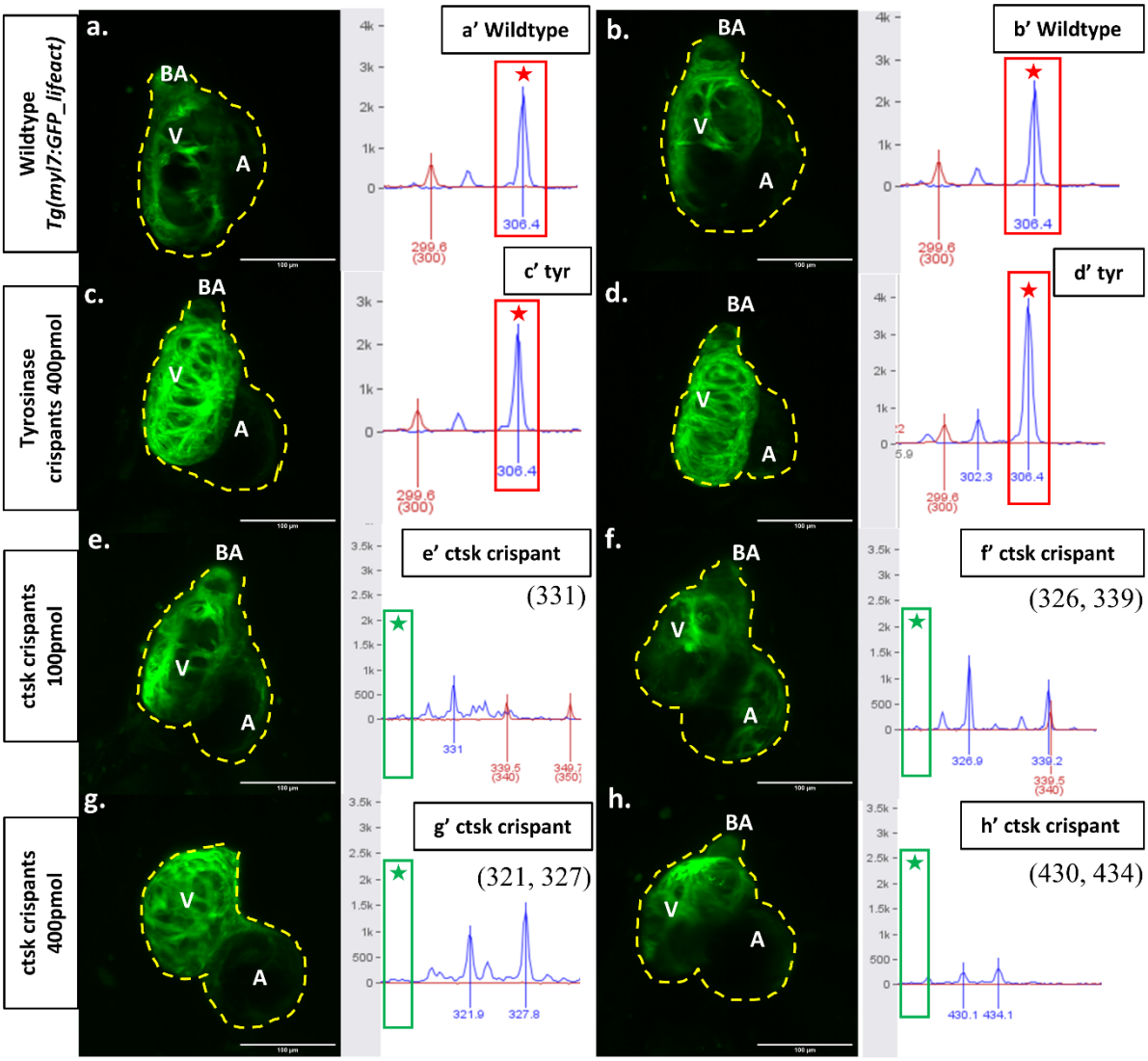


Figure 4.7: Persistence of heart shape irregularities up to 5dpf

(a&b) Wildtype un-injected larvae (n=16) representing the expected heart structure (outlined in yellow) at the 5dpf stage in the *Tg(myl7:lifeActGFP)* transgenic line, showing the heart in green fluorescence (lateral view). **(a'&b'**) CRISPR-STAT genotyping illustrates a *ctsk* wildtype peak at 306.4 (red star) **(c&d)** 400pmol tyrosinase CRISPR targeted larvae (n=8) as negative controls and their corresponding genotype **(c'&d'). (e&f)** 100pmol *ctsk* targeted larvae (n=8) appears to have smaller atrium and ventricle of irregular shape in some generated crispants (n=3), while the rest of targeted larvae display no phenotype (n=5), and confirmation of successful targeting is shown in the corresponding CRISPR-STAT **(e'&f').** Green stars indicate the expected peak position when absent. **(g&h)** 400pmol *ctsk* targeted larvae (n= 14) showing a highly disrupted heart morphology in 5 examined crispants, a mild phenotype similar to 100pmol crispants was observed in 8 of crispants, while only one larva showed no phenotype. Both chambers appear smaller than controls and the modulation of *ctsk* was confirmed (**g'&h'**). V= Ventricle, A= Atrium. BA= Bulbus Arteriosus. Scale bar= 100µm.

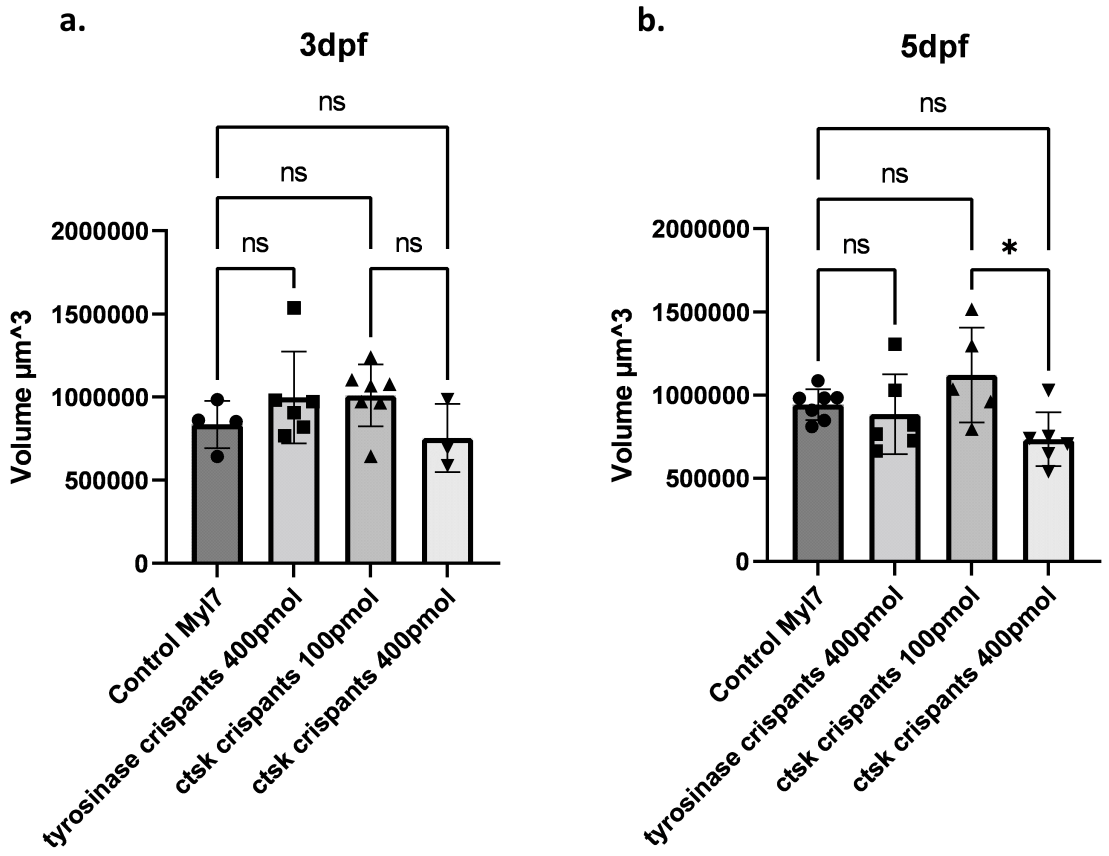


Figure 4.8: Significant decrease in heart volume of 400pmol *ctsk* targeted larvae compared to the 100pmol targeted larvae at 5dpf

No significant change in heart volume was observed at 3dpf **(a)**. Nevertheless, at 5dpf **(b),** targeted larvae with the lower concentration (100pmol) (n=5) have a non-significantly larger heart volume *(p-value= 0.4639*) in comparison to un-injected (n=7) and tyrosinase negative control (n=6). Using a higher concentration of 400pmol (n=6), the developing heart appear smaller than both controls, and only showing significant reduction in volume when compared to enlarged hearts of lower concentration *(p-value= 0.0185*). Statistical analysis was conducted as a One-way ANOVA with multiple comparisons using GraphPad Prism.

To investigate whether *ctsk* is essential for earlier heart development, a *myl7* in situ hybridization probe (Noel's Lab) was used to stain the heart at 1dpf in fixed 400pmol *ctsk* targeted larvae and uninjected wildtype controls. As shown in **Figure 4.9**, an anomalous cardiac phenotype can be detected as early as 1dpf in *ctsk* crispants. The location and orientation of the heart was consistently misaligned. It should be noted that the general head size of *ctsk* crispants appear smaller, which could potentially contribute to the observed phenotype (**Figure 4.9 c-f**). This experiment was conducted only once, with a limited sample size of 4 crispants tested against 2 wildtype controls, and therefore requires validation through repetition. Furthermore, negative controls injected with tyrosinase were not stained, which would have provided a better control for this experiment. Taken together, acquired data suggest high involvement of *ctsk* in the early stages of heart development, potentially as early as 1dpf. Targeting *ctsk* has detrimental effects that are not just exclusive to the heart structure, but also to its positioning by disrupting the looping process.

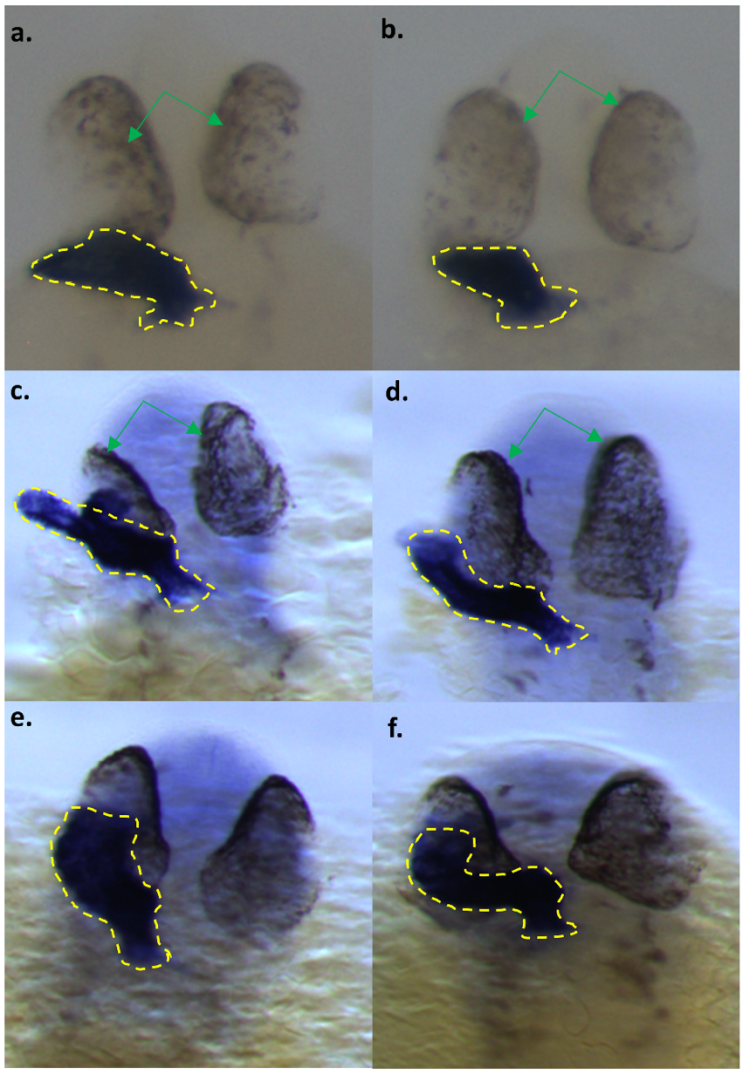


Figure 4.9: Cathepsin K is involved in early stages of heart development

In situ hybridization-stained hearts (highlighted in yellow outline) using myl7 specific RNA probe to visualise the heart structure at 1dpf (dorsal view). **(a&b)** control un-injected nacre larvae that display the anticipated heart size and orientation at the 1dpf stage. **(c-f)** 400pmol *ctsk* targeted larvae that were raised to 1dpf, fixed in PFA and stained for heart visualisation. *ctsk* targeted larvae (n=4) exhibited an abnormal heart structure that is tilted to the left side with an unexpected localization behind the left eye when compared to wildtype controls (n=2). The eye structures were unintentionally visible due to background staining and were highlighted using green arrows.

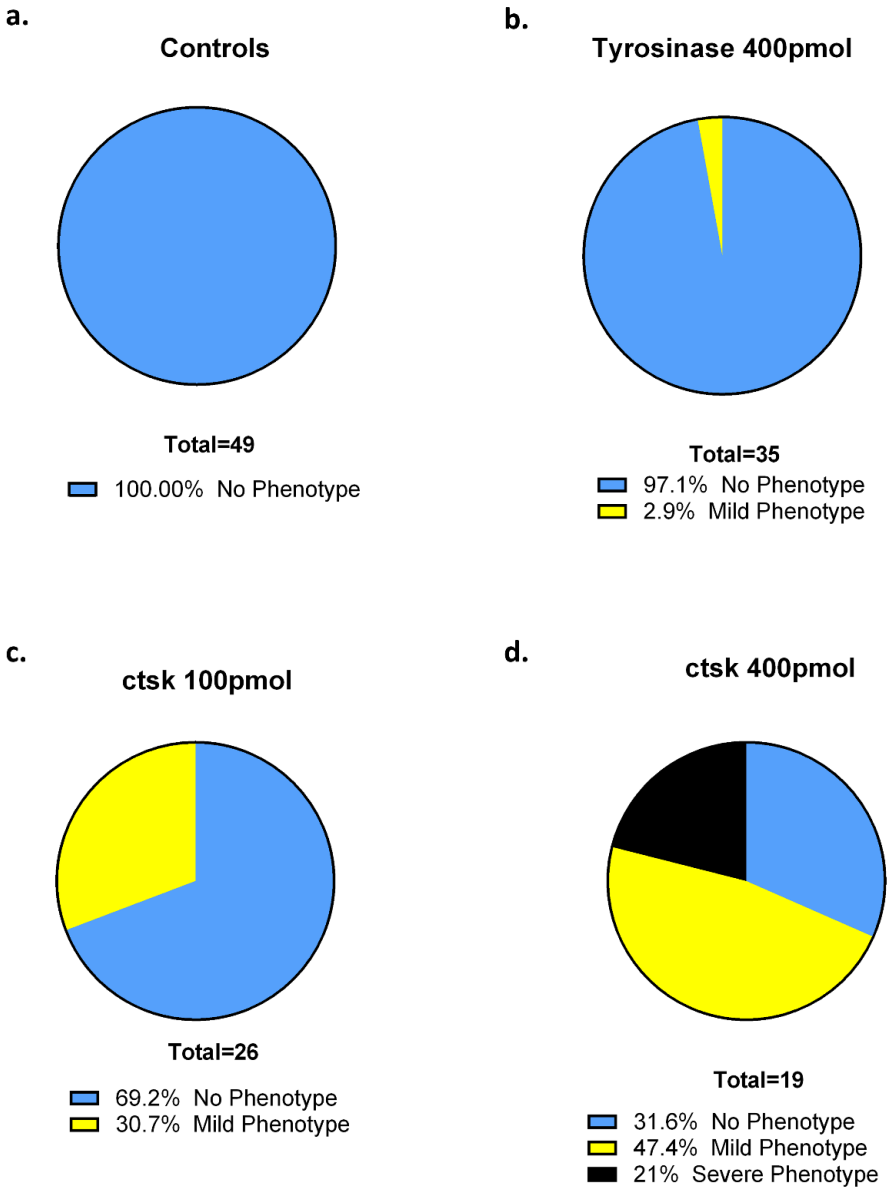


Figure 4.10: General morphology phenotype of *ctsk* targeted larvae

(a) Wildtype controls demonstrate 100% normality of whole-body structure. **(b)** Tyrosinase, used as a negative control, displays only 2.9% percentage abnormality indicating a mild impact of injection on larvae phenotype at 5dpf. **(c)** 69.2% of 100pmol injected 0.5nl display no alterations in their overall morphology, while 30.7% show a mild phenotype characterised by a shorter, curved body structure, smaller head or abnormal jaw (**Figure 4.6, F**). **(d)** 400pmol targeted larvae by injecting 2nl of injection mixture. 31.6% show no phenotype even in the presence of a mutation, 47.4% show mild phenotype such as general skeletal figure curvature, and 21% show completely disrupted whole body structure **(Figure 4.6 G).**

4.3.4. Impact of *ctsk* mutation on larval activity duration and speed

After observing a disruption in the overall structure of larvae starting at 3dpf, alterations in the swimming behaviour of larvae were suspected at 5dpf. To examine this hypothesis, 5dpf larvae were monitored over 10-minute cycles in the presence and absence of light using 16 larvae of each injected concentration, with both mild and extreme deformities, and uninjected controls. Larvae display an increase in movement in the presence of light compared to darkness. The two light states were performed separately as an additional condition to identify any significant changes in behaviour linked to a nervous system phenotype. Larvae injected with 400pmol showed a significant reduction (with light, $p= 0.0004$) (without light, $p= 0.0452$) in inactivity duration in both light conditions when compared to the 100pmol injected larvae (**Figure 4.11 a&c**), indicating an increase in general movement. However, in both light conditions, 400pmol larvae displayed a significant (with light, $p= 0.0048$) (without light, $p= 0.0443$) decrease in speed when compared to 100pmol injected larvae **(Figure 4.11 b&d**). This indicates that at 400pmol concentration, larvae are less sedentary than 100pmol larvae, but similar to wildtype movement. Although 400pmol seem to show more movement, they show significant reduction in their ability to perform rapid swimming also when compared to 100pmol larvae, but not to wildtype controls. Surprisingly, yet in correlation with previously obtained heart volume results in **section 4.3.3**, 100pmol and 400pmol show opposite observations that causes a significant change between the two groups that is not apparent when both groups are compared to wildtype controls. Furthermore, no correlation can be drawn between larval behaviour and light conditions to suggest a possible involvement of *ctsk* in the nervous system.

A possible explanation to the swimming behaviour detected in 400pmol targeted larvae in regard to the increase in general activity, yet reduction in rapid movement could be a result of the highly disrupted skeletal morphology, which disables the larval free swimming, yet larvae display more frequent subtle movements or fidgeting, which is detected by the viewpoint system as an activity. Another factor could be the reduction in heart volume of 400pmol *ctsk* targeted larvae causes a reduction in their fast movement ability. This experiment was only conducted once; therefore, further repeats may help us understand the opposite pattern displayed by the two concentrations. Additionally, sequencing of individual larvae may explain a correlation between the genotype of targeted larvae and the displayed phenotype.

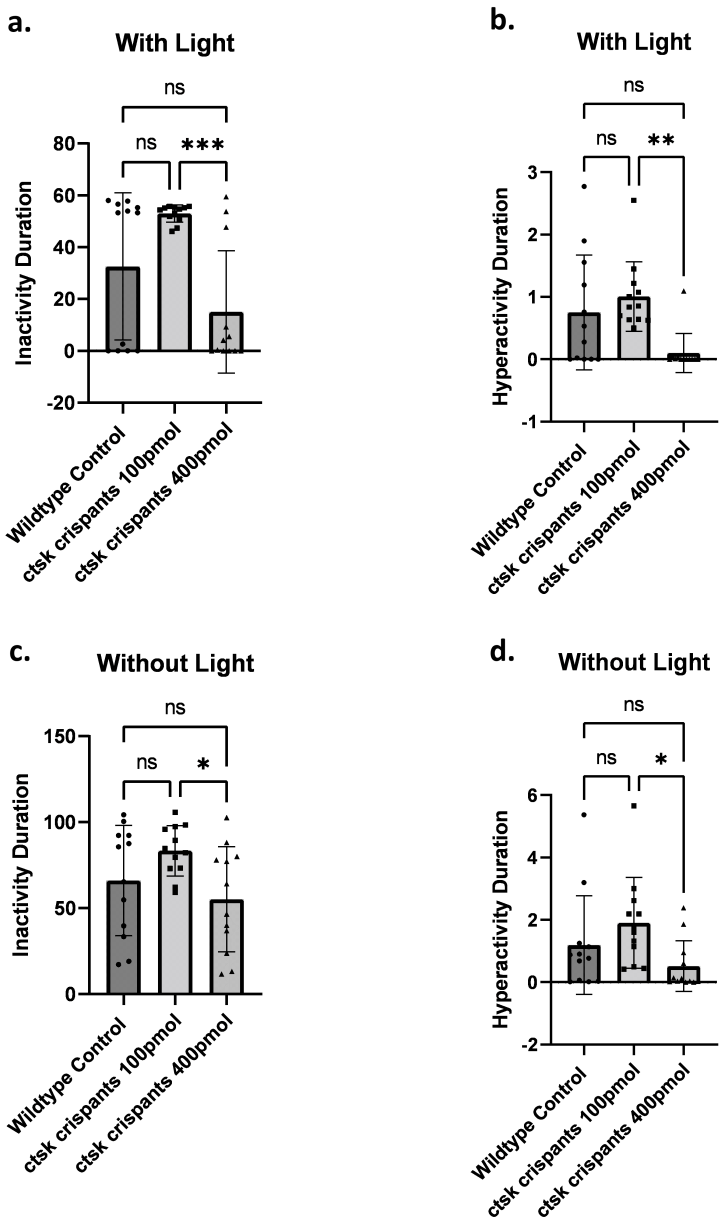


Figure 4.11: Significant increase in general activity of 400pmol compared to 100pmol *ctsk* targeted larvae, while showing a significant reduction in hyperactivity at 5dpf

(a) In the presence of light, *ctsk* targeted larvae of low concentration (100pmol) (n= 12) do not exhibit a significant change in inactivity duration compared to controls (n= 11). On the other hand, *ctsk* targeted larvae of higher concentration (400pmol) (n= 12) appear to have an increase in the general activity compared to 100pmol *(p-value= 0.0004*), although they were unable to move as fast as 100pmol larvae **(b)** *(p-value= 0.0048*). **(c&d)** in the absence of light, larvae exhibit the same movement pattern as in the presence of light with significance *(*Inactivity duration *p-value= 0.0452*) *(*Hyperactivity duration *p-value= 0.0443*). One-way ANOVA with multiple comparison was used to analyse presented data. Figure generated using Graphpad Prism.

4.3.5. Non-significant change in adult locomotion speed of 100pmol *ctsk* targeted fish

To examine the impact of the mutation on the functionality of the skeletal system, the activity levels of 12-month-old zebrafish adult crispants was monitored. Using the adult Viewpoint behavioural analysis system described in **section 2.7.2,** the time spent in an inactive state is calculated for each fish, as well as the time spent in a hyperactive state. In this section, only 100pmol crispants data is included, as injections of the higher concentration (400pmol) significantly harmed the embryonic development, preventing us from raising the larvae past 5dpf due to animal welfare concerns. However, 100pmol targeted larvae were raised, and presented data was exclusively conducted from the 100pmol *ctsk* targeted fish. All fish were fin clipped and sequenced to confirm the successful targeting of *ctsk* (**Figure 4.17**). All sequenced fish exhibited varying mutations, indicative of mosaicism and confirming mutagenesis as expected from founder crispants.

10, 12 months year-old adults, 4 wildtype controls (2 male and 2 females) and 6 crispants (3 males and 3 females), were monitored over 7 hours to detect any change to their swimming behaviour. This experiment was conducted 3 times under two different conditions of absence and presence of light separately. Initially, male and female subjects were analysed individually against their sex-matched controls, taking into account the difference in swimming speed where males typically outswim females due to their diminutive size (Leris et al., 2013). As shown in **Figure 4.12**, no change was detected in both light conditions for both sexes in 100pmol *ctsk* targeted founders. When all repeats and sexes were merged (**Figure 4.13**), no significant difference in swimming behaviour was detected. Finally, the swimming behaviour of individual fish over the 7-hour monitoring duration was analysed to determine any changes throughout the experimental period, and fish appear to have a consistent behaviour through the whole duration (**Figure 4.14**). Sequences of crispants used in this analysis are shown in **Figure 4.17**. In conclusion, *ctsk* targeting using 100pmol sgRNAs does not exhibit significant changes in adult swimming speed and efficiency of crispants.

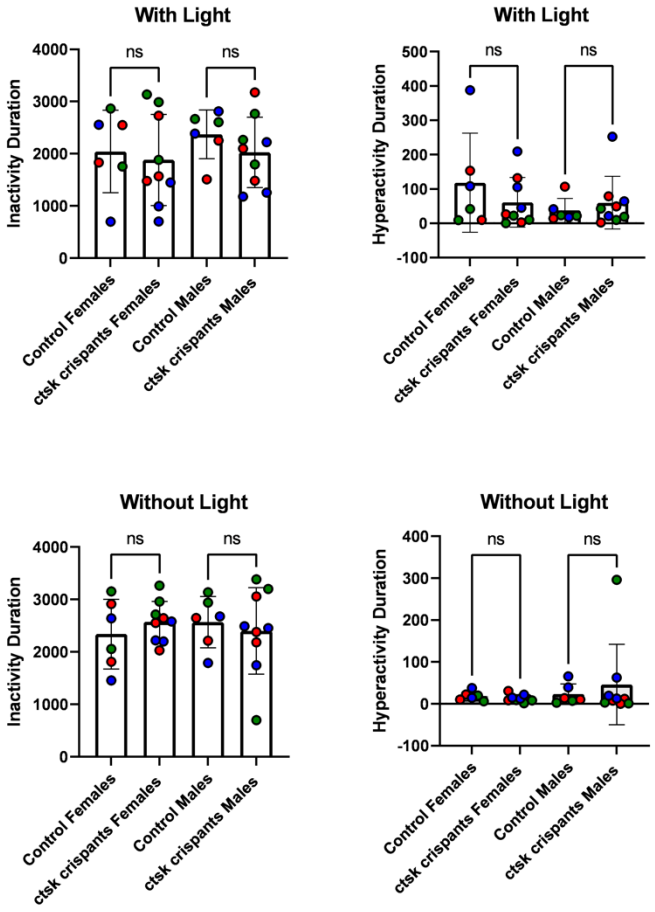


Figure 4.12: No remarkable change in adult *ctsk F0* movement duration or speed in the presence or absence of light

Figure comparing the behaviour of adult wildtype control with 100pmol *ctsk* targeted fish at the adulthood stage. **(a)** No significant change in the duration spent by adult crispants (n=6), 3 females in an inactive state when compared to wildtype controls (n=4) in the presence of light within both sexes (females p-value= 0.8967; males p-value= 0.615). **(b)** *ctsk* crispants did not exhibit a significant alteration in the duration they spent in a hyperactive swimming state when in the presence of light (females p-value= 0.4031; males p-value= 0.8628). **(c)** The same pattern observed in the presence of light was also shown in the absence of light, indicating no effect of mutation on adult inactivity of both sexes (females p-value= 0.7295; males p-value= 0.8509). **(d)** as seen in the presence of light, no change in duration spent in hyperactive state was detected in both sexes (females p-value= 0.9794; males p-value= 0.6773). Figure summarises data from three different experimental replicates within the two light conditions. Red indicating a measurement obtained in the first replicate, while green is data from second replicate experiment and blue is the third replicate. Each point represents the average time spent by a single fish through the 7-hour experimental period. Analysis was conducted as One-way ANOVA with multiple comparisons using the Graphpad prism software.

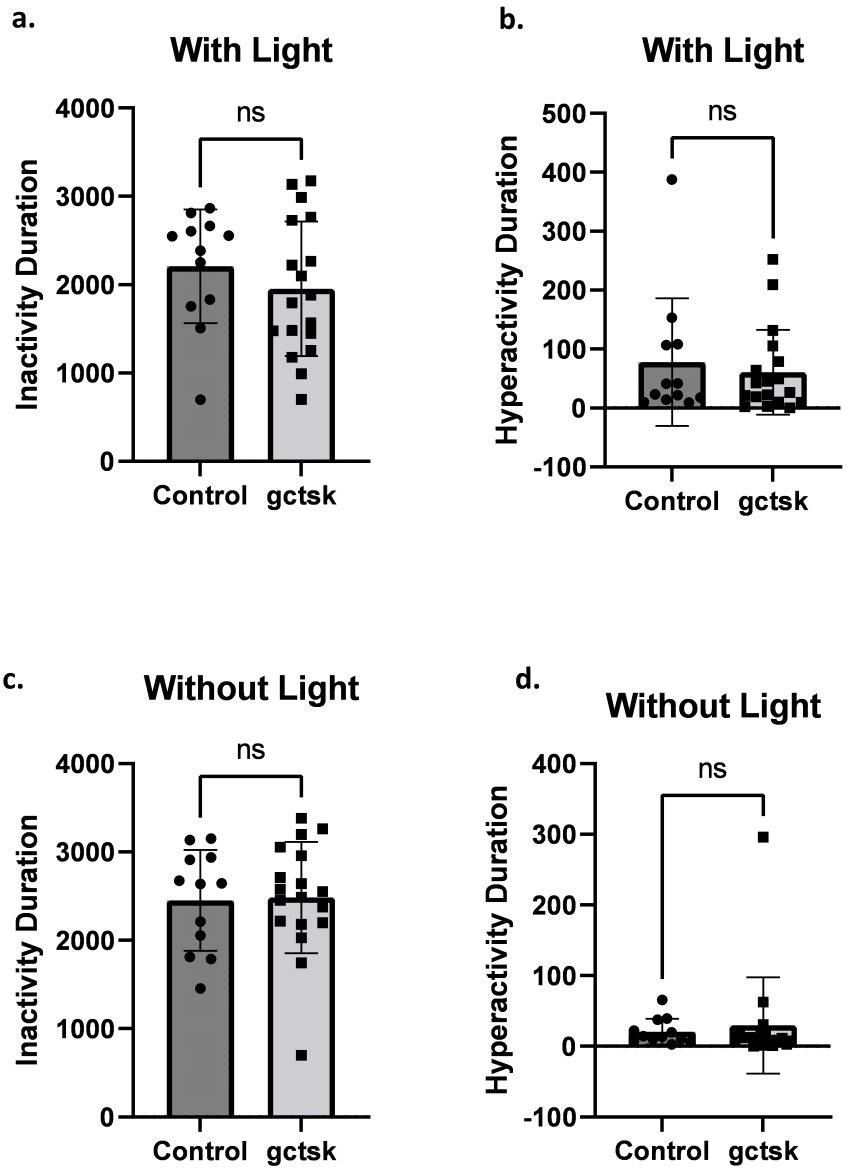


Figure 4.13: No remarkable change in total adult *ctsk* F0 movement pattern in the presence or absence of light

Figure comparing the behaviour of wildtype control adults with 100pmol *ctsk* targeted fish at the adulthood stage. **(a)** no significant change in duration spent by crispants (n=6) in an inactive form when compared to controls (n=4), both sexes combined in the presence of light *(p-value= 0.3505).* **(b)** no significant change in duration of hyperactivity movement in the presence of light *(p-value= 0.6046).* **(c)** unaffected movement in the absence of light *(p-value= 0.8825),* two sexes combined, and the same pattern observed in the hyperactivity duration **(d)** *(p-value= 0.6647).* Each point represents average time spent by an individual fish and the graph shows data for 3 replicates combined. Control= wildtype control, gctsk= *ctsk* crispant. Analysis was conducted by combining all repeats and tested using an unpaired student t-test.

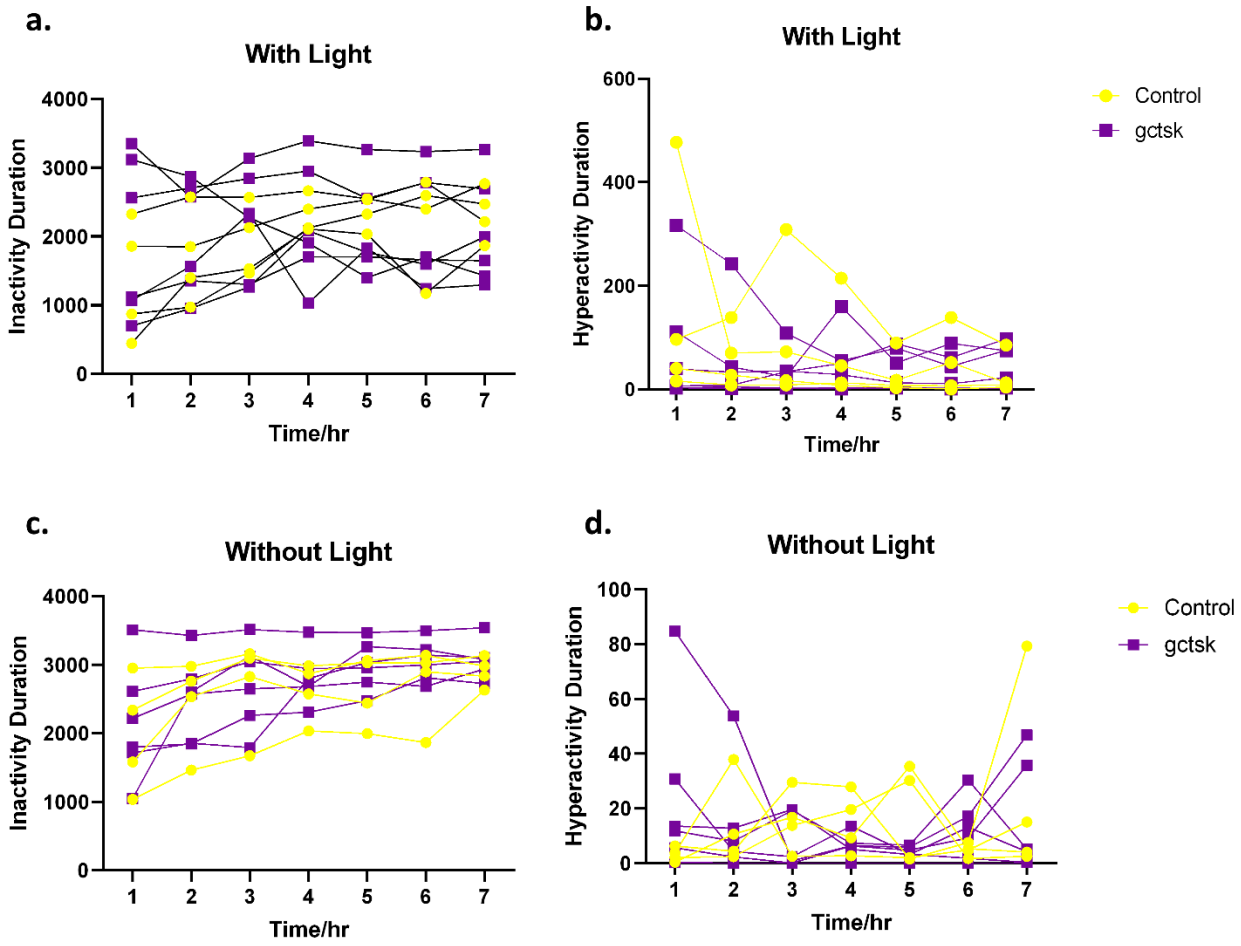


Figure 4.14: No remarkable change in movement pattern of *ctsk* targeted adult founders over the 7-hour experimental duration

Under both light conditions, *ctsk* founders appear to spend similar duration in an inactive state when compared to wildtype controls without changes overtime **(a&c).** Furthermore, **(b&d)** the fish also do not display a difference in hyperactivity duration of the 7-hour experimental period in the presence or absence of light. Control fish represented in yellow, while ctsk founders (gctsk) are shown in purple. Figure generated using Graphpad Prism.

4.3.6. Micro-CT analysis for adult *ctsk* zebrafish crispants

Skeletogenesis is a continuous process that commences with cartilage generation at 2dpf (Schilling et al., 1997), followed by the first functional joint observation at 3dpf (Brunt et al., 2017). First sign of bone development following mineralisation of developed cartilage starts at 4-6dpf and the skeletal system is typically fully formed by 2 months (Bird et al., 2003). Therefore, visualising the developed bone structure in adult *ctsk* crispants could provide novel insights into the role of *ctsk* in bone mineralisation, maintenance and degeneration at later stages. To follow the effect of *ctsk* targeting on later bone development, six 100pmol targeted fish were fixed at 12 months of age in 4%PFA, and then imaged as per the protocol detailed in **section 2.9.** Results are displayed in **Figures 4.15 & 4.16,** indicating a non-significant change in bone volume (females: p-value= 0.1783; males: p-value= 0.8757). and bone mineral density of 100pmol crispants compared to uninjected controls (females: p-value= 0.5009; males: p-value= 0.2426). However, there was a significant increase in bone surface area, which was only detected in male crispants (females: p-value= 0.9836; males: p-value= 0.0188). This finding suggests that there may be a differential effect between the sexes. Nevertheless, as the number of tested adults was limited, further repetitions are necessary to confirm this observation.

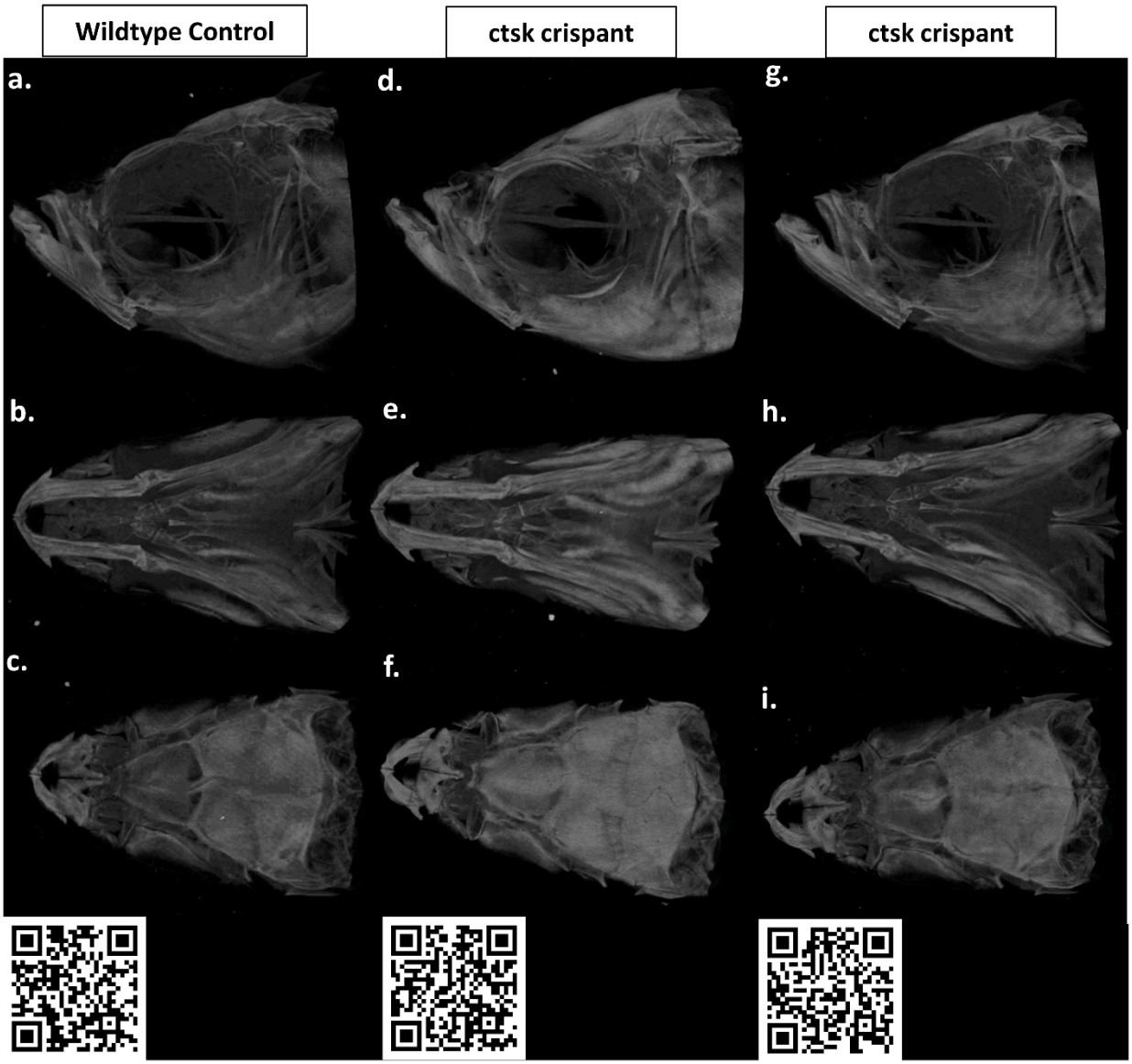


Figure 4.15: *ctsk* crispants show a significant increase in bone surface area, only in males

Micro-CT scans show the bone structure comparison of adult *ctsk* crispants (**d-i**) against wildtype control (**a-c**) at 12 months. The first row depict the lateral view (**a,d&g**), the second row displays the ventral view (**b,e&h**), and the third row shows the dorsal view of zebrafish skull (**c,f&i**). To view the entire 3D construct, please scan the QR code located at the bottom left.

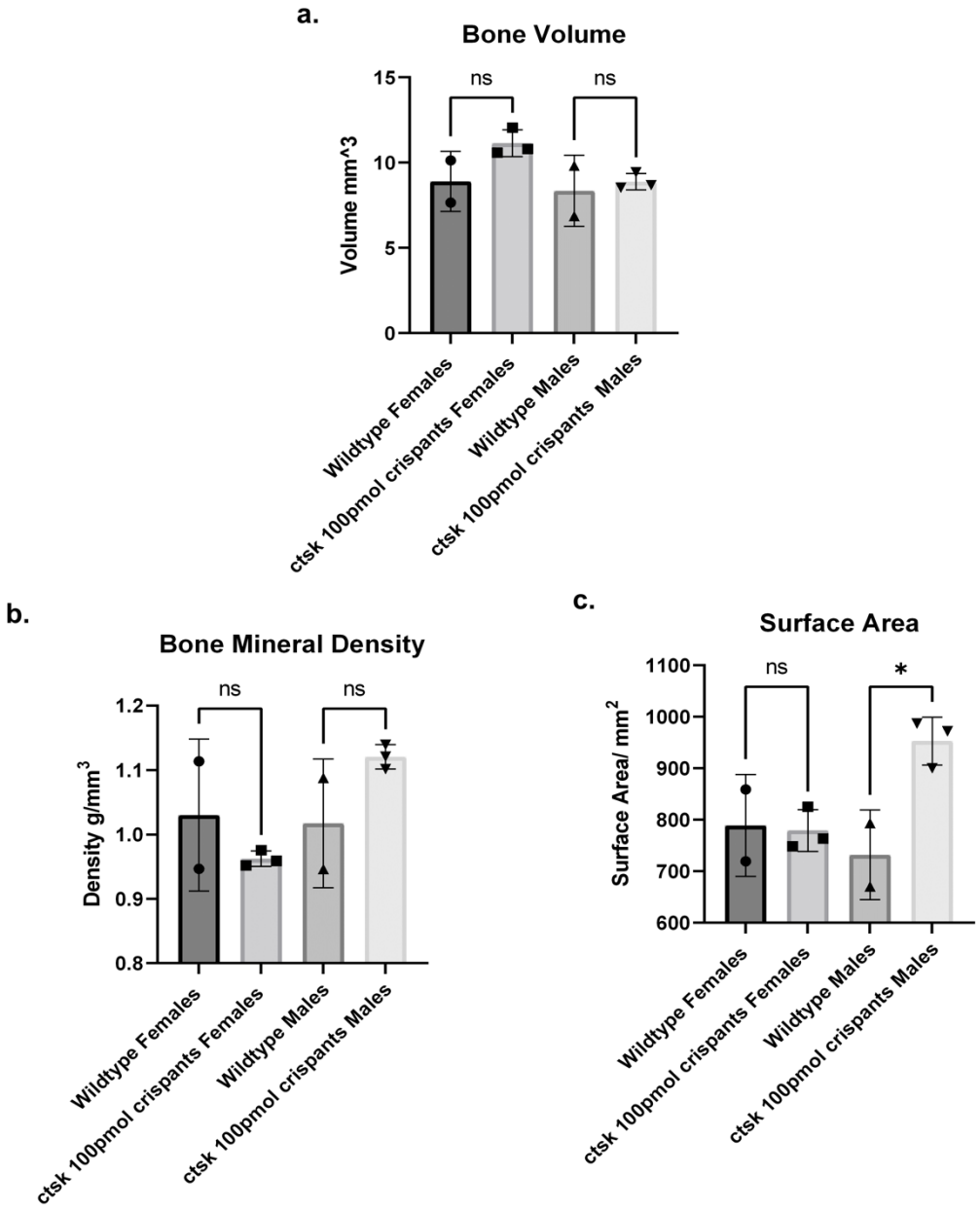


Figure 4.16: Significant increase in bone surface area of male *ctsk* adult crispants

Results obtained from micro-CT scans of 6, 12-month-old crispants. **a.** unaffected general bone volume of crispants when compared to wildtype controls (females: *p-value= 0.1783*) (males: *p-value= 0.8757*). **b.** unaffected bone mineral density of crispants when compared to controls (females: *p-value= 0.5009*) (males: *p-value= 0.2426*). **c**. significant increase in bone surface area of male crispants (females: *p-value= 0.9836*) (males: *p-value= 0.0188*). Each data point represents an individual fish and statistical analysis was conducted as One-way ANOVA with multiple comparisons using Graphpad Prism.

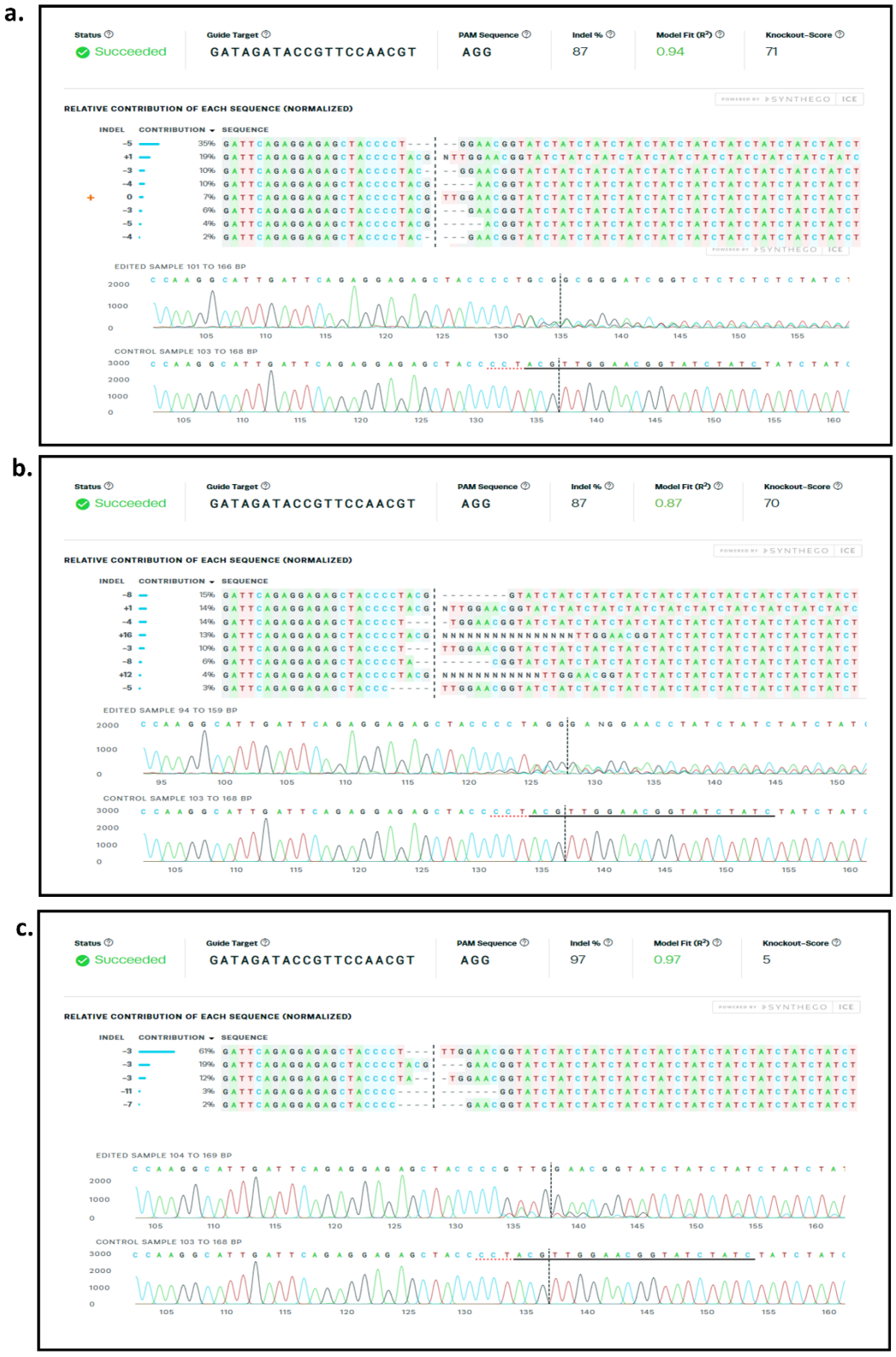


Figure 4.17: Sequencing data of *ctsk* crispants utilised in adult behavioural analysis and micro-CT scans

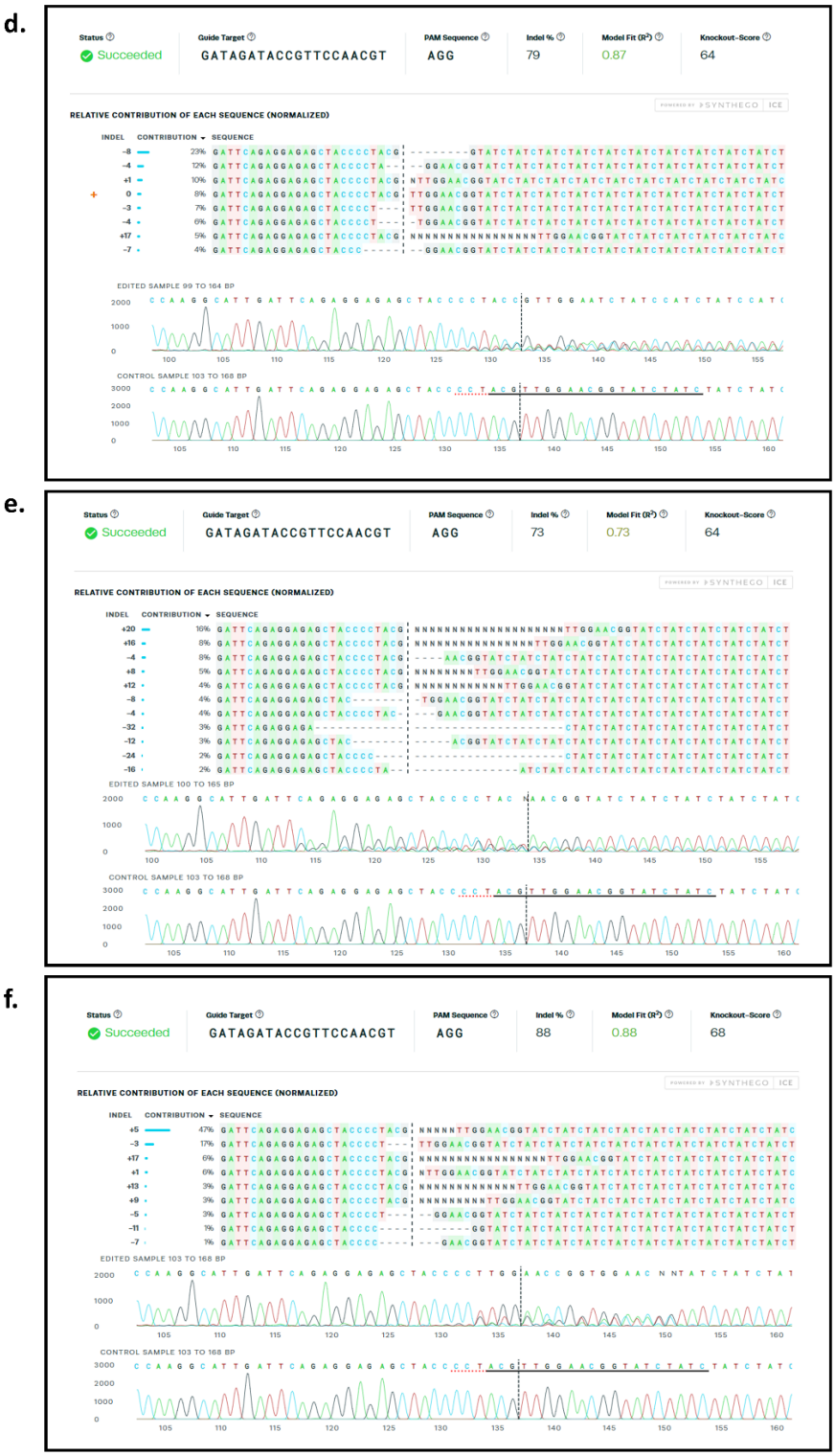


Figure 4.17: Sequencing data of *ctsk* crispants utilised in adult behavioural analysis and micro-CT scans (continued).

(a-c) 3 male crispants exhibiting a mosaic genotype. **(d-f)** 3 female crispants showing clear mosaicism, with various mutations observed in a sole organism, as anticipated in founder crispants. Top part of each panel represents the percentage of each mutation in the sequenced fish, while the bottom section shows the traces. Figure generated using Synthego ICE sequencing Analysis Tool.

4.4. Discussion

Although cathepsins have been identified for decades, yet our understanding of their mode of action remains limited due to the complexity of the system and its high involvement in numerous processes in different tissues. At present, our knowledge of the functional role of *CTSK* is primarily focused on bone development and remodelling (Gelb et al., 1996; Saftig et al., 1998; Garniero et al., 1998; Nygaad et al., 2004). This was linked to the role of *CTSK* in ECM regulation (Buck et al., 1992). *CTSK* also plays a crucial role in regulating the immune system (Riese et al., 2000; Troen., 2006). This suggests a high level of involvement in various systems (Buhling et al., 2004). As more information is revealed about *CTSK* function, the gene becomes increasingly intriguing. Much of the existing research on *CTSK* focuses solely on its role within the musculoskeletal system, with limited understanding of how it impacts non-skeletal tissues. Although elevated levels of *CTSK* have been identified in various studies related to OA and CVD, the underlying mechanism for this increase remains unknown. The most straightforward explanation would be an elevation in pro-inflammatory mediators that amplify the activation of pathways upstream of *CTSK*, like TGF-β signalling (Flanagan-Steet et al., 2018), MAPK signalling (Lee et al., 2018) and Calcium signalling pathways (Bai et al., 2018).

As discussed in **section 4.1.4**, studies revealed an increase in CTSK levels in CVD cases that was associated with the disease progression as when inhibited, signs of recovery were observed (Hua et al., 2015). Conversely, *Ctsk* knockdown in CVD mouse models worsened the symptoms and increased cardiomyocyte death indicating a protective anabolic role of CTSK in cardiac tissue (Fang et al., 2019). All investigations of CTSK functions in the cardiovascular system have been associated with diseases, neglecting the role of *CTSK* in early heart development. Therefore, in this study, we examined the role of *ctsk* in early zebrafish skeletogenesis. The effect of targeting *ctsk* on heart development was also tested for the first time, to our knowledge. Additionally, the effect of the skeletal phenotype was tested for its impact on the functionality of the musculoskeletal system by detecting changes in movement of targeted larvae, as well as in the adulthood stage of raised crispants. Our data confirmed the essential roles of *ctsk* in the early development of the musculoskeletal system. It also revealed its crucial function in the early heart development in zebrafish.

ctsk targeted larvae displayed clear disruptions in the general body morphology. At 5dpf, 30.7 % of 100pmol targeted larvae displayed a mild phenotype of a shorter curved figure with smaller eyes and skull, no signs of a severe phenotype were detected at this concentration. In the higher concentration of 400pmol, 47.4% of larvae showed described mild phenotype, while 21% displayed a severe phenotype with remarkable disruptions in the overall morphology that can be referred to as monsters. 43% of the 100pmol targeted larvae failed to develop an inflated swim bladder, while the percentage increases to 62.5% in 400pmol larvae, highlighting a role of *ctsk* in swim bladder development.

Looking at the jaw cartilage structure in the *Tg(Col2a:mCherry)* transgenic line at 3dpf, abnormal development in the Meckel's, palaquadrate and synovial joint structures was detected. However, it appears that the ceratohyal cartilage is less impacted compared to other cartilage structures. Only 400pmol *ctsk* targeted larvae were generated in the *Tg(Col2a:mCherry)* transgenic line for cartilage visualisation. Unfortunately, due to time limitations, the lower concentration tested in the heart was not performed in the cartilage line. However, in 400pmol targeted larvae, 46% of larvae showed a mild phenotype including bents in the Meckel, ceratohyal and palaquadrate structures. 20% displayed more severe defects, resulting in complete disruption to the shape of the jaw, and loss of the synovial joint.

The abnormal phenotype was persistent till 5dpf and displayed additional abnormalities in the ceratohyal cartilage and the basihyal cartilage, which develops post 3dpf and was found affected at 5dpf. 47% of targeted larvae showed persistence in the mild defects previously discussed, and 30% showed persistence of the severe phenotype, preventing us from further raising the affected larvae. The effect of targeting *ctsk* on the musculoskeletal system has been previously investigated in mice and was found to induce osteopetrosis in normal mice (Gowen et al., 1999) and delay the progression of osteoarthritis in OA models (Kozawa et al., 2012). Here we present the effect of *ctsk* targeting in zebrafish on early skeletogenesis, which results in various disruptions to the whole skeletal structure of larvae and general downsizing. Additionally, remarkable jaw cartilage defects were observed due to insufficient cartilage remodelling.

Regarding the heart, 100pmol targeted larvae showed a mild impact on the heart development starting at 3dpf, with the ventricle appearing marginally smaller than in the control group. This is a mere observation that requires a method of separate chamber volumetric measurements to confirm. However, at the higher concentration of 400pmol, the impact is more detrimental, particularly on the looping process, which occur inefficiently resulting in misalignment or inversion of the heart in 80% of 400pmol targeted larvae. At 5dpf, the phenotype remained present in both concentrations. 100pmol *ctsk* targeted larvae exhibited mild shape irregularities and a smaller ventricle in 37% of crispants. In 400pmol, an increased proportion of 57% exhibited the same deformities, with a further 36% of targeted larvae displaying severe defects in the heart, leading to complete disruption of its shape, as seen in **Figure 4.5 g&h.** Additionally, the conducted heart volumetric measurements revealed no significant change between 100pmol or 400pmol and the wildtype control. Interestingly, 400pmol larvae showed significantly lower heart volume *(p-value = 0.0185),* when compared to the 100pmol group, but not to uninjected controls *(p-value= 0.4639).* This suggests a differential effect of the two concentrations, whereby the concentration of 100pmol could possibly induce a partial knockout of the gene, compared to the 400pmol, resulting in an opposite effect on the heart volume. It must be noted that this analysis was performed on a limited number of larvae images, therefore, require further repeats to validate this hypothesis.

To detect if the targeting has an impact on the heart morphology at an earlier stage, the heart structure was examined at 1dpf in *ctsk* targeted larvae using RNA In situ hybridization. A myl7 probe was used to stain the cardiomyocytes and visualise the heart. Heart appears to be affected by the mutation as early as 1dpf, causing misalignment to the heart within the skull. This experiment was only conducted once with 4 crispants tested against 2 wildtype controls, therefore, further repetitions are necessary to confirm this observation. It should also be mentioned that the skull of injected fish appears smaller as compared to controls, which could possibly account for the misaligned position within the skull. Although I examined the detrimental effect of targeting *ctsk* on heart development, various studies previously showed how *ctsk* targeting in some defective states could rather help repair or limit disease progression. *Ctsk* knockout mice showed a restriction in ageing related cardiac dysfunction (Hua et al. 2015). *Ctsk* targeting also reduced cardiac oxidative stress in diabetic mice (Guo et al., 2017). In cardiac hypertrophy mouse models, where the heart muscles thicken resulting in increased blood pressure, *Ctsk* targeting was shown to attenuate the symptoms. The suggested mechanism for this effect is inhibiting rapamycin and ERK pathway signalling, which are typically activated during cardiac hypertrophy (Hua et al., 2013). Again, all studies focused on the association of *Ctsk* in cardiovascular disease mouse models, yet no data is currently available regarding its developmental role within the cardiovascular system. Our data suggest high involvement of *ctsk* in early heart development, as early as 1dpf, that affects its size and orientation within the organism, as apparent via the initial in situ hybridization heart staining data. Different target guide concentrations may lead to a partial or complete knockout of a gene, potentially resulting in opposing effects on the heart volume in targeted crispants. However, it is necessary to sequence individual larvae in order to confirm their genotype and establish a correlation between heart volume phenotype and genotype. This will enable us to arrive at a reliable conclusion.

Due to the abnormalities discussed above in the musculoskeletal system, a consequent impact on larval swimming abilities was suspected. To test this, the duration spent by the larvae at 5dpf in different speed categories was examined in the presence and absence of light, as additional conditions. This analysis was performed using the Viewpoint Zebrafish system. The results showed that 400pmol larvae displayed a statistically significant increase in activity in comparison with 100pmol larvae, but not with uninjected controls. Although the high concentration targeted larvae appear to move more, they exhibit a noteworthy drop in high-speed movement when compared to the 100pmol group, but not when compared to uninjected controls. Surprisingly, this outcome contradicts our expectations of observing a consistent reduction in activity with *ctsk* target concentration increase. However, these findings correspond with the observed pattern in heart volume measurements, where 100pmol larvae exhibit a mild increase in heart size followed by a sharp decrease at 400pmol. This decrease is statistically significant when compared to the lower concentration of 100pmol, but not when compared to the uninjected controls. A possible explanation for the reduction observed in activity duration of 400pmol

larvae, indicating more movement when compared to 100pmol could happen due to the limitation of movement in crispants due to severe skeletal deformities. This allows larvae to display subtle movement or fidgeting that would be detected by the system as movement. The significant reduction in rapid movement goes in line with this hypothesis. Also, the significant reduction in heart volume of 400pmol when compared to 100pmol targeted larvae would theoretically have an impact on crispants movement and add to their inability to display rapid swimming. Another aspect revealed by this data is the effect of Ctsk load on structural development. The use of two concentrations that manifested as different phenotypes highlight the differential effect of partial or complete targeting of *ctsk*. As previously mentioned, further analysis must be performed to confirm this hypothesis and check for partial and complete knockout difference between the two tested concentrations. Generation of mutant lines to create a stable mutation and investigate its specific genotypic effect using sequencing would also provide more information regarding the basis of the detected phenotypes.

As only 100pmol targeted larvae were raised due to the severe detrimental effect of the higher concentration on the embryos, the adult behavioural analysis was solely conducted on the 100pmol *ctsk* crispants at 12 months. Fish showed no alterations in their activity levels in the presence or absence of light, as part of the sex split analysis, collective data, or the 7-hour monitoring period of the experiment. However, a substantial alteration in the 400pmol targeted fish is anticipated due to the severe phenotypes that were solely detected at this concentration and not in the 100pmol *ctsk* targeted fish, but unfortunately, they remain elusive. Therefore, it is essential to repeat the injection with the two different concentrations followed by sequencing to match any detected phenotype or lethality to the genotype. This may also allow us to raise some 400pmol injected larvae if a more tolerable mutation is present.

Finally, by examining the effect of the induced mutation on later bone development using micro-CT scanning, it was observed that there were no significant alterations in terms of bone volume (females: *p-value= 0.1783*) (males: *p-value= 0.8757*), or bone mineral density (females: *p-value= 0.5009*) (males: *p-value= 0.2426*) between 100pmol targeted crispants and uninjected fish. Nevertheless, a significant elevation in bone surface area was detected among adult males *(p-value= 0.0188),* as opposed to females *(p-value= 0.9836).* Irregularities in the bone structure of adults may cause projections, which increases the surface area of the bone without demonstrating an increase in its total volume, thus resulting in an increase in bone surface area. Detailed micro-CT analysis must be performed on selective structures to examine any visible changes in the surface area. As the significant increase was only observed in males, this suggests a possible differential role of *ctsk* in bone maintenance in males, as opposed to females. However, due to the small size of tested organisms, a larger sample number must be tested to confirm the observed results.

Considering our current knowledge regarding the involvement of CTSK in various disease pathogenesis, the need for a specific CTSK modulator has been on the rise and became a target of interest in the drug development sector. Several pharmaceutical companies aim to develop a specific inhibitor for CTSK in attempt of providing a treatment for discussed diseases, especially for common musculoskeletal diseases. Balicatib (AAE581) was one of the first CTSK inhibitors developed by Novartis for the treatment of osteoporosis and knee osteoarthritis. After reaching Phase II, trials were terminated in 2010 due to the rise of undesirable skin related side effect, including rashes and morphea-like changes (Gall et al., 2008). The side effects were suggested to be a result of abnormal collagen accumulation within skin tissues due to the lack of CTSK necessary for excess collagen degradation. However, additional trials are still ongoing as balicatib has demonstrated anti-tumour effects in melanoma and breast cancer (Quintanilla-Dieck et al., 2008; Gall et al., 2007).

Another CTSK inhibitor is Odanacatib, which Merck and Co. developed over a 12-year period. Odanacatib has demonstrated promising outcomes, having a remarkable impact on the treatment of postmenopausal osteoporosis by significantly increasing bone mineral density and reducing fractures (Bone et al., 2010). The drug passed several large-scale clinical studies and was in the process of being approved as an osteoporosis treatment, before trials were stopped in 2016, due to adverse cardiovascular side effects, especially an increase in stroke incidents among study participants (Mullard, 2016). In addition to the disastrous cardiovascular side effects, skin rashes were commonly seen as with Balicatib trials (Bone et al., 2010). A rise in upper respiratory tract infections was also commonly observed (Eisman et al., 2011). Like balicatib, initial trials were terminated, however, there are currently several ongoing trials for the optimization of odanacatib as a potential cancer and anti-metastatic drug with a focus on breast cancer.

As described in **section 4.1.4.**, several studies on cardiovascular diseases have demonstrated an elevation in CTSK levels in patients suffering from CVD. This suggests that inhibiting CTSK should help limit the progression or even recover the heart function. As described above and to our surprise, treatment of OA patients using the CTSK inhibitor Odanacatib resulted in an increase in the risk of developing CVD. Once again, this observation underscores the intricacy of the system, yet it directs us towards crucial measures to be considered for future CTSK drug targets. One of which is the reduction of off-target effect by ensuring the drug is tissue specific to avoid undesirable inhibition in other organs or tissues. Secondly, the dose of administrated inhibitor should be given careful consideration. As with our heart volumetric measurements and larval behaviour where we observed different and sometimes opposite effects on heart volume and larval activity using two concentrations of CRISPR guides possibly by inducing a partial knockout or a complete knockout. Similarly, carefully considering the partial or complete inhibition of CTSK could change its effect as a treatment. Most importantly, the interactions between CTSK and other pathways should be further studied to understand how these drug

targets unintentionally affect related pathways. This will be discussed further in following chapters to display data obtained in this project which reveal genes and pathways linked to CTSK.

A proposal for future investigation regarding the negative impact of odanacatib involves exploring cardiac calcification. Cardiac calcification entails the deposition of calcium in heart valves and walls, which can lead to aortic stenosis and coronary artery calcification. Cardiac calcification has been associated with an increased risk of strokes and blood clots (Hermann et al., 2013; Wang et al., 2022), which was frequently observed in OA patients post odanacatib treatment. Some mechanisms that can lead to heart calcification include chronic inflammation resulting in a significant rise in Interleukin-1β (IL-1β). As a result, vascular smooth muscle cells differentiate into osteoblast-like cells, forcing them to act like bone cells and produce MMPs (Alves et al., 2014; Ceneri et al., 2016). Such inflammation can be caused due to the CTSK imbalance and the effect of CTSK on inflammation have been found in several conditions (Hao et al., 2015; Yue et al., 2020). Moreover, as CTSK is involved in ECM regulation, inhibition of CTSK can disrupt its regulation and allow for matrix vesicles containing phosphates, calcium, along with other ions to fuse to the heart walls, similar to bone mineralisation in bone formation and hardening (Li et al., 2022).

There are several limitations to consider in the presented data. Firstly, the range of phenotypes obtained in different crispants, which makes pinpointing a specific defect correlated to the mutation more challenging. Obtaining precise cartilage measurements could have disclosed distinct patterns in jaw cartilage morphology in *ctsk* crispants. However, these measurements were not undertaken across the extensive variety of phenotypes observed, resulting in significant detrimental effects, which in many cases impeded the identification of certain structures. Additionally, the severe detrimental effect of 400pmol *ctsk* sgRNA targeting prevented us from raising the larvae to adulthood to avoid any suffering to the organism. Therefore, adult behavioural analysis and micro-CT scans were not performed on the 400pmol targeted fish. Further repeats are necessary to validate the obtained results including testing more larvae using in situ hybridization to confirm the heart phenotype rising as early as 1dpf and replicates for larval behaviour analysis, as only one replicate was performed for this experiment. More images of crispant hearts would be beneficial to analyse their total volume and validate initial obtained data. Negative controls were siblings obtained from the same parents to injected fish. For the *Tg(myl7:lifeActGFP)* injections utilised for heart visualisation, a tyrosinase control was added to ensure detected phenotype is unrelated to the injection process. However, using tyrosinase controls would also be advantageous for the *Tg(Col2a:mCherry)* injections for cartilage phenotype comparison. Yet, cartilage *ctsk* injections were performed in parallel to other genes in separate larvae, and observed phenotype was not present in larvae targeted for other genes. Most importantly, generation of a mutant line by outcrossing raised crispants to wildtype to follow up with clean F2 generation carrying a single

mutation with elimination of any possible off targets would help us understand the effect of exact induced mutation by linking it to their phenotype.

5. Investigating the Role of *gdf5* in the Development of the Skeletal and Cardiovascular Systems

5.1. Introduction

Growth/Differentiation Factor 5 (GDF5), also referred to as, Cartilage-derived morphogenetic protein 1 (CDMP1), Lipopolysaccharide-associated Protein 4 (LAP4) or Bone Morphogenetic Protein 14 (BMP14), belongs to the TGF-β/BMP superfamily and acts as a growth factor (Ducy and Karsenty, 2000). The protein precursor encoded by the gene plays a critical role in the development and control of bone, cartilage, and tendon tissues (Hotten et al., 1996; Francis-West et al., 1999). It forms an active ligand, which binds to type I and type II serine-threonine kinase receptors such as BMPR1B. In humans, *GDF5* mutations are known for their effect on the musculoskeletal system and cause several heritable diseases, including Acromesomelic dysplasia, Brachydactyly (OMIM reference: BDA2, MIM 112600; BDC, MIM 113100; MIM 228900), and Grebe- and Hunter-Thompson-type chondrodysplasia (AMDH, MIM 201250), which are characterized by the shortening of the skeletal elements and abnormal development of some joints.

5.1.1. Cellular function of GDF5

GDF5 and BMPs are members of the TGFβ family. GDF5 has a similar structure to BMPs, enabling it to bind to their receptors and is known to be highly involved in the positive and negative regulation of the BMP signalling pathway. The BMP pathway is known for its involvement in the development of the skeletal system. In line, GDF5 has been known for its critical role during joint formation at the site of cavitation (Storm and Kingsley, 1996). In-vitro models have confirmed the anabolic function of GDF5 in maintaining cartilage. This has been achieved through the identification of a direct role of GDF5 in chondrogenic differentiation, along with chondrocyte hypertrophy, through the BMP signalling pathway. This latter is an essential transitional stage in the ossification process where cartilage develops further to form mature bones (Coleman et al., 2013).

Regarding its specific signalling, GDF5 was found to activate the BMP pathway by binding to the BMPR1B receptor, as well as BMPR1A, with a lower affinity (Nickel et al., 2005). Recently, a study found that binding of GDF5 to BMPR1B receptor compared to BMPR1A provide higher stability to the developed cartilage by limiting chondrocyte hypertrophy, commonly observed when GDF5 binds to BMPR1A (Mang et al., 2020). Both are transmembrane serine/threonine kinase receptors, which get phosphorylated following the binding process and transmit the signal through a cascade of SMADs, including SMAD1, SMAD5 and SMAD8. The 1/5/8 SMADs complex can then bind to the

unphosphorylated SMAD4 and translocate to the nucleus where it increases transcription of TGF-β target genes, including Collagen-I and Aggrecan (Juarez et al., 2010), by binding to intranuclear coactivators. This leads to a favourable regulation of chondrogenic tissue differentiation through the MAPK pathway. Subsequently, it results in an increase in cell proliferation and osteoblast differentiation, building the cartilage structure during early skeletogenesis.

The MAPK pathway was found to be a crucial downstream pathway for GDF5. When comparing *Gdf5* knockout mice to wildtypes, a significant boost in SMAD phosphorylation, accompanied by an inhibition to the MAPK pathway activity was observed in cardiomyocytes. This study revealed the practical significance of GDF5 in the MAPK pathway, while highlighting the variety of alternative cytokines/growth factors involved in SMAD complex assembly and phosphorylation (Chen et al., 2006). On the other hand, GDF5 can negatively regulate the BMP pathway upon interacting with the protein Noggin, encoded by the *NOG* gene. This interaction hinders chondrogenic differentiation, aiming to maintain equilibrium and prevent excessive muscle loss, particularly after denervation. Denervation refers to the loss of nerve supply to the muscle, leading to muscle mass loss. The observed muscle weakening indirectly affects cartilage by increasing the movement load for compensation. One missense mutation in *GDF5* was found to increase the prochondrogenic activity and create insensitivity to noggin interactions, resulting in the development of synostosis syndrome. This condition is characterised by the fusion of the carpal and tarsal bones, often leading to an irregularly shaped skull (Seemann et al., 2009).

5.1.2. Role of *GDF5* in the skeletal system function and development

Studies performed using mouse models revealed an early involvement of the *Gdf5* gene in skeletogenesis. High expression levels of *Gdf5* were detected at the beginning of the skeletogenesis process in the mesenchymal condensation stage in knee and elbow joints. This assists in the positioning of the articulation between the metacarpal and proximal phalanges bones (Storm and Kingsley, 1996). The critical involvement of GDF5 in the skeletal development was also confirmed when the overexpression of *Gdf5* showed an increase in size of the skeletal volume during early condensation till after the formation stage (Buxton et al., 2001). *Gdf5* was found to play a role in mesenchymal cell recruitment and proliferation of chondrocytes in both mice and chickens (Sumaki et al., 1999; Buxton et al., 2001). The expression of *Gdf5* is maintained at high levels during chondrogenesis. However, during joint maturation and synovial joint cavitation, *Gdf5* levels show a significant decrease, and its expression is limited to the lateral joint edges (Merino et al., 1999). Its expression is detected in the epiphyseal cartilage to increase proliferation and finally allow for the transformation of cartilage into osteoblasts to form mature bone structure in chickens (Buxton et al., 2001). One study that advanced our understanding of *Gdf5* function during early development examined the role of *Bmp4* and *Gdf5* as

two essential cartilage formation and remodelling regulators. It was found that *Bmp4* is mainly involved in the formation and maturation of cartilage nodules and can also signal to mesenchymal cells to express different cartilage regulatory markers. However, *Gdf5* had a distinctive role mainly in prechondrogenic mesenchymal condensation enhancing cell aggregation and increasing the cell response to differentiation markers, such as BMP4. Moreover, *Gdf5* did not display a direct effect on cartilage nodules or a direct regulatory effect on the expression of involved differentiation markers, as seen with BMP4 (Hatakeyama et al., 2004). This data highlights the importance of *GDF5* during early skeletogenesis.

5.1.3. *GDF5* expression pattern in osteoarthritis

One study found a significant reduction in GDF5 mRNA and protein expression in osteoarthritis cartilage when compared to healthy cartilage (Mauck et al., 2006). However, recent studies suggest the opposite observation. Kania and Colella, 2020 exhibit a rise in *GDF5* gene expression levels in the articular cartilage and synovium from OA patient samples. Additionally, they found the same expression pattern in mice following a Destabilization of the Medial Meniscus (DMM) surgery. DMM is a well-established model for the induction of osteoarthritis in mouse models that involves creating a small incision in the knee joint to destabilise the medial meniscus. Another more recent study confirmed this observation and found a direct correlation between OA severity and the increase in *GDF5* expression levels in patients' synovial tissue. In the study, OA progression was assessed using the Kellgren and Lawrence classification (KL) system in grade 3 and 4 osteoarthritis patients. *GDF5* gene expression levels were measured using quantitative Real-time PCR. An increase in *GDF5* with a median fold change of 1.81 was found in KL3 samples, while a remarkable 3.50 median expression was observed in KL4 patient samples. These findings demonstrate the correlation between GDF5 gene expression levels and disease progression and severity (Witoonpanich et al., 2022).

5.1.4. Role of *GDF5* in cardiovascular system regulation

Limited information exists on the exact function of *GDF5* in heart development or cardiovascular disease progression. Nonetheless, some studies indicated a role of *GDF5* associated with some cardiovascular diseases (Zaidi et al., 2010; Shikatani et al., 2023). A 2010 study investigated the role of *Gdf5* in cardiac repair following a surgically induced myocardial infarction (MI) in wildtype and *Gdf5* knockout (*Gdf5*-KO) mice. Wildtype mice exhibited an increase in *Gdf5* expression after the surgery, suggesting a role in tissue repair. At 4 days post-surgery, *Gdf5*-KO mice displayed an increase in cardiomyocyte apoptosis with a decline in anti-apoptotic gene expression levels. 28 days post-surgery, *Gdf5*-KO exhibited an expansion of scar tissue in the infarct region with thinning of the arteriolar wall density when compared to wildtypes and an enlargement of the left ventricle. This

indicates that its reparative mechanism of GDF5 includes limiting tissue disruption after MI, supporting cardiac repair via an anti-apoptotic effect (Zaidi et al., 2010). The same research group has recently published a follow-up paper that focuses on the role of *Gdf5* in cardiac rupture post-MI. In this study, loss of *Gdf5* reduced the risk of post-MI cardiac rupture and, as a result, decreased the mortality rate. Additionally, *Gdf5*-KO mice showed a lower heart to body weight ratio and increased expression of ECM genes. To summarise, *Gdf5* appears to be crucial in the process of cardiac repair following myocardial infarctions, potentially by inhibiting apoptosis to limit tissue scarring expansion. It can also induce the MAPK pathway, activating cardiac fibroblast differentiation (Zaidi et al., 2010). Moreover, *Gdf5* plays a vital role in regulating cardiac ECM following MI (Shikatani et al., 2023). Due to the involvement of GDF5 in critical structural development process such as ECM regulation and fibroblast cell differentiation, we hypothesise a critical role of the gene in heart development, which has not been investigated yet.

5.2. Hypothesis and aims

Based on previous *Gdf5* studies in mice, combined with several human studies and GWAS data, which indicate significant involvement of *GDF5* in the musculoskeletal system (Storm and Kingsley, 1996; Tachmazidou et al., 2019), I hypothesise a similar function of *gdf5* in the musculoskeletal system development of zebrafish. Additionally, following the identification of the role of *Gdf5* in cardiac tissue repair post myocardial infraction in mice (Zaidi et al., 2010; Shikatani et al., 2023), as well as several CVD related GWAS studies identifying *GDF5* as a hit biomarker (Evangelou et al., 2018; Xie et al., 2020; Groenland et al., 2022), I hypothesise that *gdf5* plays a crucial role in musculoskeletal and heart development in zebrafish.

To test this hypothesis, my objectives are to:

- Investigate the effect of targeting the *gdf5* gene using CRISPR-Cas9 mutagenesis on the structural development of the skeleton.
- Explore the role of *gdf5* in heart development, through its depletion, using zebrafish models.
- Examine fish movement efficiency and speed in order to determine the potential functional role of *gdf5*, and the effect of its loss across generations.
- Assess the impact of *gdf5* disruption on the developed adult bone structure.
- Investigate the role of *gdf5* in tissue repair post-injury using the caudal injury regeneration model.

5.3. Results

5.3.1. Zebrafish *gdf5* shares high homology with human *GDF5*

To determine the relevance of studying *GDF5* in zebrafish, it is necessary to have sufficient similarity between the gene and proteins, when compared to the human variants. We found that *GDF5* is considered a conserved gene, located on different chromosomes in different species. For instance, in humans, the gene can be found on chromosome 20: 35,433,347-35,454,746, while in zebrafish, it is located on chromosome 6: 52,637,308-52,661,824, showing 82% orthologue synteny. *GDF5-201* is the verified transcript found in both humans and zebrafish, comprising two exons as shown in **Figure 5.1a**. However, an additional transcript, *GDF5*-202, was also discovered, comprised of 4 exons in humans and 2 exons in zebrafish. The encoded Gdf5 protein shows 64.6% similarity in identity to human GDF5 (Zebrafish: ENSDARG00000002760 and Ensembl Genome Browser, Human: ENSG00000125965). The protein domains in both species demonstrate a TGFβ propeptide and a TGF-beta-GDF5 domain with a homodimer interface, accompanied by additional smaller domains, DUF3767, PTZ00449 and PHA03247, that are exclusively present in the human protein (**Figure 5.1b**).

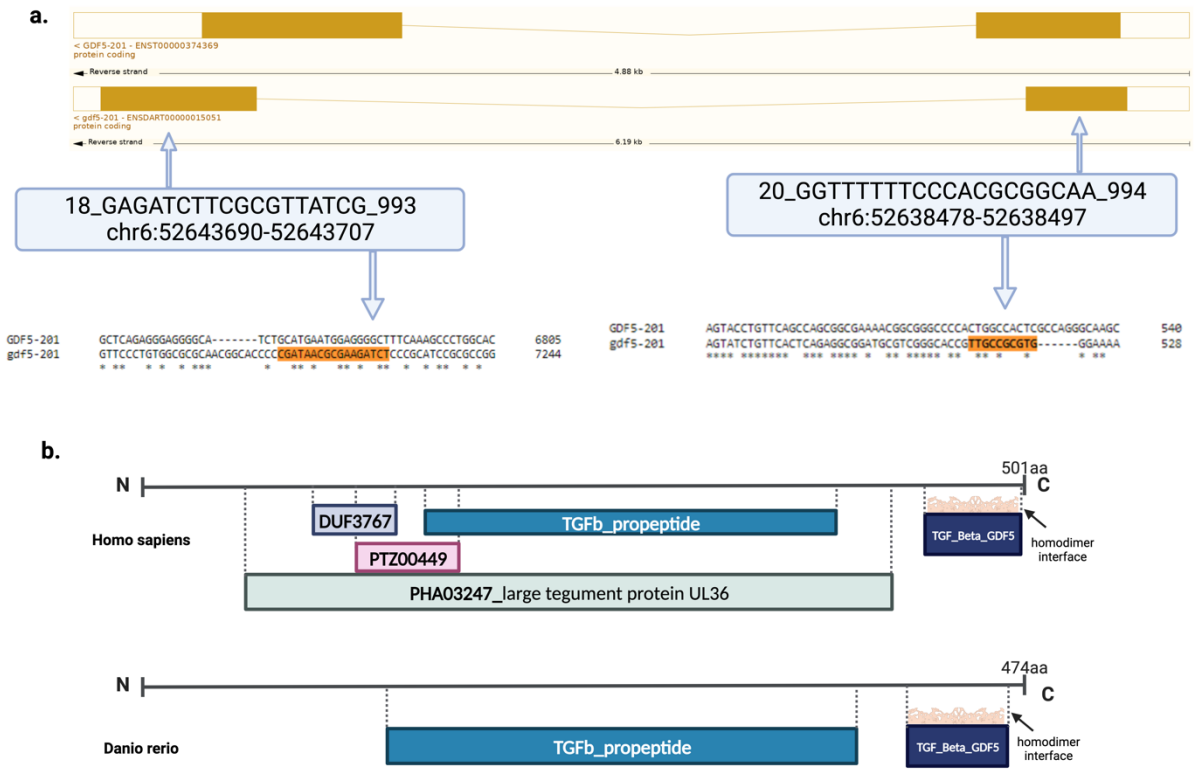


Figure 5.1: *GDF5* gene composition and protein domain

(a) *GDF5* gene composition in humans (top) and zebrafish (bottom). The blue rectangles show selected CRISPR target guides in zebrafish with their exact location. The targeting scores for guide 1 is 994 and for guide 2 is 993 respectively (blue rectangle, left) based on data from the UCSC Genome Browser (UCSC, 2019; https://genome.ucsc.edu/). **(b)** GDF5-201 protein domain in humans (top) composed of five highly conserved protein domain structures. The two main domains - TGF-β propeptide and a TGF-β domain – have also been found in zebrafish (bottom), with a similar homodimer interface for active binding to TGF-β/BMP receptors.

5.3.2. Jaw cartilage structure appear unaffected in *gdf5* crispants

Targeting *Gdf5* in mouse models led to mutants with severe skeletal defects, such as a shorter figure, curved long bones, loss of joint formation ability and rather a fusion of the bones due to disruption in articulation between the metacarpals and proximal phalanges bones (Settle et al., 2003). Based on these findings, it was anticipated that zebrafish crispants would display skeletal abnormalities, including an absence of articular joint formation, alongside deformities in chondrocyte shape and organization within the jaw. To investigate this hypothesis, two CRISPR-Cas9 guide targets were injected along with Cas9 protein to target both exons of *gdf5* in single cell embryos. To our surprise, targeted embryos showed no morphological alterations in the general body structure of *gdf5* modulated larvae (**Figure 5.2 E&F**). *gdf5* crispant larvae appear to develop normally to adulthood and were able to be raised to reach the F2 generation similar to the uninjected control larvae **(Figure 5.2 A&B)**, as well as the tyrosinase negative control (**Figure 5.2 C&D**). To detect the effect of *gdf5* targeting on early skeletogenesis, the jaw cartilage structure in CRISPR targeted larvae was examined using lightsheet microscopy and the *Tg(Col2a:mCherry*) transgenic line to visualise the cartilage at 5dpf. No significant morphological changes were detected between control uninjected group (**Figure 5.2 a&b**), tyrosinase negative injection controls (**Figure 5.2 c&d**), and the *gdf5* crispants (**Figure 5.2 e&f**), regarding the morphology of the jaw cartilage. Cartilage structures examined, including the Meckel's, ceratohyal, basihyal, synovial joint and palaquadrate cartilage, all seemed to develop normally despite successful targeting of *gdf5,* as confirmed by the corresponding genotype (**Figure 5.2 e'&f'**).

Although the above *gdf5* targeting injections has been performed four times separately to ensure sufficient repeats, the data from a single repeat is presented. The remaining three attempts were unsuccessful due to unhealthy uninjected control embryos, which by 1dpf showed 52% death in repeat 1, 64% mortality in repeat 2 and 62% mortality in repeat 3. Furthermore, abnormal phenotypes were also observed in a number of the uninjected controls, leaving the obtained data unreliable for further analysis.

Finally, an examination was conducted on the synovial joint structure within the zebrafish jaw to determine if the joint has developed with the expected gap necessary for appropriate articulation and movement or if the cartilage fused together as seen in mice (Storm and Kingsley,1996; Settle et al., 2003). **Figure 5.3** shows some representative images of synovial joint close-ups, which show the expected gaps within the joint of *gdf5* crispants (**Figure 5.3 c&d**). Although the gap size was not quantified, the initial obtained data suggest that unlike in mice, *gdf5* is unnecessary for early cartilage development or joint formation in zebrafish. However, further repeats must be performed to confirm the observation and it is vital to monitor over an extended period to detect the impact on juvenile skeletal

development. This is necessary to achieve a comprehensive understanding of the function performed by *gdf5* in skeletal development.

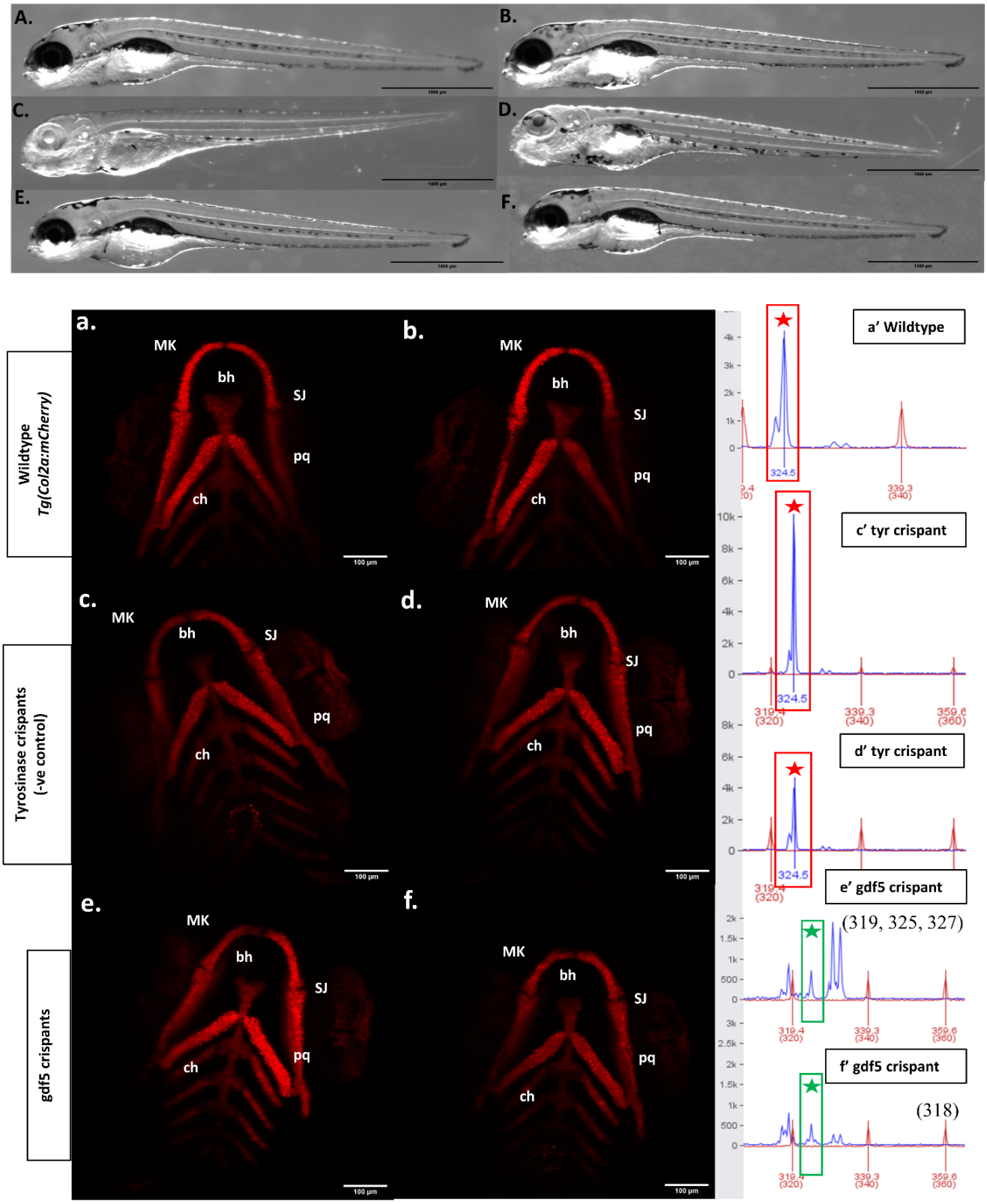

A.
B.
C.
D.
E.
F.
Wildtype
Tg(Col2a:mCherry)
Tyrosinase crispants
(-ve control)
gdf5 crispants
a.
b.
c.
d.
e.
f.
MK
bh
SJ
pq
ch
100 µm
a' Wildtype
c' tyr crispant
d' tyr crispant
e' gdf5 crispant
f' gdf5 crispant
324.5
339.3
(340)
319.4
(320)
359.6
(360)
(319, 325, 327)
(318)

Figure 5.2: Unaffected jaw cartilage development in *gdf5* crispants at 5dpf

The top panel examining the overall morphology, while the bottom panel exhibits the corresponding jaw cartilage structure within the larvae presented. **(A&B)** general morphology of wildtype controls imaged in brightfield ventral view. **(C&D**) tyrosinase targeted larvae as negative controls displaying clear loss of pigmentation as visual indication of successful targeting. **(E&F)** *gdf5* crispants at 5dpf. **(a&b)** control uninjected fish obtained from the *Tg(Col2a:mCherry)* transgenic line representing normal jaw cartilage structure in red at 5dpf (lateral view) (n=5) along with its corresponding genotyping obtained using CRISPR-STAT. **(a')** represent the wildtype allele peak highlighted by a red star at 324.5 indicating intact *gdf5.* **(c&d)** 400pmol tyrosinase CRISPR targeted larvae, as negative controls confirming normal jaw shape and expected wildtype allele peaks using *gdf5* primers (n=2). The green stars indicate the expected peak position when absent. **(c'&d')** tyrosinase targeted larvae genotyping showing wildtype peak for *gdf5*. **e&f** 400pmol *gdf5* crispants showing unaffected jaw structure development up to 5dpf even in the absence of WT peak (n=8) with other insertions in the gene confirming successful targeting **(e'&f').** MK = Meckel's cartilage, bh= Basihyal, ch= ceratohyal, pq= palaquadrate, sj= synovial joint. Scale bar = 100µm.

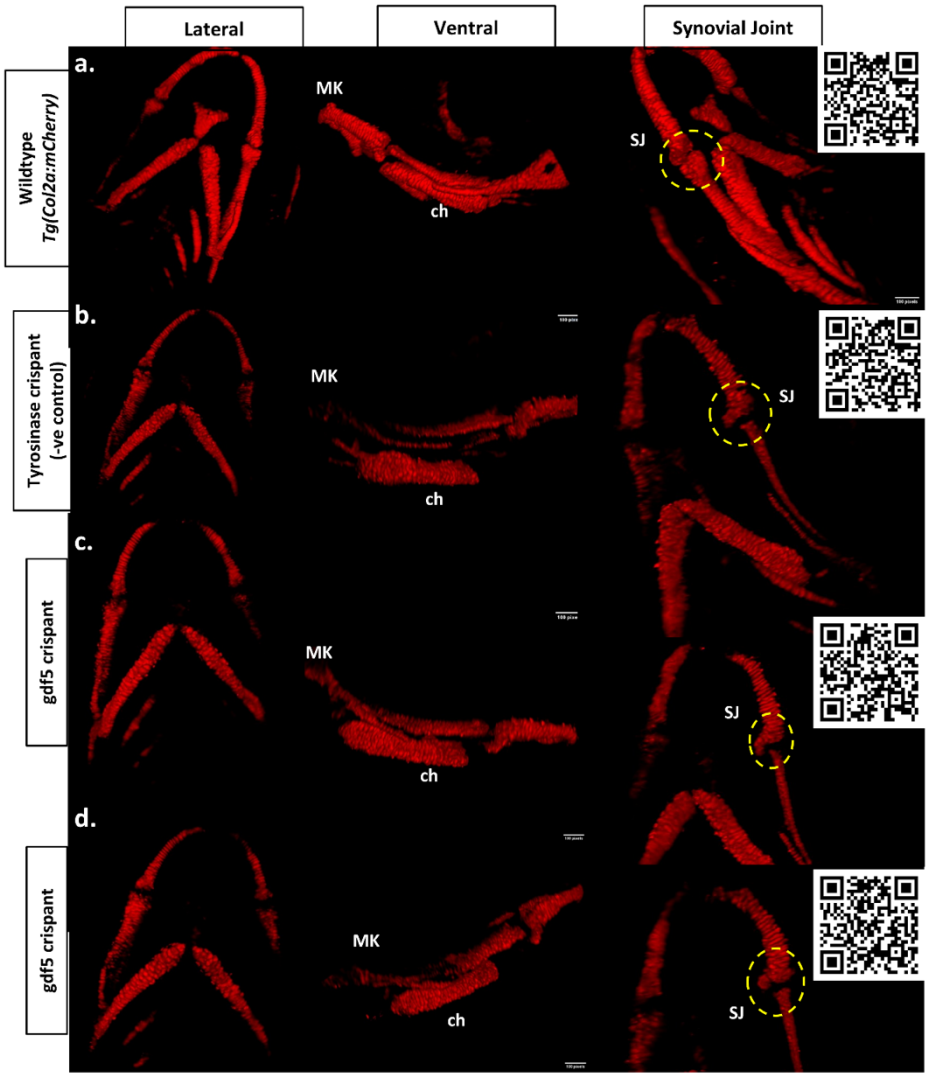
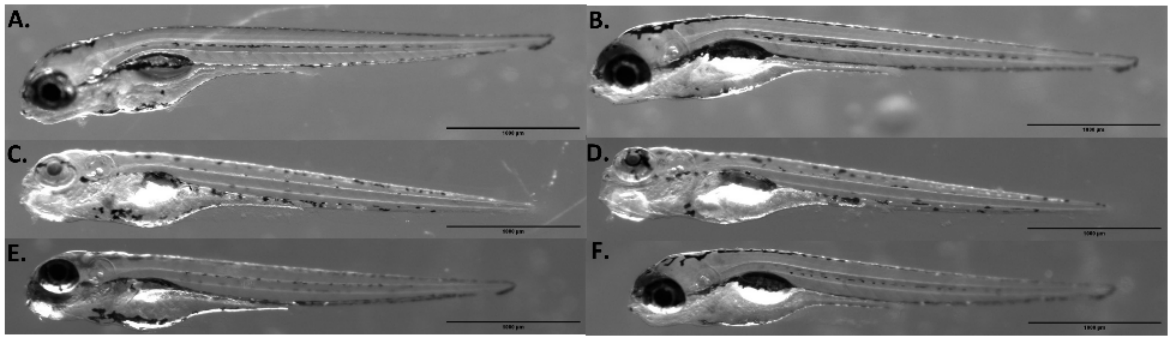


Figure 5.3: Normal development of synovial joint structure in *gdf5* crispants

Figure displays jaw cartilage with corresponding close-up images of the synovial joint structure in the jaw to examine successful separation required for articulation. **(a)** Wildtype uninjected larvae obtained from the *Tg(Col2a:mCherry)* transgenic line representing normal jaw cartilage structure in red at 5dpf with expected synovial joint structure and a gap highlighted by dotted yellow circle (n=5). **(b)** Tyrosinase CRISPR-Cas9 targeted larvae as negative controls displaying regular synovial joint development that is unaffected by the injection (n=2). **(c&d)** Two representative structures for *gdf5* crispants jaw cartilage showing normal development including synovial joint with clear expected articulation gaps (n=8). MK = Meckel's cartilage, ch= ceratohyal and SJ= synovial joint. Scale bar= 100 pixels. To view videos of the 3D constructs, please scan the corresponding QR code.

5.3.3. Morphological heart development is unaffected in *gdf5* crispants

Due to the limited research available on the role of *GDF5* on heart structural development, yet growing evidence from studies suggesting its association with cardiovascular diseases (**Section 5.1.4**), alongside analysed GWAS data demonstrating a potential role of *GDF5* in the cardiovascular system, a series of experiments were designed to explore the effect of *gdf5* targeting on heart development. The functional gene was depleted through the use of CRISPR-Cas9 genome editing in zebrafish embryos. To clearly monitor the heart, CRISPR guides were injected in the *Tg(myl7:lifeActGFP)* line to visualise the heart in green using lightsheet imaging at 5dpf. This stage was selected to visualise the general structure of the heart, and to examine the heart rotation, which would only be visible starting at 5dpf. *gdf5* crispants (**Figure 5.4 E&F**) showed no significant alterations in their general morphology nor in the heart structure when compared to either the uninjected nacre controls (**Figure 5.4 A&B**) or to the injected tyrosinase targeted negative controls (**Figure 5.4 C&D**, **c&d**). The heart develops uniformly in both chambers (**Figure 5.4 e&f**). Additionally, a volumetric measurement analysis of the heart was performed to examine any independent effects on heart size in *gdf5* crispants. The analysis revealed no significant change in the observed size of the *gdf5* targeted heart when compared to uninjected controls (*p-value= 0.0966*) (**Figure 5.5**), which suggests that *gdf5* is unnecessary for early heart development in zebrafish. Furthermore, the heart looping process proceeded normally at 5dpf, which confirms that *gdf5* has no involvement in heart looping. The acquired data provides no evidence for the involvement of *gdf5* in early heart development. However, it should be noted that this experiment was only performed once with two wildtype controls tested against 7 crispants and therefore require further repeats to validate the result.


A.
B.
C.
D.
E.
F.

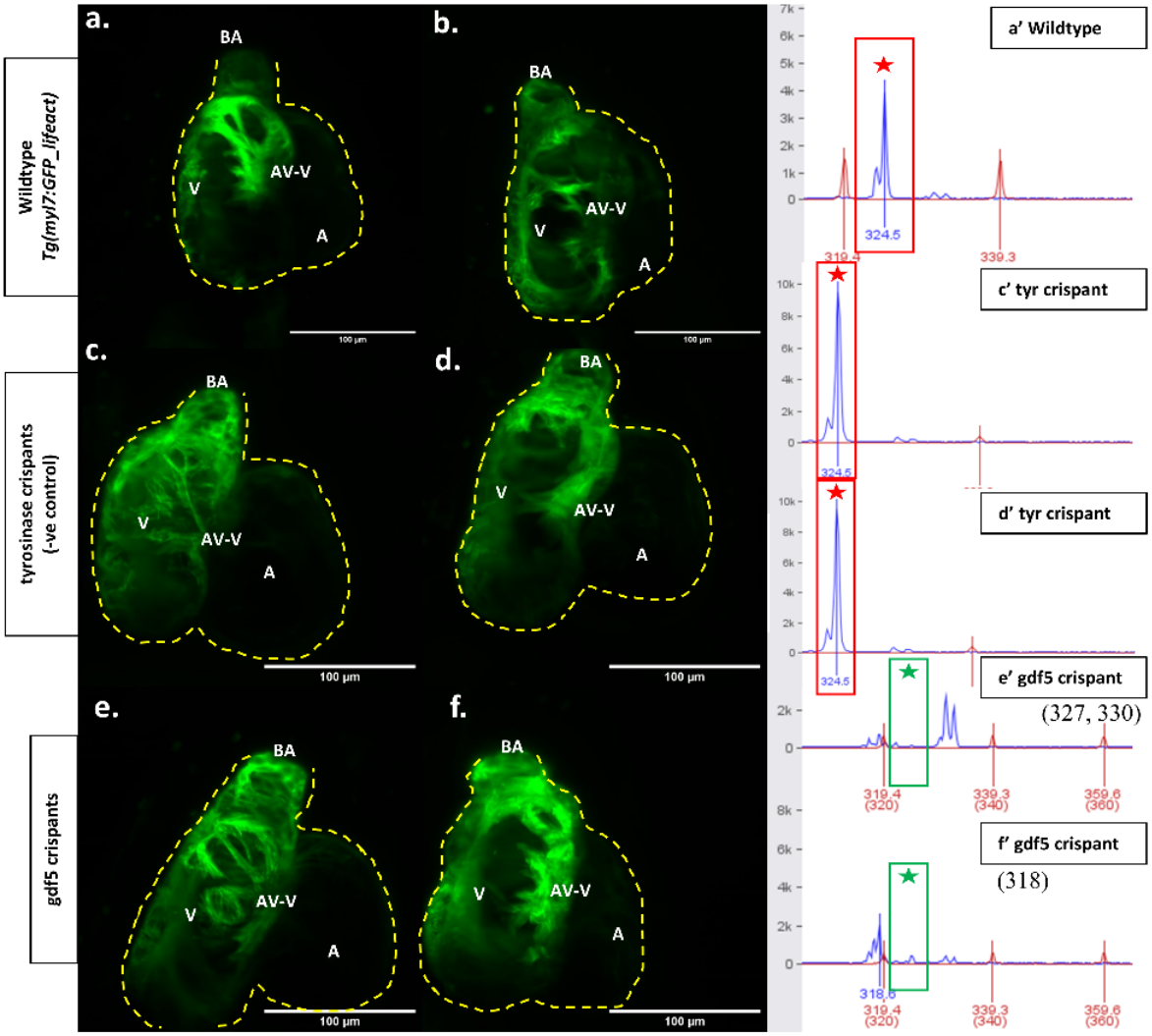

Wildtype
Tg(myl7:GFP_lifeact)
tyrosinase crispants
(-ve control)
gdf5 crispants
a.
b.
c.
d.
e.
f.
BA
AV-V
V
A
100 µm
a' Wildtype
c' tyr crispant
d' tyr crispant
e' gdf5 crispant
(327, 330)
f' gdf5 crispant
(318)
324.5
319.4
339.3
359.6
(320)
(340)
(360)
318.5

Figure 5.4: Heart morphology appears unaffected in *gdf5* crispants at 5dpf

The top panel depicts the general morphology, while the bottom panel displays the corresponding heart of the presented larvae. **(A&B)** General morphology of wildtype controls imaged in a brightfield ventral view. **(C&D)** Tyrosinase targeted larvae as negative controls that display clear loss of pigmentation, indicating successful targeting. **(E&F)** *gdf5* crispants at 5dpf displaying morphology similar to wildtype larvae. **(a&b)** Control uninjected larvae were obtained from the *Tg(myl7: lifeActGFP)* line representing normal heart structure in green at 5dpf (n=3) outlined by yellow dotted line (to simplify visualisation of low fluorescent atrium), along with its corresponding genotyping obtained using CRISPR-STAT**. (a')** represent the wildtype allele peak highlighted by a red star at 324.5 indicating intact *gdf5.* **(c&d)** Tyrosinase CRISPR targeted larvae, as negative controls confirming the normal heart shape and expected wildtype allele peaks using *gdf5* primers (n=2). **(c'&d')** Tyrosinase targeted larvae genotyping showing wildtype peak for *gdf5*. **(e&f)** Targeted *gdf5* larvae showing unaffected heart structural development (n=7) up to 5dpf even in the complete absence of WT peak as shown by the corresponding genotyping in **(e'&f')**. Green stars indicate the expected peak position when absent. A= Atrium, AV-V= Atrioventricular Valve, BA= Bulbus Arteriosus, V= Ventricle. Scale bar for the top panel = 1000µm, bottom panel scale bar = 100µm.

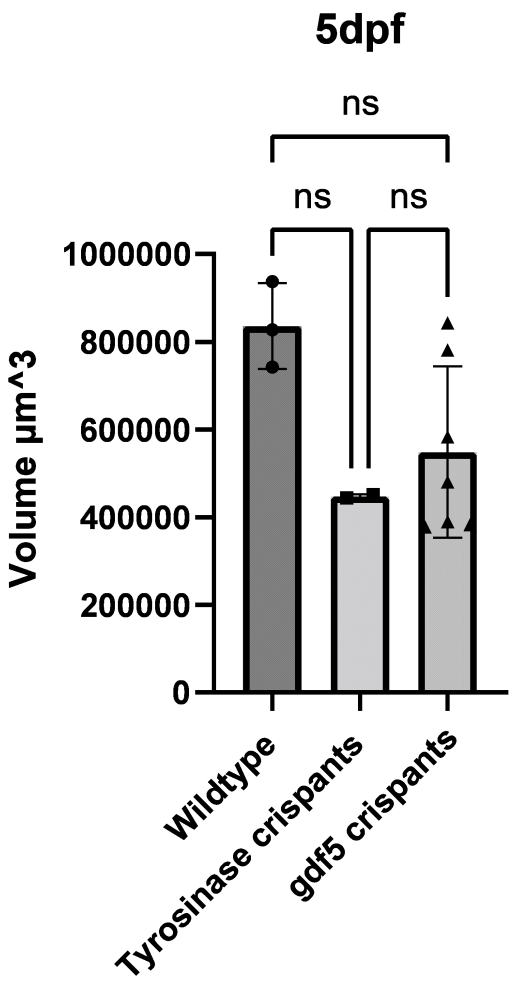


Figure 5.5: Unaffected heart volume in *gdf5* crispants at 5dpf

The figure presents the heart volume of *gdf5* crispants (n=7) compared to uninjected controls (n=3) (*p-value= 0.0966*). Tyrosinase targeted injected negative controls (n=2) were also compared to uninjected controls (*p-value= 0.0886*) and *gdf5* crispants showing no significant change in crispants heart volume (*p-value= 0.8486)*. Heart appears to develop normally without any significant changes in the total volume when compared to both controls. Data analysed using a One-way Anova statistical test with multiple comparisons on Graphpad Prism.

5.3.4. Effect of *gdf5* targeting on adult zebrafish behaviour of founders and F1 generation

Although *gdf5* targeting did not considerably affect the skeletal structure of the generated crispants, the targeting effect on skeletal functionality was tested by examining any swimming abnormalities. To test the efficiency of the crispants' swimming, crispants were raised until 11 months old to assess the effectiveness of *gdf5* modulation on the movement speed of fish in adulthood. Adult behavioural analysis protocol was performed as described in **section 2.7.2**. 10 fish were tested per experiment, initially with 3 wildtype nacre controls (2 females and 1 male) and 7 F0 crispants (5 females and 2 males) for the first two technical replicates, conducted on separate days. Nevertheless, the two male founders did not survive to the final replicate of the behavioural experiment, which took place 17 months later. Therefore, for the final replicate, 5 female nacre controls were tested versus the 5 female founders. Since *GDF5* is known to be involved in the nervous system as a regulator of neuronal axons and dendrites development, all three replicates were carried out both in the absence of light and in the presence of light, which may help detect any abnormalities in the behaviour due to a disruption to the nervous system as shown in previous studies (Sullivan et al., 2005; O'Keeffe et al., 2016). Swimming behaviour was monitored for 8 hours and the first hour was removed from the analysis to allow for adaptation time to the new environment in the Viewpoint Zebrafish Box behavioural room.

Results indicate unaffected movement of the *gdf5* crispants in both conditions of presence (**Figure 5.6**) and absence of the light (**Figure 5.7)**. *gdf5* F0 adult crispants are as active as wildtype controls, and they spend similar duration in both an inactive and hyperactive state. We also investigated whether fish display different activity according to their sex. However, no correlation was found between the sexes and movement pattern (**Figure 5.6 a&b**) (**Figure 5.7 a&b**). All *gdf5* crispants displayed a movement pattern and behaviour similar to wildtype controls highlighting that the gene is unnecessary for the development and function of adult skeletal structure.

Although no behavioural changes were observed in the gdf5 founders, I sought to investigate if the mutation would reveal a phenotype in subsequent generations. Therefore, using the same behavioural analysis experimental conditions, we examined the effect of the different mutation on the incrossed F1 generation (**Figure 5.9 & 5.10)**. Three technical replicates were performed using 4 wildtype control adults tested against 6 gdf5 F1 mutant (**Figure 5.12**). Sequencing of tested F1 mutants is displayed in **Figure 5.9** (females) **& Figure 5.10** (males). Each experiment repeated three times in two separate conditions in the presence and absence of light. No differences were observed between the controls and the F1 generation in both conditions (**Figure 5.12).**

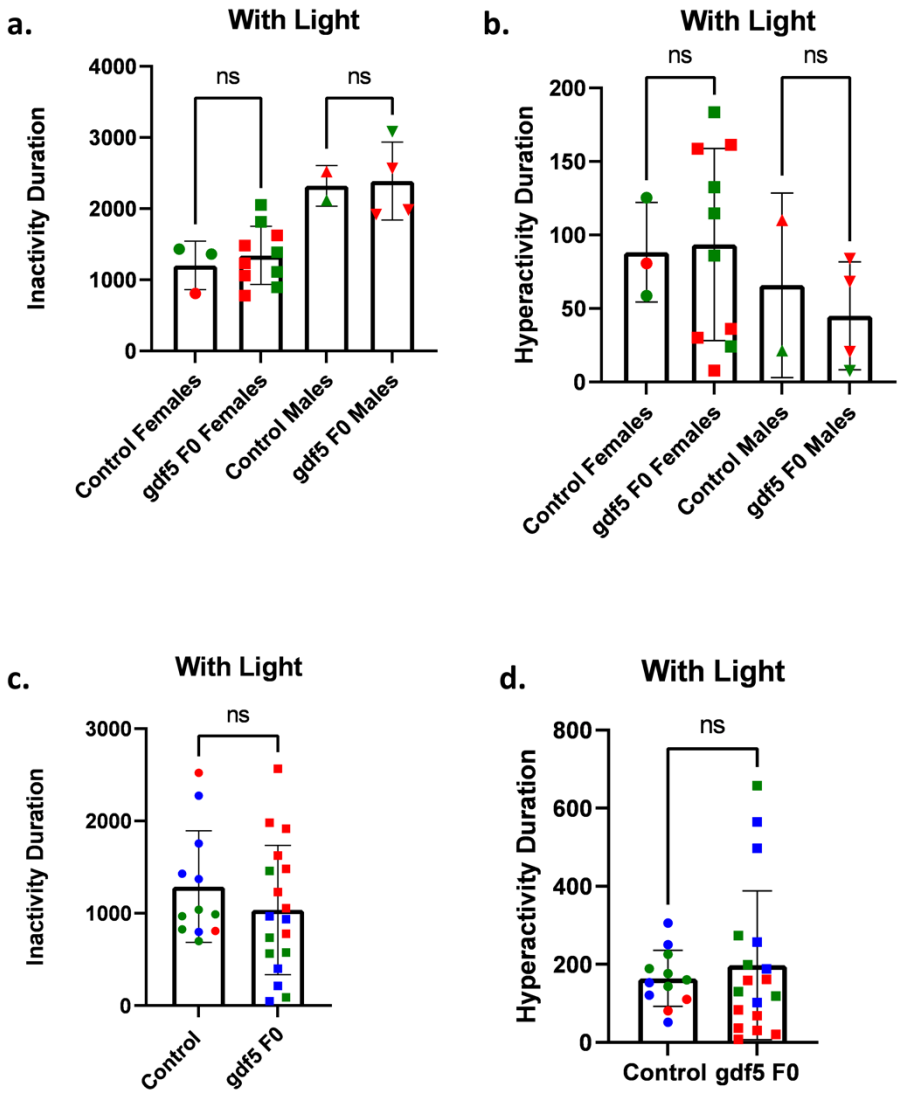


Figure 5.6: Targeting *gdf5* in zebrafish larvae does not affect the fish swimming behaviour in adulthood in the presence of light

(a) No significant change in time spent in an inactive state was observed for both sexes, females *(p-value= 0.8533)* and males *(p-value= 0.9801)*. **(b)** There was no significant difference in hyperactivity duration for both sexes, with females *(p-value= 0.9878)* and males *(p-value= 0.8975)*. **(c&d)** There was no significant change in swimming speed of total adult fish in the presence of light, inactivity duration *(p-value= 0.3126)* and hyperactivity duration *(p-value= 0.5661)*. **(a&b)** Statistical analysis was conducted as one-way Anova, while (**c&d**) through a student t-test comparing the two columns. Each data point represents the average time spent by a single fish under the given conditions during the 7-hour experimental period. Red= first replicate (n=2 uninjected nacre controls and n=8 *gdf5* crispants), Green= second replicate (n=3 uninjected nacre controls and n=7 *gdf5* crispants), Blue= third replicate (n=5 uninjected nacre controls and n=5 *gdf5* crispants). Figure generates using Graphpad Prism.

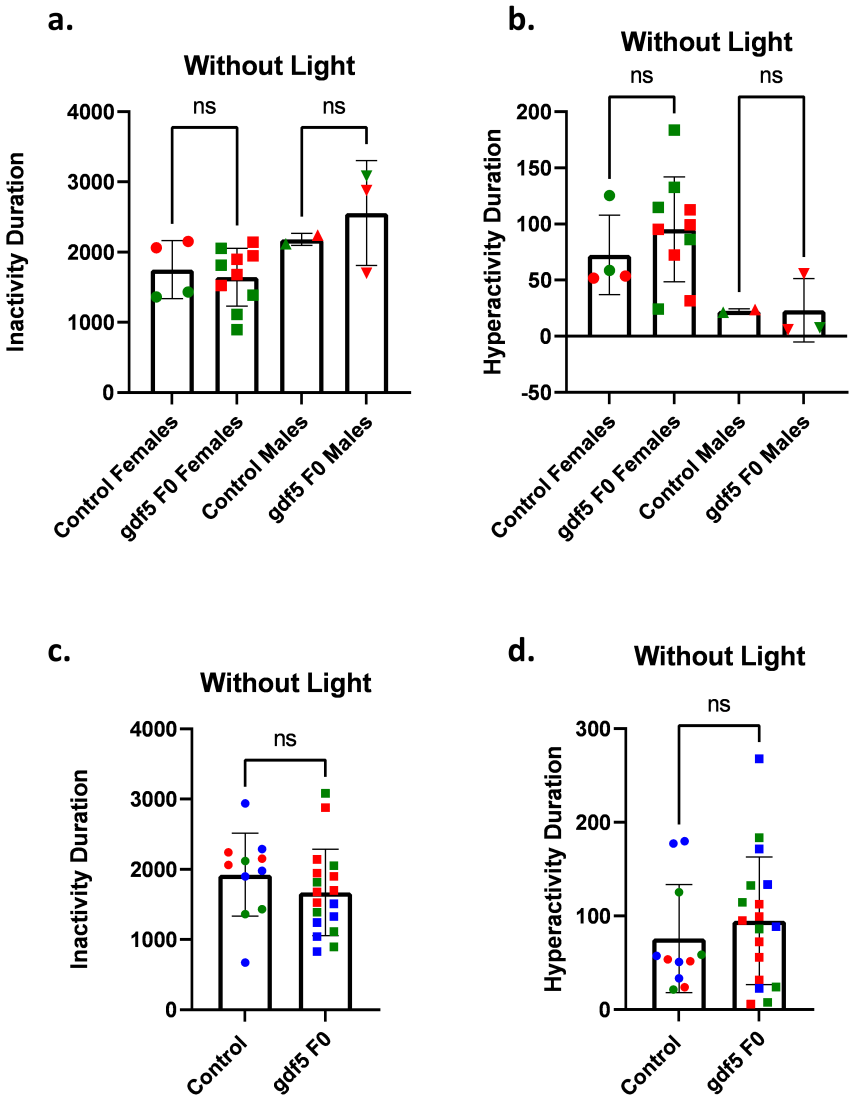


Figure 5.7: Targeting *gdf5* in zebrafish larvae does not affect the fish swimming behaviour in adulthood in the absence of light

(a) No significant difference was observed in the amount of time spent in an inactive state for both sexes, females (*p-value= 0.9095*) and males (*p-value= 0.6248*). **(b)** There was no significant change in hyperactivity duration for both sexes, females (*p-value= 0.5872)* and males (*p-value= >0.9999),* respectively. **(c&d)** no significant change in swimming speed of total adult fish in the absence of light, inactivity duration (*p-value= 0.2861*) and hyperactivity duration (*p-value= 0.4480*). Panel **(a&b)** statistical analysis was conducted as one-way Anova, while panel **(c&d)** was analysed via a student t-test comparing the two columns. Each data point represents the average time spent by a single fish in the stated conditions over the 7-hour experimental period. Red= first replicate (n= 2 uninjected nacre controls and n= 8 *gdf5* crispants), Green= second replicate (n= 3 uninjected nacre controls and 7 *gdf5* crispants), Blue= third replicate (n= 5 uninjected nacre controls and n=5 *gdf5* crispants). Figure generated using Graphpad Prism.

5.3.5. Mutant line generation

To generate a stable mutant line, a traditional approach would involve outcrossing of the founders to wildtype fish. This approach facilitates mutation dispersion and eliminates undesirable off targets that result from CRISPR-Cas9 editing. However, to generate the gdf5 mutant line, we incrossed two of the founders. **Figure 5.8** shows the followed method for gdf5 mutant line generation. Obtained embryos were allowed to grow up to 3dpf and the Zebrafish Embryonic Genotyper (ZEG) tool was utilised to extract DNA of obtained mutants in a non-invasive manner. ZEG relies on harmonic oscillation in the form subtle vibrations of a rough glass plate housing each larva in a separate well containing E3 medium (Lambert et al., 2018). Fallout cells are harvested from the medium, which is then further processed to extract individual larvae DNA without affecting the larvae. This provides a great tool for genotyping individual larvae, while maintaining the exact same larva to be raised. **Table 5.1** shows CRISPR-STAT results of some genotyped gdf5 F1 following ZEG. Samples were also sequenced and a summary of obtained results is presented in **Figure 5.11**. A 6bp deletion was the most commonly observed mutation in 26% of sequenced samples, followed by an 8bp deletion present in 23% of mutants.

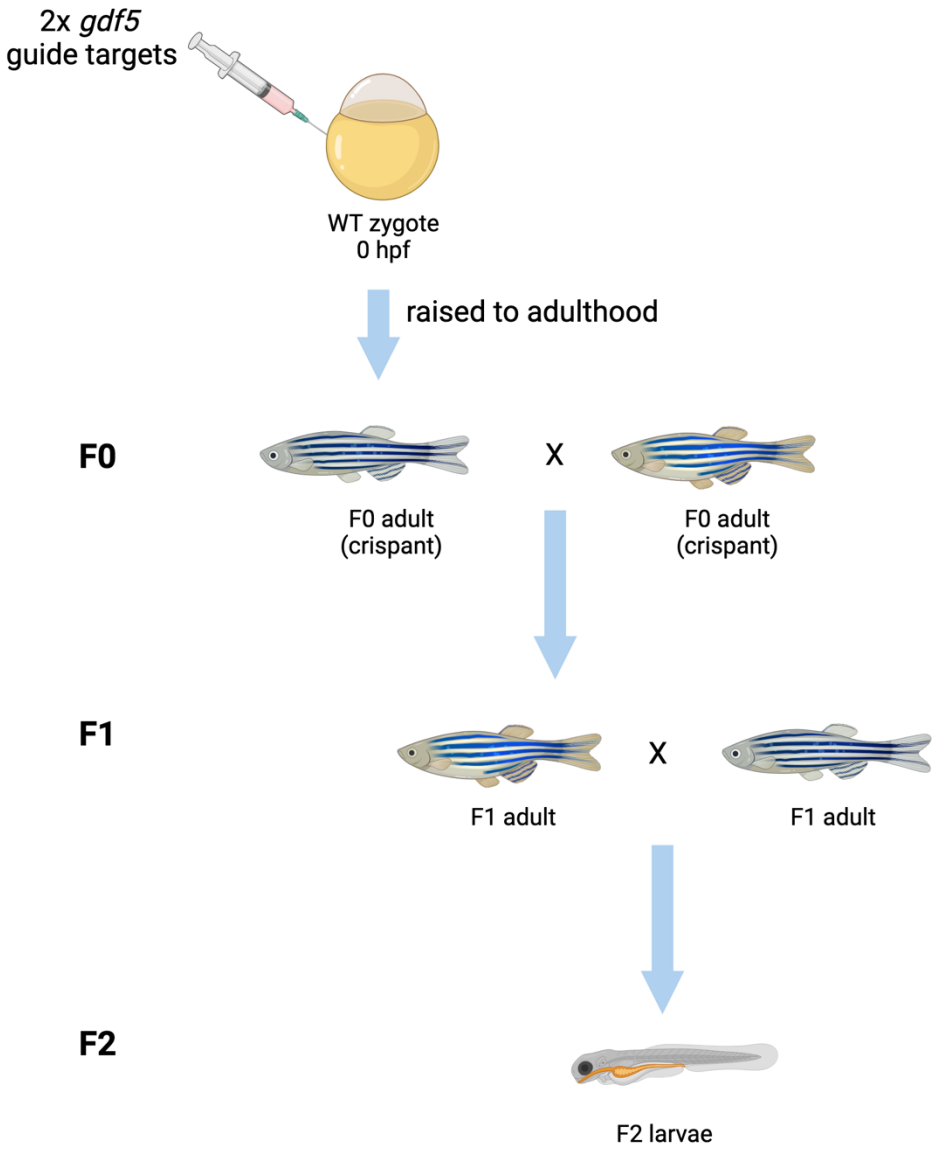


Figure 5.8: gdf5 mutant line generation

The figure illustrates the incrossing approach taken in developing the gdf5 mutant line. The process begins by injecting two single guide RNAs that target *gdf5* in a single cell embryo. Crispants were then raised to adulthood, and two injected crispants were crossed to generate the F1 generation. Two F1 mutants are then crossed to produce the F2 larvae, which were not further raised. Figure generated using Biorender.

Name	Peak
A07_GDF5_1_CTRL1	325
B07_GDF5_1_CTRL2	325
C07_GDF5_1_CTRL3	325
E07_GDF5_1_CTRL5	325
F07_GDF5_1_CTRL6	325
G07_GDF5_1_CTRL7	325
H07_GDF5_1_CTRL8	325
A01_GDF5_ZEG_F1_1A	319
A03_GDF5_ZEG_F1_2A	319
A05_GDF5_ZEG_F1_3A	331
B01_GDF5_ZEG_F1_1B	319
B03_GDF5_ZEG_F1_2B	319
B05_GDF5_ZEG_F1_3B	316
C01_GDF5_ZEG_F1_1C	317
C03_GDF5_ZEG_F1_2C	319
C05_GDF5_ZEG_F1_3C	316
D01_GDF5_ZEG_F1_1D	320
D05_GDF5_ZEG_F1_3D	322
E01_GDF5_ZEG_F1_1E	318
E03_GDF5_ZEG_F1_2E	331
E05_GDF5_ZEG_F1_3E	317
F01_GDF5_ZEG_F1_1F	317
F03_GDF5_ZEG_F1_2F	320
F05_GDF5_ZEG_F1_3F	317
G01_GDF5_ZEG_F1_1G	331
G03_GDF5_ZEG_F1_2G	319
G05_GDF5_ZEG_F1_3G	319
H01_GDF5_ZEG_F1_1H	318
H03_GDF5_ZEG_F1_2H	317
H05_GDF5_ZEG_F1_3H	317

Table 5.2: Alleles Table

The following table summarises the genotyping of incrossed gdf5 F1 larvae, highlighting the detected mutations in comparison to wildtype nacre controls. The peak size of the allele observed in individual larvae indicates a deviation from the expected wildtype peak at 324.5, confirming successful gene targeting. Data analysed and table generated using Geneious Software (https://www.geneious.com/).

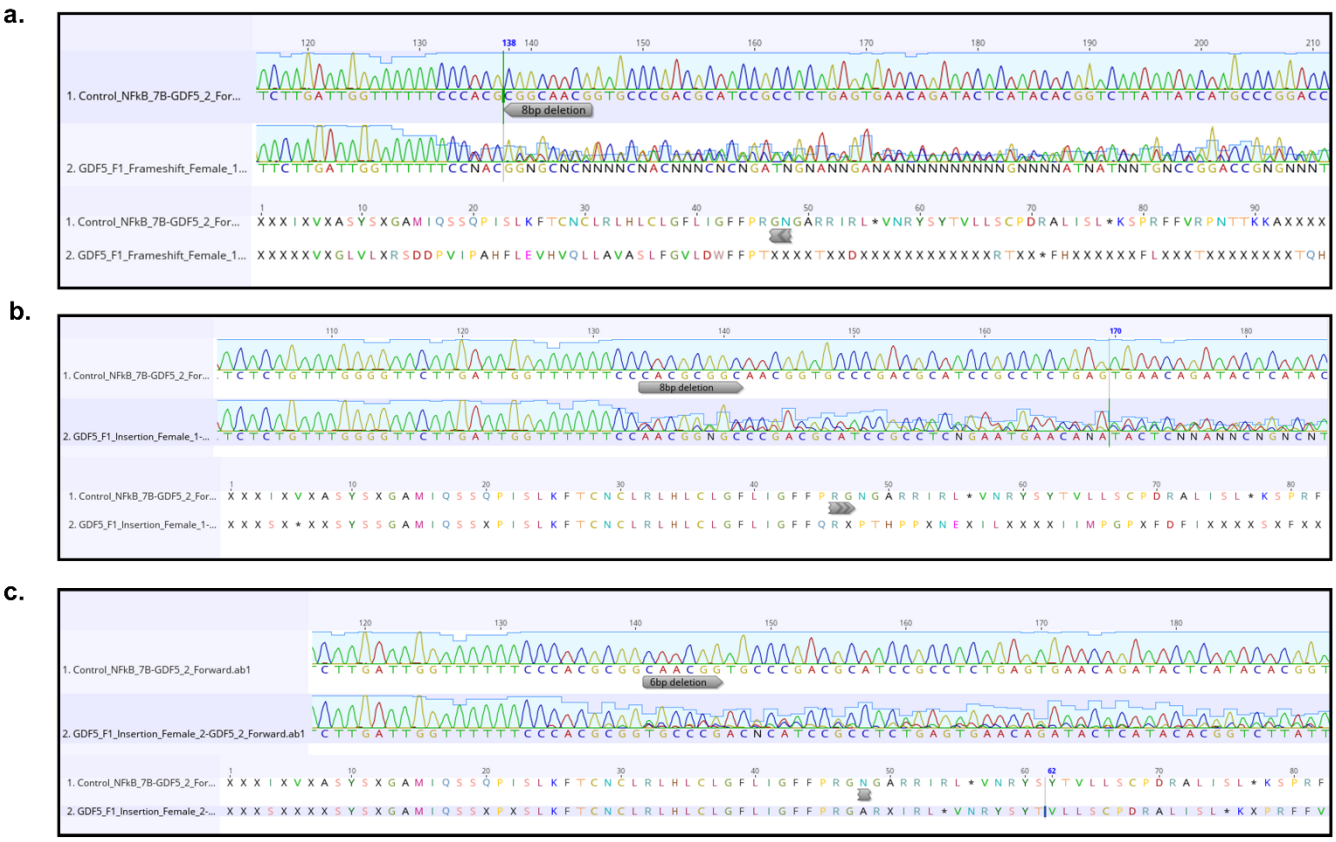


Figure 5.9: Sequencing of gdf5 F1 female adults utilised in gdf5 F1 adult behavioural analysis

The top of each panel shows the traces obtained following the sequencing of each fish, while the bottom section shows the translated sequence to predict the effect of the mutation on the encoded protein. **(a)** - 8bp deletion causing a protein frameshift. **(b)** –8bp causing a frameshift mutation. **(c)** -6bp deletion causing frameshift to encoded protein. Figure generates using the SnapGene software (https://www.snapgene.com/).

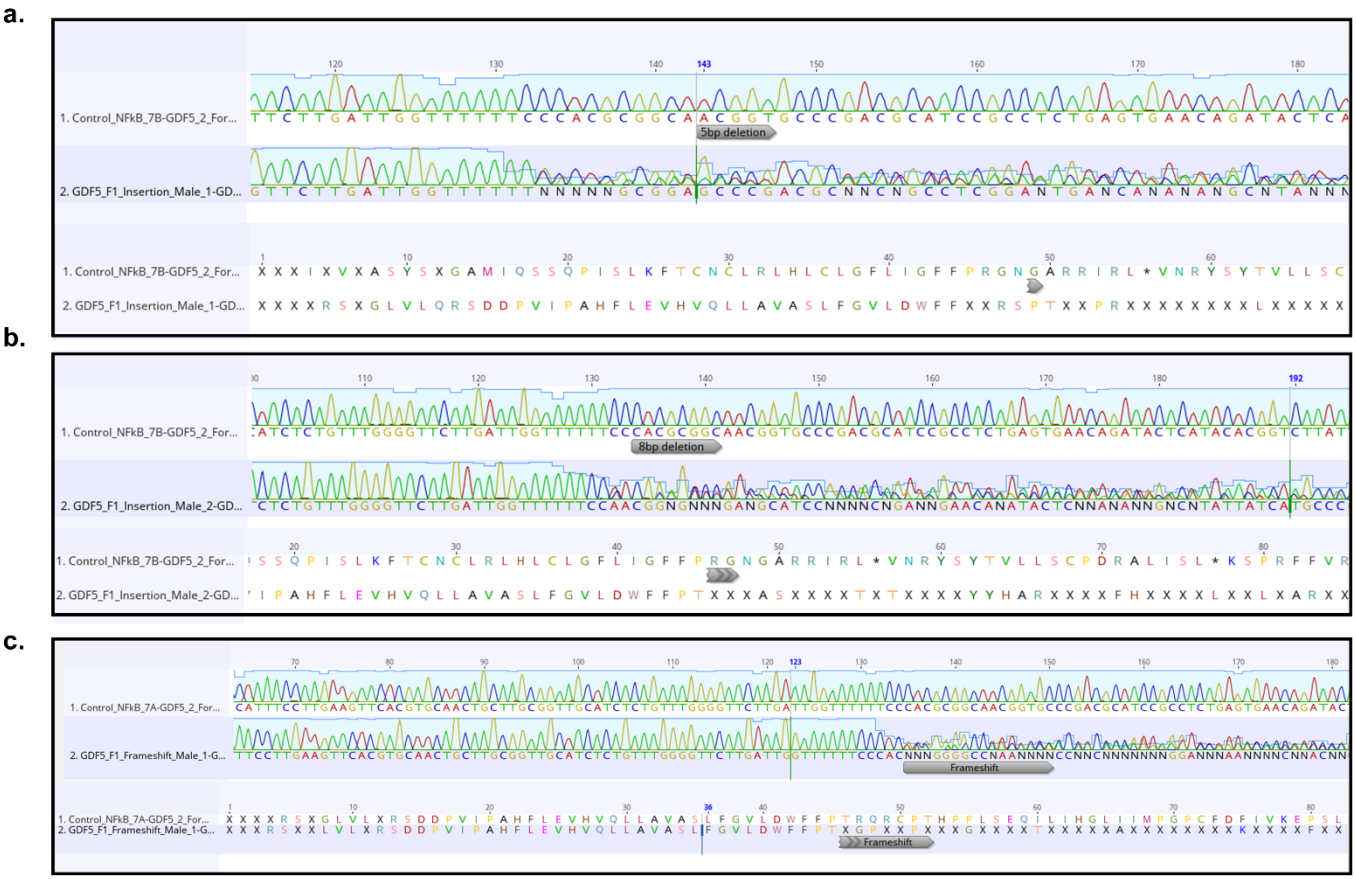


Figure 5.10: Sequencing of gdf5 F1 male adults utilised in gdf5 F1 adult behavioural analysis

The top of each panel shows the traces obtained following the sequencing of each fish, while the bottom section shows the translated sequence to predict the effect of the mutation on the encoded protein. **(a)** - 5bp deletion causing a protein frameshift**. (b)** –8bp causing a frameshift mutation. **(c)** unidentified mutation due to background in sequences traces, but translation to protein sequence suggest a frameshift effect of mutation due to the rise of random amino acids, T = Threonine and G = Glycine, which are not present in the wildtype control. Figure generates using the SnapGene software (https://www.snapgene.com/).

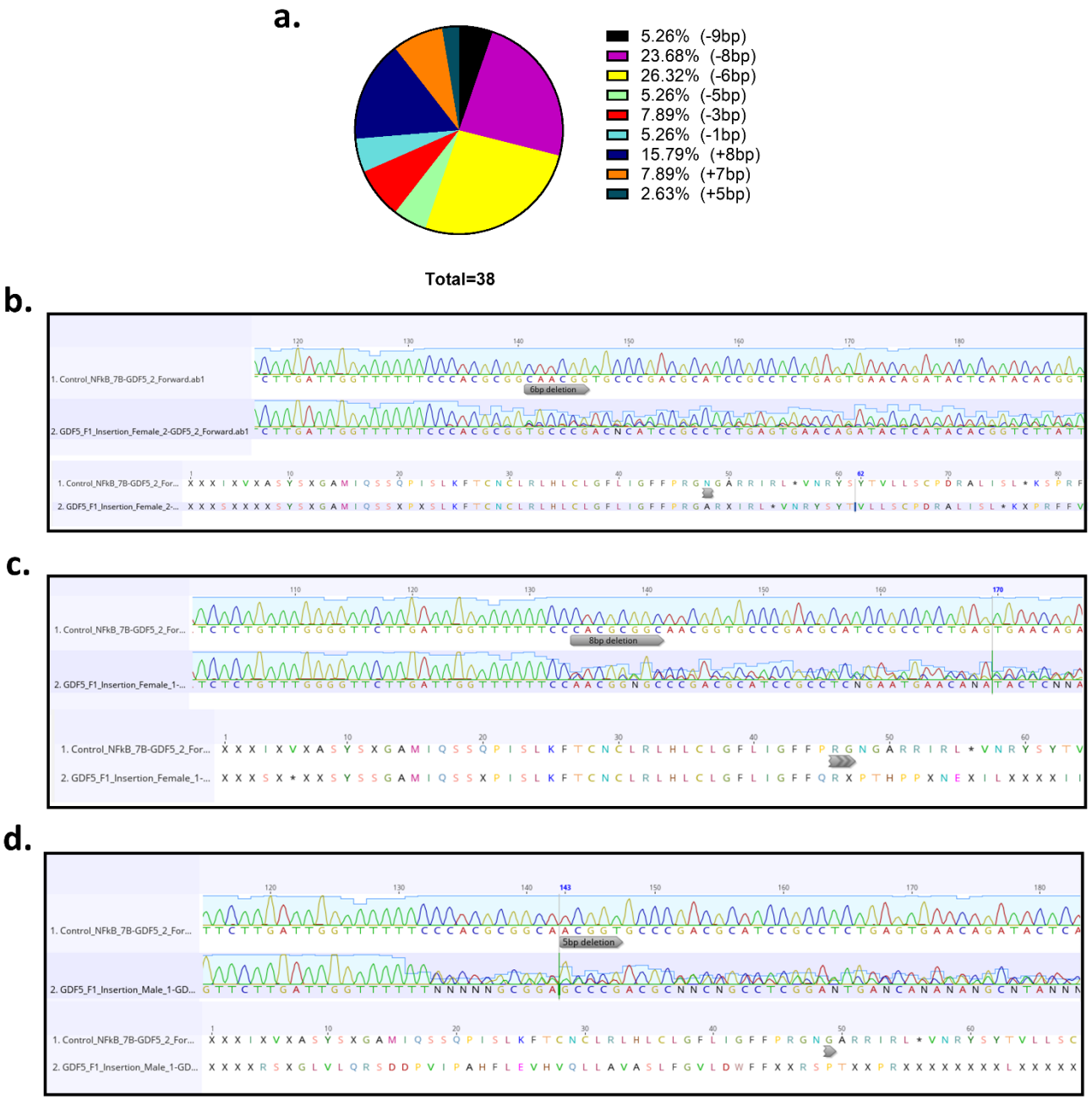


Figure 5.11: Top 3 detected mutations in gdf5 F1

(a) pie chart showing all detected mutations and their frequency. –6bp was the most commonly present mutation in 26.32% of F1, followed by –8bp deletion in 23.68% of sequenced fish, and an 8bp deletion present in 23.68% of in-crossed F1 mutants. **(b)** most commonly present 6bp deletion predicted to result in a missense mutation that caused a minor shift in the encoded protein. **(c)** second most commonly present -8bp deletion leading to a major frameshift significantly changing the predicted protein. **(d)** 5bp deletion found in 2.63% of sequenced mutants and causing a frameshift mutation.

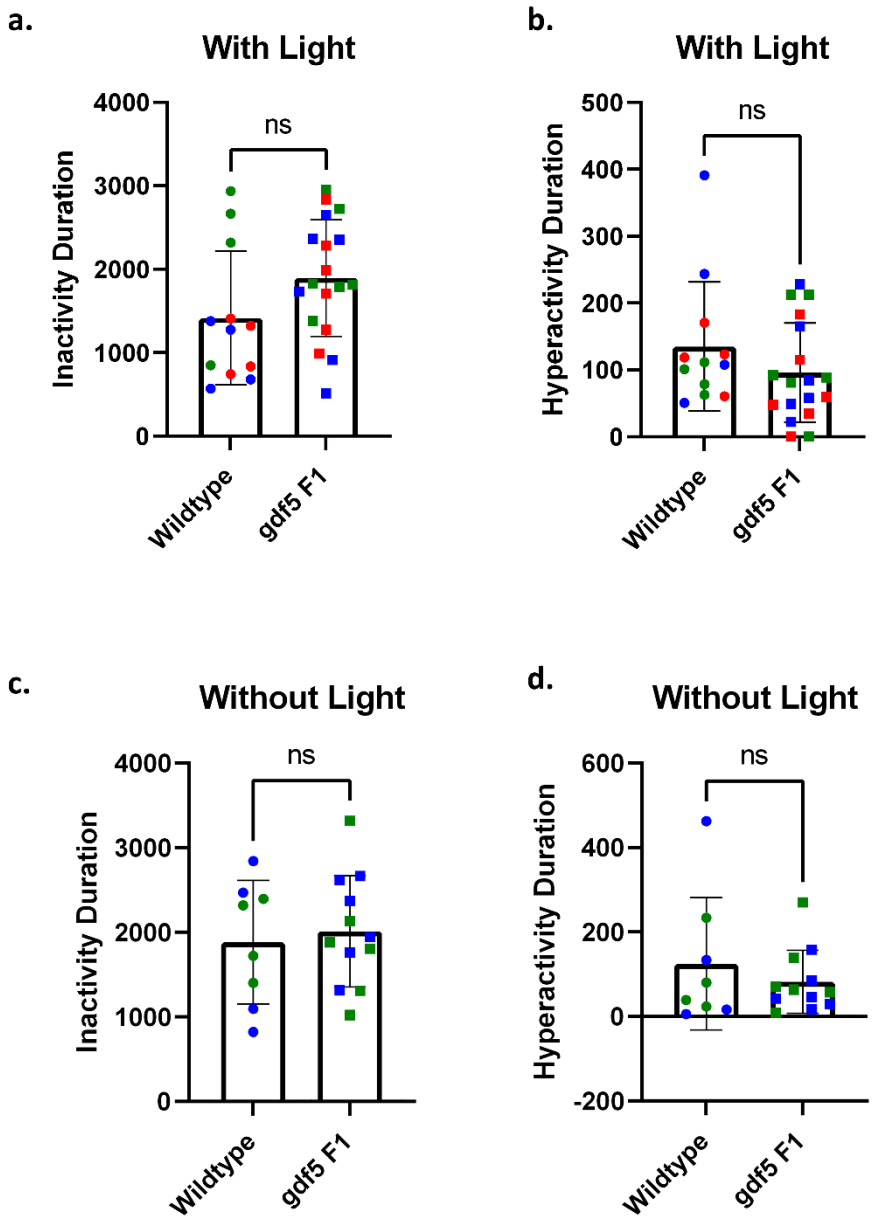


Figure 5.12: No significant change in gdf5 F1 generation behaviour in the presence and absence of light

(a) no significant change in time spent in an inactive state in the presence of light (*p-value= 0.0935).* **(b)** There was no significant change in hyperactivity duration for total fish in the presence of light (*p-value= 0.2247).* **(c)** no significant change in time spent in an inactive state in the absence of light (*p-value= 0.6859)***. (d)** no significant change in time spent in hyperactivity duration in the absence of light (*p-value= 0.4260).* All panels were analysed using a student t-test. Each data point represents the average time spent by a single fish in the stated conditions over the 7-hour experimental period. Red= first replicate, Green= second replicate, Blue= third replicate. Each replicate was performed using (n= 4) uninjected nacre controls and (n=6) gdf5 F1.

5.3.6. Effect of *gdf5* modulation on F2 larval behaviour

To further investigate the possible progression of larval swimming abilities across generations, we proceeded with testing the larval swimming activity and speed in the F2 generation. This experiment was aimed to reveal any phenotype evolution in later generations. As the F2 generation was tested during the early larval development, separating sexes was impossible at this stage. 72 gdf5 F2 mutant larvae and 24 wildtypes nacre were tested in two conditions, with and without light, for a duration of 10 minutes in the Viewpoint Zebrafish Larvae Box. The average time spent at speed categorised movement of each larva was calculated throughout the 10-minute cycle. **Figure 5.13** shows the results for both conditions with each point presenting one tested larva. A slight reduction in the inactivity duration, indicating more general movement, was observed in gdf5 F2 in the presence of light that appear to be non-significant (**a**). However, in the absence of light, gdf5 F2 show a highly significant (*p-value= 0.0008*) increase in activity (**c**) and suggest that mutant larvae appear to move more than wildtype controls. When looking at the duration they spend in a hyperactive state, it does not show an exceptional pattern in the presence or absence of light (**b&d**). As this experiment was only conducted once in each condition, further repetitions are necessary to corroborate the mentioned observation. However, the high significance of the alteration suggests a possible impact of *gdf5* targeting that disables larval movement in the F2 generation.

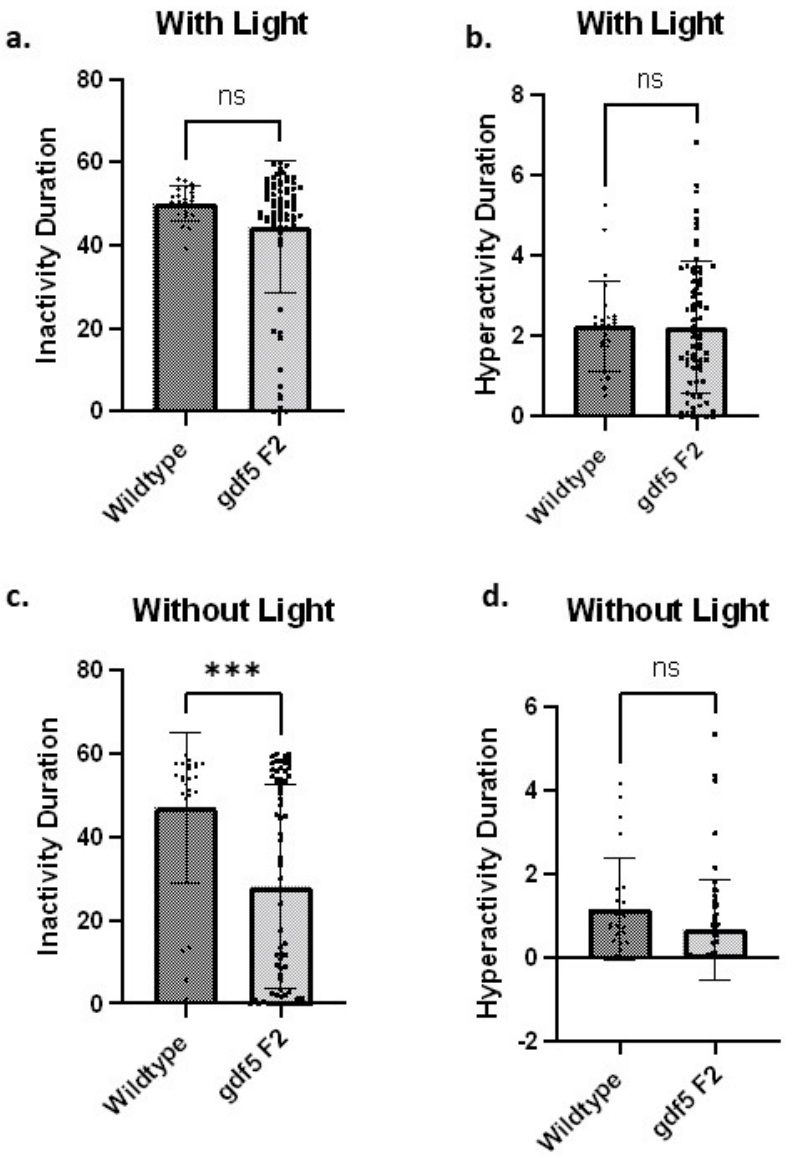


Figure 5.13: Significant increase in activity levels of gdf5 F2 larvae at 5dpf in the absence of light

(a) gdf5 F2 (n=72) spend a similar amount of time in an inactive state as the wildtype nacre controls (n=24) in the presence of light (*p-value= 0.0907*). They also do not display a change in their rapid movement duration as shown in **(b)** (*p-value= 0.9073*). **(c)** when compared to wildtype controls that appear to be more sedentary, F2 mutants show a significant reduction in inactivity duration indicting a significant increase in movement only in the absence of light (*p-value= 0.0008)*. However, larvae do not show a significant effect on rapid hyperactive movement without light (*p-value= 0.0884)*. Experiment was conducted once, and each point represents average time spent by a single larva in the stated category (inactivity or hyperactivity). A student t-test was used to determine statistical significance. Figure generated using Graphpad Prism.

5.3.7. Consistent pattern of increase in regenerative ability of gdf5 mutants that does not show significance when compared to wildtype controls

Although the function of *GDF5* in bone modulation, chondrogenesis and ECM regulation was confirmed in several studies (Hotten et al., 1996; Francis-West et al., 1999; Garciadiego-Cazares et al., 2015), the role of *GDF5* in tissue regeneration has not been investigated yet. From murine cartilage injury microarray data obtained in our laboratory, *Gdf5* has shown a remarkable increase in its expression levels post injury (**Figure 3.6**), which could be a result of the induced inflammation, or it could suggest a possible role in tissue repair. Therefore, we were keen on testing the effect of *gdf5* mutation on tissue regeneration. Using generated F2 mutants, we induced tailfin injury and measured the notochordal bud regeneration 24 hours post-injury to identify any significant change in regeneration speed and efficiency. Although a trend of increase in notochord regeneration was observed in two out of three separates technical replicates, the difference did not appear to be statistically significant (*p-value= 0.622*) **(Figure 5.14b)**. Looking at the control groups, in the final repeat, wildtype larvae appear to display an abnormal high levels of regeneration post injury, when compared to the two other replicates. This could suggest an outlier effect to the final repeat.

To confirm the efficient transmission of the mutation, and the downregulation of *gdf5* expression in the tested F2 generation, a quantitative Real-Time PCR was conducted. Samples from 5dpf larvae were tested in three technical replicates and all consistently show significantly lower levels of *gdf5* in the F2 mutants when compared to controls (*p-value= 0.0123*), **Figure 5.15**. Taken together, qPCR data confirmed the reduction in *gdf5* expression, combined with the observed non-significant increase in reparative abilities of *gdf5*, could point out a potential role of *gdf5* in limiting repair as an inhibitor to differentiation or as a pro-apoptotic factor regulated by other factors. Further research must be undertaken to confirm this suggestion and attempt to reveal hidden mechanisms of action of *gdf5*.

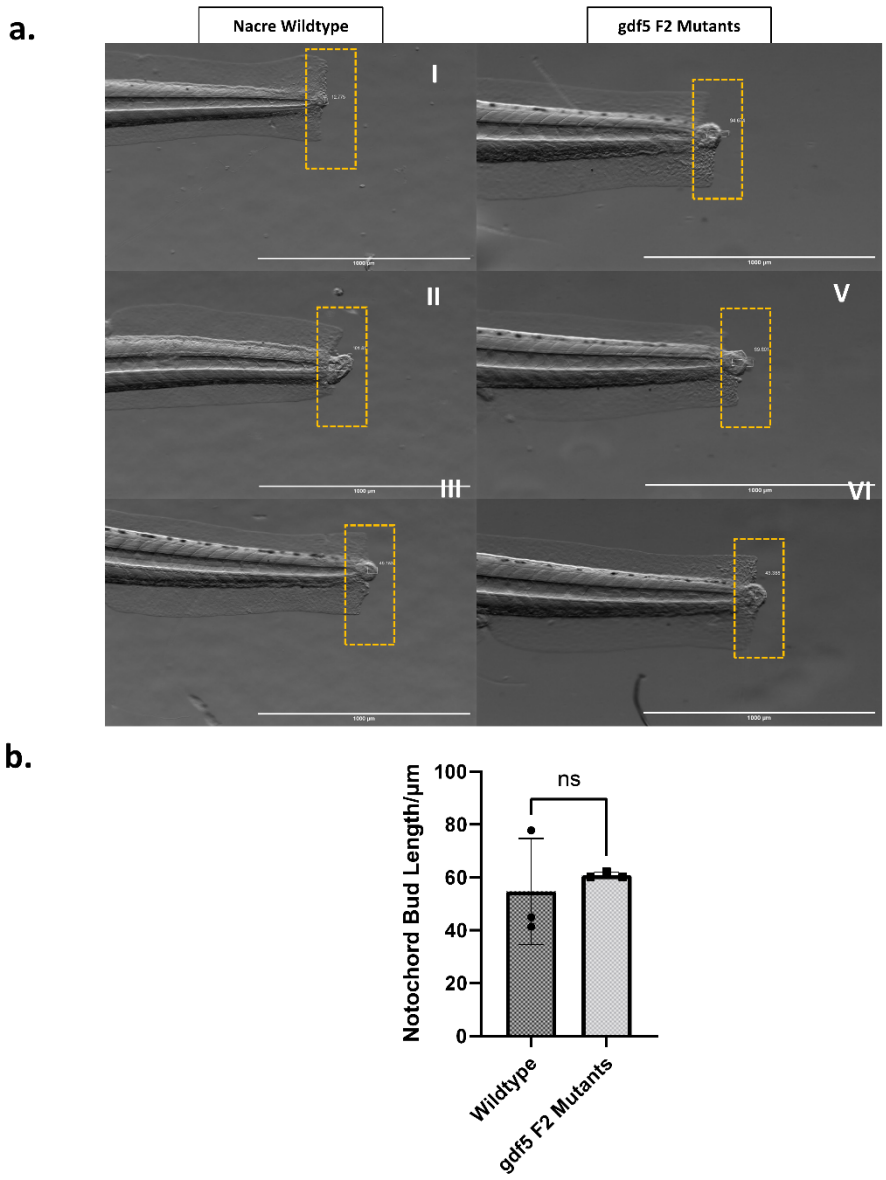


Figure 5.14: gdf5 F2 mutants display normal regenerative ability of notochord bud length following a tailfin injury that is mostly higher than wildtype nacre controls with no significance

(a) representative images of tailfin regeneration model 24 hours following injury induction to measure the regenerated notochord from the centre of cut to the tip of notochord in control wildtype larvae (I-III) and gdf5 F2 mutants (IV-VI). Yellow dotted rectangle highlights region of interest. **(b)** Analysis of notochord regenerated bud length obtained from 3 different replicates (replicate 1= 10 WT controls and 27 gdf5 F2, replicate 2= 16 controls and 27 gdf5 F2, replicate 3= 25 controls and 25 gdf5 F2) showing slightly higher regenerative ability in gdf5 F2 mutants when compared to wildtype controls, yet it appears non-significant *(p-value= 0.622).* Each point represents average growth length from all larvae in a single experiment. Statistical analysis was conducted using a student t-test. Scale bar= 1000µm.

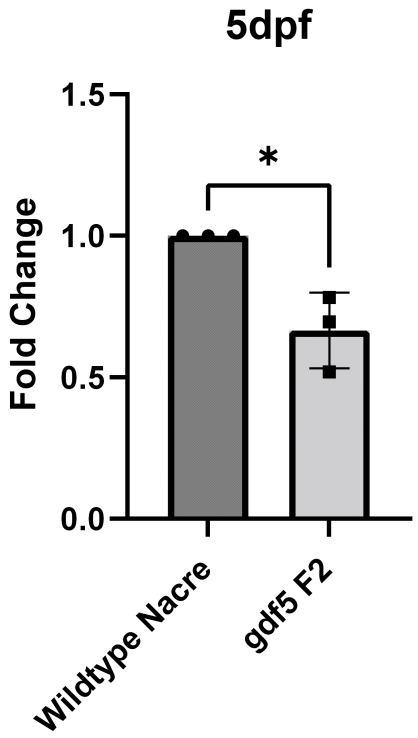


Figure 5.15: qRT-PCR data confirming the reduction of *gdf5* expression levels in the F2 generation

3 experimental replicates of quantitative Real-Time PCR reactions in wildtype nacre fish versus gdf5 F2 mutant larvae samples obtained from 5dpf fish. Data show significant reduction in *gdf5* gene expression levels (*p-value= 0.0123*) in mutants when compared to WT controls highlighting successful mutagenesis and mutation transmission through generations. Data analysed using a non-parametric student t-test.

5.3.8. Significant increase in bone mineral density and bone surface area in gdf5 founders compared to wildtype controls

Using Micro-CT scanning, the 5 female founders that were previously used in adult behavioural analysis along with 3 of wildtype nacre controls were scanned at the age of 33 months, which is considered an aged fish, to investigate any changes in mature skull bone structure. Representative images are shown in **Figure 5.16**. Our results showed non-significant change to the bone volume in *gdf5* F0 compared to wildtype controls (**Figure 5.17a)** (*p-value=0.99*). However, a significant increase in bone mineral density (**Figure 5.17b**) (*p-value= 0.01*) and bone surface area (**Figure 5.17c**) (*p-value= 0.0139*) were found in the adult crispants. This suggests that *gdf5* has a role in the regulation of bone mineralisation and is involved in later stages of skeletogenesis at 33 months.

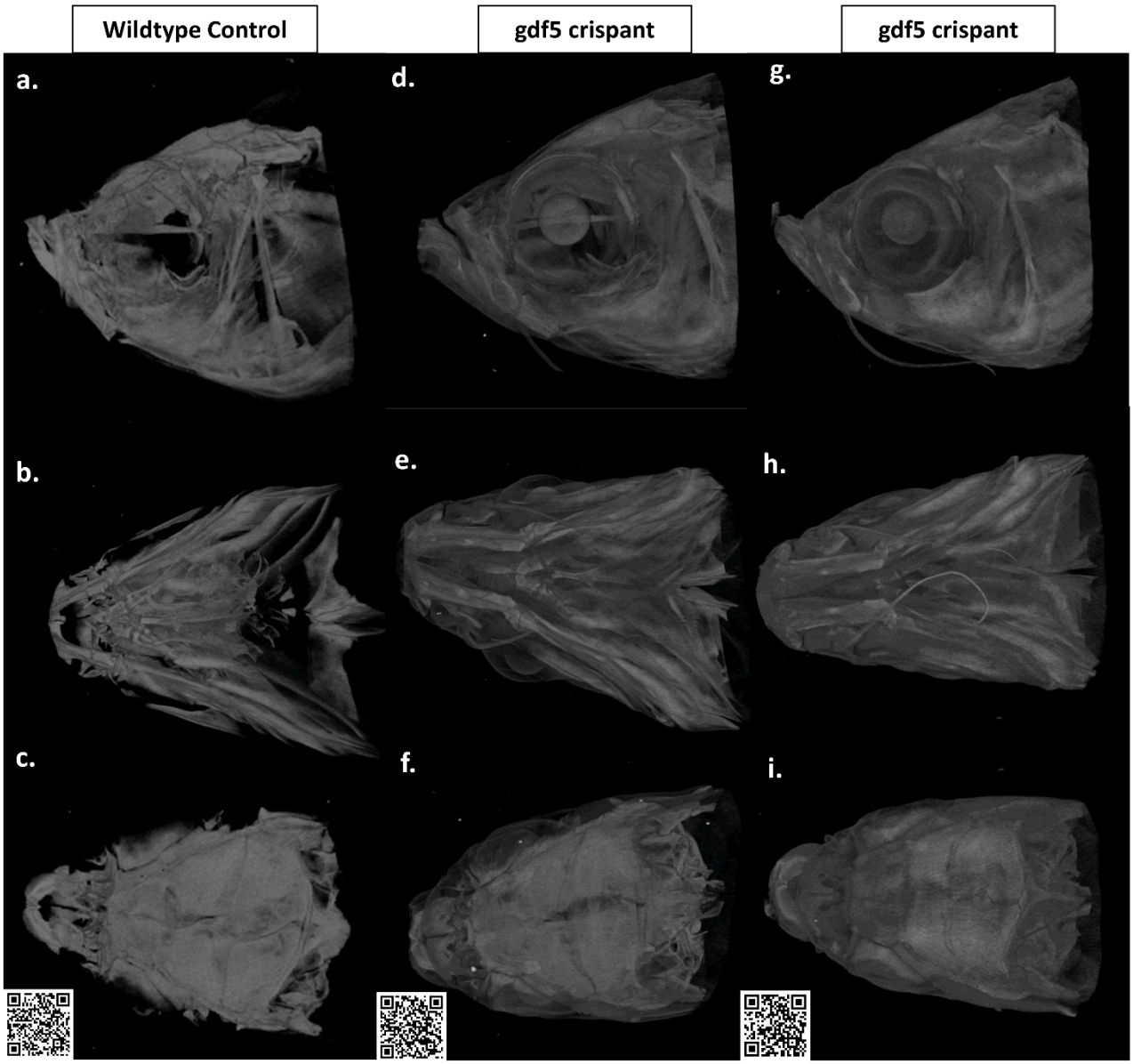


Figure 5.16: Micro-CT scans of *gdf5* crispants

Micro-CT scans displaying the bone structure of generated adult *gdf5* crispants **(d-i)** versus wildtype nacre control **(a-c)** at 33 months. First row shows the ventral view **(a,d&g),** second row is the lateral view **(b,e&h)** and third row is the dorsal view of zebrafish skull **(c,f&i).** Please scan QR code (bottom left) to see video of whole 3D construct.

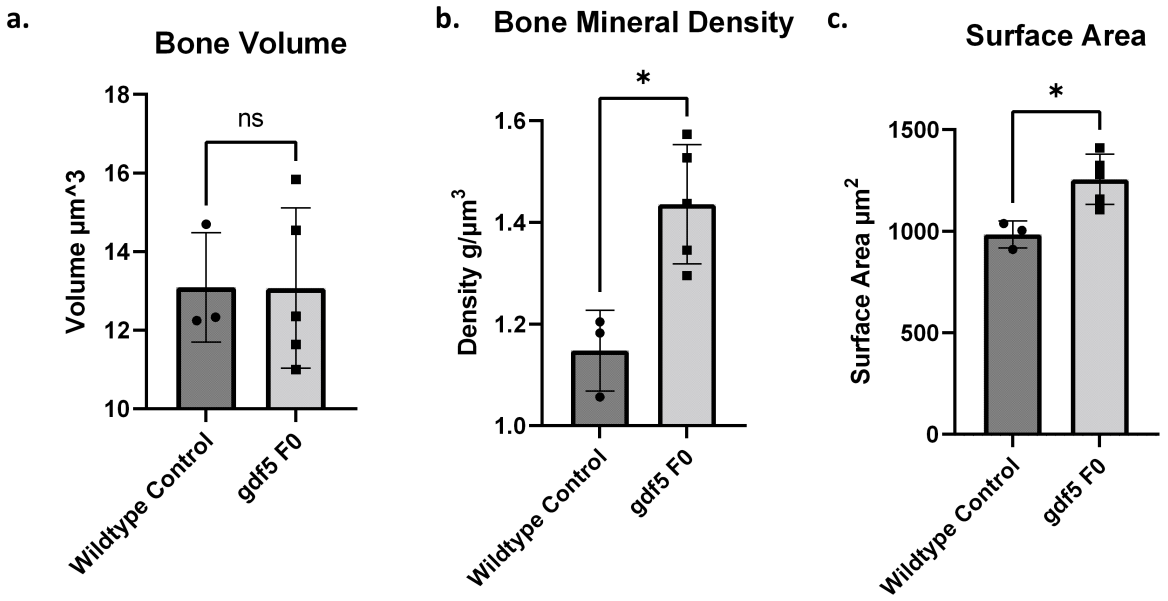


Figure 5.17: Significant increase in bone mineral density and bone surface area of *gdf5* adult crispants

Results obtained from micro-CT scans of 3 nacre female controls compared to 5 *gdf5* female crispants at the age of 33 months. **a.** unaffected general bone volume of mutants when compared to wildtype controls (*p-value= 0.99*). **b.** mutants appear to have a significantly higher bone mineral density (*p-value= 0.01*) and surface area (*p-value= 0.0139*) when compared to controls as shown in panel **c**. Each point represents a single fish, and a student t-test was conducted for statistical analysis.

5.4. Discussion

GDF5 has been identified and studied for its critical function in the musculoskeletal system since the 1990s (Storm and Kingsley, 1996). Our knowledge regarding the heavy involvement of the gene in bone formation and remodelling has been expanding. The basic signalling mechanism of the gene through the BMP signalling pathway has been identified (Baur and Dymecki, 2000), yet the exact mechanism of action of the encoded protein and its regulation through other factors remains ambiguous. Furthermore, several key interactions between GDF5 and other GDF proteins, and BMPs are still to be determined. This study aimed to further investigate the role of GDF5 by studying the effect of its targeting on zebrafish skeletal and heart development, as well as its function in the maintenance of the skeletal system at later stages. Obtained data showed that *gdf5* is not required for early skeletogenesis in the form of cartilage development in zebrafish larvae up to 5dpf. The same observation was recently confirmed in a 2021 study in zebrafish, published during the conduction of these experiments, which also found null effect on early skeletogenesis in *gdf5* homozygous mutants targeting the same exon. However, they revealed impairments in the pectoral and median fin endoskeleton development where both structures appear smaller in size at 90dpf. The total fish body length does not seem to be affected by the mutation (Waldmann et al., 2022). Our study did not incorporate this developmental stage as a consequence of the COVID-19 pandemic lockdown, which hindered our access to the aquarium, and consequently, the fish progressed beyond this stage. However, we also observed a progressive effect of the *gdf5* targeting in older adults at the age of 33 months, which has not been previously reported. The *gdf5* crispants showed a significant increase in bone mineral density *(p-value*= 0.01) and bone surface area *(p-value= 0.0139)* when compared to wildtype controls without significantly impacting the total bone volume *(p-value= 0.99).*

Although a significant change in adult fish bone structure was observed, it had no impact on the movement of the crispants in both tested light conditions in the adulthood stage indicating that observed bone changes are mild to contribute to a significant change in movement speed and duration. However, it must be noted that the behavioural analysis investigates the effect of the mutation on whole-body movement, when only the effect of the mutation on skull bone was examined using micro-CT scans. Another noteworthy observation is the death of all 3 male founders before reaching 29 months, while the 5 female founders survived till the experiment was terminated at 33 months, which may suggest a differential effect of mutations on both sexes, yet further work is required to confirm this observation and understand its cause. Currently, no studies confirmed the differential role of *GDF5* within sexes. However, OA is found more commonly in women compared to men (Tschon et al., 2021). Other musculoskeletal disorders such as rheumatoid arthritis (Gerosa et al., 2008) and osteoporosis (Alswat, 2017) are also more commonly present in women with rheumatoid arthritis showing over two-fold

frequency in women when compared to men, and osteoporosis being present four times more commonly in females. Several studies suggested the cause of this sex bias is due to hormonal differences. However, our data could suggest a potential difference in functional role of *gdf5* between the sexes, which imply a critical role for the gene in survival of males, but not as critical in females.

In gdf5 F2 larvae, mutants showed a high significant increase in general activity only in the absence of light *(p-value= 0.0008),* and the change was not significant in the presence of light *(p-value= 0.0907).* This highlights a possible cumulative effect of mutation that appears to affect the larval movement in later generations. As the change was only observed in the absence of light conditions, while the movement did not show any significant change in light, this provides a doorway on the role of *gdf5* in the nervous system. A number of studies have explored this function of *gdf5* and found it to be involved in the induction of dopaminergic neurons (Sullivan and O'Keefee, 2005) revealing a link to Parkinson's disease, and the establishment of sympathetic response (O'Keeffee et al., 2016). GDF5 presents significant promise as a target for the treatment of various neurological disorders, and potentially in the cases of co-morbidity with skeletal and neurological afflictions.

We took the lead in investigating the effect of targeting *gdf5* on zebrafish heart development and found it not to be involved in this process as *gdf5* crispants heart seem to develop normally, undergoing the essential looping process and having a normal heart volume size when compared to uninjected wildtype controls. This suggests that *gdf5* is not required for early heart development. However, it does not exclude the possibility of having a role in later heart function or development, which unfortunately was not investigated in this study due to the lack of a visualisation method of adult heart structure.

As a known building block involved in ECM osteogenesis regulation (Garciadigeo-Cazares et al., 2015) and due to the available data showing a significant increase in *Gdf5* expression post-injury **Figure 3.6**, we investigated the possibility of its involvement in tissue repair and regeneration post-injury. We found a slight trend to an increase in the regenerative abilities in gdf5 F2 when compared to wildtype controls that does not show a significant difference *(p-value= 0.622).* Even though the increase does not show significance, the consistency through two of the three replicates was intriguing. Moreover, wildtype controls in the final replicate were distinctly different when compared to controls of the other two replicates, which could indicate an outlier effect, so further replicates are required to reveal the real effect of *gdf5* targeting on the larval regenerative abilities. To confirm the success of mutation transmission through generations and examine the expression levels of *gdf5* in the F2 generation, three replicates of a qPCR reaction was performed. The qPCR analysis showed a significant reduction in *gdf5* expression in the F2 generation *(p-value= 0.0123).* This indicates successful mutation transmission which has indeed persevered through generations and it also shows that the observation of gdf5 mutants' regenerative abilities is worth further investigation. It also ties to previously discussed micro-CT data

that presents a new possible role of Gdf5 as a catabolic factor in bone and ECM homeostasis, which when targeted show an increase in tissue formation.

Previous studies have revealed the role of GDF5 in osteoblast differentiation and osteoclast inhibition (Francis-West et al., 1999; Takahara et al., 2004) indicating its functional role as an anabolic factor in bone homeostasis. However, our data suggests a complementary role of Gdf5 as a catabolic factor in bone homeostasis, concluded when its loss in generated crispants resulted in an increase in bone mineral density and surface area, as well as a potential increase in regenerative abilities and significantly enhanced swimming activity of the F2 generation. These observations unravel the binary effect proposed by the gene function regulated by other factors, which is not surprising considering the complex dual effect observed in several bone homeostatic genes including ones studied in this project. A recent study revealed a glimpse of this novel mechanism of action though the interaction of GDF5 with the long non-coding RNA, miR-525-5p. He and Lin., 2023, found LINC00313, a non-coding RNA sequence involved in gene expression regulation in various processes such as cell proliferation, reduction in pro-inflammatory cytokines and reduction in apoptosis, to be downregulated in OA patients. miR-525-5p, was found to be overexpressed in OA, which can bind to LINC00313, inhibiting its proliferative and anti-apoptotic effect, resulting in OA progression. miR-525-5p forms a complex with GDF5 regulating this process and when GDF5 is inhibited, it inhibits miR-525-5p, which consecutively release the inhibition of LINC00313, leading to an increase in cell proliferation and differentiation, which complies with our observation in the *gdf5* crispants micro-CT data. Another possible explanation for the increase in bone mineral density and surface area in *gdf5* crispants could be due to an increase in other bone homeostatic gene expression levels that gets overexpressed to counteract the reduction of osteoblast differentiation and osteoclast inhibition in crispants. However, this proposes that bone homeostasis is a highly fine-tuned process that can recalibrate itself even in the reduction or complete absence of one of its elements. In this case we would expect bone measurements to not be affected by the mutation and exhibit similar measurements to the control group, yet our data rather shows a significant increase, which could rule out this hypothesis. Finally, the induced mutation could have caused a resistance to Noggin inhibitory effect on *GDF5* as seen in synostoses syndrome (Seemann et al., 2009; Schwaerzer et al., 2012), resulting in an increase in osteoblast differentiation rather than a reduction. This could provide a plausible conclusion to the observed generally mild phenotype, which was unexpected according to our previous knowledge about the gene role in skeletal development in other organisms. Therefore, further experiments must be conducted to understand the hidden GDF5 mechanisms of action. A simple experiment could be designed to detect the expression levels of miR-525-5p, LINC00313 and Noggin in the generated mutants, to confirm or deny our hypothesis.

Unlike the traditional method of outcrossing crispants to obtain the F1 generation then incrossing the F1 fish for the generation of a stable F2 mutants, we incrossed the founders and obtained an incrossed F1, as illustrated in **Figure 5.8**. The main aim of the traditional crossing route is to reduce the off-targeting effect occasionally observed in CRISPR targeting to ensure that observed phenotypes are exclusively due to the intentionally induced mutation. As we did not observe significant phenotypic changes in our gdf5 mutants, as observed by colleagues in Sweden who took the traditional crossing approach (Waldmann et al., 2022), it does not propose as much of an issue in this case. Interestingly, the incrossing has potentially opened a novel lens of understanding the possible downstream effects of inducing a mutation within the same gene and crossing of two crispants. Even though crispants are mosaic in nature and they carry different forms of deletions, the gametes transferred to form the F1 generation will have a maximum of two mutations from the two available alleles. Taken together, F1 fish can have different combinations of mutations, yet they are limited to a maximum of four available mutant alleles passed down from the founders' gametes. When sequencing the incrossed F1 fish, we found random mutations within *gdf5* that are not part of the expected four allele mutations. This could highlight potential evolution of mutations induced by CRISPR when both founders carry mutations within the same gene. Considering the known fact that mutations can devolve and disappear through generations (Loewe and Hill., 2010; Kinnersley et al., 2021), it is plausible that the opposite would occur allowing mutations to evolve and generate new forms through generations. Although this is a mere suggestion from limited data on a single gene, I believe it should call attention to the possible disastrous effects of genome editing in humans, which have been pushed by many scientists in the field without cautiously considering the irreversible effects of this fast-growing technology on humans.

Several limitations must be taken into consideration in this study. Firstly, that obtained images of cartilage and heart structures were only produced in one experiment. Although the experiment was repeated 3 additional times, the death rate and observed defects within the control group challenged us into including them as reliable replicates. Therefore, only one repetition has been performed, which is inadequate to arrive at a conclusive result. When generating a mutant line, an incrossing approach was followed instead of the traditional method of outcrossing founders aiming to eliminate any off-target effect of the CRISPR injection. Also, generated F2 were obtained from a pool of 5 gdf5 incrossed F1, which again does not follow the traditional route of mutant line generation and may carry off-target mutations. The death of male founders prevented us from conducting further analysis on adult male founders, which appear to be more affected by the mutation than females and could possibly reveal a different aspect of Gdf5 function. Furthermore, the larval behaviour analysis showed a high significant reduction in time spent in an inactive duration, yet this experiment was only conducted once and require further replicates to confirm observation. Regarding the regeneration model, repeats of the experiment might reveal a hidden increase in the regenerative abilities of gdf5 mutants. Two out of the three

replicates showed a consistent increase in the regenerative abilities of mutant larvae post tailfin injury, yet in one replicate the control group showed a higher regenerative rate when compared to the two other replicates affecting the overall result. Finally, for adult micro-CT scans, we attempted to stain the soft tissue of the fish using Phosphotungstic acid (PTA) in the aim of visualising the heart to detect changes in adulthood. Unfortunately, PTA failed to stain the heart, yet it stained other soft tissue within the organism. Nacre controls utilised in this analysis were scanned on a different day and were not stained with PTA due to failure of staining the heart. Therefore, control group only shows the bone structure, while gdf5 founders also show some soft tissue such as the eye clearly stained. This does not have an impact on any of the conducted measurements, but it does look different in the constructed 3D models.

In summary, the data obtained revealed that *gdf5* has a varying effect on the development of the musculoskeletal system in zebrafish, which was not as detrimental as observed in mouse models. Nonetheless, abnormalities emerge during later developmental stages and become evident in the adult micro-CT scans, indicating a significant increase in bone mineral density and bone surface area. The varying effects of *GDF5* across different species must be taken into account when studying it in a model organism. Targeting *gdf5* did not result in notable changes in heart morphology or total volume, implying that the gene is dispensable in early heart development in zebrafish. The premature death observed in all adult male crispants suggests a potential role of *gdf5* in survival, with differential effects based on sex. Further investigation is essential for future drug targets involving GDF5 to determine their effect on both sexes of humans.

Additionally, there were two main reasons that the significant increase in larval swimming activity, which was only observed in the absence of light in the F2 generation, was intriguing. Firstly, it is important to consider the cumulative effect of gene modulation across generations, which allows the specific phenotype to be present only in the F2 generation, when there was no evidence of swimming alterations in the F0 and F1 generations. Secondly, the distinctive reaction demonstrating only significance in the absence of light underscores the potential function of the gene in the regulation of the nervous system, rendering GDF5 an excellent potential drug target for neurological disorders, as well as co-morbidity cases with osteoarthritis and parkinson's. The regenerative trend observed was also intriguing. Repetitions of the regenerative model experiment may allow us to determine whether the increase in *gdf5* post injury, observed in murine hip samples. is related to inflammation or tissue repair, or both. This may also help resolve the discrepancies in the expression levels of *GDF5* in OA patients, where different studies have reached varying conclusions. Finally, the increase in bone mineral density and bone surface area highlighted a role of Gdf5 as a catabolic factor, which when targeted cause an increase in bone tissue.

6. Identifying the Functional Role of *tgfβ1a* in Cartilage and Heart Structural Development

6.1. Introduction

Transforming Growth Factor Beta 1 (TGFβ1), also known as CED and DPD1, is a cytokine belonging to the TGFβ superfamily. TGFβ superfamily is a group of over 30 cytokines functioning simultaneously to regulate numerous cellular processes including proliferation, differentiation, apoptosis and inflammation (Derynck and Zhang, 2003). Members of the family include three isoforms of TGFβ, activins, inhibins, Bone Morphogenetic proteins (BMPs) and Growth Differentiation Factors (Massague, 1998), including GDF5 which was studied and discussed in the previous chapter. *TGFβ1* is one the most abundantly expressed genes involved in early developmental stages, as its expression is established by 10 weeks post-conception in humans (Graham et al., 1992). The elevated TGFβ1 expression levels persist into adulthood, rendering it a vital gene that is extensively implicated throughout most of the human lifespan. Several diseases have been associated with defective TGFβ1 signalling to name a few, Camurato-Engelmann disease (Kinoshita A et al., 2000; Janssens et al., 2000), Marfan syndrome (Matt et al., 2009), hereditary haemorrhagic telangiectasia (HHT) (Sadick et al., 2005) and inflammatory bowel disorder (IBD) (S Ihara et al., 2017; Kotlarz et al., 2018). *TGFβ1* gene is known for its role in ECM deposition by stimulating the production of various matrix components such as collagen (Chan et al., 2005), proteoglycans (Okuda et al., 1990; Humes et al., 1993) and fibronectin (Taipale et al., 1994). Mutations within *TGFβ1* can lead to various fibrotic diseases including lung (Feernandez et al., 2012), liver (Friedman et al., 2000) and kidney fibrosis (Wu et al., 2013). Similarly, it can be a contributing factor to several types of cancer including breast (Xie et al., 2002; Kang et al., 2005), pancreatic (Melzer et al., 2017) and colorectal cancer (Xu et al., 2007; Wodzinski et al., 2022). This is because the encoded protein plays a significant role in tissue growth (Simon et al., 1995), cell cycle regulation (Hannon et al., 1994; Saltis., 1996), remodelling (Bonewald et al., 1990) and apoptosis (Ramesh et al., 2008). Therefore, any abnormalities within the gene or encoded protein can have a notable impact on these functions resulting in excessive proliferation, escape of apoptosis, suppression of the immune response and activation of angiogenesis to support cancer growth and progression.

6.1.1. Cellular function and molecular mechanisms of TGFβ1

6.1.1.1. TGFβ isoforms

There are three different isoforms of TGF-β, TGFβ1, TGFβ2 and TGFβ3. The three isoforms are widely expressed in mammalian tissues, and they function through two categories of pathways, SMAD-dependent and SMAD-independent. TGFβ2, which has 71% protein sequence similarity to TGFβ1 (Ten dijke et al., 1988) was found to be involved within the nervous system by regulating neuron development (Flanders et al., 1991). Meanwhile, TGFβ3 has a more similar composition to TGFβ1, sharing 80% of its protein sequence, and is highly functional in the respiratory system through its role in lung morphogenesis, particularly through epithelial-mesenchymal interactions (Proetzel et al., 1995) (Kaartinen et al., 1995). Out of three isoforms, TGFβ1 is the most abundantly present isoform, that has a pattern of ubiquitous expression (Khalil, 1999).

6.1.1.2. TGFβ1 role in cell growth regulation

TGFβ1 was first characterised as a cell growth inhibitor (Tucker et al., 1984). It has been found to suppress c-Myc expression, thus exhibiting an anti-proliferative function (Pientenpol et al., 1990). Another known function that was identified by Hannon et al., 1994, is the role of TGFβ1 in initiating cell cycle arrest. All of which confirmed the purpose of TGFβ1 in cell growth inhibition, hinting at a possible involvement in cancer development or progression when TGFβ1 signalling is deficient. Despite this, studies have reported contradictory observations regarding the functional role of TGFβ1 in cancer development and progression. Some revealed its positive role in angiogenesis (Xiong et al., 2002), tissue invasion (Ivanovic et al., 2006), and metastasis (Kong et al., 1995), highlighting that TGFβ1 function extends beyond cell growth inhibition and cycle arrest. It wasn't until 2012, where Cerami et al. identified the increase in *TGFβ1* expression in cancer genomics data, which was later confirmed by Gao et al., in 2013, that the complexity of TGFβ1 cellular mechanisms started to unravel.

6.1.1.3. Activation of the TGFβ1 protein

When *TGFβ1* gene is expressed, it gets transcribed to generate a Large Latent Complex (LLC). This complex consists of the TGFβ subunit, attached to a latency associated protein (LAP), which is linked to a Latent_TGFβ_binding protein (LTBP) via disulphide linkage between cysteine residues present at both subunits (**Figure 6.2**). The LAP prodomain is required for accurate folding and dimerization at the carboxy-terminal growth factor domain, which forms a mature active protein upon its peptide cleavage. The protein cleavage can take place proteolytically or autocatalytically in acidic conditions. It can also be activated via other factors including thrombospondin 1, integrins and reactive oxygen species

(Grayand Mason, 1990). Once activated, the LAP subunit opens, liberating the TGFβ subunit, which can then bind to TGFβ receptors, activating the TGFβ signalling pathway (Taylor, 2009).

6.1.1.4. TGFβ1 signal transduction through the SMAD dependent pathway

Upon activation, TGFβ1 ligand binds to TGFβRII receptor, which in turn phosphorylates and activates TGFβRI. The two receptors form a homodimeric or heterodimeric structure that phosphorylates a cascade of SMAD proteins required for signal transduction (Ohta et al., 1987; Dubois et al., 1995), starting with the Receptor SMADs. SMAD proteins are a family of signal transducers that allow an external stimulus to be translated and initiate specific intranuclear gene expression (Massague et al., 1998). Receptor SMADs (R-SMADs) in this pathway are SMAD2 and SMAD3, which both get phosphorylated with SARA, and form the R-SMADs complex. Phosphorylated R-SMADs then bind to SMAD4, which allows the whole complex to migrate to the nucleus where it regulates the expression of numerous genes involved in critical pathways dictating the cell fate (Watanabe et al., 2000; Massague, 2005). **Figure 6.1** illustrates this signalling pathway.

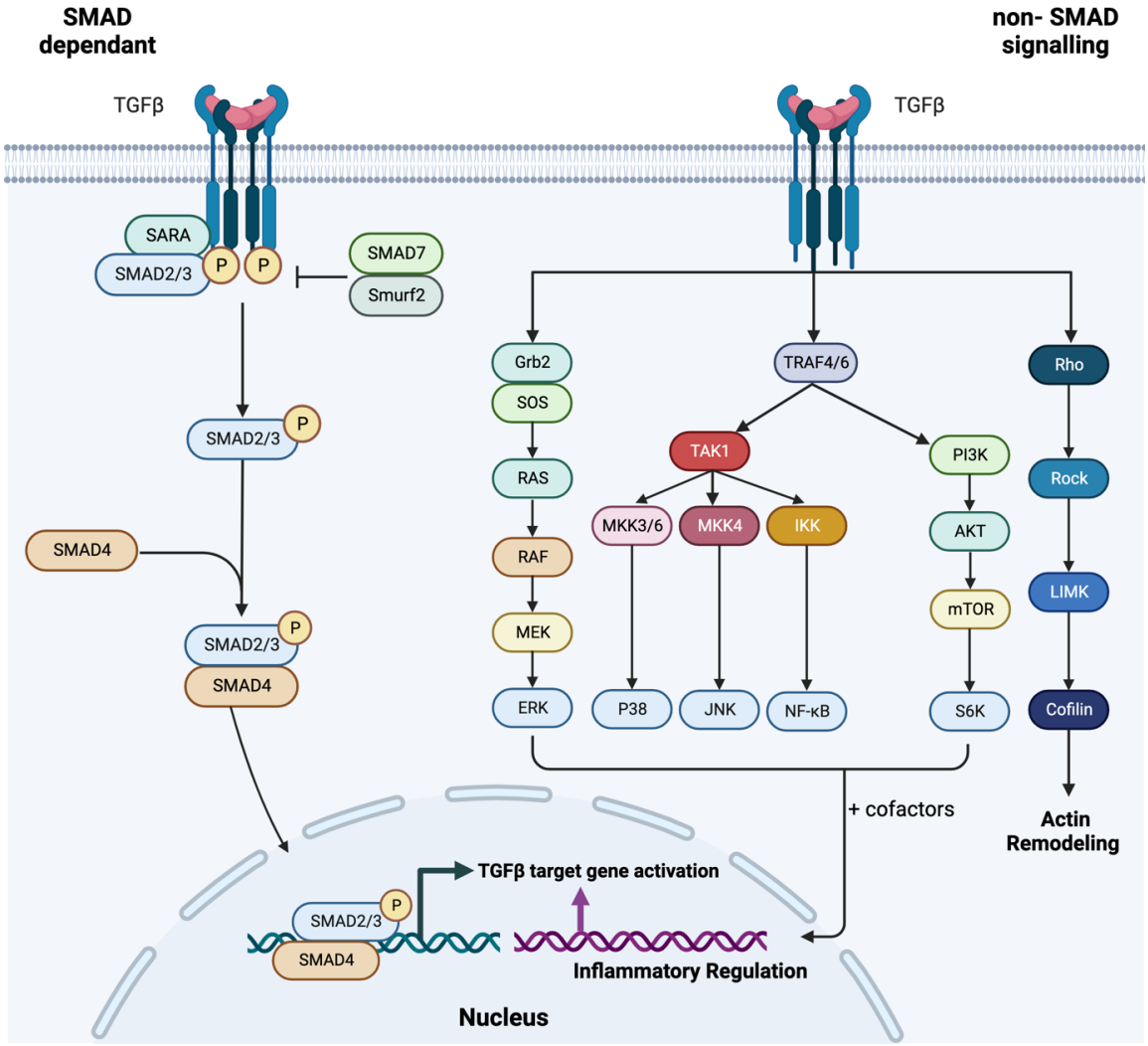


Figure 6.1: TGFβ signalling pathway

TGFβ can signal through several pathways in a SMAD dependent manner (left side), as well as non-SMAD canonical signalling (right side). In SMAD dependent signalling, a TGFβ ligand binds to a TGFβ receptor, forming a homodimer or a heterodimer that gets phosphorylated, and can subsequently phosphorylate SMAD2/3 and SARA. Phosphorylated SMAD2/3 complex can then bind to SMAD4, which allows for the complex to enter the nucleus and activate the expression of target genes. This pathway can be inhibited via the SMAD inhibitor, SMAD7 bound to Smurf2. On the other side, TGFβ signalling can take place without the presence of SMADs by activating a range of signal transducers including Grb2/SOS complex, TRAF4 or TRAF6, or Rho, which has a direct function in actin remodelling. These transducers are found upstream of various pathways including ERK, MAPK/JNK, and PI3K/AKT. Once activated, transducers can bind to a range of cofactors and enter the nucleus activating the expression of many processes, including inflammatory regulation. Figure generated using Biorender (https://www.biorender.com/).

6.1.1.5. TGFβ1 role in the inflammatory response

TGFβ1 was found to be highly expressed in the instance of an inflammatory response (Herrera-Molina and Von Bernhardi, 2005). One of functions, identified in the nervous system, is to limit neuroinflammation through regulation of the ERK signalling pathway (Grewal et al., 1999; Saud et al., 2005) and limit neurotoxicity caused by general inflammation (Tichauer et al., 2014). Several studies highlighted its association with major inflammatory pathways including ERK (Grewal et al., 1999), MAPK (Yue et al., 2000; Li et al., 2005), PI3K (Derynck and Zhang, 2003; Weiss and Attisano, 2013), and JAK/STAT (Xu et al., 2020). A link to the NF-κB pathway was also discovered, which was theoretically expected due to the presence of ERK and MAPK pathways upstream of NF-κB, which are both regulated by TGFβ1. Furthermore, it was confirmed in more recent studies that revealed the effect of NF-κB overactivation upon the increase in oxidative stress and inflammation leading to cell death. When NF-κB was blocked in aged mice, a recovery of ageing symptoms has been observed (Adler et al., 2008; Muriach et al., 2014). Another study involving the injection of TGFβ1 in the hypothalamus of mice resulted in a rise in the NF-κB inflammatory pathway (Yan et al., 2014).

6.1.2. Role of TGFβ signalling in the musculoskeletal and cardiovascular systems

In 1990, Dickinson et al., found *Tgfβ1* to be highly expressed in the cartilage, bones and skin of mice. This was initial evidence suggesting a critical role of the gene in the musculoskeletal system. It was later identified as major inducer of mesenchymal stem cell condensation in early chondrogenesis. Its involvement persists at later stages to allow for further chondrocyte proliferation (Tuli et al., 2003), and maturation by inducing the expression of various ECM proteins, including collagens and aggrecans, while simultaneously inhibiting hypertrophy for healthy cartilage development (Song et al., 2007). Another role of TGFβ in bone development and morphogenesis was identified through ALK signalling regulating the BMP pathway. TGFβ was found to activate ALK1, which subsequently affect various bone morphogenetic processes performed via the BMP pathway. ALK1, 2, 3 and 6 form a complex, which phosphorylates the 1,5,8 SMAD complex required for chondrocytes and osteoblast regulation (Davidson et al., 2009). This pathway functions in parallel to TGFβ signalling through SMADs, in the aim of attaining bone homeostasis. Other musculoskeletal disorders linked to defective TGFβ signalling and abnormal bone development in a less direct, yet a highly involved manner is through *SMAD3* mutations. *SMAD3* gene mutations are the leading cause behind aneurysm-osteoarthritis syndrome, (van der Linde et al, 2013; van der Laar et al., 2012), Marfan syndrome (Kinulainen et al., 1994), and loeys-Dietz syndrome (Loeys et al., 2006).

In mice, *Tgfβ1* double knockout was found to cause several skeletal defects, including serious midline fusion issues that lead to early death (Dunker et al., 2002). Endocardial mesenchymal transformation is an early developmental step that takes place in the heart during embryonic development to allow for heart chamber separation and valve formation. It also has displayed an essential function in outflow tract formation (Azhar et al., 2003). Additionally, targeting of *Tgfβ1* prevented neointima hyperplasia, which is a process of vascular remodelling where vascular smooth muscles relocate to reside within the tunica intima layer resulting in thicker vascular walls that could lead to heart failure. This observation highlights the role of *Tgfβ1* in smooth muscle cell development. As when *Tgfβ1* is overexpressed or overactivated, it can lead to the development of a cardiovascular disease (Kobayashi et al., 2005). Another study highlighted the role of TGFβ1 in ECM formation, where TGFβ1 upregulation was detected in patients with atherosclerosis, leading to fatty streak lesion formation (Gourdy et al., 2007). Another way *Tgfβ1* contribute to heart ECM regulation is through the reduction in collagenase production, increase in expression of several MMPs inhibitors, which in turn inhibits ECM degradation, but may lead to ECM protein accumulation, negatively impacting the heart structure and function (Verrecchia and Mauviel, 2002).

6.1.3. TGFβ1 expression pattern in osteoarthritis and cardiovascular disease

TGFβ1 was found to be overexpressed in OA patients, which is an expected observation based on its previously identified roles in various bone related processes including chondrocyte differentiation (Zhou et al., 2004), regulation of bone mineralization (Ehnert et al., 2010) and extracellular matrix (Horiguchi et al., 2012). In healthy mice cartilage, TGFβ1 is found in an active state, it functions through TAK1 kinase signalling mediated by FoxO1 to maintain cartilage homeostasis (Wang et al., 2020). It also displays a chondroprotective role by promoting ECM formation, protecting against joint degradation caused by mechanical loading (Goldring et al., 2004; Van der Kraan, 2017). In OA, TGFβ1 induces ECM formation via VEGFa signalling pathway to initiate tissue repair (Lin et al., 2022). However, due to the complexity of the system, alterations including *Tgfβ1* expression upregulation can speed up disease progression.

Several mutations in the *TGFβ1* gene have been assigned as markers to OA pathogenesis based on GWAS data. One variant of *TGFβ1,* (T29-C), commonly present in Japanese and Chinese women was found associated with OA development (Yamada et al., 2000; Woo et al., 2004). Another *TGFβ1* mutation was identified as a Camurati-Engelmann disease biomarker, as it was commonly present in patient samples (Kinoshita et al., 2000).

TGFβ1 was also found to be overexpressed in several cardiovascular diseases. In plasma samples of patients with coronary artery disease, TGFβ1 was found in higher levels (Tashiro et al., 2002). Following angioplasty, a balloon-like implant that stretches narrow arteries, a TGFβ1 increase is observed, which as a growth factor, may lead to restenosis post angioplasty causing narrowing of the target vessel (Wildgruber et al., 2007). Regarding its identified genetic mutations, one study found a polymorphism (-509C/T) and the haplotype GTGC in the *TGFβ1* gene as markers for myocardial infarction, only in men (Koch et al., 2006). Another study found two variants, rs1800469 and rs1982073, in *TGFβ1* to increase the risk of coronary heart disease development in humans (Lu et al., 2012). All these studies point out the high association between *TGFβ1* and the cardiovascular system, yet the range of functions displayed by the gene constrains our ability to fully understand its mechanism of action.

6.1.4. Other associated diseases

With ageing, an increase in TGFβ1 levels is observed. On some occasions, the transcribed gene carry defects, leading to an impairment in protein function, manifesting as different illnesses. One example is cancer, although TGFβ1 at early stages functions as a tumour suppressor, it can transform to an oncogene by initiating tumour progression processes. These processes include the promotion of angiogenesis, immunosuppression, and the conversion of epithelial cells to mesenchymal cells capable

of migration. This action forces cells to lose their anti-proliferation and apoptotic capabilities leading to cancer. It must be noted that the escape shown by tumour cells from the inhibitory system is not exclusively limited to direct mutations within *TGFβ1*, as alterations to other components within the TGF-β signalling pathway can result in similar phenotypes. This includes alterations in inhibitory SMAD proteins like SMAD7 which when mutated can lead to the activation of TGFβ signalling, causing pancreatic cancer (Kleef et al., 1999). In contrast, *SMAD7* was found to be overexpressed in endometrial cancer (Dowdy et al., 2005). Further pathogenic effects associated with TGF-β signalling via SMADs arise from heightened SMAD2 levels and mislocalization of SMAD2/3 complexes due to inadequate binding with SMAD4. This results in an oncogenic role for SMAD2/3 (Bertrand-Chapel et al., 2022).

TGFβ1 was found to be downregulated in the plasma of Alzheimer's Disease (AD) patients (Mocali et al., 2004). AD is characterised by the abnormal folding and/or accumulation of the Amyloid- β and Tau proteins. The defective protein folding occurs due to a transitional exposure of hydrophobic residues leading to protein aggregation. TAU protein aggregation is a brain-specific process, while amyloid-β accumulation can be found in different tissues, but mostly in neurons across the body. Increasing evidence suggests direct interactions between the two proteins. However, further investigations are required to understand the type of occurring interaction. In addition to the observed reduction in TGFβ1 levels in AD patients, reduction in SMAD3 or impairments to its function has been identified in some AD cases, which also leads to the accumulation of amyloid-β (Colangelo et al., 2002; Katsel et al., 2005; Tesseur et al., 2006). Evidently, TGFβ1 was found to enhance microglial function in amyloid-β clearance, reducing plaque deposition in AD mouse models (Coray et al., 2001).

Taken together, the involvement of TGFβ1 in numerous essential processes during early system/organ development, as well as in later regulatory stages points out to the expected detrimental effects in the absence of the functional gene. Even with the limited knowledge currently available, we can suggest an involvement of the gene in ECM regulation (Gourdy et al.,2007; Horiguchi et al., 2012) and its role in the inflammatory response (Hocking et al., 2006), which affects different tissues in a different manner leading to the rise of distant chronic diseases.

TGFβ1 was identified in 1983, since then it has been studied by several research groups focusing on its involvement in different diseases and body systems. However, the complexity of the gene due to its high involvement in a long list of processes is currently limiting our understanding of how the gene is regulated and how it interacts with other proteins to achieve its differential function within different tissues. Therefore, further investigations aiming to examine the function of *TGFβ1* as a broad cellular regulator rather than an independent gene and unravel interactions with other cellular proteins may help us understand how the gene itself is regulated and how it can exhibit the wide range of sometimes

opposite functions within the cell. In this project, I aim to investigate the function of *TGFβ1* in relation to other studied genes, common between OA and CVD pathogenesis.

6.2. Hypothesis and aims

Although *tgfβ1* zebrafish mutants have previously been generated, the study solely focused on the effect of *tgfβ1a* targeting on lateral line formation in zebrafish larvae, without examining any other structures (Xing et al., 2015). To our knowledge, the function of *tgfβ1a* has not yet been investigated in the skeletal and cardiovascular systems in zebrafish.

Hence, the main objective of this chapter is to investigate the impact of *tgfβ1a* on these systems. To achieve this, I aim to

- Determine the role of *tgfβ1a* in zebrafish skeletal and cardiac development by targeting the gene using CRISPR-Cas9 and monitoring the structural development of crispants using the fluorescent transgenic lines, *Tg(Col2a:mCherry)* for cartilage, and *Tg(myl7:lifeActGFP)* for heart visualisation.

- Assess the effect of *tgfβ1a* targeting on the skeletal system functionality by examining the larval and adult crispants swimming behaviour, in the presence and absence of light.

- Generate *ctsk* and *tgfβ1a* co-targeted larvae to determine any links between the two genes, based on the positive correlation in baseline expression identified through FUMA GWAS, and other data showing a possible link between the two genes (discussed further in this chapter).

6.3. Results

6.3.1. *TGFβ1a* shows high homology between zebrafish and humans

TGFβ1a gene is located on chromosome 19: 41,330,323-41,353,922 in humans, and chromosome 15: 2,776,506-2,803,313 in zebrafish. The gene is made up of 7 exons present in both species (**Figure 6.2a**). The encoded protein domain consists of two main subunits, a TGFb-propeptide and TGF_βeta_TGFβ1 with a homodimer interface, which is also present in the *GDF5* gene discussed previously. Both subunits are generally highly conserved across genes within the same superfamily and across different species. Although zebrafish only shows 45% similarity in protein sequence with the TGFβ1 variant in humans, the conservation of the two subunits, allows us to further investigate the function of the gene in zebrafish. Generated CRISPR target guides are shown in **Figure 6.2.** Another paralogue *tgfβ1b* is also found in zebrafish. However, in this study we focus on *tgfβ1a* variant in zebrafish in comparison to the human variant.

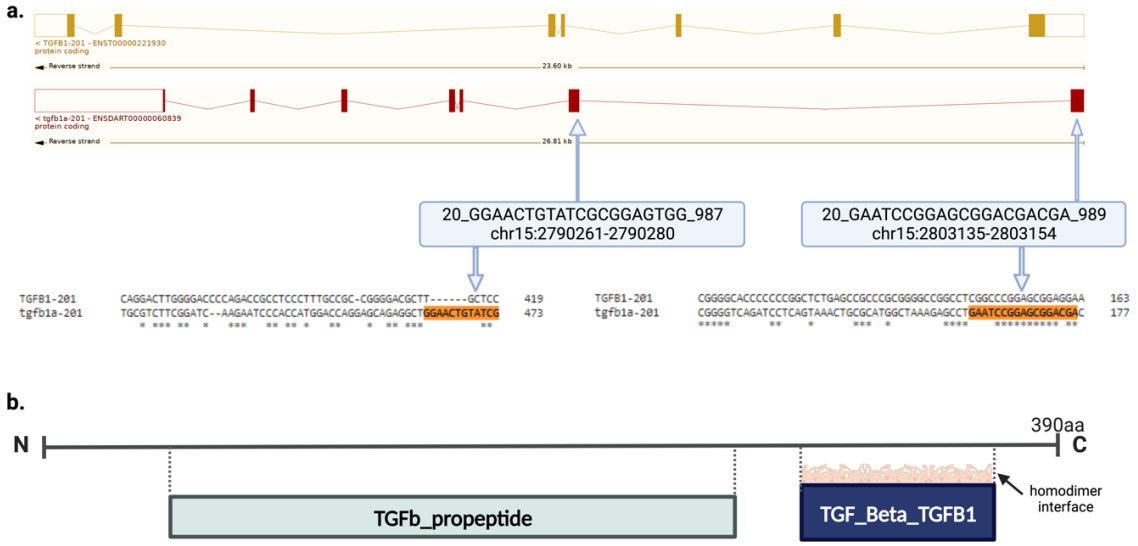


Figure 6.2: *TGFβ1a* gene composition and protein domain

(a) 7 exons can be found in *TGFβ1*, showing high conservation between humans (top) and zebrafish (bottom) genomes, although present on different chromosomes. The figure presents the gene on chromosome 19 in humans compared to the same gene on chromosome 15 in zebrafish. Blue rectangles highlight the location of the selected CRISPR target guides with their exact position and the targeting score, guide 1= 989 and guide 2= 987, based on the USCS Genome Browser tool (UCSC, 2019; https://genome.ucsc.edu/). **(b)** 390 amino acid protein domains with structures conserved across Homo sapiens and Danio rerio. A TGFβ -propeptide that requires proteolytic or autocatalytic cleavage to be activated and a TGF_BETA_TGFβ1 subunit, which is also a highly conserved domain found in most TGFβ superfamily members with a homodimer interface for active site binding to TGFβ receptors.

6.3.2. Skeletal defects observed in *tgfβ1α* crispants

At first glance, crispants appear to develop normally to 3dpf without major deformities in the general structure of the larvae (**Figure 6.3 C&D**), when compared to wildtype uninjected controls (**Figure 6.3 A&B**). However, using the *Tg(Col2a:mCherry)* transgenic line, abnormalities in the jaw cartilage structure were detected in the form of nonuniform chondrocyte organisation (**Figure 6.3 c&d**). Moreover, the palaquadrate structure show an overgrowth in the form of a thicker cartilage in some crispants, while at the same time display shrinkage of the same structure in other *tgfβ1a* targeted larvae (**green arrow**). This was a mere visual observation that requires quantification to confirm significance of the phenotype. Apart from the general changes in chondrocyte positioning, other jaw cartilage structures including the Meckel and ceratohyal cartilage, and the synovial joint seem to develop in a similar manner to wildtype larvae in the absence of *tgfβ1a*.

Most importantly, when confirming the success of CRISPR targeting, many targeted larvae seem to maintain the wildtype peak. In some cases, the wildtype peak was the only peak present, indicating no disruptions on a genetic level and failure of inducing any mutations in *tgfβ1a,* while other larvae show the presence of additional peaks along with the wildtype peak. This highlights a partial knockout, where the wildtype peak was still present in its functional state. The genotype could be matched to the observed phenotype, as healthy-looking larvae had the wildtype peak maintained within their corresponding genotype.

At 5dpf, different phenotypes appear in *tgfβ1a* targeted crispants. The overall body morphology maintains its shape **(Figure 6.4 C-F)**; however, the skull starts showing deformities at this stage visible in the brightfield view such as jaw elongation **(Figure 6.4 D)** or retraction **(Figure 6.4 E&F)**, which are the larvae with disruption in their genotype, according to the corresponding CRISPR-STAT **(d'-f')**. Meanwhile, **C** did not display any defects, and the wildtype peak was detected in the genotype, along with an additional insertion peak. This observation was confirmed by looking at the jaw cartilage in the transgenic line, displaying normal jaw cartilage development in **Figure 6.4 c**. However, a slightly protruding jaw was observed in detail **Figure 6.4 d**. More severe abnormalities were observed in **Figure 6.4 e&f,** both showing clear loss of wildtype peak (**e'&f'**). 29% of successfully targeted crispants develop a clear disruptive jaw phenotype as displayed in (**Figure 6.4 e**). 12.5% of crispants show a complete absence of the basihyal structure as presented in (**Figure 6.4 f**), highlighted by yellow dotted circle. Although CRISPR-STAT detected the initial efficiency of mutagenesis, sequencing is required to detect any correlation between the exact induced mutation and the obtained phenotype. This experiment was conducted in three replicates with 8 uninjected wildtype controls and 8 crispants, at each stage of 3 and 5dpf. Overall, the data suggests higher involvement of *tgfβ1a* at 5dpf when compared to 3dpf in cartilage development.

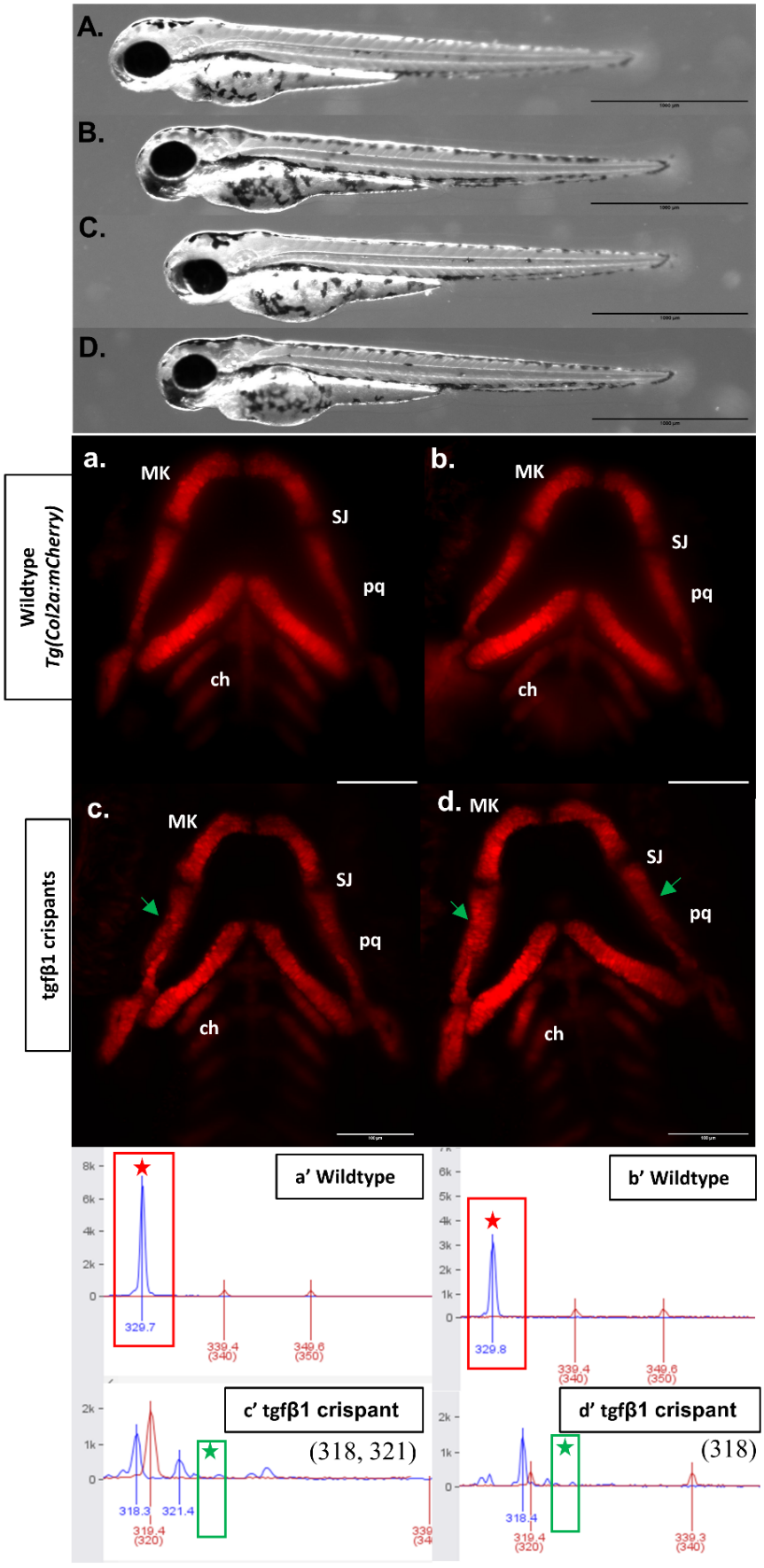
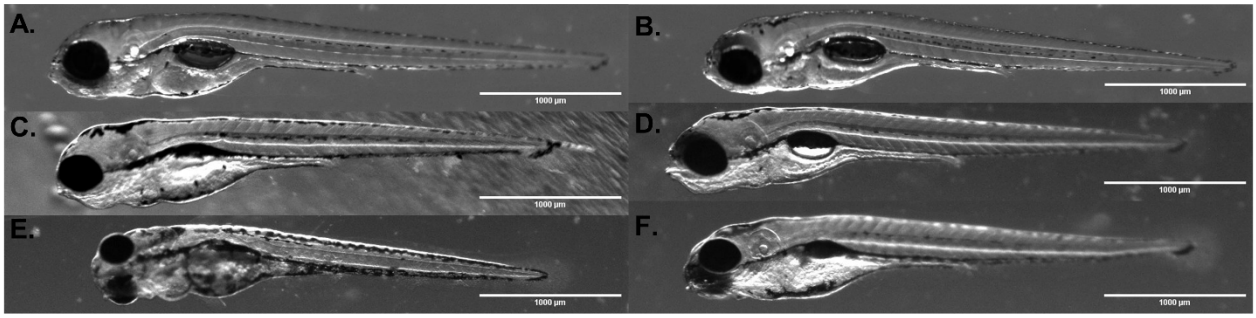

A.
B.
C.
D.
Wildtype
Tg(Col2a:mCherry)
tgfβ1 crispants
a.
MK
SJ
pq
ch
b.
MK
SJ
pq
ch
c.
MK
SJ
pq
ch
d.
MK
SJ
pq
ch
a' Wildtype
329.7
339.4
(340)
349.6
(350)
b' Wildtype
329.8
339.4
(340)
349.6
(350)
c' tgfβ1 crispant
(318, 321)
318.3
321.4
319.4
(320)
d' tgfβ1 crispant
(318)
318.4
319.4
(320)
339.3
(340)

Figure 6.3: Mild jaw cartilage defects in *tgfβ1a* crispants at 3dpf

Top panel examining the general morphology of larvae (**A-D**), while bottom panel displays the corresponding jaw cartilage structure within the presented larvae (**a-d**). **(A&B)** general morphology of wildtype controls imaged in brightfield, lateral view. **(C&D**) tgfβ1a 400pmol targeted larvae at 3dpf displaying unaffected whole-body structure when compared to wildtype controls. (**a&b)** control uninjected larvae obtained from the *Tg(Col2a:mCherry)* transgenic line representing normal general morphology and jaw cartilage structure in red at 3dpf, lateral view (n=24) along with its corresponding genotyping obtained using CRISPR-STAT. **(a'&b')** represent the wildtype allele peak highlighted by a red star at 329 indicating intact *tgfβ1a.* **(c&d)** *tgfβ1a* crispants (n=24) develop healthy body structure. However, cartilage appears slightly nonuniform in terms of chondrocyte organisation. Palaquadrate cartilage (pq-green arrow) appears to be the most affected structure showing overgrowth or shrinkage in size in different crispants. **(e'&f')** corresponding genotype confirming successful *tgfβ1a* targeting in presented 3dpf larvae. Green stars indicate the expected peak position when absent. MK = Meckel's cartilage, bh= Basihyal, ch= ceratohyal, pq= palaquadrate, sj= synovial joint. Top panel scale bar= 1000µm. Bottom panel scale bar= 100µm.


A.
B.
C.
D.
E.
F.

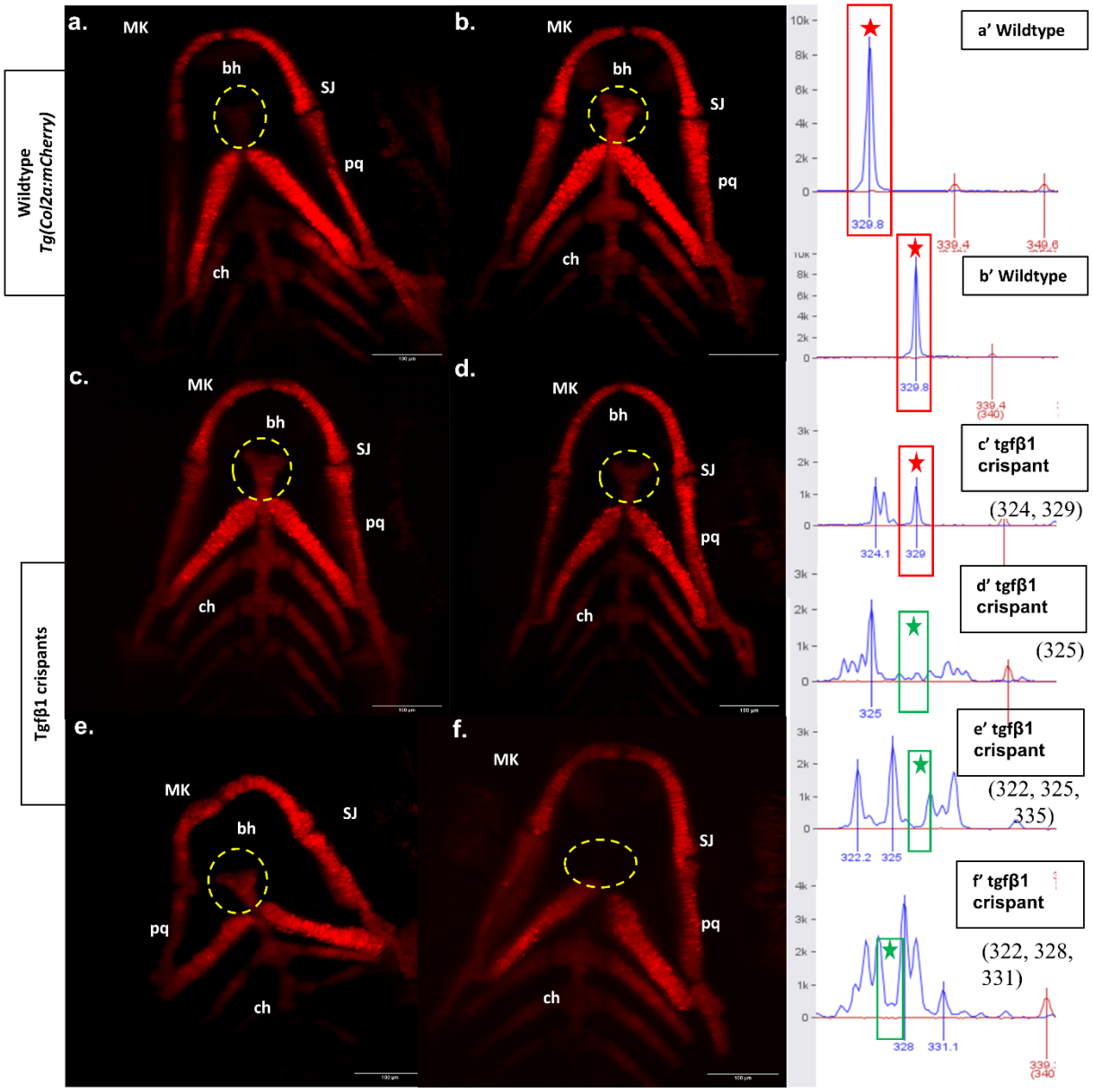

Wildtype
Tg(Col2a:mCherry)
Tgfβ1 crispants
a.
b.
c.
d.
e.
f.
MK
bh
SJ
pq
ch
a' Wildtype
b' Wildtype
c' tgfβ1 crispant
(324, 329)
d' tgfβ1 crispant
(325)
e' tgfβ1 crispant
(322, 325, 335)
f' tgfβ1 crispant
(322, 328, 331)

Figure 6.4: An array of phenotypes detected in *tgfβ1a* crispants at 5dpf

Top panel examining the general morphology of larvae (**A-F**), while bottom panel displays the corresponding jaw cartilage structure within the presented larvae (**a-f**). **(A&B)** general morphology of wildtype controls imaged in brightfield, lateral view. **(C)** Healthy-looking targeted crispant that shows unsuccessful targeting based on CRISPR-STAT **(c'). (D-F**) *tgfβ1a* crispants at 5dpf displaying a range of jaw deformities, mutagenesis confirmed **(d'-f')**. **(a&b)** control uninjected larvae obtained from the *Tg(Col2a:mCherry)* transgenic line showing normal general morphology and jaw cartilage structure in red at 5dpf (n=24) along with its corresponding genotyping obtained using CRISPR-STAT. **(a'&b')** represent the wildtype allele peak highlighted by a red star at 329 indicating intact *tgfβ1a.* **(c)** normal jaw cartilage development. ***(*d-f)** *tgfβ1a* successfully targeted larvae (n=24). Some larvae develop with mild abnormalities as shown in **(d),** with mildly disrupted jaw that is visible in brightfield view (top panel), yet the rest of the body develops normally. **(e&f)** present a more detrimental phenotype (n=7). **e.** shows complete disruption to the jaw cartilage affecting the position and orientation of the Meckel's, ceratohyal, paraquadrate and even the synovial joint. Meanwhile, **(f)** is missing the whole basihyal structure (yellow dotted shape) and shows deformities in the ceratohyal and plaquadrate (n=3). **(d'-f')** corresponding genotype confirming successful *tgfβ1a* targeting in presented 5dpf larvae. Green stars indicate the expected peak position when absent. MK = Meckel's cartilage, bh= Basihyal, ch= ceratohyal, pq= palaquadrate, sj= synovial joint. Top panel scale bar= 1000µm. Bottom panel scale bar= 100µm.

6.3.3. A range of abnormalities in the heart structure of *tgfβ1a* crispants at 5dpf

Using two different concentrations, 100pmol and 400pmol, of *tgfβ1a* CRISPR guide, the effect of targeting was tested starting from 3dpf. At this stage, many of the imaged embryos showed normal development of the heart. However, following the CRISPR-STAT analysis of the imaged larvae, unsuccessful targeting was detected in most of the imaged embryos. Therefore, we were unable to include the data for 3dpf heart development in *tgfβ1a* targeted larvae.

At 5dpf, defects were observed in 100pmol successfully targeted larvae. These defects include shape irregularities and size disproportion between the two chambers (**Figure 6.5 c&d)**. A milder phenotype with disruption in the looping process and subtle irregularities the atrium was observed in **f**, which shows a partial knockout success as the wildtype peak is still present, according to the CRISPR-STAT analysis (**f'**). It must be noted that most of the 400pmol targeted larvae imaged in this analysis, happens to display the wildtype allele, as shown in **e'**, and seem to develop normally. However, from the successfully targeted larvae a conclusion can still be drawn, even in the lower concetartion which had a remarkable impact on zebrafish heart development at 5dpf. This data indicates an indispensable functional role of *tgfβ1a* in zebrafish heart structural development at 5dpf, which leads to remarkable heart defects in its absence, which was not observed in the unsuccessfully targeted larvae. The presence of a wildtype peak along with other mutations resulted in milder defects. This experiment was conducted twice with a limited number of embryos, some of which were not successfully targeted. Therefore, further repeats must be performed to confirm the initially obtained results.

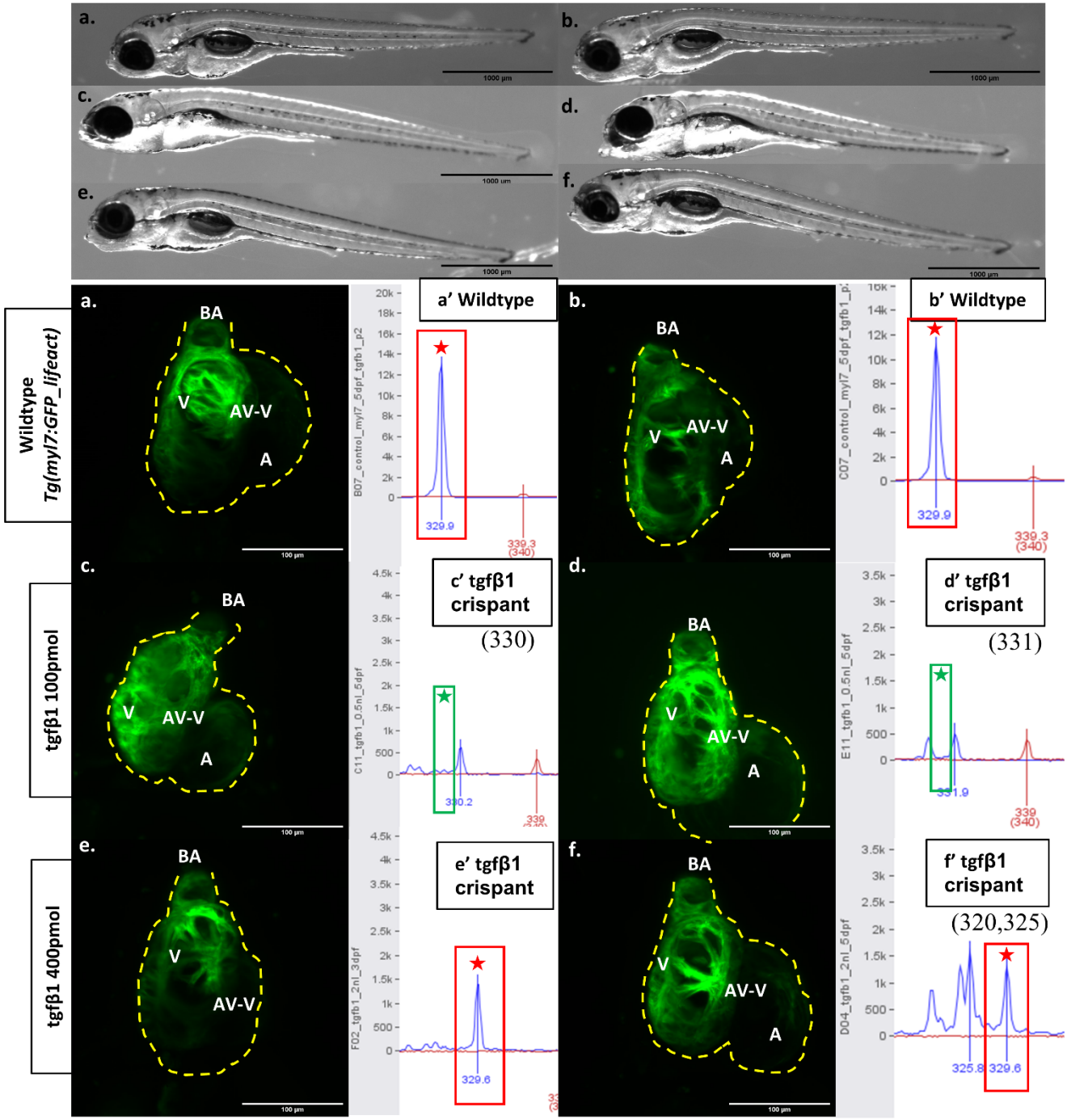

a.
b.
c.
d.
e.
f.
Wildtype
Tg(myl7:GFP_lifeact)
tgfβ1 100pmol
tgfβ1 400pmol
BA
V
AV-V
A
a' Wildtype
b' Wildtype
c' tgfβ1 crispant
(330)
d' tgfβ1 crispant
(331)
e' tgfβ1 crispant
f' tgfβ1 crispant
(320,325)
329.9
339.3
(340)
330.2
331.9
329.6
325.8

Figure 6.5: Abnormalities detected in the general morphology and the heart structure of successfully targeted *tgfβ1a* larvae at 5dpf

Top panel examining the general morphology of larvae (**A-F**), while bottom panel displays the corresponding heart structure within the presented larvae (**a-f**). **(A&B)** general morphology of wildtype controls imaged in brightfield, lateral view. **(C-D**) 100pmol successfully targeted *tgfβ1a* larvae displaying clear skull deformities and a protruding jaw. **(E)** unsuccessfully targeted *tgfβ1a* larvae showing normal development. **(F)** normal overall morphology due to the presence of WT allele based on CRISPR-STAT analysis **(f'). a'&b'** represent the wildtype allele peak highlighted by a red star at 329 indicating intact *tgfβ1a.* **c&d** 100pmol *tgfβ1a* targeted larvae (n=8) showing size disproportion between the two chambers. **c** showing a smaller atrium to ventricle ratio, while **d** present the opposite disproportion of a remarkably larger atrium and slightly smaller ventricle, which was also seen in successfully targeted larvae (**c'&d').** This was a mere observation that was not quantified to be validated. (**e)** seems to develop normally and the CRISPR-STAT analysis showed the unsuccessful targeting and the intact WT peak still present **e'. f** 400pmol *tgfβ1a* targeted larvae show milder defects, including a disruption in the looping process and a smaller ventricle, while the heart maintained its overall shape. A partial targeting was indicated by the presence of the WT peak in **f'**, accompanied by other insertions. Green stars indicate the expected peak position when absent. A= Atrium, AV-V= Atrioventricular Valve, BA= Bulbus Arteriosus, V= Ventricle. Top panel scale bar= 1000µm. Bottom panel scale bar= 100µm.

Figure 6.6 summarises the phenotype occurrence classified based on the overall body morphology of successfully targeted larvae at 5dpf. All controls appear to develop normally (**Figure 6.6 a**). Tyrosinase targeted embryos were used as a negative control and only show mild deformities in 2.9% of injected larvae (**Figure 6.6b**), indicating mild effect of injection. Mild phenotypes are characterised by the presence of tolerable structural differences such as mild disruptions to the jaw cartilage, jaw elongation or retraction, or mild curvature of the body. 14.3% of 100pmol targeted larvae display a mild phenotype, meanwhile, 29.2% of 400pmol targeted larvae display similar mild phenotype. Severe mutations refer to detrimental phenotypes that lead to monstrous larvae, which was only present in 4.2% of 400pmol targeted embryos.

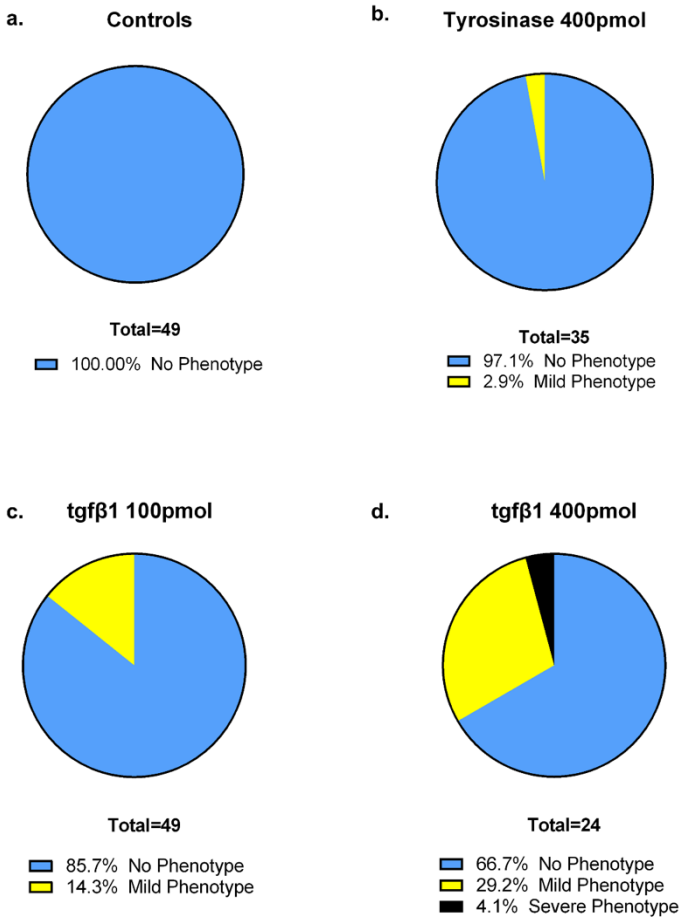


Figure 6.6: General morphology phenotype of *tgfβ1a* targeted larvae

(a) Healthy controls showing 100% normality of phenotype. **(b)** Tyrosinase as a negative control showing only 2.9% percentage abnormality indicating a mild effect of injection on larvae phenotype at 5dpf. **(c)** 85.7% of 100pmol injected display no unusual phenotype, while 14.3% show a mild phenotype characterised by mild disruption in the jaw area or slight body curvature. **(d)** 400pmol targeted larvae. 66.7% show no phenotype even in the presence of a mutation, 29.2% show mild phenotype such as general skeletal figure curvature, and only 4.2% display severe disruption in the general morphology. Figure generated using Graphpad Prism.

6.3.4. Limited effect of *tgfβ1a* targeting on the activity levels and swimming speed of adult crispants

To test the effect of *tgfβ1a* targeting on the functionality of the skeletal system in adulthood. Three experimental replicates were conducted in 1-year old adult crispants (n=6) versus wildtype controls (n=4). Consistently, no significant alterations in the activity levels or the rapid swimming abilities were detected in crispants when compared to wildtype controls (**Figure 6.7).** Fish display regular swimming activity (**Figure 6.7 a&c**) and have the ability to undergo hyperactive swimming in a slightly reduced manner that does not appear to be significantly different from the control group (**Figure 6.7 b&d**). No difference was observed in the light (**Figure 6.7 a&b**) and dark (**Figure 6.7 c&d**) conditions, indicating that *tgfβ1a* targeting does not have a direct impact on the nervous system that would affect the swimming pattern displayed by the fish. **Figure 6.8** displays the sequencing data of adult crispants utilized in the adult swimming analysis, confirming the successful targeting of *tgfβ1a.* As the tested fish were founders, they exhibit clear mosaicism with several mutations present within individual fish.

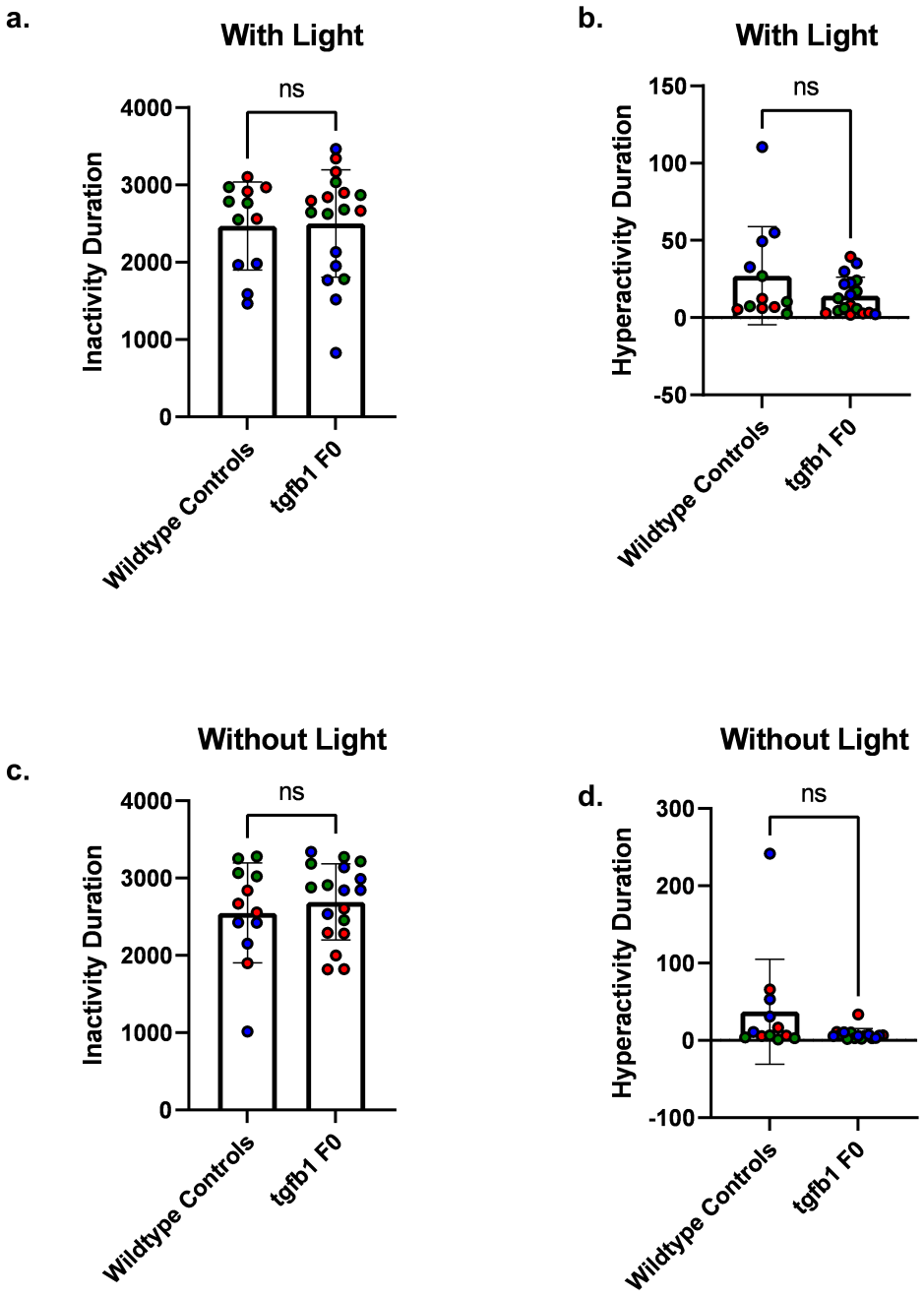


Figure 6.7: No remarkable change in adult *tgfβ1a* F0 activity duration or speed in the presence or absence of light

(a) non-significant change in general activity of *tgfβ1a* crispants (n=6) compared to wildtype controls (n=4) in the presence of light *(p-value= 0.8933).* **(b)** no significant alterations in the duration spent in a hyperactive swimming state in the presence of light *(p-value= 0.1239).* **(c)** non-significant change in general activity *(p-value= 0.5034)* or hyperactivity duration **(d)** *(p-value= 0.0952)* in the absence of light. Analysis was conducted by combining all repeats and tested using an unpaired student t-test, generated on Graphpad prism. Red= first replicate, Green= second replicate, Blue= third replicate.

a.

b.

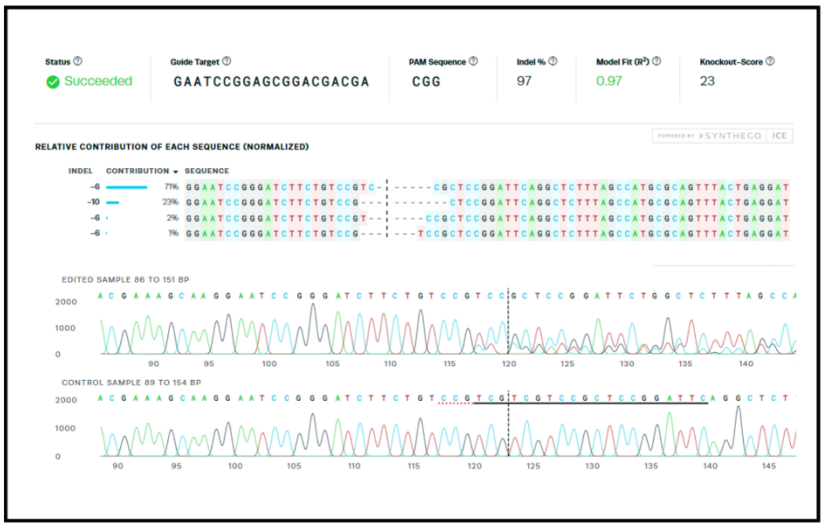
c.

Figure 6.8: Representative sequences of *tgfβ1a* crispants

d.


Status
Succeeded
Guide Target
GAATCCGGAGCGGACGACGA
PAM Sequence
CGG
Indel %
96
Model Fit (R^2)
0.96
Knockout-Score
20
RELATIVE CONTRIBUTION OF EACH SEQUENCE (NORMALIZED)
INDEL
CONTRIBUTION
SEQUENCE
EDITED SAMPLE 84 TO 149 BP
CONTROL SAMPLE 89 TO 154 BP

e.


Status
Succeeded
Guide Target
GAATCCGGAGCGGACGACGA
PAM Sequence
CGG
Indel %
97
Model Fit (R^2)
0.97
Knockout-Score
23
RELATIVE CONTRIBUTION OF EACH SEQUENCE (NORMALIZED)
INDEL
CONTRIBUTION
SEQUENCE
EDITED SAMPLE 86 TO 151 BP
CONTROL SAMPLE 89 TO 154 BP

f.


Status
Succeeded
Guide Target
GAATCCGGAGCGGACGACGA
PAM Sequence
CGG
Indel %
78
Model Fit (R^2)
0.78
Knockout-Score
51
RELATIVE CONTRIBUTION OF EACH SEQUENCE (NORMALIZED)
INDEL
CONTRIBUTION
SEQUENCE
EDITED SAMPLE 85 TO 150 BP
CONTROL SAMPLE 89 TO 154 BP

Figure 6.8: Representative sequences of *tgfβ1a* crispants (continued)

For each panel, the top section shows the percentage of each mutation present within individually sequenced fish. The bottom section shows the traces obtained from sequencing. **(a)** sequencing of an individual male generated crispant versus wildtype control showing mosaic patterns of mutations with a -3bp deletion present in 45% of sequenced sample, followed by -7bp found in 12% of sample, which may indicate a heterozygous crispant **(b)** sequencing data of another male crispant with +2bp mutation present in 32% of the sample, followed by a -6bp mutation in 8%. **(c)** male crispant sequencing with -14bp deletion in 23% of the sample, followed by a -3bp deletion in 14%. **(d)** sequencing data of an individual female crispant displaying mosaicism with -6bp mutation present in 64% of the sample, followed by a -10bp mutation in 20%. **(e)** sequencing data of a second female crispant with -6bp mutation present in 71% of the sample, followed by a -10bp mutation in 23%. **(f)** female crispnat with a -14bp deletion in 23% of the sample, followed by a -3bp deletion detected in 16% of the sequenced sample. Figure generated using Synthego ICE sequencing Analysis Tool.

6.3.5. Effect of *ctsk* and *tgfβ1a* co-targeting on zebrafish cartilage and heart development

For several years, *CTSK* has been identified as an intranuclear gene downstream of the TGF-β signalling pathway (Dole et al., 2017). Although its activation mechanism is still not fully understood, a recent study in zebrafish revealed a link between the activation of Ctsk from its pro-Ctsk state by binding to C4-S GAGs, a collagen IV derived glycosaminoglycan that has a major role in tissue membrane development and remodelling. C4-S GAGs is a known CTSK modulator, which functions by binding to inactive CTSK, proteolytically activating it. The activation of Ctsk was found to subsequently promote TGF-β signalling by activating the *chst11* gene. *chst11* (carbohydrate sulfotransferase 11) encodes for an enzyme essential for the formation of chondroitin sulphate, which is highly involved in connective tissue regulation and has shown critical function in the skeletal system (Bhattacharyya et al., 2015). This provided initial proof of a feedback loop that enables *ctsk* to demonstrate a regulatory impact on TGF-β signalling, a mechanism not yet examined.

Furthermore, detected tight co-expression pattern observed between *CTSK* and *TGFβ1* across different tissues (**Figure 3.3**), suggests a potential direct link between the two genes, worth further investigation. Therefore, in order to explore the hypothesised linkage between *CTSK* and *TGFβ1*, a co-injection of *ctsk* and *tgfβ1a* CRISPR-Cas9 guides in the same embryo was conducted. This approach aims to reveal potential crosslinks or synergy between the two genes.

Using two CRISPR-Cas9 sgRNAs for each gene, both genes were targeted simultaneously in single cell embryos. CRISPR-STAT analysis was conducted for both genes in each larva to confirm mutagenesis. **Figure 6.9** shows the effect of targeting *ctsk* and *tgfβ1a* on larval development at 3dpf. At this stage, co-injected crispants (**g&h**) appear to develop in a manner similar to both the wildtype control group (**a&b**) and *tgfβ1a* targeted group (**e&f**). Meanwhile, *ctsk* targeted larvae show clear structural abnormalities at 3dpf. At 5dpf, mild skeletal defects start evolving in just 4% of *tgfβ1a:ctsk* co-targeted crispants (**Figure 6.10, g&h**), compared to 47% of *ctsk* targeted larvae, and 29% of *tgfβ1a* targeted larvae. These mild defects include a smaller head, as shown in **h.** No severe phenotypes were observed in the overall morphology of the co-targeted larvae compared to 4% in *tgfβ1a* and 21% in ctsk targeted larvae. Co-targeted larvae fail to develop an inflated swim bladder (**g&h**) in 37.5% of tested embryos, as seen in 62.5% of *ctsk* crispants (**c&d**). The swim bladder deflation phenotype was not observed in the *tgfβ1a* crispants (**e&f**) or wildtype controls (**a&b**). This experiment was conducted twice, thus requiring an additional repetition to validate the obtained results.

Moreover, the functionality of the skeletal system was tested by monitoring the larval activity levels and hyperactive swimming abilities at 5dpf (**Figure 6.11**). *tgfβ1a:ctsk* crispants show significantly

higher activity levels, where larvae appear to spend less time in an inactive state when compared to wildtype controls of the same age (*p-value= 0.009)*. However, crispants show a highly significant reduction in the rapid movement abilities only in the presence of light *(p-value= <0.0001),* that was not apparent in the absence of light *(p-value=0.194).* Furthermore, larvae also display an increase in activity duration in the absence of light, but the change appears to be less significant when compared to the other light condition *(p-value= 0.026).*

To summarise, co-targeted larvae display a potential recovery in the overall morphology when compared to separate targeting, particularly of *ctsk* alone. A phenotype of a deflated swim bladder was observed in 37.5% of co-targeted larvae, which is significantly lower than the occurrence rate in *ctsk* crispants, but higher than *tgfβ1a* crispants, which did not display this phenotype. The initial behavioural analysis results suggest a significant impact on the larval swimming abilities in co-targeted larvae, which was not observed in the single targeting. *tgfβ1a:ctsk* targeted larvae show an increase in activity levels. However, larvae are unable to efficiently perform hyperactive swimming solely in the presence of light, indicating a direct connection between the co-targeting and the functionality of the nervous system.

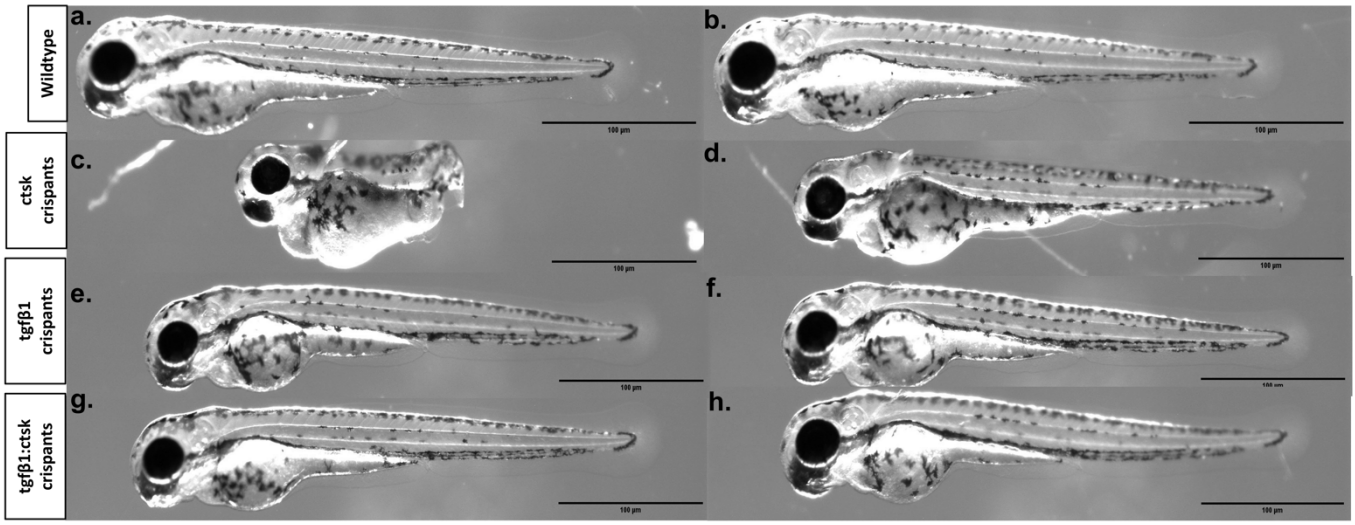


Figure 6.9: *tgfβ1a* and *ctsk* co-targeting compared to separate targeting effect on the general morphology of crispants at 3dpf

a&b wildtype control 3dpf larvae showing the expected morphology at the examined stage, lateral view (n=24). **c&d** 400pmol *ctsk* targeted larvae (n=40) showing clear phenotypical abnormalities as larvae appear to have shorter figures that is completely disrupted in **c,** while **d** looks more similar to wildtype, yet it shows clear disruption at the jaw area, enlarged yolk and abnormal retracted jaw. **e&f** *tgfβ1a* targeted larvae (n=24) that do not display clear phenotypical changes at this stage. **g&h** larvae targeted for *ctsk* and *tgfβ1a* combined (n=32) and larvae seems to develop normally as seen in wildtype controls and *tgfβ1a* separate targeting (**e&f**). Scale bar= 100µm.

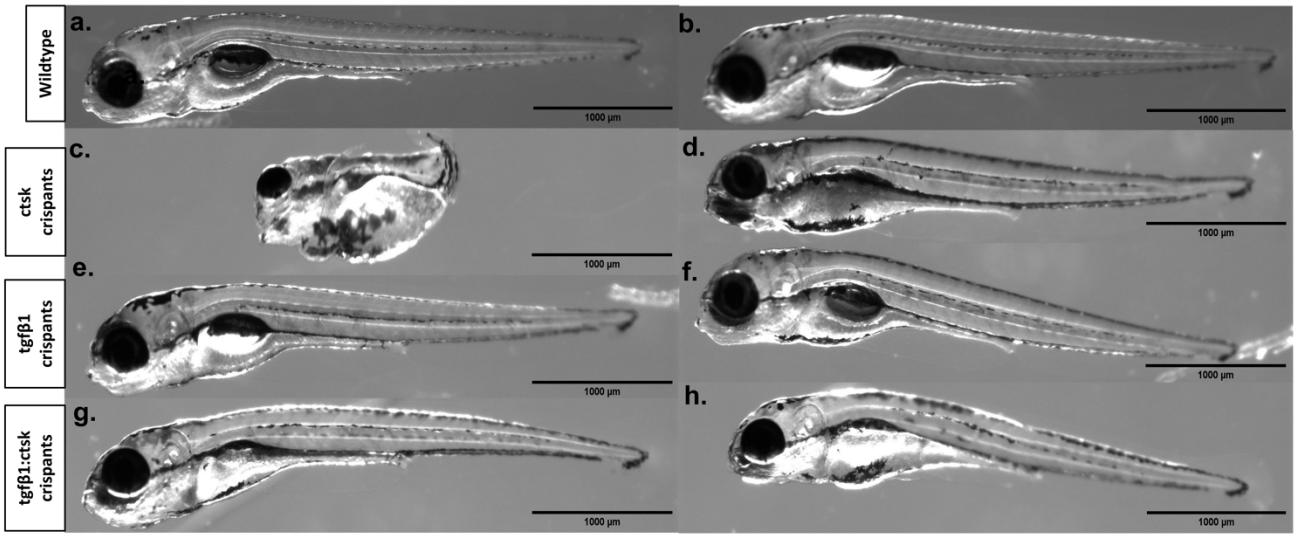


Figure 6.10: The effect of *ctsk* and *tgfβ1a* co-targeting compared to separate targeting on the overall morphology of crispants at 5dpf

(a&b) wildtype control 5dpf larvae (n=24) showing the expected morphology at the examined stage (lateral view). **(c&d)** 400pmol *ctsk* targeted larvae (n=56) showing clear phenotypical abnormalities as larvae appear to have shorter figures that is completely disrupted in (**c**). **d** shows clear disruption to the jaw and a smaller head. 62.5% of *ctsk* crispants fail to develop inflation of the swim bladder. (**e&f)** *tgfβ1a* targeted larvae (n=24) do not display remarkable abnormalities and seem to develop normally. **(g&h)** larvae targeted for *ctsk* 200pmol and *tgfβ1a* 200pmol combined, total 400pmol (n=32). Larvae have slight body curvature. 37.5% of co-injected larvae fail to develop an inflated swim bladder as seen in *ctsk* crispants **(c&d),** yet the general structure is fairly well maintained and looks phenotypically more uniform than *ctsk* targeted fish, yet more disrupted than *tgfβ1a* alone **(e&f**). Scale bar= 1000µm**.**

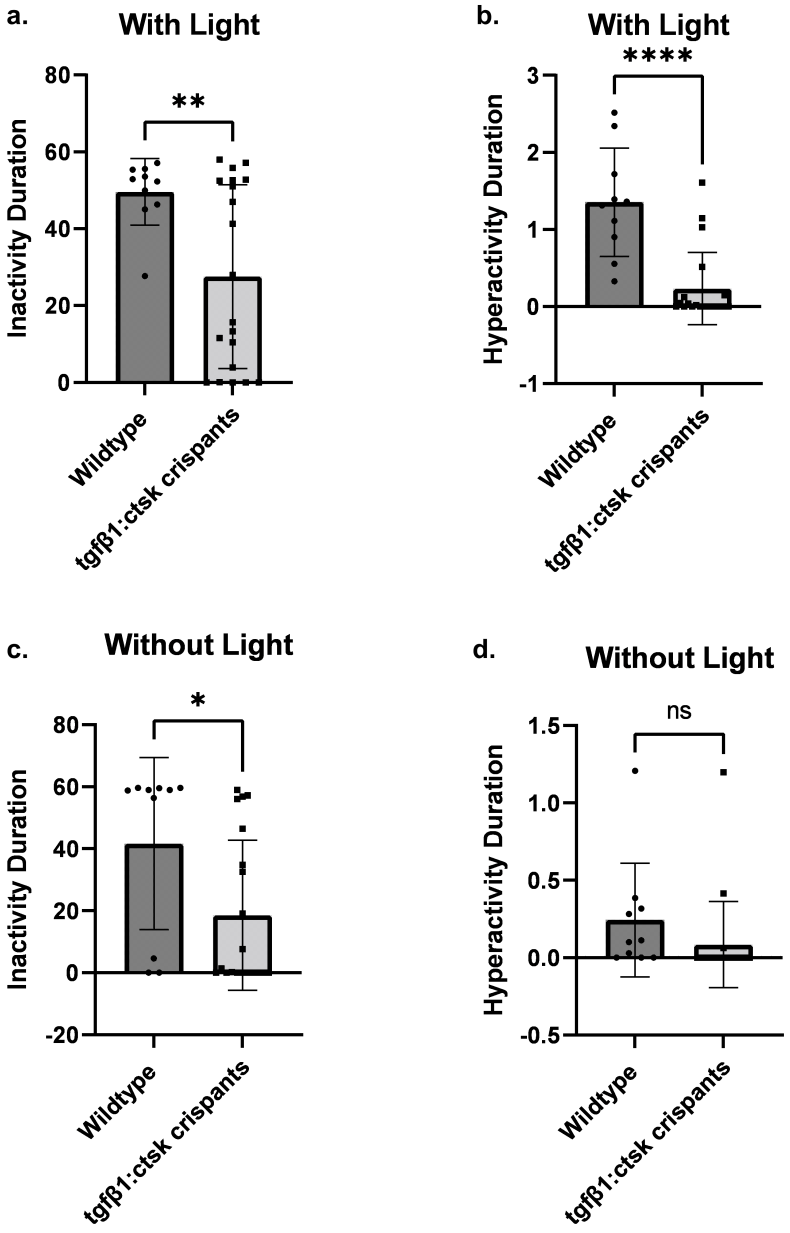


Figure 6.11: co-targeted embryos display higher activity compared to wildtype controls, and a highly significant reduction in rapid movement ability in the presence of light

a. significant reduction in the duration spent in an inactive state *(p-value= 0.009),* suggesting higher activity in *tgfβ1a:ctsk* targeted larvae (n= 20) when compared to wildtype controls (n=10). However, co-targeted crispants spend significantly less time in a hyperactive state indicating a disability to their speed in the presence of light *(p-value= <0.0001)* **(b). c.** In the absence of light, crispants appear to be less sedentary with lower significance compared to the other condition *(p-value= 0.026).* **d.** non-significant change to the swimming speed of crispants in the absence of light *(p-value= 0.193).* Statistical test performed as a student t-test using Graphpad Prism.

Using the *Tg(myl7:lifeActGFP)* transgenic line, the effect of *tgfβ1a:ctsk* co-targeting compared to separate targeting on heart development was examined. **Figure 6.12** displays the heart morphology at 3dpf. *ctsk* targeted (**c&d**) larvae show clear disruption to the heart structure, which include disruption in looping, and chamber shape and size. *tgfβ1a* crispants (**e&f**) display some abnormalities that appear milder than *ctsk* targeted larvae. Some irregularities in general shape were observed, yet the defect seemed tolerable and crispants were able to be raised up to 12 months. Interestingly, co-targeted crispants (**g&h**) display some heart defects, which appear to be milder compared to ctsk cripsants (**c&d**), and similar to *tgfβ1a* crispants (**e&f**).

At 5dpf, a similar result was observed where *ctsk* crispants displayed deformities within the heart structure, including shape irregularities and a visibly smaller ventricle (**Figure 6.13, c&d**) in 57% of imaged crispants. More severe deformities, which include significant shrinkage in size, and inversion of the heart, with disruptions to the looping process was observed in 36% of *ctsk* crispants. Such severe impact prevented us from further raising the crispants to adulthood to avoid any suffering to the organism. Mild deformities such as an incomplete heart rotation for optimum positioning was observed in 29% of t*gfβ1a* crispants (**e&f**). *tgfβ1a:ctsk* crispants (**g&h**) show deterioration when compared to the heart morphology at 3dpf. Mild defects including disproportion between the two chambers and disruption in the looping process was observed in 48% of co-targeted crispants. At this stage, co-targeted crispants again showed a phenotype that appears to be milder than *ctsk* crispants, yet more severe than *tgfβ1a* alone. Using heart images of generated crispants, heart volumetric measurements was conducted to detect any changes in the total heart volume at 5dpf associated with the induced mutation (**Figure 6.14**). A significant reduction in heart volume of *ctsk* 400pmol group *(p-value= 0.015)* and *tgfβ1a:ctsk* co-targeted crispants *(p-value= 0.028)* was detected. This experiment was only conducted once, therefore, requires further repeats to confirm the obtained result.

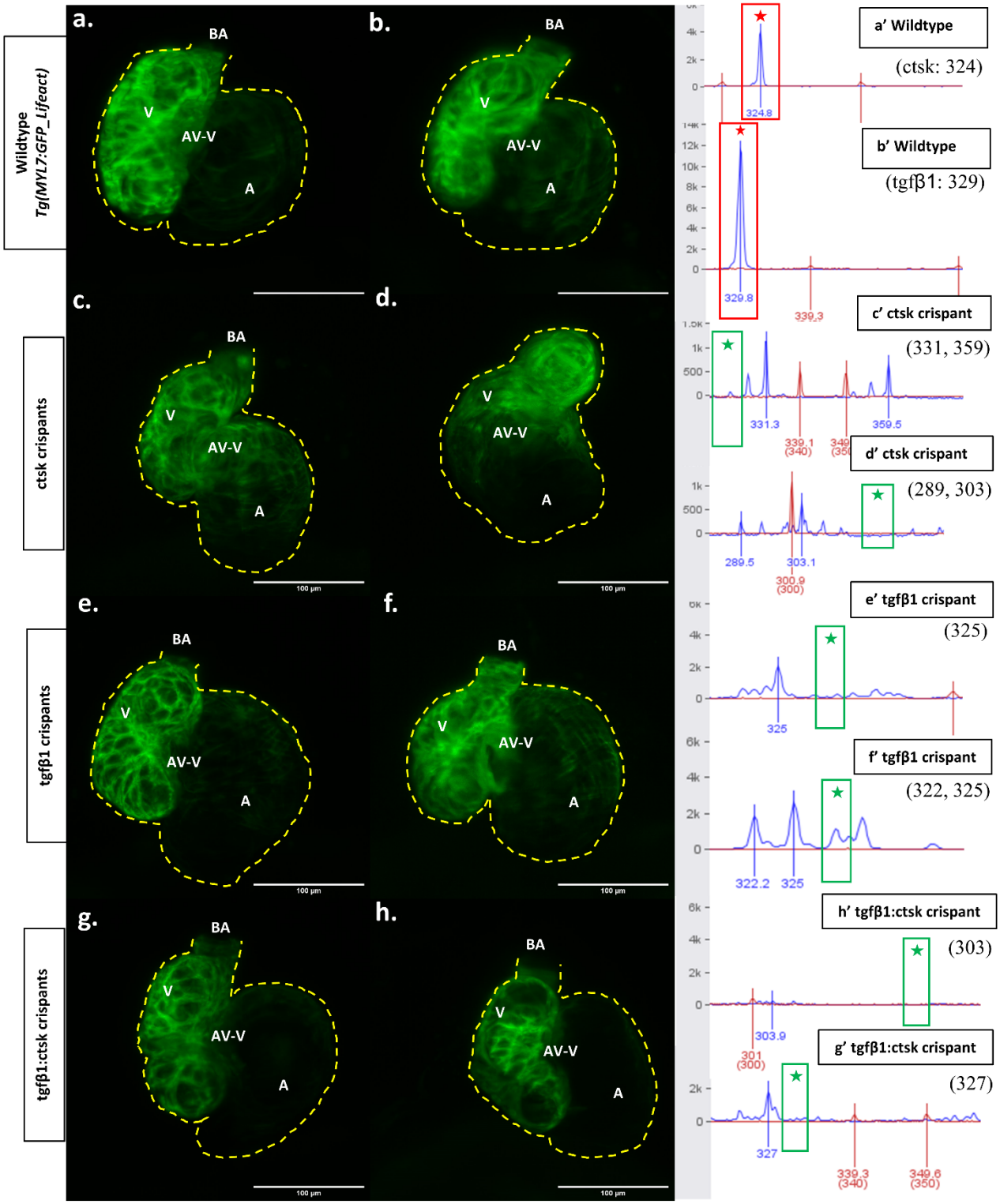

Wildtype
Tg(MYL7:GFP_Lifeact)
ctsk crispants
tgfβ1 crispants
tgfβ1:ctsk crispants
a.
b.
c.
d.
e.
f.
g.
h.
BA
V
AV-V
A
100 µm
a' Wildtype
(ctsk: 324)
324.8
b' Wildtype
(tgfβ1: 329)
329.8
339.3
c' ctsk crispant
(331, 359)
331.3
339.1
(340)
359.5
d' ctsk crispant
(289, 303)
289.5
303.1
300.9
(300)
e' tgfβ1 crispant
(325)
325
f' tgfβ1 crispant
(322, 325)
322.2
325
h' tgfβ1:ctsk crispant
(303)
303.9
301
(300)
g' tgfβ1:ctsk crispant
(327)
327
339.3
(340)
349.6
(350)

Figure 6.12: Co-targeting of *ctsk* and *tgfβ1a* show enhancement in heart structural development when compared to *ctsk* single targeting at 3dpf

(a&b) wildtype uninjected larvae (n=8) representing the expected heart structure (outlined in yellow) at the 3dpf stage in the *Tg(myl7:lifActGFP)* transgenic line, ventral view. **(a'&b')** CRISPR-STAT genotyping shows intact *ctsk* **(a')** wildtype peak at 324 and *tgfβ1a* **(b')** wildtype allele peaks at 329 (red star). **(c&d)** 400pmol *ctsk* targeted (n=14) heart morphology showing disruption to the overall shape of the heart and disproportion between the heart chambers. **(e&f)** 400pmol *tgfβ1a* targeted larvae (n=5) showing mild irregularities in the heart surface **(g&h).** *tgfβ1a:ctsk* co-targeted larvae (n=8) displaying mild phenotype similar to *tgfβ1a* targeted larvae. **(c'-h')** Green stars indicate the expected peak position when absent. V= Ventricle, A= Atrium, AV-V= Atrioventricular Valve, BA= Bulbus Arteriosus. Scale bar= 100µm.

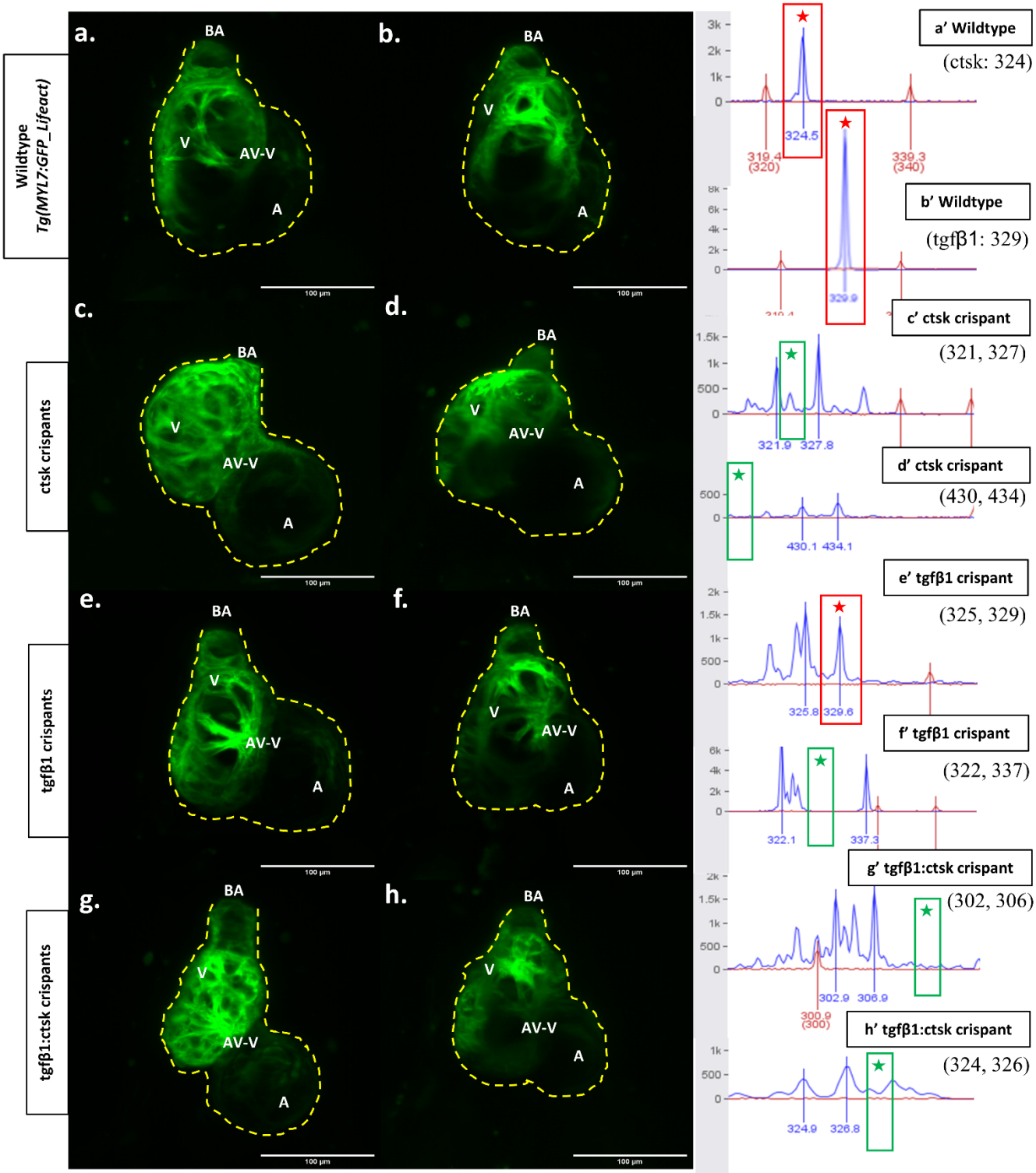

Wildtype
Tg(MYL7:GFP_Lifeact)
ctsk crispants
tgfβ1 crispants
tgfβ1:ctsk crispants
a.
b.
c.
d.
e.
f.
g.
h.
BA
V
AV-V
A
100 µm
a' Wildtype
(ctsk: 324)
b' Wildtype
(tgfβ1: 329)
c' ctsk crispant
(321, 327)
d' ctsk crispant
(430, 434)
e' tgfβ1 crispant
(325, 329)
f' tgfβ1 crispant
(322, 337)
g' tgfβ1:ctsk crispant
(302, 306)
h' tgfβ1:ctsk crispant
(324, 326)

Figure 6.13: Deterioration of the heart phenotype in *tgfβ1a:ctsk* co-targeting larvae at 5dpf

(a&b) wildtype uninjected larvae (n=8) displaying the expected heart structure (outlined in yellow) at the 5dpf stage in the *Tg(myl7:lifeActGFP)* transgenic line, ventral view. **(a'&b')** CRISPR-STAT genotyping shows intact *ctsk* **(a')** wildtype peak at 324 and *tgfβ1a* **(b')** wildtype allele peaks at 329 (red star). **(c&d)** 400pmol *ctsk* targeted (n=14) heart morphology showing clear disruption to heart morphology with smaller chambers, inefficient rotation for correct orientation, disruption in the looping process in 36% of crispants. General irregularities to the heart shape were also observed on a milder scale in 57% of imaged crispants. **(e&f)** 400pmol *tgfβ1a* targeted larvae (n=8) showing mild effect on heart morphology and incomplete rotation for correct alignment within the organism in 29% of successfully targeted crispants. **(g&h**) *tgfβ1a:ctsk* co-targeted larvae (n=8) display disruptions to the heart morphology, including smaller atrium and ventricle and disruption in looping process, present in 48% of imaged crispants. **(c'-h')** display the corresponding genotype and the green stars indicate the expected peak position when absent. V= Ventricle, A= Atrium, AV-V= Atrioventricular Valve, BA= Bulbus Arteriosus. Scale bar= 100µm.

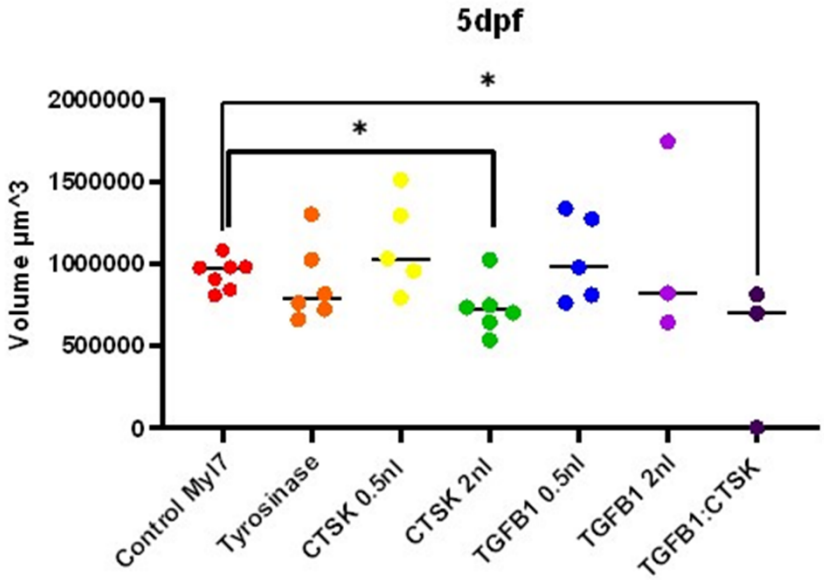


Figure 6.14: Changes in total heart volume of CRISPR targeted larvae at 5dpf

Significant reduction in 400pmol (2nl) *ctsk* (n=6) *(p-value= 0.015)* and 400pmol (2nl) *tgfβ1a:ctsk* crispants (n=3) at 5dpf *(p-value= 0.028)* when compared to wildtype controls. No significant change was detected in heart volume of *tgfβ1a* crispants when compared to wildtype controls at 5dpf. Statistical results obtained from One-way ANOVA test with multiple comparisons, performed using Graphpad prism.

Here, the overall morphology phenotype occurrence was compared between single targeted groups and the co-targeted larvae (**Figure 6.15**). *ctsk* targeted larvae show the highest percentage of abnormal larval development as 68% of larvae show an abnormal phenotype, with 21% of which a highly detrimental phenotype is present. The severe phenotype is characterised by complete disruption in the general morphology as shown in (**Figure 6.10, c**). *tgfβ1a* crispants showed defective general morphology in 29% of larvae with only 4% displaying a severe phenotype. Strikingly, *tgfβ1a:ctsk* co-targeted larvae only showed an abnormal mild phenotype in 4% of the 72 tested larvae with no evidence of severe phenotype in any of the crispants.

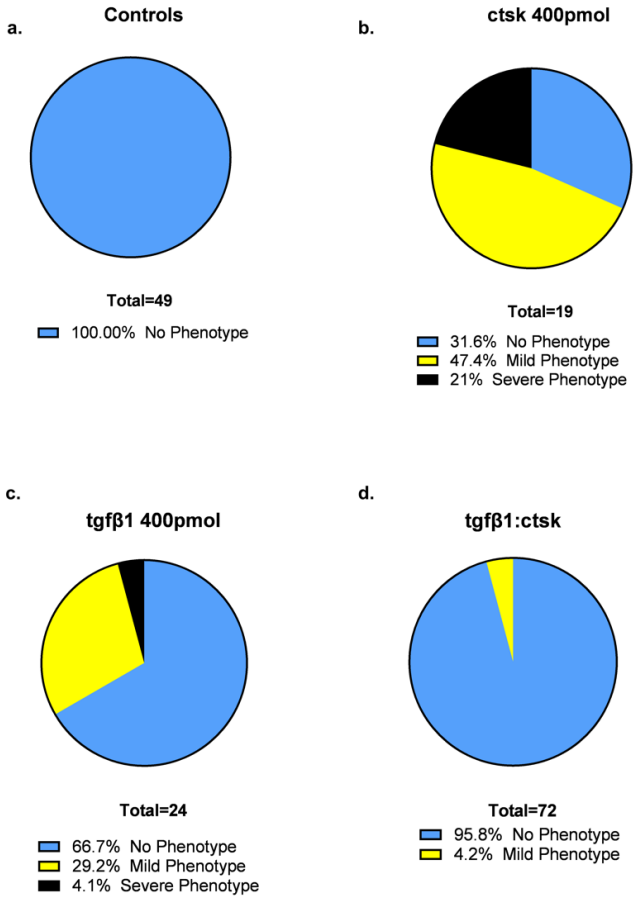


Figure 6.15: Overall morphology of *ctsk, tgfβ1a,* and *tgfβ1a:ctsk* co-targeted larvae

(a) Healthy controls showing 100% normality of phenotype. **(b)** 400pmol *ctsk* targeted larvae, with 31.6% of injected larvae showing no phenotype, 47.4% displaying mild deformities including mild spin curvature, and a small head. 21% of ctsk targeted larvae presented a severe phenotype that significantly altered their body morphology. **(c)** 66.7% of 400pmol *tgfβ1a* targeted larvae showed no phenotype, 29.2% displayed a mild phenotype characterised by a curved shorter body structure, while 4.1% of injected larvae showed severe disruption to the general morphology of the larvae. **(d)** a mere 4.2% of *tgfβ1a:ctsk* co-targeted larvae showed a mild phenotype, while the remaining 95.8% appeared healthy in terms of their general morphology. Figure generated using Graphpad Prism.

Taken together, the data indicate an enhancement in the pathogenic impact on the overall morphology of *ctsk* and *tgfßla* targeting, when both genes are co-targeted within the same embryo. Larvae appear to develop a phenotype that is more tolerable than *ctsk* crispants, and similar to that of wildtype controls, with the exception of a failure in swim bladder inflation. This phenotype was observed in 37.5% of *tgfßla:ctsk* co-targeted larvae compared to 62% in *ctsk* crispants, and was not observed in any of the wildtype controls, or *tgfßla* crispants. *ctsk* and *tgfßla* co-targeted larvae appear to have mild defects such as a smaller head in just 4% of imaged larvae, and a deflation of the swim bladder present in 37.5% of co-targeted larvae.

Looking at the heart morphology, more prominent defects are visible. The defects appear mild in nature up to 3dpf. At 5dpf, a progression in phenotype occurs, resulting in co-targeted crisprants exhibiting mild heart deformities including a disproportionate size between the two heart chambers and incomplete heart rotation for proper alignment within the organism. This phenotype is present in 48% of co-targeted larvae, compared to 29% in *tgfßla* crispants, and 57% in *ctsk* crispants, which also show more severe heart defects in 36% of targeted larvae.

Following up, heart volumetric measurements were conducted to determine the effect of co-targeting on the total heart volume. At 5dpf, both 400pmol *ctsk* targeted larvae *(p-value= 0.015)* and co-targeted larvae *(p-value= 0.028)* have a significantly smaller heart volume when compared to wildtype controls. This result suggests a progressive effect of mutation in the heart phenotype that can significantly reduce the heart volume of crispants at 5dpf. As this analysis was conducted with a limited number of crispants, further repeats should be performed to confirm the observed phenotype.

The significant decrease in the frequency of abnormal general phenotypes in the co-targeted crispants. compared to both ctsk and *tgfßla* crispants indicates a potential restorative effect on skeletal morphology. Meanwhile, the progressive detrimental effect of the co-targeting on heart development, suggests a differential impact of co-targeting in different tissues. This differential effect further adds to the complexity of interactions between the two genes and highlight the importance of studying their function in various tissues simultaneously to have a comprehensive understanding of their mechanisms of action.

Although co-targeting appears to have only a mild pathogenic impact on the skeletal structure based on the general morphology of crispants, the functionality of the musculoskeletal system seems to be highly impacted causing a significant increase in the overall activity of the larvae in both light conditions (with light: *p-value= 0.009*) (without light: *p-value= 0.026*). Meanwhile, co-targeted larvae exhibited a significant reduction in their hyperactive swimming abilities, only in the presence of light *(p-value= <0.0001)*. In the absence of light, co-targeted larvae did not display an altered swimming pattern

regarding their hyperactive swimming duration *(p-value= 0.193)*. This suggests a possible role of both genes in the nervous system, which was not detected when the genes were targeted separately, hinting at their coupled involvement in the development of the nervous system, and/or its regulation. However, this experiment was only conducted once. Therefore, further repeats must be performed to validate this initial result.

6.3.6. Effect of *gdf5* modulation on *ctsk* and *tgfβ1a* expression

After identifying a potential mechanism linking *ctsk* and *tgfβla,*and considering the established relationship between *TGFβ1* and *GDF5* as members of the same TGF-β superfamily, a question arises as to whether the modulation of *gdf5* would impact the expression pattern of *ctsk* and *tgfßla*. Previous studies have shown that TGFβ1 and GDF5 act antagonistically during early muscle development in embryos, where GDF5 induces myogenesis, while TGFβ1 inhibits this process, regulating muscle development (Wu et al., 2009). A similar function was also identified in the development of the skeletal system, whereby both proteins function antagonistically, regulating chondrogenesis to attain bone homeostasis (Wu et al., 2016). Therefore, I hypothesised a change in *tgfßla* expression levels in gdf5 F2 mutant larvae, and possibly an effect on *ctsk* levels based on the identified link between *tgfßla* and *ctsk*. To test this hypothesis, a quantitative PCR analysis of *ctsk* and *tgfßla* gene expression levels was conducted using wildtype nacre larvae compared to gdf5 F2 mutants at 5dpf. Two replicates were performed, and both display a significant reduction in the expression levels of both genes in the gdf5 mutants when compared to the wildtype controls, **Figure 6.16**. This indicates a direct correlation among all three genes, where they function concurrently, and any changes made in one gene has a direct impact on the rest of the network, potentially as an effort to counterbalance the pathogenic implication of the mutation and maintain homeostasis.

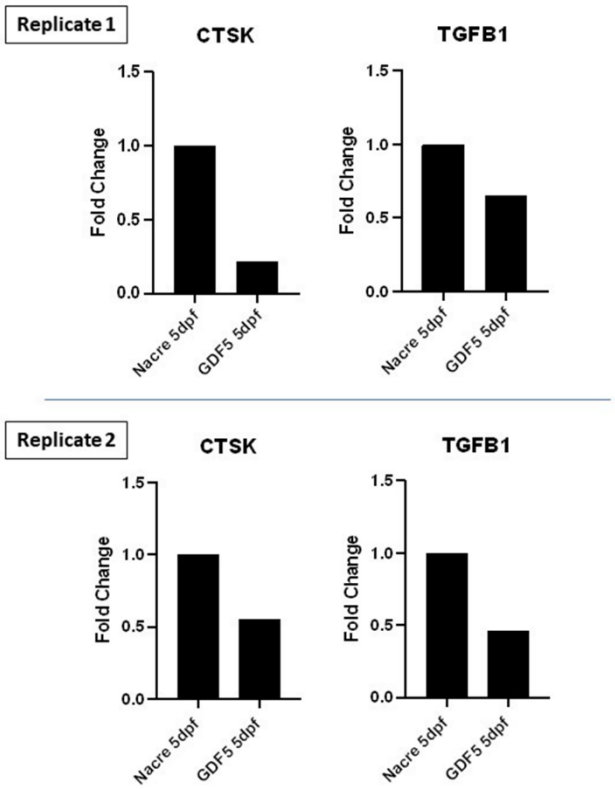


Figure 6.16: Consistent reduction in *ctsk* and *tgfβ1a* gene expression levels in gdf5 F2 mutants when compared to wildtype nacre controls

Data obtained from two repeats of qPCR analysis to investigate the gene expression levels of *ctsk* and *tgfβ1a* in gdf5 f2 mutants. β-actin was the chosen internal control. Nacre 5dpf was the calibrator for the fold change calculations. Figure generated using Graphpad Prism.

6.4. Discussion

Due to the highly complex nature of the TGFβ1 gene and its extensive involvement in essential processes, which exhibit clear dual functionality, I aim to present my observations with caution and acknowledge the limitations of my interpretation. Unfortunately, the utilised CRISPR guide targets for *tgfβ1a* did not display high targeting efficiency, leaving many of the injected larvae with the wildtype peak present, indicating unsuccessful targeting. Therefore, many of the imaged larvae had to be excluded from the analysis as they appear to maintain the wildtype peak according to the CRISPR-STAT analysis. Based on the obtained results from successfully targeted larvae, targeting of *tgfβ1a* shows mild deformities in the general structure of targeted larvae. In the 100pmol targeted larvae, 85.7% show no alterations in the overall morphology, while 14.3% show mild defects such as spine curvature and jaw elongation. Meanwhile, 400pmol targeted larvae show the same mild deformities at a higher percentage of 29. Additionally, 4% of the larvae that were targeted displayed significant disturbance to their morphology, featuring noteworthy spinal curvature and considerable reduction in whole-body length. When examining the jaw cartilage structure using the *Tg(Col2a:mCherry)* line, mild defects were observed at 3dpf. These include non-uniform chondrocyte organisation within the cartilage, and disruption to the paraquadrate structure, where in some larvae it appears thicker than the wildtype controls. However, in other cases, shrinkage of the paraquadrate structure was also observed. At 5dpf, some jaw abnormalities become clear through basic brightfield imaging showing a retraction and in some cases elongation of the jaw structure. By having a closer look at the studied structures using the transgenic line, more detrimental phenotypes within the jaw cartilage including remarkable curvature of the meckel's cartilage was present. Moreover, a complete absence of the basihyal structure was observed in 12.5% of targeted larvae. Although the targeting efficiency was detected using CRISPR-STAT, sequencing of individual crispants may reveal a link between the observed phenotype categories and the nature of the induced mutation.

The detrimental effect observed in just 4% of crispants was our primary expected result. According to published literature, TGFβ1 plays a critical role in ECM regulation by controlling the synthesis of collagen (Chan et al., 2005), fibronectin (Taipale et al., 1994) and proteoglycans (Okuda et al., 1990); all of which are critical components of skeletogenesis. Furthermore, it regulates other key bone-related growth factors such as bone morphogenetic proteins (BMPs) through ALK signalling (Davidson et al., 2009). Taken together, this highlights the critical functions of *TGFβ1* in skeletogenesis, which comes with no surprise following the findings obtained while studying other bone related genes, *ctsk* and *gdf5*. However, remarkable skeletal abnormalities were only present in 4% of generated crispants. This variability demonstrates the complexity of bone homeostasis, especially in the case of *TGFβ1* where its function goes beyond bone morphogenesis due to its involvement in structural development and cell

cycle regulation in numerous systems. Even when inducing a mutation within such a critical gene, other bone regulatory proteins function in parallel to the TGFβ signalling pathway through other pathways including Wnt/B-catenin (Duan et al., 2016), JAK/STAT (Li et al., 2013; Sanpaolo et al., 2020) and BMP signalling pathway (Cao et al., 2009), in an attempt to attain bone homeostasis, which compensate for *TGFβ1* loss of function. However, it must also be taken into account that *tgfβ1* has two paralogues, *tgfβ1a* and *tgfβ1b* in zebrafish. Therefore, the expected phenotype might be attenuated due to compensation provided by the second paralogue, as we only targeted *tgfβ1a.* To confirm or deny this hypothesis, generating double crispants targeting both paralogues would be necessary.

Apart from the previously discussed role of TGFβ1 in the regulation of the ECM that would justify its involvement within the cardiovascular system, several studies established direct links between TGFβ1 and several types of CVD, including atherosclerosis (Mallat et al., 2001) and coronary heart disease (Lu et al., 2012). TGFβ1 displays other functions within the cardiovascular system development and regulation including vascular remodelling. A study found cardiac remodelling enhancement when targeting *Tgfβ1*. Knockdown of *Tgfβ1* limited neoinitma hyperplasia by inhibiting the migration of vascular smooth muscle cells to the tunica intima layer, which prevented tissue accumulation that when overstimulated would lead to heart failure (Suwanabo et al., 2012). Our results from *tgfβ1a* targeted larvae in the *Tg(myl7:lifeActGFP)* transgenic line showed only mild defects that could be explained by one the two presented functions of *tgfβ1a* in the cardiovascular regulation. Two different CRISPR guide target concentrations were tested, 100pmol and 400pmol. However, only a few of 400pmol targeted larvae showed successful targeting. Therefore, a correlation between the observed phenotype for the two concentrations and their genotype could not be determined following the CRISPR-STAT analysis. However, repetition of the injection followed by sequencing of individual larvae may reveal a correlation between the phenotype and genotype. Moreover, generation of a stable mutant lines, after targeting both paralogues, would be ideal to provide more information regarding the effect of the specific induced mutation. At 5dpf, defects start appearing within the heart, resulting in chamber shape irregularities, which appear as bumps within the heart wall. However, to confirm the exact affected tissue, injection for targeting *tgfβ1a* must be performed in heart tissue specific transgenic lines. The irregularities and bumps observed in the heart chambers could be a result of disruptions in the ECM formation or due to defects in cardiac remodelling. Also, disproportion regarding the size of both chambers in relation to each other was observed, but not quantified for validation purposes. Analysis of separate chamber sizes would have to be performed to confirm the disproportionate factor.

TGFβ1 is not only involved in early chondrogenesis by initiating mesenchymal stem cell condensation (Tuli et al., 2003), it also has a role in chondrocyte maintenance by regulating cell proliferation (Simon et al., 1995) and inhibiting cartilage hypertrophy (Song et al., 2007). Thus, even if the mutagenesis impact on early development has been fairly mild, progression of the phenotype is expected in *tgfβ1a*

targeted adult fish and may affect the functionality of the skeletal system. Therefore, to test the increase of possible defects in adulthood regarding the fish swimming abilities, crispants were raised up to 12 months old and the effect of the targeting on the adult swimming abilities was tested in three replicates in the stated two conditions of absence and presence of light. Targeting of *tgfβ1a* does not appear to have an impact on the fish swimming duration or speed, in both light and dark conditions. Sequencing data of adult crispants utilised in this analysis suggest successful targeting of *tgfβ1a*. However, it must be noted that in several cases of sequenced crispants, a 3bp and a 6bp deletion was present, which may result in a functional protein, if the mutation happens to be missense. Even in the presence of other mutations, the fish may be heterozygous.

Unfortunately, due to time constraints, we were unable to perform micro-CT scans to adult crispants utilised in the behavioural analysis. Based on detected larval skeletal phenotype, which manifests as subtle deformities sometimes undetectable in the brightfield view, yet visible in the transgenic lines, we suspect persistence of skeletal abnormalities to the adulthood stage. This hypothesis is supported by multiple studies of TGFβ1 function in patients with OA, where TGFβ1 plays a crucial role in regulating extracellular matrix development (Ota et al., 2012) and bone mineralisation (Ehnert et al., 2010). Therefore, conducting micro-CT analysis to the adult crispants may reveal the role of *tgfβ1a* in later skeletal development or maintenance.

The complexity of *TGFβ1* due to its tight regulation through numerous factors, in addition to its involvement in many processes with sometimes opposite functions, makes *TGFβ1* a truly challenging gene to investigate. However, additional experiments may reveal deeper insights into the function of *TGFβ1*. Generation of a stable mutant line with a specific point mutation would allow us to track the effect of the studied mutation on the skeletal and heart structural development. Additionally, generation of double crispants mutants targeting both *tgfβ1* paralogues would reveal key information regarding the hypothesised compensation mechanism. Regarding the heart examinations, there were visible abnormalities such as irregularities in chamber shape and differences in size ratio between the two chambers. Although these abnormalities were clear, quantifying them would add a wealth of information on the exact effect of *tgfβ1a* modulation. Expanding the volumetric analysis of the heart to incorporate additional samples could unveil the potential impact of targeting on heart volume, which is suggested by the bumps visible in the obtained lightsheet images. Finally, assessing the skeletal and heart structures in adult crispants could provide key information on the effect of the mutation at later developmental stages.

Based on the detected correlation between *CTSK* and *TGFβ1* co-expression pattern, in addition to recent evidence suggesting a correlation between the two genes (Flanagen-Steet et al., 2018), an experiment was designed to investigate the effect of co-targeting both genes within the same organism. In this experiment, a severe detrimental impact on the larval health was hypothesised due to the array of critical functions displayed by both genes. Surprisingly, co-targeted larvae appear less affected by the targeting of both genes when compared to their separate targeting, when looking at the overall morphology of targeted larvae. By comparing the overall morphology of crispants, a complete eradication of the severe phenotype in co-targeted larvae was found, and only 4% of crispants showed a mild phenotype compared to 29% in *tgfβ1a* crispants and 47% of *ctsk* crispants. This suggests a possible recovery mechanism that must be cautiously considered. It should be noted that the larvae targeted using both *ctsk* and *tgfβ1a* guide targets were injected with a lower dose of 200pmol compared to the separate targeting, which used 400pmol. Despite this, even at the minimum concentration of 100pmol, 37.5% of the *ctsk* targeted larvae displayed a mild phenotype, and 14.29% of the *tgfβ1a* targeted larvae also exhibited a mild phenotype. With that in mind, the improvement in the overall phenotype of the co-targeted crispants, accounting for only 4% displaying a mild phenotype, still suggests a possible recovery mechanism in the co-targeted larvae.

Due to time limitations, co-targeting injections in the *Tg(Col2a:mCherry)* cartilage line were not performed. However, a single repeat was conducted for the larval behavioural analysis, in the presence and absence of light, to examine the functionality of the skeletal system in the co-targeted larvae. A significant increase in general activity was observed in both light conditions. Regarding their hyperactive swimming abilities, a high significant decrease was observed, only in the presence of light. Hyperactive swimming of co-targeted larvae does not seem to be affected in the absence of light. Both genes were identified separately as regulators of the nervous system (Kiefer et al., 1995; Dauth et al., 2020). However, based on our combined results, where targeting *ctsk* and *tgfβ1* separately did not show a significant change in movement in the two light conditions, yet co-targeting them revealed a differential effect between light and dark environment, suggest a possible functional link between the two genes in the nervous system. Thus, targeting both genes could be a great potential area for further investigation to confirm their coupled role in the nervous system.

Regarding the effect of co-targeting on the heart morphology, co-targeted crispants appear to initially develop with mild irregularities in the heart shape. However, the detected phenotype seems to worsen at 5dpf, allowing the larvae to display a heart morphology, similar to that of *ctsk* crispants. The heart display clear alterations in size and disproportion between the two chambers. The looping process is also affected and result in misalignment of the heart within the organism. Additionally, the co-targeted larvae showed a significant reduction in the total heart volume *(p-value= 0.028)*, as seen in *ctsk*

crispants *(p-value= 0.015)*. This suggests a functional role of both genes in the development of the cardiovascular system.

After exploring the possible link between *ctsk* and *tgfβ1a,* and due to the known link between *GDF5* and *TGFβ1* as members of the TGF-β superfamily, I suspected an effect of modulating *gdf5* on *ctsk* and *tgfβ1a*. Using qPCR, the gene expression levels of *ctsk* and *tgfβ1a* were measured in nacre wildtype controls and gdf5 F2 mutants. Throughout the two performed replicates, both genes showed a consistent reduction in their expression in the gdf5 mutants. This highlights a direct linkage between the three studied genes, which forces a change in gene expression levels of *ctsk* and *tgfβ1a* when *gdf5* was the targeted gene. As this experiment was performed twice, another replicate would be required to validate this result. Also, generating stable mutant lines for *ctsk* and *tgfβ1a*, and examining the other genes expression levels would be essential to confirm this hypothesis.

The obtained data highlights direct interactions between *ctsk* and *tgfβ1a*. It also shows the differential effect of their targeting in different systems. The co-targeting seems to counteract the severe detrimental effect of targeting *ctsk* on its own, improving the overall morphology of the larvae. An enhancement in the functionality of the skeletal system was also observed as co-targeted crispants show higher activity levels. However, they also show a significant reduction in their hyperactive swimming when compared to wildtype controls, only in the presence of light. Hence, co-targeting appears to differentially affect the musculoskeletal, cardiovascular, and nervous systems. This finding emphasises the genetic relatedness of apparently distinct chronic conditions, whilst also highlighting the intricacy of multimorbidity. It also reveals additional factors that should be considered in the investigation of multimorbidity which include: the necessity of using an in-vivo system, the importance of examining multiple structures simultaneously, the consequences of administering drug inhibitors targeting the mentioned pathways on other involved proteins, and the usefulness of broad comprehensive studies that focus on genes within systems, rather than identifying a single function.

Figure 6.17 presents the hypothesis regarding the molecular mechanism behind the enhancement of the skeletal phenotype in co-targeted crispants compared to the separate targeting. Bone homeostasis is an essential ongoing process to maintain healthy bones. Homeostasis is attained through a delicate balance between osteoblast activation for bone formation, and osteoclast activation for bone resorption (Rodan, 1998). *CTSK* is primarily expressed by osteoclasts (Ho et al., 1999) and is highly involved in their catabolic function. Therefore, when targeted a disruption to osteoclasts is expected. On the other hand, *TGFβ1* is known for its function in inducing osteoblast differentiation (Janssens et al., 2005) and therefore, in the loss or reduction of functional osteoclasts with the presence of osteoblast inducing protein, TGFβ1, severe disruption to bone homeostasis would be expected due to skewness of the two components. Thus, when both genes are targeted, balance may be achieved via the simultaneous

reduction in building and breaking of bone tissue. To validate this hypothesis, a quantitative real-time PCR must be performed to detect the levels of *tgfβ1a* in *ctsk* crispants and *ctsk* levels in *tgfβ1a* crispants to detect any changes in their expression levels when the other gene is targeted. However, due to time limitations, these experiments are yet to be performed. However, a similar experiment was performed using gdf5 F2 mutants to detect changes in expression levels of *ctsk* and *tgfβ1a*. As discussed, a remarkable reduction in both gene expression levels was detected in the two conducted replicates suggesting a direct link to the three studied genes that forces a change in their expression when one gene has been modulated.

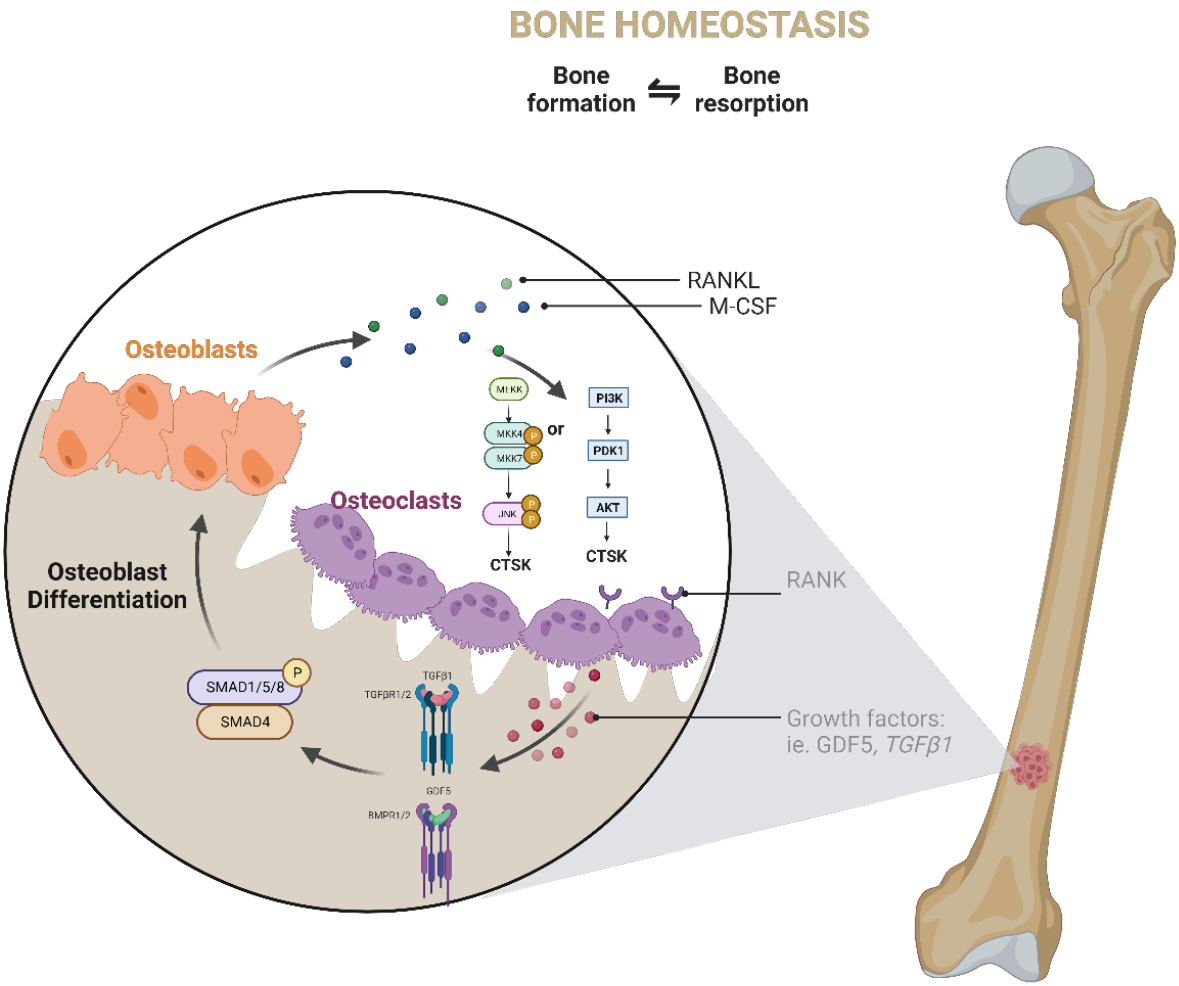
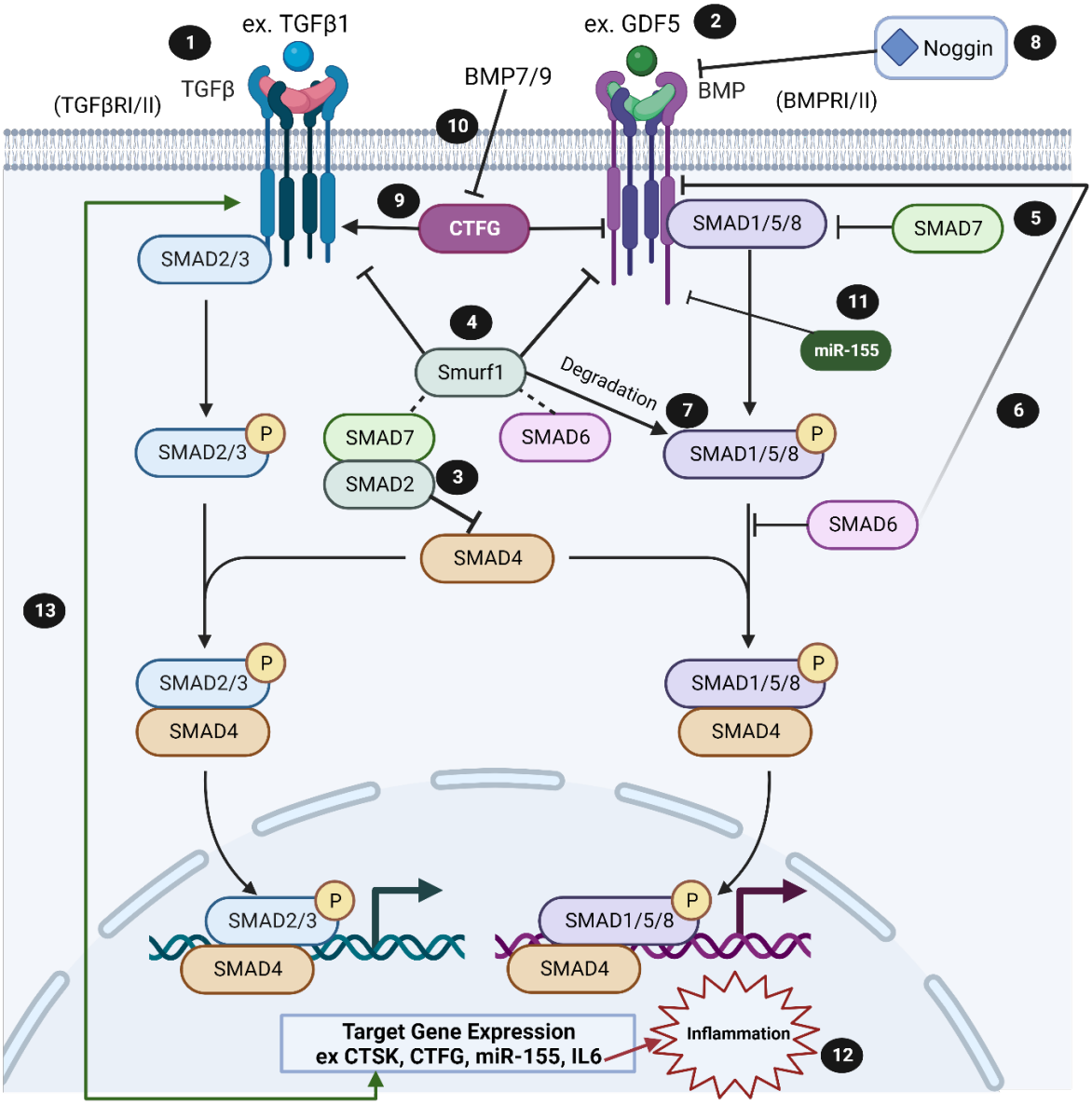


Figure 6.17: Suggested molecular mechanism behind the observed improvement in the skeletal phenotype of co-targeted crispants when compared to single targeted crispants

TGFβ1 and GDF5 are two growth factors that are highly involved in skeletal system regulation. Both cytokines signal through a SMAD-dependent pathway to induce osteoblast differentiation. Consecutively, osteoblasts release RANK ligand and M-CSF to activate osteoclast differentiation, which are highly abundant in CTSK and utilise it as one of the main regulators in bone resorption. The release of CTSK then signals back to TGFβ1 inhibiting its function on activating osteoblast differentiation in the aim of attaining bone homeostasis. Figure generated using Biorender (https://www.biorender.com/).

Figure 6.18 summarises molecular mechanisms to provide insight on the possible link between the three studied genes via a SMAD-dependent pathway. However, it must be noted that other SMAD-independent pathways were also found common between the two diseases and identified based on our bioinformatic analysis pipeline, yet functional analysis is required to validate their involvement through experimental investigation.



Please see figure legend on the following page.

Figure 6.18: Schematic diagram of molecular interactions linking three studied genes

1) TGFβ1 ligand binds to TGFβRI/II receptors, phosphorylating SMAD2/3 complex. SMAD2/3 the bind to SMAD4 for intranuclear transportation to activate target genes. 2) GDF5 ligand activate the BMP pathway signalling in parallel to TGF-β signalling pathway. BMPRI/II receptors form a homodimer or heterodimer complex, which allow for subsequent phosphorylation of SMAD1/5/8 complex that competes with TGF- β on SMAD4 binding to enter the nucleus and regulate downstream genes. 3) SMAD7 can bind to SMAD2 resulting in TGF-β signalling inhibition by preventing its binding to SMAD4. 4) Smurf1 can inhibit both pathways by inhibiting receptor homodimer/heterodimer binding, or through binding to SMAD6/7 inhibitory proteins. 5) SMAD7 can also bind to SMAD1/5/8 preventing their phosphorylation resulting in downregulation of the BMP pathway. 6) similar to SMAD7, SMAD6 can bind to BMPRI/II inhibiting the BMP signalling pathway. 7) SMAD6 can bind to Smurf1 inducing degradation of SMAD1/5/8 complex, inhibiting the BMP pathway. 8) Noggin can also inhibit the BMP pathway by blocking binding of growth factors such as GDF5. 9) Connective tissue growth factor (CTFG) gets activated through downstream signalling of both TGF-β and BMP pathways, however, the encoded protein has a regulatory effect on both pathways by activating TGFβRI/II, and inhibiting BMPRI/II binding, favouring TGF-β signalling. 10) Furthermore, BMP7 and BMP9 can both inhibit CTFG, stopping its inhibitory effect on the BMP pathway. 11) miR-155 is a microRNA molecule that is expressed downstream of both pathways and also show an inhibitory effect on BMP signalling by preventing receptor binding. 12) IL6 is another downstream expressed gene that is highly involved in pathogenic inflammatory side of bone homeostasis. 13) suggested feedback regulatory effect of CTSK on TGF-β signalling pathway (green arrow).

7. Discussion

7.1. Common mechanisms underlying the pathogenesis of OA and CVD

37 risk genes were identified common between the two diseases based on 133 OA and 2884 CVD GWAS data. By combining the candidate common risk genes for bioinformatic analysis, genes were found to be involved in two main functions. Firstly, they display a structural regulatory role through tissue maintenance, cell cycle and extracellular matrix regulation. Secondly, a function in inflammation regulation was identified via cytokine-cytokine interactions.

Through the bioinformatic analysis, 3 SNPs were found common between OA and CVD. Two out of the 3 SNPs are located on the two genes, *DPEP1* and *SLC39A8*. Both genes are involved in zinc transportation (Kawasaki et al., 1993; Begum et al., 2002). DPEP1 has also been found to regulate iron trafficking in kidney cells (Guan et al., 2021). This observation was intriguing and provided initial evidence to the possible underlying pathogenic effect of altered metal ion regulation on the development of OA and CVD. Using STRING to identify protein-protein interactions, significant direct interactions were found between TGFβ1 and SMAD3, and CAMK2B and MAPT proteins. TGFβ1 and SMAD3 interaction has been extensively studied and known for its function in signal transduction through TGFβ1 cytokine to intranuclear target genes through interactions with SMAD2 and SMAD4. Interestingly, CAMK2B is a calcium dependant protein kinase. CAMK2B was found to regulate dendritic cell formation and actin cytoskeleton in neurons (Koller et al., 2020). It also has a function in inflammation regulation by signalling through Ras/ERK pathway (Mohanan et al., 2022). One study discovered the co-expression pattern of CAMK2B and MAPT to be essential for TAU protein phosphorylation in Alzheimer's disease (Panda et al., 2003). However, the exact mechanism of action of the two proteins remains unknown. In relation to the two studied diseases, CAMK2B was identified for its role in cartilage and subchondral bone maintenance, as when inhibited, remarkable cartilage and bone degradation was observed (Nalesso et al., 2021). CAMK2B was also found to regulate the heart rate by catalysing the activation through phosphorylation of proteins involved in heart contraction-relaxation (Schulman and Anderson, 2011). MAPT was also identified for its role in cardiovascular performance, as one study found significant decline in heart efficiency in the absence of *Mapt* in knockout mouse models (Betrie et al., 2017). Currently, no link has been found between MAPT and OA. The significance of calcium signalling within the two diseases, provides support to the suggested pathogenic effect of Odanacatib, discussed in **section 4.4**, in CTSK inhibition that result in an increased risk of CVD development.

According to results obtained through the designed bioinformatic analysis, combined with literature searches, four signalling pathways were found to be involved in OA and CVD, possibly underlying their co-morbidity. These include the JAK/STAT (Mascareno et al., 2001; Malemud, 2018), MAPK/JNK (Muslin, 2008; Loeser et al., 2008; Sadoshima et al., 2002; Ge et al., 2017), PI3K/AKT/mTOR (Aoyagi and Matsui, 2011; Chen et al., 2012), and NF-κB (van der Heiden et al., 2010; Rigoglou and Papavassiliou, 2013) signalling pathways. This information highlights the complexity of multimorbidity, and the importance of investigating the role of candidate genes with consideration to their impact on these critical pathways. It also points to the necessity of investigating the impact of any potential drug inhibitor on these pathways to avoid detrimental side effects that prevent us from utilising these drugs after years of research and development.

Finally, three genes were selected for further functional analysis using the pipeline describe in **section 1.8**. *CTSK*, *GDF5* and *TGFβ1a* were selected due to consistent data indicating a significant involvement in both the musculoskeletal and cardiovascular systems. The genes were also found to be involved in other chronic diseases including Alzheimer's disease (Dauth et al., 2011; Wu et al., 2021; Bernhardi et al., 2015), Parkinson's (McGlinchey et al., 2020; Goulding et al., 2022; Comino et al., 2022) and various types of cancer, including prostate cancer (Wu et al., 2022; Conti et al., 2021; Shiota et al., 2021). Therefore, examining these genes would encompass a variety of prevalent chronic illnesses, facilitating the investigation of the pathogenic mechanisms underlying multimorbidity. Additionally, *CTSK* and *TGFβ1* show a clear positive correlation in their baseline expression levels in different tissues, based on the FUMA GWAS analysis, suggesting a direct link that is worth exploring. Finally, based on the microarray data obtained in our laboratory, exhibiting a significant increase in fold change of *Gdf5* expression levels post murine hip injury, *gdf5* was selected for further investigation. Another reason for selecting *GDF5* is its association with *TGFβ1* and established involvement in skeletal development, while its role in the cardiovascular system remains unknown.

7.2. Cathepsin K is required for early skeletogenesis and heart development

Cathepsin K was first identified in 1996 by Gelb et al., who discovered that pycnodysostosis disorder was a result of a mutation in *CTSK* that prevents the encoded protein from degrading collagen I, causing defects in bone homeostasis (Garnero et al., 1998). Since then, CTSK has been identified for its role in bone remodelling and resorption through the induction of bone matrix protein degeneration (Troen, 2004; Christensen and Shastri, 2015; Bonnet et al., 2017; Yoshida et al., 2018). It is characterised as a lysosomal protease mainly expressed and stored in the lysosomes. The gene can be activated downstream of several pathways including SMAD-dependent and SMAD-independent pathways (Troen, 2006). In cartilage tissue, most CTSK is expressed by the osteoclasts. However, more recent

studies found the gene to be expressed in the osteoblasts and osteocytes (Mandelin et al., 2006). The mechanism regulating the expression pattern within the different cells is currently unknown. Nevertheless, mechanical loading was one factor identified to particularly induce CTSK production in osteoblasts and osteocytes (Bonnet et al., 2018), yet the question arises of whether the CTSK produced by the osteoblasts and osteocytes function in a similar manner to the CTSK expressed by osteoclasts. Differential functions may explain the different roles exhibited by CTSK, because although CTSK was identified as catabolic factor, recently, it has been found to exhibit cardioprotective role by inhibiting degradation of the ECM (Fang et al., 2019). It must be noted that the cardioprotective role was found in mice models with myocardial infraction and the knockout of *Ctsk* worsened the symptoms by inducing cardiomyocyte death. From this data, Fang concluded that *Ctsk* exhibit cardioprotective roles, yet the pathogenic context of the study in pre-diseased mice must be considered.

Looking at the musculoskeletal role of *Ctsk*, knockout mice developed osteopetrosis like symptoms (Saftig et al., 1998), characterised by bone overgrowth. Moreover, overexpression of *Ctsk* resulted in a significant increase in bone turnover confirming the catabolic function of CTSK within the bone tissue (Kiviranta et al., 2001). Additionally, a recent study observed a delay in OA progression in *Ctsk* knockout mice (Sokki et al., 2018). This provides another indicator to the importance of studying the roles of genes within the context of a disease, but also the importance of studying the intrinsic role of the gene in an organ or system development. In osteoarthritis, *CTSK* was found overexpressed, which contribute to the excessive cartilage and underlying bone degeneration that eventually leads to OA development or progression (Logar et al., 2007). In cardiovascular disease, the gene is also found to be overexpressed. In the case of atherosclerosis, CTSK responds to the disturbed flow observed in atherosclerosis cases leading to its overexpression. When *CTSK* is overexpressed, it contributes to the rise in inflammation within endothelial walls, and cause disturbance to vascular remodelling. These changes eventually contribute to plaque accumulation seen in atherosclerosis (Fang et al., 2023). In aortic aneurysm, the increase in *CTSK* levels was found to induce muscle cell apoptosis, elastin degradation and T-cell proliferation (Zhao et al., 2015). Cell proliferation and apoptosis activation are two processes that were not previously identified to be regulated by CTSK. However, both processes are known to be regulated through TGF-β cytokines, including TGFβ1 (Simon et al., 1995; Ramesh et al., 2008). The TGF-β signalling pathway is found upstream of *CTSK* and can regulate its expression levels, which could suggest that observed phenotype in aortic aneurysm is rather a change in TGF-β signalling that subsequently increase *CTSK* levels and lead to disease progression. Another study in 2015 came up with the same conclusion suggesting the direct involvement of CTSK in apoptosis. In this case, they found knocking out *Ctsk* in aged mice to reduce impairments in heart function by reducing apoptosis (Hua et al., 2015). Yet, based on data obtained in this study, in addition to recent studies suggesting a direct crosstalk between TGFβ1 and CTSK (Bhattacharyya et al., 2015; Flanagan-

Steet et al., 2018; Lu et al., 2020), a plausible suggestion to the observed phenotype could rather be an impairment to TGF-β signalling pathway rather than CTSK itself. The relation of the two genes is not only limited to their co-expression pattern as seen in FUMA GWAS analysis, the localisation of CTSK can also impact TGF-β. Lu et al., 2020 revealed the negative effect of CTSK mislocalisation on TGF-β signalling disruption in *gnptab* knockout zebrafish. Mucolipidosis (MII) zebrafish models are generated via the induction of a mutation to *gnptab* gene, several lysosomal proteases are found to be affected, including Ctsk which get improperly mislocalised, resulting in its overactivation. Such increase was found to have a direct disruptive effect on TGF-β signalling, resulting in defective formation of myocardial and valvular structures leading to heart failure. In this study, inhibition of CTSK restored heart development (Lu et al., 2020).

Targeting *ctsk* in zebrafish embryos confirmed the essential function of the gene in the development of the skeletal and heart structures. Skeletal defects were detected as early as 3dpf causing mild defects in 47.4% of targeted larvae, while 21% targeted larvae display severe deformities. Clear abnormalities in the jaw structure were visible through basic brightfield imaging. When having a closer look at the structure using the *Tg(Col2a:mCherry)* transgenic line, disruptions to the meckel, palaquadrate and synovial joint were visible at 3dpf, leading to mild defects in 46% and severe defects in 20% of targeted larvae. The phenotype progressed with the skeletal development progression, affecting additional structures including the ceratohyal and basihyal structures at 5dpf, resulting in an increase to 30% of targeted larvae displaying severe abnormalities.

The heart also displayed clear defects in *ctsk* targeted larvae, as shown previously in *Ctsk* knockout mice (Guo et al., 2018). At 3dpf, irregularities within the chamber shape were observed, along with disruption to the looping process of the heart leading to abnormal looping. Initial in situ hybridization data suggest the disruption in the heart development to be initiated as early as 1dpf. However, due to the limited number of examined embryos, further repeats must be performed to confirm *ctsk* involvement at early heart development of 1dpf. The remarkable impact of *ctsk* targeting on both structures confirm the role of Ctsk in the skeletal and heart structural development in zebrafish. A possible explanation to the observed effect of what may seem unrelated structures, could be a simple link through connective tissues. As both the skeletal and heart structures are rich in connective tissues, which are also rich in collagen that is directly regulated by CTSK. Therefore, any disruption to CTSK will have an impact on both structures.

As part of the investigation of *ctsk* targeting effect on the studied structures, two concentrations of CRISPR guide targets were injected to detect any differential impact of targeting load on observed phenotype. Results reveal different phenotype in correlation with injected guide concentration. The lower 100pmol concentration seem to have a milder effect on the skeletal structure and heart

development when compared to 400pmol targeted larvae. Additionally, the two concentrations showed an opposite effect instead of a cumulative effect in some cases. In heart volumetric measurements, 100pmol showed non-significant increase in heart volume, while 400pmol showed a non-significant reduction in heart volume, when compared to wildtype controls. The difference in heart volume of the two concentrations allowed for a significant change (p-value = 0.0185) to be detected when comparing the two concentrations to each other, but not to the uninjected controls. Initially, this result was disregarded due to the small number of larvae included in the analysis. However, a similar pattern was observed in the larval behaviour analysis, where *ctsk* crispants showed a high significant difference between the two concentrations, especially in the presence of light. In this analysis, 400pmol targeted larvae appear to exhibit higher activity levels, with a significant reduction in hyperactive swimming when compared to 100pmol targeted larvae, but not to the uninjected controls. This experiment was also conducted once, thus require further repeats for validation, yet the high significant difference observed between the two concentrations, combined with the heart volumetric data, suggest a possible opposite impact of mutation based on guide target load. Therefore, generation and sequencing of stable mutant lines with the two different concentrations would be essential to confirm if the lower guide concentration result in lower targeting efficiency. This may help explain the discrepancies observed in the effect of *ctsk* targeting or protein inhibition on CVD progression, where in some cases *ctsk* targeting seem to contribute to CVD progression, while its inhibition improved the existing CVD condition (Hua et al., 2015; Fang et al., 2019).

As mentioned, the inhibition of CTSK attenuated CVD progression in aged mice (Fang et al., 2019) and in this study, researchers suggested that CTSK makes a good target for the treatment of CVD. However, the observed effect of inhibiting CTSK in OA patients using the same drug, odanacatib suggest otherwise. Although the drug has proved to significantly improve cartilage degeneration and attenuate OA progression, drug trials had to be terminated due to the increased risk of CVD development in OA patients undertaking the drug trial (Mullard, 2016). This highlights the importance of tight regulation between CTSK and other factors that function to attain homeostasis. Any intervention to one factor would subsequently affect other involved proteins and may manifest into another disease. One possible explanation to the pathogenic effect of CTSK inhibition, using odanacatib, on the cardiovascular system could be related to calcium signalling. One study generated *Ctsk* knockout mice and examined the effect of the knockout on CVD progression (Hua et al., 2015). In Ctsk mutants, they observed reduction in calcium signals in response to electric stimulus, in addition to an extension in time required for return of calcium levels to baseline, which depends on calcium export from the cytoplasm to the periphery outside the cell. This suggests that the targeting or inhibition of CTSK disrupt the intracellular calcium homeostasis, resulting in an increase of calcium levels, which may lead to heart calcification if the inhibition took place in a non-tissue specific manner.

To examine the effect of *ctsk* targeting on the later musculoskeletal system development, crispants were raised to adulthood, and a series of experiments were performed to examine the bone structure at later stages, and the functionality of the system by monitoring the fish swimming abilities. Due to the severe detrimental impact observed in 400pmol targeted larvae, only 100pmol targeted larvae were raised and used in the adult experiments. In 100pmol *ctsk* targeted adult fish, no change was observed in the swimming behaviour displayed by the fish in the presence or absence of light. In line, micro-CT scans revealed a non-significant effect of *ctsk* targeting on the bone volume and bone mineral density of adult fish bones. However, a significant increase in bone surface area was observed only in male crispants. Due to the limited number of scanned fish, it is hard to reach a conclusion to the increase in bone surface area. Nonetheless, based on our previous knowledge about CTSK function, it could be the case of disruption in remodelling and resorption in the absence of *ctsk*, which result in changes in bone morphology with irregularities in the bone shape. Such irregularities would not affect the total bone volume, yet it may result in grooves or projections within the bone structure, increasing its surface area.

To summarise, data confirmed the high involvement of *ctsk* in the skeletal and heart structural development possibly through its role in connective tissue, especially collagen regulation. It is worth noting that the *Tg(Col2a:mCherry)* transgenic line was employed for cartilage visualisation, which displays collagen II in red fluorescence. Meanwhile, collagen I makes up 90% of the bone matrix. Therefore, some information may be hidden about the exact structural changes and the use of a collagen I line may reveal additional information about the targeting effect on early skeletogenesis. Different CRISPR target concentration was used for the injections and showed different phenotypes that sometimes appear to have an opposite effect rather than a cumulative phenotypic effect. This highlights the effect of targeting or inhibition load on observed defects and may explain the discrepancies in knockout effect conducted by different studies. 100pmol targeting caused milder defects that do not affect the adult swimming behaviour and showed limited effect on bone structure of adult fish, that was only significant in male crispants. Therefore, generating different genetic alleles with varying strengths would be essential to confirm the impact of Ctsk load on the manifested phenotype.

7.3. *gdf5* is required for later stages of bone development, with no remarkable effect on early cartilage and heart development

GDF5 has been known for its function as a bone growth hormone (Hotten et al., 1996; Francis-West et al., 1999). Similar to CTSK, it is involved in SMAD-dependent and SMAD-independent pathways through the BMP signalling pathway. GDF5 was primarily though to have an activating role on BMP signalling as an anabolic factor, but recently Mang et al., 2020 discovered the differential effect of GDF5 binding to the two BMP receptors. When GDF5 binds to BMPR1B, it results in stable cartilage formation. On the other hand, when it binds to BMPR1A, with lower affinity, it signals with other BMPs

for osteogenic differentiation. Therefore, binding of GDF5 to BMPR1B is favoured to stabilise cartilage development and limit hypertrophy.

Gdf5 was discovered to exhibit high expression levels in mice during the initial stages of skeletogenesis. Its overexpression results in an increased size of skeletal structures (Sumaki et al., 1999; Buxton et al., 2001). Furthermore, the knockout of *Gdf5* in mouse models resulted in a shorter total body length with curvature in the long bones and loss of joint formation ability required in articular cartilage formation (Settle et al., 2003). Therefore, a similar detrimental effect was suspected in the generated *gdf5* crispants. To our surprise, *gdf5* crispants did not display abnormalities in early skeletogenesis up to 5dpf. The jaw structure appears to develop normally with expected articulation gap in the synovial joint structure. However, crispants displayed a defective bone structure in adulthood as they show a significant increase in bone mineral density and surface area without changing the total bone volume. The defect appears to have a tolerable effect on the functionality of the skeletal system as fish do not display an altered swimming behaviour when compared to wildtype controls. This suggest that *gdf5* is dispensable for early skeletogenesis, but is involved in further bone maturation and regulation, affecting the bone development at later stages.

Limited information is currently available regarding the functional role of *gdf5* in heart development. One study generated *Gdf5* knockout mice and showed an increase in apoptosis following a myocardial infarction, suggesting a reparative role of the gene (Huang et al., 2010). However, the intrinsic role of *GDF5* in heart development has not been investigated before. To study the effect of *gdf5* targeting on heart development, *gdf5* crispants were generated in the *Tg(myl7:lifeActGFP)* transgenic line. *gdf5* crispants did not display a significant change in heart structural development up to 5dpf. The total heart volume of crispants does not seem to be affected as well. Unfortunately, I was unable to investigate the heart structure or function in adults due to the lack of a heart visualisation method in adult zebrafish.

In the adult *gdf5* crispants, all males did not survive over 29 months. Meanwhile, all female crispants survived to the termination of the experiment. This suggests a distinct role of *gdf5* in the two sexes. Although only three males were monitored to adulthood, which require the use of a larger sample size to confirm this hypothesis. However, it goes in line with the difference in pathogenesis rate of bone diseases between the two sexes. Females have a higher risk of developing OA (Tschon et al., 2021). Additionally, rheumatoid arthritis shows a two-fold higher risk in females compared to males (Gerosa et al., 2008), and osteoporosis is four times more common in females (Alswat et al., 2017). Thus, gene mutation pathogenic effect may impact the sexes differently. Although in this case the difference does not manifest in a bone phenotype, and rather cause early death in males. However, it highlights the possible differential functions of *gdf5* in different sexes.

As gdf5 founders developed in a healthy manner similar to wildtype controls, we were able to raise the crispants and obtain F1 embryos for raising. At the adulthood stage, F1 fish did not display a change in their swimming abilities even though a mutation was confirmed by sequencing the F1 fish. Regarding obtained mutations, a 6bp deletion, an 8bp deletion and a 5bp deletion were the most commonly detected mutations that would theoretically have different impact on the encoded protein. Therefore, different mutants were tested in a behavioural analysis. Interestingly, unintended mutations were detected within the F1 generation, in addition to the four expected mutations in the gene. This could highlight potential evolution of the mutation when breeding two organisms carrying a mutation within the same gene. However, further investigation through sequencing of incrossed crispants with a mutation in the same gene must be performed to confirm this hypothesis.

Although *gdf5* targeting showed no remarkable effect on founders and F1 swimming abilities when compared to wildtype controls, F2 mutants displayed a high significant alteration in their swimming ability. Mutants showed a significant increase in general activity, only in the absence of light. No significant change was detected in their rapid hyperactive movement in the absence or presence of light. Due to the high significant change observed in the absence of light that was not detected in the presence of light, a role of *gdf5* in the regulation of the nervous system may be present. This finding is unsurprising given that previous research has identified the neurological functions of *GDF5*, notably its influence on the induction of dopaminergic neurons (Sullivan and O'Keefee, 2005), as well as its association with Parkinson's disease (O'Keeffee et al., 2016). However, as this effect was not observed in the founders or F1 generation, a cumulative effect of the mutation may lead to such change in later generations. It must be noted that these mutants were generated through incrossing. Therefore, observed changes in F2 mutants' behaviour may be a result of off-target effect due to the induced CRISPR editing, but it also might be the actual impact of the mutation as the reduction in *gdf5* expression in F2 mutants was confirmed using qPCR. In this series of experiments, controls were obtained from the same injected line, but they are not considered siblings to the tested larvae, as they do not come from the same parents.

GDF5 was found to exhibit an anabolic role in ECM osteogenesis regulation (Garciadigeo-Cazares et al., 2015). This study suggests a role of GDF5 in articular cartilage maintenance by inducing the expression of α5 integrin, an essential ECM protein. α5 integrin supports the attachment of other structural proteins to the ECM, enhancing the stability of cartilage and underlying bones. Based on microarray data conducted in our laboratory, a significant increase in *Gdf5* expression of 4.50 folds post-injury of murine hip cartilage was detected. This suggests a role of *Gdf5* in inflammation or tissue repair, or both. To investigate this hypothesis, a tailfin regeneration model was utilised in wildtype larvae and gdf5 f2 mutants. A consistent pattern of increase in regenerative abilities in gdf5 F2 was detected in two of the three conducted replicates. One replicate showed a particularly enhanced

regenerative abilities of control group, that was not observed in the other two replicates, which could indicate an outlier effect. Further repeats must be performed to confirm the possible enhancement in regenerative abilities of gdf5 mutants. Preferably using control siblings to confirm the observed difference is exclusively based on the induced mutation through generations. If found significant, this would indicate a potential role of Gdf5 as a cellular differentiation inhibitor or a pro-apoptotic factor. The detected increase in bone mineral density and bone surface area observed in micro-CT scans of *gdf5* crispants, also points out the possible role of *gdf5* in one or both of these cellular regulatory processes. Moreover, testing the tailfin injury model in the *Tg(mpx:GFP)* transgenic line to visualise neutrophil migration to injury site in green fluorescence, may reveal an additional functional role of *gdf5* in the injury-induced inflammatory response.

Previous OA studies examining the expression levels of *GDF5* in OA patients found reported contradicting expression levels. Mauck et al., 2006 initially showed a reduction in GDF5 mRNA and protein levels in cartilage samples of OA patients when compared to healthy cartilage. However, several recent studies report the opposite expression pattern. Kania and Colella., 2020 reported an increase in GDF5 levels in the synovium and articular cartilage of OA patients. They also showed the same pattern in mice post a DMM surgery. The same increase in expression was also observed in our microarray data following a murine hip injury. More recently, another study confirmed the increase in GDF5 associated with OA pathogenesis and also provided evidence to a correlation between *GDF5* expression and OA severity (Jinawath et al., 2022). In this case, excluding the initially reported reduction in GDF5 levels of OA patients may be tempting. However, considering this observation may help us understand the pathogenic basis of the gene in context to other bone homeostatic factors. As previously discussed, GDF5 shows high correlation to TGFβ1 as they signal simultaneously through TGF-β and BMP signalling pathways. This suggests that even in the presence of an impairment to one factor, the other pathway may compensate for the pathogenic effect though its regulation. As the impairment of any of the three studied genes could induce OA development or progression, it would be challenging to confidently pinpoint the actual cause of the disease. A possible route to be followed for further investigation would involve examining multiple mutations identified in GWAS in a single individual, rather than the single gene examination approach. Also, co-targeting of multiple genes within the same model organism may reveal correlations between the targeted genes.

7.4. *tgfβ1a* exhibits key roles in early skeletogenesis and heart development

TGFβ1 is a cytokine belonging to the TGF-β superfamily. It is abundantly expressed through early development starting from 10 weeks post conception and remains highly involved in adulthood (Graham et al., 1992). It plays a crucial role in multiple processes including cell proliferation,

differentiation, apoptosis and inflammation (Simon et al., 1995; Hannon et al., 1994; Saltis, 1996; Ramesh et al., 2009; Herrera-Molina and Bernhardi, 2005). As seen in GDF5, TGFβ1 also functions through a SMAD-dependent manner or SMAD-independent pathways. The gene was found to display a critical role in ECM regulation by modulating the expression of collagen (Chen et al., 2005) and proteoglycans (Okuda et al., 1990). Additionally, it was found to be involved in the regulation of inflammation as a pro-inflammatory mediator, as seen in Th17 cell activation (Veldhoen et al., 2006), and also as an anti-inflammatory regulator following brain injuries (Rustenhoven et al., 2016).

In the skeletal system, TGFβ1 was found to induce chondrogenesis, maintain chondrocyte proliferation, while inhibiting chondrocyte hypertrophy (Tuli et al., 2003). Not only does the protein function by activating TGF-β signalling, but it can also activate the BMP signalling pathway by binding to ALK1 protein (Davidson et al., 2009). Therefore, TGFβ1 protein can exhibit a range of phenotypes due to its involvement in the mentioned critical processes. To test the role of *tgfβ1* in the cartilage and heart development in zebrafish embryos, *tgfβ1a* crispants were generated. In zebrafish, two paralogues of *tgfβ1* are present, *tgfβ1a* and *tgfβ1b*. Each is located on a different chromosome, yet both show similar composition to each other. Tgfβ1a shows 45% protein sequence similarity to the human sequence, while Tgfβ1b show 44.5% similarity. In this study, *tgfβ1a* was the targeted paralogue due to availability of CRISPR target guides with higher expected efficiency. Yet, the selected CRISPR guides utilised in this analysis did not display a high targeting efficiency, leading to many of the injected and imaged larvae to maintain the wildtype peak. This suggests that *tgfβ1a* targeting was either unsuccessful or only partially achieved, and a significant amount of data had to be excluded as a result. Nonetheless, by examining the successfully targeted larvae, *tgfβ1a* crispants appear to develop in a similar manner to wildtype controls at 3dpf. However, the *Tg(Col2a:mCherry)* revealed mild jaw cartilage abnormalities such as non-uniform chondrocytes and bent ceratohyal cartilage. At 5dpf, progression in phenotype is observed and 29% of crispants show remarkable deterioration within the meckel's cartilage, and a complete absence of the basihyal cartilage in 12.5% of targeted larvae. The tolerance of *tgfβ1a* targeting by the larvae was unexpected due to the high involvement of the gene throughout early skeletogenesis. Yet as discussed in previous chapters, bone homeostasis is a complex process that involve various factors working simultaneously that even a mutation in such an essential player may be counteracted by other genes. It must also be noted that the presence of *tgfβ1b* paralogue may provide this compensation in the loss of *tgfβ1a*. To further investigate this hypothesis, generation of double mutants targeting both paralogues would reveal the mechanism behind the tolerable targeting effect.

When the effect of *tgfβ1a* targeting on heart development was examined, mild abnormalities were detected in the examined crispants, including irregularities within the chamber shape, along with bumps appearing on the heart wall and a disproportion in the size of the two chambers were visible, but not

quantified. Analysis of separate chamber volumes may reveal a statistically significant difference in atrial or ventricular size.

Regarding the functionality of the musculoskeletal system following the *tgfβ1a* targeting, no alterations in adult crispants activity levels or speed were detected. Adult cispants seem to display a similar swimming pattern to the control group, and no correlation to light conditions was detected.

To summarise, *tgfβ1a* targeting does not seem to have a severe detrimental impact on early skeletogenesis and heart development in zebrafish larvae. crispants show mild defects that are visible using tissue specific transgenic lines. The skeletal defects do not impact the swimming efficiency of crispants in adulthood. Unfortunately, micro-CT scans were not conducted to examine the bone structure in adults, but according to the observed progression of phenotype between 3 and 5dpf, alternation in adult bone structure would be expected, due to the high involvement of TGFβ1 in ECM regulation (Song et al., 2007) and bone mineralisation (Ehnert et al., 2010). Thus, micro-CT analysis could enable the detection of milder defects.

7.5. Co-targeting of *ctsk* and *tgfβ1a* reveals a possible direct crosstalk between the two genes

Based on the co-expression pattern of *CTSK* and *TGFβ1* detected through FUMA GWAS analysis, in addition to studies hinting a possible relation between *CTSK* and *TGFβ1* (Flanagen-Steet et al., 2018), co-targeted crispants for *ctsk* and *tgfβ1a* were generated in an attempt to reveal a possible crosstalk mechanism between the two encoded proteins. Results suggest a remarkable improvement in the overall morphology of the co-targeted larvae when compared to phenotype observed in single targeted larvae. A mere 4% of co-targeted crispants show a mild phenotype in general morphology of larvae at 5dpf, compared to 29% in *tgfβ1a* crispants and 47% of *ctsk* crispants. A complete eradication of the severe phenotype was observed. This suggests a possible rescue mechanism in the co-targeting of the two genes compared to their single targeting. However, the targeting had a different effect on heart morphology. At 3dpf, co-targeted crispants showed mild heart abnormities similar to *tgfβ1a* crispants, characterised by shape irregularities. When the larvae reached 5dpf, the phenotype seems to progress causing severe heart defects, similar to those observed in *ctsk* crispants, including complete heart inversion, shrinkage in both chambers, and disruptions in the looping process. Taken together, this highlight the differential effect of co-targeting on different tissues as the modulation resulted in an improvement in the general skeletal morphology, which appears to be rescued compared to *ctsk* crispants alone. Meanwhile, the heart initially showed milder deformities when compared to *ctsk* crispants but progressed to have more detrimental effects at 5dpf.

Larval behavioural analysis of co-targeted crispants displayed a significant increase in activity levels of larvae in both light conditions. The significance was higher in the presence of light compared to absence of light. Also, co-targeted crispants showed a high significant reduction in their hyperactive swimming, only in the presence of light, that was not detected in the absence of light. This suggests a coupled role of *ctsk* and *tgfβ1a* in the regulation of the nervous system, as their separate targeting did not display a significant change in their behaviour within the two light conditions, but their co-targeting did show a high significant change that was different in the two tested light conditions.

Following this series of experiments that suggested a direct link between *ctsk* and *tgfβ1a*, and in line with rising evidence to the correlation between the three studied genes that function simultaneously to attain homeostasis of different tissue, a subsequent effect of modulating one gene is predicted to affect the expression levels of the other genes. To investigate this hypothesis, a qPCR analysis was conducted to detect the expression levels of *ctsk* and *tgfβ1a* in gdf5 F2 mutants. In the two performed replicates, a consistent reduction in *ctsk* and *tgfβ1a* was detected in gdf5 F2 mutants when compared to wildtype controls, suggesting strong correlation between the studied genes. **Figure 6.18** illustrated pre-identified molecular mechanisms to provide insight to the possible link between the three studied genes via the SMAD-dependent signalling pathway. However, it must be noted that other SMAD-independent pathways were also found common between the two diseases and identified based on our bioinformatic analysis pipeline. These include the JAK/STAT, MAPK/JNK, PIK3/AKT/mTOR, and NF-kB. Yet, functional analysis is required to validate their involvement through experimental investigation.

Understanding the underlying mechanisms behind OA and CVD, and the interactions between studied factors, may allow us to identify potential drug targets or repurpose available drugs known to target one protein, in order to attenuate another protein that is correlated to our initial target. Metformin is a well-established type 2 diabetes drug. It was recently found to inhibit TGFβ1 by blocking its binding to TGF-β signalling receptor, suppressing the TGF-β signalling pathway (Xiao et al., 2016; Yang et al., 2021). The drug is known to be well tolerated with minimal side effects. Therefore, metformin may provide a great potential osteoarthritis treatment by avoiding direct CTSK inhibition, which has led to several disastrous side effects, and opting for indirect inhibition through TGFβ1. Although the drug is widely used in the treatment of type 2 diabetes and is considered to be a safe drug, the recently identified role in TGFβ1 suppression suggests possible unwanted off-target effects that may not yet have been identified but is definitely worth exploring.

7.6. Study limitations and future directions

Several limitations to the study must be considered. Firstly, the use of crispants provide a great tool for understanding the general functional role of the studied gene. However, the generation of stable mutant lines to track a specific effect of a mutation would provide more detailed information about the effect of the exact mutation. This is particularly important when investigating single nucleotide polymorphism (SNPs) identified in GWAS and are present within the coding region of the gene. In this case, specific CRISPR knock-in is the preferable method to mimic the studied SNP in zebrafish. Yet mapping the same SNPs on zebrafish genome could be challenging due to the presence of genes on different chromosomes with differences in target sequence. Additionally, generation of stable mutant lines allows us to limit the off-target effect of CRISPR genome editing by outcrossing the founders to wildtype fish. The efficiency of the induced indels must also be considered, as 6bp indels detected in gdf5 mutants, and 3bp /6bp deletions present in *tgfβ1a* crispants may lead to silent mutations that could still translate to a functional protein. Partial mutations in a percentage of the genome may also result in the presence of the functional protein, indicating that generated fish may present as heterozygous mutants. Therefore, for future investigation, I aim to generate stable a mutant line to each of the candidate genes, ideally using CRISPR knock-in to closely mimic the identified SNPs linked to OA and CVD. Each fish must be sequenced to confirm the exact mutation and link it to the observed phenotype.

Although zebrafish provides an exceptional model for studying early structural development through available fluorescent transgenic line, a method of visualising adult heart structure is currently unavailable and therefore limited our investigation to the early development of the heart. Additionally, the use of a system to track the functionality of the heart would have been beneficial to understand the effect of the induced mutation on the functionality of the system, as studied in the musculoskeletal system using swimming behaviour analysis.

This work was carried out during the COVID-19 pandemic, which limited the time available for further investigation, disrupted time-sensitive experiments and limited the resources available to extend the project. The gdf5 mutant line generated was severely affected by the lockdown, which led us to take a different crossing route. Investigation of the targeting effect was limited to early and late development in older adults due to lack of access to the aquarium during critical developmental stages. In addition, training plans that required travel were cancelled. This training would have enabled me to carry out analysis of skeletal functionality specific to the jaw area in larvae and adults. It would also have provided me with guidance on a specific image analysis protocol to obtain detailed measurements from collected data, which would have been possible with the help of colleagues at the University of Bristol.

Due to time constraints, several replicates were not performed, yet they would have added great value and allowed us to draw solid conclusions in several experiments. These include:

- In situ hybridization heart staining of *ctsk* targeted larvae at 1dpf, to determine the earliest involvement of *ctsk* in heart development.
- Performing micro-CT scans of more *ctsk* crispants to determine the differential sex-specific effect.
- Analysis of a larger number of heart images to include in the heart volumetric measurements.
- Repeats of *gdf5* targeting in the *Tg(Col2a:mCherry)* and *Tg(myl7:lifeActGFP)* transgenic lines. As well as an additional repeat to the regenerative abilities experiment of gdf5 F2 mutants.
- Co-targeting of both *tgfβ1a* paralogues, which may reveal the hidden reason behind the tolerability observed following *tgfβ1a* targeting.
- Repetition of *tgfβ1a:ctsk* co-targeted injections to confirm the recovery observed in the skeletal system. Conducting the injections in the *Tg(Col2a:mCherry*) transgenic line would be essential to examine the exact impact of the co-targeting on cartilage development.
- An additional repeat to the qPCR, testing the expression levels of *ctsk* and *tgfβ1a* in gdf5 F2 mutants.
- Performing qPCR analysis to detect the expression levels of *tgfβ1a* in ctsk mutants, and *ctsk* levels in tgfβ1a mutants. This would allow us to confirm the knockout effect of each gene on the other, aiming to confirm the suspected crosstalk.
- Conducting micro-CT scans on *tgfβ1a* adult crispants to identify the function of the gene in later stages of bone development.
- Generation of CRISPR-targeted lines for the identified common SNPs in the *dpep1* and *slc39a8* genes in zebrafish. This experiment may expand our understanding of the role of defective metal ion processing in the pathogenesis of multimorbidity.

8. References

Abifadel, M., Varret, M., Rabes, J. P., Allard, D., Ouguerram, K., Devillers, M., Cruaud, C., Benjannet, S., Wickham, L., Erlich, D., Derre, A., Villeger, L., Farnier, M., Beucler, I., Bruckert, E., Chambaz, J., Chanu, B., Lecerf, J. M., Luc, G., . . . Boileau, C. (2003). Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet*, *34*(2), 154-156. https://doi.org/10.1038/ng1161

Abisambra, J. F., Jinwal, U. K., Blair, L. J., O'Leary, J. C., 3rd, Li, Q., Brady, S., Wang, L., Guidi, C. E., Zhang, B., Nordhues, B. A., Cockman, M., Suntharalingham, A., Li, P., Jin, Y., Atkins, C. A., & Dickey, C. A. (2013). Tau accumulation activates the unfolded protein response by impairing endoplasmic reticulum-associated degradation. *J Neurosci*, *33*(22), 9498-9507. https://doi.org/10.1523/JNEUROSCI.5397-12.2013

Adler, A. S., Kawahara, T. L., Segal, E., & Chang, H. Y. (2008). Reversal of aging by NFkappaB blockade. *Cell Cycle*, *7*(5), 556-559. https://doi.org/10.4161/cc.7.5.5490

Aguilar-Pineda, J. A. et al. (2021) 'Vascular smooth muscle cell dysfunction contribute to neuroinflammation and Tau hyperphosphorylation in Alzheimer disease', iScience, 24(9). doi: 10.1016/J.ISCI.2021.102993.

Ait-Oufella, H., Horvat, B., Kerdiles, Y., Herbin, O., Gourdy, P., Khallou-Laschet, J., Merval, R., Esposito, B., Tedgui, A., & Mallat, Z. (2007). Measles virus nucleoprotein induces a regulatory immune response and reduces atherosclerosis in mice. *Circulation*, *116*(15), 1707-1713. https://doi.org/10.1161/CIRCULATIONAHA.107.699470

Alimena, G., Morra, E., Lazzarino, M., Liberati, A. M., Montefusco, E., Inverardi, D., Bernasconi, P., Mancini, M., Donti, E., Grignani, F., & et al. (1988). Interferon alpha-2b as therapy for Ph'-positive chronic myelogenous leukemia: a study of 82 patients treated with intermittent or daily administration. *Blood*, *72*(2), 642-647. https://www.ncbi.nlm.nih.gov/pubmed/3165298

Alnahdi, A. H., Zeni, J. A., & Snyder-Mackler, L. (2012). Muscle impairments in patients with knee osteoarthritis. *Sports Health*, *4*(4), 284-292. https://doi.org/10.1177/1941738112445726

Alswat, K. A. (2017). Gender Disparities in Osteoporosis. *J Clin Med Res*, *9*(5), 382-387. https://doi.org/10.14740/jocmr2970w

Alves, R. D., Eijken, M., van de Peppel, J., & van Leeuwen, J. P. (2014). Calcifying vascular smooth muscle cells and osteoblasts: independent cell types exhibiting extracellular matrix and

biomineralization-related mimicries. *BMC Genomics*, *15*(1), 965. https://doi.org/10.1186/1471-2164-15-965

Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nat Protoc*, *5*(9), 1564-1573. https://doi.org/10.1038/nprot.2010.116

Andrault, P. M., Panwar, P., Mackenzie, N. C. W., & Bromme, D. (2019). Elastolytic activity of cysteine cathepsins K, S, and V promotes vascular calcification. *Sci Rep*, *9*(1), 9682. https://doi.org/10.1038/s41598-019-45918-1

Aoyagi, T., & Matsui, T. (2011). Phosphoinositide-3 kinase signaling in cardiac hypertrophy and heart failure. *Curr Pharm Des*, *17*(18), 1818-1824. https://doi.org/10.2174/138161211796390976

Apschner, A., Schulte-Merker, S., & Witten, P. E. (2011). Not all bones are created equal - using zebrafish and other teleost species in osteogenesis research. *Methods Cell Biol*, *105*, 239-255. https://doi.org/10.1016/B978-0-12-381320-6.00010-2

Arroll, B., & Goodyear-Smith, F. (2004). Corticosteroid injections for osteoarthritis of the knee: meta-analysis. *BMJ*, *328*(7444), 869. https://doi.org/10.1136/bmj.38039.573970.7C

Asagiri, M., Hirai, T., Kunigami, T., Kamano, S., Gober, H. J., Okamoto, K., Nishikawa, K., Latz, E., Golenbock, D. T., Aoki, K., Ohya, K., Imai, Y., Morishita, Y., Miyazono, K., Kato, S., Saftig, P., & Takayanagi, H. (2008). Cathepsin K-dependent toll-like receptor 9 signaling revealed in experimental arthritis. *Science*, *319*(5863), 624-627. https://doi.org/10.1126/science.1150110

Atefi, M., von Euw, E., Attar, N., Ng, C., Chu, C., Guo, D., Nazarian, R., Chmielowski, B., Glaspy, J. A., Comin-Anduix, B., Mischel, P. S., Lo, R. S., & Ribas, A. (2011). Reversing melanoma cross-resistance to BRAF and MEK inhibitors by co-targeting the AKT/mTOR pathway. *PLoS One*, *6*(12), e28973. https://doi.org/10.1371/journal.pone.0028973

Aubourg, G., Rice, S. J., Bruce-Wootton, P., & Loughlin, J. (2022). Genetics of osteoarthritis. *Osteoarthritis Cartilage*, *30*(5), 636-649. https://doi.org/10.1016/j.joca.2021.03.002

Avdesh, A., Chen, M., Martin-Iverson, M. T., Mondal, A., Ong, D., Rainey-Smith, S., Taddei, K., Lardelli, M., Groth, D. M., Verdile, G., & Martins, R. N. (2012). Regular care and maintenance of a zebrafish (Danio rerio) laboratory: an introduction. *J Vis Exp*(69), e4196. https://doi.org/10.3791/4196

Azhar, M., Schultz Jel, J., Grupp, I., Dorn, G. W., 2nd, Meneton, P., Molin, D. G., Gittenberger-de Groot, A. C., & Doetschman, T. (2003). Transforming growth factor beta in cardiovascular

development and function. *Cytokine Growth Factor Rev*, *14*(5), 391-407. https://doi.org/10.1016/s1359-6101(03)00044-3

Azhar, M., Yin, M., Bommireddy, R., Duffy, J. J., Yang, J., Pawlowski, S. A., Boivin, G. P., Engle, S. J., Sanford, L. P., Grisham, C., Singh, R. R., Babcock, G. F., & Doetschman, T. (2009). Generation of mice with a conditional allele for transforming growth factor beta 1 gene. *Genesis*, *47*(6), 423-431. https://doi.org/10.1002/dvg.20516

Balkan, W., Martinez, A. F., Fernandez, I., Rodriguez, M. A., Pang, M., & Troen, B. R. (2009). Identification of NFAT binding sites that mediate stimulation of cathepsin K promoter activity by RANK ligand. *Gene*, *446*(2), 90-98. https://doi.org/10.1016/j.gene.2009.06.013

Barbour, K. E., Lui, L. Y., Nevitt, M. C., Murphy, L. B., Helmick, C. G., Theis, K. A., Hochberg, M. C., Lane, N. E., Hootman, J. M., Cauley, J. A., & Study of Osteoporotic Fractures Research, G. (2015). Hip Osteoarthritis and the Risk of All-Cause and Disease-Specific Mortality in Older Women: A Population-Based Cohort Study. *Arthritis Rheumatol*, *67*(7), 1798-1805. https://doi.org/10.1002/art.39113

Barnett, K. et al. (2012) "Epidemiology of multimorbidity and implications for health care, research, and Medical Education: A cross-sectional study," The Lancet, 380(9836), pp. 37–43. Available at: https://doi.org/10.1016/s0140-6736(12)60240-2.

Barnes, J. N. (2015). Exercise, cognitive function, and aging. *Adv Physiol Educ*, *39*(2), 55-62. https://doi.org/10.1152/advan.00101.2014

Begum, N. A., Kobayashi, M., Moriwaki, Y., Matsumoto, M., Toyoshima, K., & Seya, T. (2002). Mycobacterium bovis BCG cell wall and lipopolysaccharide induce a novel gene, BIGM103, encoding a 7-TM protein: identification of a new protein family having Zn-transporter and Zn-metalloprotease signatures. *Genomics*, *80*(6), 630-645. https://doi.org/10.1006/geno.2002.7000

Berenbaum, F., Eymard, F., & Houard, X. (2013). Osteoarthritis, inflammation and obesity. *Curr Opin Rheumatol*, *25*(1), 114-118. https://doi.org/10.1097/BOR.0b013e32835a9414

Bertrand-Chapel, A., Caligaris, C., Fenouil, T., Savary, C., Aires, S., Martel, S., Huchede, P., Chassot, C., Chauvet, V., Cardot-Ruffino, V., Morel, A. P., Subtil, F., Mohkam, K., Mabrut, J. Y., Tonon, L., Viari, A., Cassier, P., Hervieu, V., Castets, M., . . . Bartholin, L. (2022). SMAD2/3 mediate oncogenic effects of TGF-beta in the absence of SMAD4. *Commun Biol*, *5*(1), 1068. https://doi.org/10.1038/s42003-022-03994-6

Betrie, A. H. et al. (2017) ‘Evidence of a Cardiovascular Function for Microtubule-Associated Protein Tau’, Journal of Alzheimer’s disease : JAD, 56(2), pp. 849–860. doi: 10.3233/JAD-161093.

Bhattacharyya, S., Zhang, X., Feferman, L., Johnson, D., Tortella, F. C., Guizzetti, M., & Tobacman, J. K. (2015). Decline in arylsulfatase B and Increase in chondroitin 4-sulfotransferase combine to increase chondroitin 4-sulfate in traumatic brain injury. *J Neurochem*, *134*(4), 728-739. https://doi.org/10.1111/jnc.13156

Bird, N. C., & Mabee, P. M. (2003). Developmental morphology of the axial skeleton of the zebrafish, Danio rerio (Ostariophysi: Cyprinidae). *Dev Dyn*, *228*(3), 337-357. https://doi.org/10.1002/dvdy.10387

Bitzer, M., von Gersdorff, G., Liang, D., Dominguez-Rosales, A., Beg, A. A., Rojkind, M., & Bottinger, E. P. (2000). A mechanism of suppression of TGF-beta/SMAD signaling by NF-kappa B/RelA. *Genes Dev*, *14*(2), 187-197. https://www.ncbi.nlm.nih.gov/pubmed/10652273

Blaney Davidson, E. N., Remst, D. F., Vitters, E. L., van Beuningen, H. M., Blom, A. B., Goumans, M. J., van den Berg, W. B., & van der Kraan, P. M. (2009). Increase in ALK1/ALK5 ratio as a cause for elevated MMP-13 expression in osteoarthritis in humans and mice. *J Immunol*, *182*(12), 7937-7945. https://doi.org/10.4049/jimmunol.0803991

Blaney Davidson, E. N., Vitters, E. L., van Beuningen, H. M., van de Loo, F. A., van den Berg, W. B., & van der Kraan, P. M. (2007). Resemblance of osteophytes in experimental osteoarthritis to transforming growth factor beta-induced osteophytes: limited role of bone morphogenetic protein in early osteoarthritic osteophyte formation. *Arthritis Rheum*, *56*(12), 4065-4073. https://doi.org/10.1002/art.23034

Bloom, G. S. (2014). Amyloid-beta and tau: the trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurol*, *71*(4), 505-508. https://doi.org/10.1001/jamaneurol.2013.5847

Boer, C. G., Hatzikotoulas, K., Southam, L., Stefansdottir, L., Zhang, Y., Coutinho de Almeida, R., Wu, T. T., Zheng, J., Hartley, A., Teder-Laving, M., Skogholt, A. H., Terao, C., Zengini, E., Alexiadis, G., Barysenka, A., Bjornsdottir, G., Gabrielsen, M. E., Gilly, A., Ingvarsson, T., . . . Zeggini, E. (2021). Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell*, *184*(18), 4784-4818 e4717. https://doi.org/10.1016/j.cell.2021.07.038

Bohmer, F. D., & Friedrich, K. (2014). Protein tyrosine phosphatases as wardens of STAT signaling. *JAKSTAT*, *3*(1), e28087. https://doi.org/10.4161/jkst.28087

Bone, H. G., McClung, M. R., Roux, C., Recker, R. R., Eisman, J. A., Verbruggen, N., Hustad, C. M., DaSilva, C., Santora, A. C., & Ince, B. A. (2010). Odanacatib, a cathepsin-K inhibitor for osteoporosis: a two-year study in postmenopausal women with low bone density. *J Bone Miner Res*, *25*(5), 937-947. https://doi.org/10.1359/jbmr.091035

Bonnet, N., Brun, J., Rousseau, J. C., Duong, L. T., & Ferrari, S. L. (2017). Cathepsin K Controls Cortical Bone Formation by Degrading Periostin. *J Bone Miner Res*, *32*(7), 1432-1441. https://doi.org/10.1002/jbmr.3136

Brunt, L. H., Begg, K., Kague, E., Cross, S., & Hammond, C. L. (2017). Wnt signalling controls the response to mechanical loading during zebrafish joint development. *Development*, *144*(15), 2798-2809. https://doi.org/10.1242/dev.153528

Buck, M. R., Karustis, D. G., Day, N. A., Honn, K. V., & Sloane, B. F. (1992). Degradation of extracellular-matrix proteins by human cathepsin B from normal and tumour tissues. *Biochem J*, *282 (Pt 1)*(Pt 1), 273-278. https://doi.org/10.1042/bj2820273

Buckwalter, J. A., Saltzman, C., & Brown, T. (2004). The impact of osteoarthritis: implications for research. *Clin Orthop Relat Res*(427 Suppl), S6-15. https://doi.org/10.1097/01.blo.0000143938.30681.9d

Bueno, O. F., De Windt, L. J., Tymitz, K. M., Witt, S. A., Kimball, T. R., Klevitsky, R., Hewett, T. E., Jones, S. P., Lefer, D. J., Peng, C. F., Kitsis, R. N., & Molkentin, J. D. (2000). The MEK1-ERK1/2 signaling pathway promotes compensated cardiac hypertrophy in transgenic mice. *EMBO J*, *19*(23), 6341-6350. https://doi.org/10.1093/emboj/19.23.6341

Buhling, F., Waldburg, N., Reisenauer, A., Heimburg, A., Golpon, H., & Welte, T. (2004). Lysosomal cysteine proteases in the lung: role in protein processing and immunoregulation. *Eur Respir J*, *23*(4), 620-628. https://doi.org/10.1183/09031936.04.00105304

Buxton, P., Edwards, C., Archer, C. W., & Francis-West, P. (2001). Growth/differentiation factor-5 (GDF-5) and skeletal development. *J Bone Joint Surg Am*, *83-A Suppl 1*(Pt 1), S23-30. https://www.ncbi.nlm.nih.gov/pubmed/11263662

Cao, X. and Chen, Y. (2009) “Mitochondria and calcium signaling in embryonic development,” Seminars in Cell & Developmental Biology, 20(3), pp. 337–345.

Cerami, E. et al. (2012) “The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.” Cancer discovery, 2(5), pp.401-404.

Chan, C.P. et al. (2005) “Effects of TGF-βs on the growth, collagen synthesis and collagen lattice contraction of human dental pulp fibroblasts in vitro,” Archives of Oral Biology, 50(5), pp. 469–479.

Chang, A. S., Hathaway, C. K., Smithies, O., & Kakoki, M. (2016). Transforming growth factor-beta1 and diabetic nephropathy. *Am J Physiol Renal Physiol*, *310*(8), F689-F696. https://doi.org/10.1152/ajprenal.00502.2015

Chen, H. B., Shen, J., Ip, Y. T., & Xu, L. (2006). Identification of phosphatases for Smad in the BMP/DPP pathway. *Genes Dev*, *20*(6), 648-653. https://doi.org/10.1101/gad.1384706

Chen, J., Crawford, R., & Xiao, Y. (2013). Vertical inhibition of the PI3K/Akt/mTOR pathway for the treatment of osteoarthritis. *J Cell Biochem*, *114*(2), 245-249. https://doi.org/10.1002/jcb.24362

Chen, X., Chen, W., Aung, Z. M., Han, W., Zhang, Y., & Chai, G. (2021). LY3023414 inhibits both osteogenesis and osteoclastogenesis through the PI3K/Akt/GSK3 signalling pathway. *Bone Joint Res*, *10*(4), 237-249. https://doi.org/10.1302/2046-3758.104.BJR-2020-0255.R2

Cheng, C., Tempel, D., van Haperen, R., van der Baan, A., Grosveld, F., Daemen, M. J., Krams, R., & de Crom, R. (2006). Atherosclerotic lesion size and vulnerability are determined by patterns of fluid shear stress. *Circulation*, *113*(23), 2744-2753. https://doi.org/10.1161/CIRCULATIONAHA.105.590018

Christensen, J., & Shastri, V. P. (2015). Matrix-metalloproteinase-9 is cleaved and activated by cathepsin K. *BMC Res Notes*, *8*, 322. https://doi.org/10.1186/s13104-015-1284-8

Cleveland, D. W., Hwo, S. Y., & Kirschner, M. W. (1977). Purification of tau, a microtubule-associated protein that induces assembly of microtubules from purified tubulin. *J Mol Biol*, *116*(2), 207-225. https://doi.org/10.1016/0022-2836(77)90213-3

Colangelo, V., Schurr, J., Ball, M. J., Pelaez, R. P., Bazan, N. G., & Lukiw, W. J. (2002). Gene expression profiling of 12633 genes in Alzheimer hippocampal CA1: transcription and neurotrophic factor down-regulation and up-regulation of apoptotic and pro-inflammatory signaling. *J Neurosci Res*, *70*(3), 462-473. https://doi.org/10.1002/jnr.10351

Coleman, C. M., Vaughan, E. E., Browe, D. C., Mooney, E., Howard, L., & Barry, F. (2013). Growth differentiation factor-5 enhances in vitro mesenchymal stromal cell chondrogenesis and hypertrophy. *Stem Cells Dev*, *22*(13), 1968-1976. https://doi.org/10.1089/scd.2012.0282

Conti, D. V., Darst, B. F., Moss, L. C., Saunders, E. J., Sheng, X., Chou, A., Schumacher, F. R., Olama, A. A. A., Benlloch, S., Dadaev, T., Brook, M. N., Sahimi, A., Hoffmann, T. J., Takahashi, A., Matsuda,

K., Momozawa, Y., Fujita, M., Muir, K., Lophatananon, A., . . . Haiman, C. A. (2021). Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. *Nat Genet*, *53*(1), 65-75. https://doi.org/10.1038/s41588-020-00748-0

Costa, A. G., Cusano, N. E., Silva, B. C., Cremers, S., & Bilezikian, J. P. (2011). Cathepsin K: its skeletal actions and role as a therapeutic target in osteoporosis. *Nat Rev Rheumatol*, *7*(8), 447-456. https://doi.org/10.1038/nrrheum.2011.77

Collier, S. and Ghosh, P. (1995) "Effects of transforming growth factor beta on proteoglycan synthesis by cell and explant cultures derived from the knee joint meniscus," Osteoarthritis and Cartilage, 3(2), pp. 127–138.

Comstock, K. L. et al. (1998) 'LPS-Induced TNF-αRelease from and Apoptosis in Rat Cardiomyocytes: Obligatory Role for CD14 in Mediating the LPS Response', Journal of Molecular and Cellular Cardiology, 30(12), pp. 2761–2775. doi: 10.1006/JMCC.1998.0851.

Cui, X., Liu, X., Han, Q., Zhu, J., Li, J., Ren, Z., Liu, L., Luo, Y., Wang, Z., Zhang, D., Fan, Y., Zhang, D., & Dong, G. (2019). DPEP1 is a direct target of miR-193a-5p and promotes hepatoblastoma progression by PI3K/Akt/mTOR pathway. *Cell Death Dis*, *10*(10), 701. https://doi.org/10.1038/s41419-019-1943-0

da Costa, B. R., Reichenbach, S., Keller, N., Nartey, L., Wandel, S., Juni, P., & Trelle, S. (2017). Effectiveness of non-steroidal anti-inflammatory drugs for the treatment of pain in knee and hip osteoarthritis: a network meta-analysis. *Lancet*, *390*(10090), e21-e33. https://doi.org/10.1016/S0140-6736(17)31744-0

da Rocha, A. L., Pinto, A. P., Kohama, E. B., Pauli, J. R., de Moura, L. P., Cintra, D. E., Ropelle, E. R., & da Silva, A. S. R. (2019). The proinflammatory effects of chronic excessive exercise. *Cytokine*, *119*, 57-61. https://doi.org/10.1016/j.cyto.2019.02.016

Dalton, S. O., Johansen, C., Mellemkjaer, L., Norgard, B., Sorensen, H. T., & Olsen, J. H. (2003). Use of selective serotonin reuptake inhibitors and risk of upper gastrointestinal tract bleeding: a population-based cohort study. *Arch Intern Med*, *163*(1), 59-64. https://doi.org/10.1001/archinte.163.1.59

Dauth, S., Rakov, H., Sirbulescu, R. F., Ilies, I., Weber, J., Batbajar Dugershaw, B., Braun, D., Rehders, M., Wirth, E. K., Fuhrer, D., Schweizer, U., & Brix, K. (2020). Function of Cathepsin K in the Central Nervous System of Male Mice is Independent of Its Role in the Thyroid Gland. *Cell Mol Neurobiol*, *40*(5), 695-710. https://doi.org/10.1007/s10571-019-00765-6

de Bakker, P. I., Yelensky, R., Pe'er, I., Gabriel, S. B., Daly, M. J., & Altshuler, D. (2005). Efficiency and power in genetic association studies. *Nat Genet*, *37*(11), 1217-1223. https://doi.org/10.1038/ng1669

de Lange, K. M., Moutsianas, L., Lee, J. C., Lamb, C. A., Luo, Y., Kennedy, N. A., Jostins, L., Rice, D. L., Gutierrez-Achury, J., Ji, S. G., Heap, G., Nimmo, E. R., Edwards, C., Henderson, P., Mowat, C., Sanderson, J., Satsangi, J., Simmons, A., Wilson, D. C., . . . Barrett, J. C. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet*, *49*(2), 256-261. https://doi.org/10.1038/ng.3760

DeGroot, J., Verzijl, N., Bank, R. A., Lafeber, F. P., Bijlsma, J. W., & TeKoppele, J. M. (1999). Age-related decrease in proteoglycan synthesis of human articular chondrocytes: the role of nonenzymatic glycation. *Arthritis Rheum*, *42*(5), 1003-1009. https://doi.org/10.1002/1529-0131(199905)42:5<1003::AID-ANR20>3.0.CO;2-K

DeLaurier, A., Eames, B. F., Blanco-Sanchez, B., Peng, G., He, X., Swartz, M. E., Ullmann, B., Westerfield, M., & Kimmel, C. B. (2010). Zebrafish sp7:EGFP: a transgenic for studying otic vesicle formation, skeletogenesis, and bone regeneration. *Genesis*, *48*(8), 505-511. https://doi.org/10.1002/dvg.20639

Derynck, R., & Zhang, Y. E. (2003). Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature*, *425*(6958), 577-584. https://doi.org/10.1038/nature02006

Diabetes Genetics Initiative of Broad Institute of, H., Mit, L. U., Novartis Institutes of BioMedical, R., Saxena, R., Voight, B. F., Lyssenko, V., Burtt, N. P., de Bakker, P. I., Chen, H., Roix, J. J., Kathiresan, S., Hirschhorn, J. N., Daly, M. J., Hughes, T. E., Groop, L., Altshuler, D., Almgren, P., Florez, J. C., Meyer, J., . . . Purcell, S. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*, *316*(5829), 1331-1336. https://doi.org/10.1126/science.1142358

Dillon, C. F., Rasch, E. K., Gu, Q., & Hirsch, R. (2006). Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94. *J Rheumatol*, *33*(11), 2271-2279. https://www.ncbi.nlm.nih.gov/pubmed/17013996

Ding, C., Martel-Pelletier, J., Pelletier, J. P., Abram, F., Raynauld, J. P., Cicuttini, F., & Jones, G. (2007). Knee meniscal extrusion in a largely non-osteoarthritic cohort: association with greater loss of cartilage volume. *Arthritis Res Ther*, *9*(2), R21. https://doi.org/10.1186/ar2132

Diseases, G. B. D., & Injuries, C. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*, *396*(10258), 1204-1222. https://doi.org/10.1016/S0140-6736(20)30925-9

Dole, N. S., Mazur, C. M., Acevedo, C., Lopez, J. P., Monteiro, D. A., Fowler, T. W., Gludovatz, B., Walsh, F., Regan, J. N., Messina, S., Evans, D. S., Lang, T. F., Zhang, B., Ritchie, R. O., Mohammad, K. S., & Alliston, T. (2017). Osteocyte-Intrinsic TGF-beta Signaling Regulates Bone Quality through Perilacunar/Canalicular Remodeling. *Cell Rep*, *21*(9), 2585-2596. https://doi.org/10.1016/j.celrep.2017.10.115

Dowdy, S. C., Mariani, A., Reinholz, M. M., Keeney, G. L., Spelsberg, T. C., Podratz, K. C., & Janknecht, R. (2005). Overexpression of the TGF-beta antagonist Smad7 in endometrial cancer. *Gynecol Oncol*, *96*(2), 368-373. https://doi.org/10.1016/j.ygyno.2004.10.006

Drake, F. H., Dodds, R. A., James, I. E., Connor, J. R., Debouck, C., Richardson, S., Lee-Rykaczewski, E., Coleman, L., Rieman, D., Barthlow, R., Hastings, G., & Gowen, M. (1996). Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. *J Biol Chem*, *271*(21), 12511-12516. https://doi.org/10.1074/jbc.271.21.12511

Dubland, J. A., & Francis, G. A. (2015). Lysosomal acid lipase: at the crossroads of normal and atherogenic cholesterol metabolism. *Front Cell Dev Biol*, *3*, 3. https://doi.org/10.3389/fcell.2015.00003

Dubois, C. M., Laprise, M. H., Blanchette, F., Gentry, L. E., & Leduc, R. (1995). Processing of transforming growth factor beta 1 precursor by human furin convertase. *J Biol Chem*, *270*(18), 10618-10624. https://doi.org/10.1074/jbc.270.18.10618

Ducy, P., & Karsenty, G. (2000). The family of bone morphogenetic proteins. *Kidney Int*, *57*(6), 2207-2214. https://doi.org/10.1046/j.1523-1755.2000.00081.x

Duncan, L., Yilmaz, Z., Gaspar, H., Walters, R., Goldstein, J., Anttila, V., Bulik-Sullivan, B., Ripke, S., Eating Disorders Working Group of the Psychiatric Genomics, C., Thornton, L., Hinney, A., Daly, M., Sullivan, P. F., Zeggini, E., Breen, G., & Bulik, C. M. (2017). Significant Locus and Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa. *Am J Psychiatry*, *174*(9), 850-858. https://doi.org/10.1176/appi.ajp.2017.16121402

Dunker, N., Schmitt, K., & Krieglstein, K. (2002). TGF-beta is required for programmed cell death in interdigital webs of the developing mouse limb. *Mech Dev*, *113*(2), 111-120. https://doi.org/10.1016/s0925-4773(02)00015-1

Ehnert, S., Baur, J., Schmitt, A., Neumaier, M., Lucke, M., Dooley, S., Vester, H., Wildemann, B., Stockle, U., & Nussler, A. K. (2010). TGF-beta1 as possible link between loss of bone mineral density and chronic inflammation. *PLoS One*, *5*(11), e14073. https://doi.org/10.1371/journal.pone.0014073

Eisenach, P. A., Soeth, E., Roder, C., Kloppel, G., Tepel, J., Kalthoff, H., & Sipos, B. (2013). Dipeptidase 1 (DPEP1) is a marker for the transition from low-grade to high-grade intraepithelial neoplasia and an adverse prognostic factor in colorectal cancer. *Br J Cancer*, *109*(3), 694-703. https://doi.org/10.1038/bjc.2013.363

Eisenberg, C. A., & Eisenberg, L. M. (1999). WNT11 promotes cardiac tissue formation of early mesoderm. *Dev Dyn*, *216*(1), 45-58. https://doi.org/10.1002/(SICI)1097-0177(199909)216:1<45::AID-DVDY7>3.0.CO;2-L

Eisman, J. A., Bone, H. G., Hosking, D. J., McClung, M. R., Reid, I. R., Rizzoli, R., Resch, H., Verbruggen, N., Hustad, C. M., DaSilva, C., Petrovic, R., Santora, A. C., Ince, B. A., & Lombardi, A. (2011). Odanacatib in the treatment of postmenopausal women with low bone mineral density: three-year continued therapy and resolution of effect. *J Bone Miner Res*, *26*(2), 242-251. https://doi.org/10.1002/jbmr.212

Ellman, M. B., Yan, D., Ahmadinia, K., Chen, D., An, H. S., & Im, H. J. (2013). Fibroblast growth factor control of cartilage homeostasis. *J Cell Biochem*, *114*(4), 735-742. https://doi.org/10.1002/jcb.24418

Everts, V., van der Zee, E., Creemers, L., & Beertsen, W. (1996). Phagocytosis and intracellular digestion of collagen, its role in turnover and remodelling. *Histochem J*, *28*(4), 229-245. https://doi.org/10.1007/BF02409011

Faiz, A., Tjin, G., Harkness, L., Weckmann, M., Bao, S., Black, J. L., Oliver, B. G., & Burgess, J. K. (2013). The expression and activity of cathepsins D, H and K in asthmatic airways. *PLoS One*, *8*(3), e57245. https://doi.org/10.1371/journal.pone.0057245

Fang, W., He, A., Xiang, M. X., Lin, Y., Wang, Y., Li, J., Yang, C., Zhang, X., Liu, C. L., Sukhova, G. K., Barascuk, N., Larsen, L., Karsdal, M., Libby, P., & Shi, G. P. (2019). Cathepsin K-deficiency impairs mouse cardiac function after myocardial infarction. *J Mol Cell Cardiol*, *127*, 44-56. https://doi.org/10.1016/j.yjmcc.2018.11.010

Felson, D. T., Gale, D. R., Elon Gale, M., Niu, J., Hunter, D. J., Goggins, J., & Lavalley, M. P. (2005). Osteophytes and progression of knee osteoarthritis. *Rheumatology (Oxford)*, *44*(1), 100-104. https://doi.org/10.1093/rheumatology/keh411

Fernandez, I. E., & Eickelberg, O. (2012). The impact of TGF-beta on lung fibrosis: from targeting to biomarkers. *Proc Am Thorac Soc*, *9*(3), 111-116. https://doi.org/10.1513/pats.201203-023AW

Fernandez-Figueroa, E. A., Imaz-Rosshandler, I., Castillo-Fernandez, J. E., Miranda-Ortiz, H., Fernandez-Lopez, J. C., Becker, I., & Rangel-Escareno, C. (2016). Down-Regulation of TLR and JAK/STAT Pathway Genes Is Associated with Diffuse Cutaneous Leishmaniasis: A Gene Expression Analysis in NK Cells from Patients Infected with Leishmania mexicana. *PLoS Negl Trop Dis*, *10*(3), e0004570. https://doi.org/10.1371/journal.pntd.0004570

Fernandez-Moreno, M., Rego, I., Carreira-Garcia, V., & Blanco, F. J. (2008). Genetics in osteoarthritis. *Curr Genomics*, *9*(8), 542-547. https://doi.org/10.2174/138920208786847953

Flanagan-Steet, H., Christian, C., Lu, P. N., Aarnio-Peterson, M., Sanman, L., Archer-Hartmann, S., Azadi, P., Bogyo, M., & Steet, R. A. (2018). TGF-ss Regulates Cathepsin Activation during Normal and Pathogenic Development. *Cell Rep*, *22*(11), 2964-2977. https://doi.org/10.1016/j.celrep.2018.02.066

Flanders, K. C., Ludecke, G., Engels, S., Cissel, D. S., Roberts, A. B., Kondaiah, P., Lafyatis, R., Sporn, M. B., & Unsicker, K. (1991). Localization and actions of transforming growth factor-beta s in the embryonic nervous system. *Development*, *113*(1), 183-191. https://doi.org/10.1242/dev.113.1.183

Francis-West, P. H., Abdelfattah, A., Chen, P., Allen, C., Parish, J., Ladher, R., Allen, S., MacPherson, S., Luyten, F. P., & Archer, C. W. (1999). Mechanisms of GDF-5 action during skeletal development. *Development*, *126*(6), 1305-1315. https://doi.org/10.1242/dev.126.6.1305

Freudlsperger, C., Bian, Y., Contag Wise, S., Burnett, J., Coupar, J., Yang, X., Chen, Z., & Van Waes, C. (2013). TGF-beta and NF-kappaB signal pathway cross-talk is mediated through TAK1 and SMAD7 in a subset of head and neck cancers. *Oncogene*, *32*(12), 1549-1559. https://doi.org/10.1038/onc.2012.171

Fukada, T., Civic, N., Furuichi, T., Shimoda, S., Mishima, K., Higashiyama, H., Idaira, Y., Asada, Y., Kitamura, H., Yamasaki, S., Hojyo, S., Nakayama, M., Ohara, O., Koseki, H., Dos Santos, H. G., Bonafe, L., Ha-Vinh, R., Zankl, A., Unger, S., . . . Hirano, T. (2008). The zinc transporter SLC39A13/ZIP13 is required for connective tissue development; its involvement in BMP/TGF-beta signaling pathways. *PLoS One*, *3*(11), e3642. https://doi.org/10.1371/journal.pone.0003642

Garciadiego-Cazares, D., Aguirre-Sanchez, H. I., Abarca-Buis, R. F., Kouri, J. B., Velasquillo, C., & Ibarra, C. (2015). Regulation of alpha5 and alphaV Integrin Expression by GDF-5 and BMP-7 in Chondrocyte Differentiation and Osteoarthritis. *PLoS One*, *10*(5), e0127166. https://doi.org/10.1371/journal.pone.0127166

Garg, V., Muth, A. N., Ransom, J. F., Schluterman, M. K., Barnes, R., King, I. N., Grossfeld, P. D., & Srivastava, D. (2005). Mutations in NOTCH1 cause aortic valve disease. *Nature*, *437*(7056), 270-274. https://doi.org/10.1038/nature03940

Garnero, P., Borel, O., Byrjalsen, I., Ferreras, M., Drake, F. H., McQueney, M. S., Foged, N. T., Delmas, P. D., & Delaisse, J. M. (1998). The collagenolytic activity of cathepsin K is unique among mammalian proteinases. *J Biol Chem*, *273*(48), 32347-32352. https://doi.org/10.1074/jbc.273.48.32347

Ge, H. X., Zou, F. M., Li, Y., Liu, A. M., & Tu, M. (2017). JNK pathway in osteoarthritis: pathological and therapeutic aspects. *J Recept Signal Transduct Res*, *37*(5), 431-436. https://doi.org/10.1080/10799893.2017.1360353

Gelb, B. D., Shi, G. P., Chapman, H. A., & Desnick, R. J. (1996). Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. *Science*, *273*(5279), 1236-1238. https://doi.org/10.1126/science.273.5279.1236

Gerosa, M., De Angelis, V., Riboldi, P., & Meroni, P. L. (2008). Rheumatoid arthritis: a female challenge. *Womens Health (Lond)*, *4*(2), 195-201. https://doi.org/10.2217/17455057.4.2.195

Gialeli, C., Shami, A., & Goncalves, I. (2021). Extracellular matrix: paving the way to the newest trends in atherosclerosis. *Curr Opin Lipidol*, *32*(5), 277-285. https://doi.org/10.1097/MOL.0000000000000775

Godoi, M. A., Camilli, A. C., Gonzales, K. G. A., Costa, V. B., Papathanasiou, E., Leite, F. R. M., & Guimaraes-Stabili, M. R. (2023). JAK/STAT as a Potential Therapeutic Target for Osteolytic Diseases. *Int J Mol Sci*, *24*(12). https://doi.org/10.3390/ijms241210290

Goedert, M., Spillantini, M. G., Jakes, R., Rutherford, D., & Crowther, R. A. (1989). Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron*, *3*(4), 519-526. https://doi.org/10.1016/0896-6273(89)90210-9

Goldring, M. B., & Marcu, K. B. (2009). Cartilage homeostasis in health and rheumatic diseases. *Arthritis Res Ther*, *11*(3), 224. https://doi.org/10.1186/ar2592

Gomez-Bernal, F., Quevedo-Abeledo, J. C., Garcia-Gonzalez, M., Fernandez-Cladera, Y., Gonzalez-Rivero, A. F., Martin-Gonzalez, C., Gonzalez-Gay, M. A., & Ferraz-Amaro, I. (2023). Transforming growth factor beta 1 is associated with subclinical carotid atherosclerosis in patients with systemic lupus erythematosus. *Arthritis Res Ther*, *25*(1), 64. https://doi.org/10.1186/s13075-023-03046-2

Gowen, M., Lazner, F., Dodds, R., Kapadia, R., Feild, J., Tavaria, M., Bertoncello, I., Drake, F., Zavarselk, S., Tellis, I., Hertzog, P., Debouck, C., & Kola, I. (1999). Cathepsin K knockout mice develop osteopetrosis due to a deficit in matrix degradation but not demineralization. *J Bone Miner Res*, *14*(10), 1654-1663. https://doi.org/10.1359/jbmr.1999.14.10.1654

Graham, C. H., Lysiak, J. J., McCrae, K. R., & Lala, P. K. (1992). Localization of transforming growth factor-beta at the human fetal-maternal interface: role in trophoblast growth and differentiation. *Biol Reprod*, *46*(4), 561-572. https://doi.org/10.1095/biolreprod46.4.561

Gray, A. M., & Mason, A. J. (1990). Requirement for activin A and transforming growth factor--beta 1 pro-regions in homodimer assembly. *Science*, *247*(4948), 1328-1330. https://doi.org/10.1126/science.2315700

Greten, F. R., Arkan, M. C., Bollrath, J., Hsu, L. C., Goode, J., Miething, C., Goktuna, S. I., Neuenhahn, M., Fierer, J., Paxian, S., Van Rooijen, N., Xu, Y., O'Cain, T., Jaffee, B. B., Busch, D. H., Duyster, J., Schmid, R. M., Eckmann, L., & Karin, M. (2007). NF-kappaB is a negative regulator of IL-1beta secretion as revealed by genetic and pharmacological inhibition of IKKbeta. *Cell*, *130*(5), 918-931. https://doi.org/10.1016/j.cell.2007.07.009

Guan, Y., Liang, X., Ma, Z., Hu, H., Liu, H., Miao, Z., Linkermann, A., Hellwege, J. N., Voight, B. F., & Susztak, K. (2021). A single genetic locus controls both expression of DPEP1/CHMP1A and kidney disease development via ferroptosis. *Nat Commun*, *12*(1), 5078. https://doi.org/10.1038/s41467-021-25377-x

Guo, R., Hua, Y., Ren, J., Bornfeldt, K. E., & Nair, S. (2018). Cardiomyocyte-specific disruption of Cathepsin K protects against doxorubicin-induced cardiotoxicity. *Cell Death Dis*, *9*(6), 692. https://doi.org/10.1038/s41419-018-0727-2

Guttridge, D. C., Albanese, C., Reuther, J. Y., Pestell, R. G., & Baldwin, A. S., Jr. (1999). NF-kappaB controls cell growth and differentiation through transcriptional regulation of cyclin D1. *Mol Cell Biol*, *19*(8), 5785-5799. https://doi.org/10.1128/MCB.19.8.5785

Hammond, C. L., & Moro, E. (2012). Using transgenic reporters to visualize bone and cartilage signaling during development in vivo. *Front Endocrinol (Lausanne)*, *3*, 91. https://doi.org/10.3389/fendo.2012.00091

Hampe, J., Franke, A., Rosenstiel, P., Till, A., Teuber, M., Huse, K., Albrecht, M., Mayr, G., De La Vega, F. M., Briggs, J., Gunther, S., Prescott, N. J., Onnie, C. M., Hasler, R., Sipos, B., Folsch, U. R., Lengauer, T., Platzer, M., Mathew, C. G., . . . Schreiber, S. (2007). A genome-wide association scan of

nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet*, *39*(2), 207-211. https://doi.org/10.1038/ng1954

Hanlon, P., Quinn, T. J., Gallacher, K. I., Myint, P. K., Jani, B. D., Nicholl, B. I., Lowrie, R., Soiza, R. L., Neal, S. R., Lee, D., & Mair, F. S. (2020). Assessing Risks of Polypharmacy Involving Medications With Anticholinergic Properties. *Ann Fam Med*, *18*(2), 148-155. https://doi.org/10.1370/afm.2501

Hannon, G. J., & Beach, D. (1994). p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature*, *371*(6494), 257-261. https://doi.org/10.1038/371257a0

Harvey, R. J., Skelton-Robinson, M., & Rossor, M. N. (2003). The prevalence and causes of dementia in people under the age of 65 years. *J Neurol Neurosurg Psychiatry*, *74*(9), 1206-1209. https://doi.org/10.1136/jnnp.74.9.1206

Hatakeyama, Y., Tuan, R. S., & Shum, L. (2004). Distinct functions of BMP4 and GDF5 in the regulation of chondrogenesis. *J Cell Biochem*, *91*(6), 1204-1217. https://doi.org/10.1002/jcb.20019

Hawker, G. A., Croxford, R., Bierman, A. S., Harvey, P. J., Ravi, B., Stanaitis, I., & Lipscombe, L. L. (2014). All-cause mortality and serious cardiovascular events in people with hip and knee osteoarthritis: a population based cohort study. *PLoS One*, *9*(3), e91286. https://doi.org/10.1371/journal.pone.0091286

Heidari, B. (2011). Knee osteoarthritis prevalence, risk factors, pathogenesis and features: Part I. *Caspian J Intern Med*, *2*(2), 205-212. https://www.ncbi.nlm.nih.gov/pubmed/24024017

Henrotin, Y., Mathy, M., Sanchez, C., & Lambert, C. (2010). Chondroitin sulfate in the treatment of osteoarthritis: from in vitro studies to clinical recommendations. *Ther Adv Musculoskelet Dis*, *2*(6), 335-348. https://doi.org/10.1177/1759720X10383076

Heo, S. C., Kim, Y. N., Choi, Y., Joo, J. Y., Hwang, J. J., Bae, M. K., & Kim, H. J. (2021). Elevated Expression of Cathepsin K in Periodontal Ligament Fibroblast by Inflammatory Cytokines Accelerates Osteoclastogenesis via Paracrine Mechanism in Periodontal Disease. *Int J Mol Sci*, *22*(2). https://doi.org/10.3390/ijms22020695

Hermann, D. M., Gronewold, J., Lehmann, N., Moebus, S., Jockel, K. H., Bauer, M., Erbel, R., & Heinz Nixdorf Recall Study Investigative, G. (2013). Coronary artery calcification is an independent stroke predictor in the general population. *Stroke*, *44*(4), 1008-1013. https://doi.org/10.1161/STROKEAHA.111.678078

Herrera-Molina, R., & von Bernhardi, R. (2005). Transforming growth factor-beta 1 produced by hippocampal cells modulates microglial reactivity in culture. *Neurobiol Dis*, *19*(1-2), 229-236. https://doi.org/10.1016/j.nbd.2005.01.003

Hinoi, E., Iezaki, T., Fujita, H., Watanabe, T., Odaka, Y., Ozaki, K., & Yoneda, Y. (2014). PI3K/Akt is involved in brown adipogenesis mediated by growth differentiation factor-5 in association with activation of the Smad pathway. *Biochem Biophys Res Commun*, *450*(1), 255-260. https://doi.org/10.1016/j.bbrc.2014.05.108

Hirschhorn, J. N. (2009). Genomewide association studies--illuminating biologic pathways. *N Engl J Med*, *360*(17), 1699-1701. https://doi.org/10.1056/NEJMp0808934

Horiguchi, K., Sakamoto, K., Koinuma, D., Semba, K., Inoue, A., Inoue, S., Fujii, H., Yamaguchi, A., Miyazawa, K., Miyazono, K., & Saitoh, M. (2012). TGF-beta drives epithelial-mesenchymal transition through deltaEF1-mediated downregulation of ESRP. *Oncogene*, *31*(26), 3190-3201. https://doi.org/10.1038/onc.2011.493

Hotten, G. C., Matsumoto, T., Kimura, M., Bechtold, R. F., Kron, R., Ohara, T., Tanaka, H., Satoh, Y., Okazaki, M., Shirai, T., Pan, H., Kawai, S., Pohl, J. S., & Kudo, A. (1996). Recombinant human growth/differentiation factor 5 stimulates mesenchyme aggregation and chondrogenesis responsible for the skeletal development of limbs. *Growth Factors*, *13*(1-2), 65-74. https://doi.org/10.3109/08977199609034567

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., Collins, J. E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J. C., Koch, R., Rauch, G. J., White, S., . . . Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, *496*(7446), 498-503. https://doi.org/10.1038/nature12111

Hua, Y., Robinson, T. J., Cao, Y., Shi, G. P., Ren, J., & Nair, S. (2015). Cathepsin K knockout alleviates aging-induced cardiac dysfunction. *Aging Cell*, *14*(3), 345-351. https://doi.org/10.1111/acel.12276

Hua, Y., Xu, X., Shi, G. P., Chicco, A. J., Ren, J., & Nair, S. (2013). Cathepsin K knockout alleviates pressure overload-induced cardiac hypertrophy. *Hypertension*, *61*(6), 1184-1192. https://doi.org/10.1161/HYPERTENSIONAHA.111.00947

Huang, J., Zhang, L., Wan, D., Zhou, L., Zheng, S., Lin, S., & Qiao, Y. (2021). Extracellular matrix and its therapeutic potential for cancer treatment. *Signal Transduct Target Ther*, *6*(1), 153. https://doi.org/10.1038/s41392-021-00544-0

Huang, T. C., Chang, W. T., Hu, Y. C., Hsieh, B. S., Cheng, H. L., Yen, J. H., Chiu, P. R., & Chang, K. L. (2018). Zinc Protects Articular Chondrocytes through Changes in Nrf2-Mediated Antioxidants, Cytokines and Matrix Metalloproteinases. *Nutrients*, *10*(4). https://doi.org/10.3390/nu10040471

Humes, H. D., Nakamura, T., Cieslinski, D. A., Miller, D., Emmons, R. V., & Border, W. A. (1993). Role of proteoglycans and cytoskeleton in the effects of TGF-beta 1 on renal proximal tubule cells. *Kidney Int*, *43*(3), 575-584. https://doi.org/10.1038/ki.1993.85

Hunter, D. J., McDougall, J. J., & Keefe, F. J. (2008). The symptoms of osteoarthritis and the genesis of pain. *Rheum Dis Clin North Am*, *34*(3), 623-643. https://doi.org/10.1016/j.rdc.2008.05.004

Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., Tung, J. Y., Hinds, D. A., Perlis, R. H., & Winslow, A. R. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet*, *48*(9), 1031-1036. https://doi.org/10.1038/ng.3623

Ihara, S., Hirata, Y., & Koike, K. (2017). TGF-beta in inflammatory bowel disease: a key regulator of immune cells, epithelium, and the intestinal microbiota. *J Gastroenterol*, *52*(7), 777-787. https://doi.org/10.1007/s00535-017-1350-1

Ismail, H. M., Miotla-Zarebska, J., Troeberg, L., Tang, X., Stott, B., Yamamoto, K., Nagase, H., Fosang, A. J., Vincent, T. L., & Saklatvala, J. (2016). Brief Report: JNK-2 Controls Aggrecan Degradation in Murine Articular Cartilage and the Development of Experimental Osteoarthritis. *Arthritis Rheumatol*, *68*(5), 1165-1171. https://doi.org/10.1002/art.39547

Ivanovic, V., Demajo, M., Krtolica, K., Krajnovic, M., Konstantinovic, M., Baltic, V., Prtenjak, G., Stojiljkovic, B., Breberina, M., Neskovic-Konstantinovic, Z., Nikolic-Vukosavljevic, D., & Dimitrijevic, B. (2006). Elevated plasma TGF-beta1 levels correlate with decreased survival of metastatic breast cancer patients. *Clin Chim Acta*, *371*(1-2), 191-193. https://doi.org/10.1016/j.cca.2006.02.027

Ivashkiv, L. B., & Hu, X. (2004). Signaling by STATs. *Arthritis Res Ther*, *6*(4), 159-168. https://doi.org/10.1186/ar1197

Jansen, P. R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A. R., de Leeuw, C. A., Benjamins, J. S., Munoz-Manchado, A. B., Nagel, M., Savage, J. E., Tiemeier, H., White, T., andMe Research, T., Tung, J. Y., Hinds, D. A., Vacic, V., Wang, X., Sullivan, P. F., . . . Posthuma, D. (2019). Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet*, *51*(3), 394-403. https://doi.org/10.1038/s41588-018-0333-3

Janssen, I., Heymsfield, S. B., Wang, Z. M., & Ross, R. (2000). Skeletal muscle mass and distribution in 468 men and women aged 18-88 yr. *J Appl Physiol (1985)*, *89*(1), 81-88. https://doi.org/10.1152/jappl.2000.89.1.81

Janssens, K., ten Dijke, P., Janssens, S., & Van Hul, W. (2005). Transforming growth factor-beta1 to the bone. *Endocr Rev*, *26*(6), 743-774. https://doi.org/10.1210/er.2004-0001

Kaartinen, V., Voncken, J. W., Shuler, C., Warburton, D., Bu, D., Heisterkamp, N., & Groffen, J. (1995). Abnormal lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of epithelial-mesenchymal interaction. *Nat Genet*, *11*(4), 415-421. https://doi.org/10.1038/ng1295-415

Kain, D., Simon, A. J., Greenberg, A., Ben Zvi, D., Gilburd, B., & Schneiderman, J. (2018). Cardiac leptin overexpression in the context of acute MI and reperfusion potentiates myocardial remodeling and left ventricular dysfunction. *PLoS One*, *13*(10), e0203902. https://doi.org/10.1371/journal.pone.0203902

Kainulainen, K., Karttunen, L., Puhakka, L., Sakai, L., & Peltonen, L. (1994). Mutations in the fibrillin gene responsible for dominant ectopia lentis and neonatal Marfan syndrome. *Nat Genet*, *6*(1), 64-69. https://doi.org/10.1038/ng0194-64

Kamolmatyakul, S., Chen, W., & Li, Y. P. (2001). Interferon-gamma down-regulates gene expression of cathepsin K in osteoclasts and inhibits osteoclast formation. *J Dent Res*, *80*(1), 351-355. https://doi.org/10.1177/00220345010800011001

Kang, Y., He, W., Tulley, S., Gupta, G. P., Serganova, I., Chen, C. R., Manova-Todorova, K., Blasberg, R., Gerald, W. L., & Massague, J. (2005). Breast cancer bone metastasis mediated by the Smad tumor suppressor pathway. *Proc Natl Acad Sci U S A*, *102*(39), 13909-13914. https://doi.org/10.1073/pnas.0506517102

Kania, K., Colella, F., Riemen, A. H. K., Wang, H., Howard, K. A., Aigner, T., Dell'Accio, F., Capellini, T. D., Roelofs, A. J., & De Bari, C. (2020). Regulation of Gdf5 expression in joint remodelling, repair and osteoarthritis. *Sci Rep*, *10*(1), 157. https://doi.org/10.1038/s41598-019-57011-8

Katsel, P., Davis, K. L., & Haroutunian, V. (2005). Variations in myelin and oligodendrocyte-related gene expression across multiple brain regions in schizophrenia: a gene ontology study. *Schizophr Res*, *79*(2-3), 157-173. https://doi.org/10.1016/j.schres.2005.06.007

Kennedy, B. K., Berger, S. L., Brunet, A., Campisi, J., Cuervo, A. M., Epel, E. S., Franceschi, C., Lithgow, G. J., Morimoto, R. I., Pessin, J. E., Rando, T. A., Richardson, A., Schadt, E. E., Wyss-Coray,

T., & Sierra, F. (2014). Geroscience: linking aging to chronic disease. *Cell*, *159*(4), 709-713. https://doi.org/10.1016/j.cell.2014.10.039

Kettleborough, R. N., Busch-Nentwich, E. M., Harvey, S. A., Dooley, C. M., de Bruijn, E., van Eeden, F., Sealy, I., White, R. J., Herd, C., Nijman, I. J., Fenyes, F., Mehroke, S., Scahill, C., Gibbons, R., Wali, N., Carruthers, S., Hall, A., Yen, J., Cuppen, E., & Stemple, D. L. (2013). A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature*, *496*(7446), 494-497. https://doi.org/10.1038/nature11992

Khalil, N. (1999). TGF-beta: from latent to active. *Microbes Infect*, *1*(15), 1255-1263. https://doi.org/10.1016/s1286-4579(99)00259-2

Khandelwal, N., Simpson, J., Taylor, G., Rafique, S., Whitehouse, A., Hiscox, J., & Stark, L. A. (2011). Nucleolar NF-kappaB/RelA mediates apoptosis by causing cytoplasmic relocalization of nucleophosmin. *Cell Death Differ*, *18*(12), 1889-1903. https://doi.org/10.1038/cdd.2011.79

King, L. K., March, L., & Anandacoomarasamy, A. (2013). Obesity & osteoarthritis. *Indian J Med Res*, *138*(2), 185-193. https://www.ncbi.nlm.nih.gov/pubmed/24056594

Kinnersley, M., Schwartz, K., Yang, D. D., Sherlock, G., & Rosenzweig, F. (2021). Evolutionary dynamics and structural consequences of de novo beneficial mutations and mutant lineages arising in a constant environment. *BMC Biol*, *19*(1), 20. https://doi.org/10.1186/s12915-021-00954-0

Kinoshita, A., Saito, T., Tomita, H., Makita, Y., Yoshida, K., Ghadami, M., Yamada, K., Kondo, S., Ikegawa, S., Nishimura, G., Fukushima, Y., Nakagomi, T., Saito, H., Sugimoto, T., Kamegaya, M., Hisa, K., Murray, J. C., Taniguchi, N., Niikawa, N., & Yoshiura, K. (2000). Domain-specific mutations in TGFB1 result in Camurati-Engelmann disease. *Nat Genet*, *26*(1), 19-20. https://doi.org/10.1038/79128

Kirschke, H., Barrett, A. J., & Rawlings, N. D. (1995). Proteinases 1: lysosomal cysteine proteinases. *Protein Profile*, *2*(14), 1581-1643. https://www.ncbi.nlm.nih.gov/pubmed/8771190

Kiviranta, I., Tammi, M., Jurvelin, J., Arokoski, J., Saamanen, A. M., & Helminen, H. J. (1992). Articular cartilage thickness and glycosaminoglycan distribution in the canine knee joint after strenuous running exercise. *Clin Orthop Relat Res*(283), 302-308. https://www.ncbi.nlm.nih.gov/pubmed/1395265

Krum, S. A., Chang, J., Miranda-Carboni, G., & Wang, C. Y. (2010). Novel functions for NFkappaB: inhibition of bone formation. *Nat Rev Rheumatol*, *6*(10), 607-611. https://doi.org/10.1038/nrrheum.2010.133

Kone, A. P., & Scharf, D. (2021). Prevalence of multimorbidity in adults with cancer, and associated health service utilization in Ontario, Canada: a population-based retrospective cohort study. *BMC Cancer*, *21*(1), 406. https://doi.org/10.1186/s12885-021-08102-1

Kong, F. M., Anscher, M. S., Murase, T., Abbott, B. D., Iglehart, J. D., & Jirtle, R. L. (1995). Elevated plasma transforming growth factor-beta 1 levels in breast cancer patients decrease after surgical removal of the tumor. *Ann Surg*, *222*(2), 155-162. https://doi.org/10.1097/00000658-199508000-00007

Konttinen, Y. T., Mandelin, J., Li, T. F., Salo, J., Lassus, J., Liljestrom, M., Hukkanen, M., Takagi, M., Virtanen, I., & Santavirta, S. (2002). Acidic cysteine endoproteinase cathepsin K in the degeneration of the superficial articular hyaline cartilage in osteoarthritis. *Arthritis Rheum*, *46*(4), 953-960. https://doi.org/10.1002/art.10185

Kotlarz, D., Marquardt, B., Baroy, T., Lee, W. S., Konnikova, L., Hollizeck, S., Magg, T., Lehle, A. S., Walz, C., Borggraefe, I., Hauck, F., Bufler, P., Conca, R., Wall, S. M., Schumacher, E. M., Misceo, D., Frengen, E., Bentsen, B. S., Uhlig, H. H., . . . Klein, C. (2018). Human TGF-beta1 deficiency causes severe inflammatory bowel disease and encephalopathy. *Nat Genet*, *50*(3), 344-348. https://doi.org/10.1038/s41588-018-0063-6

Kozawa, E., Nishida, Y., Cheng, X. W., Urakawa, H., Arai, E., Futamura, N., Shi, G. P., Kuzuya, M., Hu, L., Sasaki, T., & Ishiguro, N. (2012). Osteoarthritic change is delayed in a Ctsk-knockout mouse model of osteoarthritis. *Arthritis Rheum*, *64*(2), 454-464. https://doi.org/10.1002/art.33398

Lambert, C. J., Freshner, B. C., Chung, A., Stevenson, T. J., Bowles, D. M., Samuel, R., Gale, B. K., & Bonkowsky, J. L. (2018). An automated system for rapid cellular extraction from live zebrafish embryos and larvae: Development and application to genotyping. *PLoS One*, *13*(3), e0193180. https://doi.org/10.1371/journal.pone.0193180

Lau, H. H., Ho, A. Y., Luk, K. D., & Kung, A. W. (2004). Transforming growth factor-beta1 gene polymorphisms and bone turnover, bone mineral density and fracture risk in southern Chinese women. *Calcif Tissue Int*, *74*(6), 516-521. https://doi.org/10.1007/s00223-004-0163-4

Laurie, C. C., Doheny, K. F., Mirel, D. B., Pugh, E. W., Bierut, L. J., Bhangale, T., Boehm, F., Caporaso, N. E., Cornelis, M. C., Edenberg, H. J., Gabriel, S. B., Harris, E. L., Hu, F. B., Jacobs, K. B., Kraft, P., Landi, M. T., Lumley, T., Manolio, T. A., McHugh, C., . . . Investigators, G. (2010). Quality

control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol*, *34*(6), 591-602. https://doi.org/10.1002/gepi.20516

Lavine, K. J., Yu, K., White, A. C., Zhang, X., Smith, C., Partanen, J., & Ornitz, D. M. (2005). Endocardial and epicardial derived FGF signals regulate myocardial proliferation and differentiation in vivo. *Dev Cell*, *8*(1), 85-95. https://doi.org/10.1016/j.devcel.2004.12.002

Le Gall, C., Bellahcene, A., Bonnelye, E., Gasser, J. A., Castronovo, V., Green, J., Zimmermann, J., & Clezardin, P. (2007). A cathepsin K inhibitor reduces breast cancer induced osteolysis and skeletal tumor burden. *Cancer Res*, *67*(20), 9894-9902. https://doi.org/10.1158/0008-5472.CAN-06-3940

LeBlanc, M., Zuber, V., Andreassen, B. K., Witoelar, A., Zeng, L., Bettella, F., Wang, Y., McEvoy, L. K., Thompson, W. K., Schork, A. J., Reppe, S., Barrett-Connor, E., Ligthart, S., Dehghan, A., Gautvik, K. M., Nelson, C. P., Schunkert, H., Samani, N. J., Consortium, C. A., . . . Andreassen, O. A. (2016). Identifying Novel Gene Variants in Coronary Artery Disease and Shared Genes With Several Cardiovascular Risk Factors. *Circ Res*, *118*(1), 83-94. https://doi.org/10.1161/CIRCRESAHA.115.306629

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linner, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., . . . Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*, *50*(8), 1112-1121. https://doi.org/10.1038/s41588-018-0147-3

Lee, K., Seo, I., Choi, M. H., & Jeong, D. (2018). Roles of Mitogen-Activated Protein Kinases in Osteoclast Biology. *Int J Mol Sci*, *19*(10). https://doi.org/10.3390/ijms19103004

Li, D., Achkar, J. P., Haritunians, T., Jacobs, J. P., Hui, K. Y., D'Amato, M., Brand, S., Radford-Smith, G., Halfvarson, J., Niess, J. H., Kugathasan, S., Buning, C., Schumm, L. P., Klei, L., Ananthakrishnan, A., Aumais, G., Baidoo, L., Dubinsky, M., Fiocchi, C., . . . Duerr, R. H. (2016). A Pleiotropic Missense Variant in SLC39A8 Is Associated With Crohn's Disease and Human Gut Microbiome Composition. *Gastroenterology*, *151*(4), 724-732. https://doi.org/10.1053/j.gastro.2016.06.051

Li, D., Ni, S., Miao, K. S., & Zhuang, C. (2019). PI3K/Akt and caspase pathways mediate oxidative stress-induced chondrocyte apoptosis. *Cell Stress Chaperones*, *24*(1), 195-202. https://doi.org/10.1007/s12192-018-0956-4

Li, G., Cheng, T., & Yu, X. (2021). The Impact of Trace Elements on Osteoarthritis. *Front Med (Lausanne)*, *8*, 771297. https://doi.org/10.3389/fmed.2021.771297

Li, N., Yang, P., Tang, M., Liu, Y., Guo, W., Lang, B., Wang, J., Wu, H., Tang, H., Yu, Y., Wu, X., Zeng, C., Cao, T., & Cai, H. (2022). Reduced erythrocyte membrane polyunsaturated fatty acid levels indicate diminished treatment response in patients with multi- versus first-episode schizophrenia. *Schizophrenia (Heidelb)*, *8*(1), 7. https://doi.org/10.1038/s41537-022-00214-2

Li, T., Yu, H., Zhang, D., Feng, T., Miao, M., Li, J., & Liu, X. (2022). Matrix Vesicles as a Therapeutic Target for Vascular Calcification. *Front Cell Dev Biol*, *10*, 825622. https://doi.org/10.3389/fcell.2022.825622

Li, Z., Chen, J., Yu, H., He, L., Xu, Y., Zhang, D., Yi, Q., Li, C., Li, X., Shen, J., Song, Z., Ji, W., Wang, M., Zhou, J., Chen, B., Liu, Y., Wang, J., Wang, P., Yang, P., . . . Shi, Y. (2017). Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat Genet*, *49*(11), 1576-1583. https://doi.org/10.1038/ng.3973

Liau, N. P. D., Laktyushin, A., Lucet, I. S., Murphy, J. M., Yao, S., Whitlock, E., Callaghan, K., Nicola, N. A., Kershaw, N. J., & Babon, J. J. (2018). The molecular basis of JAK/STAT inhibition by SOCS1. *Nat Commun*, *9*(1), 1558. https://doi.org/10.1038/s41467-018-04013-1

Lin, A., Zhu, L., Jiang, A., Mou, W., Zhang, J., & Luo, P. (2022). Activation of the TGF-beta Pathway Enhances the Efficacy of Platinum-Based Chemotherapy in Small Cell Lung Cancer Patients. *Dis Markers*, *2022*, 8766448. https://doi.org/10.1155/2022/8766448

Lin, A. Y., Ding, Y., Vanselow, D. J., Katz, S. R., Yakovlev, M. A., Clark, D. P., Mandrell, D., Copper, J. E., van Rossum, D. B., & Cheng, K. C. (2018). Rigid Embedding of Fixed and Stained, Whole, Millimeter-Scale Specimens for Section-free 3D Histology by Micro-Computed Tomography. *J Vis Exp*(140). https://doi.org/10.3791/58293

Liu, H., Kato, Y., Erzinger, S. A., Kiriakova, G. M., Qian, Y., Palmieri, D., Steeg, P. S., & Price, J. E. (2012). The role of MMP-1 in breast cancer growth and metastasis to the brain in a xenograft model. *BMC Cancer*, *12*, 583. https://doi.org/10.1186/1471-2407-12-583

Loeser, R. F., Erickson, E. A., & Long, D. L. (2008). Mitogen-activated protein kinases as therapeutic targets in osteoarthritis. *Curr Opin Rheumatol*, *20*(5), 581-586. https://doi.org/10.1097/BOR.0b013e3283090463

Loewe, L., & Hill, W. G. (2010). The population genetics of mutations: good, bad and indifferent. *Philos Trans R Soc Lond B Biol Sci*, *365*(1544), 1153-1167. https://doi.org/10.1098/rstb.2009.0317

Loeys, B. L., Schwarze, U., Holm, T., Callewaert, B. L., Thomas, G. H., Pannu, H., De Backer, J. F., Oswald, G. L., Symoens, S., Manouvrier, S., Roberts, A. E., Faravelli, F., Greco, M. A., Pyeritz, R. E., Milewicz, D. M., Coucke, P. J., Cameron, D. E., Braverman, A. C., Byers, P. H., . . . Dietz, H. C. (2006). Aneurysm syndromes caused by mutations in the TGF-beta receptor. *N Engl J Med*, *355*(8), 788-798. https://doi.org/10.1056/NEJMoa055695

Logar, D. B., Komadina, R., Prezelj, J., Ostanek, B., Trost, Z., & Marc, J. (2007). Expression of bone resorption genes in osteoarthritis and in osteoporosis. *J Bone Miner Metab*, *25*(4), 219-225. https://doi.org/10.1007/s00774-007-0753-0

Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S., & Hirschhorn, J. N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet*, *33*(2), 177-182. https://doi.org/10.1038/ng1071

Long, H., Liu, Q., Yin, H., Wang, K., Diao, N., Zhang, Y., Lin, J., & Guo, A. (2022). Prevalence Trends of Site-Specific Osteoarthritis From 1990 to 2019: Findings From the Global Burden of Disease Study 2019. *Arthritis Rheumatol*, *74*(7), 1172-1183. https://doi.org/10.1002/art.42089

Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The hallmarks of aging. *Cell*, *153*(6), 1194-1217. https://doi.org/10.1016/j.cell.2013.05.039

Lotinun, S., Kiviranta, R., Matsubara, T., Alzate, J. A., Neff, L., Luth, A., Koskivirta, I., Kleuser, B., Vacher, J., Vuorio, E., Horne, W. C., & Baron, R. (2013). Osteoclast-specific cathepsin K deletion stimulates S1P-dependent bone formation. *J Clin Invest*, *123*(2), 666-681. https://doi.org/10.1172/JCI64840

Lu, P. N., Moreland, T., Christian, C. J., Lund, T. C., Steet, R. A., & Flanagan-Steet, H. (2020). Inappropriate cathepsin K secretion promotes its enzymatic activation driving heart and valve malformation. *JCI Insight*, *5*(20). https://doi.org/10.1172/jci.insight.133019

Lui, H., Denbeigh, J., Vaquette, C., Tran, H. M., Dietz, A. B., Cool, S. M., Dudakovic, A., Kakar, S., & van Wijnen, A. J. (2021). Fibroblastic differentiation of mesenchymal stem/stromal cells (MSCs) is enhanced by hypoxia in 3D cultures treated with bone morphogenetic protein 6 (BMP6) and growth and differentiation factor 5 (GDF5). *Gene*, *788*, 145662. https://doi.org/10.1016/j.gene.2021.145662

Maldonado, M., & Nam, J. (2013). The role of changes in extracellular matrix of cartilage in the presence of inflammation on the pathology of osteoarthritis. *Biomed Res Int*, *2013*, 284873. https://doi.org/10.1155/2013/284873

Malemud, C. J. (2018). The role of the JAK/STAT signal pathway in rheumatoid arthritis. *Ther Adv Musculoskelet Dis*, *10*(5-6), 117-127. https://doi.org/10.1177/1759720X18776224

Mallat, Z., Fornes, P., Costagliola, R., Esposito, B., Belmin, J., Lecomte, D., & Tedgui, A. (2001). Age and gender effects on cardiomyocyte apoptosis in the normal human heart. *J Gerontol A Biol Sci Med Sci*, *56*(11), M719-723. https://doi.org/10.1093/gerona/56.11.m719

Mang, T., Kleinschmidt-Dorr, K., Ploeger, F., Lindemann, S., & Gigout, A. (2020). The GDF-5 mutant M1673 exerts robust anabolic and anti-catabolic effects in chondrocytes. *J Cell Mol Med*, *24*(13), 7141-7150. https://doi.org/10.1111/jcmm.15149

Marees, A. T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., & Derks, E. M. (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int J Methods Psychiatr Res*, *27*(2), e1608. https://doi.org/10.1002/mpr.1608

Mascareno, E., El-Shafei, M., Maulik, N., Sato, M., Guo, Y., Das, D. K., & Siddiqui, M. A. (2001). JAK/STAT signaling is associated with cardiac dysfunction during ischemia and reperfusion. *Circulation*, *104*(3), 325-329. https://doi.org/10.1161/01.cir.104.3.325

Massague, J. (1998). TGF-beta signal transduction. *Annu Rev Biochem*, *67*, 753-791. https://doi.org/10.1146/annurev.biochem.67.1.753

Massague, J., Blain, S. W., & Lo, R. S. (2000). TGFbeta signaling in growth control, cancer, and heritable disorders. *Cell*, *103*(2), 295-309. https://doi.org/10.1016/s0092-8674(00)00121-5

Massague, J., Seoane, J., & Wotton, D. (2005). Smad transcription factors. *Genes Dev*, *19*(23), 2783-2810. https://doi.org/10.1101/gad.1350705

Matt, P., Schoenhoff, F., Habashi, J., Holm, T., Van Erp, C., Loch, D., Carlson, O. D., Griswold, B. F., Fu, Q., De Backer, J., Loeys, B., Huso, D. L., McDonnell, N. B., Van Eyk, J. E., Dietz, H. C., & Gen, T. A. C. C. (2009). Circulating transforming growth factor-beta in Marfan syndrome. *Circulation*, *120*(6), 526-532. https://doi.org/10.1161/CIRCULATIONAHA.108.841981

Mauck, R. L., Yuan, X., & Tuan, R. S. (2006). Chondrogenic differentiation and functional maturation of bovine mesenchymal stem cells in long-term agarose culture. *Osteoarthritis Cartilage*, *14*(2), 179-189. https://doi.org/10.1016/j.joca.2005.09.002

McGlinchey, R. P., & Lee, J. C. (2015). Cysteine cathepsins are essential in lysosomal degradation of alpha-synuclein. *Proc Natl Acad Sci U S A*, *112*(30), 9322-9327. https://doi.org/10.1073/pnas.1500937112

McGrath, M. E., Klaus, J. L., Barnes, M. G., & Bromme, D. (1997). Crystal structure of human cathepsin K complexed with a potent inhibitor. *Nat Struct Biol*, *4*(2), 105-109. https://doi.org/10.1038/nsb0297-105

McPhail, S. M. (2016). Multimorbidity in chronic disease: impact on health care resources and costs. *Risk Manag Healthc Policy*, *9*, 143-156. https://doi.org/10.2147/RMHP.S97248

Mealer, R. G., Jenkins, B. G., Chen, C. Y., Daly, M. J., Ge, T., Lehoux, S., Marquardt, T., Palmer, C. D., Park, J. H., Parsons, P. J., Sackstein, R., Williams, S. E., Cummings, R. D., Scolnick, E. M., & Smoller, J. W. (2020). The schizophrenia risk locus in SLC39A8 alters brain metal transport and plasma glycosylation. *Sci Rep*, *10*(1), 13162. https://doi.org/10.1038/s41598-020-70108-9

Melzer, C., Hass, R., von der Ohe, J., Lehnert, H., & Ungefroren, H. (2017). The role of TGF-beta and its crosstalk with RAC1/RAC1b signaling in breast and pancreas carcinoma. *Cell Commun Signal*, *15*(1), 19. https://doi.org/10.1186/s12964-017-0175-0

Mengshol, J. A., Vincenti, M. P., & Brinckerhoff, C. E. (2001). IL-1 induces collagenase-3 (MMP-13) promoter activity in stably transfected chondrocytic cells: requirement for Runx-2 and activation by p38 MAPK and JNK pathways. *Nucleic Acids Res*, *29*(21), 4361-4372. https://doi.org/10.1093/nar/29.21.4361

Merino, R., Rodriguez-Leon, J., Macias, D., Ganan, Y., Economides, A. N., & Hurle, J. M. (1999). The BMP antagonist Gremlin regulates outgrowth, chondrogenesis and programmed cell death in the developing limb. *Development*, *126*(23), 5515-5522. https://doi.org/10.1242/dev.126.23.5515

Messier, S. P., Resnik, A. E., Beavers, D. P., Mihalko, S. L., Miller, G. D., Nicklas, B. J., deVita, P., Hunter, D. J., Lyles, M. F., Eckstein, F., Guermazi, A., & Loeser, R. F. (2018). Intentional Weight Loss in Overweight and Obese Patients With Knee Osteoarthritis: Is More Better? *Arthritis Care Res (Hoboken)*, *70*(11), 1569-1575. https://doi.org/10.1002/acr.23608

Migliore, A., & Procopio, S. (2015). Effectiveness and utility of hyaluronic acid in osteoarthritis. *Clin Cases Miner Bone Metab*, *12*(1), 31-33. https://doi.org/10.11138/ccmbm/2015.12.1.031

Mikic, B. (2004). Multiple effects of GDF-5 deficiency on skeletal tissues: implications for therapeutic bioengineering. *Ann Biomed Eng*, *32*(3), 466-476. https://doi.org/10.1023/b:abme.0000017549.57126.51

Milne, R. L., Kuchenbaecker, K. B., Michailidou, K., Beesley, J., Kar, S., Lindstrom, S., Hui, S., Lemacon, A., Soucy, P., Dennis, J., Jiang, X., Rostamianfar, A., Finucane, H., Bolla, M. K., McGuffog,

L., Wang, Q., Aalfs, C. M., Investigators, A., Adams, M., . . . Simard, J. (2017). Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nat Genet*, *49*(12), 1767-1778. https://doi.org/10.1038/ng.3785

Mitchell, R. E., Huitema, L. F., Skinner, R. E., Brunt, L. H., Severn, C., Schulte-Merker, S., & Hammond, C. L. (2013). New tools for studying osteoarthritis genetics in zebrafish. *Osteoarthritis Cartilage*, *21*(2), 269-278. https://doi.org/10.1016/j.joca.2012.11.004

Mocali, A., Cedrola, S., Della Malva, N., Bontempelli, M., Mitidieri, V. A., Bavazzano, A., Comolli, R., Paoletti, F., & La Porta, C. A. (2004). Increased plasma levels of soluble CD40, together with the decrease of TGF beta 1, as possible differential markers of Alzheimer disease. *Exp Gerontol*, *39*(10), 1555-1561. https://doi.org/10.1016/j.exger.2004.07.007

Morko, J. P., Soderstrom, M., Saamanen, A. M., Salminen, H. J., & Vuorio, E. I. (2004). Up regulation of cathepsin K expression in articular chondrocytes in a transgenic mouse model for osteoarthritis. *Ann Rheum Dis*, *63*(6), 649-655. https://doi.org/10.1136/ard.2002.004671

Moschen, A. R., Tilg, H., & Raine, T. (2019). IL-12, IL-23 and IL-17 in IBD: immunobiology and therapeutic targeting. *Nat Rev Gastroenterol Hepatol*, *16*(3), 185-196. https://doi.org/10.1038/s41575-018-0084-8

Muliyala, K. P., & Varghese, M. (2010). The complex relationship between depression and dementia. *Ann Indian Acad Neurol*, *13*(Suppl 2), S69-73. https://doi.org/10.4103/0972-2327.74248

Mullard, A. (2016). Merck &Co. drops osteoporosis drug odanacatib. *Nat Rev Drug Discov*, *15*(10), 669. https://doi.org/10.1038/nrd.2016.207

Muriach, M., Flores-Bellver, M., Romero, F. J., & Barcia, J. M. (2014). Diabetes and the brain: oxidative stress, inflammation, and autophagy. *Oxid Med Cell Longev*, *2014*, 102158. https://doi.org/10.1155/2014/102158

Muslin, A. J. (2008). MAPK signalling in cardiovascular health and disease: molecular mechanisms and therapeutic targets. *Clin Sci (Lond)*, *115*(7), 203-218. https://doi.org/10.1042/CS20070430

Nebreda, A. R., & Porras, A. (2000). p38 MAP kinases: beyond the stress response. *Trends Biochem Sci*, *25*(6), 257-260. https://doi.org/10.1016/s0968-0004(00)01595-4

Nelson, C. E., Wu, Y., Gemberling, M. P., Oliver, M. L., Waller, M. A., Bohning, J. D., Robinson-Hamm, J. N., Bulaklak, K., Castellanos Rivera, R. M., Collier, J. H., Asokan, A., & Gersbach, C. A.

(2019). Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. *Nat Med*, *25*(3), 427-432. https://doi.org/10.1038/s41591-019-0344-3

Nickel, J., Kotzsch, A., Sebald, W., & Mueller, T. D. (2005). A single residue of GDF-5 defines binding specificity to BMP receptor IB. *J Mol Biol*, *349*(5), 933-947. https://doi.org/10.1016/j.jmb.2005.04.015

Nielsen, J., Skadhede, S., & Correll, C. U. (2010). Antipsychotics associated with the development of type 2 diabetes in antipsychotic-naive schizophrenia patients. *Neuropsychopharmacology*, *35*(9), 1997-2004. https://doi.org/10.1038/npp.2010.78

Nikpay, M., Goel, A., Won, H. H., Hall, L. M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C. P., Hopewell, J. C., Webb, T. R., Zeng, L., Dehghan, A., Alver, M., Armasu, S. M., Auro, K., Bjonnes, A., Chasman, D. I., Chen, S., . . . Farrall, M. (2015). A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*, *47*(10), 1121-1130. https://doi.org/10.1038/ng.3396

Niu, G. J., Xu, J. D., Yuan, W. J., Sun, J. J., Yang, M. C., He, Z. H., Zhao, X. F., & Wang, J. X. (2018). Protein Inhibitor of Activated STAT (PIAS) Negatively Regulates the JAK/STAT Pathway by Inhibiting STAT Phosphorylation and Translocation. *Front Immunol*, *9*, 2392. https://doi.org/10.3389/fimmu.2018.02392

Nuesch, E., Dieppe, P., Reichenbach, S., Williams, S., Iff, S., & Juni, P. (2011). All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ*, *342*, d1165. https://doi.org/10.1136/bmj.d1165

Ohta, M., Greenberger, J. S., Anklesaria, P., Bassols, A., & Massague, J. (1987). Two forms of transforming growth factor-beta distinguished by multipotential haematopoietic progenitor cells. *Nature*, *329*(6139), 539-541. https://doi.org/10.1038/329539a0

O'Keeffe, G. W., Gutierrez, H., Howard, L., Laurie, C. W., Osorio, C., Gavalda, N., Wyatt, S. L., & Davies, A. M. (2016). Region-specific role of growth differentiation factor-5 in the establishment of sympathetic innervation. *Neural Dev*, *11*, 4. https://doi.org/10.1186/s13064-016-0060-3

Okuda, S., Languino, L. R., Ruoslahti, E., & Border, W. A. (1990). Elevated expression of transforming growth factor-beta and proteoglycan production in experimental glomerulonephritis. Possible role in expansion of the mesangial extracellular matrix. *J Clin Invest*, *86*(2), 453-462. https://doi.org/10.1172/JCI114731

Ong, K. L., Wu, B. J., Cheung, B. M., Barter, P. J., & Rye, K. A. (2013). Arthritis: its prevalence, risk factors, and association with cardiovascular diseases in the United States, 1999 to 2008. *Ann Epidemiol*, *23*(2), 80-86. https://doi.org/10.1016/j.annepidem.2012.11.008

O'Shea, J. J., & Plenge, R. (2012). JAK and STAT signaling molecules in immunoregulation and immune-mediated disease. *Immunity*, *36*(4), 542-550. https://doi.org/10.1016/j.immuni.2012.03.014

Osiri, M., Welch, V., Brosseau, L., Shea, B., McGowan, J., Tugwell, P., & Wells, G. (2000). Transcutaneous electrical nerve stimulation for knee osteoarthritis. *Cochrane Database Syst Rev*(4), CD002823. https://doi.org/10.1002/14651858.CD002823

Osztovits, J., Horvath, T., Littvay, L., Steinbach, R., Jermendy, A., Tarnoki, A., Tarnoki, D., Metneki, J., Kollai, M., & Jermendy, G. (2011). Effects of genetic vs. environmental factors on cardiovascular autonomic function: a twin study. *Diabet Med*, *28*(10), 1241-1248. https://doi.org/10.1111/j.1464-5491.2011.03363.x

Padmalayam, I., & Suto, M. (2013). Role of adiponectin in the metabolic syndrome: current perspectives on its modulation as a treatment strategy. *Curr Pharm Des*, *19*(32), 5755-5763. https://doi.org/10.2174/13816128113199990360

Panda, D., Samuel, J. C., Massie, M., Feinstein, S. C., & Wilson, L. (2003). Differential regulation of microtubule dynamics by three- and four-repeat tau: implications for the onset of neurodegenerative disease. *Proc Natl Acad Sci U S A*, *100*(16), 9548-9553. https://doi.org/10.1073/pnas.1633508100

Pandur, P., Lasche, M., Eisenberg, L. M., & Kuhl, M. (2002). Wnt-11 activation of a non-canonical Wnt signalling pathway is required for cardiogenesis. *Nature*, *418*(6898), 636-641. https://doi.org/10.1038/nature00921

Parrish, W. R., Byers, B. A., Su, D., Geesin, J., Herzberg, U., Wadsworth, S., Bendele, A., & Story, B. (2017). Intra-articular therapy with recombinant human GDF5 arrests disease progression and stimulates cartilage repair in the rat medial meniscus transection (MMT) model of osteoarthritis. *Osteoarthritis Cartilage*, *25*(4), 554-560. https://doi.org/10.1016/j.joca.2016.11.002

Pedram, A., Razandi, M., Hu, R. M., & Levin, E. R. (1997). Vasoactive peptides modulate vascular endothelial cell growth factor production and endothelial cell proliferation and invasion. *J Biol Chem*, *272*(27), 17097-17103. https://doi.org/10.1074/jbc.272.27.17097

Petrey, A. C., Flanagan-Steet, H., Johnson, S., Fan, X., De la Rosa, M., Haskins, M. E., Nairn, A. V., Moremen, K. W., & Steet, R. (2012). Excessive activity of cathepsin K is associated with cartilage

defects in a zebrafish model of mucolipidosis II. *Dis Model Mech*, *5*(2), 177-190. https://doi.org/10.1242/dmm.008219

Petrie, J. R., Guzik, T. J., & Touyz, R. M. (2018). Diabetes, Hypertension, and Cardiovascular Disease: Clinical Insights and Vascular Mechanisms. *Can J Cardiol*, *34*(5), 575-584. https://doi.org/10.1016/j.cjca.2017.12.005

Pfander, D., Heinz, N., Rothe, P., Carl, H. D., & Swoboda, B. (2004). Tenascin and aggrecan expression by articular chondrocytes is influenced by interleukin 1beta: a possible explanation for the changes in matrix synthesis during osteoarthritis. *Ann Rheum Dis*, *63*(3), 240-244. https://doi.org/10.1136/ard.2002.003749

Pickrell, J. K., Berisa, T., Liu, J. Z., Segurel, L., Tung, J. Y., & Hinds, D. A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet*, *48*(7), 709-717. https://doi.org/10.1038/ng.3570

Pirinen, M., Donnelly, P., & Spencer, C. C. (2012). Including known covariates can reduce power to detect genetic effects in case-control studies. *Nat Genet*, *44*(8), 848-851. https://doi.org/10.1038/ng.2346

Pollard, T. C., Gwilym, S. E., & Carr, A. J. (2008). The assessment of early osteoarthritis. *J Bone Joint Surg Br*, *90*(4), 411-421. https://doi.org/10.1302/0301-620X.90B4.20284

Prasadam, I., Mao, X., Shi, W., Crawford, R., & Xiao, Y. (2013). Combination of MEK-ERK inhibitor and hyaluronic acid has a synergistic effect on anti-hypertrophic and pro-chondrogenic activities in osteoarthritis treatment. *J Mol Med (Berl)*, *91*(3), 369-380. https://doi.org/10.1007/s00109-012-0953-5

Proetzel, G., Pawlowski, S. A., Wiles, M. V., Yin, M., Boivin, G. P., Howles, P. N., Ding, J., Ferguson, M. W., & Doetschman, T. (1995). Transforming growth factor-beta 3 is required for secondary palate fusion. *Nat Genet*, *11*(4), 409-414. https://doi.org/10.1038/ng1295-409

Puenpatom, R. A., & Victor, T. W. (2009). Increased prevalence of metabolic syndrome in individuals with osteoarthritis: an analysis of NHANES III data. *Postgrad Med*, *121*(6), 9-20. https://doi.org/10.3810/pgm.2009.11.2073

Qiu, C., Zeng, P., Li, X., Zhang, Z., Pan, B., Peng, Z. Y. F., Li, Y., Ma, Y., Leng, Y., & Chen, R. (2017). What is the impact of PCSK9 rs505151 and rs11591147 polymorphisms on serum lipids level and cardiovascular risk: a meta-analysis. *Lipids Health Dis*, *16*(1), 111. https://doi.org/10.1186/s12944-017-0506-6

Quintanilla-Dieck, M. J., Codriansky, K., Keady, M., Bhawan, J., & Runger, T. M. (2008). Cathepsin K in melanoma invasion. *J Invest Dermatol*, *128*(9), 2281-2288. https://doi.org/10.1038/jid.2008.63

Rafieian-Kopaei, M., Setorki, M., Doudi, M., Baradaran, A., & Nasri, H. (2014). Atherosclerosis: process, indicators, risk factors and new hopes. *Int J Prev Med*, *5*(8), 927-946. https://www.ncbi.nlm.nih.gov/pubmed/25489440

Rahman, M. M., Kopec, J. A., Anis, A. H., Cibere, J., & Goldsmith, C. H. (2013). Risk of cardiovascular disease in patients with osteoarthritis: a prospective longitudinal study. *Arthritis Care Res (Hoboken)*, *65*(12), 1951-1958. https://doi.org/10.1002/acr.22092

Ramesh, S., Qi, X. J., Wildey, G. M., Robinson, J., Molkentin, J., Letterio, J., & Howe, P. H. (2008). TGF beta-mediated BIM expression and apoptosis are regulated through SMAD3-dependent expression of the MAPK phosphatase MKP2. *EMBO Rep*, *9*(10), 990-997. https://doi.org/10.1038/embor.2008.158

Rawlings, J. S., Rosler, K. M., & Harrison, D. A. (2004). The JAK/STAT signaling pathway. *J Cell Sci*, *117*(Pt 8), 1281-1283. https://doi.org/10.1242/jcs.00963

Reischauer, S., Arnaout, R., Ramadass, R., & Stainier, D. Y. (2014). Actin binding GFP allows 4D in vivo imaging of myofilament dynamics in the zebrafish heart and the identification of Erbb2 signaling as a remodeling factor of myofibril architecture. *Circ Res*, *115*(10), 845-856. https://doi.org/10.1161/CIRCRESAHA.115.304356

Rigoglou, S., & Papavassiliou, A. G. (2013). The NF-kappaB signalling pathway in osteoarthritis. *Int J Biochem Cell Biol*, *45*(11), 2580-2584. https://doi.org/10.1016/j.biocel.2013.08.018

Ripoll, E., de Ramon, L., Draibe Bordignon, J., Merino, A., Bolanos, N., Goma, M., Cruzado, J. M., Grinyo, J. M., & Torras, J. (2016). JAK3-STAT pathway blocking benefits in experimental lupus nephritis. *Arthritis Res Ther*, *18*(1), 134. https://doi.org/10.1186/s13075-016-1034-x

Rodan, G. A. (1998). Bone homeostasis. *Proc Natl Acad Sci U S A*, *95*(23), 13361-13362. https://doi.org/10.1073/pnas.95.23.13361

Rose, J., Kraft, T., Brenner, B., & Montag, J. (2020). Hypertrophic cardiomyopathy MYH7 mutation R723G alters mRNA secondary structure. *Physiol Genomics*, *52*(1), 15-19. https://doi.org/10.1152/physiolgenomics.00100.2019

Sadick, H., Riedel, F., Naim, R., Goessler, U., Hormann, K., Hafner, M., & Lux, A. (2005). Patients with hereditary hemorrhagic telangiectasia have increased plasma levels of vascular endothelial growth

factor and transforming growth factor-beta1 as well as high ALK1 tissue expression. *Haematologica*, *90*(6), 818-828. https://www.ncbi.nlm.nih.gov/pubmed/15951295

Sadoshima, J., Montagne, O., Wang, Q., Yang, G., Warden, J., Liu, J., Takagi, G., Karoor, V., Hong, C., Johnson, G. L., Vatner, D. E., & Vatner, S. F. (2002). The MEKK1-JNK pathway plays a protective role in pressure overload but does not mediate cardiac hypertrophy. *J Clin Invest*, *110*(2), 271-279. https://doi.org/10.1172/JCI14938

Saftig, P., Hunziker, E., Wehmeyer, O., Jones, S., Boyde, A., Rommerskirch, W., Moritz, J. D., Schu, P., & von Figura, K. (1998). Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin-K-deficient mice. *Proc Natl Acad Sci U S A*, *95*(23), 13453-13458. https://doi.org/10.1073/pnas.95.23.13453

Saito, T., Fukai, A., Mabuchi, A., Ikeda, T., Yano, F., Ohba, S., Nishida, N., Akune, T., Yoshimura, N., Nakagawa, T., Nakamura, K., Tokunaga, K., Chung, U. I., & Kawaguchi, H. (2010). Transcriptional regulation of endochondral ossification by HIF-2alpha during skeletal growth and osteoarthritis development. *Nat Med*, *16*(6), 678-686. https://doi.org/10.1038/nm.2146

Salminen, A., Kauppinen, A., Suuronen, T., & Kaarniranta, K. (2008). SIRT1 longevity factor suppresses NF-kappaB -driven immune responses: regulation of aging via NF-kappaB acetylation? *Bioessays*, *30*(10), 939-942. https://doi.org/10.1002/bies.20799

Saltis, J. (1996). TGF-beta: receptors and cell cycle arrest. *Mol Cell Endocrinol*, *116*(2), 227-232. https://doi.org/10.1016/0303-7207(95)03721-7

Sanpaolo, E. R., Rotondo, C., Cici, D., Corrado, A., & Cantatore, F. P. (2020). JAK/STAT pathway and molecular mechanism in bone remodeling. *Mol Biol Rep*, *47*(11), 9087-9096. https://doi.org/10.1007/s11033-020-05910-9

Saud, K., Herrera-Molina, R., & Von Bernhardi, R. (2005). Pro- and anti-inflammatory cytokines regulate the ERK pathway: implication of the timing for the activation of microglial cells. *Neurotox Res*, *8*(3-4), 277-287. https://doi.org/10.1007/BF03033981

Schilling, T. F., & Kimmel, C. B. (1997). Musculoskeletal patterning in the pharyngeal segments of the zebrafish embryo. *Development*, *124*(15), 2945-2960. https://doi.org/10.1242/dev.124.15.2945

Seemann, P., Brehm, A., Konig, J., Reissner, C., Stricker, S., Kuss, P., Haupt, J., Renninger, S., Nickel, J., Sebald, W., Groppe, J. C., Ploger, F., Pohl, J., Schmidt-von Kegler, M., Walther, M., Gassner, I., Rusu, C., Janecke, A. R., Dathe, K., & Mundlos, S. (2009). Mutations in GDF5 reveal a key residue

mediating BMP inhibition by NOGGIN. *PLoS Genet*, *5*(11), e1000747. https://doi.org/10.1371/journal.pgen.1000747

Selfe, T. K., & Taylor, A. G. (2008). Acupuncture and osteoarthritis of the knee: a review of randomized, controlled trials. *Fam Community Health*, *31*(3), 247-254. https://doi.org/10.1097/01.FCH.0000324482.78577.0f

Settle, S. H., Jr., Rountree, R. B., Sinha, A., Thacker, A., Higgins, K., & Kingsley, D. M. (2003). Multiple joint and skeletal patterning defects caused by single and double mutations in the mouse Gdf6 and Gdf5 genes. *Dev Biol*, *254*(1), 116-130. https://doi.org/10.1016/s0012-1606(02)00022-2

Shakibaei, M., John, T., Schulze-Tanzil, G., Lehmann, I., & Mobasheri, A. (2007). Suppression of NF-kappaB activation by curcumin leads to inhibition of expression of cyclo-oxygenase-2 and matrix metalloproteinase-9 in human articular chondrocytes: Implications for the treatment of osteoarthritis. *Biochem Pharmacol*, *73*(9), 1434-1445. https://doi.org/10.1016/j.bcp.2007.01.005

Sharma, L., Lou, C., Felson, D. T., Dunlop, D. D., Kirwan-Mellis, G., Hayes, K. W., Weinrach, D., & Buchanan, T. S. (1999). Laxity in healthy and osteoarthritic knees. *Arthritis Rheum*, *42*(5), 861-870. https://doi.org/10.1002/1529-0131(199905)42:5<861::AID-ANR4>3.0.CO;2-N

Sheridan, P. E., Mair, C. A., & Quinones, A. R. (2019). Associations between prevalent multimorbidity combinations and prospective disability and self-rated health among older adults in Europe. *BMC Geriatr*, *19*(1), 198. https://doi.org/10.1186/s12877-019-1214-z

Shi, G. P., Bryant, R. A., Riese, R., Verhelst, S., Driessen, C., Li, Z., Bromme, D., Ploegh, H. L., & Chapman, H. A. (2000). Role for cathepsin F in invariant chain processing and major histocompatibility complex class II peptide loading by macrophages. *J Exp Med*, *191*(7), 1177-1186. https://doi.org/10.1084/jem.191.7.1177

Shikatani, E. A., Wang, T., Dingwell, L. S., White-Dzuro, C., Momen, A., & Husain, M. (2023). GDF5 deficiency prevents cardiac rupture following acute myocardial infarction in mice. *Cardiovasc Pathol*, 107581. https://doi.org/10.1016/j.carpath.2023.107581

Shiota, M., Fujimoto, N., Matsumoto, T., Tsukahara, S., Nagakawa, S., Ueda, S., Ushijima, M., Kashiwagi, E., Takeuchi, A., Inokuchi, J., Uchiumi, T., & Eto, M. (2021). Differential Impact of TGFB1 Variation by Metastatic Status in Androgen-Deprivation Therapy for Prostate Cancer. *Front Oncol*, *11*, 697955. https://doi.org/10.3389/fonc.2021.697955

Singer, L., Green, M., Rowe, F., Ben-Shlomo, Y., Kulu, H., & Morrissey, K. (2019). Trends in multimorbidity, complex multimorbidity and multiple functional limitations in the ageing population of England, 2002-2015. *J Comorb*, *9*, 2235042X19872030. https://doi.org/10.1177/2235042X19872030

Skou, S. T., Pedersen, B. K., Abbott, J. H., Patterson, B., & Barton, C. (2018). Physical Activity and Exercise Therapy Benefit More Than Just Symptoms and Impairments in People With Hip and Knee Osteoarthritis. *J Orthop Sports Phys Ther*, *48*(6), 439-447. https://doi.org/10.2519/jospt.2018.7877

Snow, W. M., & Albensi, B. C. (2016). Neuronal Gene Targets of NF-kappaB and Their Dysregulation in Alzheimer's Disease. *Front Mol Neurosci*, *9*, 118. https://doi.org/10.3389/fnmol.2016.00118

Sokolove, J., & Lepus, C. M. (2013). Role of inflammation in the pathogenesis of osteoarthritis: latest findings and interpretations. *Ther Adv Musculoskelet Dis*, *5*(2), 77-94. https://doi.org/10.1177/1759720X12467868

Srivastava, D. (2006). Making or breaking the heart: from lineage determination to morphogenesis. *Cell*, *126*(6), 1037-1048. https://doi.org/10.1016/j.cell.2006.09.003

Staudt, D., & Stainier, D. (2012). Uncovering the molecular and cellular mechanisms of heart development using the zebrafish. *Annu Rev Genet*, *46*, 397-418. https://doi.org/10.1146/annurev-genet-110711-155646

Storm, E. E., & Kingsley, D. M. (1996). Joint patterning defects caused by single and double mutations in members of the bone morphogenetic protein (BMP) family. *Development*, *122*(12), 3969-3979. https://doi.org/10.1242/dev.122.12.3969

Storm, E. E., & Kingsley, D. M. (1999). GDF5 coordinates bone and joint formation during digit development. *Dev Biol*, *209*(1), 11-27. https://doi.org/10.1006/dbio.1999.9241

Sud, A., Kinnersley, B., & Houlston, R. S. (2017). Genome-wide association studies of cancer: current insights and future perspectives. *Nat Rev Cancer*, *17*(11), 692-704. https://doi.org/10.1038/nrc.2017.82

Sullivan, A. M., & O'Keeffe, G. W. (2005). The role of growth/differentiation factor 5 (GDF5) in the induction and survival of midbrain dopaminergic neurones: relevance to Parkinson's disease treatment. *J Anat*, *207*(3), 219-226. https://doi.org/10.1111/j.1469-7580.2005.00447.x

Sun, J., Sukhova, G. K., Zhang, J., Chen, H., Sjoberg, S., Libby, P., Xiang, M., Wang, J., Peters, C., Reinheckel, T., & Shi, G. P. (2011). Cathepsin L activity is essential to elastase perfusion-induced

abdominal aortic aneurysms in mice. *Arterioscler Thromb Vasc Biol*, *31*(11), 2500-2508. https://doi.org/10.1161/ATVBAHA.111.230201

Tachmazidou, I., Hatzikotoulas, K., Southam, L., Esparza-Gordillo, J., Haberland, V., Zheng, J., Johnson, T., Koprulu, M., Zengini, E., Steinberg, J., Wilkinson, J. M., Bhatnagar, S., Hoffman, J. D., Buchan, N., Suveges, D., arc, O. C., Yerges-Armstrong, L., Smith, G. D., Gaunt, T. R., . . . Zeggini, E. (2019). Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. *Nat Genet*, *51*(2), 230-236. https://doi.org/10.1038/s41588-018-0327-1

Taipale, J., Miyazono, K., Heldin, C. H., & Keski-Oja, J. (1994). Latent transforming growth factor-beta 1 associates to fibroblast extracellular matrix via latent TGF-beta binding protein. *J Cell Biol*, *124*(1-2), 171-181. https://doi.org/10.1083/jcb.124.1.171

Takahara, M., Harada, M., Guan, D., Otsuji, M., Naruse, T., Takagi, M., & Ogino, T. (2004). Developmental failure of phalanges in the absence of growth/differentiation factor 5. *Bone*, *35*(5), 1069-1076. https://doi.org/10.1016/j.bone.2004.06.020

Takito, J., Inoue, S., & Nakamura, M. (2018). The Sealing Zone in Osteoclasts: A Self-Organized Structure on the Bone. *Int J Mol Sci*, *19*(4). https://doi.org/10.3390/ijms19040984

Tang, L. Y., Heller, M., Meng, Z., Yu, L. R., Tang, Y., Zhou, M., & Zhang, Y. E. (2017). Transforming Growth Factor-beta (TGF-beta) Directly Activates the JAK1-STAT3 Axis to Induce Hepatic Fibrosis in Coordination with the SMAD Pathway. *J Biol Chem*, *292*(10), 4302-4312. https://doi.org/10.1074/jbc.M116.773085

Tashiro, H., Shimokawa, H., Sadamatu, K., & Yamamoto, K. (2002). Prognostic significance of plasma concentrations of transforming growth factor-beta in patients with coronary artery disease. *Coron Artery Dis*, *13*(3), 139-143. https://doi.org/10.1097/00019501-200205000-00001

Taylor, A. W. (2009). Review of the activation of TGF-beta in immunity. *J Leukoc Biol*, *85*(1), 29-33. https://doi.org/10.1189/jlb.0708415

Thisse, B., Heyer, V., Lux, A., Alunni, V., Degrave, A., Seiliez, I., Kirchner, J., Parkhill, J. P., & Thisse, C. (2004). Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. *Methods Cell Biol*, *77*, 505-519. https://doi.org/10.1016/s0091-679x(04)77027-2

Thouverey, C., & Caverzasio, J. (2015). Focus on the p38 MAPK signaling pathway in bone development and maintenance. *Bonekey Rep*, *4*, 711. https://doi.org/10.1038/bonekey.2015.80

Tian, B., Nowak, D. E., Jamaluddin, M., Wang, S., & Brasier, A. R. (2005). Identification of direct genomic targets downstream of the nuclear factor-kappaB transcription factor mediating tumor necrosis factor signaling. *J Biol Chem*, *280*(17), 17435-17448. https://doi.org/10.1074/jbc.M500437200

Tian, M., Chang, X., Zhang, Q., Li, C., Li, S., & Sun, Y. (2019). TGF-beta1 mediated MAPK signaling pathway promotes collagen formation induced by Nano NiO in A549 cells. *Environ Toxicol*, *34*(6), 719-727. https://doi.org/10.1002/tox.22738

Tichauer, J. E., Flores, B., Soler, B., Eugenin-von Bernhardi, L., Ramirez, G., & von Bernhardi, R. (2014). Age-dependent changes on TGFbeta1 Smad3 pathway modify the pattern of microglial cell activation. *Brain Behav Immun*, *37*, 187-196. https://doi.org/10.1016/j.bbi.2013.12.018

Tilstra, J. S., Clauson, C. L., Niedernhofer, L. J., & Robbins, P. D. (2011). NF-kappaB in Aging and Disease. *Aging Dis*, *2*(6), 449-465. https://www.ncbi.nlm.nih.gov/pubmed/22396894

Troen, B. R. (2004). The role of cathepsin K in normal bone resorption. *Drug News Perspect*, *17*(1), 19-28. https://doi.org/10.1358/dnp.2004.17.1.829022

Troen, B. R. (2006). The regulation of cathepsin K gene expression. *Ann N Y Acad Sci*, *1068*, 165-172. https://doi.org/10.1196/annals.1346.018

Tschon, M., Contartese, D., Pagani, S., Borsari, V., & Fini, M. (2021). Gender and Sex Are Key Determinants in Osteoarthritis Not Only Confounding Variables. A Systematic Review of Clinical Data. *J Clin Med*, *10*(14). https://doi.org/10.3390/jcm10143178

Tucker, R. F., Branum, E. L., Shipley, G. D., Ryan, R. J., & Moses, H. L. (1984). Specific binding to cultured cells of 125I-labeled type beta transforming growth factor from human platelets. *Proc Natl Acad Sci U S A*, *81*(21), 6757-6761. https://doi.org/10.1073/pnas.81.21.6757

Tuli, R., Tuli, S., Nandi, S., Huang, X., Manner, P. A., Hozack, W. J., Danielson, K. G., Hall, D. J., & Tuan, R. S. (2003). Transforming growth factor-beta-mediated chondrogenesis of human mesenchymal progenitor cells involves N-cadherin and mitogen-activated protein kinase and Wnt signaling cross-talk. *J Biol Chem*, *278*(42), 41227-41236. https://doi.org/10.1074/jbc.M305312200

Turk, B., Turk, V., & Turk, D. (1997). Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem*, *378*(3-4), 141-150. https://www.ncbi.nlm.nih.gov/pubmed/9165064

Turk, V., Stoka, V., Vasiljeva, O., Renko, M., Sun, T., Turk, B., & Turk, D. (2012). Cysteine cathepsins: from structure, function and regulation to new frontiers. *Biochim Biophys Acta*, *1824*(1), 68-88. https://doi.org/10.1016/j.bbapap.2011.10.002

van de Laar, I. M., van der Linde, D., Oei, E. H., Bos, P. K., Bessems, J. H., Bierma-Zeinstra, S. M., van Meer, B. L., Pals, G., Oldenburg, R. A., Bekkers, J. A., Moelker, A., de Graaf, B. M., Matyas, G., Frohn-Mulder, I. M., Timmermans, J., Hilhorst-Hofstee, Y., Cobben, J. M., Bruggenwirth, H. T., van Laer, L., . . . Wessels, M. W. (2012). Phenotypic spectrum of the SMAD3-related aneurysms-osteoarthritis syndrome. *J Med Genet*, *49*(1), 47-57. https://doi.org/10.1136/jmedgenet-2011-100382

van den Brule, S., Misson, P., Buhling, F., Lison, D., & Huaux, F. (2005). Overexpression of cathepsin K during silica-induced lung fibrosis and control by TGF-beta. *Respir Res*, *6*(1), 84. https://doi.org/10.1186/1465-9921-6-84

Van der Heiden, K., Cuhlmann, S., Luong le, A., Zakkar, M., & Evans, P. C. (2010). Role of nuclear factor kappaB in cardiovascular health and disease. *Clin Sci (Lond)*, *118*(10), 593-605. https://doi.org/10.1042/CS20090557

van der Kraan, P. M. (2017). The changing role of TGFbeta in healthy, ageing and osteoarthritic joints. *Nat Rev Rheumatol*, *13*(3), 155-163. https://doi.org/10.1038/nrrheum.2016.219

van der Kraan, P. M., Goumans, M. J., Blaney Davidson, E., & ten Dijke, P. (2012). Age-dependent alteration of TGF-beta signalling in osteoarthritis. *Cell Tissue Res*, *347*(1), 257-265. https://doi.org/10.1007/s00441-011-1194-6

van der Linde, D., Konings, E. E., Slager, M. A., Witsenburg, M., Helbing, W. A., Takkenberg, J. J., & Roos-Hesselink, J. W. (2011). Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol*, *58*(21), 2241-2247. https://doi.org/10.1016/j.jacc.2011.08.025

Varshney, G. K., Carrington, B., Pei, W., Bishop, K., Chen, Z., Fan, C., Xu, L., Jones, M., LaFave, M. C., Ledin, J., Sood, R., & Burgess, S. M. (2016). A high-throughput functional genomics workflow based on CRISPR/Cas9-mediated targeted mutagenesis in zebrafish. *Nat Protoc*, *11*(12), 2357-2375. https://doi.org/10.1038/nprot.2016.141

Veronese, N., Trevisan, C., De Rui, M., Bolzetta, F., Maggi, S., Zambon, S., Musacchio, E., Sartori, L., Perissinotto, E., Crepaldi, G., Manzato, E., & Sergi, G. (2016). Association of Osteoarthritis With Increased Risk of Cardiovascular Diseases in the Elderly: Findings From the Progetto Veneto Anziano Study Cohort. *Arthritis Rheumatol*, *68*(5), 1136-1144. https://doi.org/10.1002/art.39564

Verrecchia, F., & Mauviel, A. (2002). Transforming growth factor-beta signaling through the Smad pathway: role in extracellular matrix gene expression and regulation. *J Invest Dermatol*, *118*(2), 211-215. https://doi.org/10.1046/j.1523-1747.2002.01641.x

Vincenti, M. P., Coon, C. I., & Brinckerhoff, C. E. (1998). Nuclear factor kappaB/p50 activates an element in the distal matrix metalloproteinase 1 promoter in interleukin-1beta-stimulated synovial fibroblasts. *Arthritis Rheum*, *41*(11), 1987-1994. https://doi.org/10.1002/1529-0131(199811)41:11<1987::AID-ART14>3.0.CO;2-8

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *Am J Hum Genet*, *90*(1), 7-24. https://doi.org/10.1016/j.ajhg.2011.11.029

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet*, *101*(1), 5-22. https://doi.org/10.1016/j.ajhg.2017.06.005

Wada, T., Joza, N., Cheng, H. Y., Sasaki, T., Kozieradzki, I., Bachmaier, K., Katada, T., Schreiber, M., Wagner, E. F., Nishina, H., & Penninger, J. M. (2004). MKK7 couples stress signalling to G2/M cell-cycle progression and cellular senescence. *Nat Cell Biol*, *6*(3), 215-226. https://doi.org/10.1038/ncb1098

Wahlberg, K. E., Guazzetti, S., Pineda, D., Larsson, S. C., Fedrighi, C., Cagna, G., Zoni, S., Placidi, D., Wright, R. O., Smith, D. R., Lucchini, R. G., & Broberg, K. (2018). Polymorphisms in Manganese Transporters SLC30A10 and SLC39A8 Are Associated With Children's Neurodevelopment by Influencing Manganese Homeostasis. *Front Genet*, *9*, 664. https://doi.org/10.3389/fgene.2018.00664

Waldmann, L., Leyhr, J., Zhang, H., Allalou, A., Ohman-Magi, C., & Haitina, T. (2022). The role of Gdf5 in the development of the zebrafish fin endoskeleton. *Dev Dyn*, *251*(9), 1535-1549. https://doi.org/10.1002/dvdy.399

Wan, C. Y., Ruber, T., Hohmann, A., & Schlaug, G. (2010). The Therapeutic Effects of Singing in Neurological Disorders. *Music Percept*, *27*(4), 287-295. https://doi.org/10.1525/mp.2010.27.4.287

Wang, H., Bai, J., He, B., Hu, X., & Liu, D. (2016). Osteoarthritis and the risk of cardiovascular disease: a meta-analysis of observational studies. *Sci Rep*, *6*, 39672. https://doi.org/10.1038/srep39672

Wang, K., Zhang, H., Kugathasan, S., Annese, V., Bradfield, J. P., Russell, R. K., Sleiman, P. M., Imielinski, M., Glessner, J., Hou, C., Wilson, D. C., Walters, T., Kim, C., Frackelton, E. C., Lionetti, P., Barabino, A., Van Limbergen, J., Guthery, S., Denson, L., . . . Hakonarson, H. (2009). Diverse

genome-wide association studies associate the IL12/IL23 pathway with Crohn Disease. *Am J Hum Genet*, *84*(3), 399-405. https://doi.org/10.1016/j.ajhg.2009.01.026

Wang, L., Jin, F., Wang, P., Hou, S., Jin, T., Chang, X., & Zhao, L. (2022). Adropin Inhibits Vascular Smooth Muscle Cell Osteogenic Differentiation to Alleviate Vascular Calcification via the JAK2/STAT3 Signaling Pathway. *Biomed Res Int*, *2022*, 9122264. https://doi.org/10.1155/2022/9122264

Wang, W., Gao, W., Zhu, Q., Alasbahi, A., Seki, E., & Yang, L. (2021). TAK1: A Molecular Link Between Liver Inflammation, Fibrosis, Steatosis, and Carcinogenesis. *Front Cell Dev Biol*, *9*, 734749. https://doi.org/10.3389/fcell.2021.734749

Watanabe, M., Masuyama, N., Fukuda, M., & Nishida, E. (2000). Regulation of intracellular dynamics of Smad4 by its leucine-rich nuclear export signal. *EMBO Rep*, *1*(2), 176-182. https://doi.org/10.1093/embo-reports/kvd029

Waterworth, D. M., Ricketts, S. L., Song, K., Chen, L., Zhao, J. H., Ripatti, S., Aulchenko, Y. S., Zhang, W., Yuan, X., Lim, N., Luan, J., Ashford, S., Wheeler, E., Young, E. H., Hadley, D., Thompson, J. R., Braund, P. S., Johnson, T., Struchalin, M., . . . Sandhu, M. S. (2010). Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol*, *30*(11), 2264-2276. https://doi.org/10.1161/ATVBAHA.109.201020

Weiss, A., & Attisano, L. (2013). The TGFbeta superfamily signaling pathway. *Wiley Interdiscip Rev Dev Biol*, *2*(1), 47-63. https://doi.org/10.1002/wdev.86

Weston, C. R., & Davis, R. J. (2007). The JNK signal transduction pathway. *Curr Opin Cell Biol*, *19*(2), 142-149. https://doi.org/10.1016/j.ceb.2007.02.001

Wetterling, T. (2021). Pathogenesis of multimorbidity-what is known? *Z Gerontol Geriatr*, *54*(6), 590-596. https://doi.org/10.1007/s00391-020-01752-z (Pathogenese der Multimorbiditat - was ist bekannt?)

Wilson, S., Hashamiyan, S., Clarke, L., Saftig, P., Mort, J., Dejica, V. M., & Bromme, D. (2009). Glycosaminoglycan-mediated loss of cathepsin K collagenolytic activity in MPS I contributes to osteoclast and growth plate abnormalities. *Am J Pathol*, *175*(5), 2053-2062. https://doi.org/10.2353/ajpath.2009.090211

Witoonpanich, B., Jinawath, A., Wongtawan, T., & Tawonsawatruk, T. (2022). Association of synovial expression of growth and differentiation factor 5 (GDF5) with radiographic severity of knee osteoarthritis. *Heliyon*, *8*(11), e11798. https://doi.org/10.1016/j.heliyon.2022.e11798

Witten, P. E., Hansen, A., & Hall, B. K. (2001). Features of mono- and multinucleated bone resorbing cells of the zebrafish Danio rerio and their contribution to skeletal development, remodeling, and growth. *J Morphol*, *250*(3), 197-207. https://doi.org/10.1002/jmor.1065

Woo, J., Lau, E., Lee, P., Kwok, T., Lau, W. C., Chan, C., Chiu, P., Li, E., Sham, A., & Lam, D. (2004). Impact of osteoarthritis on quality of life in a Hong Kong Chinese population. *J Rheumatol*, *31*(12), 2433-2438. https://www.ncbi.nlm.nih.gov/pubmed/15570647

Wu, C. F., Chiang, W. C., Lai, C. F., Chang, F. C., Chen, Y. T., Chou, Y. H., Wu, T. H., Linn, G. R., Ling, H., Wu, K. D., Tsai, T. J., Chen, Y. M., Duffield, J. S., & Lin, S. L. (2013). Transforming growth factor beta-1 stimulates profibrotic epithelial signaling to activate pericyte-myofibroblast transition in obstructive kidney fibrosis. *Am J Pathol*, *182*(1), 118-131. https://doi.org/10.1016/j.ajpath.2012.09.009

Wu, M., Chen, G., & Li, Y. P. (2016). TGF-beta and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. *Bone Res*, *4*, 16009. https://doi.org/10.1038/boneres.2016.9

Wu, M. Y., & Hill, C. S. (2009). Tgf-beta superfamily signaling in embryonic development and homeostasis. *Dev Cell*, *16*(3), 329-343. https://doi.org/10.1016/j.devcel.2009.02.012

Wu, N., Wang, Y., Wang, K., Zhong, B., Liao, Y., Liang, J., & Jiang, N. (2022). Cathepsin K regulates the tumor growth and metastasis by IL-17/CTSK/EMT axis and mediates M2 macrophage polarization in castration-resistant prostate cancer. *Cell Death Dis*, *13*(9), 813. https://doi.org/10.1038/s41419-022-05215-8

Wu, P. F., Zhang, X. H., Zhou, P., Yin, R., Zhou, X. T., & Zhang, W. (2021). Growth Differentiation Factor 15 Is Associated With Alzheimer's Disease Risk. *Front Genet*, *12*, 700371. https://doi.org/10.3389/fgene.2021.700371

Xie, W., Mertens, J. C., Reiss, D. J., Rimm, D. L., Camp, R. L., Haffty, B. G., & Reiss, M. (2002). Alterations of Smad signaling in human breast carcinoma are associated with poor outcome: a tissue microarray study. *Cancer Res*, *62*(2), 497-505. https://www.ncbi.nlm.nih.gov/pubmed/11809701

Xing, C., Gong, B., Xue, Y., Han, Y., Wang, Y., Meng, A., & Jia, S. (2015). TGFbeta1a regulates zebrafish posterior lateral line formation via Smad5 mediated pathway. *J Mol Cell Biol*, *7*(1), 48-61. https://doi.org/10.1093/jmcb/mjv004

Xu, Y., & Pasche, B. (2007). TGF-beta signaling alterations and susceptibility to colorectal cancer. *Hum Mol Genet*, *16 Spec No 1*(SPEC), R14-20. https://doi.org/10.1093/hmg/ddl486

Yadati, T., Houben, T., Bitorina, A., & Shiri-Sverdlov, R. (2020). The Ins and Outs of Cathepsins: Physiological Function and Role in Disease Management. *Cells*, *9*(7). https://doi.org/10.3390/cells9071679

Yamada, Y. (2000). Association of a Leu(10)-->Pro polymorphism of the transforming growth factor-beta1 with genetic susceptibility to osteoporosis and spinal osteoarthritis. *Mech Ageing Dev*, *116*(2-3), 113-123. https://doi.org/10.1016/s0047-6374(00)00131-7

Yamashita, H., Ten Dijke, P., Heldin, C. H., & Miyazono, K. (1996). Bone morphogenetic protein receptors. *Bone*, *19*(6), 569-574. https://doi.org/10.1016/s8756-3282(96)00259-1

Yamauchi-Takihara, K., Ihara, Y., Ogata, A., Yoshizaki, K., Azuma, J., & Kishimoto, T. (1995). Hypoxic stress induces cardiac myocyte-derived interleukin-6. *Circulation*, *91*(5), 1520-1524. https://doi.org/10.1161/01.cir.91.5.1520

Yan, B., Chen, G., Saigal, K., Yang, X., Jensen, S. T., Van Waes, C., Stoeckert, C. J., & Chen, Z. (2008). Systems biology-defined NF-kappaB regulons, interacting signal pathways and networks are implicated in the malignant phenotype of head and neck cancer cell lines differing in p53 status. *Genome Biol*, *9*(3), R53. https://doi.org/10.1186/gb-2008-9-3-r53

Yang, C., Coker, K. J., Kim, J. K., Mora, S., Thurmond, D. C., Davis, A. C., Yang, B., Williamson, R. A., Shulman, G. I., & Pessin, J. E. (2001). Syntaxin 4 heterozygous knockout mice develop muscle insulin resistance. *J Clin Invest*, *107*(10), 1311-1318. https://doi.org/10.1172/JCI12274

Yoshida, E., Suzuki, T., Morita, M., Taguchi, K., Tsuchida, K., Motohashi, H., Doita, M., & Yamamoto, M. (2018). Hyperactivation of Nrf2 leads to hypoplasia of bone in vivo. *Genes Cells*, *23*(5), 386-392. https://doi.org/10.1111/gtc.12579

Yusuf, S., Hawken, S., Ounpuu, S., Dans, T., Avezum, A., Lanas, F., McQueen, M., Budaj, A., Pais, P., Varigos, J., Lisheng, L., & Investigators, I. S. (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet*, *364*(9438), 937-952. https://doi.org/10.1016/S0140-6736(04)17018-9

Zaidi, D., & Wine, E. (2018). Regulation of Nuclear Factor Kappa-Light-Chain-Enhancer of Activated B Cells (NF-kappabeta) in Inflammatory Bowel Diseases. *Front Pediatr*, *6*, 317. https://doi.org/10.3389/fped.2018.00317

Zaidi, S. H., Huang, Q., Momen, A., Riazi, A., & Husain, M. (2010). Growth differentiation factor 5 regulates cardiac repair after myocardial infarction. *J Am Coll Cardiol*, *55*(2), 135-143. https://doi.org/10.1016/j.jacc.2009.08.041

Zhang, B., Calado, D. P., Wang, Z., Frohler, S., Kochert, K., Qian, Y., Koralov, S. B., Schmidt-Supprian, M., Sasaki, Y., Unitt, C., Rodig, S., Chen, W., Dalla-Favera, R., Alt, F. W., Pasqualucci, L., & Rajewsky, K. (2015). An oncogenic role for alternative NF-kappaB signaling in DLBCL revealed upon deregulated BCL6 expression. *Cell Rep*, *11*(5), 715-726. https://doi.org/10.1016/j.celrep.2015.03.059

Zhang, L., Zhou, F., & ten Dijke, P. (2013). Signaling interplay between transforming growth factor-beta receptor and PI3K/AKT pathways in cancer. *Trends Biochem Sci*, *38*(12), 612-620. https://doi.org/10.1016/j.tibs.2013.10.001

Zhang, P., Lewinger, J. P., Conti, D., Morrison, J. L., & Gauderman, W. J. (2016). Detecting Gene-Environment Interactions for a Quantitative Trait in a Genome-Wide Association Study. *Genet Epidemiol*, *40*(5), 394-403. https://doi.org/10.1002/gepi.21977

Zhang, W., & Liu, H. T. (2002). MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Res*, *12*(1), 9-18. https://doi.org/10.1038/sj.cr.7290105

Zhang, W., Wu, X., Pei, Z., Kiess, W., Yang, Y., Xu, Y., Chang, Z., Wu, J., Sun, C., & Luo, F. (2019). GDF5 Promotes White Adipose Tissue Thermogenesis via p38 MAPK Signaling Pathway. *DNA Cell Biol*, *38*(11), 1303-1312. https://doi.org/10.1089/dna.2019.4724

Zhang, X., Fu, Z., Meng, L., He, M., & Zhang, Z. (2018). The Early Events That Initiate beta-Amyloid Aggregation in Alzheimer's Disease. *Front Aging Neurosci*, *10*, 359. https://doi.org/10.3389/fnagi.2018.00359

Zhang, X., Luo, S., Wang, M., & Shi, G. P. (2020). Cysteinyl cathepsins in cardiovascular diseases. *Biochim Biophys Acta Proteins Proteom*, *1868*(4), 140360. https://doi.org/10.1016/j.bbapap.2020.140360

Zhao, W., Rasheed, A., Tikkanen, E., Lee, J. J., Butterworth, A. S., Howson, J. M. M., Assimes, T. L., Chowdhury, R., Orho-Melander, M., Damrauer, S., Small, A., Asma, S., Imamura, M., Yamauch, T., Chambers, J. C., Chen, P., Sapkota, B. R., Shah, N., Jabeen, S., . . . Saleheen, D. (2017). Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nat Genet*, *49*(10), 1450-1457. https://doi.org/10.1038/ng.3943

Zhou, B., Ma, Q., Rajagopal, S., Wu, S. M., Domian, I., Rivera-Feliciano, J., Jiang, D., von Gise, A., Ikeda, S., Chien, K. R., & Pu, W. T. (2008). Epicardial progenitors contribute to the cardiomyocyte lineage in the developing heart. *Nature*, *454*(7200), 109-113. https://doi.org/10.1038/nature07060

Zhou, S., Eid, K., & Glowacki, J. (2004). Cooperation between TGF-beta and Wnt pathways during chondrocyte and adipocyte differentiation of human marrow stromal cells. *J Bone Miner Res*, *19*(3), 463-470. https://doi.org/10.1359/JBMR.0301239

Zhuo, Q., Yang, W., Chen, J., & Wang, Y. (2012). Metabolic syndrome meets osteoarthritis. *Nat Rev Rheumatol*, *8*(12), 729-737. https://doi.org/10.1038/nrrheum.2012.135